

**Online Appendix**

**For**

**“The Natural Gas Announcement Day Puzzle”**

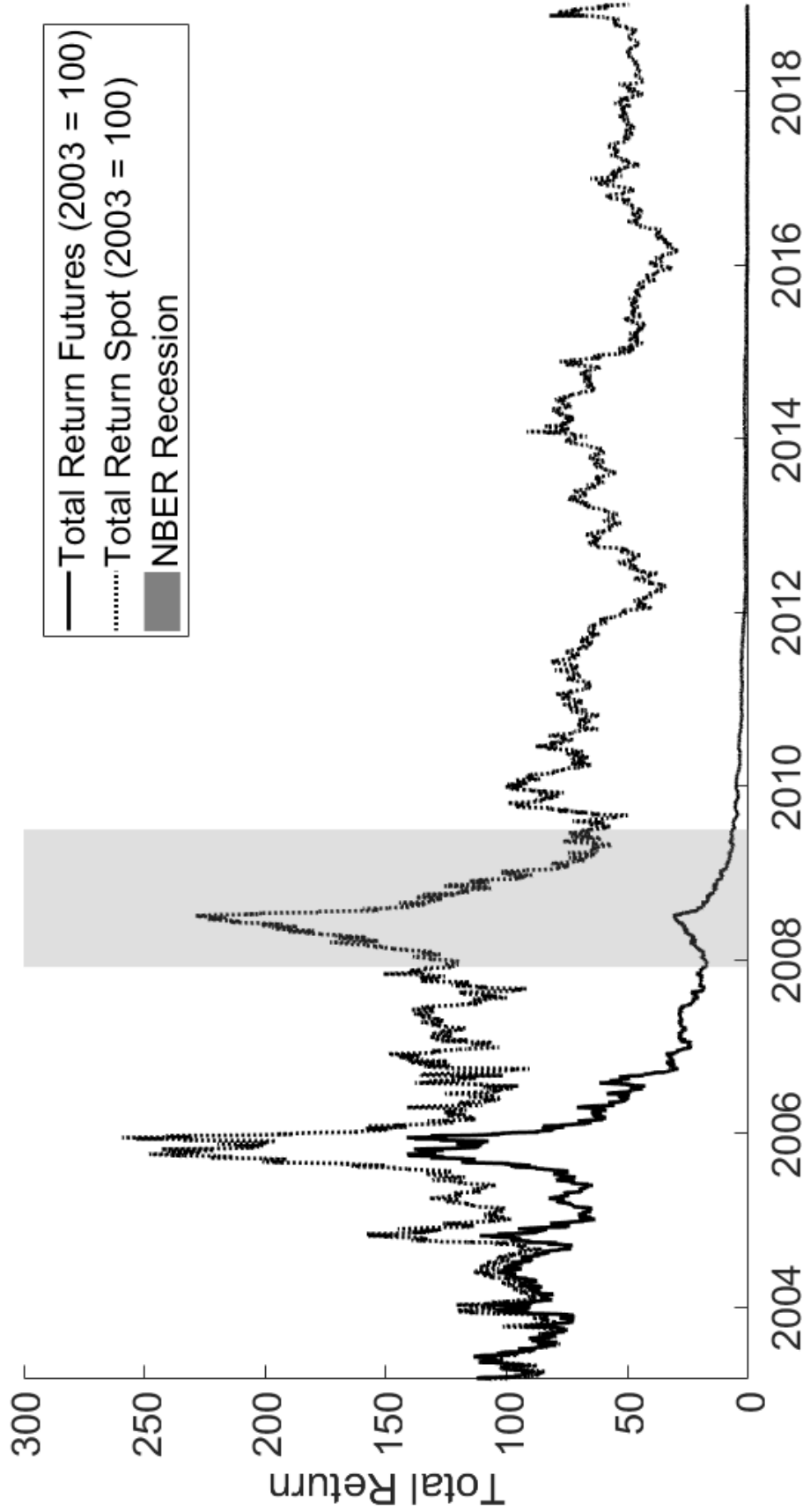
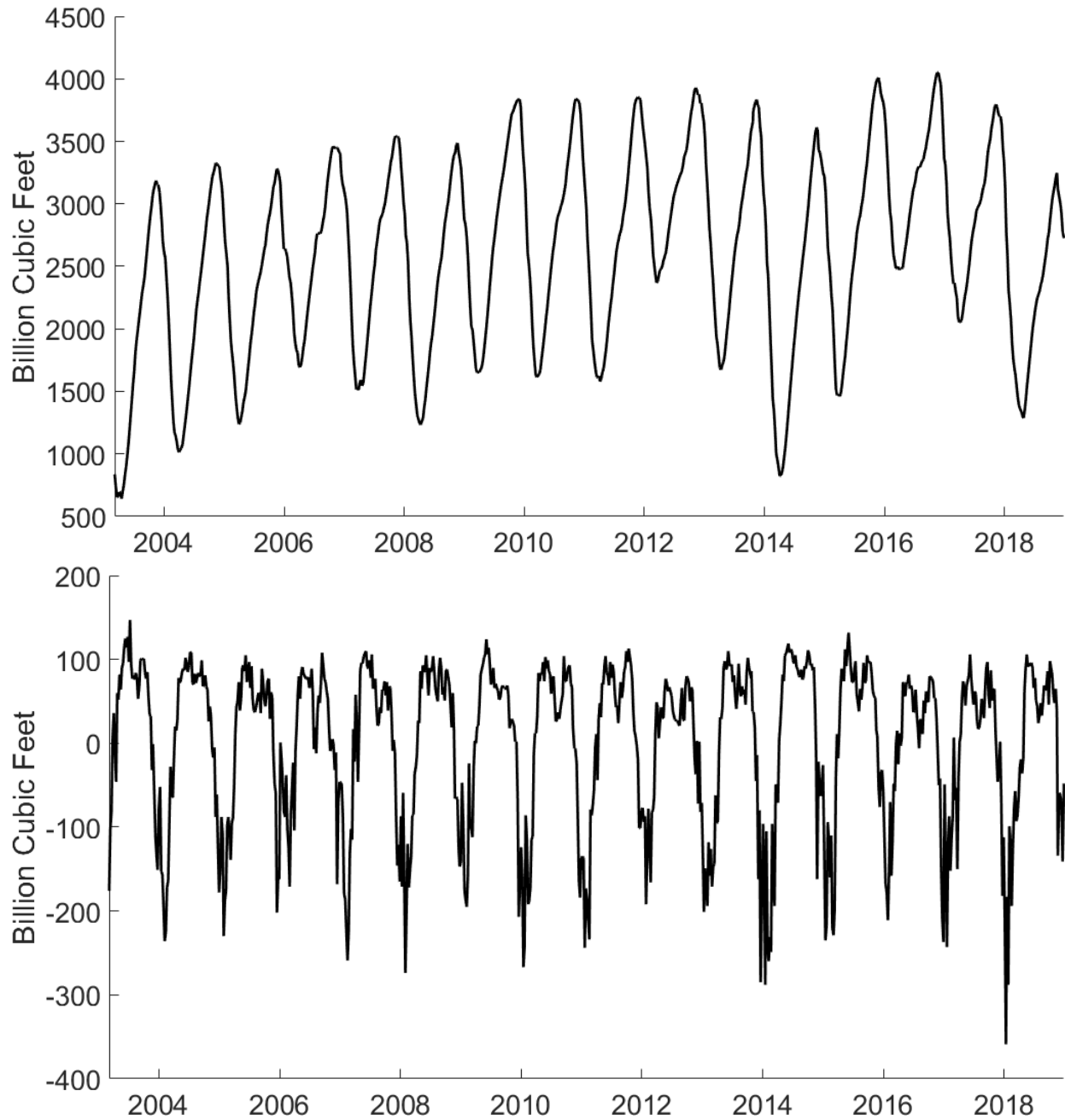


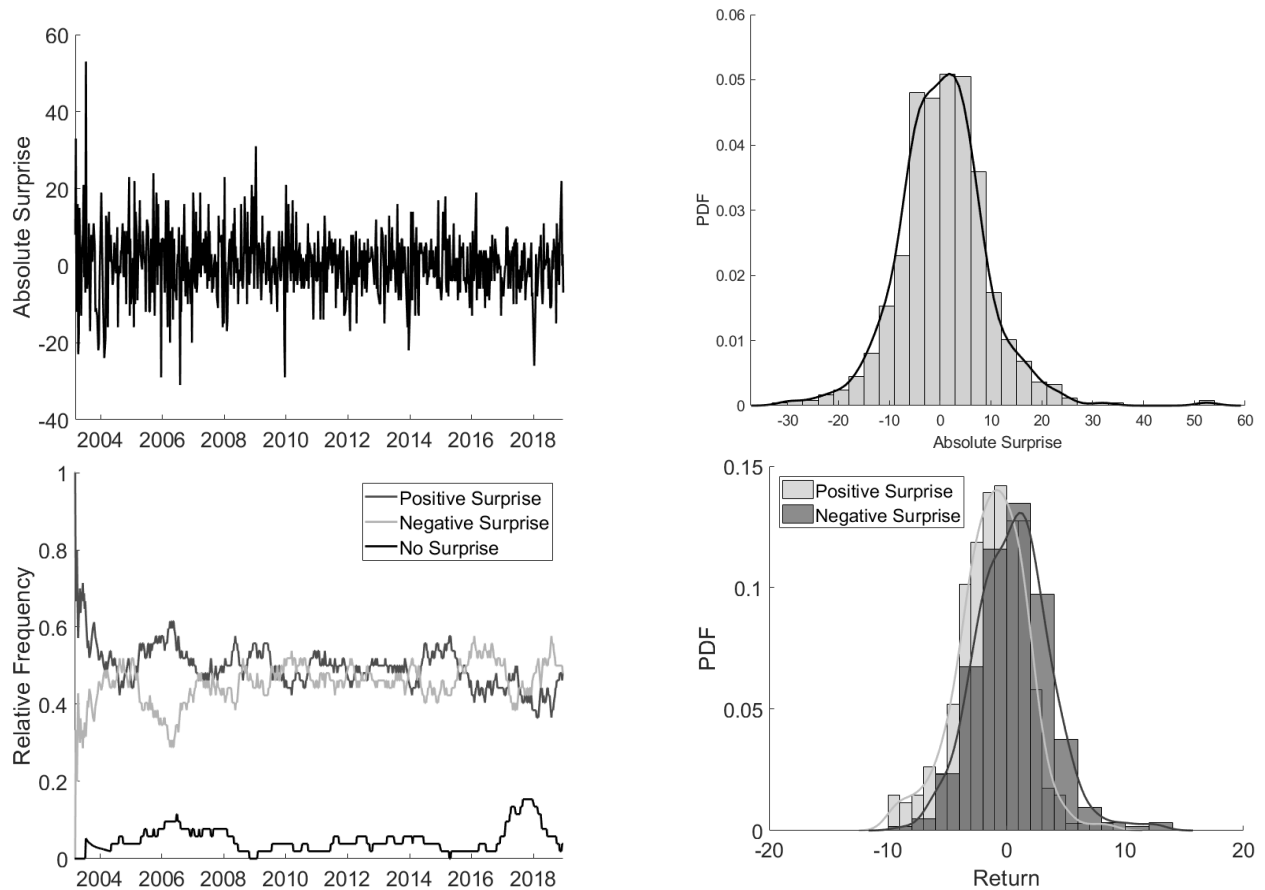
Figure A.1: Natural Gas Total Return Series

*This figure shows the total return series of the front contract in Henry Hub natural gas futures for the period from March 2003 to December 2018 obtained from Bloomberg. The dotted line represents the returns without accounting for the rolling of contracts, i.e., on rolling days returns refer to two different contracts. The solid line accounts for the rolling over in the returns. Contracts are rolled over at the end of the month preceding the month prior to delivery and scaled to have value 100 at the beginning of March 2003. The grey shaded area represents the NBER recession (December 2007 - June 2009).*



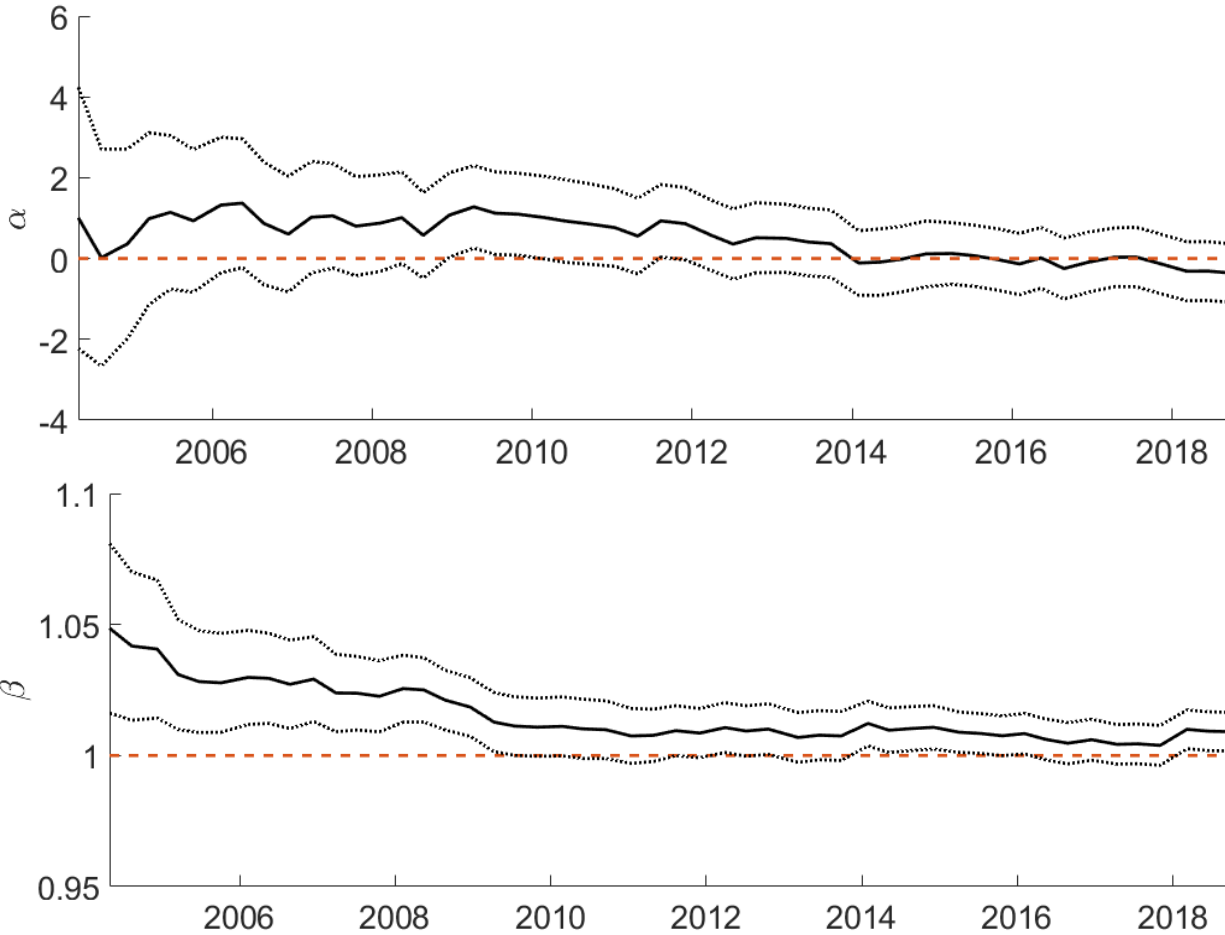
**Figure A.2: Inventory Level and First Difference**

*This figure shows the inventory level as announced by the Energy Information Administration (EIA) in the Weekly Natural Gas Storage Report every Thursday at 10:30 AM ET. The report tracks U.S. natural gas inventories held in underground storage facilities in five regions of the 48 lower states. The upper panel shows the level of inventories and the lower panel shows the change in inventory levels. Both figures are measured in billion cubic feet over the sample period from March 2003 to December 2018.*



**Figure A.3: Histogram and Densities for Inventory Surprise**

*This figure shows the time series of inventory surprises as the difference between actual and forecasted value (left-hand side, upper panel) as well as a histogram and density estimation (right-hand side, upper panel) for the distribution of the surprise. The lower left-hand side panel shows the relative frequency of positive and negative surprises and the lower right-hand side panel shows the distribution of returns for positive (blue) and negative (red) surprises.*

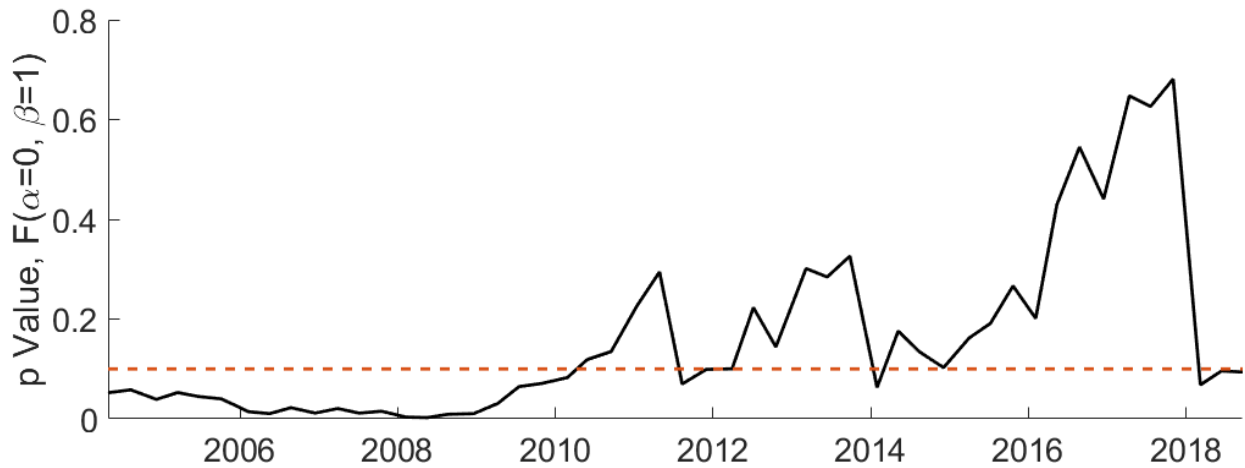


**Figure A.4: Regression Results on Forecast Accuracy**

*This figure presents the regression results of the regression*

$$A_t = \alpha + \beta E_t + u_t, \tag{1}$$

*where  $A_t$  is the actual storage reported by the EIA,  $\alpha$  is the intercept,  $\beta$  is the regression coefficient,  $E_t$  is the Bloomberg median forecast of the storage level and  $u_t$  is the residual. Regressions are run over a rolling window of 5 years (ca. 260 observations). The first panel shows the coefficient estimates for  $\alpha$  together with the bounds on a 5% confidence interval as a dotted line. The second panel shows the coefficient estimates for  $\beta$  together with the lower bound of the 5% confidence interval as a dotted line. The values for an unbiased forecast (no forecast error on average,  $\alpha = 0$  and  $\beta = 1$ ) are marked with red dashed lines.*



**Figure A.5: Regression Results on Forecast Accuracy**

*This figure presents the results of a hypothesis test on the regression*

$$A_t = \alpha + \beta E_t + u_t, \quad (2)$$

*where  $A_t$  is the actual storage reported by the EIA,  $\alpha$  is the intercept,  $\beta$  is the regression coefficient,  $E_t$  is the Bloomberg median forecast of the storage level and  $u_t$  is the residual. Regressions are run over a rolling window of 5 years (ca. 260 observations). The reported value is the p-value of a F-test on the hypothesis of an unbiased forecast ( $\alpha = 0, \beta = 1$ ). The red dashed line marks the 10% confidence level.*

## Table A.1: General Information on Natural Gas Trading and Storage Reports

*Panel A of this table summarises contract specifications of the Henry Hub natural gas futures contract used as data source in the article. Panel B gives an overview of the development of the natural gas market and the storage announcement.*

### Panel A: Henry Hub Natural Gas Futures (NYMEX)

Contract Unit	10,000 million British thermal units (MMBtu)
Minimum Tick Size	\$0.001 per MMBtu
Minimum Tick Value	\$10
Delivery Month	January- December
End of Trading	Three business days prior to the first day of the delivery month
Position Limit	1,000 contracts
Settlement	At the volume-weighted average price of all trades that are executed between 14:28:00 and 14:30:00 ET

### Panel B: History of Natural Gas Trading

1990/04	Introduction of the Henry Hub Natural Gas Futures Contract (NG) at the New York Mercantile Exchange (NYMEX)
1994/01 - 2003/02	American Gas Association (AGA) Report is released every Wednesday after market closure
2002/03 - 2002/04	AGA report is released earlier on Wednesdays during trading hours
2002/05 - today	EIA overtakes responsibility for the report and releases it Thursdays during trading hours
2003/03 - today	Bloomberg starts to publish analyst survey estimates for the Storage Report

**Table A.2: Bloomberg Data Summary**

*This table lists the data obtained from Bloomberg by ticker. The upper part before the horizontal line lists the tickers for which price series are obtained. For the lower part after the horizontal line only the release dates are obtained except the natural gas storage report for which median, average, high and low forecast, forecast dispersion and number of analysts is collected.*

Ticker	Description
SPGSCITR Index	S&P GSCI Total Return
BCOMTR Index	Bloomberg Commodity Index Total Return
USGG3M Index	US 3-month rate
USGG10YR Index	US 10-year rate
US00O/N Index	Overnight LIBOR
DOENUSCH Index	EIA Weekly Natural Gas Storage Report
DOEASCRD Index	EIA Petroleum Report Crude Storage
IPMGCHNG Index	US Industrial Production Industry Groups Manufacturing MoM
USTBTOT Index	US Trade Balance of Goods and Services SA
NHSLTOT Index	US New One Family Houses Sold Annual Total SAAR
NHSPSTOT Index	US New Privately Owned Housing Units Started by Structure Total
CICRTOT Index	Federal Reserve Consumer Credit Total Net Change SA
DGNOCHNG Index	US Durable Goods New Orders Industries MoM SA
MWINCHNG Index	Merchant Wholesalers Inventories Total Monthly % Change
CPI YOY Index	US CPI Urban Consumers YoY NSA
USPHTMOM Index	US Pending Home Sales Index MoM SA
NHSPATOT Index	Private Housing Authorized by Bldg Permits by Type Total
FDTR Index	Federal Funds Target Rate - Upper Bound
IMP1YOY% Index	US Import Price Index by End Use All YoY NSA
ETSLTOTL Index	US Existing Homes Sales SAAR
GDPCTOT% Index	US GDP Total YoY NSA



**Table A.3: Summary Statistics – Subsample Analysis**

*This table reports summary statistics of the returns on the first nearby contracts in Henry Hub Natural Gas Futures on announcement days of the EIA Weekly Gas Storage Report (Columns ‘EIA’) and non-announcement days (Columns ‘Non-EIA’). Column ‘t-Test’ reports the t-statistic and p-value in parentheses for a two-sample t-test on equal means assuming unequal variances. Column ‘F-Test’ reports the F-statistic and p-value in parentheses for a F-test on equal variances. Daily mean returns and annualized standard deviations are reported in percentage points. We split the whole sample into three subsamples in 2007 and 2014.*

Subsample	Mean			Standard Deviation		
	EIA	Non-EIA	t-Test	EIA	Non-EIA	F-Test
2003 – 2018	-0.37	-0.09	-2.19 (0.029)	49.2	43.3	1.29 (0.000)
2003 – 2007	-0.52	-0.06	-1.75 (0.081)	54.0	48.3	1.25 (0.039)
2007 – 2014	-0.48	-0.09	-1.98 (0.049)	52.4	40.7	1.65 (0.000)
2014 – 2018	-0.05	-0.14	0.46 (0.642)	37.5	42.1	0.79 (0.044)

**Table A.4: Summary Statistics – Excluding Observations**

*This table reports summary statistics of the returns on the first to sixth nearby contracts in Henry Hub Natural Gas Futures on announcement days of the EIA Weekly Gas Storage Report (Columns ‘EIA’) and non-announcement days (Columns ‘Non-EIA’). Column ‘t-Test’ reports the t-statistic and p-value in parentheses for a two-sample t-test on equal means assuming unequal variances. Column ‘F-Test’ reports the F-statistic and p-value in parentheses for a F-test on equal variances. Daily mean returns and annualized standard deviations are reported in percentage points. In Panel A, we exclude days on which the EIA has revised their estimate. In Panel B, daily returns that are larger than 10% in absolute value are excluded.*

**Panel A: Excluding First Year and Revision Dates**

Subsample	Mean			Standard Deviation		
	EIA	Non-EIA	t-Test	EIA	Non-EIA	F-Test
1	-0.37	-0.10	-2.11 (0.035)	48.0	43.2	1.23 (0.000)
2	-0.28	-0.08	-1.73 (0.084)	44.3	39.2	1.28 (0.000)
3	-0.24	-0.04	-1.82 (0.069)	39.8	35.8	1.24 (0.000)
4	-0.22	-0.04	-1.83 (0.067)	36.6	31.9	1.32 (0.000)
5	-0.20	-0.04	-1.75 (0.080)	34.3	29.7	1.33 (0.000)
6	-0.17	-0.03	-1.66 (0.097)	32.2	28.1	1.32 (0.000)

**Panel B: Excluding Extreme Returns**

Nearby	Mean			Standard Deviation		
	EIA	Non-EIA	t-Test	EIA	Non-EIA	F-Test
1	-0.42	-0.11	-2.58 (0.010)	47.6	41.3	1.33 (0.000)
2	-0.33	-0.08	-2.24 (0.025)	43.7	37.5	1.36 (0.000)
3	-0.29	-0.05	-2.33 (0.020)	39.3	34.0	1.33 (0.000)
4	-0.26	-0.05	-2.27 (0.024)	36.2	31.1	1.36 (0.000)
5	-0.23	-0.04	-2.19 (0.029)	33.8	28.9	1.37 (0.000)
6	-0.20	-0.03	-2.08 (0.038)	31.7	27.3	1.35 (0.000)

**Table A.5: Regression Alternative Surprise Measures**

*This table reports regression results of the regression*

$$R_t = \alpha_0 + \alpha_1 I_{EIA,t} + \beta_0 S_t + \beta_1' X_t + \epsilon_t, \quad (3)$$

where  $R_t$  is the first nearby log return,  $I_{EIA}$  is an indicator variable, equal to 1 on EIA days and 0 otherwise,  $S_t$  is the announcement surprise,  $X_t$  are additional exogenous variables and  $\epsilon_t$  is the residual. Hence  $\alpha_0$  is the average return on non-EIA days and  $\alpha_1$  is the return difference between EIA and non-EIA days. Column (I) includes the baseline constant and the surprise variable. In Column (II) uses the lagged surprise  $S_{t-1}$ , and column (III) uses both. Column (IV) and (V) use the alternative surprise measure

$$S_t^{disp} := \frac{A_t - E_t}{\sigma(E_t)}, \quad S_t^{rel} := \frac{A_t - E_t}{A_{t-1}}, \quad (4)$$

where  $\sigma(E_t)$  is the dispersion among forecasters for the announcement on day  $t$ , and  $A_{t-1}$  is the previous inventory level. Returns are in percentage points and  $p$ -values in parentheses are based on Newey-West standard errors with two lags.

Variables	(I)	(II)	(III)	(IV)	(V)
Intercept	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)
$I_{EIA}$	-0.24 (0.05)	-0.28 (0.03)	-0.24 (0.04)	-0.23 (0.06)	-0.21 (0.08)
$S_t$	-1.04 (0.00)		-1.04 (0.00)		
$S_{t-1}$		0.05 (0.67)	0.02 (0.83)		
$S^{disp}$				-0.88 (0.00)	
$S^{rel}$					-1.99 (0.00)
$R^2$	0.03	0.00	0.03	0.03	0.02
Obs	3982	3981	3981	3982	3982

**Table A.6: Regression – Surprise and Interacted Indicator Variables**

*This table reports regression results of the regression*

$$R_t = \alpha_0 + \alpha_1 I_{EIA,t} + \beta_0 S_t + \beta_1' X_t + \epsilon_t, \quad (5)$$

where  $R_t$  is the first nearby log return,  $I_{EIA}$  is an indicator variable, equal to 1 on EIA days and 0 otherwise,  $S_t$  is the announcement surprise,  $X_t$  are additional exogenous variables and  $\epsilon_t$  is the residual. Hence  $\alpha_0$  is the average return on non-EIA days and  $\alpha_1$  is the return difference between EIA and non-EIA days. Column (I) includes only a constant and column (II) adds the surprise variable. In Columns (III) to (VIII), the regression is augmented with the surprise variable interacted with indicator variables for low forecast dispersion ( $I_{lowSD}$ ), NBER recessions ( $I_{NBER}$ ), the injection period from April to October ( $I_{inject}$ ), the post-2009 period ( $I_{post09}$ ), and the hurricane seasons from June to November,  $I_{hurricane}$ . Column (VIII) includes all variables that were significant at the 10% level in columns (I) to (VII). Returns are in percentage points and  $p$ -values in parentheses are based on Newey-West standard errors with two lags.

Variables	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Intercept	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)	-0.09 (0.04)
$I_{EIA}$	-0.24 (0.05)	-0.24 (0.05)	-0.23 (0.05)	-0.23 (0.06)	-0.25 (0.04)	-0.23 (0.05)	-0.23 (0.05)	-0.23 (0.05)
$S$	-1.04 (0.00)	-1.02 (0.00)	-0.97 (0.00)	-0.79 (0.00)	-0.88 (0.00)	-0.92 (0.00)	-0.41 (0.13)	-0.41 (0.13)
$S \times I_{lowSD}$		-1.06 (0.12)					-0.65 (0.33)	-0.65 (0.33)
$S \times I_{NBER}$			-0.57 (0.12)				-0.96 (0.02)	-0.96 (0.02)
$S \times I_{inject}$				-0.54 (0.03)			-0.83 (0.01)	-0.83 (0.01)
$S \times I_{post09}$					-0.43 (0.07)		-0.64 (0.02)	-0.64 (0.02)
$S \times I_{hurricane}$						-0.28 (0.26)	0.27 (0.36)	0.27 (0.36)
$R^2$	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Obs	3982	3982	3982	3982	3982	3982	3982	3982

**Table A.7: Regression – Supply and Demand Variables**

*This table reports regression results of the regression using macro variables related to the economics of natural gas markets*

$$R_t = \alpha_0 + \alpha_1 I_{EIA,t} + \beta_0 S_t + \beta'_1 X_t + \epsilon_t, \quad (6)$$

where  $R_t$  is the first nearby log return,  $I_{EIA}$  is an indicator variable, equal to 1 on EIA days and 0 otherwise,  $S_t$  is the announcement surprise,  $X_t$  are additional exogenous variables and  $\epsilon_t$  is the residual. Hence  $\alpha_0$  is the average return on non-EIA days and  $\alpha_1$  is the return difference between EIA and non-EIA days. In columns (I) to (V), the regression is augmented with the deviation from the 5-year average Heating (Cooling) Degree Days,  $\Delta HDD$  ( $\Delta CDD$ ), the change in monthly US natural production,  $\Delta Production$ , the change in term spread ( $\Delta TERM$ ), which is the difference between the 3-month and 10-year U.S. Treasury Bill rate, and the change in the CBOE Volatility Index,  $\Delta VIX$ , respectively. All variables are scaled to have unit standard deviation. Returns are in percentage points and p-values in parentheses are based on Newey-West standard errors with two lags.

Variables	(I)	(II)	(III)	(IV)	(V)	(VI)
Intercept	-0.10 (0.04)	-0.10 (0.04)	-0.08 (0.08)	-0.09 (0.05)	-0.09 (0.04)	-0.08 (0.08)
$I_{EIA}$	-0.23 (0.07)	-0.24 (0.06)	-0.25 (0.04)	-0.25 (0.04)	-0.24 (0.05)	-0.25 (0.05)
$S$	-1.03 (0.00)	-1.02 (0.00)	-1.05 (0.00)	-1.04 (0.00)	-1.04 (0.00)	-1.03 (0.00)
$\Delta HDD$	0.11 (0.02)					0.11 (0.03)
$\Delta_{t+1}HDD$	-0.02 (0.72)					-0.01 (0.86)
$\Delta CDD$		0.09 (0.03)				0.09 (0.03)
$\Delta_{t+1}CDD$		-0.08 (0.06)				-0.08 (0.06)
$\Delta Production$			-0.14 (0.86)			-0.07 (0.93)
$\Delta TERM$				-0.07 (0.23)		-0.06 (0.34)
$\Delta VIX$					-0.11 (0.01)	-0.09 (0.03)
$R^2$	0.03	0.03	0.03	0.03	0.03	0.03
Obs	3669	3669	3962	3981	3981	3649

**Table A.8: Regression – Spillover Effects**

*This table reports regression results of the regression using variables related to spillover effects*

$$R_t = \alpha_0 + \alpha_1 I_{EIA,t} + \beta_0 S_t + \beta_1' X_t + \epsilon_t, \quad (7)$$

*where  $R_t$  is the first nearby log return,  $I_{EIA}$  is an indicator variable, equal to 1 on EIA days and 0 otherwise,  $S_t$  is the announcement surprise,  $X_t$  are additional exogenous variables and  $\epsilon_t$  is the residual. Hence  $\alpha_0$  is the average return on non-EIA days and  $\alpha_1$  is the return difference between EIA and non-EIA days. In columns (I) to (V), the regression is augmented with the lagged return on natural gas futures,  $R_{t-1}$ , the return on WTI crude oil futures,  $R^{WTI}$ , the return on the Goldman Sachs Commodity Index,  $R^{GSCI}$ , and the excess return on the value-weighted stock market index from CRSP,  $R^{CRSP}$ , respectively. Returns are in percentage points and  $p$ -values in parentheses are based on Newey-West standard errors with two lags.*

Variables	(I)	(II)	(III)	(IV)	(V)
Intercept	-0.10 (0.03)	-0.08 (0.06)	-0.07 (0.10)	-0.10 (0.03)	-0.05 (0.20)
$I_{EIA}$	-0.23 (0.05)	-0.29 (0.01)	-0.30 (0.01)	-0.24 (0.05)	-0.23 (0.02)
$S$	-1.05 (0.00)	-1.01 (0.00)	-0.93 (0.00)	-1.03 (0.00)	-0.78 (0.00)
$r_{t-1}$	-0.06 (0.00)				-0.07 (0.00)
$r^{WTI}$		0.35 (0.00)			-1.36 (0.00)
$r^{GSCI}$			0.74 (0.00)		2.72 (0.00)
$r^{CRSP}$				0.11 (0.00)	-0.19 (0.00)
$R^2$	0.03	0.10	0.18	0.03	0.29
Obs	3981	3982	3982	3981	3980

**Table A.9: Regression – Commodity Return Predictors**

*This table reports regression results of the regression using commodity trading signals*

$$R_t = \alpha_0 + \alpha_1 I_{EIA,t} + \beta_0 S_t + \beta_1' X_t + \epsilon_t, \quad (8)$$

where  $R_t$  is the first nearby log return,  $I_{EIA}$  is an indicator variable, equal to 1 on EIA days and 0 otherwise,  $S_t$  is the announcement surprise,  $X_t$  are additional exogenous variables and  $\epsilon_t$  is the residual. Hence  $\alpha_0$  is the average return on non-EIA days and  $\alpha_1$  is the return difference between EIA and non-EIA days. In columns (I) to (V), the regression is augmented with the front slope of the futures curve,  $B_{1,2}$ , the slope between the futures contracts with same expiry month one year ahead,  $B_{1,13}$ , the hedging pressure, HP, the idiosyncratic volatility, IVOL, and the change in trading volume in thousand transactions,  $\Delta$ Volume, respectively. Returns are in percentage points and  $p$ -values in parentheses are based on Newey-West standard errors with two lags.

Ind. Var.	(I)	(II)	(III)	(IV)	(V)	(VI)
Intercept	-0.02 (0.64)	0.01 (0.80)	-0.08 (0.18)	0.28 (0.03)	-0.08 (0.07)	0.46 (0.00)
$I_{EIA}$	-0.23 (0.05)	-0.24 (0.05)	-0.24 (0.05)	-0.24 (0.04)	-0.27 (0.03)	-0.27 (0.03)
$S$	-1.04 (0.00)	-1.05 (0.00)	-1.04 (0.00)	-1.04 (0.00)	-1.04 (0.00)	-1.04 (0.00)
$b^{(1,2)}$	0.44 (0.00)					0.42 (0.00)
$b^{(1,13)}$		1.99 (0.00)				
HP			0.15 (0.70)			0.31 (0.43)
IVOL				-0.14 (0.01)		-0.18 (0.00)
$\Delta$ Volume					0.08 (0.13)	0.08 (0.15)
$R^2$	0.03	0.03	0.03	0.03	0.03	0.03
Obs	3982	3982	3982	3982	3844	3844

**Table A.10: Summary Statistics Excluding Macro News**

*This table reports mean and standard deviation of the returns on the first nearby contract in Henry Hub Natural Gas Futures on EIA announcement days excluding days on which the report coincides with other macroeconomic news releases. The first column lists the event to be excluded. Columns ‘News’ represent those days where only EIA reports are released, columns ‘Rest’ include all other days including those where the EIA report coincides with the release mentioned in the first column. Column ‘t-Test’ reports the t-statistic and p-value in parentheses for a two-sample t-test on equal means assuming unequal variances. Column ‘F-Test’ reports the F-statistic and p-value in parentheses for a F-test on equal variances. The last column reports the number of announcements excluding the event. The first row reports the base line results only excluding coinciding release days of the EIA Petroleum Report. The last three rows represent exclude days on which any news on the housing market, consumption or the macro economy are excluded. Daily mean returns and annualized standard deviations are reported in percentage points. The sample ranges from March 2003 to December 2018.*

Excluded Event	Mean			Standard Deviation			Obs
	News	Rest	t-Test	News	Rest	F-Test	
EIA Petroleum Report	-0.37	-0.09	-2.19 (0.029)	49.15	43.28	1.29 (0.000)	699
Industrial Production	-0.38	-0.09	-2.30 (0.022)	48.76	43.40	1.26 (0.000)	679
Trade Balance	-0.37	-0.10	-2.08 (0.038)	49.60	43.23	1.32 (0.000)	670
New House Sales	-0.39	-0.09	-2.28 (0.023)	49.62	43.21	1.32 (0.000)	677
New Housing Units	-0.34	-0.10	-1.85 (0.065)	49.24	43.32	1.29 (0.000)	674
Consumer Credit	-0.37	-0.09	-2.21 (0.028)	48.91	43.37	1.27 (0.000)	679
Durable Goods	-0.35	-0.10	-1.94 (0.052)	49.82	43.19	1.33 (0.000)	670
Wholesale Inventories	-0.39	-0.09	-2.32 (0.020)	49.17	43.29	1.29 (0.000)	688
Consumer Price Index	-0.38	-0.09	-2.23 (0.026)	48.46	43.49	1.24 (0.000)	674
Pending Home Sales	-0.36	-0.10	-2.02 (0.044)	49.10	43.35	1.28 (0.000)	672
Housing Permits	-0.34	-0.10	-1.85 (0.065)	49.24	43.32	1.29 (0.000)	674
Fed Announcements	-0.33	-0.10	-1.80 (0.072)	49.30	43.30	1.30 (0.000)	674
Import Index	-0.39	-0.09	-2.27 (0.024)	48.79	43.43	1.26 (0.000)	661
Existing Home Sales	-0.40	-0.09	-2.42 (0.016)	49.31	43.29	1.30 (0.000)	672
GDP	-0.35	-0.10	-1.95 (0.052)	49.10	43.34	1.28 (0.000)	680
Housing Market	-0.37	-0.10	-2.00 (0.046)	49.89	43.32	1.33 (0.000)	598
Consumption	-0.37	-0.09	-2.21 (0.028)	48.91	43.37	1.27 (0.000)	679
Macro Economy	-0.34	-0.11	-1.71 (0.088)	49.82	43.33	1.32 (0.000)	604



**Table A.11: Intraday Return Regressions**

*This table reports regression results of the regressio using intraday returns*

$$R_t = \alpha_0 + \alpha_1 I_{EIA,t} + \beta_0 S_t + \beta_1' X_t + \epsilon_t, \quad (9)$$

*where  $R_t$  is the first nearby log return,  $I_{EIA}$  is an indicator variable, equal to 1 on EIA days and 0 otherwise,  $S_t$  is the announcement surprise,  $X_t$  are additional exogenous variables and  $\epsilon_t$  is the residual. Hence  $\alpha_0$  is the average return on non-EIA days and  $\alpha_1$  is the return difference between EIA and non-EIA days. The dependent variable changes in every column, starting with the intraday return from 90 minutes before the announcement to 30 minutes after the announcement,  $(-90,30)$ . The second, third and fourth column use the intraday return from 60, 30 and 5 minutes before the announcement to 30 minutes after the announcement as dependent variable, or  $(-60,30)$ ,  $(-30,30)$  and  $(-5,30)$ , respectively. Results for the exogenous variables are not reported, they include the crude oil returns, commodity index returns, the basis, idiosyncratic volatility and changes in Volume. Returns are in percentage points and  $p$ -values in parentheses are based on Newey-West standard errors with two lags.*

Dep. Var.	(-90,30)	(-60,30)	(-30,30)	(-5,30)
Constant	0.14 (0.04)	0.13 (0.03)	0.12 (0.03)	0.09 (0.07)
$I_{EIA}$	-0.30 (0.00)	-0.22 (0.00)	-0.16 (0.00)	-0.15 (0.00)
Surprise	-1.01 (0.00)	-0.97 (0.00)	-0.93 (0.00)	-0.87 (0.00)
Basis	-0.03 (0.51)	0.01 (0.87)	0.00 (0.96)	-0.03 (0.38)
IVOL	-0.06 (0.01)	-0.05 (0.02)	-0.05 (0.02)	-0.03 (0.04)
Control	Yes	Yes	Yes	Yes
$R^2$	0.15	0.16	0.17	0.16
Obs	3979	3979	3979	3979

**Table A.12: Regression Bloomberg Survey Forecast Accuracy**

*This table reports regression results for the forecast accuracy of the Bloomberg median survey forecast as described*

$$A_t = \alpha + \beta E_t + u_t, \tag{10}$$

*where  $A_t$  denotes the weekly storage reported by the EIA,  $\alpha$  is the intercept,  $\beta$  is the regression coefficient,  $E_t$  is the Bloomberg median forecast of the storage level, and  $u_t$  is the residual. Column (I) uses the raw figure for  $A_t$  and  $E_t$ . Column (II) uses the seasonally-adjusted figures, removing the 5-year average for the specific week. The second to last row reports the  $F$ -statistic for the hypothesis of  $\alpha = 0$  and  $\beta = 1$  with the  $p$ -value reported in parentheses.*

Variables	(I)	(II)
Intercept ( $\alpha$ )	0.2385 (0.460)	0.3243 (0.313)
Forecast ( $E_t$ )	1.0119 (0.000)	1.0132 (0.000)
$F(\alpha = 0, \beta = 1)$	6.32 (0.002)	7.72 (0.000)
$R^2$	0.99	0.99

**Table A.13: Regression Spread Returns**

*This table reports regression results of the regression*

$$R_t = \alpha_0 + \alpha_1 I_{EIA,t} + \beta_0 S_t + \beta_1' X_t + \epsilon_t, \quad (11)$$

where  $R_t$  is the log return on the spread between the first and second nearby,  $I_{EIA}$  is an indicator variable, equal to 1 on EIA days and 0 otherwise,  $S_t$  is the announcement surprise,  $X_t$  are additional exogenous variables and  $\epsilon_t$  is the residual. Hence  $\alpha_0$  is the average return on non-EIA days and  $\alpha_1$  is the return difference between EIA and non-EIA days. Column (I) includes only a constant and column (II) adds the surprise variable. In column (III) we also control for crude oil returns, commodity index returns, idiosyncratic volatility, and changes in Volume. Returns are in percentage points and  $p$ -values in parentheses are based on Newey-West standard errors with two lags.

Variables	(I)	(II)	(III)
Constant	-0.02 (0.01)	-0.02 (0.01)	-0.02 (0.04)
$I_{EIA}$	-0.06 (0.00)	-0.06 (0.00)	-0.07 (0.00)
$S$		-0.06 (0.00)	-0.04 (0.05)
Control	No	No	Yes
$R^2$	0.00	0.01	0.07
No. of Obs.	3982	3982	3844

## A Factor Construction

This section contains instruction on how to construct the factor returns for the computation of the idiosyncratic volatility measure, IVOL.

The market factor is computed as the equally-weighted sum of 26 commodity market returns, excluding natural gas itself, i.e.,

$$R_{\text{MRKT},t} = \frac{1}{27} \sum_{i=1}^{27} R_{i,t}, \quad (12)$$

where  $R_{i,t}$  is the daily return on commodity market  $i$ . For the basis, momentum, and basis-momentum factors the 26 markets are sorted regarding the respective signal and divided along the median into two portfolios of 13 commodities. The factor returns evolves as the long-short return on the equally-weighted portfolio returns for the upper and lower half. The signals for basis, momentum, and basis-momentum are computed as follows:

$$B_{i,t} = \left( \frac{F_{i,t}^1}{F_{i,t}^2} \right)^{\frac{365}{M_{i,t}^2 - M_{i,t}^1}} - 1,$$

$$\text{MOM}_{i,t} = \sum_{j=1}^{252} R_{i,t-j}^1$$

$$\text{BMOM}_{i,t} = \sum_{j=1}^{252} R_{i,t-j}^1 - \sum_{j=1}^{252} R_{i,t-j}^2$$

where  $F_{i,t}^1$  ( $F_{i,t}^2$ ) is the futures price of the first (second) nearby,  $M_{i,t}^1$  ( $M_{i,t}^2$ ) is the time to maturity in days of the first (second) nearby, and  $R_{i,t}^1$  ( $R_{i,t}^2$ ) is the return on the first (second) nearby.