The natural gas announcement day puzzle

Article

Accepted Version

Prokopczuk, M., Wese Simen, C. and Wichmann, R. (2021) The natural gas announcement day puzzle. Energy Journal, 42 (2). ISSN 1944-9089 doi: https://doi.org/10.5547/01956574.42.2.mpro Available at https://centaur.reading.ac.uk/90003/

The Natural Gas Announcement Day Puzzle*

Marcel Prokopczuk[†,§] Chardin Wese Simen[‡] Robert Wichmann[§]

November 25, 2019

Abstract

This paper studies natural gas futures returns on EIA storage announcement days. More than 50% of the annual return is earned on these days. We find a significant difference between announcement and non-announcement day returns, which cannot be explained by the announcement surprise or other control variables. At the intraday level, the return splits half into a pre- and post-announcement part. The pre-announcement return is entirely generated on days when storage levels exceed analysts' expectations casting doubt on explanations based on informed trading. After transaction and funding cost, a simple trading strategy yields substantial returns.

JEL Classifciation: G14, Q02, Q43

Keywords: Gas Markets, Announcement Effect, Storage News, Intraday Analysis

*We would like to thank four anonymous referees for their thoughtful comments. We also benefited from constructive comments by the participants of the Annual Meeting of the German Finance Association 2019 in Essen, the Commodity and Energy Markets Annual Meeting 2019 in Pittsburgh, the Commodity Winter Workshop 2019 in Hannover, and the Henley Doctoral Conference 2018 in Reading. Contact: prokopczuk@fmt.uni-hannover.de (M. Prokopczuk), c.wese-simen@liverpool.ac.uk (C. Wese Simen), and r.c.wichmann@pgr.icmacentre.ac.uk (R. Wichmann).

†School of Economics and Management, Leibniz University Hannover, Königsworther Platz 1, 30167 Hannover, Germany

‡Management School, University of Liverpool, L69 7ZH, Liverpool, United Kingdom

§ICMA Centre, Henley Business School, University of Reading, RG6 6BA, Reading, United Kingdom

I Introduction

The natural gas market has undergone massive changes throughout the last decades, starting with its deregulation in the 1980s, the inception of the futures market in 1990, the inflow of financial investors at the beginning of the twenty-first century, and recent shifts in supply and demand due to the introduction of shale gas, a growing industry for liquefied natural gas (LNG) as well as increased attention related to climate change. Natural gas storage levels have always been an important indicator of changes due to their natural role as a buffer between supply and demand. As such, release of the Weekly Natural Gas Storage Report by the Energy Information Administration (EIA), which contains information about the current storage level, draws attention from all market participants. When new information is released to an efficient market, participants adjust their expectations and prices accordingly. Figure 1 shows that more than 50% of the annual return of natural gas futures is generated on weekly EIA announcement days. Therefore, returns on natural gas futures are significantly different on EIA announcement days compared to non-announcement days. However, after controlling for the information of the announcement this difference should disappear.

This article documents a significant difference between the average returns observed on EIA announcement days and non-announcement days. Puzzlingly, this difference in returns between announcement days and non-announcement days cannot be explained by the information content of the announcement. Indeed, we find a strong significant negative relationship between natural gas futures returns and the announcement surprise, but we cannot explain the return difference between announcement and non-announcement days. This result is robust after augmenting the model with supply and demand measures, spillover effects from options, energy or equity markets as well as commodity specific variables such as the slope of the futures curve, hedging pressure, liquidity or volatility measures.

At the intraday level, we decompose the return within a two hour window surrounding the announcement into a pre- and post-announcement part. Curiously, the overall return divides equally into the pre-announcement part (49.4%) and the post-announcement part (50.6%). Albeit modest evidence for the leakage of information, this can only be a partial explanation as there is still a significant effect from the announcement. Lastly, we document that the pre-announcement return is entirely realized on days where the announcement surprise is positive, i.e., the published inventory exceeds analysts' expectations. The asymmetry of this result casts doubt on a simple explanation based on informed trading.

From the perspective of an investor, this puzzling result raises the question whether the newly documented premium is economically large once transaction and funding costs are accounted for. Our results show that the simple strategy of opening a short position 90 minutes before the announcement and closing it 30 minutes afterwards yields a significant annual return of 12% (t-stat = 2.93) translating into a Sharpe ratio of 1.76 after transaction and funding costs. However, the time series of strategy returns and the accuracy of analysts' forecasts suggests that the anomaly has decreased in magnitude and efficiency has returned to natural gas markets, leaving open the possibility that our strategy was new to investors who are now arbitraging it away.

Our work contributes to the literature on storage effects in energy markets. Linn and Zhu (2004) show that the intradaily volatility of natural gas futures is significantly higher in the hour surrounding the American Gas Association (AGA) report and this effect has carried on after the EIA took on the reporting. Gay et al. (2009) show that the announcement return is negatively related to the inventory surprise, i.e., when the reported inventory level is higher than analysts' expectations, futures returns tend to be negative and vice versa. Halova et al. (2014) find seasonal patterns relating to the withdrawal period from November to March and the injection period from April to October. During winter, when inventories are lower than on average, inventory shocks have a smaller effect on futures returns, while the effect is stronger in summer. They also find that the effect is weaker, when forecast dispersion is higher which in general is the case in winter where demand shocks due to weather are an important driver of energy prices. Chiou-Wei et al. (2014) show that the announcement effect is unique to the day of the announcement. Bu (2014), Ye and Karali (2016), and Miao et al. (2018) find similar results for oil and gasoline using the EIA Petroleum Report announcements. Ederington et al. (2019) revisit these studies, and find that analysts' natural gas forecasts efficiently impound the available time-series information but crude oil forecasts do not. Demirer and Kutan (2010) and Schmidbauer and Rösch (2012) study the effect of OPEC announcements on crude oil markets. Wolfe and Rosenman (2014) show that announcements in oil and gas markets cause spillover effects to each other. Compared to studies for other energy markets and studies on the effect of crop reports on agricultural commodities (Adjemian, 2012; Mattos and Silveira, 2016), focussing on the EIA Weekly Natural Gas Storage Report provides a unique setting. The report only publishes storage information without any supplementary information on supply, demand, or future prospects of production, hence the effect can be clearly referred to the changes in inventory.

Our work also relates to the broader literature on the effects of scheduled news on energy prices. Basistha and Kurov (2015) study the effect of Federal Open Market Committee (FOMC) announcements on energy prices. For crude oil, Kilian and Vega (2011) and Chatrath et al. (2012) find no evidence to suggest that energy prices respond to macroeconomic news. They conclude that crude oil prices are predetermined with respect to macro aggregates, confirming the view of Kilian (2009), that prices are determined by flow supply and flow demand.

Moreover, our work adds to the growing literature on analyzing risk premia on announcement days. Savor and Wilson (2013) find that 60% of the annual equity risk premium can be earned by only investing when important macroeconomic news is released. They interpret this finding as the premium investors demand for bearing macroeconomic risk. Ai and Bansal (2018) develop a theoretical framework that explains the announcement premium with the generalised risk sensitivity of investors used as evidence for a class of non-expected utility models. Relative to these studies, we focus on announcements of natural gas inventories, which are presumably asset specific news. We find a sizeable premium on these days suggesting a risk premium for idiosyncratic news.

The intraday analysis in this article is related to the work of Lucca and Moench (2015), who document the pre-FOMC announcement drift in the US equity market. As Brusa et al. (2019) show, the Fed is unique in channelling such an effect compared to other central banks. The EIA report plays a similar role for the US natural gas market. Gu and Kurov (2018) find a pre-announcement drift, and link it to informed trading caused by superior forecasting abilities of certain participants. Rousse and Sévi (2019) find evidence of an asymmetric response of crude oil returns to the EIA Petroleum Report. Our study reveals that the documented pre-announcement effect in the natural gas market is asymmetric and only accounts for half of the entire return, and therefore casts doubt on an explanation based on informed trading.

Lastly, our work relates to the literature on the pricing of commodity futures. Brown and Yücel (2008) show that natural gas markets are driven by weather, inventories and spill-overs from crude oil markets. Besides these supply and demand driven factors, we relate to the growing literature on factor models for commodity futures that include hedging pressure (De Roon et al., 2000), open interest (Hong and Yogo, 2012), idiosyncratic volatility (Fernandez-Perez et al., 2016) or the slope of the futures curve (Szymanowska et al., 2014). We confirm earlier studies that the listed variables affect natural gas returns. However, they are not able to explain the EIA announcement effect.

The remainder of this article is organised as follows. Section II describes the data and introduces the main variables. Section III documents the EIA announcement effect and explores possible explanations. Section IV looks at the intraday frequency. Section V discusses robustness checks. Finally, Section V concludes.

II Data & Variables

From Bloomberg, we obtain the daily price, trading volume and open interest series of 499 Henry Hub Natural Gas futures contracts (Ticker NG) from March 2003 to December 2018.[1] Since we are dealing with futures contracts, we need to construct an investable price series by rolling over contracts before expiry.[2] We follow Szymanowska et al. (2014) and rollover the entire curve at the end of the month preceding the month prior to delivery, i.e., the return on the futures price series is defined as

$$R_t^n := \begin{cases} P_t^n - P_{t-1}^{n+1}, & \text{if } t-1 \text{ is a roll over day} \\ P_t^n - P_{t-1}^n, & \text{otherwise,} \end{cases} \tag{1}$$

where P_t^n is the log price of the n^{th} nearby on day t. We provide summary statistics on the returns of the first six nearby contracts in Table 1, that confirm common characteristics of natural gas markets. We find a strongly negative average return of -31.65%, which is in line with other studies (de Groot et al., 2014; Paschke et al., 2017). Further, we see high volatility of up to 45% per annum and a decreasing pattern of volatility in line with the Samuelson effect.

We collect the main variables of interest, the EIA figure for the total storage level and the respective median forecast also from Bloomberg. Natural gas storage levels and changes show strong seasonal patterns resulting from demand cycles. Storage levels decrease during the withdrawal period from November to March as the demand of gas for heating in winter exceeds the less volatile supply, before inventories build up again during the injection period from April to October. Since the economics of natural gas markets are known to market participants, these patterns are also included in analysts' forecasts. Therefore, the actual new information of the announcement is the deviation of the announced figures from the market's expectation. As a proxy of this market expectation on future storage levels, we use the Bloomberg median survey forecast. It is regarded as a proxy for the market expectation by academics and practitioners (Chiou-Wei et al., 2014) and provides additional information beyond seasonal and historic patterns to market participants (Ederington et al., 2019). We

[1] The start of the sample is motivated by the inception of the Bloomberg forecast for the EIA report. For more details on the contract specifications and a time line, see Table A.1 of the Online Appendix. An alternative approach would be to use physical spot data. Unfortunately, spot trading involves a number of costs that can affect the response of spot prices to news. In order to guard against this criticism, we focus on the futures market.

[2] This is an important distinction to the pure price series of the front contract or the spot price (Singleton, 2014). It leads to large differences between the actual price series and a constructed total return series that uses the realized returns on a rolled series. For illustration, Figure A.1 of the Online Appendix shows the spot price series as well as the futures price series constructed from the rolled return series.

define the non-scaled announcement surprise, S^{level}, as

$$S_t^{level} := A_t - E_t \tag{2}$$

where A_t is the announced inventory level on day t and E_t is the market expectation as measured by the median survey forecast. To normalise the surprise measure, we follow Andersen et al. (2003) and define the standardised announcement surprise as

$$S_t := \frac{S_t^{level}}{\sigma(S^{level})} \tag{3}$$

where $\sigma(S^{level})$ is the standard deviation of the forecast error. For robustness, we also use a relative and a dispersion-adjusted surprise measure defined as

$$S_t^{\text{rel}} := \frac{S_t^{level}}{A_{t-1}}, \qquad S_t^{\text{disp}} := \frac{S_t^{level}}{\sigma(E_t)}, \tag{4}$$

where A_{t-1} is the previous inventory level and $\sigma(E_t)$ is the dispersion among forecasters for the announcement on day t.[3] We provide summary statistics for inventory reports, the median survey forecast and the surprise measures in Table 2.[4]

Further, we obtain weather data on Heating Degree Days (HDD) and Cooling Degree Days (CDD) from the American Gas Association (AGA). Heating Degree Days are a measure of the coldness of the weather experienced, based on the extent to which the daily mean temperature falls below a reference temperature (65° F).[5] Cooling Degree Days are a measure of the need for air conditioning (cooling) based on the extent to which the daily mean temperature rises above a reference temperature. The AGA Heating Degree Day Report contains heating and cooling degree data aggregated on a weekly basis for nine census regions and the US.

Lastly, we collect financial variables and macroeconomic variables as well as other macroeconomic announcements and their survey forecasts from Bloomberg. A detailed list of tickers is provided in Table A.2 in the Online Appendix. We also collect the Commitment of Traders (CoT) report for natural gas from the Commodity Futures Trading Commission (CFTC).

[3]Note that in 3, $\sigma(S^{level}) = \sigma(A_t - E_t)$ is the standard deviation over the whole sample, hence a constant that scales S to have unit standard deviation, while A_{t-1} and $\sigma_t(E)$ are not a constant, but a varying denominator in Equation (4).

[4]For further information see also histograms and density plots in Figure A.3 of the Online Appendix.

[5]The daily mean temperature is computed as the sum of the high and the low readings divided by two.

III The EIA Announcement Effect

III.A The Facts

Every Thursday at 10:30 AM Eastern Time (ET) the EIA releases the Weekly Natural Gas Storage Report, which lists the underground storage net changes for five regions of the United States.[6] The report provides fundamental information to the natural gas markets. Therefore, it is interesting to see how prices behave on such EIA announcement days compared to non-announcement days.

Table 3 reports summary statistics on EIA announcement days and non-announcement days, respectively. For the first nearby contract, the mean daily return on EIA announcement days is −0.37%, while it is only −0.09% on the remaining days. The difference is statistically significant at the 5% level. Annualized volatility on announcement days is 49.2% compared to 43.3% on non-announcement days, with the difference being statistically significant at the 1% level.[7] The return and volatility differentials prevail also for longer maturities. Figure 1 shows that more than 50% of the negative average return on natural gas futures is earned on EIA announcement days, with this figure being even larger for more deferred contracts. A two–sample Kolmogorov–Smirnov test in Panel B of Table 3, rejects the hypothesis of both subsamples stemming from the same distribution. Overall, the data provide strong evidence of natural gas prices behaving differently on EIA announcement days as opposed to non-announcement days.

The fact, that natural gas markets behave differently on EIA announcement days, itself is to be expected as these days provide fundamental information to the market. However, the interesting question is whether the observed differences can be explained by information unique to these days.

[6]If national holidays like Thanksgiving, Christmas or Independence day fall on a Thursday, the report is released on Wednesday or Friday. However, the release schedule is known in advance, so all announcements are scheduled. We exclude observations on which the announcement of the report coincides with the Weekly Petroleum Status Report by the EIA, that is usually published on Wednesdays.

[7]We also provide a subsample analysis of the effect in Table A.3 of the Online Appendix that shows that the effect is not present in the most recent period (2014–2018), which saw a sharp decline in energy prices. In the former periods (2003–2007 and 2007–2014), however, the effect was even stronger. The results are further robust towards excluding extreme returns or revision dates as reported in Table A.4 of the Online Appendix.

III.B Potential Explanations

III.B.1 The Announcement Surprise

We start with three intuitive explanations for the documented return difference. First, one could think that the news on EIA announcement days are on average 'bad', i.e. positive surprises, and hence the more negative effect. If that were the case, the surprise should be significantly different from zero. Although we find a slightly positive surprise on average, it is not significantly different from zero (t-stat = 0.95). Second, it could be the case that positive surprises are on average larger, and therefore have a stronger effect. If that were the case, we should find a significant difference in the absolute value of positive and negative surprises. However, we find the difference in means to be indistinguishable from zero (t-stat = 0.69). Third, there could be a different effect on returns between negative and positive surprises. If this was the case, the demeaned announcement returns should be larger in absolute value on days with a positive surprise. In fact, we do find the opposite, returns are slightly larger on negative surprise days, although the difference is not significant (t-stat = 1.37).

Having ruled out these explanations, we want to quantify the effect of the announcement surprise on the return difference in the following regression,

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

where R_t is the log return on the first nearby contract, $I_{EIA,t}$ is an indicator variable with value 1, if t is an EIA announcement day and 0 otherwise, S_t is the announcement surprise as defined in Equation (3) and set to zero for non-announcement days, X_t is a set of control variables, and ϵ_t is the error term.[8] The coefficient of interest is α_1, as it represents the difference in average returns between announcement and non-announcement days after controlling for the announcement surprise and other potential explanatory variables.

We use several sets of control variables for X_t in Equation (5), and report the main results in Table 4. The first column only includes a constant and the indicator variable I_{EIA}, α_1 therefore represents the difference in means as documented before ($-0.37\% - (-0.09\%) = -0.28\%$). In the second column, we add the surprise variable, and confirm earlier results by Halova et al. (2014). A one standard deviation surprise in inventories decreases futures prices by 1.04%, the return differnce however reduces only slightly to -0.24%.[9] The remaining

[8]Note that because returns also differ in volatility between announcement and non-announcement days, we adjust for this heteroskedasticity by scaling the residuals on non-EIA days by the fraction of volatilities between EIA and non-EIA days. Therefore, the standard errors we obtain are most conservative.

[9]These results are robust to the definition of the announcement surprise, see Table A.5 in the Online Appendix.

columns present similar results while controlling for several alternative channels. We discuss the different channels in the following paragraphs, and present detailed regression results for the columns (III) to (VI) of Table 4 in the Online Appendix, Table A.6 to Table A.9.

III.B.2 Asymmetric Effects

Table A.6 of the Online Appendix reports the result of the baseline regression in Equation (5) using the surprise interacted with indicator variables as controls. This way, we test whether the return difference is due to asymmetric price reactions to the surprise under specific conditions. The results show that in times of low forecast dispersion, i.e. when analysts' opinions are less diverging, the surprise effect doubles, since a given deviation comes as a bigger surprise. The same is true during the injection period from April to October, when demand is rather stable compared to higher volatility in winter. During these times, markets are more supply-driven and production is naturally easier to foresee than the demand side which is heavily dependent on weather. Therefore, an equally-sized surprise will have a larger effect during the injection period. We further find stronger effects during recessions and after 2009.[10] We do not find significant differences during the hurricane season from June to November.

Altogether, we find that the announcement surprise explains part of the announcement return, but still leaves a significant negative average return on announcement days.

III.B.3 Supply, Demand and Market Conditions

As an important energy market, the natural gas market is affected by general economic conditions as well as supply and demand forces. By augmenting our model with the necessary macroeconomic variables, we can see whether these general market conditions are driving the return difference on EIA announcement days. The demand on natural gas is highly dependent on weather because of its use for heating and cooling. To measure the influence of weather, we use Heating Degree Days (HDD) and Cooling Degree Days (CDD) obtained from the American Gas Association (AGA). Since these variables are highly seasonal due to winter and summer periods, we deseasonalise the variables using the five year average for

[10]Halova et al. (2014) identify December 2009 as a structural break point following the modification of the sample selection and estimation procedure by the EIA. This finding is also in line with Dehnavi et al. (2015) who study changes in the natural gas market due to the increasing role of LNG.

each week, i.e.,

$$\Delta \text{HDD}_t = \text{HDD}_t - \frac{1}{5}\sum_{j=1}^{5} \text{HDD}_{t-52\cdot j}, \tag{6}$$

$$\Delta \text{CDD}_t = \text{CDD}_t - \frac{1}{5}\sum_{j=1}^{5} \text{CDD}_{t-52\cdot j}, \tag{7}$$

where t is a weekly subscript. Because Cooling Degree Days are only measured from April to October and Heating Degree Days only for November to March, we set all remaining values to zero. To account for the fact that expected temperature might be more important than current temperature, we also include the variable led by one week. For the supply side, we collect the monthly change in US natural gas production from the EIA.[11]

Energy markets, as a fundamental part of the economy, are exposed to general financial conditions. Therefore, we augment the model with the term spread (TERM), defined as the difference in yields between a 3-month and a 10-year US Treasury bill, to measure economic conditions and we include changes in the Chicago Board Options Exchange (CBOE) Volatility Index (VIX) to capture general market uncertainty. All variables are scaled by their standard deviation to obtain comparable coefficient estimates.

Table A.7 of the Online Appendix summarizes the results adding the above mentioned variables one by one. We find some evidence for the effect of weather on natural gas prices as well as a negative effect of increased market volatility. However, the annual return difference remains significantly negative at around -0.25%.

III.B.4 Spillover Effects from other Markets

The growing integration of commodity markets also referred to as financialization (Tang and Xiong, 2012; Cheng and Xiong, 2014; Basak and Pavlova, 2016) as well as links between commodity markets and equity markets increasingly affects also natural gas markets. Further, its role as a substitute for energy production links natural gas markets to the oil market. Wolfe and Rosenman (2014) find a bidirectional causal relationship between the two markets, which could also bias the announcement effect. Brown and Yücel (2009) and Dehnavi et al. (2015) study gas-to-gas arbitrage between Europe and the US through the development of the market for liquified natural gas (LNG).

Therefore, the effect might be driven by events in larger markets, that spill over to the

[11] We acknowledge, that the use of these variables is limited due to their weekly and monthly frequency, reducing the likelihood of influencing returns at the daily frequency. We thank a reviewer for suggesting to use the leading weather variables to account for forecasts.

natural gas market. To account for such an effect, we augment the model in Equation (5) with the return on the first nearby contract of West Texas Intermediate (WTI) crude oil futures, as traded on the New York Mercantile Exchange (Ticker CL). On a broader level, we also include the return on the Standard & Poors Goldman Sachs Commodity Index (SPGSCI) as a proxy for overall commodity market returns.[12] Within the natural gas market, we account for reversal or momentum effects by including the lagged return into the regression. Lastly, we add the excess return on the value-weighted market index from the Center for Research in Security Prices (CRSP) as a proxy for stock market returns.

Table A.8 of the Online Appendix presents the regression results. All variables have a highly significant effect on natural gas returns. We find a reversal effect within the natural gas markets, while crude oil, commodity index and stock market returns are positively related to natural gas return. Altogether, the variables can explain nearly 30% of the variation in natural gas returns. However, they do not crowd out the EIA announcement effect. We still find a significant daily return difference of -0.23% between announcement and non-announcement days.

III.B.5 Commodity Return Predictors

Apart from market integration, the financialization of commodity markets has also increased index investing and the rise of rule-based strategies. The literature has identified several factors that can predict commodity returns and serve as trading signals. The slope of the futures curve or basis serves as an indicator for the scarcity of the commodity and hence predicts commodity returns (Szymanowska et al., 2014). In times of backwardation, when the spot price exceeds the futures price, inventories shrink as it is more profitable to sell than to store. We define the basis as

$$B_{1,2}(t) = \left(\frac{F_t^1}{F_t^2}\right)^{\frac{365}{M_t^2 - M_t^1}} - 1, \tag{8}$$

where F_t^1 (F_t^2) is the price of the first (second) nearby, and M_t^1 (M_t^2) denotes the time to maturity of the first (second) nearby.[13]

Another force driving futures risk premia is the positions of traders (Hirshleifer, 1990). Producers (hedgers) who need to hedge their production are offering a risk premium to

[12] We favour the SPGSCI for its larger loading on energy markets here. As a robustness check, we also include the less energy-weighted Bloomberg Commodity Index. We find similar results.

[13] Alternatively, we also use a seasonality-adjusted basis by replacing the second with the thirteenth nearby such that, $B_{1,13}(t) = \left(\frac{F_t^1}{F_t^{13}}\right)^{\frac{365}{M_t^{13} - M_t^1}} - 1$, where F_t^{13} and M_t^{13} are the price and time to maturity of the thirteenth nearby contract. This way the basis measures the price differential between two contracts with the same expiry month.

speculators for agreeing to enter the futures contract. Using the Commodity Futures Trading Commission (CFTC) report on the Commitment of Traders (CoT), we contruct the hedging pressure following De Roon et al. (2000)

$$\mathrm{HP}_t = \frac{\#\text{short hedge positions}_t - \#\text{long hedge positions}_t}{\#\text{hedge positions}_t}, \tag{9}$$

such that HP_t is a relative measure of the direction in which producers are hedging, with $\mathrm{HP}_t = 1$ indicating only short positions and $\mathrm{HP}_t = -1$ indicating only long positions.[14]

We include a measure of idiosyncratic volatility, which has shown to have predictive power for commodity returns by Fernandez-Perez et al. (2016). The factor is constructed from the residuals of the time-series regression,

$$R_t = a + b' F_t + \epsilon_t, \tag{10}$$

where a is the intercept, b is the vector of sensitivities towards the factors F_t, ϵ_t is the residual, and $\mathrm{IVOL}_t = \sigma(\epsilon_t)$ is the 30-day standard deviation of the residuals. We use a 4-factor model including an equally-weighted commodity market return, as well as the returns on long-short portfolios sorted by basis, momentum, and basis-momentum.[15] Lastly, we also control for changes in the volume traded to include a possible liquidity channel.

The results are presented in Table A.9 of the Online Appendix. We find a strong positive relationship between the basis and returns, i.e. returns are higher, when the market is in backwardation. This is in line with the literature relating the basis to inventory and thus reflecting the markets' expectation on prices (Gorton et al., 2013). We also confirm the negative pricing of idiosyncratic volatility as reported by Fernandez-Perez et al. (2016). However, none of the variables carries enough daily variation to explain the difference in returns between EIA announcement and non-announcement days.

III.B.6 Macroeconomic News

To isolate the effect of the EIA storage report on natural gas returns, we excluded those days on which the report coincides with the EIA petroleum report. However, another possibility is that other important news are coinciding with the EIA report and hence the econometrician might mistake the observed effect for the EIA announcement effect when in fact it is other

[14]Since the CFTC only reports these figures on a weekly basis, again this variable is only measured at the weekly frequency.

[15]This way, we include the most recent studies on commodity return predictors by Bakshi et al. (2019) and Boons and Prado (2019). Details on how the factors are constructed can be found in Section A of the Online Appendix.

macro news.

Table A.10 of the Online Appendix reports summary statistics of the returns on EIA announcement days and non-announcement days excluding days on which EIA announcements coincide with other events. We find that the return difference remains significant after excluding other events, even when aggregating news on certain topics, e.g., all days with news on consumption and prices or macroeconomic indices.

IV Intraday Analysis

Having established, that there is a significant return difference between EIA announcement and non-announcement days, it is interesting to see how this return emerges throughout an announcement day. For this purpose, in this section we employ 5-minute data obtained from Thomson Reuters Tick History over the same sample period, and decompose the return into different parts.

IV.A Intraday Return Decomposition

The graphs in Figure 2 show the average volume traded and the return volatility across time for every 5-minute interval. On EIA annuoncement days there is a clear spike in volume and volatility at exactly 10:30 AM with volumes sixfold and volatility fivefold compared to non-announcement days.[16] The patterns indicate an immediate and short-lived reaction.

To see how the return is distributed around the announcement, we decompose the daily announcement return into the return from 90 minutes before to 30 minutes after the announcement, $(-90, 30)$, and the remaining parts from market closure of the previous day to 90 minutes before the announcement, $(C_{t-1}, -90)$, and from 30 minutes after the announcement to market closure of the announcement day, $(30, C_t)$. Panel A of Table 5 shows that the entire effect (99%) stems from the two hour window surrounding the announcement.

Panel B of Table 5 decomposes the intraday return from 90 minutes before to 30 minutes after the announcement into a pre-announcement return, $(-90, -5)$, and a post-announcement return $(-5, 30)$.[17] If the reason for the announcement return were a pre-announcement drift because of informed trading (Gu and Kurov, 2018) or information leakage (Rousse and Sévi, 2019), we should find that the post-announcement return is not significantly different from

[16]Note that the volume also spikes at 14:30 p.m., but this is due to the fact that daily prices are settled at the volume-weighted average price of all trades that are executed between 14:28:00 and 14:30:00 p.m. ET. Another indication for this not having any price effect is that there is no complementing spike in volatility during the same period.

[17]To clarify the wording, we will always refer to the post-announcement return as including the actual announcement.

zero, since the information is already priced before it is officially announced and prices do not react. Instead the effect should be entirely absorbed before the announcement. However, Panel B of Table 5 reveals that 49.4% of the return is generated before the announcement and 50.6% after the announcement. This bisection of the return is puzzling as it neither proves a pre-announcement drift nor does it show that the effect is only generated after the announcement.[18]

IV.B Regression Analysis

To control for the effects discussed in the previous section, we repeat the regression analysis and regress intraday returns on a constant, the indicator variable I_{EIA}, the announcement surprise, S_t and control variables X_t as in Equation (5).[19]

We report the results using different dependent variables in Table 6. For the $(-90, 30)$ return, we find an even stronger result than for the daily returns with a return difference of -0.30% between EIA and non-EIA days after controlling for announcement surprise and other effects (-0.24% for daily returns). Splitting up the return into a pre- and post-announcement part, we again find that the return difference halves into -0.15% for the pre-announcement return, $(-90, -5)$, and -0.15% for the post-announcement return, $(-5, 30)$.[20] More surprisingly, we find a significant negative relationship between the announcement surprise and the pre-announcement return, indicating leakage of the information.[21] This is puzzling as previous literature does not find evidence of leakage (Bjursell et al., 2015; Ederington et al., 2019). However, this can only be a partial explanation as there is still a significant surprise effect when the actual announcement is made.

Taking a closer look at the pre-announcement return, $(-90, -5)$, we report the returns conditional on the sign of the surprise in Panel A of Table 7. Surprisingly, we find that the pre-announcement return is only significantly different from non-EIA intraday returns when the surprise is positive. There is a negative effect of -0.36%, significant at the 1% level, when the announcement surprise is positive. If the surprise is negative, there is no significant price reaction before the event, suggesting that the pre-announcement drift is only identifiable, when storage levels exceed expectations. This result is puzzling as it does not

[18] There is no evidence that this effect has been caused by price limits being hit before the EIA announcement.

[19] We use the variables that have shown significant effects on returns in the previous section, i.e., the basis, idiosyncratic volatility, as well as spillovers from the previous day, oil markets, commodity markets and stock markets.

[20] Results for smaller windows of 60 or 30 minutes in Table A.11 of the Online Appendix show that the effect steadily decreases.

[21] Control variables such as the basis or idiosyncratic volatility, that are significant at the daily level, do not influence intraday returns.

align with a story of superior forecasting ability. Assuming informed traders extract additional information from the Bloomberg forecast that leads them to anticipate the surprise, the pre-announcement drift should show up independent of the sign of the surprise.

Since the phenomenon is unique to positive surprises, we want to investigate how it evolved over time by looking into three subsamples from 2003 to the beginning of the financial crisis in December 2007, from then until the peak of oil prices after the crisis in June 2014, as well as the most recent period.[22] The results in Panel B of Table 7 show, that the pre-announcement effect has been even stronger in the past with average values of −0.38% and −0.47% for the first and second period, respectively. However, it has more than halved to only −0.15% in the most recent period after 2014. At the same time returns on negative surprise days have increased from −0.08% to 0.08%.

V What About ...?

V.A Forecasting Accuracy

The previous section has shown that the pre-announcement drift only occurs for positive surprises. Therefore, it is crucial to see how accurate the Bloomberg median survey forecast is, as it decides, whether the surprise is positive, i.e., the reported storage level exceeds the forecasted value. We test the accuracy of the median forecast by regressing the actual reported values on the median forecast such that

$$A_t = \alpha + \beta E_t + u_t, \tag{11}$$

where A_t is the actual reported value, α is the intercept, β is the regression coefficient, E_t is expected value or median forecast, and u_t is the residual.

If the analysts predicted the natural gas storage without any bias, the intercept of the regression should be equal to zero, and the coefficient for the median forecast should be one. The results in Table A.13 of the Online Appendix reject the hypothesis of α being significantly different from zero, and find a regression coefficient β that is slightly larger than one.[23] Although we find the intercept not to be significantly different from zero, and the estimates for β being close to unity, a joint F-test of the hypothesis, $\alpha = 0$ and $\beta = 1$, is rejected at the 1% level.

Considering that the EIA has changed the sample selection procedure in 2008 (Halova et al., 2014) and industry forecasts have improved throughout, we also carry out the above

[22] We thank an anonymous referee for the suggestion of splitting the sample.

[23] Note that the p-value for β refers to a test on whether β equals zero and is therefore not informative.

analysis using a rolling window of 5 years (approximately 260 observations) to see how estimates have changed over time. Figure A.4 of the Online Appendix shows the estimates for α and β and Figure A.5 the p-value of the F-test on $\alpha = 0$ and $\beta = 1$. The graphs for the coefficient estimates both show a trend towards the values for unbiased prediction highlighted in red. At the same time, the p-value of the F-test on $\alpha = 0$ and $\beta = 1$ is increasing throughout the sample and is not rejected at the 10% for the first time at the beginning of 2010 coinciding with growing attention of the literature on pre-announcement drifts (Lucca and Moench, 2015).

V.B Spreads

An interesting question is whether the documented return difference is related to the maturity of the contract. From the previous analysis, we know that while the magnitude of the effect is smaller in absolute terms for deferred contracts (see Table 3), it amounts to a higher part of the annual return (see Figure 1). Therefore, it is not clear a priori whether we should find a significant return difference on EIA days, if we were to repeat the analysis from before using the spread return between the first and second nearby as the dependent variable.

Results are reported in Table A.12 of the Online Appendix. We find a strongly significant negative return difference for spreads on EIA days of -0.06% resulting in a five times larger return on EIA days. This magnitude is similar in relative terms to what we find for the first nearby return. Again this return remains significant after controlling for the announcement surprise and other channels. This is another interesting result because the first and second nearby contract are written on the same underlying only differing by expiry date.

V.C Limits to Arbitrage

For practitioners, it is especially interesting to investigate whether the observed effect can be exploited and the return withstands funding and transaction costs. We follow the strategy, to open a short position 90 minutes before the announcement, and to close it 30 minutes after the announcement. This way, we can harvest the pre-announcement return and the effect from the announcement while minimizing the investment window, reducing funding costs.

Table 8 reports the returns on the described strategy. The first row of Panel A reports the raw return on the futures, which amounts to an annual average of 17.86% with a Sharpe ratio of 2.57. However, this is based on returns from the settlement prices. To incorporate trading cost, we instead use the bid and ask quotes, hence when we open the short position, we sell the futures contract at the last bid, and when we close the position, we buy back at the last ask. The return in the second row in Panel A of Table 8 incorporates these costs and shows, that the return is reduced to 12.21% per annum, still securing a Sharpe ratio of 1.79. Lastly, we also take into account the funding cost of the position. Although futures contracts can be entered holding only a fraction of the contract value, since we are opening a short position and for robustness, we consider a fully funded position. We use the Overnight London Interbank Offer Rate (ONR), which serves as a globally accepted key benchmark interest rate that indicates borrowing costs between banks. The reported return in the third row drops further to 12.01%, with a Sharpe ratio of 1.76.[24]

Panel B and C of Table 8 show that the strategy has worked much better in the past, securing annual returns of 25% after transaction and funding costs with a Sharpe ratio of 3.26. In the more recent period, this has declined to only 3% and a Sharpe ratio of 0.5, which do not withstand transaction and funding costs. Figure 4 shows the risk-adjusted five year moving average return of the strategy, controlling for the returns on a commodity market portfolio and long-short portfolios using basis, momentum and basis-momentum factors. We see a constantly significant excess return during the first 10 years of the sample, gradually decreasing to become insignificantly different from zero in the most recent period.

VI Conclusion

We study the relationship between inventory news and the natural gas market, and find a significant return difference between announcement and non-announcement days, that can neither be explained by the announcement surprise, nor after controlling for general market conditions, spillover effects, commodity return predictors, or coinciding macroeconomic news. One half of the return is generated as a pre-announcement effect, that is unique to positive surprises, while the other half is realized after the announcement. These results are puzzling and have three interesting implications for academics and practitioners.

First, the fact that the announcement days of the EIA storage report account for more than 50% of the annual return on natural gas futures and even more than 60% for more deferred contracts, should attract the attention of investors and regulators. It opens the possibility to harvest a significant amount of the annual return on natural gas futures without committing capital for more than 20% of the year.[25] At the same time, it calls for increased attention of regulators towards ensuring that the information is not released before the

[24]Note that while returns are annualized using a multiplier of 52 to represent the realisable return within a year, the corresponding Sharpe ratio is annualized with a multiplier of $\sqrt{252}$, since means and standard deviation are based on daily returns.

[25]This percentage is based on investing only 1 of 5 days a week and could be even further reduced to less than 5%, if we allow for intraday trading around the announcement.

announcement and the information is gathered avoiding any possible bias.

Second, the significantly negative average return for natural gas returns on EIA announcement days poses a challenge to the academic literature, as it cannot be explained by the announcement surprise, market specific variables, spillover effects or factor investing. The negative sign is important since it also opposes the interpretation as a premium investors demand for bearing the risk of holding the asset during uncertain events (Savor and Wilson, 2013).

Third, the time-series dimension of the effect suggests a decline in the recent period coinciding with an improved forecasting accuracy. However, the simple strategy of opening and closing a short position before and after the announcement yields an annual return of 12% after transaction and funding costs. The gradual decrease of this return suggests that it has not been exploited by arbitrageurs but rather disappeared over time, challenging the idea of an efficient market that does not allow for such anomalies and once they are encountered, adapts immediately.

References

Adjemian, M. K. (2012). Quantifying the WASDE announcement effect. *American Journal of Agricultural Economics*, 94(1):238–256.

Ai, H. and Bansal, R. (2018). Risk preferences and the macroeconomic announcement premium. *Econometrica*, 86(4):1383–1430.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review*, 93(1):38–62.

Bakshi, G., Gao, X., and Rossi, A. (2019). Understanding the sources of risk underlying the cross section of commodity returns. *Management Science*, 65(2):619–641.

Basak, S. and Pavlova, A. (2016). A model of financialization of commodities. *Journal of Finance*, 71(4):1511–1556.

Basistha, A. and Kurov, A. (2015). The impact of monetary policy surprises on energy prices. *Journal of Futures Markets*, 35(1):87–103.

Bjursell, J., Gentle, J. E., and Wang, G. H. (2015). Inventory announcements, jump dynamics, volatility and trading volume in US energy futures markets. *Energy Economics*, 48:336–349.

Boons, M. and Prado, M. P. (2019). Basis-momentum. *Journal of Finance*, 74(1):239–279.

Brown, S. P. and Yücel, M. K. (2008). What drives natural gas prices? *Energy Journal*, 29(2):45–60.

Brown, S. P. and Yücel, M. K. (2009). Market arbitrage: European and North American natural gas prices. *Energy Journal*, 30(Special Issue):167–185.

Brusa, F., Savor, P. G., and Wilson, M. I. (2019). One central bank to rule them all. *Saïd Business School WP 2015-13; Fox School of Business Research Paper, Available at SSRN 2658953*, 13.

Bu, H. (2014). Effect of inventory announcements on crude oil price volatility. *Energy Economics*, 46:485–494.

Chatrath, A., Miao, H., and Ramchander, S. (2012). Does the price of crude oil respond to macroeconomic news? *Journal of Futures Markets*, 32(6):536–559.

Cheng, I.-H. and Xiong, W. (2014). Financialization of commodity markets. *Annual Review of Financial Economics*, 6(1):419–441.

Chiou-Wei, S.-Z., Linn, S. C., and Zhu, Z. (2014). The response of US natural gas futures and spot prices to storage change surprises: Fundamental information and the effect of escalating physical gas production. *Journal of International Money and Finance*, 42:156–173.

de Groot, W., Karstanje, D., and Zhou, W. (2014). Exploiting commodity momentum along the futures curves. *Journal of Banking & Finance*, 48:79–93.

De Roon, F. A., Nijman, T. E., and Veld, C. (2000). Hedging pressure effects in futures markets. *Journal of Finance*, 55(3):1437–1456.

Dehnavi, J., Wirl, F., and Yegorov, Y. (2015). Arbitrage in natural gas markets? *International Journal of Energy and Statistics*, 3(4):1550018.

Demirer, R. and Kutan, A. M. (2010). The behavior of crude oil spot and futures prices around OPEC and SPR announcements: An event study perspective. *Energy Economics*, 32(6):1467–1476.

Ederington, L. H., Lin, F., Linn, S. C., and Yang, L. Z. (2019). EIA storage announcements, analyst storage forecasts, and energy prices. *Energy Journal*, 40(5):121–142.

Fernandez-Perez, A., Fuertes, A.-M., and Miffre, J. (2016). Is idiosyncratic volatility priced in commodity futures markets? *International Review of Financial Analysis*, 46:219–226.

Gay, G. D., Simkins, B. J., and Turac, M. (2009). Analyst forecasts and price discovery in futures markets: The case of natural gas storage. *Journal of Futures Markets*, 29(5):451–477.

Gorton, G. B., Hayashi, F., and Rouwenhorst, K. G. (2013). The fundamentals of commodity futures returns. *Review of Finance*, 17(1):35–105.

Gu, C. and Kurov, A. (2018). What drives informed trading before public releases? Evidence from natural gas inventory announcements. *Journal of Futures Markets*, 38(9):1079–1096.

Halova, M. W., Kurov, A., and Kucher, O. (2014). Noisy inventory announcements and energy prices. *Journal of Futures Markets*, 34(10):911–933.

Hirshleifer, D. (1990). Hedging pressure and futures price movements in a general equilibrium model. *Econometrica*, 58(2):411–428.

Hong, H. and Yogo, M. (2012). What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics*, 105(3):473–490.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–69.

Kilian, L. and Vega, C. (2011). Do energy prices respond to us macroeconomic news? A test of the hypothesis of predetermined energy prices. *Review of Economics and Statistics*, 93(2):660–671.

Linn, S. C. and Zhu, Z. (2004). Natural gas prices and the gas storage report: Public news and volatility in energy futures markets. *Journal of Futures Markets*, 24(3):283–313.

Lucca, D. O. and Moench, E. (2015). The pre-FOMC announcement drift. *Journal of Finance*, 70(1):329–371.

Mattos, F. L. and Silveira, R. L. (2016). Futures price response to crop reports in grain markets. *Journal of Futures Markets*, 36(10):923–942.

Miao, H., Ramchander, S., Wang, T., and Yang, J. (2018). The impact of crude oil inventory announcements on prices: Evidence from derivatives markets. *Journal of Futures Markets*, 38(1):38–65.

Newey, W. K. and West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 28(3):777–787.

Paschke, R., Prokopczuk, M., and Wese Simen, C. (2017). Curve momentum. *Working Paper*.

Rousse, O. and Sévi, B. (2019). Informed trading in the WTI oil futures market. *Energy Journal*, 40(2):139–160.

Savor, P. and Wilson, M. (2013). How much do investors care about macroeconomic risk? Evidence from scheduled economic announcements. *Journal of Financial and Quantitative Analysis*, 48(2):343–375.

Schmidbauer, H. and Rösch, A. (2012). OPEC news announcements: Effects on oil price expectation and volatility. *Energy Economics*, 34(5):1656–1663.

Singleton, K. J. (2014). Investor flows and the 2008 boom/bust in oil prices. *Management Science*, 60(2):300–318.

Szymanowska, M., Roon, F., Nijman, T., and Goorbergh, R. (2014). An anatomy of commodity futures risk premia. *Journal of Finance*, 69(1):453–482.

Tang, K. and Xiong, W. (2012). Index investment and the financialization of commodities. *Financial Analysts Journal*, 68(6):54–74.

Wolfe, M. H. and Rosenman, R. (2014). Bidirectional causality in oil and gas markets. *Energy Economics*, 42:325–331.

Ye, S. and Karali, B. (2016). The informational content of inventory announcements: Intraday evidence from crude oil futures market. *Energy Economics*, 59:349–364.

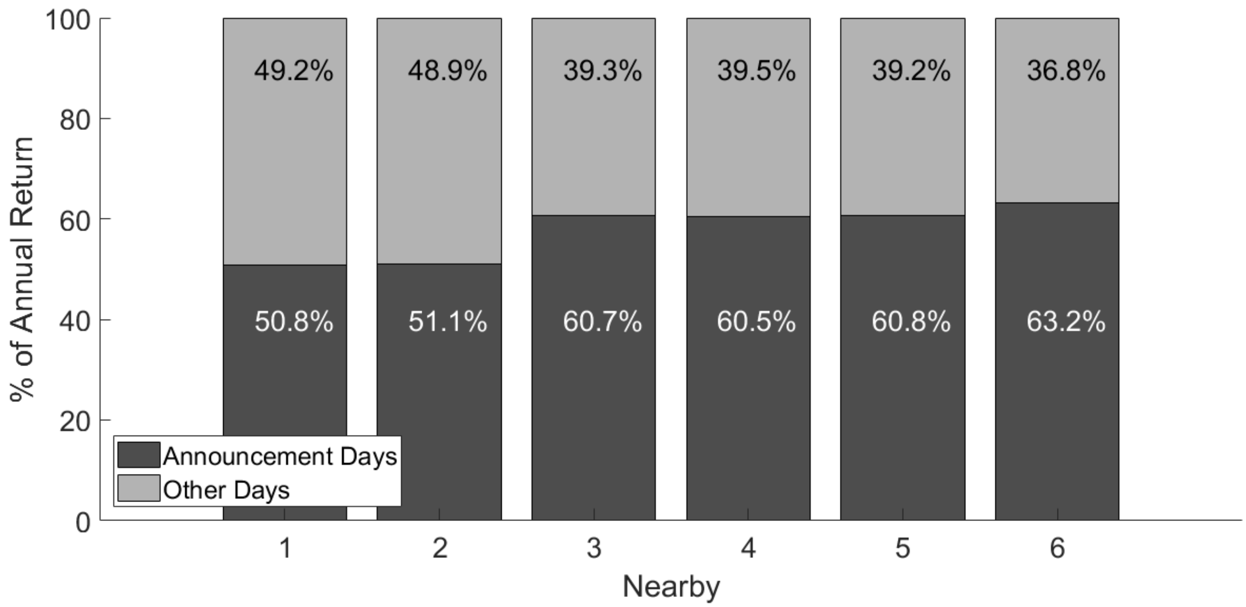

Figure 1: Decomposition of Annual Natural Gas Futures Returns

This figure shows the decomposition of the annual return on the first to sixth nearby in Henry Hub Natural Gas Futures between days on which the Weekly Natural Gas Storage Report is published by the Energy Information Administration (EIA) (dark lower bar) and non-announcement days (light upper bar), i.e., the percentages are computed as

$$\blacksquare = \frac{52 \cdot \bar{R}_{EIA}}{52 \cdot \bar{R}_{EIA} + 200 \cdot \bar{R}_{Non\text{-}EIA}} \quad \textit{and} \quad \square = \frac{200 \cdot \bar{R}_{Non\text{-}EIA}}{52 \cdot \bar{R}_{EIA} + 200 \cdot \bar{R}_{Non\text{-}EIA}}$$

where $\bar{R}_{EIA}$ and $\bar{R}_{Non\text{-}EIA}$ are the average daily returns on EIA days and Non-EIA days, respectively, and 52 is the number of EIA days per year (weekly), which leaves 200 other trading days. The percentage contribution of announcement days is written in white inside the lower bar and the percentage contribution of non-announcement days is written in black inside the top bars. The sample period comprises daily returns from March 2003 to December 2018 (3982 days), that decompose into 699 EIA announcement days, excluding days, where the report coincides with the EIA Petroleum Report, and 3283 non-announcement days.

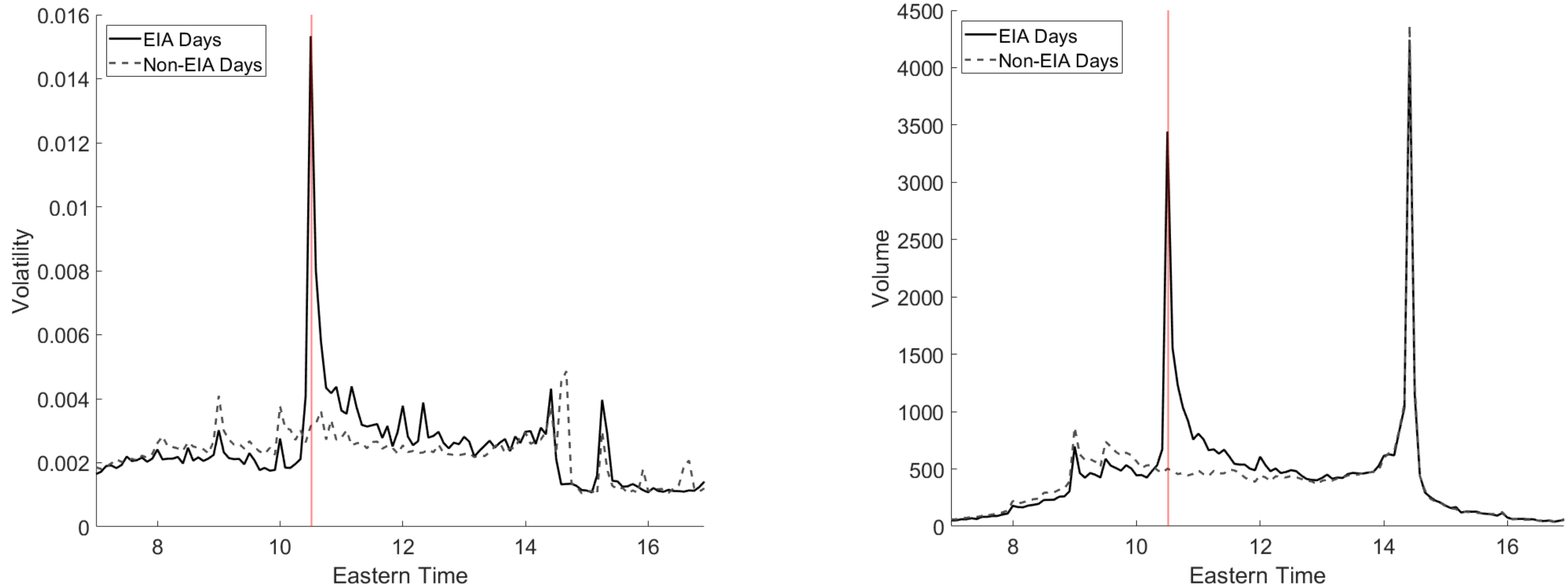

Figure 2: Volatility and Volume on EIA Storage Report Announcement Days

The left panel of this figure shows the return volatility across time for every 5-minute interval on EIA announcement days (black solid line) and non-announcement days (grey dahsed line). The right panel of this figure shows the average volume traded for every 5-minute interval on EIA announcement days (black solid line) and non-announcement days (grey dahsed line). The horizontal axis covers the main trading window from 7:00 AM ET to 17:00 ET. The vertical red line marks the publication of the report at 10:30 AM ET. The underlying dataset comprises the period from March 2003 to December 2018, which includes 699 announcements.

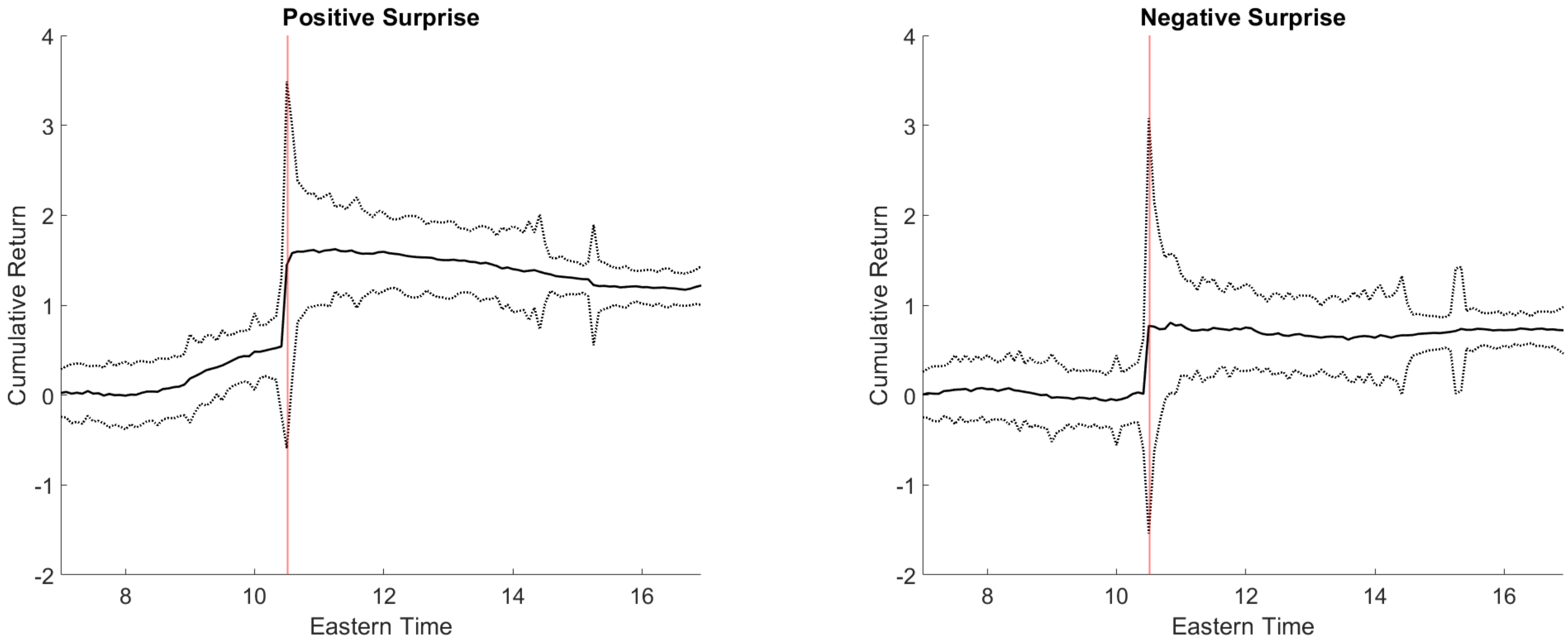

Figure 3: Intraday Cumulated Returns for Positive and Negative Surprise

This figure shows the cumulated average 5-minute interval return of natural gas futures on EIA announcement days. The left panel shows the returns for days on which the announced figure exceeded the Bloomberg median forecast (positive surprise) and the right panel shows the returns for days on which the announced figure fell short of the Bloomberg median forecast (negative surprise). For comparison the return for positive surprises is mirrored at the x-axis. The horizontal axis covers the main trading window from 7:00 AM ET to 17:00 ET. The vertical red line marks the publication of the report at 10:30 AM ET. The underlying dataset comprises the period from March 2003 to December 2018, which includes 699 announcements from which 345 mark a positive and 319 mark a negative surprise.

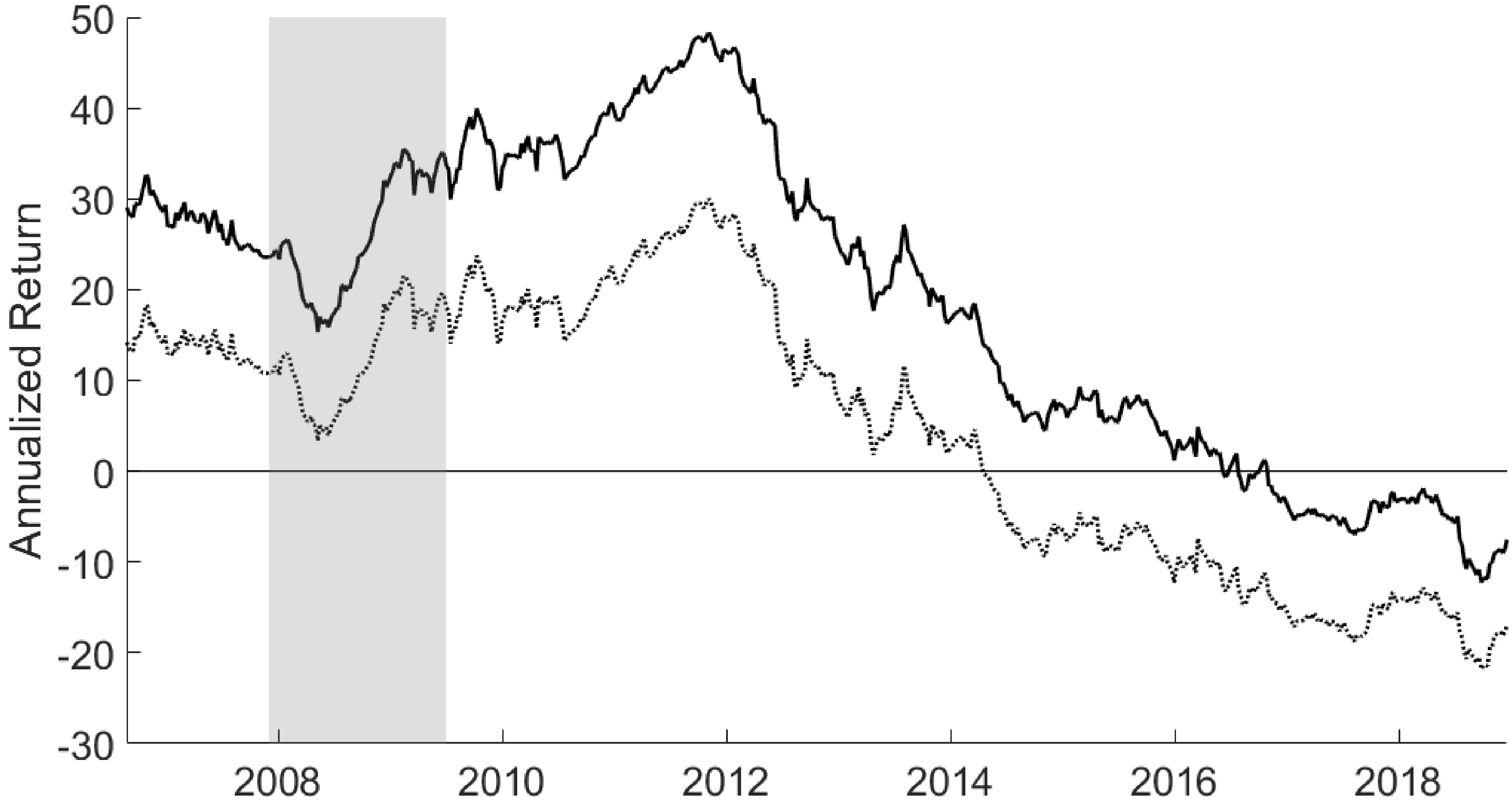

Figure 4: Three Year Moving Average Risk-Adjusted Returns on Investment Strategy

This figure shows the three year moving average risk-adjusted return on the investment strategy that opens a short position 90 minutes before the announcement and closes the position 30 minutes after the announcement. The returns are adjusted for their exposure to a 4-factor model including an equally-weighted commodity market return, as well as the returns on long-short portfolios sorted by basis, momentum, and basis-momentum, see Bakshi et al. (2019) and Boons and Prado (2019) for details. The black line shows the annualized moving average return and the dashed line shows the one-sided 5% confidence interval.

Table 1: Summary Statistics of Natural Gas Returns

This table reports summary statistics of daily log returns for the first six nearby contracts in Henry Hub Natural Gas Futures for the period from March 2003 to December 2018. Contracts have been rolled over at the end of the month preceding the month prior to delivery. Column 'n' denotes the order of nearby, column 'Mean' reports the annualized mean return, columns 'Min' and 'Max' report the minimal and maximal daily return, column 'Std. Dev.' reports the annualized standard deviation and column 'SR' the annualized Sharpe ratio. Column 'AR(1)' reports the first order autocorrelation. The columns 'Skew' and 'Kurt' report skewness and kurtosis of the returns and column 'JB' reports the p-value of the Jarque-Bera test for normality. Returns and standard deviations are reported in percentage points.

n	Mean	Min	Max	Std. Dev.	SR	AR(1)	Skew	Kurt	JB
1	-31.65	-19.18	18.76	44.99	-0.70	-0.06	0.15	5.88	0.00
2	-24.74	-20.21	17.13	40.71	-0.61	-0.05	0.07	6.21	0.00
3	-17.32	-21.80	18.63	36.81	-0.47	-0.05	0.08	7.49	0.00
4	-15.89	-12.05	10.58	32.95	-0.48	-0.04	0.07	4.55	0.00
5	-14.13	-11.03	10.31	30.67	-0.46	-0.04	0.04	4.57	0.00
6	-11.43	-11.08	10.13	28.96	-0.39	-0.04	0.02	4.75	0.00

Table 2: Summary Statistics of Inventory, Forecast and Surprise

This table reports summary statistics on the EIA Natural Gas Storage report and its Bloomberg survey forecast for the period from March 2003 to December 2018 (699 observations). The first two rows report statistics on the level and change of the natural gas storage in billion cubic feet. The third and fourth row report statistics for the median and average forecast values of the Bloomberg median survey forecast. The fifth row reports on the forecast dispersion between the different survey analysts. The last three rows report the statistics for the non-scaled announcement surprise, S^{level}, the normalised surprise, S, the relative surprise, S^{rel}, and the dispersion-adjusted surprise, S^{disp}, as defined in Equations (2),(3) *and* (4)*:*

$$S_t^{level} = A_t - E_t, \qquad S_t = \frac{S_t^{level}}{\sigma(S_t^{level})}, \qquad S_t^{rel} := \frac{S_t^{level}}{A_{t-1}}, \qquad S_t^{disp} := \frac{S_t^{level}}{\sigma(E_t)}, \qquad (2,\ 3\ \&\ 4)$$

where A_t is the announced inventory level on day t, E_t is the median forecast, $\sigma(S^{level})$ is the standard deviation of the forecast error, A_{t-1} is the previous inventory level and $\sigma(E_t)$ is the dispersion among forecasters on day t. The columns report mean, median, standard deviation, first order autocorrelation, skewness, and kurtosis.

	Mean	Median	Std. Dev.	AR(1)	Skew	Kurt
Inventory Level	2553.92	2614.00	793.69	0.98	-0.24	2.23
Inventory Change	6.35	43.00	93.93	0.88	-1.10	3.25
Median Forecast	6.04	45.00	92.44	0.90	-1.09	3.25
Average Forecast	6.06	44.00	92.21	0.90	-1.09	3.24
Forecast Dispersion	6.87	6.00	3.52	0.54	1.88	8.98
Surprise Level	0.31	0.00	8.57	-0.03	0.29	5.77
Normalised Surprise	0.04	0.00	1.00	-0.03	0.29	5.77
Relative Surprise	0.03	0.00	0.48	-0.04	0.79	18.25
Dispersion-adjusted Surprise	0.06	0.00	1.32	-0.05	0.81	8.59

Table 3: Summary Statistics on EIA Announcement Days

This table reports summary statistics of the log returns on the first six 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. Column 'KS-Test' reports the test-statistic and p-value in parentheses for a two–sample Kolmogorov–Smirnov test on different distributions. Means and standard deviations are reported in percentage points, standard deviations are annualized. The sample includes 3982 daily returns from March 2003 to December 2018.

Panel A: First and Second Moment

Nearby	**Mean**			**Standard Deviation**		
	EIA	Non-EIA	t-Test	EIA	Non-EIA	F-Test
1	-0.37	-0.09	-2.19 (0.029)	49.2	43.3	1.29 (0.000)
2	-0.28	-0.07	-1.84 (0.067)	45.2	39.2	1.33 (0.000)
3	-0.24	-0.04	-1.94 (0.052)	40.6	35.6	1.30 (0.000)
4	-0.22	-0.04	-1.93 (0.054)	37.3	31.7	1.38 (0.000)
5	-0.20	-0.03	-1.87 (0.062)	34.7	29.5	1.39 (0.000)
6	-0.17	-0.03	-1.74 (0.082)	32.6	27.9	1.37 (0.000)

Panel B: Third and Fourth Moment

Nearby	**Skewness**		**Kurtosis**		KS-Test
	EIA	Non-EIA	EIA	Non-EIA	
1	0.10	0.11	4.17	5.79	0.09 (0.000)
2	0.14	0.09	4.22	6.42	0.09 (0.000)
3	0.10	0.14	4.21	8.32	0.08 (0.001)
4	0.02	0.14	4.13	4.23	0.08 (0.002)
5	0.01	0.10	4.12	4.29	0.08 (0.003)
6	0.06	0.06	4.30	4.37	0.07 (0.008)

Table 4: Regression Announcement Surprise

This table reports regression results of the regression in Equation (5)

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

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 control variables and ϵ_t is the residual. Column (I) includes only a constant and the dummy for EIA days, and column (II) adds the surprise variable. Columns (III) to (VI) add several sets of control variables such as dummy variables interacted with the surprise (III), macroeconomic measures for supply and demand of natural gas (IV), return spillovers from other markets (V), and commodity return predictors (VI). The last column combines all set of control variables. Returns are in percentage points and p-values in parentheses are based on Newey and West (1987) standard errors with two lags.

Variables	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Constant	-0.09	-0.09	-0.09	-0.09	-0.05	0.46	0.10
	(0.04)	(0.04)	(0.04)	(0.07)	(0.20)	(0.00)	(0.50)
I_{EIA}	-0.28	-0.24	-0.23	-0.25	-0.23	-0.27	-0.27
	(0.03)	(0.05)	(0.05)	(0.05)	(0.02)	(0.03)	(0.01)
S		-1.04	-0.41	-1.03	-0.78	-1.04	-0.31
		(0.00)	(0.13)	(0.00)	(0.00)	(0.00)	(0.09)
Control Dummies	No	No	Yes	No	No	No	Yes
Control Macro	No	No	No	Yes	No	No	Yes
Control Spillover	No	No	No	No	Yes	No	Yes
Control Predictors	No	No	No	No	No	Yes	Yes
R^2	0.00	0.03	0.03	0.03	0.29	0.03	0.30
Obs	3982	3982	3982	3649	3980	3844	3529

Table 5: Intraday Return Decomposition

This table reports the average returns on Henry Hub natural gas futures on EIA announcement days. Panel A decomposes the daily return from market closure on the previous day to market closure on the announcement day, (C_{t-1}, C_t), into an intraday component from 90 minutes before the announcement to 30 minutes after the announcement, $(-90, 30)$, and the sum of the return from the close price of the previous day to 90 minutes before the announcement, $(C_{t-1}, -90)$, and the return from 30 minutes after the announcement to the close price of the announcement day, $(30, C_t)$. Panel B decomposes the intraday return, $(-90, 30)$, into the pre-announcement return, $(-90, -5)$, and the announcement return, $(-5, 30)$.

Panel A: Daily and Intraday Return

	(C_{t-1}, C_t)	$(-90, 30)$	$(C_{t-1}, -90)\&(30, C_t)$
Average Return	-0.37	-0.37	-0.00
p-value	(0.002)	(0.000)	(0.963)
% of Daily Return	100%	99.0%	1.0%

Panel B: Pre- and Post-Announcement Return

	$(-90, 30)$	$(-90, -5)$	$(-5, 30)$
Average Return	-0.37	-0.18	-0.19
p-value	(0.000)	(0.000)	(0.010)
% of (-90,30) Return	100%	49.4%	50.6%

Table 6: Intraday Return Regressions

This table reports regression results of the regression in Equation (5) using intraday returns

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

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 ϵ_t is the residual. The dependent variable changes in every column, starting with the daily return, (C_{t-1}, C_t), the return from 90 minutes before the announcement to 30 minutes after the announcement, (-90,30), the pre-announcement return from 90 before until 5 minutes before the announcement, (-90,-5), and the post-announcement return from 5 minutes before the announcement until 30 minutes after the announcement, (-5,30), respectively. Results for the basis and idiosyncratic volatility are reported, the control variables are not reported, they include spillovers from the previous day, oil markets, commodity markets and stock markets. Returns are in percentage points and p-values in parentheses are based on Newey and West (1987) standard errors with two lags.

Dependent Variable	Daily Return (C_{t-1}, C_t)	Intraday Return (-90,30)	Pre-Announce (-90,-5)	Post-Announce (-5,30)
Constant	0.34	0.14	0.05	0.09
	(0.00)	(0.04)	(0.28)	(0.07)
I_{EIA}	-0.24	-0.30	-0.15	-0.15
	(0.01)	(0.00)	(0.00)	(0.00)
Surprise	-0.77	-1.01	-0.14	-0.87
	(0.00)	(0.00)	(0.00)	(0.00)
Basis	0.42	-0.03	-0.00	-0.03
	(0.00)	(0.51)	(0.96)	(0.38)
IVOL	-0.12	-0.06	-0.03	-0.03
	(0.00)	(0.01)	(0.09)	(0.04)
Control	Yes	Yes	Yes	Yes
R^2	0.29	0.15	0.03	0.16
Obs	3979	3979	3979	3979

Table 7: Summary Statistics Pre-Announcement Return

This table reports summary statistics on the pre-announcement returns of natural gas futures on EIA announcement days. Panel A reports the daily mean returns on Non-EIA days, days with a positive surprise and days with a negative surprise. The pre-announcement returns are measured from 90 minutes before the announcement until 5 minutes before the announcement. Panel B reports the mean returns for different subsamples from March 2003 until November 2007, from December 2007 until June 2014, as well as from July 2014 until December 2018. The p-value of a two-sample t-test on equal means is reported in parentheses. Rows 'No. of Obs.' denote the number of observations.

Panel A: Positive and Negative Surprise

	Non-EIA	Positive Surprise	Negative Surprise
Average Return	-0.03	-0.36 (0.000)	0.02 (0.300)
No. of Obs.	3419	345	319

Panel B: Subsample Analysis

Subsample	Non-EIA	Positive Surprise	Negative Surprise
2003–2007	-0.08	-0.38 (0.000)	-0.08 (0.981)
No. of Obs.	918	102	80
2007–2014	-0.00	-0.47 (0.000)	0.04 (0.609)
No. of Obs.	1519	153	142
2014–2018	-0.02	-0.15 (0.070)	0.08 (0.147)
No. of Obs.	981	90	97

Table 8: Returns on Investment Strategy

This table reports the average annualized return and the Sharpe ratio on an investment strategy in Henry Hub natural gas futures. The strategy opens a short position 90 minutes before the EIA storage report announcement, usually Thursdays at 10:30 AM ET, and closes the position 30 minutes after the announcement. Panel A reports the statistics for the whole sample, Panel B for the period before 2011, and Panel C for the period after 2011. The first row of each panel reports the raw return based on mid prices. The second row takes into account transaction cost (TC) by using the bid and ask prices for buying and selling. The third row also subtracts funding costs (FC), assuming a fully funded futures position funded at the Overnight London Interbank Offered Rate (ONR). The p-value for a t-test on difference to zero is reported in parentheses. Returns are reported in percentage points and Sharpe ratios are annualized using 252 days.

Panel A: Whole Sample

	Average Return	Sharpe Ratio
Raw (-90,30)	17.86 (0.000)	2.57
Raw + TC	12.21 (0.003)	1.79
Raw + TC + FC	12.01 (0.003)	1.76

Panel B: Before 2011

	Average Return	Sharpe Ratio
Raw (-90,30)	32.85 (0.000)	4.22
Raw + TC	25.36 (0.000)	3.30
Raw + TC + FC	25.02 (0.000)	3.26

Panel C: After 2011

	Average Return	Sharpe Ratio
Raw (-90,30)	2.91 (0.558)	0.50
Raw + TC	-0.90 (0.854)	-0.16
Raw + TC + FC	-0.97 (0.843)	-0.17

Online Appendix

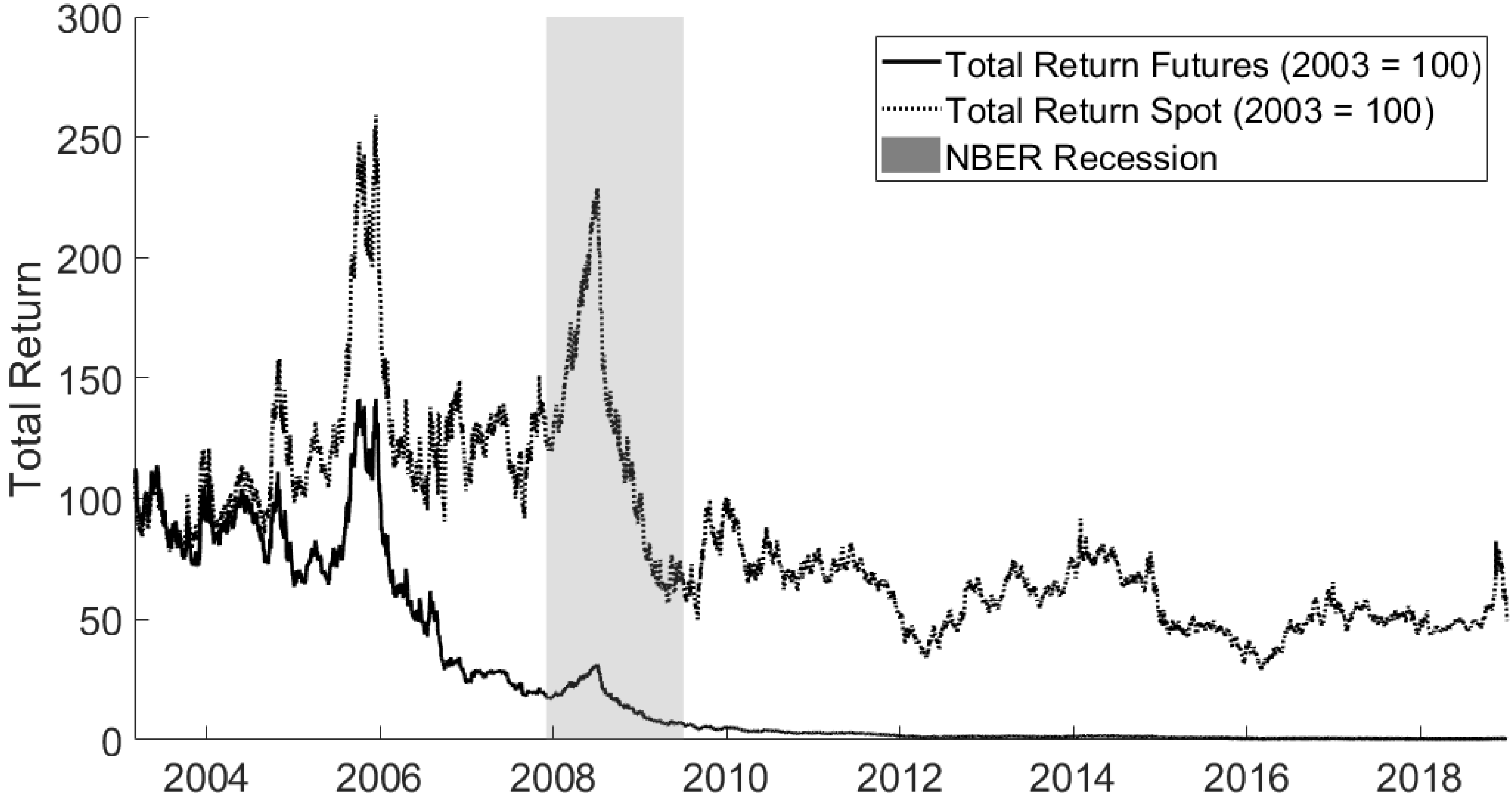

Figure A.1: Natural Gas Total Return Series

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., $R_t^n = P_t^n - P_{t-1}^n$*, which refers to two different contracts on rolling days. The solid line accounts for the rolling over in the returns as explained in Equation* (1)*. 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).*

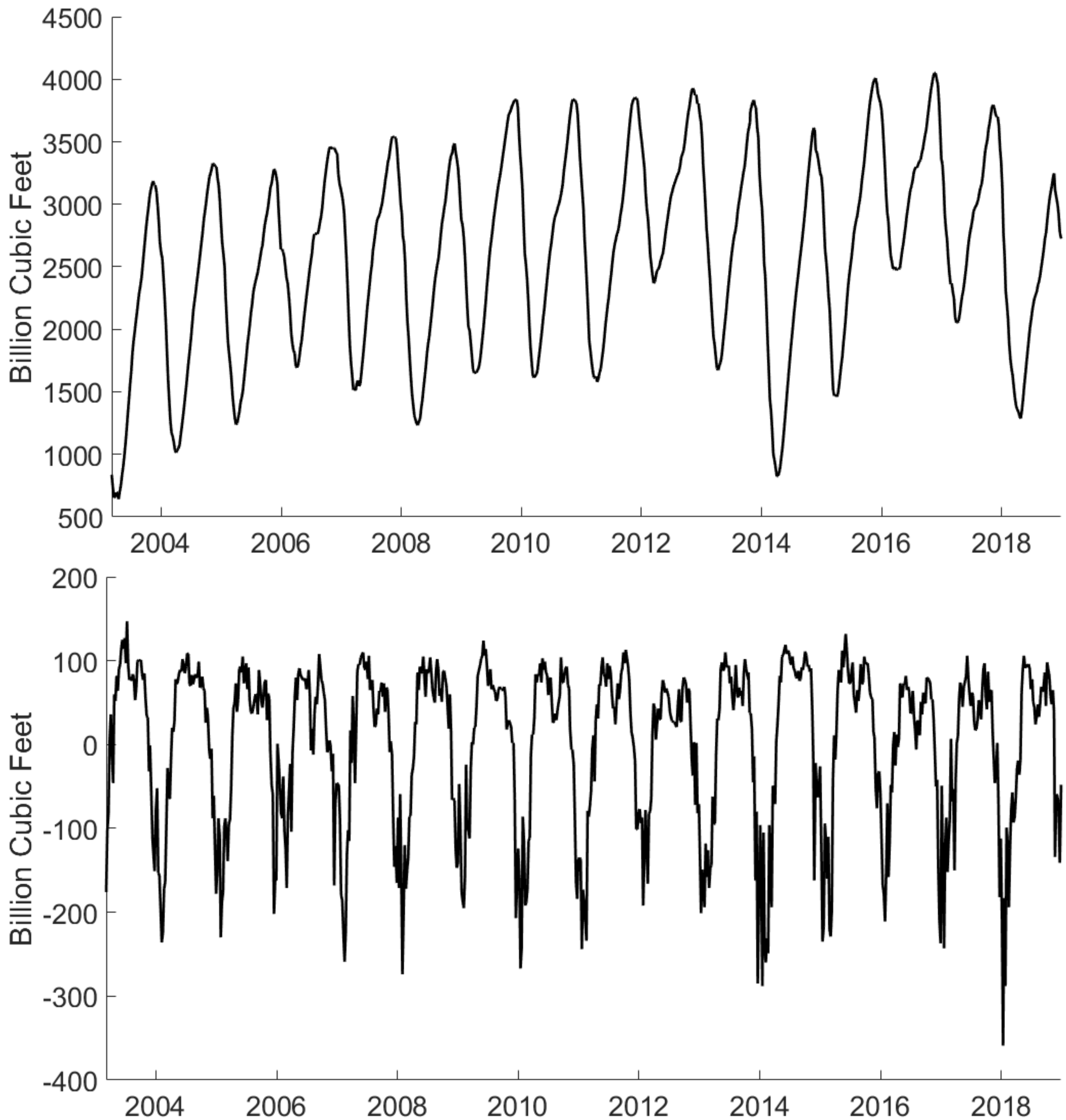

Figure A.2: Inventory Level and First Difference

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.

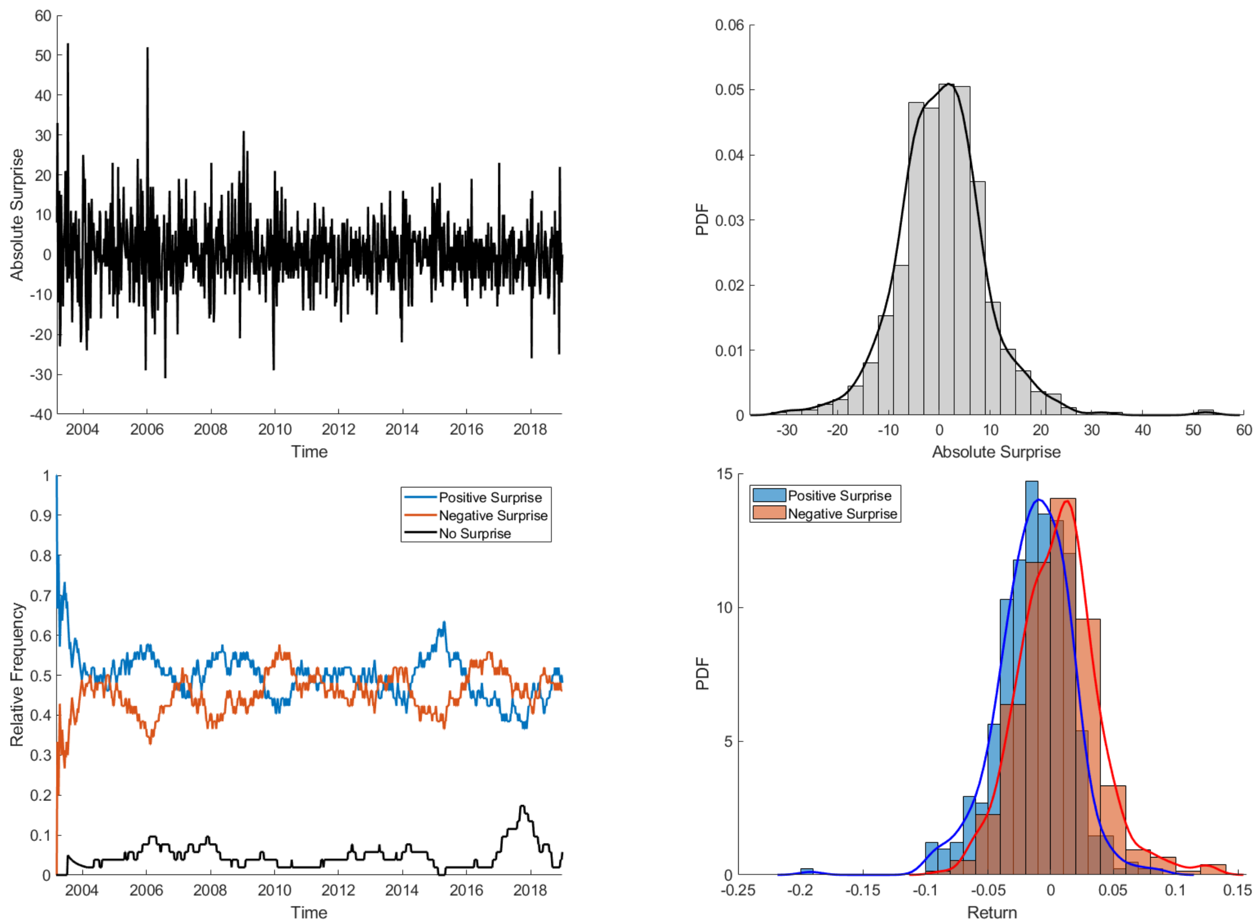

Figure A.3: Histogram and Densities for Inventory Surprise

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 ditribution of returns for positive (blue) and negative (red) surprises.

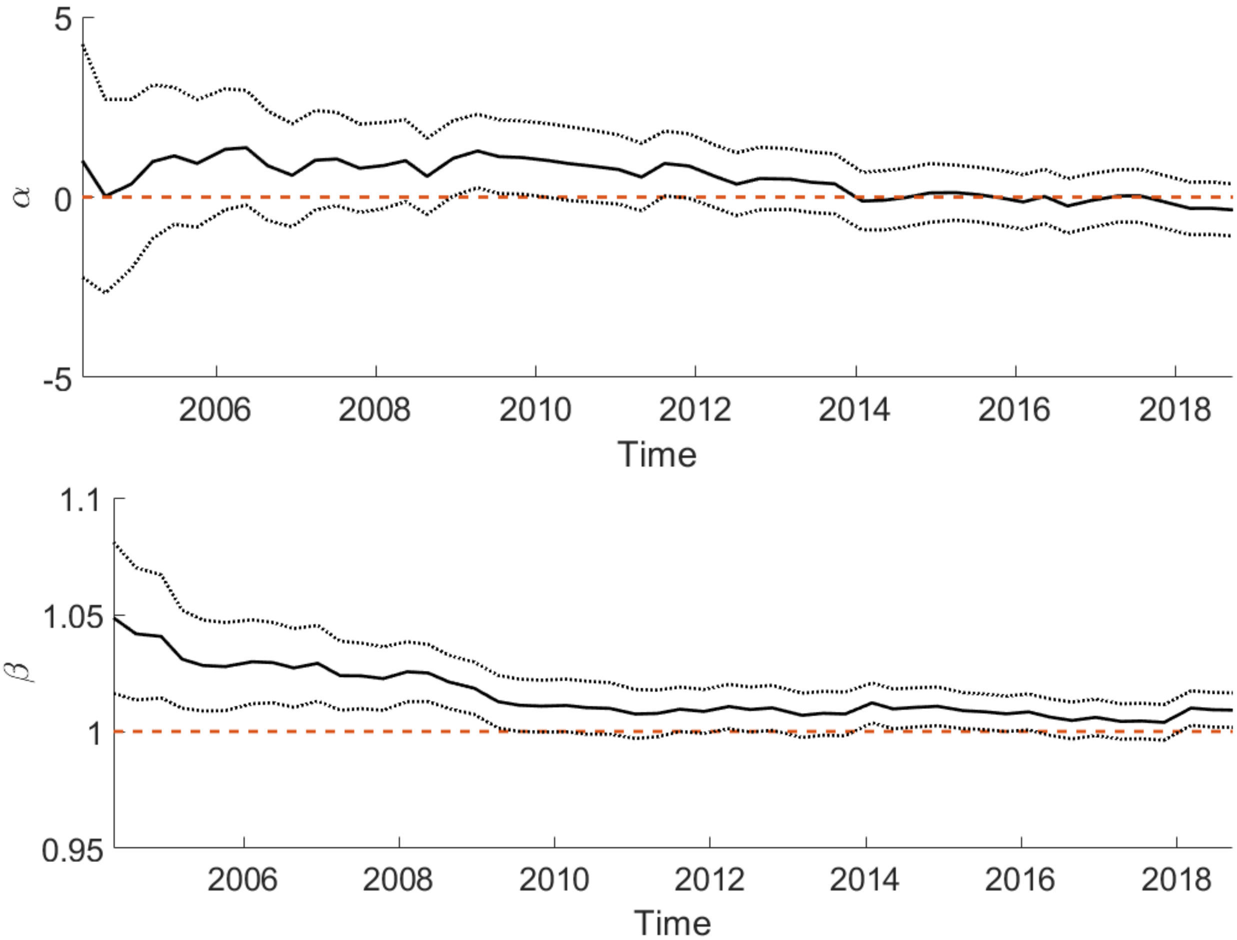

Figure A.4: Regression Results on Forecast Accuracy

This figure presents the regression results of the regression in Equation (11)

$$A_t = \alpha + \beta E_t + u_t, \tag{11}$$

where A_t is the actual storage reported by the EIA, α is the intercept, β 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 α together with the bounds on a 5% confidence interval as a dotted line. The second panel shows the coefficient estimates for β 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 wtih red dashed lines.

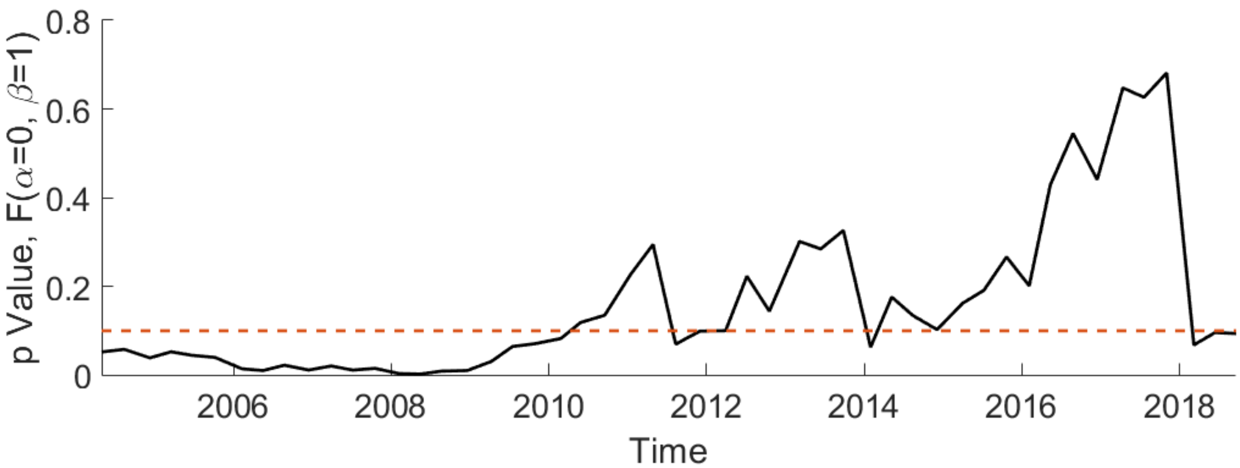

Figure A.5: Regression Results on Forecast Accuracy

This figure presents the results of a hypothesis test on the regression in Equation (11)

$$A_t = \alpha + \beta E_t + u_t, \tag{11}$$

where A_t is the actual storage reported by the EIA, α is the intercept, β 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 in Equation (5)

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

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 ϵ_t is the residual. Hence α_0 is the average return on non-EIA days and α_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 from Equation (4)

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

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 and West (1987) standard errors with two lags.

Variables	(I)	(II)	(III)	(IV)	(V)
Constant	-0.09	-0.09	-0.09	-0.09	-0.09
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
I_{EIA}	-0.24	-0.28	-0.24	-0.23	-0.21
	(0.05)	(0.03)	(0.04)	(0.06)	(0.08)
S_t	-1.04		-1.04		
	(0.00)		(0.00)		
S_{t-1}		0.05	0.02		
		(0.67)	(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 in Equation (5)

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

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 ϵ_t is the residual. Hence α_0 is the average return on non-EIA days and α_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 and West (1987) standard errors with two lags.

Variables	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Constant	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
I_{EIA}	-0.28	-0.24	-0.24	-0.23	-0.23	-0.25	-0.23	-0.23
	(0.03)	(0.05)	(0.05)	(0.05)	(0.06)	(0.04)	(0.05)	(0.05)
S		-1.04	-1.02	-0.97	-0.79	-0.88	-0.92	-0.41
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.13)
$S \times I_{lowSD}$			-1.06					-0.65
			(0.12)					(0.33)
$S \times I_{\mathrm{NBER}}$				-0.57				-0.96
				(0.12)				(0.02)
$S \times I_{inject}$					-0.54			-0.83
					(0.03)			(0.01)
$S \times I_{post09}$						-0.43		-0.64
						(0.07)		(0.02)
$S \times I_{hurricane}$							-0.28	0.27
							(0.26)	(0.36)
R^2	0.00	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 in Equation (5) *using macro variables related to the economics of natural gas markets*

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

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 ϵ_t is the residual. Hence α_0 is the average return on non-EIA days and α_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, ΔHDD (ΔCDD), the change in monthly US natural production, ΔProduction, the change in term spread (ΔTERM), which is the difference between the 3-month and 10-year U.S. Treasury Bill rate, and the change in the CBOE Volatility Index, ΔVIX, respectively. All variables are scaled to have unit standard deviation. Returns are in percentage points and p-values in parentheses are based on Newey and West (1987) standard errors with two lags.

Variables	(I)	(II)	(III)	(IV)	(V)	(VI)
Constant	-0.10	-0.10	-0.08	-0.09	-0.09	-0.08
	(0.04)	(0.04)	(0.08)	(0.05)	(0.04)	(0.08)
I_{EIA}	-0.23	-0.24	-0.25	-0.25	-0.24	-0.25
	(0.07)	(0.06)	(0.04)	(0.04)	(0.05)	(0.05)
S	-1.03	-1.02	-1.05	-1.04	-1.04	-1.03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ΔHDD	0.11					0.11
	(0.02)					(0.03)
Δ_{t+1}HDD	-0.02					-0.01
	(0.72)					(0.86)
ΔCDD		0.09				0.09
		(0.03)				(0.03)
Δ_{t+1}CDD		-0.08				-0.08
		(0.06)				(0.06)
ΔProduction			-0.14			-0.07
			(0.86)			(0.93)
ΔTERM				-0.07		-0.06
				(0.23)		(0.34)
ΔVIX					-0.11	-0.09
					(0.01)	(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 in Equation (5) *using variables related to spillover effects*

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

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 ϵ_t is the residual. Hence α_0 is the average return on non-EIA days and α_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 and West (1987) standard errors with two lags.

Variables	(I)	(II)	(III)	(IV)	(V)
Constant	-0.10	-0.08	-0.07	-0.10	-0.05
	(0.03)	(0.06)	(0.10)	(0.03)	(0.20)
I_{EIA}	-0.23	-0.29	-0.30	-0.24	-0.23
	(0.05)	(0.01)	(0.01)	(0.05)	(0.02)
S	-1.05	-1.01	-0.93	-1.03	-0.78
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R_{t-1}	-0.06				-0.07
	(0.00)				(0.00)
R^{WTI}		0.35			-1.36
		(0.00)			(0.00)
R^{GSCI}			0.74		2.72
			(0.00)		(0.00)
R^{CRSP}				0.11	-0.19
				(0.00)	(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 in Equation (5) using commodity trading signals

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

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 ϵ_t is the residual. Hence α_0 is the average return on non-EIA days and α_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, ΔVolume, respectively. Returns are in percentage points and p-values in parentheses are based on Newey and West (1987) standard errors with two lags.

Ind. Var.	(I)	(II)	(III)	(IV)	(V)	(VI)
Constant	-0.02	0.01	-0.08	0.28	-0.08	0.46
	(0.64)	(0.80)	(0.18)	(0.03)	(0.07)	(0.00)
I_{EIA}	-0.23	-0.24	-0.24	-0.24	-0.27	-0.27
	(0.05)	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)
S	-1.04	-1.05	-1.04	-1.04	-1.04	-1.04
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$B_{1,2}$	0.44					0.42
	(0.00)					(0.00)
$B_{1,13}$		1.99				
		(0.00)				
HP			0.15			0.31
			(0.70)			(0.43)
IVOL				-0.14		-0.18
				(0.01)		(0.00)
ΔVolume					0.08	0.08
					(0.13)	(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 inclduing 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.

	Mean			**Standard Deviation**			
Excluded Event	News	Rest	t-Test	News	Rest	F-Test	Obs
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
Gross Domestic Product	-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 and Prices	-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 regression in Equation (5) *using intraday returns*

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

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 ϵ_t is the residual. Hence α_0 is the average return on non-EIA days and α_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 and West (1987) standard errors with two lags.

Dep. Var.	(-90,30)	(-60,30)	(-30,30)	(-5,30)
Constant	0.14	0.13	0.12	0.09
	(0.04)	(0.03)	(0.03)	(0.07)
I_{EIA}	-0.30	-0.22	-0.16	-0.15
	(0.00)	(0.00)	(0.00)	(0.00)
Surprise	-1.01	-0.97	-0.93	-0.87
	(0.00)	(0.00)	(0.00)	(0.00)
Basis	-0.03	0.01	0.00	-0.03
	(0.51)	(0.87)	(0.96)	(0.38)
IVOL	-0.06	-0.05	-0.05	-0.03
	(0.01)	(0.02)	(0.02)	(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 Spread Returns

This table reports regression results of the regression in Equation (5)

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

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 ϵ_t is the residual. Hence α_0 is the average return on non-EIA days and α_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 and West (1987) standard errors with two lags.

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

Table A.13: Regression Bloomberg Survey Forecast Accuracy

This table reports regression results for the forecast accuracy of the Bloomberg median survey forecast as described in Equation (11)

$$A_t = \alpha + \beta E_t + u_t, \tag{11}$$

where A_t denotes the weekly storage reported by the EIA, α is the intercept, β 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 (α)	0.2385	0.3243
	(0.460)	(0.313)
Forecast (E_t)	1.0119	1.0132
	(0.000)	(0.000)
$F(\alpha = 0, \beta = 1)$	6.32	7.72
	(0.002)	(0.000)
R^2	0.99	0.99

A Factor Construction

This section contains instruction on how to construct the factor returns for the computation of the idiosyncratic volatility measure, IVOL, in Equation (10).

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

$$R_{\mathrm{MRKT},t} = \frac{1}{27} \sum_{i=1}^{27} R_{i,t}, \tag{12}$$

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:

$$\begin{aligned}
B_{i,t} &= \left(\frac{F_{i,t}^1}{F_{i,t}^2} \right)^{\frac{365}{M_{i,t}^2 - M_{i,t}^1}} - 1, \\
\mathrm{MOM}_{i,t} &= \sum_{j=1}^{252} R_{i,t-j}^1 \\
\mathrm{BMOM}_{i,t} &= \sum_{j=1}^{252} R_{i,t-j}^1 - \sum_{j=1}^{252} R_{i,t-j}^2
\end{aligned}$$

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.