When Do Markets Fully React to Monetary Policy Announcements?

Paul L. Tran*

University of Texas at Austin

September 29, 2025

*Email: pltran@utexas.edu. Website: https://paulletran.com/.

Motivation

- News is released $\longrightarrow$ Financial markets react to news
  - If change in price $\approx$ change in expectations $\longrightarrow$ Unanticipated news/news shock

Motivation

- News is released $\longrightarrow$ Financial markets react to news
- Useful for empirical macro: Obtain shocks from news to infer causal effects
  - Ex: Monetary policy (MP) announcements

Motivation

- News is released → Financial markets react to news
- Useful for empirical macro: Obtain shocks from news to infer causal effects
- Method: High-frequency Identification (HFI) of MP shocks
  - Measure price change within event window around MP announcement
  - Most popular choice in literature: 30 minutes

Motivation

- News is released $\longrightarrow$ Financial markets react to news
- Useful for empirical macro: Obtain shocks from news to infer causal effects
- Method: High-frequency Identification (HFI) of MP shocks
- **Research Q**: What size should the window length around MP announcements be?
  - Too short: Markets might not fully react to policy news yet
  - Too long: Change in price $\approx$ Unanticipated MP with unrelated news
  - Just right: Change in price $\approx$ Unanticipated MP with minimised noise

Motivation

- News is released $\longrightarrow$ Financial markets react to news
- Useful for empirical macro: Obtain shocks from news to infer causal effects
- Method: High-frequency Identification (HFI) of MP shocks
- **Research Q**: What size should the window length around MP announcements be?
  - Too short: Markets might not fully react to policy news yet
  - Too long: Change in price $\approx$ Unanticipated MP with unrelated news
  - Just right: Change in price $\approx$ Unanticipated MP with minimised noise
- **Wrong A**: Contributes to MP shocks lacking precision $\because$ noise

Summary: Previous of Results

FOMC Statement Ex

- **This Paper**: Estimate optimal window size for FOMC statements using NLP:
  - By combining text-based signal with observed price dynamics
  - By isolating market price changes to "full" text of FOMC statements

Summary: Previous of Results

FOMC Statement Ex

- **This Paper**: Estimate optimal window size for FOMC statements using NLP

Summary: Previous of Results

FOMC Statement Ex

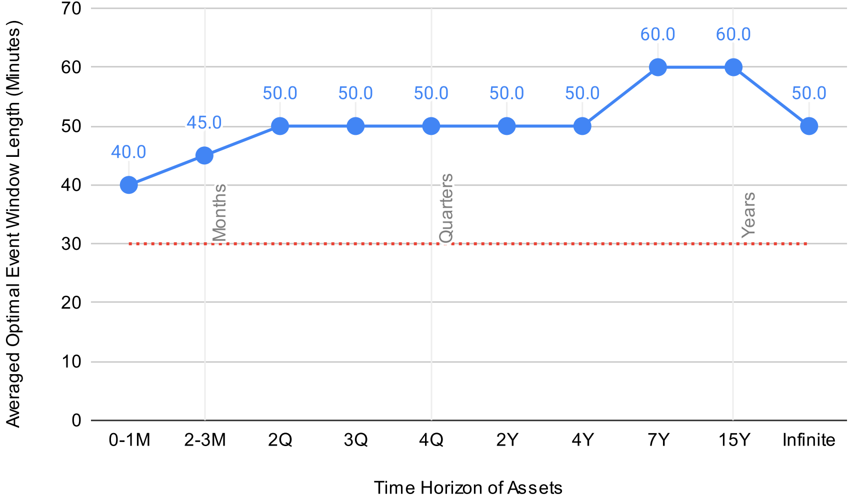
- **This Paper**: Estimate optimal window size for FOMC statements using NLP
- **How Long?** Longer than 30 minutes:
  - On avg, markets fully react within window 10 min before and 30+ min after
  - Time horizon of assets $\uparrow \rightarrow$ Avg optimal window length $\uparrow$
  - Time horizon of asset at least 2 quarters out $\rightarrow$ 50- to 60-min window
  - Complex/dissimilar/dissent statements $\rightarrow$ Relatively longer windows

Summary: Previous of Results

- **This Paper**: Estimate optimal window size for FOMC statements using NLP
- **How Long?** Longer than 30 minutes:
  - On avg, markets fully react within window 10 min before and 30+ min after
  - Time horizon of assets $\uparrow\rightarrow$ Avg optimal window length $\uparrow$
  - Time horizon of asset at least 2 quarters out $\rightarrow$ 50- to 60-min window
  - Complex/dissimilar/dissent statements $\rightarrow$ Relatively longer windows
- **Effects**: By changing only event window choice:
  - Time horizon of assets $\uparrow\rightarrow$ corr. between MP surprise sets $\downarrow$
  - MP shocks about forward guidance have $\uparrow$ impact on yields and stock prices

Summary: Diff Horizons, Diff Window Lengths

Summary Text Summary Table Recap



Related Literature and Contributions

1. **Measuring Appropriate Event Window Lengths**
   - Hillmer and Yu (1979); Chang and Chen (1989); Krivin et al. (2003); Das and King (2021); An et al. (2025); Boehm and Kroner (2025); and others...
   - Contributions: (1) If news = text → Method can estimate optimal window; (2) Introduce noise min. approach

Related Literature and Contributions

1. **Measuring Appropriate Event Window Lengths**
2. **Text Analysis in Monetary Policy Communication**
    - Lucca and Trebbi (2009); Hansen and McMahon (2016); Hansen, McMahon, and Prat (2017); Cieslak and Vissing-Jorgensen (2020); Husted et al. (2020); Handlan (2022a); Handlan (2022b); Acosta (2023); Doh et al. (2023); Cieslak and McMahon (2023); Cieslak, Hansen, et al. (2023); Gorodnichenko et al. (2023); Aruoba and Drechsel (2024); Gáti and Handlan (2025a); Gáti and Handlan (2025b); Pillar et al. (2025); and others...
    - Contributions: ↑ adoption of NLP to quantify dimensions of MP communication

Related Literature and Contributions

1. **Measuring Appropriate Event Window Lengths**
2. **Text Analysis in Monetary Policy Communication**
3. **Event Window Lengths in Monetary Policy**
   - Examples: Gürkaynak, Sack, et al. (2005); Nakamura and Steinsson (2018); Swanson and Jayawickrema (2023); An et al. (2025); Boehm and Kroner (2025); and others...
   - Contributions: (1) Optimal window length around FOMC statements > 30-min; (2) diff. markets, diff. window lengths; (3) MP effects less dampened

Presentation Roadmap

Motivation: Why the Need for NLP?

- News is released $\implies$ Markets react to news
- Because of cognitive noise, markets might need more time to react
- But too much time can introduce unrelated news to measured reaction
- **Therefore**: How to choose optimal time horizon with minimal noise?
- **Purpose**: Motivate the need for NLP method with FOMC statements
  - Simple framework of asset price movements around news

Conceptual Framework of Asset Market Prices (1/4)

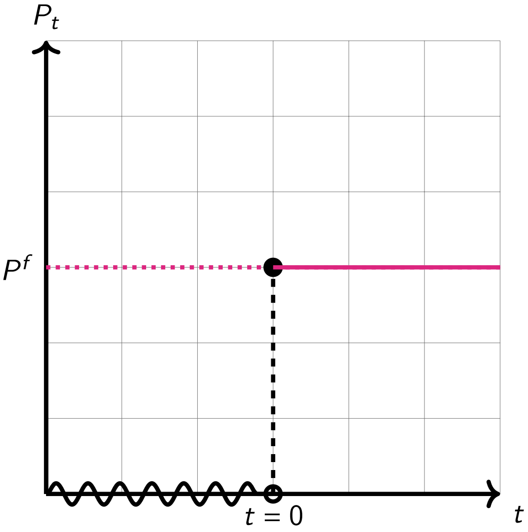
$$\{P_t = P_t^f + \varepsilon_t^c + \varepsilon_t^n | t \geq 0\}, \tag{1}$$

- Full price component: $P_t^f = P^f \in \mathbb{R}$
- Interpretation: Price $\because$ news

Impulse Response Scenarios of Asset Prices (1/4)

Scenario 1. No cognitive noise + No unrelated news

- $P_t \rightarrow P^f$ $\because$ no cognitive noise
- P_t moves anywhere over time $\because$ unrelated news
- $\therefore$ Choose shortest event window



Conceptual Framework of Asset Market Prices (2/4)

$$\{P_t = P_t^f + \varepsilon_t^c + \varepsilon_t^n | t \geq 0\}, \tag{2}$$

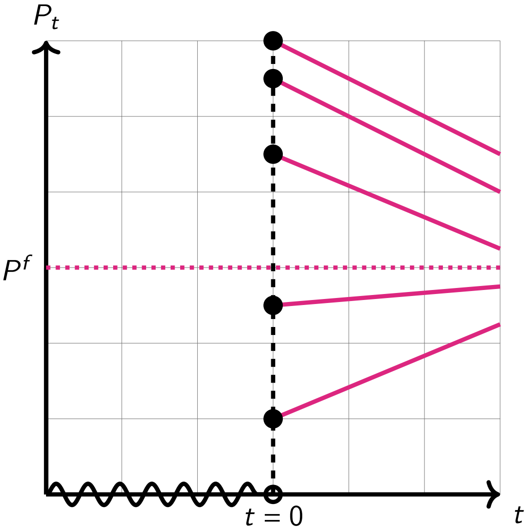
- Full price component: $P_t^f = P^f \in \mathbb{R}$
- Cognitive noise: $\varepsilon_t^c = \rho^c \varepsilon_{t-1}^c + e^{-\mathcal{D}t} \nu_t^c$ Interpretations
  - $\nu_t^c \sim \mathcal{N}(0, \sigma_c^2)$
  - $|\rho_c| < 1$
  - Decay: $\mathcal{D} \in \mathbb{R}^+$
  - $|\frac{\rho_c}{\mathcal{D}}| < 1$
  - Assumption: $\mathrm{Var}(\varepsilon_0^c) = \sigma_c^2$
- ε_t^c and error decay to zero

Impulse Response Scenarios of Asset Prices (2/4)

Scenario 2. Cognitive noise + No unrelated news

Interpretations

- P_t jumps anywhere $\because$ cognitive noise
- $P_t \rightarrow P^f$ $\because$ no unrelated news
- $\therefore$ Choose long event window



Conceptual Framework of Asset Market Prices (3/4)

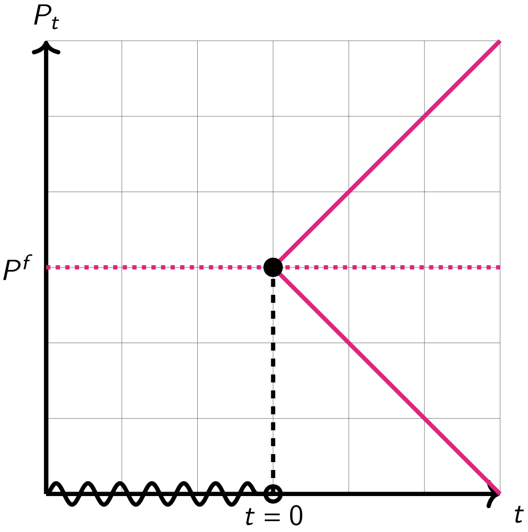
$$\{P_t = P_t^f + \varepsilon_t^c + \varepsilon_t^n | t \geq 0\}, \tag{3}$$

- Full price component: $P_t^f = P^f \in \mathbb{R}$
- Unrelated news: $\varepsilon_t^n = \varepsilon_{t-1}^n + \nu_t^n$
  - $\nu_t^n \sim \mathcal{N}(0, \sigma_n^2)$
  - Assumption: $\mathrm{Var}(\varepsilon_0^n) = 0$

Impulse Response Scenarios of Asset Prices (3/4)

Scenario 1. No cognitive noise + Unrelated news

- $P_t \rightarrow P^f \because$ no cognitive noise
- P_t moves anywhere over time $\because$ unrelated news
- $\therefore$ Choose short event window



Conceptual Framework of Asset Market Prices (4/4)

$$\{P_t = P_t^f + \varepsilon_t^c + \varepsilon_t^n | t \geq 0\}, \tag{4}$$

- Full price component: $P_t^f = P^f \in \mathbb{R}$
- Cognitive noise: $\varepsilon_t^c = \rho^c \varepsilon_{t-1}^c + e^{-\mathcal{D}t} \nu_t^c$

  Interpretations

  - $\nu_t^c \sim \mathcal{N}(0, \sigma_c^2)$
  - $|\rho_c| < 1$
  - Decay: $\mathcal{D} \in \mathbb{R}^+$
  - $|\frac{\rho_c}{\mathcal{D}}| < 1$
  - Assumption: $\mathrm{Var}(\varepsilon_0^c) = \sigma_c^2$
- Unrelated news: $\varepsilon_t^n = \varepsilon_{t-1}^n + \nu_t^n$
  - $\nu_t^n \sim \mathcal{N}(0, \sigma_n^2)$
  - Assumption: $\mathrm{Var}(\varepsilon_0^n) = 0$

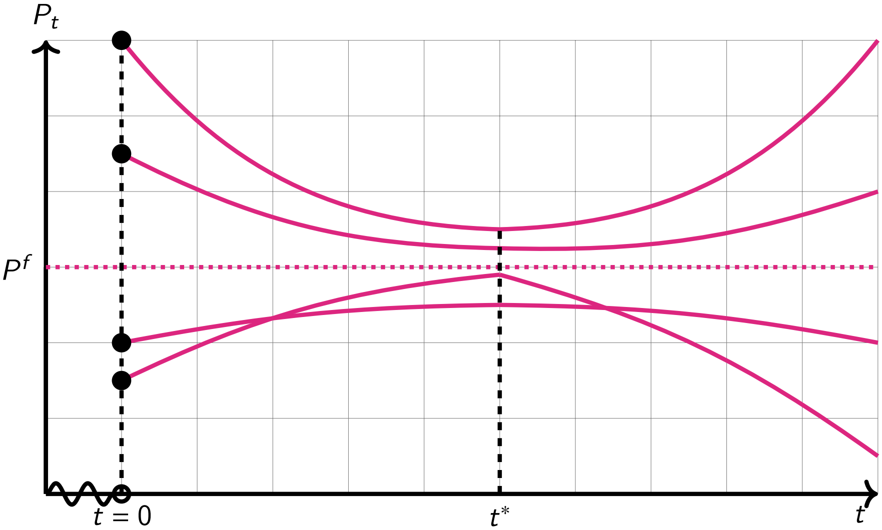
Conceptual Framework of Asset Market Prices (4/4)

$$\{P_t = P_t^f + \varepsilon_t^c + \varepsilon_t^n | t \geq 0\}, \tag{4}$$

- Full price component: $P_t^f = P^f \in \mathbb{R}$
- Cognitive noise: $\varepsilon_t^c = \rho^c \varepsilon_{t-1}^c + e^{-\mathcal{D}t} \nu_t^c$ Interpretations
- Unrelated news: $\varepsilon_t^n = \varepsilon_{t-1}^n + \nu_t^n$
- **Goal**: If $\exists$ “good” signal $\rightarrow$ Estimate time horizon reflecting full market reactions

Impulse Response Scenarios of Asset Prices (4/4)

Scenario 3. Cognitive noise + Unrelated news



Single News: Analytical Expressions of $\mathrm{Var}(P_t|t \geq 0)$ and t^*

Derivation

$$\mathrm{Var}(P_t|t \geq 0) = \overbrace{\left[\frac{\rho_c^{2(t+1)} - e^{-2(t+1)\mathcal{D}}}{\rho_c^2 - e^{-2\mathcal{D}}}\right]}^{\lim_{t\to\infty} \text{ is } 0} \sigma_c^2 + t\sigma_n^2 \tag{5}$$

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

Single News: Analytical Expressions of $\mathrm{Var}(P_t|t \geq 0)$ and t^*

Derivation

$$\mathrm{Var}(P_t|t \geq 0) = \overbrace{\left[\frac{\rho_c^{2(t+1)} - e^{-2(t+1)\mathcal{D}}}{\rho_c^2 - e^{-2\mathcal{D}}}\right]}^{\lim_{t\to\infty} \text{ is } 0} \sigma_c^2 + t\sigma_n^2 \tag{5}$$

$$\implies t^{one} : \mathcal{D}\left[e^{-2(t+1)\mathcal{D}}\right] + ln(\rho_c)\rho_c^{2(t+1)} = \left[\frac{(e^{-2\mathcal{D}} - \rho_c^2)}{2}\right]\frac{\sigma_n^2}{\sigma_c^2} \tag{6}$$

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

Single News: Analytical Expressions of $\mathrm{Var}(P_t|t \geq 0)$ and t^*

Derivation

$$\mathrm{Var}(P_t|t \geq 0) = \overbrace{\left[\frac{\rho_c^{2(t+1)} - e^{-2(t+1)\mathcal{D}}}{\rho_c^2 - e^{-2\mathcal{D}}}\right]}^{\lim_{t\to\infty} \text{ is } 0} \sigma_c^2 + t\sigma_n^2 \tag{5}$$

$$\implies t^{one} : \mathcal{D}\left[e^{-2(t+1)\mathcal{D}}\right] + ln(\rho_c)\rho_c^{2(t+1)} = \left[\frac{(e^{-2\mathcal{D}} - \rho_c^2)}{2}\right]\frac{\sigma_n^2}{\sigma_c^2} \tag{6}$$

$$\implies \frac{\partial t^{one}}{\partial \sigma_n^2} < 0, \frac{\partial t^{one}}{\partial \sigma_c^2} > 0^\dagger$$

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

Multiple News: Estimator Form

- Current expressions for $\mathrm{Var}(P_t|t \geq 0), t^{one}$: One news event
- **Problem**: N announcements and one asset price:
- **Goal**: Choose time horizon t^* such that

$$t^* : \min_t \sum_{i=1}^{N} \frac{1}{N} \left(P_{i,t} - P_{i,t}^f\right)^2$$

- However, assume $P_{i,t}^f$ is unobservable. Instead, noisy signal $s_i = P_i^f + \xi_i$ is observed
  - $\xi_i \sim \mathcal{N}(0, \sigma_s^2)$

Multiple News: MSE Minimisation Problem with Signal

Derivation

$$t^* : \min_t \frac{1}{N} \sum_{i=1}^{N} (P_{i,t} - s_i)^2 \implies \min_t \left[\frac{1}{N} \sum_{i=1}^{N} \left(\varepsilon^c_{i,t} + \varepsilon^n_{i,t} \right)^2 + \sigma_s^2 \right] \quad (7)$$

- With noisy signal s_i, MSE minimisation problem is the same as that with $P^f_{i,t}$
  - Asymptotic result: Quality of signal doesn't matter
- $\implies$ Possible to estimate optimal t^* $(\hat{t})$ with s_i
  - Small samples: Precision of s_i matters $\rightarrow$ "good" signal matters

Conceptual Framework Takeaways

- Simulated MSEs using $P^f_{i,t}, s_i$ for different market scenarios Simulations
  - Scenario 1 $\sim$ High presence of cognitive noise, little unrelated news
  - Scenario 2 $\sim$ Little cognitive noise, high presence of unrelated news
  - Scenario 3 $\sim$ Presence of both cognitive noise and unrelated news
- $\hat{t} \approx t^*$ in all scenarios
  - "Good" signal $\longrightarrow$ Possible to estimate time horizon reflecting market full reactions
  - MP shocks $=$ Small sample problem $\longrightarrow$ "Good" signal matters

Conceptual Framework Takeaways

- Simulated MSEs using $P^f_{i,t}, s_i$ for different market scenarios Simulations
  - Scenario 1 $\sim$ High presence of cognitive noise, little unrelated news
  - Scenario 2 $\sim$ Little cognitive noise, high presence of unrelated news
  - Scenario 3 $\sim$ Presence of both cognitive noise and unrelated news
- $\hat{t} \approx t^*$ in all scenarios
  - "Good" signal $\longrightarrow$ Possible to estimate time horizon reflecting market full reactions
  - MP shocks $=$ Small sample problem $\longrightarrow$ "Good" signal matters
- **Q: How to get "good" signal for MP announcements?**
  - How to approximate relationship from FOMC statement text to asset price changes?

Presentation Roadmap

Estimating Optimal Event Windows from FOMC Statements: Overview

1. Apply text-analysis neural network from computer science literature
   - Isolates Δasset prices within given event window to "full" FOMC statement text

$\Longrightarrow$ "Using only the entire FOMC statement, what is your predicted price change?"

ΔAsset Prices

FOMC Statement: Monetary Policy, Target FFR Range, LSAP, Forward Guidance, Analysis of Current Economy, etc.

Other Stuff: Cognitive Noise, Unrelated News

Estimating Optimal Event Windows from FOMC Statements: Overview

1. Apply text-analysis neural network from computer science literature
    - Isolates Δasset prices within given event window to "full" FOMC statement text

   $\Longrightarrow$ "Using only the entire FOMC statement, what is your predicted price change?"

2. Regress Δasset prices within different event windows on FOMC statements

Estimating Optimal Event Windows from FOMC Statements: Overview

1. Apply text-analysis neural network from computer science literature
   - Isolates Δasset prices within given event window to "full" FOMC statement text

   $\Longrightarrow$ "Using only the entire FOMC statement, what is your predicted price change?"
2. Regress Δasset prices within different event windows on FOMC statements
3. Find event window where $\widehat{\Delta\text{asset prices}}$ is closest to Δasset prices
   - Optimal window around FOMC statements: Δasset prices has min noise on avg

Estimating Optimal Event Windows: Variables and Approach

Input/Output Visual

- **Approach**: Approximate $f(\text{Inputs}) = \text{Outputs}$
  - Nonparametric regression approximated by many linear + non-linear combos

Estimating Optimal Event Windows: Variables and Approach

Input/Output Visual

- **Approach**: Approximate $f(\text{Inputs}) = \text{Outputs}$
  - Nonparametric regression approximated by many linear + non-linear combos
- **Inputs**: FOMC statements from scheduled FOMC meetings FOMC Statement Text Prep
  - 165 statements from May 1999 - October 2019 FOMC Statement Ex

Estimating Optimal Event Windows: Variables and Approach

Input/Output Visual

- **Approach**: Approximate $f(\text{Inputs}) = \text{Outputs}$
  - Nonparametric regression approximated by many linear + non-linear combos
- **Inputs**: FOMC statements from scheduled FOMC meetings

  FOMC Statement Text Prep
  - 165 statements from May 1999 - October 2019

  FOMC Statement Ex
- **Output**: $DP_{t+n} = ln\left(\frac{P_{t+n}}{P_{t-10}}\right)$ for interest-rate and equity futures

  Futures Overview
  - Price lvls at 10-min-intervals from 10 min before - 18 hrs after statement release
  - Fed Fund Futures: $FF1, FF2, FF3, FF4$
  - Eurodollar Futures: $EDcm2, EDcm3, EDcm4$
  - 2-Year Treasury Futures: $TUc1, TUc2$
  - 5-Year Treasury Futures: $FVc1, FVc2$
  - 10-Year Treasury Futures: $TYc1,\ TYc2$
  - 30-year Treasury Futures: $USc1, USc2$
  - S&P 500 Index and E-mini Futures: $SPX, ESc1, ESc2$

Estimating Optimal Event Windows: Approach

- **At the core**: $f(\text{FOMC statement text}) = DP_{t+n}$: Nonparametric mapping

Estimating Optimal Event Windows: Approach

- **At the core**: f(FOMC statement text) $= DP_{t+n}$: Nonparametric mapping
- Popular text analysis methods in empirical macro:
  - "Fitting predictive models on simple counts of text features" (Gentzkow et al., 2019)

Estimating Optimal Event Windows: Approach

- **At the core**: f(FOMC statement text) = DP_{t+n}: Nonparametric mapping
- Popular text analysis methods in empirical macro:
  - "Fitting predictive models on simple counts of text features" (Gentzkow et al., 2019)
- Popular methods cannot realistically:

  Popular Method Ex Issues
  - Quantify all info. in text (e.g., Word interdependencies, context, long-term memory)

$\Longrightarrow$ Approximate "full" relationship from FOMC statement text to asset price changes

Estimating Optimal Event Windows: Approach

- **At the core**: f(FOMC statement text) = DP_{t+n}: Nonparametric mapping
- Popular methods cannot quantify "full" FOMC statement
- **Foundation**: Text-analysis neural network XLNet (Yang et al., 2019) can quantify:
  - Transfer learning: Fine-tune pre-trained XLNet on FOMC language
  - Features: Bi-directional learning, recurrency memory, permutation modelling
    UAT + Layers XLNet Details Addressing Look-ahead Bias
  - Represent entire text numerically for diff tasks (e.g., Gmail/Google, academia)

$\Longrightarrow$ "Good" signal based on FOMC statement text

Estimating Optimal Event Windows: Prepare for Train + Eval XLNet

- **Goal**: "Good" signal from XLNet for every FOMC statements
  - Method from ML literature: Train XLNet on splits/folds of data

Estimating Optimal Event Windows: Prepare for Train + Eval XLNet

- **Goal**: "Good" signal from XLNet for every FOMC statements
  - Method from ML literature: Train XLNet on splits/folds of data
- Split data into training (132) and testing (33) samples:
  - By stratified sampling k-fold cross validation Why Stratified? Stratified Visual Why CV?
  - Conditions: Word count (200-word bins), FFR decision, FOMC chair, pre/post 2007
  - $k = 5$
  - Every testing subsample share NO FOMC statements

Estimating Optimal Event Windows: Prepare for Train + Eval XLNet

- **Goal**: "Good" signal from XLNet for every FOMC statements
  - Method from ML literature: Train XLNet on splits/folds of data
- Split data into training (132) and testing (33) samples:
  - By stratified sampling k-fold cross validation Why Stratified? Stratified Visual Why CV?
  - Conditions: Word count (200-word bins), FFR decision, FOMC chair, pre/post 2007
  - $k = 5$
  - Every testing subsample share NO FOMC statements
- **Result**: XLNet learns $f(\text{FOMC statement text}) = DP_{t+n}$ for each fold:
  - With equal dist. of FOMC statements based on characteristics

Estimating Optimal Window Lengths: Accuracy Metrics

- For each fold, primary metric to judge NN = generalised $R^2 := R^2_{OOS}$ R^2 Details
- Make adjustments from typical definition because:

Estimating Optimal Window Lengths: Accuracy Metrics

- For each fold, primary metric to judge NN = generalised $R^2 := R^2_{OOS}$ R^2 Details
- Make adjustments from typical definition because:
  1. NN is a non-linear regression $\implies \rho^2 \neq R^2$

Estimating Optimal Window Lengths: Accuracy Metrics

- For each fold, primary metric to judge NN = generalised $R^2 := R^2_{OOS}$ R^2 Details
- Make adjustments from typical definition because:
  1. NN is a non-linear regression $\implies \rho^2 \neq R^2$
  2. Judging out-of-sample performance, not in-sample

Estimating Optimal Window Lengths: Accuracy Metrics

- For each fold, primary metric to judge NN = generalised $R^2 := R^2_{OOS}$ R^2 Details
- Make adjustments from typical definition because:
  1. NN is a non-linear regression $\implies \rho^2 \neq R^2$
  2. Judging out-of-sample performance, not in-sample
- **Other tracked metrics**: ρ_{OOS}, $\widehat{MAE_{OOS}}$, $\widehat{MSE_{IS}}$

Estimating Optimal Window Lengths: Loop "Diagram"

For each interest-rate and equity futures contract:

Estimating Optimal Window Lengths: Loop "Diagram"

For each interest-rate and equity futures contract:

- For each DP_{t+n} up to $t + 60$:

Estimating Optimal Window Lengths: Loop "Diagram"

For each interest-rate and equity futures contract:

- For each DP_{t+n} up to $t+60$:
  - For each $\mathbf{k} = \mathbf{5}$ fold:

1. Fine-tune NN parameters and hyperparameters to fit training data
   NN Training Overview Hyperparameter Tuning Addressing Look-ahead Bias
2. Evaluate NN on testing data $\longrightarrow$ Choose hyperparameters that yield highest R^2_{OOS}

Estimating Optimal Window Lengths: Loop "Diagram"

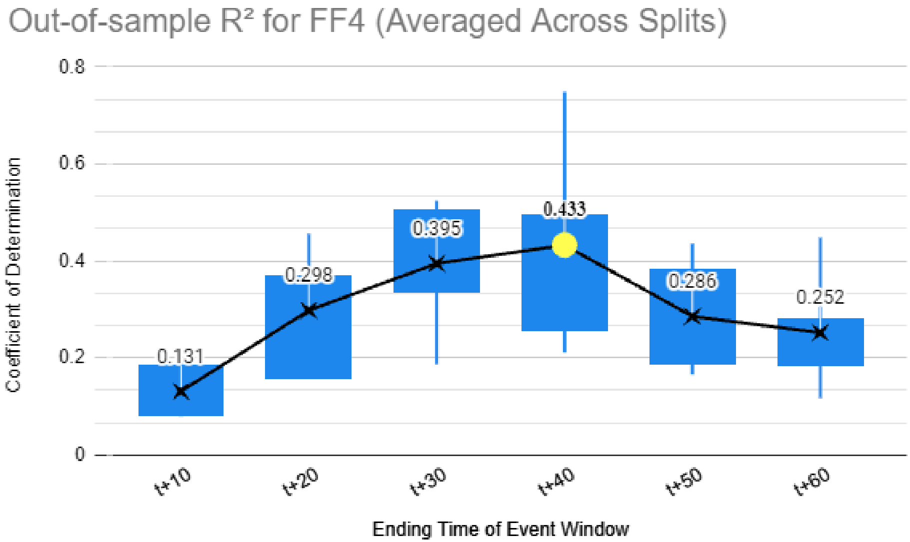
For each interest-rate and equity futures contract:

- For each DP_{t+n} up to $t + 60$:
  - For each **k** = **5** fold:

1. Fine-tune NN parameters and hyperparameters to fit training data
   NN Training Overview | Hyperparameter Tuning | Addressing Look-ahead Bias
2. Evaluate NN on testing data → Choose hyperparameters that yield highest R^2_{OOS}
3. **Final Output**: $\overline{R^2_{OOS}}$:= Average R^2_{OOS} across k folds
   - Other R^2_{OOS} metrics: Min, max, $75^{th}, 25^{th}$ prctiles

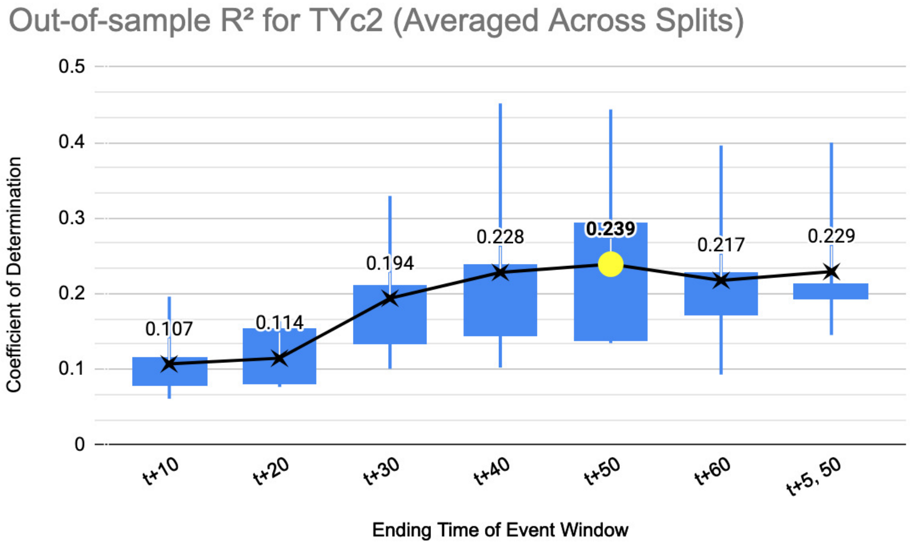
Optimal Event Windows: *FF*4

Summary Visual One Signal



Estimating Optimal Event Windows: *TYc2*

Summary Visual One Signal

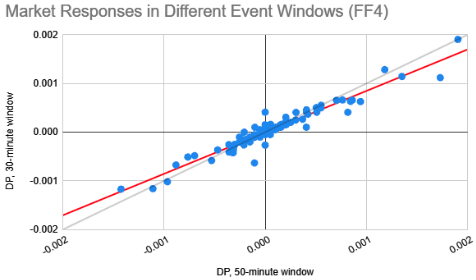


Optimal Event Windows: Summary

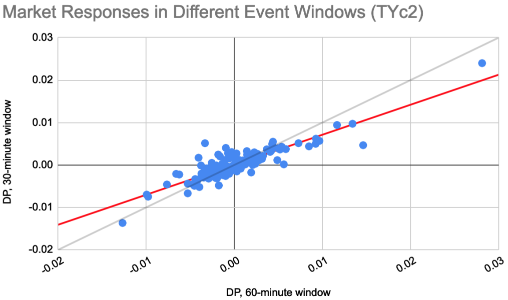
- **How Long?** Longer than 30 minutes: Other Assets Summary Visual Summary Table Recap
  - On avg, markets fully react within window 10 min before and 30+ min after
  - $\overline{R^2_{OOS}} \uparrow$ by 2–17% when event window $\uparrow$ to 40+ min
  - Time horizon of assets $\uparrow \rightarrow$ Avg optimal window length $\uparrow$
  - Time horizon of asset at least 2 quarters out $\rightarrow$ 50- to 60-min window

Optimal Event Windows: Diff Windows, Diff Responses

Other Assets



(a) *FF4****



(b) *TYc2****

- **Takeaway**: On avg, markets under-react, ex-post, to FOMC statement text

Presentation Roadmap

Estimating Optimal Event Windows: “Joint” and “One Signal” Approaches

- **Recap**: XLNet approx f(FOMC Statement Text) = DP_{t+5}, $\forall$Folds of $\forall DP_{t+n}$
  - $\longrightarrow$ “Joint” estimation of signal and optimal event window length

[‡]Signal from XLNet is likely to change $\because$ Changing LHS DP_{t+n} $\rightarrow$ retraining NN + “Joint” estimation was performed on “general” sample of FOMC statements, not specific types of statements.

Estimating Optimal Event Windows: "Joint" and "One Signal" Approaches

- **Recap**: XLNet approx $f(\text{FOMC Statement Text}) = DP_{t+5}$, $\forall$Folds of $\forall DP_{t+n}$
  - $\rightarrow$ "Joint" estimation of signal and optimal event window length

‡Signal from XLNet is likely to change $\because$ Changing LHS $DP_{t+n} \rightarrow$ retraining NN + "Joint" estimation was performed on "general" sample of FOMC statements, not specific types of statements.

Estimating Optimal Event Windows: "Joint" and "One Signal" Approaches

- **Recap**: XLNet approx f(FOMC Statement Text) = DP_{t+5}, $\forall$Folds of $\forall DP_{t+n}$

  $\longrightarrow$ "Joint" estimation of signal and optimal event window length

- Fine-tuning XLNet for "joint" estimation = Computationally intensive
  - GPU + Financial constraints = Estimate optimal window lengths only up to $\mathbf{t+60}$
  - Current computation time: 249+ days

‡Signal from XLNet is likely to change $\because$ Changing LHS DP_{t+n} $\rightarrow$ retraining NN + "Joint" estimation was performed on "general" sample of FOMC statements, not specific types of statements.

Estimating Optimal Event Windows: "Joint" and "One Signal" Approaches

- **Recap**: XLNet approx f(FOMC Statement Text) = DP_{t+5}, $\forall$Folds of $\forall DP_{t+n}$
  - $\rightarrow$ "Joint" estimation of signal and optimal event window length
- Fine-tuning XLNet for "joint" estimation = Computationally intensive
  - GPU + Financial constraints = Estimate optimal window lengths only up to $\mathbf{t+60}$
  - Current computation time: 249+ days
- **Assumption**: NN Predictions in "joint-estimated" event window = Constant $\forall t$[‡]
  1. Much less computationally intensive
  2. Can check if FOMC statement characteristics affect optimal window length
  3. Can check if $\exists$ greater out-of-sample R^2_{OOS} for $t+n > t+60$

Robustness Check

[‡] Signal from XLNet is likely to change $\because$ Changing LHS DP_{t+n} $\rightarrow$ retraining NN + "Joint" estimation was performed on "general" sample of FOMC statements, not specific types of statements.

Effects of FOMC Statement Characteristics on Event Windows

- Use "one signal" approach to compare MSEs computed based on:
  1. Complexity of FOMC statements
  2. Similarity of FOMC statements
  3. Presence of Dissents

FOMC Statement Characteristics: Text Complexity (1/3)

- Condition FOMC statements based on text complexity
  - Hernandez-Murillo and Shell (2014); Hansen and McMahon (2016); Smales and Apergis (2017); Haldane and McMahon (2018); De Pooter (2021); McMahon (2023); and others...

FOMC Statement Characteristics: Text Complexity (1/3)

- Condition FOMC statements based on text complexity
  - Hernandez-Murillo and Shell (2014); Hansen and McMahon (2016); Smales and Apergis (2017); Haldane and McMahon (2018); De Pooter (2021); McMahon (2023); and others...
- Measured based on Flesch Kincaid Grade Level
  - Based on sentence structure, word structure, and word phonology

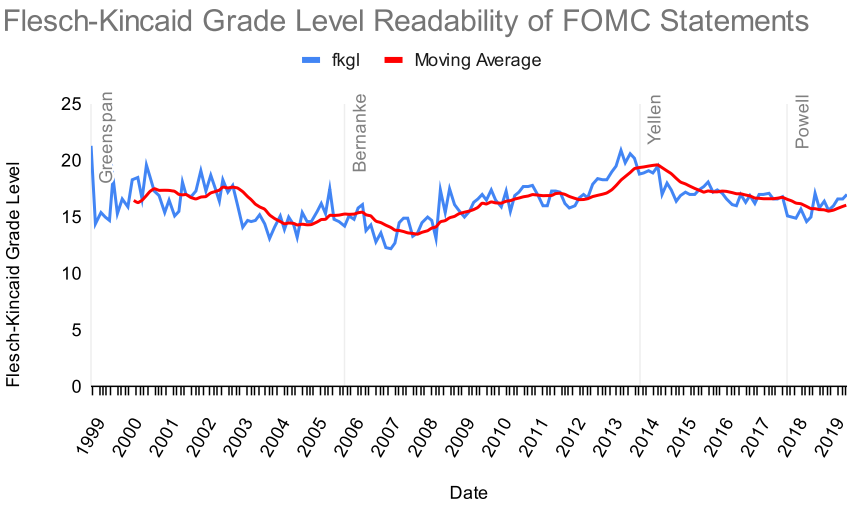
FOMC Statement Characteristics: Text Complexity (1/3)

- Condition FOMC statements based on text complexity
  - Hernandez-Murillo and Shell (2014); Hansen and McMahon (2016); Smales and Apergis (2017); Haldane and McMahon (2018); De Pooter (2021); McMahon (2023); and others...
- Measured based on Flesch Kincaid Grade Level
  - Based on sentence structure, word structure, and word phonology
  - Range of reading Levels: 12.2–21.3
  - Median Reading Level: 16.5

FOMC Statement Characteristics: Text Complexity (1/3)

- Condition FOMC statements based on text complexity
  - Hernandez-Murillo and Shell (2014); Hansen and McMahon (2016); Smales and Apergis (2017); Haldane and McMahon (2018); De Pooter (2021); McMahon (2023); and others...
- Measured based on Flesch Kincaid Grade Level
  - Based on sentence structure, word structure, and word phonology
  - Range of reading Levels: 12.2–21.3
  - Median Reading Level: 16.5
- Split sample conditioned on being <= or > 16.5
- Calculate sub-set MSEs and event window lengths

FOMC Statement Characteristics: Text Complexity (2/3)



FOMC Statement Characteristics: Text Complexity (3/3)

Metric	Simple	Complicated
Minimised MSE		
Average	1.26e-5	1.03e-5
Event Window Length (Minutes)		
Average	59	71

Table 1: Complexity of FOMC statements measured by the Flesch-Kincaid Grade level, defined as: $FKGL = 0.39 \times \text{average sentence length} + 11.8 \times \text{average number of syllables per word} - 15.59$. "Simple" are statements with grade level up to 16.5. "Complicated" are statements with grade levels above 16.5. In order to lessen the effects of outliers, the event window length for the 3-month-ahead federal funds future under the "one signal" approach is reduced from its original value and set to equal the median of the sub-set window lengths for the asset type.

$\longrightarrow$ FOMC statements with $\uparrow$ complexity $\rightarrow$ Longer event window on avg

FOMC Statement Characteristics: Text Similarity (1/4)

- Condition FOMC statements based on text similarity
  - Acosta and Meade (2015); Handlan (2022a); and others...

FOMC Statement Characteristics: Text Similarity (1/4)

- Condition FOMC statements based on text similarity
  - Acosta and Meade (2015); Handlan (2022a); and others...
- Measured based on Term Frequency-Inverse Document Frequency (TFIDF)
  - Weighted frequency assigned to terms based on: TFIDF Equation
    1. Number of times term appears in a document
    2. Number of documents terms appears in

FOMC Statement Characteristics: Text Similarity (1/4)

- Condition FOMC statements based on text similarity
  - Acosta and Meade (2015); Handlan (2022a); and others...
- Measured based on Term Frequency-Inverse Document Frequency (TFIDF)
  - Weighted frequency assigned to terms based on: TFIDF Equation
    1. Number of times term appears in a document
    2. Number of documents terms appears in
- Terms with ↑ $TFIDF_{d,t}$ = Informative terms at distinguishing documents d

TFIDF Terms

FOMC Statement Characteristics: Text Similarity (2/4)

- TFIDF matrix dimensions: D documents $\times$ T terms

FOMC Statement Characteristics: Text Similarity (2/4)

- TFIDF matrix dimensions: D documents $\times$ T terms
- Row vector: All terms t of document d represented by $TFIDF_{d,t}$ values

FOMC Statement Characteristics: Text Similarity (2/4)

- TFIDF matrix dimensions: D documents $\times$ T terms
- Row vector: All terms t of document d represented by $TFIDF_{d,t}$ values
- All row vectors normalised by number of d and t in matrix

FOMC Statement Characteristics: Text Similarity (2/4)

- TFIDF matrix dimensions: D documents $\times$ T terms
- Row vector: All terms t of document d represented by $TFIDF_{d,t}$ values
- All row vectors normalised by number of d and t in matrix
- $\implies$ $TFIDF \cdot TFIDF^T$ = Dot product between every pair of FOMC statements

FOMC Statement Characteristics: Text Similarity (2/4)

- TFIDF matrix dimensions: D documents $\times$ T terms
- Row vector: All terms t of document d represented by $TFIDF_{d,t}$ values
- All row vectors normalised by number of d and t in matrix
- $\implies$ $TFIDF \cdot TFIDF^T$ = Dot product between every pair of FOMC statements
- $\implies$ Degree of similarity between 2 FOMC statements $=$ Cosine similarity:

Similarity Matrix

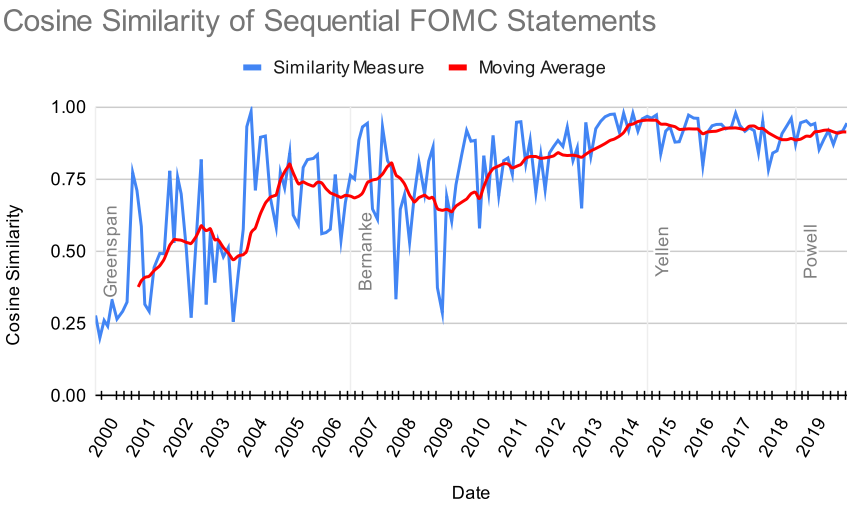
$$S^{A,B} := \cos\theta = \frac{\mathbf{A}\cdot\mathbf{B}}{||\mathbf{A}||\,||\mathbf{B}||},$$

- **Scale**: Entirely different $= 0 \leq$ Cosine Similarity $\leq 1 =$ Exact same

FOMC Statement Characteristic: Text Similarity (3/4)

- $S^1 := (d, d-1)$: Degree of similarity between sequential FOMC statements
  - Range of S^1: 0.02–0.984
  - Median of S^1: 0.826
- Split sample conditioned on being $\leq$ or $>$ 0.826
- Calculate sub-set MSEs and event window lengths

FOMC Statement Characteristic: Text Similarity (3/4)



FOMC Statement Characteristics: Text Similarity (4/4)

Metric	Different	Similar
Minimised MSE		
Average	1.14e-5	1.14e-5
Event Window Length (Minutes)		
Average	61	51

Table 2: Changes in FOMC statements are measured using a pairwise-statement cosine dissimilarity measure. The event window lengths are displayed in minutes. "Different" are sequential statements with a cosine similarity of less than to 0.885. "Similar" are sequential statements with a cosine similarity of more than 0.885.

→ ↓ Similar FOMC statements → Longer event windows on avg

FOMC Statement Characteristics: Presence of Dissents (1/2)

- Condition FOMC statements based presence of dissent votes or not[§]
  - Riboni and Ruge-Murcia (2014); Tsang and Yang (2024); Bobrov et al. (2025); and others...

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

FOMC Statement Characteristics: Presence of Dissents (1/2)

- Condition FOMC statements based presence of dissent votes or not[§]
  - Riboni and Ruge-Murcia (2014); Tsang and Yang (2024); Bobrov et al. (2025); and others...
- Roughly 40% of FOMC statement sample has recorded dissents
- By Fed tradition, dissents usually recorded if majority opinion = unacceptable
- Presence of dissents provides additional info. for markets to process

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

FOMC Statement Characteristics: Presence of Dissents (1/2)

- Condition FOMC statements based presence of dissent votes or not[§]
  - Riboni and Ruge-Murcia (2014); Tsang and Yang (2024); Bobrov et al. (2025); and others...
- Roughly 40% of FOMC statement sample has recorded dissents
- By Fed tradition, dissents usually recorded if majority opinion = unacceptable
- Presence of dissents provides additional info. for markets to process
- Calculate sub-set MSEs and event window lengths

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

FOMC Statement Characteristics: Presence of Dissents (2/2)

Metric	Unity	Dissents
Minimised MSE		
Average	9.21e-6	1.44e-5
Event Window Length (Minutes)		
Average	61	83

Table 3: Changes in FOMC statements are measured using a pairwise-statement cosine dissimilarity measure. The event window lengths are displayed in minutes. "Unity" statements are those without votes of dissent. "Dissents" are statements with recorded dissent votes.

→ FOMC statements with dissents → longer event windows on avg

Overall Recap

- Optimal event window lengths longer than 30 minutes
- Diff time horizons of assets → Diff optimal windows
- Complex/dissimilar/dissent statements → Relative longer windows

→ **What happens to MP surprises and shocks?**

Presentation Roadmap

Monetary Policy Surprises: Overview

- Found optimal event window lengths: 40-, 50-, 60-minutes

1. Pick an interest-rate futures contract

Monetary Policy Surprises: Overview

- Found optimal event window lengths: 40-, 50-, 60-minutes

1. Pick an interest-rate futures contract
2. Construct MP surprises within 30-minute and optimal windows

DP → Surprise

Monetary Policy Surprises: Overview

- Found optimal event window lengths: 40-, 50-, 60-minutes

1. Pick an interest-rate futures contract
2. Construct MP surprises within 30-minute and optimal windows   *DP* → Surprise
3. Calculate ρ between MP surprise sets

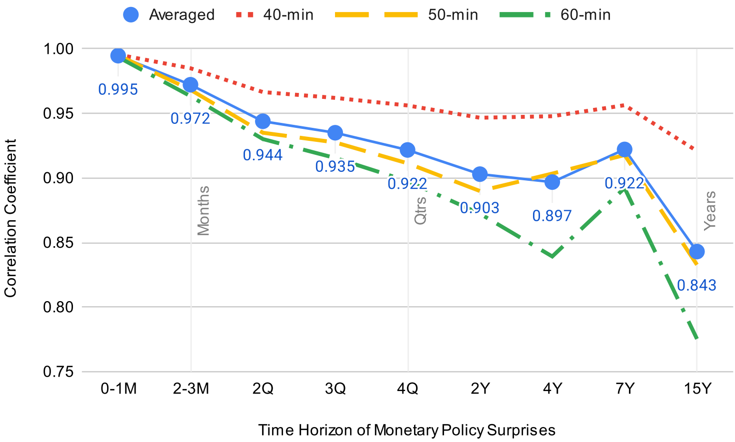
Monetary Policy Surprises: Overview

- Found optimal event window lengths: 40-, 50-, 60-minutes

1. Pick an interest-rate futures contract
2. Construct MP surprises within 30-minute and optimal windows

   DP → Surprise
3. Calculate ρ between MP surprise sets
4. Back to step 1

MP Surprises: ρ Along the Yield Curve



→ Changing only window length has ↑ effect at farther horizons

MP Shocks: Nominal Interest Rates

	Both	30-minute Window			Optimal Window			**Difference**		
	ΔTY_1	ΔTY_2	ΔTY_5	ΔTY_{10}	ΔTY_2	ΔTY_5	ΔTY_{10}	ΔTY_2	ΔTY_5	ΔTY_{10}
GSS_T	1.00***	0.82***	0.15	-0.37	0.78***	0.08	-0.42	**-0.04**	**-0.07**	**-0.05**
	(0.29)	(0.38)	(0.51)	(0.53)	(0.31)	(0.41)	(0.42)	**(-0.06)**	**(-0.11)**	**(-0.11)**
GSS_P	1.00***	1.46***	1.89***	1.64***	1.51***	1.92***	1.66***	**+0.05**	**+0.04**	**+0.02**
	(0.11)	(0.12)	(0.26)	(0.35)	(0.09)	(0.20)	(0.29)	**(-0.03)**	**(-0.05)**	**(-0.06)**
NS_{MP}	1.00***	1.24***	1.29***	0.94***	1.30***	1.39***	1.06***	**+0.06**	**+0.11**	**+0.11**
	(0.09)	(0.12)	(0.21)	(0.25)	(0.13)	(0.21)	(0.25)	**(+0.01)**	**(-0.00)**	**(+0.01)**
JK_{MP}	1.00***	1.30***	1.39***	0.99***	1.35***	1.52***	1.16***	**+0.04**	**+0.13**	**+0.17**
	(0.14)	(0.18)	(0.28)	(0.33)	(0.16)	(0.30)	(0.39)	**(-0.02)**	**(+0.02)**	**(+0.06)**
JK_{CBI}	1.00***	1.04***	1.00***	0.82***	1.20***	1.14***	0.85***	**+0.16**	**+0.14**	**+0.03**
	(0.31)	(0.37)	(0.39)	(0.34)	(0.22)	(0.26)	(0.27)	**(-0.15)**	**(-0.13)**	**(-0.07)**

Table 4: Diff in Responses of Nominal Interest Rates to Shocks from Event Window Choice

$\longrightarrow$ Using optimal window length $\rightarrow\uparrow$ Effects for MP shocks about forward guidance

MP Shocks: Real Interest Rates

	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**
	(1.66)	(0.65)	(0.58)	(1.72)	(0.53)	(0.46)	**(+0.05)**	**(-0.12)**	**(-0.13)**
GSS_P	2.21***	1.96***	1.74***	2.20***	2.03***	1.75***	**-0.00**	**+0.06**	**+0.01**
	(0.49)	(0.40)	(0.40)	(0.36)	(0.32)	(0.33)	**(-0.13)**	**(-0.08)**	**(-0.07)**
NS_{MP}	1.17***	1.29***	1.08***	1.31***	1.47***	1.20***	**+0.14**	**+0.18**	**+0.13**
	(0.80)	(0.30)	(0.27)	(0.63)	(0.27)	(0.26)	**(-0.17)**	**(-0.02)**	**(-0.00)**
JK_{MP}	1.40***	1.40***	1.15***	1.66***	1.64***	1.38***	**+0.26**	**+0.24**	**+0.23**
	(0.92)	(0.39)	(0.35)	(0.66)	(0.42)	(0.41)	**(-0.27)**	**(+0.03)**	**(+0.05)**
JK_{CBI}	0.51	0.99***	0.85***	0.60	1.13***	0.84***	**+0.09**	**+0.14**	**-0.01**
	(0.87)	(0.37)	(0.29)	(0.85)	(0.33)	(0.26)	**(-0.02)**	**(-0.04)**	**(-0.02)**

Table 5: Diff in Responses of Real Interest Rates to Shocks from Event Window Choice

→ Using optimal window length →↑ Effects for MP shocks about forward guidance

MP Shocks: Stock Prices

	$DP_{SPX,t+20}$	$DP_{SPX,t+40}$	**Difference**
GSS_T	-8.40^{***}	-7.39^{***}	**+1.01**
	(2.78)	(2.95)	**(+0.17)**
GSS_P	-6.14^{***}	-6.85^{***}	**-0.71**
	(1.69)	(2.61)	**(+0.92)**
NS_{MP}	-6.92^{***}	-7.00^{***}	**-0.09**
	(1.27)	(1.84)	**(+0.57)**
JK_{MP}	-14.76^{***}	-17.46^{***}	**-2.69**
	(0.74)	(1.03)	**(+0.28)**
JK_{CBI}	15.19^{***}	14.08^{***}	**-1.12**
	(2.07)	(2.07)	**(-0.00)**

Table 6: Diff in Responses of Stock Prices to Shocks from Event Window Choice

$\rightarrow$ Using optimal window length $\rightarrow\uparrow$ Effects for MP shocks about forward guidance

Conclusion

- **This Paper**: Estimate optimal window size for FOMC statements using NLP:
  - By combining text-based signal with observed price dynamics
  - By isolating market price changes to "full" text of FOMC statements
- **How Long?** Longer than 30 minutes:
  - On avg, markets fully react within window 10 min before and 30+ min after
  - Time horizon of assets $\uparrow \rightarrow$ Avg optimal window length $\uparrow$
  - Time horizon of asset at least 2 quarters out $\rightarrow$ 50- to 60-min window
  - Complex/dissimilar/dissent statements $\rightarrow$ Relatively longer windows
- **Effects**: By changing only event window choice:
  - Time horizon of assets $\uparrow \rightarrow$ corr. between MP surprise sets $\downarrow$
  - MP shocks about forward guidance have $\uparrow$ impact on yields and stock prices

Next Steps

- **Next steps**:
  1. Estimate optimal event window lengths for other MP communication
  2. Analyse how deeper changes in MP communication affect optimal windows

Thank you!

pltran@utexas.edu

https://paulletran.com/

References I

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

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*. Working Paper 25-03. Federal Reserve Bank of Kansas City.

Antweiler, Werner and Murray A. Frank (2022). "Do US Stock Markets Typically Overreact to Corporate News Stories?" SSRN Working Paper No 878091.

Aruoba, S. Borağan and Thomas Drechsel (2024). *Identifying Monetary Policy Shocks: A Natural Language Approach*. Working Paper 32417. National Bureau of Economic Research.

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

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

Ben-David, Itzhak, Francesco Franzoni, Byungwook Kim, and Rabih Moussawi (2022). "Competition for Attention in the ETF Space". In: *The Review of Financial Studies* 36.3, pp. 987–1042. eprint: https://academic.oup.com/rfs/article-pdf/36/3/987/49288527/hhac048.pdf.

Bernard, Victor L. and Jacob K. Thomas (1989). "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?" In: *Journal of Accounting Research* 27, pp. 1–36.

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

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar (2009). "How Noise Trading Affects Markets: An Experimental Analysis". In: *The Review of Financial Studies* 22.6, pp. 2275–2302.

Bobrov, Anton, Rupal Kamdar, and Mauricio Ulate (2025). "Regional Dissent: Do Local Economic Conditions Influence FOMC Votes?" In: *American Economic Review: Insights* 7.2, pp. 268–84.

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

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

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2019). "Diagnostic Expectations and Stock Returns". In: *The Journal of Finance* 74.6, pp. 2839–2874. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12833.

References II

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020). "Overreaction in Macroeconomic Expectations". In: *American Economic Review* 110.9, pp. 2748–82.

Caivano, Valeria (2015). "The Impact of High-Frequency Trading on Volatility. Evidence from the Italian Market". CONSOB Working Papers No 80.

Chan, Louis K. C., Narasimhan Jegadeesh, and Josef Lakonishok (1996). "Momentum Strategies". In: *The Journal of Finance* 51.5, pp. 1681–1713. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1996.tb05222.x.

Chan, Wesley S. (2003). "Stock price reaction to news and no-news: drift and reversal after headlines". In: *Journal of Financial Economics* 70.2, pp. 223–260.

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.

Cieslak, Anna, Stephen Hansen, Michael McMahon, and Song Xiao (2023). *Policymakers' Uncertainty*. Working Paper 31849. National Bureau of Economic Research.

Cieslak, Anna and Michael McMahon (2023). "Tough Talk: The Fed and the Risk Premium". SSRN Working Paper No 4560220.

Cieslak, Anna and Annette Vissing-Jorgensen (2020). "The Economics of the Fed Put". In: *The Review of Financial Studies* 34.9, pp. 4045–4089.

Da, Zhi, Joseph Engelberg, and Pengjie Gao (2011). "In Search of Attention". In: *The Journal of Finance* 66.5, pp. 1461–1499. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2011.01679.x.

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.

De Bondt, Werner F. M. and Richard Thaler (1985). "Does the Stock Market Overreact?" In: *The Journal of Finance* 40.3, pp. 793–805. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1985.tb05004.x.

De Pooter, Michiel (2021). *Questions and Answers: The Information Content of the Post-FOMC Meeting Press Conference*. Tech. rep. FEDS Notes. Board of Governors of the Federal Reserve System.

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.

Gáti, Laura and Amy Handlan (2025a). "Monetary Communication Rules". ECB Working Paper No. 2022/2759.

Gáti, Laura and Amy Handlan (2025b). "Reputation for Confidence". Working Paper.

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

Gider, Jasmin, Simon Schmickler, and Christian Westheide (2019). "High-Frequency Trading and Price Informativeness". SAFE Working Paper No 248.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera (2023). "The Voice of Monetary Policy". In: *American Economic Review* 113.2, pp. 548–84.

References III

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.

Haldane, Andrew and Michael McMahon (2018). "Central Bank Communications and the General Public". In: *AEA Papers and Proceedings* 108, pp. 578–83.

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.

Hansen, Stephen and Michael McMahon (2016). "Shocking language: Understanding the macroeconomic effects of central bank communication". In: *Journal of International Economics* 99. 38th Annual NBER International Seminar on Macroeconomics, S114–S133.

Hansen, Stephen, Michael McMahon, and Andrea Prat (2017). "Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach". In: *The Quarterly Journal of Economics* 133.2, pp. 801–870. URL: https://doi.org/10.1093/qje/qjx045.

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

Hernandez-Murillo, Ruben and Hannah Shell (2014). "The Rising Complexity of the FOMC Statement". In: *Economic Synopses* 23.

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

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

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

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

La Porta, Rafael (1996). "Expectations and the Cross-Section of Stock Returns". In: *The Journal of Finance* 51.5, pp. 1715–1742.

Lucca, David O and Francesco Trebbi (2009). *Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements*. Working Paper 15367. National Bureau of Economic Research.

Lucca, David O. and Emanuel Moench (2015). "The Pre-FOMC Announcement Drift". In: *The Journal of Finance* 70.1, pp. 329–371.

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.

References IV

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

Riboni, Alessandro and Francisco Ruge-Murcia (2014). "Dissent in monetary policy decisions". In: *Journal of Monetary Economics* 66, pp. 137–154.

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

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

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

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

Weller, Brian M. (2017). "Does Algorithmic Trading Reduce Information Acquisition?" In: *The Review of Financial Studies* 31.6, pp. 2184–2226. eprint: https://academic.oup.com/rfs/article-pdf/31/6/2184/24833081/hhx137.pdf.

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. arXiv: 1906.08237.

Interpretations of Cognitive Noise

Back

1. Over-reaction: De Bondt and Thaler (1985); La Porta (1996); Antweiler and Frank (2022); Da et al. (2011); Bordalo, Gennaioli, La Porta, et al. (2019); Bordalo, Gennaioli, Ma, et al. (2020); Boguth et al. (2023); Bianchi et al. (2024)
2. Under-reaction: Bernard and Thomas (1989); LKC Chan et al. (1996); WS Chan (2003); Lucca and Moench (2015); Boguth et al. (2023); Bianchi et al. (2024)
3. Noise trading: Bloomfield et al. (2009); Hervé et al. (2019); Ben-David et al. (2022)
4. Algorithmic trading: Caivano (2015); Weller (2017); Gider et al. (2019); Bazzana and Collini (2020); Ben Ammar and Hellara (2022)

Derivation of $\mathrm{Var}(P_t|t \geq 0)$ and $\frac{\partial \mathrm{Var}(P_t|t\geq 0)}{\partial t}$ (1/2)

Back

$$\begin{aligned}
\mathrm{Var}(P_0) &= \mathrm{Var}(\varepsilon_0^c) + \mathrm{Var}(\varepsilon_0^n) \\
&= \sigma_c^2 \\
\mathrm{Var}(P_1) &= \mathrm{Var}(\varepsilon_1^c) + \mathrm{Var}(\varepsilon_1^n) \\
&= \sigma_c^2(\rho_c^2 + e^{-2\mathcal{D}}) + \sigma_n^2 \\
\mathrm{Var}(P_2) &= \mathrm{Var}(\varepsilon_2^c) + \mathrm{Var}(\varepsilon_2^n) \\
&= \sigma_c^2(\rho_c^4 + \rho_c^2 e^{-2\mathcal{D}} + e^{-4\mathcal{D}}) + 2\sigma_n^2 \\
&\vdots \\
\mathrm{Var}(P_t|t \geq 0) &= \left[\sum_{i=0}^{t} \rho_c^{2(t-i)} e^{-2\mathcal{D}i}\right] \sigma_c^2 + t\sigma_n^2 \\
\implies \mathrm{Var}(P_t|t \geq 0) &= \left[\frac{\rho_c^{2(t+1)} - e^{-2(t+1)\mathcal{D}}}{\rho_c^2 - e^{-2\mathcal{D}}}\right] \sigma_c^2 + t\sigma_n^2
\end{aligned}$$

Derivation of $\mathrm{Var}(P_t|t \geq 0)$ and $\frac{\partial \mathrm{Var}(P_t|t\geq 0)}{\partial t}$ (2/2)

Back

$$\mathrm{Var}(P_t|t \geq 0) = \left[\frac{\rho_c^{2(t+1)} - e^{-2(t+1)\mathcal{D}}}{\rho_c^2 - e^{-2\mathcal{D}}}\right]\sigma_c^2 + t\sigma_n^2$$

$$\frac{\partial \mathrm{Var}(P_t|t \geq 0)}{\partial t} = \left\{\frac{2\left[ln(\rho_c)\rho_c^{2(t+1)} + \mathcal{D}\left[(e^{-2(t+1)\mathcal{D}}\right]\right]}{\rho_c^2 - e^{-2\mathcal{D}}}\right\}\sigma_c^2 + \sigma_n^2$$

Derivation of MSE Minimisation Problem with Signal

Back

$$
\begin{aligned}
t^* : \min_t \frac{1}{N}\sum_{i=1}^{N}\left(P_{i,t}-s_i\right)^2 &= \min_t \frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon^c_{i,t}+\varepsilon^n_{i,t}-P^f_i-\xi_i\right)^2 \\
&= \min_t \frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon^c_{i,t}+\varepsilon^n_{i,t}-\xi_i\right)^2 \\
&= \min_t \frac{1}{N}\sum_{i=1}^{N}\left[\left(\varepsilon^c_{i,t}+\varepsilon^n_{i,t}\right)^2+\xi_i^2-2\xi_i\left(\varepsilon^c_{i,t}+\varepsilon^n_{i,t}\right)\right]^2 \\
&= \min_t \left\{\mathbb{E}\left[\left(\varepsilon^c_{i,t}+\varepsilon^n_{i,t}\right)^2\right]+\mathbb{E}\left[\xi_i^2\right]-2\,\mathbb{E}\left[\xi_i\right]\mathbb{E}\left[\left(\varepsilon^c_{i,t}+\varepsilon^n_{i,t}\right)\right]\right\} \\
\implies t^* : \min_t \frac{1}{N}\sum_{i=1}^{N}\left(P_{i,t}-s_i\right)^2 &= \min_t \left[\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon^c_{i,t}+\varepsilon^n_{i,t}\right)^2+\sigma_s^2\right] \qquad (8)
\end{aligned}
$$

Simulation Setup (1/3): Initial Conditions

Back

- $t = 0$: Release of one FOMC announcement
  - $P^f_{t,i} = P^f_i \in [-100, 100]$
  - $\varepsilon^c_{i,0} \in [-100, 100]$
  - $\varepsilon^n_{i,0} = 0$
  - $\sigma_s \in \mathbb{R}$

Simulation Setup (2/3): MSEs

Back

- For single news $i \in N = 10,000$:
  - Simulate $P_{i,t}$ (and components) and s_i up to $t = 100$
  - Calculate $\left(P_{i,t} - P_{i,t}^f\right)^2$ and $(P_{i,t} - s_i)^2$
- Across all N news:
  - Calculate MSEs $\sum_{i=1}^{N} \frac{1}{N}\left(P_{i,t} - P_{i,t}^f\right)^2$ and $\sum_{i=1}^{N} \frac{1}{N}\left(P_{i,t} - s_{i,t}\right)^2$
  - Calculate t^* and $\hat{t}$

Simulation Setup (3/3): Market Scenarios

Back

- Calculate $t^*, \hat{t}$ under 3 asset market scenarios:

	Scenario 1	Scenario 2	Scenario 3
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	$\in \mathbb{R}$	$\in \mathbb{R}$	$\in \mathbb{R}$

Table 7: Framework Parameters for Simulations

- Scenario 1 $\sim$ High presence of cognitive noise, little unrelated news
- Scenario 2 $\sim$ Little cognitive noise, high presence of unrelated news
- Scenario 3 $\sim$ Presence of both cognitive noise and unrelated news

Simulation Results

Back

	Scenario 1	Scenario 2	Scenario 3
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 8: Framework Parameters and Results from 10,000 Simulations

Preprocessing FOMC Statement Text

Back to Variables

- Remove:
  - URLs and hyperlinks from statement's HTML file
  - FOMC member voting record from end of statement
  - List of regional bank request approvals
  - Release timestamp (e.g., "For immediate release")
- Change:
  - Statement file type to text
  - Text coding into standardised UTF-8 format (e.g., change length of "-")
  - Spacing between words to be one space

Cleaned FOMC Statement (09/2006)

Back to Results Preview | Back to Variables

1. The Federal Open Market Committee decided today to keep its target for the federal funds rate at 5-1/4 percent.
2. The moderation in economic growth appears to be continuing, partly reflecting a cooling of the housing market.
3. Readings on core inflation have been elevated, and the high levels of resource utilization and of the prices of energy and other commodities have the potential to sustain inflation pressures
4. However, inflation pressures seem likely to moderate over time, reflecting reduced impetus from energy prices, contained inflation expectations, and the cumulative effects of monetary policy actions and other factors restraining aggregate demand.
5. Nonetheless, the Committee judges that some inflation risks remain.
6. The extent and timing of any additional firming that may be needed to address these risks will depend on the evolution of the outlook for both inflation and economic growth, as implied by incoming information.

Futures Contract Overview (1/2)

Back to Variables

- Federal Fund futures: For a given expiry month, on last day of expiry month, contract pays out 100 minus the average federal funds rate over the final month
- 2-year Treasure futures: Quarterly contracts that obligate seller to deliver Treasury notes within a range of maturities up to 2 years on the last day of the expiry month to the buyer (i.e., March, June, September, and December). Average maturity is 2 years (Gürkaynak, Kisacikoğlu, et al., 2020)
- 5-year Treasure futures: Quarterly contracts that obligate seller to deliver Treasury notes within a range of maturities up to 5 years on the last day of the expiry month to the buyer (i.e., March, June, September, and December). Average maturity is 4 years (Gürkaynak, Kisacikoğlu, et al., 2020)
- 10-year Treasure futures: Quarterly contracts that obligate seller to deliver Treasury notes within a range of maturities up to 10 years on the last day of the expiry month to the buyer (i.e., March, June, September, and December). Average maturity is 7 years (Gürkaynak, Kisacikoğlu, et al., 2020)

Futures Contract Overview (2/2)

Back to Variables

- 30-year Treasure futures: Quarterly contracts that obligate seller to deliver Treasury notes within a range of maturities up to 30 years on the last day of the expiry month to the buyer (i.e., March, June, September, and December). Average maturity is 15 years (Gürkaynak, Kisacikoğlu, et al., 2020)
- Eurodollar futures: Quarterly contracts which pay out 100 minus the 3-month US dollar British Banks Association London Interbank Offered Rate (BBA LIBOR) interest rate at the time of expiry month (i.e., March, June, September, and December).
- S&P 500 E-mini futures: Quarterly contracts that pay out 50 USD $\times$ S&P 500 value on the last day of the expiry month (i.e., March, June, September, and December)

NN Input/Output Visual

Back to Variables

- Each FOMC statement is paired with DP_{t+n} for each asset
- Input $X_i = 768 \times j$ matrix: Columns = j words in order, rows = 768 word-features

Statement Text	Dec 12, 2006: "The Federal Open Market Committee decided today to keep its target for the federal funds rate at 5 1/4 percent..."
Input Matrix 768 word-features × 512 words	$\left[\underbrace{x_i^1}_{\textit{The}} \; \underbrace{x_i^2}_{\textit{Federal}} \; \underbrace{x_i^3}_{\textit{Open}} \; \underbrace{x_i^4}_{\textit{Market}} \; \underbrace{x_i^5}_{\textit{Committee}} \; \underbrace{x_i^6}_{\textit{decided}} \; \dots \; \underbrace{x_i^{512}}_{.}\right]$

- x_t^0 = Dummy vector that gets updated with intermediate layers of X
- Output = $DP_{i,t+n}$ for each asset
- Update XLNet parameters to minimise $\sum_{i \in N} \frac{1}{N} \left(DP_{i,t+n} - \widehat{DP}_{i,t+n}\right)^2$

Popular Text Analysis Methods in Macro

Back to Approach

1. Counts of single words
   - "*employment* went up, but *inflation* did not"
   - "*inflation* went up, but *employment* did not"
   - **Problem**: Method produces same measure from both sentences
2. Counts of n-grams
   - "*economic growth slowed*, but is expected to pick up pace later this year"
   - **Problem**: Method doesn't quantify full sentence context

Universal Approximation Theorem

Back to Approach Back to NN Training Overview

- Universal Approximation Theorem from ML literature:
  - Neural networks with at least 1 hidden layer can approximate any function
  - Existence theorem → Nothing about finding structure and training
- In reality, adding more layers:
  - ↓ number of parameters for each node function
  - ↓ computational, data, and training requirements

Details about XLNet from Yang et al. (2019)

Hyperparameters Back to Approach

- **Overview**: Open-source, pretrained NN for text analysis
  - Paper version: xlnet-base-cased
- **Design**: Permutation- and autoregressive-based learning
  - 12 hidden layers (each of size 768)
  - 12 self-attention heads (each of size 64)
  - Vocabulary size of 32,000 word tokens
  - 110 million network parameters
- **Text Input**: Sequence of numerical vectors representing words and document
- **Transfer Learning**: "pretrained" parameters = Starting point for new task
  - SOTA status in translation, Q&A, word prediction, classification/regression
- **Initial Task**: Predict words using all permutations of text
- **Initial Data**: BookCorpus (11,038 books), English Wikipedia (6 mil articles)
  - Additional data for original, larger-scale NN:
    - Giga5 (9.9 mil news articles), ClueWeb12 (733 mil web pages), Common Crawl (1K+ TB text from web pages)

Hyperparameters for Fine-tuning XLNet (1/2)

XLNet Details

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

Table 9: The symbol "*" denotes the hyperparameter that undergoes tuning during the fine-tuning process. The random seed is set as a constant to ensure replicability by always initalising random components (e.g., weights and biases) to the same value. (cont.)

Hyperparameters for Fine-tuning XLNet (2/2)

XLNet Details

Hyperparameter	Value
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 10: The symbol "*" denotes the hyperparameter that undergoes tuning during the fine-tuning process. The random seed is set as a constant to ensure replicability by always initalising random components (e.g., weights and biases) to the same value.

Addressing Look-ahead Bias

Back to Approach | Back to Loop

- Sarkar and Vafa (2024): NNs predict values in past using info. in the future.
    - NNs for text analysis trained with large amounts of data
    - High probability of future info. used in initial training of NN weights
    - Look-ahead bias addressed for 2 reasons:

1. xlnet-base-cased Initially trained only with BookCorpus and English Wikipedia
    - Very low probability of XLNet initially trained on FOMC statements and futures data
2. Pre-processed FOMC statements have no references to relevant times t and $t+1$

XLNet Details

NN Training Overview

Back to Loop

- Train NN → Fine-tune parameters and hyperparameters to fit training data Small NN Ex

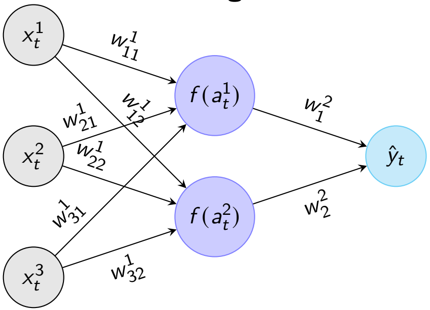
  1. Fix network structure (layers and nodes) + non-tuned hyperparameters UAT + Layers
  2. Choose value for hyperparameter that will be tuned Hyperparameter Tuning
  3. Iteratively update parameters to $\downarrow \widehat{MSE_{IS}}$
  4. Evaluate NN → Judge based on R^2_{OOS}
  5. Poor performance → Go back to step 1

Small NN Example

Back to NN Training Overview

- **Data:** 4 variables x_t^1, x_t^2, x_t^3, y_t
- **Goal:** Predict y_t from $X \equiv x_t^1, x_t^2, x_t^3$
- **Example:** 2 layers, 2 "hidden" nodes
- From X_t to $\hat{y}_t$ for observation $t \in T$:
  - Linearly combine $x_t^1, x_t^2, x_t^3 \rightarrow a_t^1, a_t^2$
  - f is a non-linear function
  - $\hat{y}_t$ is predicted output
- **Training** prediction error $\rightarrow$ update weights w
- **Testing** prediction error $\rightarrow$ update network structure

NN Figure



NN Matrix Algebra

$$\begin{bmatrix} x_t^1 & x_t^2 & x_t^3 \end{bmatrix} \begin{bmatrix} w_{11}^1 & w_{12}^1 \\ w_{21}^1 & w_{22}^1 \\ w_{31}^1 & w_{32}^1 \end{bmatrix} = \begin{bmatrix} a_t^1 & a_t^2 \end{bmatrix}$$

$$\begin{bmatrix} f(a_t^1) & f(a_t^2) \end{bmatrix} \begin{bmatrix} w_1^2 \\ w_2^2 \end{bmatrix} = \hat{y}_t$$

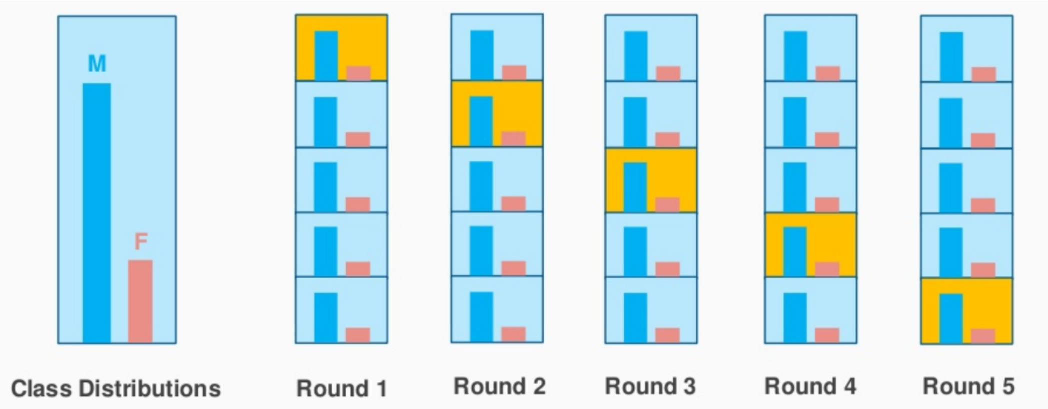
Why Stratified Sampling?

Back to CV

- Why stratified over random splitting?
  1. Transfer learning $\longrightarrow$ Lower data requirements for NNs BUT
  2. Large sample size for NNs $\longrightarrow$ Fold $\approx$ Population for characteristics
     - $\longrightarrow$ Can use random k-fold cross validation
  3. Small sample size for NNs $\nrightarrow$ Fold $\approx$ Population
     - Create folds conditioned on class dist can help
  4. Minimises diff between pop and fold distributions of FOMC statement characteristics
  5. **Result**: Better learning and predictive performance from NN

Stratified Cross Validation Visual

Back to CV



Estimating Optimal Event Windows: Accuracy Metrics

Back

- For each fold, primary metric to judge NN from Hawinkel et al. (2024):

$$R^2_{OOS} = 1 - \frac{\widehat{MSE}}{\widehat{MST}} = 1 - \frac{T^{-1}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2}{\frac{T+1}{T(T-1)}\sum_{i=1}^{T}(y_i - \overline{y_{IS}})^2}, \tag{9}$$

Estimating Optimal Event Windows: Accuracy Metrics

Back

- For each fold, primary metric to judge NN from Hawinkel et al. (2024):

$$R^2_{OOS} = 1 - \frac{\widehat{MSE}}{\widehat{MST}} = 1 - \frac{T^{-1}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2}{\frac{T+1}{T(T-1)}\sum_{i=1}^{T}(y_i - \overline{y_{IS}})^2}, \tag{9}$$

- **Definition**: Comparison between two models: NN and null model
  - Null model: $\overline{y_{IS}}$ as prediction

Estimating Optimal Event Windows: Accuracy Metrics

Back

- For each fold, primary metric to judge NN from Hawinkel et al. (2024):

$$R^2_{OOS} = 1 - \frac{\widehat{MSE}}{\widehat{MST}} = 1 - \frac{T^{-1}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2}{\frac{T+1}{T(T-1)}\sum_{i=1}^{T}(y_i - \overline{y_{IS}})^2}, \tag{9}$$

- **Definition**: Comparison between two models: NN and null model
  - Null model: $\overline{y_{IS}}$ as prediction
- **Interpretation**: % of null model's MSE explained by NN
  - NOT % of DP_{t+n} variance explained by NN $\because$ nonlinearity

Estimating Optimal Event Windows: Accuracy Metrics

Back

- For each fold, primary metric to judge NN from Hawinkel et al. (2024):

$$R^2_{OOS} = 1 - \frac{\widehat{MSE}}{\widehat{MST}} = 1 - \frac{T^{-1}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2}{\frac{T+1}{T(T-1)}\sum_{i=1}^{T}(y_i - \overline{y_{IS}})^2}, \tag{9}$$

- **Definition**: Comparison between two models: NN and null model
  - Null model: $\overline{y_{IS}}$ as prediction
- **Interpretation**: % of null model's MSE explained by NN
  - NOT % of DP_{t+n} variance explained by NN $\because$ nonlinearity
- **Explicit objective function**: Minimise $\widehat{MSE}$ during fine-tuning
  - $\min \widehat{MSE} = \max R^2_{OOS}$

Estimating Optimal Event Windows: Accuracy Metrics

Back

- For each fold, primary metric to judge NN from Hawinkel et al. (2024):

$$R^2_{OOS} = 1 - \frac{\widehat{MSE}}{\widehat{MST}} = 1 - \frac{T^{-1}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2}{\frac{T+1}{T(T-1)}\sum_{i=1}^{T}(y_i - \overline{y_{IS}})^2}, \tag{9}$$

- **Definition**: Comparison between two models: NN and null model
  - Null model: $\overline{y_{IS}}$ as prediction
- **Interpretation**: % of null model's MSE explained by NN
  - NOT % of DP_{t+n} variance explained by NN $\because$ nonlinearity
- **Explicit objective function**: Minimise $\widehat{MSE}$ during fine-tuning
  - $\min \widehat{MSE} = \max R^2_{OOS}$
- **Other tracked metrics**: ρ_{OOS}, $\widehat{MAE}_{OOS}$, $\widehat{MSE}_{IS}$

Tuning XLNet Hyperparameters

Back to Loop

- Hyperparameters: Variables outside NN structure that affect training performance
  - "Tuned" hyperparameter: Learning rate
  - Analogy: CPU & GPU overclocking (e.g., voltage, temperature, clock frequencies)
- For each 10-minute interval, "tune" learning rate of XLNet:
  - "Best chance" of approximating $f(\text{FOMC statement text}) = DP_{t+n}$
  - Tuning process takes 1 computation day for each DP_{t+n}

Optimal Event Windows: R^2_{OOS} Table (1/2)

Back to Summary Text

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.**

Table 11: Differences of $\overline{R^2_{OOS}}$ between 30-minute and Optimal Event Windows (cont.)

Optimal Event Windows: R^2_{OOS} Table (2/2)

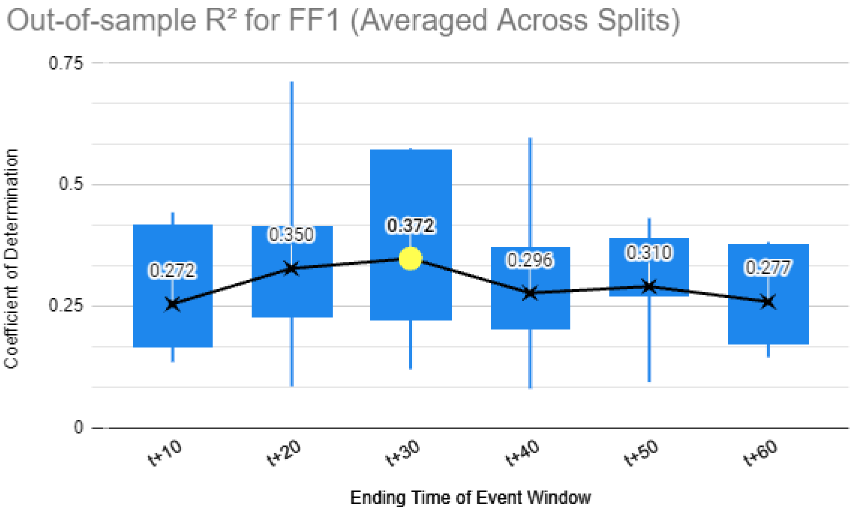
Back to Summary Text

Asset	$\overline{R^2_{OOS}}$, 30-min	$\overline{R^2_{OOS}}$, Optimal	**Difference**
*FVc*1	11.7%	21.4%	**+9.7 p.p.**
*FVc*2	12.7%	19.2%	**+6.5 p.p.**
*TYc*1	11.7%	25.5%	**+13.8 p.p.**
*TYc*2	11.4%	23.9%	**+12.5 p.p.**
*USc*1	15.7%	28.5%	**+12.8 p.p.**
*USc*2	18.7%	32.1%	**+13.4 p.p.**
SPX	18.4%	23.2%	**+4.8 p.p.**
*ESc*1	22.9%	26.2%	**+3.3 p.p.**
*ESc*2	19.3%	23.5%	**+4.2 p.p.**

Table 12: Differences of $\overline{R^2_{OOS}}$ between 30-minute and Optimal Event Windows

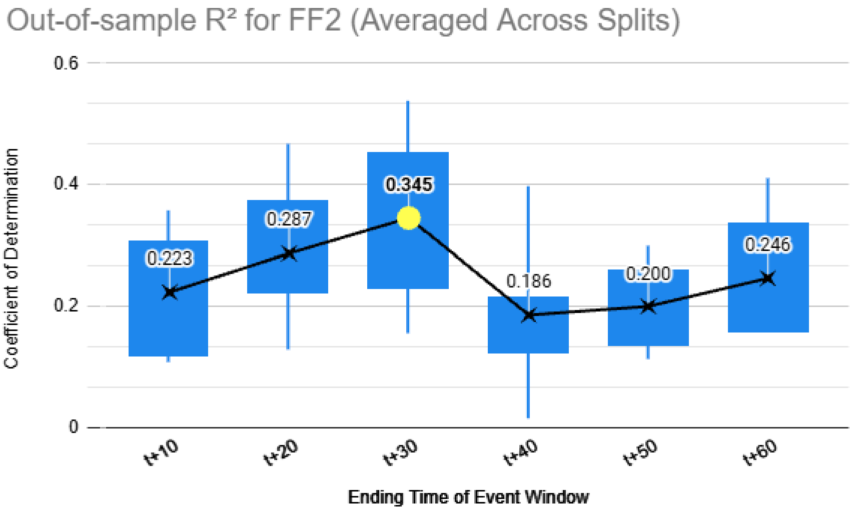
Optimal Event Windows: *FF*1

Back to Summary Text



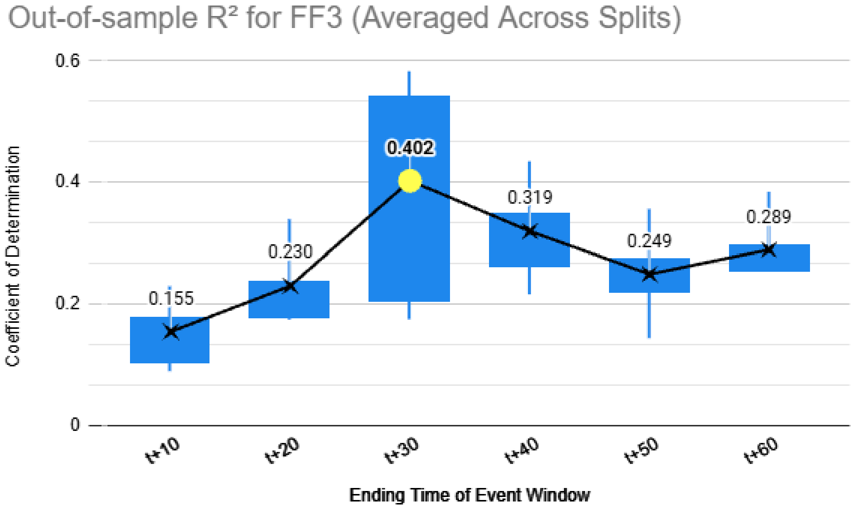
Optimal Event Windows: *FF*2

Back to Summary Text One Signal



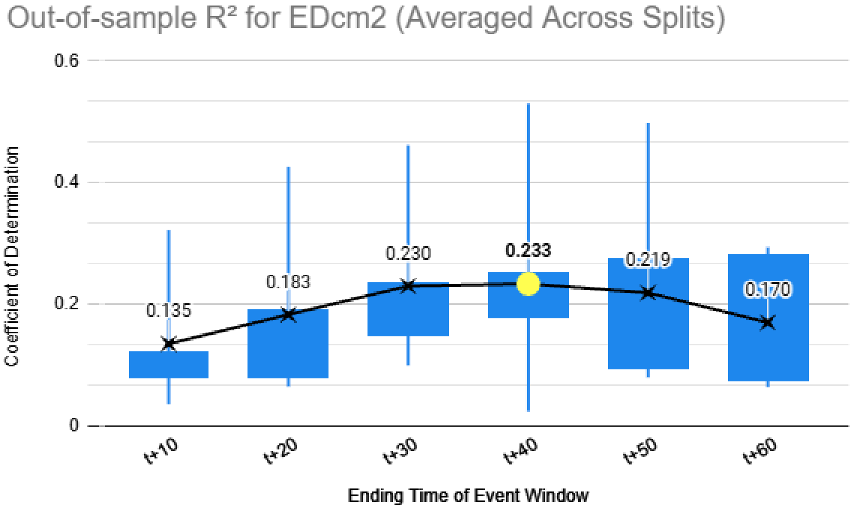
Optimal Event Windows: *FF3*

Back to Summary Text



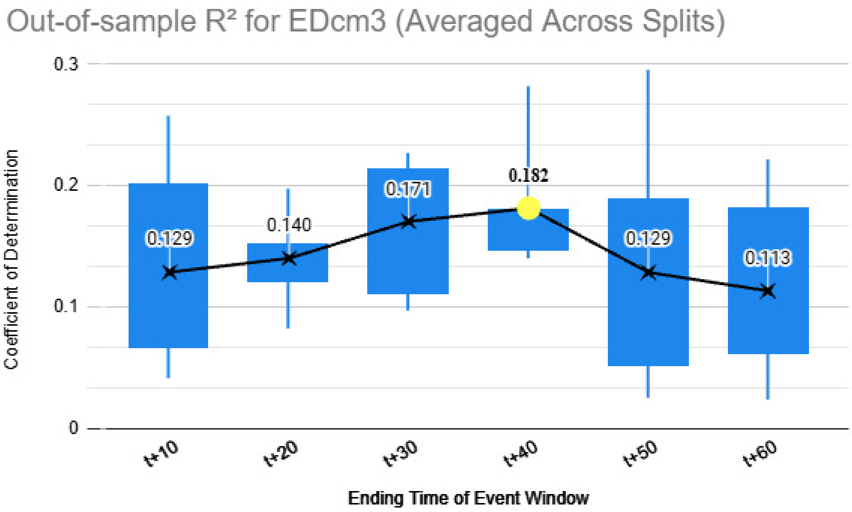
Optimal Event Windows: *EDcm2*

Back to Summary Text



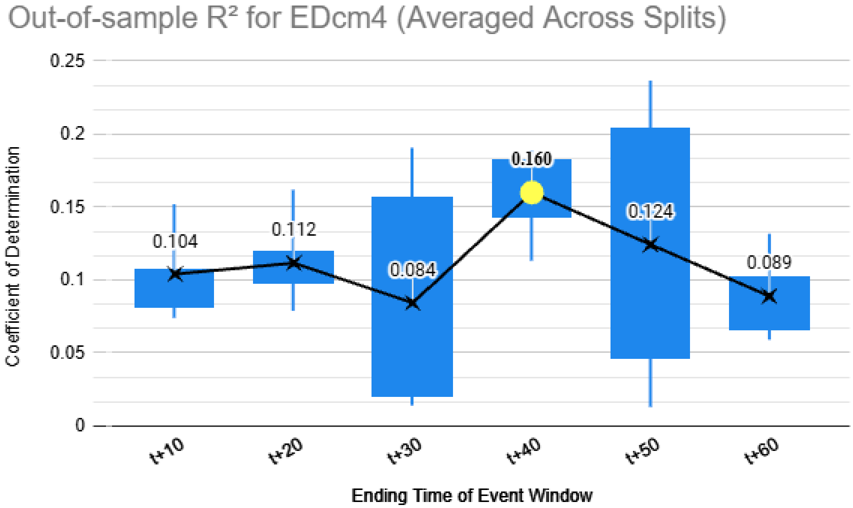
Optimal Event Windows: *EDcm3*

Back to Summary Text



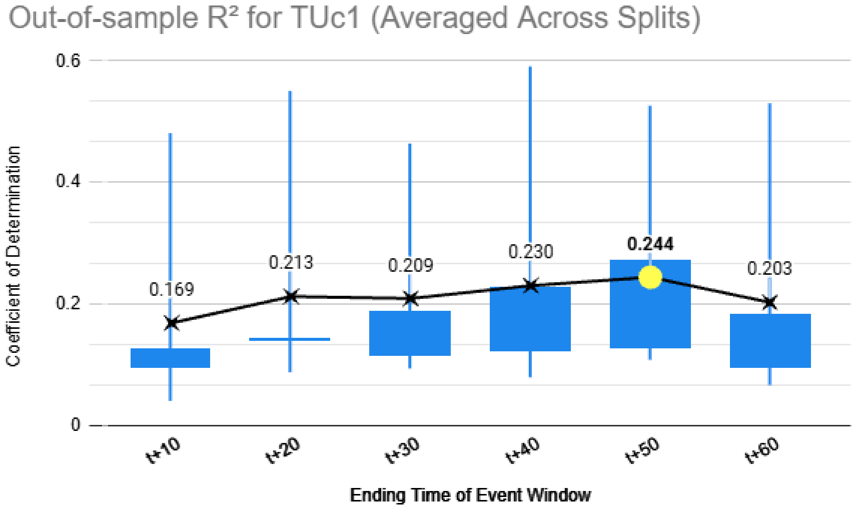
Optimal Event Windows: *EDcm4*

Back to Summary Text



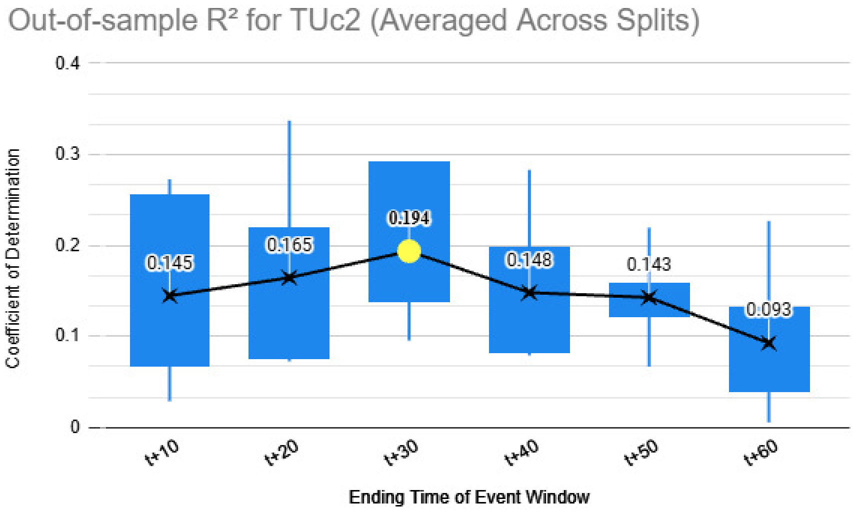
Optimal Event Windows: *TUc*1

Back to Summary Text One Signal



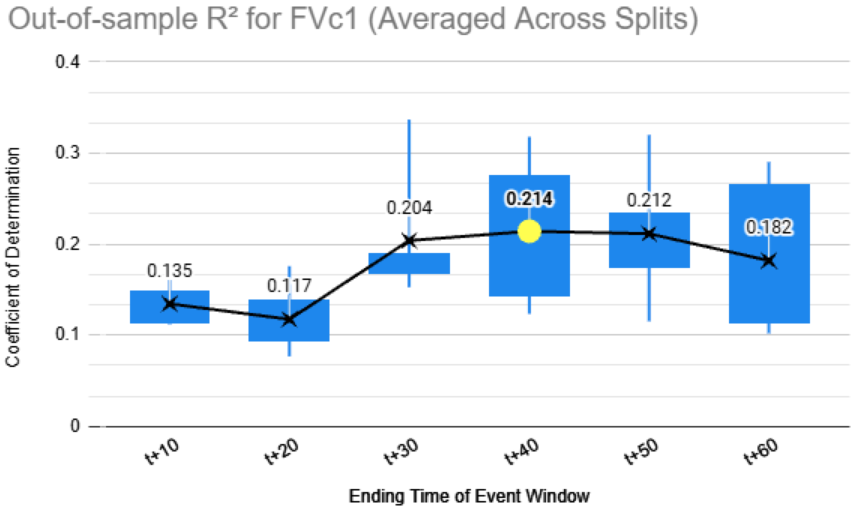
Optimal Event Windows: *TUc2*

Back to Summary Text



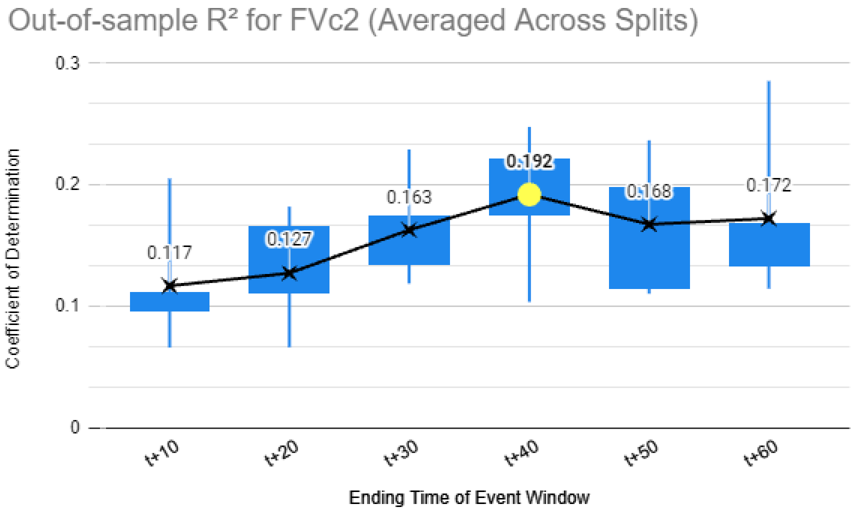
Optimal Event Windows: *FVc*1

Back to Summary Text



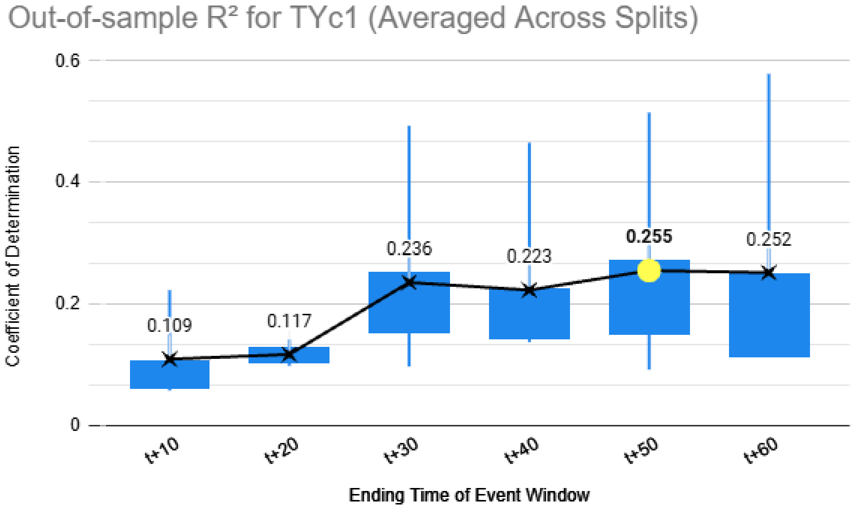
Optimal Event Windows: *FVc2*

Back to Summary Text



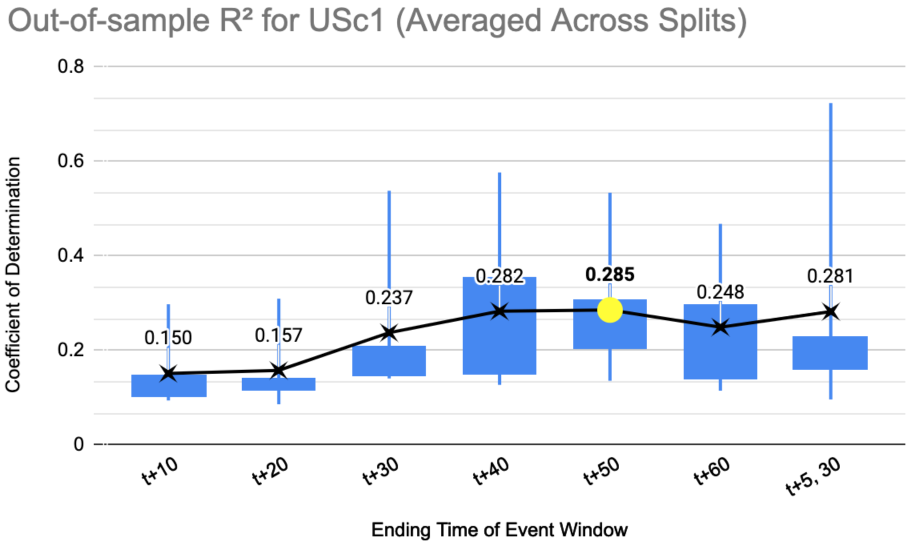
Optimal Event Windows: *TYc*1

Back to Summary Text



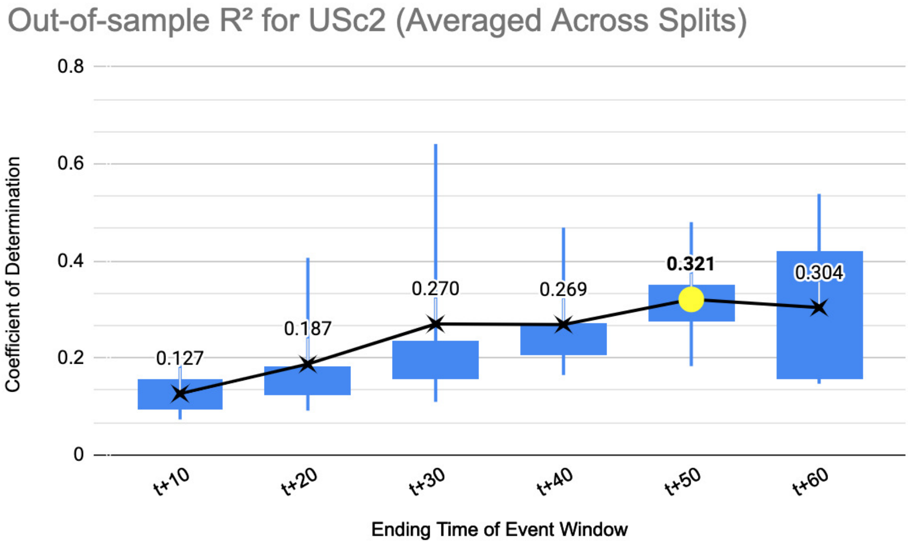
Optimal Event Windows: *USc*1

Back to Summary Text One Signal



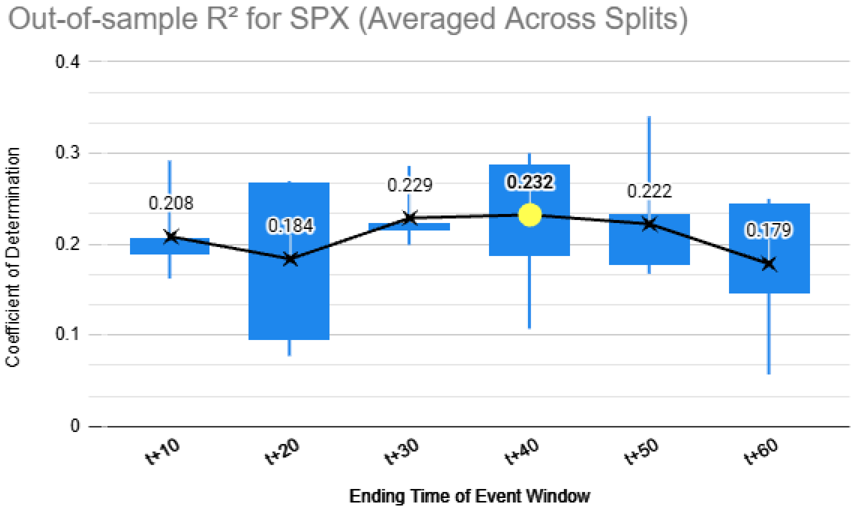
Optimal Event Windows: *USc2*

Back to Summary Text



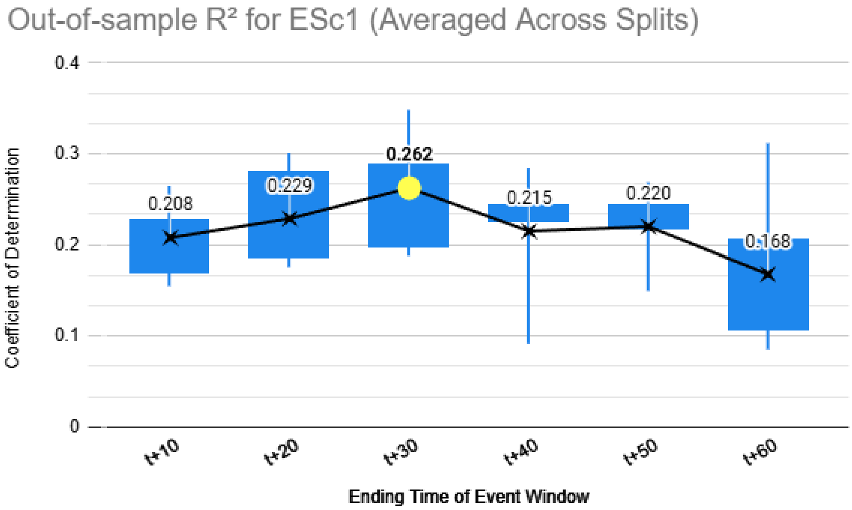
Optimal Event Windows: *SPX*

Back to Summary Text



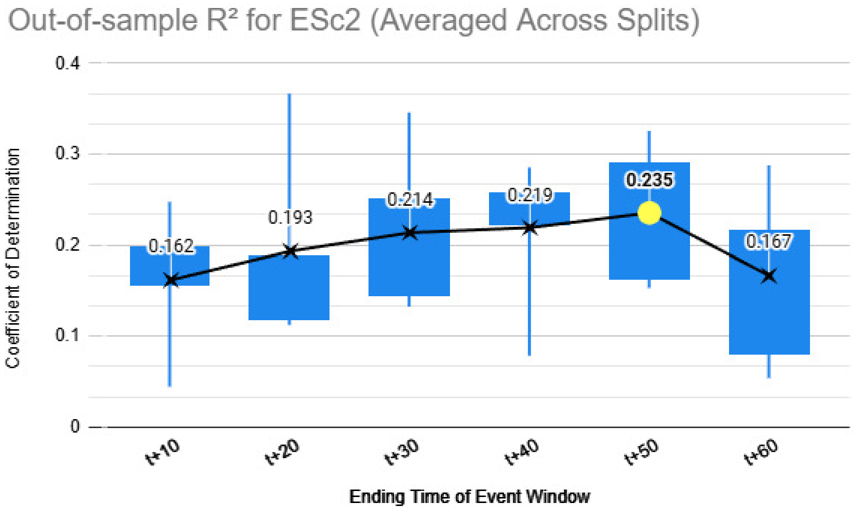
Optimal Event Windows: *ESc*1

Back to Summary Text



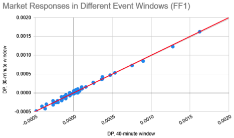
Optimal Event Windows: *ESc2*

Back to Summary Text

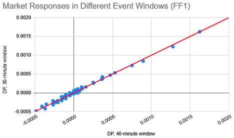


Optimal Event Windows: Diff Windows, Diff Responses (1/2)

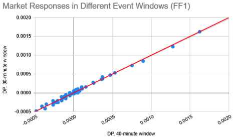
Back



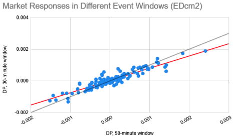
(a) $FF1^{**}$



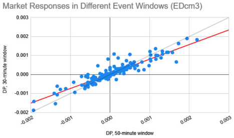
(b) $FF2^{***}$



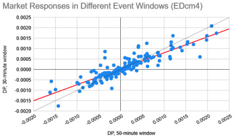
(c) $FF3^{***}$



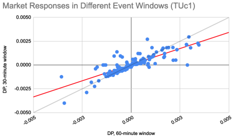
(d) $EDcm2^{***}$



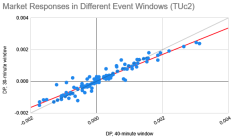
(e) $EDcm3^{***}$



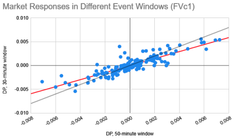
(f) $EDcm4^{***}$



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



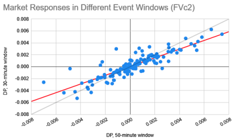
(h) $TUc2^{***}$



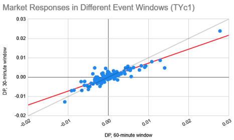
(i) $FVc1^{***}$

Optimal Event Windows: Diff Windows, Diff Responses (2/2)

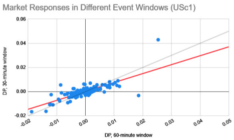
Back



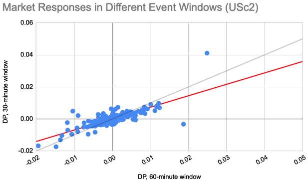
(a) *FVc2****



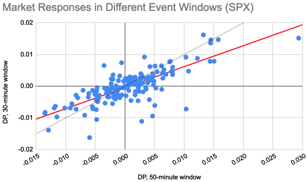
(b) *TYc1****



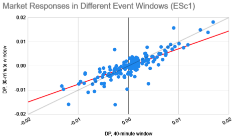
(c) *USc1****



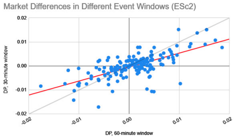
(d) *USc2****



(e) *SPX****



(f) *ESc1****



(g) *ESc2****

Robustness Check of Optimal Event Windows

Back to One Signal

1. Pick an interest-rate or equity futures contract

¶Performed for $FF2$, $FF4$, $TUc1$, $TYc2$, $USc1$.

Robustness Check of Optimal Event Windows

Back to One Signal

1. Pick an interest-rate or equity futures contract
2. Take predictions $\widehat{DP_{t+n}}$ for each $k = 5$ fold from optimal event window

¶Performed for $FF2$, $FF4$, $TUc1$, $TYc2$, $USc1$.

Robustness Check of Optimal Event Windows

Back to One Signal

1. Pick an interest-rate or equity futures contract
2. Take predictions $\widehat{DP_{t+n}}$ for each $k = 5$ fold from optimal event window
3. Check if $\overline{R^2_{OOS}} \forall t + n \geq \overline{R^2_{OOS}}$ in optimal window length
   - Yes: Perform "joint" estimation in that window length[¶]
   - No: Go back to step 1

[¶] Performed for $FF2$, $FF4$, $TUc1$, $TYc2$, $USc1$.

Robustness Check of Optimal Event Windows

Back to One Signal

1. Pick an interest-rate or equity futures contract
2. Take predictions $\widehat{DP_{t+n}}$ for each $k = 5$ fold from optimal event window
3. Check if $\overline{R^2_{OOS}} \forall t + n \geq \overline{R^2_{OOS}}$ in optimal window length
   - Yes: Perform "joint" estimation in that window length[¶]
   - No: Go back to step 1

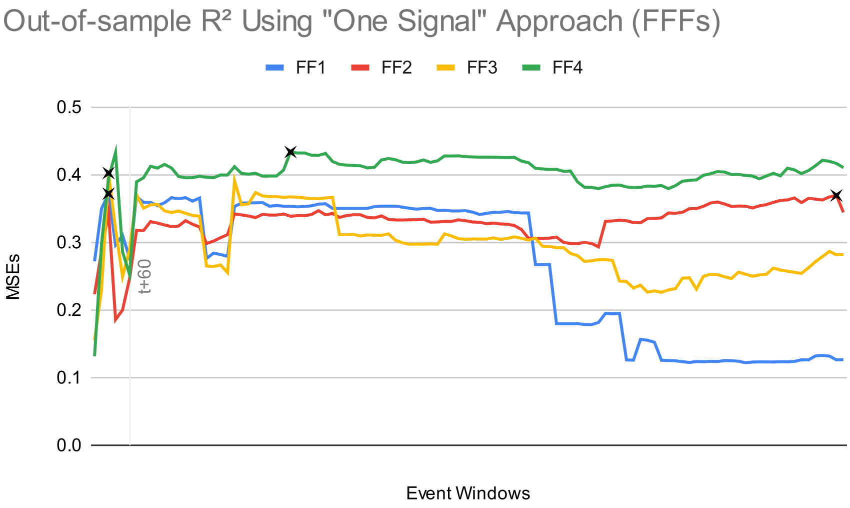
- **Results**:
  1. Optimal window length has highest $\overline{R^2_{OOS}}$
  2. "Jointly" estimated $\overline{R^2_{OOS}}$ for window $> t + 60$ greater than "" for window $t + 20$

  $\rightarrow$ Event window with global maximum $\overline{R^2_{OOS}}$ could be in window length $> t + 60$

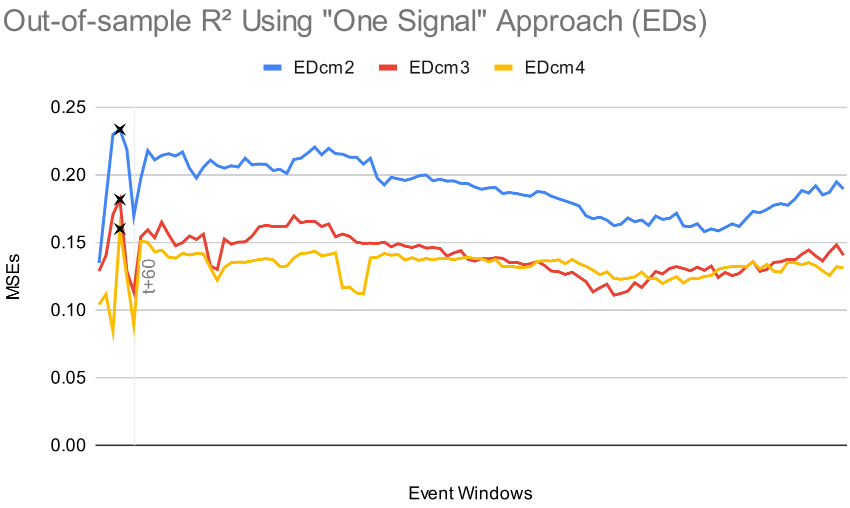
[¶] Performed for $FF2$, $FF4$, $TUc1$, $TYc2$, $USc1$.

Testing R^2 Using "One Signal" Approach for Federal Funds Futures

FF2 FF4

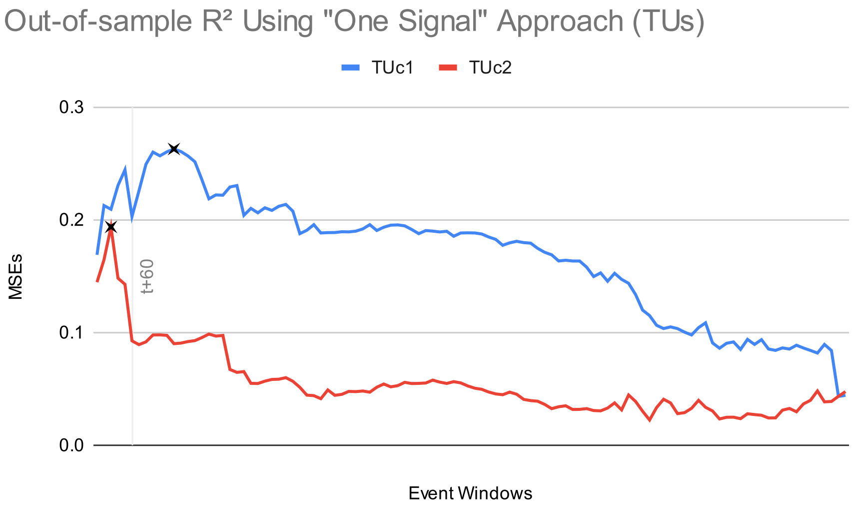


Testing R^2 Using "One Signal" Approach for Eurodollar Futures

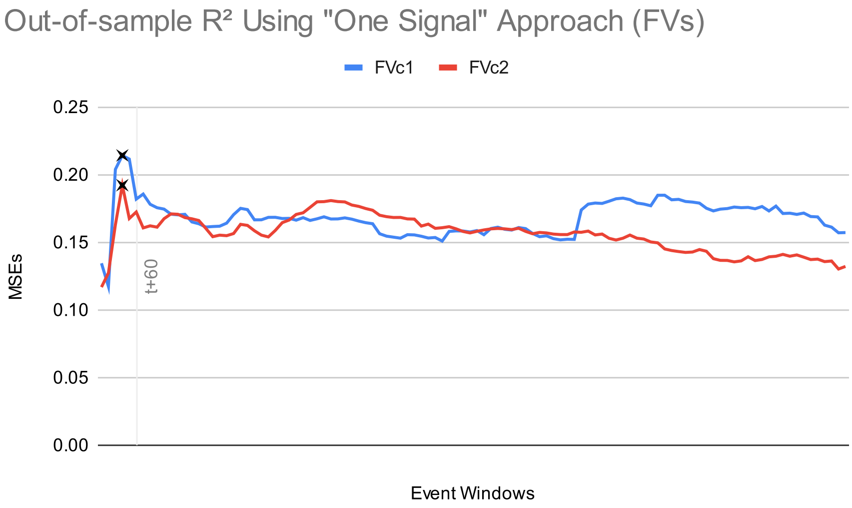


Testing R^2 Using "One Signal" Approach for 2-Year Treasury Futures

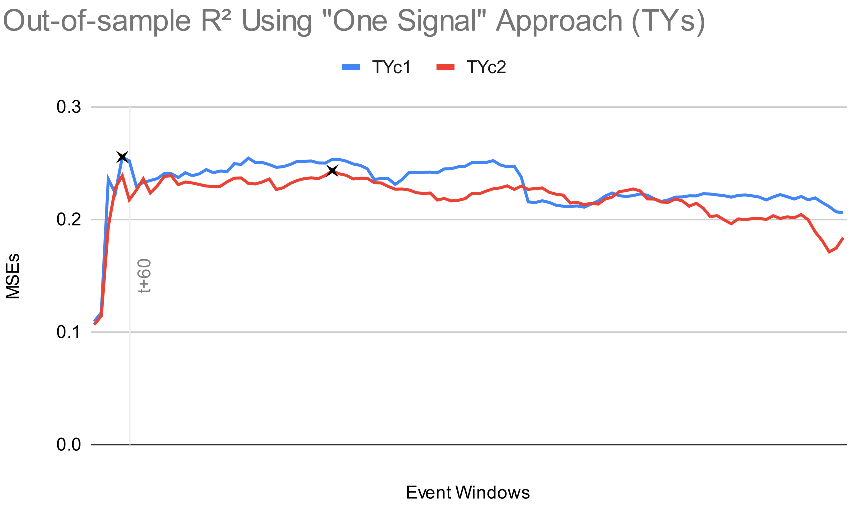
TUc1



Testing R^2 Using "One Signal" Approach for 5-Year Treasury Futures

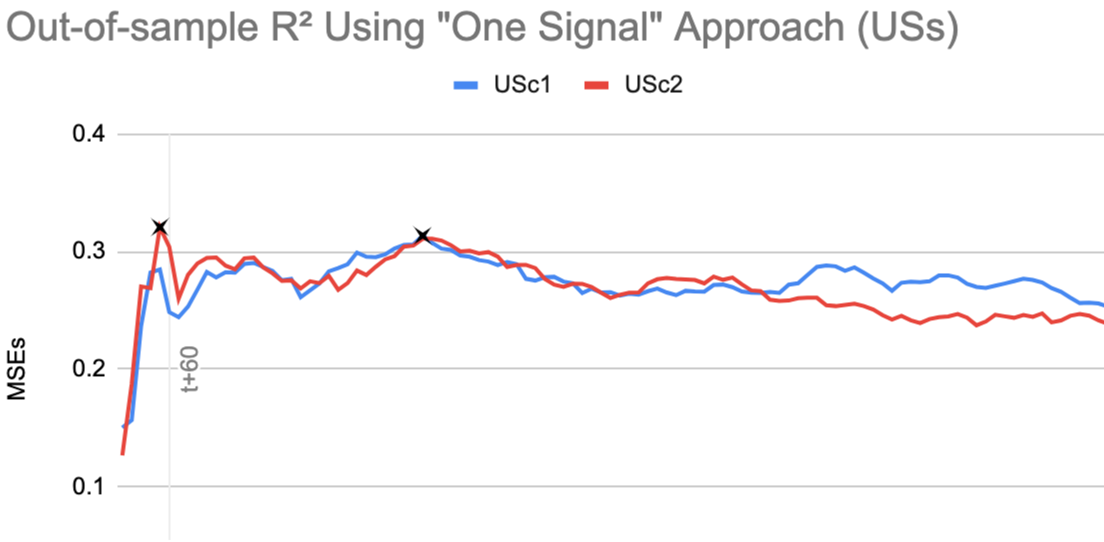


Testing R^2 Using "One Signal" Approach for 10-Year Treasury Futures TYc2

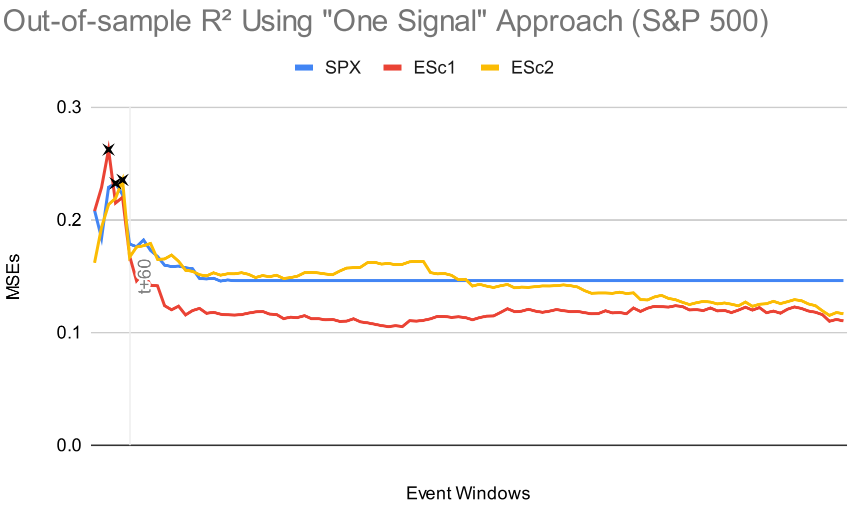


Testing R^2 Using "One Signal" Approach for 30-Year Treasury Futures

USc1



Testing R^2 Using "One Signal" Approach for S&P Index



TFIDF Equation

Back to Similarity

$$
\begin{aligned}
TFIDF_{d,t} &= tf_{d,t} * idf_{d,t} \\
&= \left[\ln\left(\frac{tc_{d,t}}{nt_d}\right) + 1\right] * \left[\ln\left(\frac{nd}{df_{d,t} + 1}\right) + 1\right]
\end{aligned}
$$

- $tf_{d,t}$: Number of times term t is in document d
- nt_d: Number of terms in document d
- nd: Number of documents
- $df_{d,t}$: Number of documents term t appears in

TFIDF Informative Terms

Back to Similarity

- Additional pre-processing steps on FOMC statements:
    1. Make all words lowercase
    2. Remove words with little semantic meaning (e.g., articles)
    3. Convert all words into base terms (e.g., "increas")

TFIDF Informative Terms

Back to Similarity

- Additional pre-processing steps on FOMC statements:
  1. Make all words lowercase
  2. Remove words with little semantic meaning (e.g., articles)
  3. Convert all words into base terms (e.g., “increas”)

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 13: FOMC Statement Base Terms with Top 30 TFIDF Scores

Cosine Similarity Matrix

Back to Similarity