Appendices

A. Data Sources and Construction

Our data sources and basic information on the releases are presented in the table below.

Data Release	Source	Frequency	Release time	Surprise St. Dev.	Units
Non-farm	BLS	Monthly	8:30	90.81	Thousands
Init. Claims	ETA	Weekly	8:30	17.82	Thousands
Durable	Census	Monthly	8:30	2.74	Percentage change mom
Emp. Cost	BLS	Monthly	8:30	0.19	Percentage change mom
Retail	Census	Monthly	8:30	0.55	Percentage change mom
Retail Ex. Auto	Census	Monthly	8:30	0.42	Percentage change mom
GDP (advance)	BEA	Quarterly	8:30	0.75	Percentage change qoq, ar
CPI	BLS	Monthly	8:30	0.12	Percentage change mom
Core CPI	BLS	Monthly	8:30	0.09	Percentage change mom
PPI	BLS	Monthly	8:30	0.40	Percentage change mom
Core PPI	BLS	Monthly	8:30	0.25	Percentage change mom
Hourly Earn.	BLS	Monthly	8:30	0.15	Dollars per hour
Unemp.	BLS	Monthly	8:30	0.14	Percent
FOMC	Fed	8 per year	14:15*	8.1	Basis points

(*) We incorporate some minor deviations of timing to accommodate FOMC announcement times in the early sample. However, in the majority of our sample the announcements are made around 14:15.

Notes: Acronyms for the sources are as follows: BEA (Bureau of Economic Analysis), BLS (Bureau of Labor Statistics), Census (Bureau of the Census), ETA (Employment and Training Administration), Fed (Federal Reserve Board of Governors). Acronyms of the unite are: mom (month-on-month), qoq (quarter-on-quarter) and ar (annualized rate). Standard deviations are for the sample 1992-2017. For the FOMC, the sample is 1992-2007.

To calculate the macroeconomic data release surprises used in the study we proceed as follows. Let $R_{j,t}$ be the released value of a variable j at time t. Let $E_{j,t}$ be the expectation (or the survey) of this release. Then the surprise is defined as:

$$S_{j,t} = R_{j,t} - E_{j,t}$$

Then we standardize the surprises to so that units are comparable across different types of announcements, and transmission coefficients capture per standard deviation effects:

$$s_{j,t} = \frac{S_{j,t}}{\sigma_{S_j}}$$

where σ_{S_j} is the standard deviation of the surprise for the announcement type j. For expectations, we use the median prediction from the survey conducted by MMS/Action Economics on the previous Friday of a release.

Monetary policy surprises are measured using intraday changes of Fed Funds Futures implied yield changes around FOMC announcements, following the methodology of Kuttner (2001).

For the yields, our high frequency data consists of 5-minute quotes of first Eurodollar (ED1), fourth Eurodollar (ED4), on the run 2-year, 5-year, 10-year and 30-year Treasury futures from Chicago Mercantile Exchange (CME). Eurodollar futures prices are converted to interest rates by subtracting the price of ED1 and ED4 from 100. We calculate 20-minute changes in future prices around macroeconomic and FOMC releases:

$$\Delta P_{j,d} = P_{j,d,t-5min} - P_{j,d,t+15min}$$

where $P_{j,d}$ is the futures price of an asset $j \in${2-year, 5-year, 10-year, 30-year} on the day d of a specific announcement and t is the time of that announcement (e.g. 8:30am). For Eurodollar futures, we use implied interest rates to calculate announcement window changes. For the Treasury futures, we divide the price changes by the approximate duration of the bonds and flip the sign to convert them to yield changes.

B. Heteroskedasticity-Based Estimation Applied to the OLS Residuals

An event study regression with a latent factor and no measurement error has the form:

$$y_t = \beta s_t + \gamma d_t f_t + \varepsilon_t$$

where $s_t = s_t^*$. In the usual event study setup, β can be separately identified by OLS run on data from event days. The residual of this regression is:

$${\phi_t}^E = \gamma f_t + \varepsilon_t$$

The counterpart for non-event days is:

$${\phi_t}^{NE} = \varepsilon_t$$

We then have the following event and non-event variance-covariance matrices for ϕ_t:

$$\Omega^{\phi_E} = \begin{pmatrix} \gamma^2 + \sigma_\varepsilon^2 & 0 \\ . & \sigma_s^2 \end{pmatrix}$$

$$\Omega^{\phi_{NE}} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{pmatrix}$$

Thus, the heteroskedasticity-based estimator for γ is given by $\sqrt{\hat{\Omega}_{1,1}^{\phi_E} - \hat{\Omega}_{1,1}^{\phi_{NE}}}$. Below we show that this two-step estimation procedure produces similar coefficients to the one step estimation we employed.

We demonstrate this point by considering FOMC announcements. To make sure that our results are not influenced by the different number of observations, we drop the days with at least one missing yield change. Then, we estimate equation (3.2) around FOMC announcement days and compare the estimates of γ from the one step estimation with that of the two step estimates.

	ED1	ED4	2-year	5-year	10-year	30-year
Kalman Filter	2.10	6.96	5.62	6.00	4.24	2.44
Two-step	2.84	6.64	5.05	5.20	3.89	2.54

Notice that the estimated coefficients are very close, implying that Kalman filter and the (two step) heteroskedasticity-based estimates are very similar. But the estimates are not exactly equal. The Kalman filter takes into account the covariance between yield changes around announcements, since the filter uses all assets at once. However, the two step estimation is done asset by asset. Due to this information loss, coefficients are slightly different.

C. OLS and Heteroskedasticity-based Estimators

We consider a general model which incorporates both measurement error and an unobservable latent factor, nesting both cases. The model is:

$$
\begin{aligned}
y_t &= \beta s_t^* + \gamma d_t f_t + \varepsilon_t \\
s_t &= s_t^* + \eta_t
\end{aligned}
$$

where y_t is a log return or yield change (a scalar, without loss of generality), s_t is the observed surprise, s_t^* is the true headline surprise, d_t is a dummy that is 1 on an announcement day and 0 otherwise, f_t is an iid $N(0,1)$ latent variable, and ε_t and η_t are processes measuring noise in yields and measurement error of the headline surprise. We assume that s_t, ε_t and η_t are iid, mutually uncorrelated, have mean zero, and variances σ_*^2, σ_ε^2 and σ_η^2, respectively. To estimate β, the parameter of interest in event studies, using OLS and identification through heteroskedasticity, we need the variance-covariance matrices for event (Ω^E) and non-event (Ω^{NE}) windows:

$$
\Omega^E = \begin{pmatrix} \beta^2\sigma_*^2 + \gamma^2 + \sigma_\varepsilon^2 & \beta\sigma_*^2 \\ . & \sigma_*^2 + \sigma_\eta^2 \end{pmatrix}, \; \Omega^{NE} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{pmatrix}
$$

In this general model, the OLS estimate for β is:

$$
\hat{\beta}^{OLS} = \frac{[\hat{\Omega}^E]_{1,2}}{[\hat{\Omega}^E]_{2,2}}
$$

and the identification through heteroskedasticity estimate of β is:

$$
\hat{\beta}^{HET} = \frac{[\hat{\Omega}^E]_{1,1} - [\hat{\Omega}^{NE}]_{1,1}}{[\hat{\Omega}^E]_{1,2}}
$$

Below we derive the OLS and heteroskedasticity-based estimates in four possible cases:

1. $\gamma = 0$, $\sigma_\eta^2 = 0$ This is the case where there is neither measurement error nor a latent factor.

   Since $s_t = s_t^*$, the model simplifies to:

   $$
   y_t = \beta s_t^* + \varepsilon_t
   $$

   The variance-covariance matrices around event and non-event windows are as follows:

   $$
   \Omega^E = \begin{pmatrix} \beta^2\sigma_*^2 + \sigma_\varepsilon^2 & \beta\sigma_*^2 \\ . & \sigma_*^2 \end{pmatrix}
   $$
   $$
   \Omega^{NE} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{pmatrix}
   $$

The OLS coefficient is given by:

$$\frac{\beta\sigma_*^2}{\sigma_*^2} = \beta$$

Heteroskedasticity-based estimate is given by:

$$\frac{\beta^2\sigma_*^2 + \sigma_\varepsilon^2 - \sigma_\varepsilon^2}{\sigma_*^2} = \beta$$

In this case both estimates are consistent and should produce the same result.

2. $\gamma = 0$, $\sigma_\eta^2 \neq 0$

This case is the classical errors in variables problem for survey-based surprises that Rigobon and Sack (2006) consider. Now the model takes the following form:

$$\begin{aligned} y_t &= \beta s_t^* + \varepsilon_t \\ s_t &= s_t^* + \eta_t \end{aligned}$$

Variance-covariance matrices around event and non-event windows are given as follows:

$$\Omega^E = \begin{pmatrix} \beta^2\sigma_*^2 + \sigma_\varepsilon^2 & \beta\sigma_*^2 \\ . & \sigma_s^2 \end{pmatrix}$$
$$\Omega^{NE} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{pmatrix}$$

The OLS coefficient is given by:

$$\frac{\beta\sigma_*^2}{\sigma_s^2} = \frac{\beta\sigma_*^2}{\sigma_*^2 + \sigma_\eta^2} = \beta\left(1 - \frac{\sigma_\eta^2}{\sigma_*^2 + \sigma_\eta^2}\right)$$

Heteroskedasticity-based estimator is given by:

$$\frac{\beta^2\sigma_*^2 + \sigma_\varepsilon^2 - \sigma_\varepsilon^2}{\beta\sigma_*^2} = \beta$$

In this case OLS has attenuation bias but heteroskedasticity-based estimate is consistent.

3. $\gamma \neq 0$, $\sigma_\eta^2 = 0$

In this case, since $s_t = s_t^*$ the model takes the following form:

$$y_t = \beta s_t^* + \gamma d_t f_t + \varepsilon_t$$

Model implied variance-covariance matrices around event and non-event windows are given by:

$$\Omega^E = \begin{pmatrix} \beta^2\sigma_*^2 + \gamma^2 + \sigma_\varepsilon^2 & \beta\sigma_*^2 \\ . & \sigma_*^2 \end{pmatrix}$$

$$\Omega^{NE} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{pmatrix}$$

The OLS coefficient is given by:

$$\frac{\beta\sigma_*^2}{\sigma_*^2} = \beta$$

Using the variance-covariance matrices we can derive the heteroskedasticity-based estimator:

$$\frac{\beta^2\sigma_*^2 + \gamma^2 + \sigma_\varepsilon^2 - \sigma_\varepsilon^2}{\beta\sigma_*^2} = \beta + \frac{\gamma^2}{\beta\sigma_*^2} = \beta\left(1 + \frac{\gamma^2}{\beta^2\sigma_*^2}\right)$$

This time OLS is consistent and heteroskedasticty-based estimate is increased in absolute value due to the variance of the latent factor. The paper shows that this is the relevant case.

4. $\gamma \neq 0$, $\sigma_\eta^2 \neq 0$

   Now we are back to the general model:

$$y_t = \beta s_t^* + \gamma d_t f_t + \varepsilon_t$$
$$s_t = s_t^* + \eta_t$$

   Event and non-event window variance-covariance matrices are given as follows:

$$\Omega^E = \begin{pmatrix} \beta^2\sigma_*^2 + \gamma^2 + \sigma_\varepsilon^2 & \beta\sigma_*^2 \\ . & \sigma_s^2 \end{pmatrix}$$

$$\Omega^{NE} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{pmatrix}$$

   Using the event window variance covariance matrix, we derive the OLS coefficient:

$$\frac{\beta\sigma_*^2}{\sigma_s^2} = \frac{\beta\sigma_*^2}{\sigma_*^2 + \sigma_\eta^2} = \beta\left(1 - \frac{\sigma_\eta^2}{\sigma_*^2 + \sigma_\eta^2}\right)$$

   The heteroskedasticity-based estimate is given as follows:

$$\frac{\beta^2\sigma_*^2 + \gamma^2 + \sigma_\varepsilon^2 - \sigma_\varepsilon^2}{\beta\sigma_*^2} = \beta + \frac{\gamma^2}{\beta\sigma_*^2} = \beta\left(1 + \frac{\gamma^2}{\beta^2\sigma_*^2}\right)$$

   The table below summarizes the four cases and their implications for the coefficients:

Case	$\hat{\beta}^{OLS} \rightarrow$	$\hat{\beta}^{HET} \rightarrow$
1. $\gamma = 0, \sigma_\eta^2 = 0$	β	β
2. $\gamma = 0, \sigma_\eta^2 \neq 0$	$\beta(1 - \frac{\sigma_\eta^2}{\sigma_*^2 + \sigma_\eta^2})$	β
3. $\gamma \neq 0, \sigma_\eta^2 = 0$	β	$\beta(1 + \frac{\gamma^2}{\beta^2 \sigma_*^2})$
4. $\gamma \neq 0, \sigma_\eta^2 \neq 0$	$\beta(1 - \frac{\sigma_\eta^2}{\sigma_*^2 + \sigma_\eta^2})$	$\beta(1 + \frac{\gamma^2}{\beta^2 \sigma_*^2})$

In the paper, we rule out cases 1, 2 and 4. Furthermore, if the interpretation offered by case 3 is correct, the heteroskedasticity-based estimator should provide an estimate approximately equal to the sum of the OLS event study estimate, and the variation caused due to the unobservable component of the news. We check this in the table below. Here γ^2 is identified following the methodology in Appendix B. The OLS estimates for the announcements differ from Table 1 because days with multiple releases are dropped. It is striking that the sum in all cases is about equal to the heteroskedasticity-based estimator. The difference (for some coefficients) is caused by small sample issues (verified by a Monte Carlo exercise) and they are economically insignificant. This validates that the extra term in the heteroskedasticity-based estimator is indeed the unobserved news effect and that this estimator finds the combined effect of the headline surprise and the latent factor.

(a) β^{HET}

IDHET

	ED1	ED4	TU	FV	TY	US
Non-Farm	7.98	13.23	10.25	12.23	9.45	6.32
Initial Claims	-1.69	-2.66	-1.81	-2.36	-1.99	-1.47
Durable	1.02	3.43	3.32	3.46	2.86	1.80
Emp Cost	2.74	3.39	2.40	3.14	2.38	1.55
Retail Ex. Auto	5.19	7.33	3.96	3.96	3.00	1.81
GDP (advance)	7.59	8.08	5.76	6.77	5.46	2.91
Core CPI	2.79	5.75	4.02	5.22	4.12	2.90
Core PPI	4.22	5.33	4.41	5.59	4.70	3.20
FOMC	0.94	2.70	2.57	3.95	6.25	-7.46

(b) β^{OLS}

OLS

ED1	ED4	TU	FV	TY	US
3.27	5.99	4.81	5.49	4.10	2.48
-0.36	-0.77	-0.62	-0.72	-0.57	-0.35
0.52	1.03	0.84	1.08	0.79	0.50
0.45	1.25	0.82	1.21	0.93	0.54
0.36	1.04	0.92	1.23	0.95	0.67
0.98	2.22	1.75	1.99	1.42	0.78
0.76	1.73	1.06	1.77	1.51	1.06
0.95	1.40	1.17	1.49	1.27	0.96
0.53	0.43	0.21	0.15	0.05	-0.01

(c) γ^2

	ED1	ED4	2-year	5-year	10-year	30-year
Non-Farm	15.47	43.43	26.33	36.98	21.99	9.53
Initial Claims	0.49	1.52	0.76	1.18	0.81	0.39
Durable	0.26	2.72	2.10	2.58	1.64	0.65
Emp Cost	1.07	2.77	1.29	2.34	1.35	0.55
Retail Ex. Auto	1.73	6.79	2.78	3.39	1.96	0.77
GDP (advance)	6.49	13.01	7.07	9.51	5.73	1.66
Core CPI	1.53	7.23	3.11	6.15	3.96	1.95
Core PPI	3.15	5.22	3.99	6.11	4.35	2.14
FOMC	10.95	48.99	26.33	26.64	13.53	4.80

(d) $\beta^{OLS} + \frac{\gamma^2}{\beta\sigma_s^2}$

Sum

ED1	ED4	TU	FV	TY	US
8.00	13.24	10.29	12.23	9.46	6.32
-1.70	-2.75	-1.85	-2.36	-1.99	-1.47
1.03	3.68	3.34	3.46	2.87	1.80
2.81	3.46	2.40	3.14	2.38	1.55
5.22	7.57	3.94	3.98	3.02	1.82
7.59	8.08	5.79	6.77	5.46	2.91
2.78	5.91	3.99	5.24	4.14	2.90
4.26	5.13	4.59	5.59	4.70	3.20
0.94	2.70	2.57	3.95	6.25	-7.47

Table A1: a) Heteroskedasticity-based Event Study Estimates b) OLS-based Event Study Estimates c) Effect of the latent factor (identified following Appendix B) d) Reconstructing the heteroksedasticity-based estimates from the OLS-based estimates and the latent factor effects