

# DO ENERGY PRICES RESPOND TO U.S. MACROECONOMIC NEWS? A TEST OF THE HYPOTHESIS OF PREDETERMINED ENERGY PRICES

Lutz Kilian and Clara Vega\*

*Abstract*—We propose a formal test of the hypothesis that energy prices are predetermined with respect to U.S. macroeconomic aggregates. The test is based on regressing changes in daily energy prices on daily news from U.S. macroeconomic data releases. Using a wide range of macroeconomic news, we find no compelling evidence of feedback at daily or monthly horizons, contradicting the view that energy prices respond instantaneously to macroeconomic news and consistent with the commonly used identifying assumption that there is no feedback from U.S. macroeconomic aggregates to monthly innovations in energy prices.

## I. Introduction

A standard identifying assumption in vector autoregressive (VAR) models of the transmission of energy price shocks is that energy prices are predetermined with respect to U.S. macroeconomic aggregates such as real output, consumption, investment, or interest rates.<sup>1</sup> In a monthly model, for example, this assumption rules out feedback from domestic macroeconomic shocks to the price of energy within the same month. This assumption is not testable within the VAR framework.<sup>2</sup>

Although the identifying assumption of predetermined energy prices is widely accepted in empirical work, its rationale is less than obvious. Given that crude oil and gasoline, in particular, are storable and relatively homogeneous, they may be viewed alternatively as an asset, the price of which is determined by the supply of and demand for stocks, or as a good, the price of which is determined by flow supply and flow demand (see Frankel & Rose, 2010; Kilian, 2009; Alquist & Kilian, 2010). Hence, an obvious concern is that oil and gasoline prices in practice may behave like asset prices and jump in response to any news about future supply and demand, including domestic macroeconomic news. To the extent that they do, commonly used empirical models based on the assumption of predetermined energy prices would be invalid.

Received for publication October 13, 2008. Revision accepted for publication December 9, 2009.

\* Kilian: University of Michigan and CEPR; Vega: Federal Reserve Board.

We thank Christian S. Miller for providing research assistance. The views expressed in this paper do not necessarily reflect the views of the Federal Reserve Board or its members. We have benefited from comments by Roger Gordon, Eric Sims, Daniel Beltran, the editor, and two anonymous referees.

<sup>1</sup> See Bernanke, Gertler, and Watson (1997, 2004); Blanchard and Galí (2010); Davis and Haltiwanger (2010); Davis and Kilian (forthcoming); Edelstein and Kilian (2007, 2009); Hamilton and Herrera (2004); Herrera (2009); Herrera and Pesavento (2009); Leduc and Sill (2004); Lee and Ni (2002); Kilian and Park (2009); Rotemberg and Woodford (1996). The same assumption has been used implicitly in VAR models that disentangle demand and supply shocks in energy markets (Kilian, 2009, 2010; Kilian & Park, 2009).

<sup>2</sup> Kilian (2009) provides some empirical evidence in support of this assumption, but that evidence does not cover all possible forms of instantaneous feedback and hence is suggestive only.

In this paper, we propose a formal test of the view that oil prices respond without delay to macroeconomic news. Our approach is based on a methodology pioneered by Andersen et al. (2003, 2007) in the related context of studying price discovery in asset markets (also see Faust et al., 2007). We utilize daily data on crude oil and gasoline prices for 1983 to 2008 in conjunction with daily data on the news component of thirty U.S. macroeconomic data releases. Our first result is that unlike stock prices, bond prices, or exchange rates, the price of West Texas Intermediate (WTI) crude oil does not respond significantly to any one of the U.S. macroeconomic news within the day, contradicting the view that oil prices should be thought of as asset prices. Our second result is that there is no compelling evidence of feedback within the month from U.S. macroeconomic news to the price of crude oil. Specifically, there is no significant evidence of feedback from any one news item. Only joint tests for a set of forward-looking news variables reveal any evidence of statistically significant feedback within the month, but the extent of that feedback appears small enough to be ignored. Ninety-nine percent of the monthly variation in oil prices remains unexplained by all thirty macroeconomic news combined, making the assumption of predetermined oil prices a reasonable approximation in practice. Broadly similar results hold for gasoline prices.

The remainder of the paper is organized as follows. In section II, we describe the data and econometric methodology. Section III contains a detailed analysis of the impact response of energy prices to U.S. macroeconomic news. In section IV, we extend the analysis to monthly horizons and directly test the assumption that energy prices are predetermined with respect to U.S. macroeconomic aggregates at monthly frequency. Section V contains additional results based on joint tests for subsets of news related to the same economic concept. In section VI, we discuss the power of news-based tests of no feedback. The concluding remarks are in section VII.

## II. Methodology

### A. Macroeconomic News

Macroeconomic news is defined as the difference between ex ante survey expectations and the subsequently announced realizations of macroeconomic aggregates. Real-time data on expected and realized U.S. macroeconomic fundamentals are available from Money Market Services (MMS). Our longest sample period extends from January 1983 through April 2008, but not all of the announcements are available from MMS from the beginning

of this sample period.<sup>3</sup> Table 1 provides a description of the announcement releases, including the number of observations, the agency reporting the news, and the time of the release. Our data set includes quarterly announcements for GDP; monthly announcements for various measures of real activity, consumption, investment, fiscal and trade balances, prices, the Fed target rate, and forward-looking indicators; and weekly announcements of initial unemployment claims. The units of measurement obviously differ across the macroeconomic indicators, as is apparent from the last column of table 1, which shows the standard deviations. To allow meaningful comparisons of the estimated news response coefficients across indicators and asset classes, we follow Andersen et al. (2003) in that we use standardized news measures. Specifically, we divide the surprise component of the announcement by its sample standard deviation, defining the standardized news associated with indicator  $i$  at time  $t$  as

$$S_{it} = \frac{A_{it} - E_{it}}{\hat{\sigma}_i},$$

where  $A_{it}$  denotes the announced value of indicator  $i$ ,  $E_{it}$  refers to the market's expectation of indicator  $i$  prior to the announcement (represented by the MMS median forecast), and  $\hat{\sigma}_i$  is the sample standard deviation of the surprise component,  $A_{it} - E_{it}$ . Because  $\hat{\sigma}_i$  is constant for each indicator  $i$ , this standardization affects neither the statistical significance of the estimated response coefficients nor the fit of the regressions compared to the results based on the raw surprises.

### B. Energy Prices

Our oil price series is the daily spot price for WTI crude oil for delivery (freight on board) in Cushing, Oklahoma, expressed in dollars per barrel. The data source is the U.S. Energy Information Administration (see <http://tonto.eia.doe.gov/dnav/pet/hist/rwtcd.htm>). This price is identical to the front-month oil futures price on the New York Mercantile Exchange with the exception of dates on which the front-month contract expires. The daily gasoline price series is based on credit card transactions obtained by the Oil Price Information Service (OPIS) from gas stations in the United States. (For more detailed information see <http://www.opisretail.com/methodology.html>.) Both price series are expressed in cumulative percentage changes.

### C. Estimating the Effect of U.S. Macroeconomic News on Energy Prices

We model energy prices as daily percentage changes, permitting news to have a permanent effect on the level of nominal energy prices. The baseline model in section III focuses on the impact effect of news. We fit the model

$$R_{t+1} = \alpha + \beta_i S_{it} + \varepsilon_{t+1}, \quad (1)$$

where  $R_{t+1} = 100 \times \ln(P_{t+1}/P_t)$  denotes the daily return on holding regular gas or WTI crude oil from the end of day  $t-1$  to the end of day  $t$ , and  $S_{it}$  refers to the standardized news for announcement  $i$ ,  $i = 1, \dots, 30$ , on day  $t$ . The regression estimates are based only on data for those days on which a news announcement was made. Inference is based on White standard errors to allow for the possibility of time-varying variances. We do not control for serial correlation because the daily price changes are not consecutive, rendering the residuals serially uncorrelated under the null hypothesis. Moreover, there is no evidence of serial correlation in the unrestricted model residuals. The parameter  $\beta_i$  measures the response of  $R_{t+1}$  to a 1-standard deviation news shock. An estimate of  $\hat{\beta}_5 = 0.027$ , for example, would imply that an unexpected increase of nonfarm payroll employment by 111,153 jobs would cause an increase in the price of oil by 0.027%.

In addition to the regressions involving one news shock predictor at a time, we also consider the joint regression

$$R_{t+1} = \alpha + \sum_{i=1}^{30} \beta_i S_{it} + \varepsilon_{t+1}, \quad (2)$$

for all date  $t$  observations. In that case, inference is based on Newey-West standard errors to allow for the possibility of both serial correlation and heteroskedasticity under the null hypothesis. Allowing for serial correlation is advisable here because, unlike in model (1), the dependent variable is likely to be serially correlated under the null hypothesis when all date  $t$  observations are included.

Focusing on daily asset price changes around the time of the announcement, and estimating the immediate news reaction of asset prices helps isolate the effect of the news announcement among the effect of a myriad of other changes in the economy. This strategy has already been applied successfully to numerous financial assets in the literature. If traders are slow to appreciate the significance of news shocks, however, the reaction of oil prices to news shocks may be delayed; hence, the focus on daily data may cause us to miss the impact of news on oil prices. In section IV, we allow for a delayed reaction of oil prices to news by regressing cumulative daily returns on crude oil for a horizon of up to one month on current macroeconomic news (the monthly returns are calculated from the end of day  $t-1$  to the end of day  $t+h-1$ , where  $h$  is equal to 20 business days in the monthly regression, and the announcement occurs on day  $t$ ).

Our sample period is dictated by data availability constraints. The full-sample regression results for WTI crude oil prices rely on data from May 1983 to April 2008 (see table 2), and the full sample regression results for regular gasoline prices are based on data from January 2003 to April 2008 (see table 3). We report the coefficient estimates,  $t$ -statistics, and  $p$ -values calculated using robust standard errors. We also

<sup>3</sup> From 2003 onward, we use survey data provided by Bloomberg rather than MMS, since MMS ceased to exist. Both the MMS and Bloomberg surveys have been shown to be unbiased expectations measures.

TABLE 1.—U.S. NEWS ANNOUNCEMENTS

Announcement	Observations <sup>a</sup>	Source <sup>b</sup>	Dates <sup>c</sup>	Release Time <sup>d</sup>	s.d. <sup>e</sup>
<b>Quarterly Announcements</b>					
1. GDP advance	83	BEA	4/1987–4/2008 <sup>f</sup>	8:30	0.771
2. GDP preliminary	82	BEA	4/1987–4/2008 <sup>g</sup>	8:30	0.418
3. GDP final	83	BEA	4/1987–4/2008 <sup>h</sup>	8:30	0.310
<b>Monthly Announcements</b>					
<b>Real activity</b>					
4. Unemployment rate	304	BLS	1/1983–4/2008 <sup>i</sup>	8:30	0.156
5. Nonfarm payroll employment	279	BLS	2/1985–4/2008 <sup>j</sup>	8:30	111.153
6. Retail sales	258	BC	12/1986–4/2008	8:30	0.604
7. Industrial production	257	FRB	12/1986–4/2008	9:15	0.273
8. Capacity utilization	240	FRB	4/1988–4/2008 <sup>k</sup>	9:15	0.320
9. Personal income	253	BEA	12/1986–4/2008 <sup>l</sup>	10:00/8:30 <sup>m</sup>	0.252
10. Consumer credit	241	FRB	4/1988–4/2008 <sup>n</sup>	15:00 <sup>o</sup>	4.243
<b>Consumption</b>					
11. New home sales	239	BEA	3/1988–4/2008 <sup>p</sup>	10:00/8:30	62.946
12. Personal consumption expenditures	256	BC	12/1986–4/2008 <sup>q</sup>	10:00 <sup>r</sup>	0.208
<b>Investment</b>					
13. Durable goods orders	299	BC	4/1983–4/2008 <sup>s</sup>	8:30/9:00/10:00 <sup>t</sup>	2.906
14. Construction spending	240	BC	4/1988–4/2008 <sup>u</sup>	10:00	1.007
15. Factory orders	240	BC	3/1988–4/2008 <sup>v</sup>	10:00	0.714
16. Business inventories	240	BC	4/1988–4/2008 <sup>w</sup>	10:00/8:30 <sup>x</sup>	0.273
<b>Fiscal balance</b>					
17. Net government purchases	236	FMS	4/1988–4/2008 <sup>y</sup>	14:00	8.646
<b>Net exports</b>					
18. Trade balance	256	BEA	12/1986–4/2008 <sup>z</sup>	8:30	2.337
<b>Prices</b>					
19. Producer Price Index	257	BLS	12/1986–4/2008	8:30	0.399
20. Core PPI	195	BLS	1/1992–4/2008 <sup>aa</sup>	8:30	0.266
21. Consumer Price Index	304	BLS	1/1983–4/2008	8:30	0.127
22. Core CPI	195	BLS	1/1992–4/2008 <sup>bb</sup>	8:30	0.214
<b>Forward looking</b>					
23. Michigan CCI preliminary	110	UM	5/1999–7/2008 <sup>cc</sup>	10:00	10.677
24. Michigan CCI final	110	UM	5/1999–6/2008	10:00	10.843
25. Board CCI index	200	CB	7/1991–4/2008	8:30	4.960
26. NAPM index	220	NAPM	2/1990–4/2008	10:00	2.008
27. Housing starts	303	BC	1/1983–4/2008 <sup>dd</sup>	8:30	0.135
28. Index of leading indicators	304	CB	1/1983–4/2008	8:30	0.243
<b>Six-Week Announcements</b>					
<b>FOMC</b>					
29. Target federal funds rate	185	FRB	1/1983–4/2008	14:15 <sup>ee</sup>	0.089
<b>Weekly Announcements</b>					
30. Initial unemployment claims	870	ETA	7/1991–4/2008	8:30	12.996

We partition the U.S. monthly news announcements into seven groups: aggregate real activity, the components of real GDP (consumption, investment, fiscal balance, and net exports), prices, and forward-looking variables. Within each group, we list U.S. news announcements in chronological order of their release. CCI denotes the Consumer Confidence Index.

<sup>a</sup>Total number of observations in our announcements and expectations data sample.

<sup>b</sup>Bureau of Labor Statistics (BLS), Bureau of the Census (BC), Bureau of Economic Analysis (BEA), Federal Reserve Board (FRB), National Association of Purchasing Managers (NAPM), Conference Board (CB), Financial Management Office (FMO), Employment and Training Administration (ETA), University of Michigan (UM).

<sup>c</sup>Starting and ending dates of our announcements and expectations data sample.

<sup>d</sup>Eastern Standard Time. Daylight savings time starts on the first Sunday of April and ends on the last Sunday of October.

<sup>e</sup>Standard deviation of the macroeconomic news surprise before we standardize it.

<sup>f</sup>7/87 and 1/88 are missing observations.

<sup>g</sup>11/87 and 11/95 are missing observations.

<sup>h</sup>12/87 is a missing observation.

<sup>i</sup>7/93 is a missing observation.

<sup>j</sup>4/85 and 10/98 are missing observations.

<sup>k</sup>11/03 is a missing observation.

<sup>l</sup>11/95, 2/96, 3/97, and 12/07 are missing observations.

<sup>m</sup>In 01/94, the personal income announcement time moved from 10:00 a.m. EST to 8:30 a.m. EST.

<sup>n</sup>11/03 is a missing observation.

<sup>o</sup>Beginning in 01/96, consumer credit was released regularly at 3:00 p.m. EST. Prior to this date the release times varied.

<sup>p</sup>4/88, 1/89, and 12/95 are missing observations.

<sup>q</sup>11/95 and 2/96 are missing observations.

<sup>r</sup>In 12/93, the personal consumption expenditures announcement time moved from 10:00 a.m. EST to 8:30 a.m. EST.

<sup>s</sup>12/95 and 1/96 are missing observations.

<sup>t</sup>Whenever GDP is released on the same day as durable goods orders, the durable goods orders announcement is moved to 10:00 a.m. EST. On 07/96 the durable goods orders announcement was released at 9:00 a.m. EST.

<sup>u</sup>1/96, 10/98, 12/03, and 12/07 are missing observations.

<sup>v</sup>1/96 and 11/03 are missing observations.

<sup>w</sup>11/03 is a missing observation.

<sup>x</sup>In 01/97, the business inventory announcement was moved from 10:00 a.m. EST to 8:30 a.m. EST.

<sup>y</sup>5/88, 6/88, 11/89, 12/89, 1/90, and 1/96 are missing observations.

<sup>z</sup>3/87 is a missing observation.

<sup>aa</sup>11/92 is a missing observation.

<sup>bb</sup>11/92 and 12/98 are missing observations.

<sup>cc</sup>7/99 is a missing observation.

<sup>dd</sup>12/95 is a missing observation.

<sup>ee</sup>Beginning in 3/28/94, the fed funds rate was released regularly at 2:15 p.m. EST. Prior to this date the release times varied.

TABLE 2.—DAILY WTI CRUDE OIL PRICES

A. Individual Regressions, 1983–2008 <sup>a</sup>							
Announcement	$\hat{\beta}_i$	$\hat{t}_i$	Standard <i>p</i> -Value	Robust <i>p</i> -Value	<i>R</i> <sup>2</sup> Percent	Observations	Alternative Hypothesis
GDP advanced	-0.224	-1.08	0.86	1.00	0.89	83	$H_1 : \beta_i > 0$
GDP preliminary	0.348	1.54	<b>0.06</b>	0.86	3.22	82	$H_1 : \beta_i > 0$
GDP final	0.166	0.50	0.31	1.00	0.46	82	$H_1 : \beta_i > 0$
Unemployment rate	-0.087	-0.63	0.27	1.00	0.19	288	$H_1 : \beta_i < 0$
Nonfarm payroll	0.027	0.19	0.42	1.00	0.02	268	$H_1 : \beta_i > 0$
Retail sales	-0.276	-1.06	0.86	1.00	1.86	257	$H_1 : \beta_i > 0$
Industrial production	0.011	0.08	0.47	1.00	0.00	255	$H_1 : \beta_i > 0$
Capacity utilization	0.056	0.38	0.35	1.00	0.07	238	$H_1 : \beta_i > 0$
Personal income	-0.120	-0.82	0.79	1.00	0.23	247	$H_1 : \beta_i > 0$
Consumer credit	0.057	0.49	0.31	1.00	0.10	238	$H_1 : \beta_i > 0$
New home sales	0.202	1.36	<b>0.09</b>	0.94	0.93	237	$H_1 : \beta_i > 0$
Personal consumption	-0.116	-0.51	0.70	1.00	0.23	249	$H_1 : \beta_i > 0$
Durable goods orders	-0.102	-0.75	0.77	1.00	0.20	298	$H_1 : \beta_i > 0$
Construction spending	0.005	0.04	0.49	1.00	0.00	237	$H_1 : \beta_i > 0$
Factory orders	-0.008	-0.04	0.52	1.00	0.00	239	$H_1 : \beta_i > 0$
Business inventories	-0.035	-0.20	0.42	1.00	0.03	238	$H_1 : \beta_i < 0$
Government budget deficit	0.329	2.40	<b>0.01</b>	0.23	1.37	232	$H_1 : \beta_i > 0$
Trade balance	-0.026	-0.19	0.57	1.00	0.01	255	$H_1 : \beta_i > 0$
PPI	-0.205	-1.66	0.95	1.00	1.04	257	$H_1 : \beta_i > 0$
Core PPI	-0.089	-0.62	0.73	1.00	0.19	195	$H_1 : \beta_i > 0$
CPI	-0.045	-0.30	0.62	1.00	0.03	299	$H_1 : \beta_i > 0$
Core CPI	0.190	2.39	<b>0.01</b>	0.24	0.87	194	$H_1 : \beta_i > 0$
CCI preliminary (Michigan)	0.246	1.07	0.14	0.99	1.34	107	$H_1 : \beta_i > 0$
CCI final (Michigan)	0.061	0.41	0.34	1.00	0.12	108	$H_1 : \beta_i > 0$
CCI (board)	0.181	1.36	<b>0.09</b>	0.94	0.77	199	$H_1 : \beta_i > 0$
NAPM index	-0.017	-0.11	0.55	1.00	0.00	218	$H_1 : \beta_i > 0$
Housing starts	0.243	2.17	<b>0.02</b>	0.38	0.55	298	$H_1 : \beta_i > 0$
Index of leading indicators	-0.053	-0.27	0.61	1.00	0.03	298	$H_1 : \beta_i > 0$
Target rate surprises	0.112	0.67	0.75	1.00	0.20	185	$H_1 : \beta_i < 0$
Initial claims	0.038	0.50	0.69	1.00	0.03	869	$H_1 : \beta_i < 0$
B. Joint Regression, 1983–2008							
Announcement	$\hat{\beta}_i$	$\hat{t}_i$	Standard <i>p</i> -Value	Robust <i>p</i> -Value	Alternative Hypothesis	Observations	<i>R</i> <sup>2</sup> Percent
GDP advanced	-0.225	-1.07	0.86	1.00	$H_1 : \beta_i > 0$	6,214	0.38
GDP preliminary	0.332	1.55	<b>0.06</b>	1.00	$H_1 : \beta_i > 0$		
GDP final	0.117	0.38	0.35	1.00	$H_1 : \beta_i > 0$		
Unemployment rate	-0.099	-0.78	0.22	1.00	$H_1 : \beta_i < 0$		
Nonfarm payroll	0.048	0.33	0.37	1.00	$H_1 : \beta_i > 0$		
Retail sales	-0.266	-1.03	0.85	1.00	$H_1 : \beta_i > 0$		
Industrial production	-0.056	-0.28	0.61	1.00	$H_1 : \beta_i > 0$		
Capacity utilization	0.098	0.45	0.33	1.00	$H_1 : \beta_i > 0$		
Personal income	-0.114	-0.72	0.77	1.00	$H_1 : \beta_i > 0$		
Consumer credit	0.051	0.44	0.33	1.00	$H_1 : \beta_i > 0$		
New home sales	0.199	1.32	<b>0.09</b>	0.95	$H_1 : \beta_i > 0$		
Personal consumption	-0.103	-0.45	0.67	1.00	$H_1 : \beta_i > 0$		
Durable goods orders	-0.104	-0.76	0.78	1.00	$H_1 : \beta_i > 0$		
Construction spending	0.000	0.00	0.50	1.00	$H_1 : \beta_i > 0$		
Factory orders	-0.003	-0.02	0.51	1.00	$H_1 : \beta_i > 0$		
Business inventories	-0.004	-0.02	0.49	1.00	$H_1 : \beta_i < 0$		
Government budget deficit	0.321	2.45	<b>0.01</b>	0.19	$H_1 : \beta_i > 0$		
Trade balance	-0.009	-0.07	0.53	1.00	$H_1 : \beta_i > 0$		
PPI	-0.190	-1.46	0.93	1.00	$H_1 : \beta_i > 0$		
Core PPI	-0.004	-0.03	0.51	1.00	$H_1 : \beta_i > 0$		
CPI	-0.098	-0.65	0.74	1.00	$H_1 : \beta_i > 0$		
Core CPI	0.234	2.52	<b>0.01</b>	0.16	$H_1 : \beta_i > 0$		
CCI preliminary (Michigan)	0.162	0.68	0.25	1.00	$H_1 : \beta_i > 0$		
CCI final (Michigan)	0.088	0.55	0.29	1.00	$H_1 : \beta_i > 0$		
CCI (board)	0.172	1.28	0.10	0.96	$H_1 : \beta_i > 0$		
NAPM index	0.029	0.19	0.42	1.00	$H_1 : \beta_i > 0$		
Housing starts	0.250	2.30	<b>0.01</b>	0.28	$H_1 : \beta_i > 0$		
Index of leading indicators	0.028	0.14	0.44	1.00	$H_1 : \beta_i > 0$		
Target rate surprises	0.126	0.80	0.79	1.00	$H_1 : \beta_i < 0$		
Initial claims	0.040	0.57	0.72	1.00	$H_1 : \beta_i < 0$		

All regressions in part A include a constant, as does the regression in part B. Data mining robust *p*-values were computed based on a parametric bootstrap approach under the null hypothesis of no predictability. Standard *p*-values based on  $N(0,1)$  distribution. Boldface indicates statistical significance at the 10% level.

TABLE 3.—DAILY U.S. GASOLINE PRICES

A. Individual Regressions, 2003–2008							
Announcement	$\hat{\beta}_i$	$\hat{t}_i$	Standard $p$ -Value	Robust $p$ -Value	$R^2$ Percent	Observations	Alternative Hypothesis
GDP advanced	0.024	0.33	0.37	1.00	0.25	21	$H_1 : \beta_i > 0$
GDP preliminary	-0.039	-0.18	0.57	1.00	0.04	21	$H_1 : \beta_i > 0$
GDP final	0.073	0.51	0.31	1.00	1.38	21	$H_1 : \beta_i > 0$
Unemployment rate	-0.096	-0.74	0.23	1.00	1.05	64	$H_1 : \beta_i < 0$
Nonfarm payroll	0.002	0.02	0.49	1.00	0.00	63	$H_1 : \beta_i > 0$
Retail sales	-0.079	-1.16	0.88	1.00	2.02	64	$H_1 : \beta_i > 0$
Industrial production	0.076	1.49	<b>0.07</b>	0.89	3.53	64	$H_1 : \beta_i > 0$
Capacity utilization	-0.015	-0.23	0.59	1.00	0.10	63	$H_1 : \beta_i > 0$
Personal income	-0.135	-1.02	0.84	1.00	2.51	62	$H_1 : \beta_i > 0$
Consumer credit	-0.052	-1.00	0.84	1.00	1.40	63	$H_1 : \beta_i > 0$
New home sales	-0.041	-1.13	0.87	1.00	2.11	64	$H_1 : \beta_i > 0$
Personal consumption	0.023	0.34	0.37	1.00	0.06	63	$H_1 : \beta_i > 0$
Durable goods orders	-0.010	-0.14	0.56	1.00	0.03	64	$H_1 : \beta_i > 0$
Construction spending	-0.063	-0.54	0.71	1.00	0.27	63	$H_1 : \beta_i > 0$
Factory orders	0.105	1.35	<b>0.09</b>	0.94	2.68	63	$H_1 : \beta_i > 0$
Business inventories	0.029	0.29	0.61	1.00	0.17	63	$H_1 : \beta_i < 0$
Government budget deficit	-0.058	-0.88	0.81	1.00	1.30	63	$H_1 : \beta_i > 0$
Trade balance	0.002	0.05	0.48	1.00	0.00	64	$H_1 : \beta_i > 0$
PPI	0.018	0.66	0.25	1.00	0.38	64	$H_1 : \beta_i > 0$
Core PPI	0.053	1.51	<b>0.07</b>	0.88	2.27	64	$H_1 : \beta_i > 0$
CPI	-0.126	-2.28	0.99	1.00	5.05	64	$H_1 : \beta_i > 0$
Core CPI	-0.003	-0.18	0.57	1.00	0.01	64	$H_1 : \beta_i > 0$
CCI preliminary (Michigan)	-0.099	-1.64	0.95	1.00	3.56	64	$H_1 : \beta_i > 0$
CCI final (Michigan)	-0.031	-0.60	0.73	1.00	0.52	64	$H_1 : \beta_i > 0$
CCI (board)	-0.005	-0.06	0.53	1.00	0.01	60	$H_1 : \beta_i > 0$
NAPM index	-0.187	-1.25	0.89	1.00	5.68	63	$H_1 : \beta_i > 0$
Housing starts	0.008	0.13	0.45	1.00	0.02	64	$H_1 : \beta_i > 0$
Index of leading indicators	-0.036	-0.26	0.60	1.00	0.09	64	$H_1 : \beta_i > 0$
Target rate surprises	-0.018	-0.57	0.28	1.00	0.14	43	$H_1 : \beta_i < 0$
Initial claims	-0.022	-0.54	0.29	1.00	0.12	277	$H_1 : \beta_i < 0$
B. Joint Regression, 2003–2008							
Announcement	$\hat{\beta}_i$	$\hat{t}_i$	Standard $p$ -Value	Robust $p$ -Value	Alternative Hypothesis	Observations	$R^2$ Percent
GDP advanced	0.003	0.05	0.48	1.00	$H_1 : \beta_i > 0$	1,385	1.91
GDP preliminary	-0.042	-0.22	0.59	1.00	$H_1 : \beta_i > 0$		
GDP final	-0.007	-0.08	0.53	1.00	$H_1 : \beta_i > 0$		
Unemployment rate	-0.123	-0.95	0.17	1.00	$H_1 : \beta_i < 0$		
Nonfarm payroll	-0.023	-0.29	0.61	1.00	$H_1 : \beta_i > 0$		
Retail sales	-0.102	-1.53	0.94	1.00	$H_1 : \beta_i > 0$		
Industrial production	0.190	3.03	<b>0.00</b>	<b>0.04</b>	$H_1 : \beta_i > 0$		
Capacity utilization	-0.184	-2.63	1.00	1.00	$H_1 : \beta_i > 0$		
Personal income	-0.125	-1.01	0.84	1.00	$H_1 : \beta_i > 0$		
Consumer credit	-0.046	-0.91	0.82	1.00	$H_1 : \beta_i > 0$		
New home sales	-0.047	-1.23	0.89	1.00	$H_1 : \beta_i > 0$		
Personal consumption	0.050	0.74	0.23	1.00	$H_1 : \beta_i > 0$		
Durable goods orders	-0.013	-0.18	0.57	1.00	$H_1 : \beta_i > 0$		
Construction spending	-0.068	-0.57	0.72	1.00	$H_1 : \beta_i > 0$		
Factory orders	0.112	1.29	0.10	0.95	$H_1 : \beta_i > 0$		
Business inventories	0.023	0.24	0.60	1.00	$H_1 : \beta_i < 0$		
Government budget deficit	-0.071	-1.07	0.86	1.00	$H_1 : \beta_i > 0$		
Trade balance	-0.009	-0.20	0.58	1.00	$H_1 : \beta_i > 0$		
PPI	-0.002	-0.07	0.53	1.00	$H_1 : \beta_i > 0$		
Core PPI	0.068	1.67	<b>0.05</b>	0.77	$H_1 : \beta_i > 0$		
CPI	-0.137	-2.49	0.99	1.00	$H_1 : \beta_i > 0$		
Core CPI	0.010	0.39	0.35	1.00	$H_1 : \beta_i > 0$		
CCI preliminary (Michigan)	-0.106	-1.64	0.95	1.00	$H_1 : \beta_i > 0$		
CCI final (Michigan)	-0.023	-0.43	0.67	1.00	$H_1 : \beta_i > 0$		
CCI (board)	0.006	0.08	0.47	1.00	$H_1 : \beta_i > 0$		
NAPM index	-0.169	-1.29	0.90	1.00	$H_1 : \beta_i > 0$		
Housing starts	-0.005	-0.08	0.53	1.00	$H_1 : \beta_i > 0$		
Index of leading indicators	-0.042	-0.32	0.63	1.00	$H_1 : \beta_i > 0$		
Target rate surprises	-0.013	-0.42	0.34	1.00	$H_1 : \beta_i < 0$		
Initial claims	-0.017	-0.34	0.37	1.00	$H_1 : \beta_i < 0$		

All regressions in part A include a constant, as does the regression in part B. Data mining robust  $p$ -values were computed based on a parametric bootstrap approach under the null hypothesis of no predictability. Standard  $p$ -values based on  $N(0,1)$  distribution. Boldface indicates statistical significance at 10% level.

report the  $R^2$  of the regression and the number of observations in each regression. In the case of the individual regressions, that sample size corresponds to the number of news announcements over the sample period.

We test  $H_0 : \beta_i = 0$  against the one-sided alternative hypotheses suggested by economic theory. In particular, a positive news shock about measures of current or future output (and its components) or about employment should be associated with a positive response. The same is true for unanticipated increases in the price level. In contrast, positive news shocks about the unemployment rate and initial claims should be associated with declining energy prices. Similarly, positive interest rate shocks tend to be associated with a decline of economic activity and hence lower energy prices.<sup>4</sup> Finally, an unanticipated increase of business inventories is interpreted as evidence of an economic slowdown and is associated with a negative sign. The use of one-sided  $t$ -tests not only makes economic sense in this context, but it also improves substantially the power of tests of predictability, as discussed in Inoue and Kilian (2004).

### III. Should We Think of Oil Prices as Asset Prices?

It is well documented that stock prices and exchange rates fully and systematically respond within the same day to macroeconomic news announcements (see Andersen et al., 2003, 2007). Given the perception that oil prices behave much like asset prices that respond instantaneously to all news, it is natural to contrast the response of oil and gasoline prices to macroeconomic news shocks to that of commonly studied asset prices. A useful starting point is the nonfarm payroll report. Of the thirty macroeconomic news announcements we analyze, the nonfarm payroll report is one of the most closely observed U.S. macroeconomic announcements. Andersen and Bollerslev (1998), among others, refer to this announcement as the “king” of announcements because of the sensitivity of most asset prices to its release. Tables 2A and 3A, however, suggest that nonfarm payroll announcements have no effect on retail gas prices and crude oil prices using conventional asymptotic  $p$ -values. This is a first indication that oil and gasoline prices should not be thought of as asset prices. In fact, even the  $R^2$  estimates of 0.02% and 0.37% for nonfarm payroll employment are strikingly low compared with the estimates that would be obtained for the corresponding sample periods using other asset returns. The latter estimates may be as high as 20% in some cases.

A second indication is that the  $R^2$  of the joint regressions in tables 2B and 3B tends to be very low. In the joint regression, all macroeconomic news shocks combined explain

only 0.38% of the variation in oil prices (and only 1.91% of the variation in gasoline prices). Put differently, more than 99% (or more than 98%) of the variation in daily energy prices is driven by factors not correlated with domestic macroeconomic aggregates. These  $R^2$  estimates also tend to be lower than those for similar regressions for other asset prices. For example, for daily S&P 500 returns, which are known to be among the least predictable asset returns, the  $R^2$  estimates for the corresponding sample periods and joint regressions are 2.0% and 3.2%, respectively. At the other extreme, for daily ten-year bond returns, we obtained  $R^2$  estimates of 4.9% and 8.2%, respectively, confirming the impression that macroeconomic news is less informative for oil and gasoline prices than for financial asset prices.

The  $R^2$  estimates for the individual regressions, at least for gasoline prices, seem at first sight to paint a more favorable picture for the asset market interpretation.<sup>5</sup> Whereas for the price of crude oil the individual  $R^2$  exceeds 2% only in one case, in the case of gasoline prices, the individual  $R^2$  estimates tend to be higher in general, with nine estimates exceeding 2%, of which two even exceed 5%. There is reason to be cautious in interpreting these individual  $R^2$  results, however. For example, the National Association of Purchasing Managers (NAPM) index appears to explain 5.58% of the variation in U.S. gasoline prices, but it has a coefficient of the wrong sign. The fact that quite frequently the estimated coefficients are of the wrong sign is an indication that the regression fit is likely to be spurious. For example, in tables 2A and 2B, unanticipated increases in retail sales, personal income or consumption, or durables goods orders, should increase the price of oil, not lower it. The same is true for inflation surprises, yet three of four inflation news shocks have negative coefficients. Moreover, the signs of different inflation measures are mutually contradictory.

If we focus on the statistical significance of the one-sided  $t$ -tests using conventional asymptotic critical values, a somewhat different picture emerges. For the price of crude oil, only six predictors appear statistically significant at the 10% level in the individual regressions (see table 2A). In the joint regression, only five predictors remain statistically significant at the 10% level (see table 2B). For the price of gasoline, there are three rejections using individual regressions in table 3A and two rejections for the joint regression in table 3B. The statistically significant predictors are not the same in both markets, which again suggests that the results are likely to be spurious. For gasoline prices, industrial production and core PPI are most significant (with mixed results for factory orders), whereas for crude oil preliminary GDP announcements, new home sales, net government purchases, the core CPI, and housing starts are most significant (with mixed results for the Conference Board's consumer confidence measure).

<sup>4</sup> An alternative view is that interest rate cuts may signal weaker-than-expected economic growth to financial markets (Bernanke & Kuttner, 2005). This interpretation would suggest a positive sign for the interest rate coefficient. However, there does not appear to be empirical evidence in support of that alternative view, and theoretical models overwhelmingly predict a negative sign (see, e.g., Barsky & Kilian, 2002).

<sup>5</sup> In interpreting the results, it is useful to keep in mind that the individual regressions are based on a different data set than the joint regressions, so the magnitude of the  $R^2$  estimates is not comparable.

Although many of these variables are not part of the regression models that have been used to study the transmission of energy prices shocks, it may be tempting to interpret these rejections as evidence that the assumption of predetermined energy prices is suspect. This interpretation, however, is questionable. For one thing, it is odd that among news variables that are conceptually closely related, only some appear to have predictive power. For example, we would expect GDP and industrial production news to have similar effects on energy prices.

More importantly, that interpretation would ignore that we have conducted not one  $t$ -test in assessing the evidence against that assumption but thirty  $t$ -tests. Conventional critical values do not account for repeated applications of the same test to alternative regressors. The failure to account for such data mining is known to cause spurious rejections of the null of no feedback (see, e.g., Inoue & Kilian, 2004). The problem of data mining is well recognized in the literature (see, e.g., Denton, 1985). If we investigate whether at least one of many predictors is statistically significant, the probability of rejecting the null hypothesis of no predictability at conventional significance levels increases with the number of predictors considered, resulting in spurious rejections of the null hypothesis of no predictability when that null hypothesis is in fact true. Such data mining problems have been shown to be practically important in a variety of related contexts, including the search for calendar effects in stock returns and the search for profitable technical trading rules (White, 2000; Sullivan, Timmermann, & White, 2001).

Inoue and Kilian (2004) discuss appropriate adjustments to the null distribution of predictability tests in the presence of data mining. The basic idea is to compute data mining robust critical values for the supremum of the  $t$ -statistic across the thirty alternative regressors. In practice, this may be accomplished by bootstrap methods. We simulate the finite sample distribution of the supremum of the  $t$ -statistic under the null hypothesis of no feedback. For simplicity, we postulate that returns and news shocks are independent and identically distributed (i.i.d.) normally distributed with the variances found in the actual data. We abstract from the possibility of fat tails, heteroskedasticity, or serial correlation under the null hypothesis. Accounting for these possible departures from i.i.d. normality, if anything, would tend to increase further the data mining robust critical values constructed below. We treat the news shocks as mutually independent.<sup>6</sup> The empirical distribution of the supremum  $t$ -statistic is constructed by estimating the regression models in question in each bootstrap sample and tabulating the distribution of the largest  $t$ -statistic among the thirty alternative predictors. All results are based on 100,000 bootstrap

replications. The bootstrap replicates of the individual regressions take account of the differences in sample size across regressions. When bootstrapping the joint regression, we treat the timing of the news shocks as exogenously given in repeated sampling. This makes sense because the announcements are prescheduled.

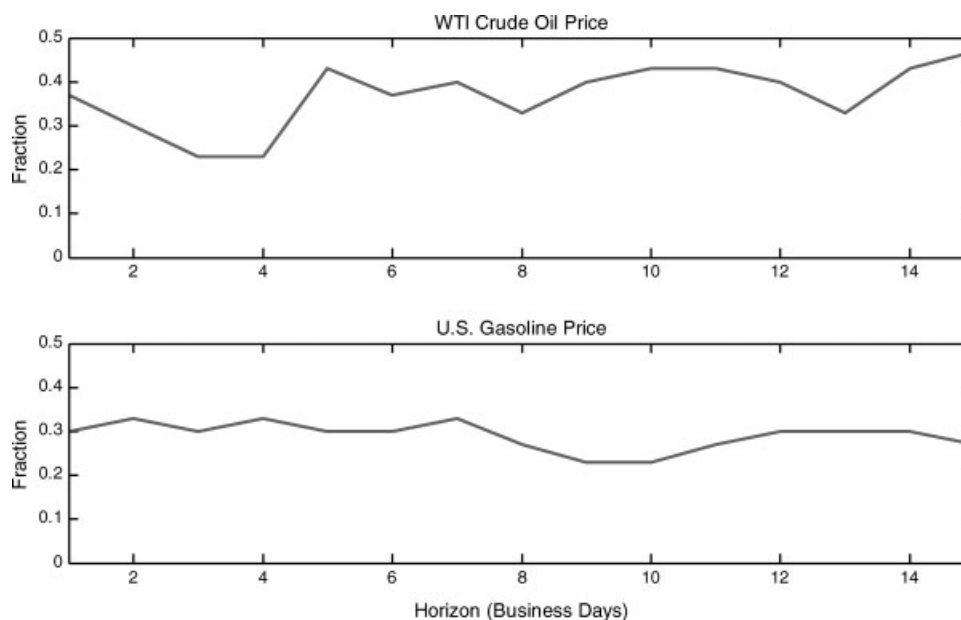
After adjusting for data mining, none of the statistically significant results in table 2 remains. For example, the 5% data mining robust critical value for table 2A is 2.96, and the 10% critical value rises to 2.72. For the price of crude oil, the lowest  $p$ -value is 0.23 in the individual regressions and 0.46 in the joint regression. These results suggest that there is no empirical evidence that daily WTI crude oil prices respond to macroeconomic news shocks on impact.<sup>7</sup>

For gasoline prices in table 3A, the lowest  $p$ -value is 0.88 in the individual regressions, and no result remains statistically significant. In table 3B, only the result for industrial production remains marginally statistically significant at the 5% level. Given that this result is inconsistent with the corresponding  $p$ -value in table 3A and with results for economically closely related types of news, caution seems called for in interpreting this evidence. The next lowest data mining robust  $p$ -value in table 3B is 0.77.

The finding of virtually no significant feedback at the daily frequency is also consistent with the observation that the distribution of the thirty  $t$ -statistics in tables 2A and 3A is roughly what we would expect in the absence of feedback in that approximately one-third of the  $t$ -statistics exceed unity in absolute value. The reason that we focus on daily returns as our starting point is that one would expect our statistical test to have the highest possible power to detect asset price dynamics immediately following the macroeconomic news. To the extent that oil markets respond to macroeconomic news only with a delay or gradually, however, one might expect the strength of the feedback to the price of oil to grow at longer horizons. Figure 1 investigates this point. It shows the fraction of  $t$ -statistics in excess of unity in absolute value as a function of the horizon (expressed in days) over which the cumulative returns are computed. In the case of the price of oil, there is a slight, if nonmonotonic, tendency for that fraction to increase with the horizon, although not beyond what sample variation could explain under the null of no feedback. For the price of gasoline, there is no indication that the fraction increases within the first fifteen business days. These patterns are suggestive of a very slow diffusion of macroeconomic news, unlike in typical asset price models. In the next section, we investigate in more detail how strong that feedback is at the monthly horizon relevant to the identifying assumption of predetermined energy prices in VAR models of the transmission of energy price shocks.

<sup>6</sup> This assumption is empirically plausible except for news announcements on closely related series that occur on the same day (such as news announcements for the core CPI and the CPI). The latter situation is an exception. We experimented with alternative assumptions that account for the possible dependence of these announcements. The results reported are robust to these alternative assumptions.

<sup>7</sup> Note that many of the data-mining robust  $p$ -values are effectively 1.000. The reason is that we compare all individual  $t$ -test statistics to the null distribution of the maximum  $t$ -statistic. Alternatively, one could focus on the largest of the thirty  $t$ -statistics only. The substantive interpretation of the results would be the same.

FIGURE 1.—FRACTION OF  $t$ -STATISTICS IN EXCESS OF UNITY AT INCREASING HORIZONS

The plot shows the fraction of  $t$ -statistics exceeding unity in absolute value as a function of the length of time (in days) since the macroeconomic news were made public. Under the null of independence and no feedback, one would expect a fraction of 0.32.

#### IV. Testing the Assumption of Predetermined Energy Prices at Monthly Horizons

The formal evidence we have presented so far was based on the reaction of energy prices to news shocks within the day. This approach made sense because financial asset prices are known to adjust fully to news announcements within the day (Andersen et al., 2007), and a systematic rejection of the no-feedback null hypothesis at daily horizons would have sufficed to reject the assumption of predetermined energy prices at monthly frequency. Since we did not reject the null for any news shock, some additional analysis is required. The reason is that even if energy prices are not asset prices in the same sense as exchange rates or stock prices, they may still significantly respond to macroeconomic news shocks within the month, invalidating the commonly used identifying assumption of predetermined oil prices. In this section, we address this concern by specifying a regression for the percentage change in energy prices between close of business on the trading day preceding the news shock  $S_{it}$  and thirty calendar days later:

$$R_{t+1}^h = \alpha + \beta_i S_{it} + \varepsilon_{t+1}^h, \quad (3)$$

where  $R_{t+1}^h = 100 \times \ln(P_{t+h}/P_t)$  denotes the monthly return on energy from the end of day  $t-1$  to the end of day  $t+h-1$ ,  $h = 20$  (since there are five business days per week), and the one-step ahead predictive error  $\varepsilon_{t+1}^h$  is serially correlated under  $H_0 : \beta_i = 0$ , necessitating the use of Newey-West standard errors. As before, the estimates are

based only on data for those dates for which an announcement was made on day  $t$ . Alternatively, we consider the joint regression:

$$R_{t+1}^h = \alpha + \sum_{i=1}^{30} \beta_i S_{it} + \varepsilon_{t+1}^h. \quad (4)$$

One concern is that one-month-ahead regressions may lack the power to detect predictability, because we need to estimate the effect of news shocks among a myriad of other changes that take place over the course of one month. We address this concern by focusing on the WTI price of crude oil, for which 6,214 observations spanning 25 years of data are available. The comparatively large sample size helps increase the power of the test. As table 4 shows, conventional  $p$ -values indicate about as many rejections of the null of no predictability at the monthly horizon as in the earlier daily analysis, suggesting that low power is not a concern. The first two columns of  $p$ -values in table 4 refer to the results from the thirty individual regressions and the next two columns to the results from the joint regression. The  $R^2$  estimate from the joint regression is somewhat larger at the monthly horizon than at the daily horizon. It rises from 0.38% to 0.69%. This pattern is consistent with the increasing importance of feedback from macroeconomic news shocks at longer horizons. In absolute terms, however, the feedback continues to be negligible, even abstracting from the dangers of overfitting.

Conventional  $t$ -tests indicate five rejections at the 5% level in the individual regressions (capacity utilization, net



TABLE 4.—MONTHLY WTI CRUDE OIL PRICES: REGRESSIONS, 1983–2008

Announcement	Individual Regressions		Joint Regression		Alternative Hypothesis
	Standard $p$ -Value	Robust $p$ -Value	Standard $p$ -Value	Robust $p$ -Value	
GDP advanced	0.97	1.00	0.97	1.00	$H_1 : \beta_i > 0$
GDP preliminary	0.16	1.00	0.17	1.00	$H_1 : \beta_i > 0$
GDP final	<b>0.10</b>	0.95	0.12	0.98	$H_1 : \beta_i > 0$
Unemployment rate	0.19	1.00	0.18	1.00	$H_1 : \beta_i < 0$
Nonfarm payroll	0.38	1.00	0.43	1.00	$H_1 : \beta_i > 0$
Retail sales	0.63	1.00	0.53	1.00	$H_1 : \beta_i > 0$
Industrial production	0.30	1.00	0.81	1.00	$H_1 : \beta_i > 0$
Capacity utilization	<b>0.05</b>	0.79	<b>0.04</b>	0.72	$H_1 : \beta_i > 0$
Personal income	0.20	1.00	0.15	0.99	$H_1 : \beta_i > 0$
Consumer credit	0.69	1.00	0.69	1.00	$H_1 : \beta_i > 0$
New home sales	0.40	1.00	0.46	1.00	$H_1 : \beta_i > 0$
Personal consumption	0.69	1.00	0.66	1.00	$H_1 : \beta_i > 0$
Durable goods orders	0.17	1.00	0.23	1.00	$H_1 : \beta_i > 0$
Construction spending	0.36	1.00	0.33	1.00	$H_1 : \beta_i > 0$
Factory orders	0.28	1.00	0.27	1.00	$H_1 : \beta_i > 0$
Business inventories	0.15	0.99	0.17	1.00	$H_1 : \beta_i < 0$
Government budget deficit	<b>0.03</b>	0.64	<b>0.04</b>	0.74	$H_1 : \beta_i > 0$
Trade balance	<b>0.10</b>	0.96	0.13	0.98	$H_1 : \beta_i > 0$
PPI	0.98	1.00	0.93	1.00	$H_1 : \beta_i > 0$
Core PPI	0.83	1.00	0.48	1.00	$H_1 : \beta_i > 0$
CPI	0.42	1.00	0.41	1.00	$H_1 : \beta_i > 0$
Core CPI	0.38	1.00	0.64	1.00	$H_1 : \beta_i > 0$
CCI preliminary (Michigan)	<b>0.01</b>	0.21	<b>0.02</b>	0.39	$H_1 : \beta_i > 0$
CCI final (Michigan)	0.45	1.00	0.41	1.00	$H_1 : \beta_i > 0$
CCI (board)	<b>0.01</b>	0.18	<b>0.01</b>	0.26	$H_1 : \beta_i > 0$
NAPM index	0.44	1.00	0.47	1.00	$H_1 : \beta_i > 0$
Housing starts	0.83	1.00	0.79	1.00	$H_1 : \beta_i > 0$
Index of leading indicators	<b>0.01</b>	0.25	<b>0.01</b>	0.17	$H_1 : \beta_i > 0$
Target rate surprises	0.63	1.00	0.60	1.00	$H_1 : \beta_i < 0$
Initial claims	0.41	1.00	0.44	1.00	$H_1 : \beta_i < 0$

See tables 2A and 2B. The  $R^2$  of the joint regression is 0.69%.

government purchases, preliminary Michigan consumer confidence, Conference Board consumer confidence, and index of leading indicators) and two additional rejections at the 10% level (GDP final and trade balance). In the joint regression, we obtain the same five rejections at the 5% level, with no additional rejections at the 10% level.

As in the daily analysis, there is reason to distrust these  $p$ -values. It is not uncommon for the point estimates underlying table 4 to be of the wrong sign—in some cases, even significantly so. For example, GDP (advanced) in both the individual and joint regression has a  $t$ -statistic of about  $-1.9$ . Using more appropriate data-mining robust critical values constructed along the lines described in section III, none of the  $t$ -statistics remains statistically significant. The 5% critical value rises to 2.965, the 10% critical value to 2.726. The lowest  $p$ -value in the joint regression is obtained for the index of leading indicators with 0.17; for the individual regressions, it is 0.18 for the Conference Board's index of consumer confidence. There is no evidence of within-the-month feedback from industrial production, consumer expenditures, the unemployment rate, consumer prices, or interest rates, in particular. These are the variables most widely used in monthly regressions aimed at uncovering the effects of oil price shocks on domestic aggregates. The results in table 4 support the common practice of treating

oil prices as predetermined with respect to U.S. macroeconomic aggregates.

## V. Joint Tests for Subsets of Macroeconomic News

An alternative approach to addressing the potential for data mining based on the joint regression is to construct tests for the joint statistical significance of subsets of news shocks related to the same economic concept. For example, the first ten news shocks jointly with the last shock all represent news about *domestic aggregate real activity*. If we add news shocks 11 through 18 to this set, we obtain the set of all *aggregate and disaggregate measures of domestic real activity*. News shocks 19 through 22 represent *inflation shocks*, and news shocks 23 through 28 represent *forward-looking indicators*.

A natural approach is to focus on the sum of all coefficients (suitably normalized to account for the expected sign) in each subset of news variables. Since there are only four sets of predictors, the scope for data mining is limited, and conventional critical values are likely to be only mildly downward biased. For the monthly horizon relevant to the assumption of predetermined energy prices, table 5 shows that the feedback from the set of news about aggregate real activity and the set of domestic inflation news to the WTI

TABLE 5.—*p*-VALUES OF JOINT SIGNIFICANCE TESTS

Announcement	WTI Price of Crude Oil		U.S. Gasoline Price	
	Daily	Monthly	Daily	Monthly
Aggregate real activity $i = 1,2,3,4,5,6,7,8,9,10,30$	0.47	0.25	0.71	<b>0.05</b>
Aggregate and disaggregate real activity $i = 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,30$	0.32	0.06	0.73	0.12
Inflation $i = 19,20,21,22$	0.58	0.86	0.72	0.66
Forward-looking variables $i = 23,24,25,26,27,28$	0.06	<b>0.00</b>	0.93	<b>0.05</b>

The index  $i$  refers to the news shocks in the order listed in tables 2 and 3. Boldface indicates significance using standard asymptotic critical values at the 5% level. The underlying test is based on the sum of the coefficients in question (suitably normalized to account for the expected sign). For a similar approach in a different context, see Hamilton (2003).

price of oil is not statistically significant at conventional significance levels. The combined set of aggregate and disaggregate real activity news, however, is jointly significant at the 10% level, as is the set of forward-looking news variables at the 1% level. For the price of gasoline, two of the four sets are statistically significant at the 5% level. Unlike our earlier results, this evidence is supportive of the existence of feedback at the monthly horizon at least from some types of macroeconomic news. This feedback could potentially invalidate the assumption of predetermined energy prices.

One reason that suggests caution in interpreting these results is that, especially at the monthly horizon, one would expect oil and gasoline prices to respond similarly to the same macroeconomic news. This is the case for the set of forward-looking news variables but not for aggregate real activity, for example. Another way of gauging the plausibility and, more importantly, the practical relevance of these test results is to focus on the explanatory power of these news shocks as measured by  $R^2$  rather than statistical significance alone. Quantitatively important feedback at the monthly horizon should be reflected in nontrivial regression fits. Our results, however, imply that only 0.69% of the monthly variation in oil prices is explained by all news shocks combined, and hence even less by any subset of these predictors. Even if we restrict ourselves to days on which announcements about forward-looking variables took place, which tends to result in larger  $R^2$  estimates as shown in table 2, the  $R^2$  of the set of all forward-looking variables is only 0.38%. In other words, if there is feedback within the month, it is so weak that we can ignore it in practice. Thus, the results in table 5 do little to overturn our earlier evidence in favor of the assumption that oil prices can be treated as predetermined with respect to monthly measures of domestic macroeconomic aggregates. Similarly, only 1.6% of the monthly variation in gasoline prices can be explained by all macroeconomic news shocks combined. Restricting ourselves to announcement dates, the combined  $R^2$  of all forward-looking variables is 2.28%. This is much larger than in the case of crude oil prices but still represents only a small fraction of the month-to-month variability in gasoline prices. As in the case of the price of oil, these results are broadly supportive of the assumption of predetermined gasoline prices as a reasonable approximation.

At the daily horizon, table 5 shows no significant rejections of the null hypothesis of no feedback to the retail price of gasoline for any of the four sets, suggesting that the one rejection found in table 3B was indeed a fluke. The corresponding results for the WTI price of oil show no evidence of feedback either except perhaps for the set of forward-looking variables. That set is significant at the 10% level but not the 5% level. Although the evidence for the price of oil is stronger than for the tests based on one news item at a time, it is very weak compared with the results of similar regressions for other asset returns (such as the ten-year bond yield) that tend to show rejections at the 1% level for all four subsets of macroeconomic news. Moreover, the marginal rejection for the set of forward-looking news variables is not robust to small increases in the horizon. Combined with the poor regression fit documented in table 2B, the evidence of feedback at the daily horizon from forward-looking news to WTI prices in table 5 must be considered extremely tenuous.

## VI. Are News-Based Tests Powerful Enough for Testing the No-Feedback Hypothesis?

An important question is how informative news-based tests of the no-feedback hypothesis are. There is reason to expect our statistical test to have enough power against feedback occurring within the day at least. Any response in the price of oil within a few minutes or hours should be reflected in the daily return by construction. Certainly, to the extent that the common perception is correct that these news shocks are one of the main explanations of daily oil and gasoline price movements, one would expect our test to detect this price response. The power of news-based tests of the null hypothesis of no feedback may be verified by conducting similar statistical tests using other asset returns (such as bond yields or exchange rate returns) for the same sample period and the same set of macroeconomic news. As discussed in section III, under the null of no feedback, one would expect about 32% of the thirty two-sided  $t$ -statistics in model (1) to exceed unity. It can be shown that in analogous daily regressions for the ten-year bond yield 70% of the two-sided  $t$ -statistics exceed unity. Likewise, for the DM-USD exchange rate, 53% of the two-sided  $t$ -statistics exceed unity. These examples demonstrate that our test

of no feedback does have statistical power at short horizons and that the absence of significant rejections for oil and gasoline prices is informative.

The power of our test at longer horizons, in contrast, is more doubtful. The premise of our approach is that few other shocks occur near announcement dates, allowing us to get a good estimate of the effect of that news item. While this assumption is credible at daily horizons, the longer the horizon is, the more the variability of other shocks in the intervening period is likely to obscure that causal link. Hence, with increasing horizons, we expect estimates of the strength of the feedback to become unreliable and tests of the null hypothesis of no feedback to lose power. It can be shown that at horizons of twenty business days, for example, the fraction of two-sided *t*-statistics exceeding unity is not systematically higher than under the null hypothesis even for the ten-year bond yield and the DM-USD exchange rate return, suggesting that the power of the tests on individual news items has all but vanished at that horizon. The corresponding joint test in table 5, however, has higher power than individual *t*-tests and does generate rejections for other asset returns—not just at the daily horizon but even at the monthly horizon. For example, we find that both sets of real activity news have jointly significant effects at the 5% significance level on the ten-year bond yield even at the horizon of twenty business days. This evidence suggests that the joint test retains enough power to be informative at the monthly horizon of interest in testing the assumption of predetermined energy prices.<sup>8</sup>

## VII. Conclusion

Our analysis in this paper established that oil prices, unlike financial asset prices, do not respond instantaneously to domestic macroeconomic news. We showed that there is no compelling evidence of such feedback in daily WTI oil price data for the period 1983 to 2008. Ninety-nine percent of the variation in crude oil prices is left unexplained by domestic macroeconomic news. Similar results were obtained for U.S. gasoline prices using a much shorter sample. There was no evidence of a statistically significant, strong, and systematic response to domestic macroeconomic news at daily horizons.

Our analysis also shed light on the validity of the commonly used identifying assumption that energy price shocks

in VAR models are predetermined with respect to domestic macroeconomic aggregates. Testing this assumption is complicated by the fact that exactly identifying assumptions are inherently untestable. We overcame that problem by estimating the response of daily WTI crude oil prices and U.S. gasoline prices at various horizons to U.S. macroeconomic news shocks. Because these shocks are exogenous by construction, we were able to estimate their effect on energy prices and test for feedback from U.S. macroeconomic aggregates to energy prices within the month. For a wide range of macroeconomic aggregates commonly used in studies of the transmission of energy price shocks (including U.S. real output and consumption, interest rates, and inflation), we found no evidence at all of statistically significant feedback within the month from macroeconomic news to the price of crude oil or the price of gasoline.

Somewhat stronger results were obtained from tests of the joint significance of broader sets of news variables. The results most favorable to the hypothesis that there is feedback from the U.S. economy to energy prices within the month were obtained with a set of forward-looking news variables. For example, although none of the forward-looking news variables (including the index of leading indicators) were individually significant in predicting the price of crude oil at monthly horizons, they were jointly statistically significant at the 1% level. Considering the low overall explanatory power of all news shocks combined of less than 1%, the extent of the feedback to the price of crude oil within the same month seems minimal, however, suggesting that the assumption of predetermined oil prices is a good approximation in practice.

Similar evidence of feedback from the set of forward-looking news variables was obtained for gasoline prices. The latter results are necessarily more tentative given the much smaller sample size. In any case, the overall explanatory power of all macroeconomic news shocks combined for gasoline prices is below 2% at the monthly horizon, suggesting that the assumption of no contemporaneous feedback provides a good approximation at monthly frequency, even for gasoline prices.

We concluded that the widely used assumption that energy prices are predetermined at monthly frequency is broadly consistent with the data, lending support to empirical as well as theoretical models of the transmission of energy price shocks based on that assumption. At the same time, our results cast doubt on empirical work based on the alternative assumption that energy prices should be ordered below domestic macroeconomic aggregates in recursively identified VAR models.

Our analysis focused appropriately on the spot price of oil and on transaction prices for gasoline. The WTI spot price is essentially equal to the one-month NYMEX oil futures price. We did not investigate oil futures prices of longer maturities because such futures prices have not been included in any of the VAR models that motivated our analysis. Although a natural conjecture is that such oil futures

<sup>8</sup> It is widely accepted that oil prices are endogenous with respect to U.S. macroeconomic aggregates (Barsky & Kilian, 2004; Hamilton, 2008; Kilian, 2009). If the price of oil is endogenous, there must be feedback from exogenous variation in macroeconomic variables to the price of oil at sufficiently long lags. Thus, it may seem that we could test endogeneity by simply increasing the horizon in model (3) beyond one month until we detect significant evidence of feedback, but our methodology is not designed to detect feedback at such long horizons. Even the power of the joint test will dissipate at horizons in excess of one month. Hence, we do not take a stand on the issue of endogeneity in this paper. Because VAR models impose no restrictions on lag coefficients, VAR estimates obtained under the identifying assumption of predetermined energy prices will be consistent whether the price of oil is endogenous or not.

prices may be more forward looking than the spot price of oil and may provide stronger evidence of feedback from macroeconomic news, there is no evidence to support that view. Additional estimates based on three-month NYMEX oil futures prices produced results very similar to the WTI price. The evidence for much longer maturities is more difficult to assess, because even at horizons as short as one year, oil futures markets become increasingly illiquid. Sometimes only one contract is traded on a given day, making the price data unreliable. At even longer horizons, there is no trading at all for extended periods, making it impossible to estimate the regressions of interest.

## REFERENCES

- Alquist, Ron, and Lutz Kilian, "What Do We Learn from the Price of Crude Oil Futures?" *Journal of Applied Econometrics* 25 (2010), 539–573.
- Andersen, Torben G., and Tim Bollerslev, "DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies," *Journal of Finance* 53 (1998), 219–265.
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Clara Vega, "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review* 93 (2003), 38–62.
- "Real-Time Price Discovery in Stock, Bond, and Foreign Exchange Markets," *Journal of International Economics* 73 (2007), 251–277.
- Barsky, Robert B., and Lutz Kilian (2002), "Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative" (pp. 137–183) in B. S. Bernanke and K. Rogoff (Eds.), *NBER Macroeconomics Annual* (Cambridge, MA: MIT Press, 2002).
- "Oil and the Macroeconomy since the 1970s," *Journal of Economic Perspectives* 18 (2004), 115–134.
- Bernanke, Ben S., Mark Gertler, and Mark W. Watson, "Systematic Monetary Policy and the Effects of Oil Price Shocks," *Brookings Papers on Economic Activity* 1 (1997), 91–157.
- "Reply to Oil Shocks and Aggregate Economic Behavior: The Role of Monetary Policy," *Journal of Money, Credit and Banking* 36 (2004), 287–291.
- Bernanke, Ben S., and Kenneth N. Kuttner, "What Explains the Stock Market's Reaction to Federal Reserve Policy?" *Journal of Finance* 60 (2005), 1221–1257.
- Blanchard, Olivier J., and Jordi Galí, "The Macroeconomic Effects of Oil Price Shocks: Why Are the 2000s So Different from the 1970s?" in J. Galí and M. Gertler (Eds.), *International Dimensions of Monetary Policy* (Chicago: University of Chicago Press, 2010), 373–428.
- Davis, Steven J., and John Haltiwanger, "Sectoral Job Creation and Destruction Responses to Oil Price Changes," *Journal of Monetary Economics* 48 (2001), 465–512.
- Davis, Lucas W., and Lutz Kilian, "Estimating the Effect of a Gasoline Tax on Carbon Emissions," *Journal of Applied Econometrics*, forthcoming.
- Denton, Frank T., "Data Mining as an Industry," this REVIEW 67 (1985), 124–127.
- Edelstein, Paul, and Lutz Kilian, "The Response of Business Fixed Investment to Changes in Energy Prices: A Test of Some Hypotheses About the Transmission of Energy Price Shocks," *B. E. Journal of Macroeconomics* 7 (2007). <http://www.bepress.com/bejm/vol7/iss1/art35>.
- "How Sensitive Are Consumer Expenditures to Retail Energy Prices?" *Journal of Monetary Economics* 56 (2009), 766–779.
- Faust, Jon, John Rogers, Shing-Yi Wang, and Jonathan Wright, "The High Frequency Response of Exchange Rates and Interest Rates to Macroeconomic Announcements," *Journal of Monetary Economics* 54 (2007), 1051–1068.
- Frankel, Jeffrey, and Andrew Rose, "Determinants of Agricultural and Mineral Commodity Prices," in R. Fry, C. Jones, and C. Kent (Eds.), *Inflation in an Era of Relative Price Shocks* (Sydney: Reserve Bank of Australia, 2010), 9–51.
- Hamilton, James D., "What Is an Oil Shock?" *Journal of Econometrics* 113 (2003), 363–398.
- "Oil and the Macroeconomy," in S. Durlauf and L. Blume (Eds.), *The New Palgrave Dictionary of Economics*, 2nd ed. (New York: Palgrave Macmillan, 2008).
- Hamilton, James D., and Ana María Herrera, "Oil Shocks and Aggregate Economic Behavior: The Role of Monetary Policy: Comment," *Journal of Money, Credit, and Banking* 36 (2004), 265–286.
- Herrera, Ana María, "Oil Price Shocks, Inventories, Oil Shocks and Macroeconomic Dynamics," mimeograph, Department of Economics, Wayne State University (2009).
- Herrera, Ana, María, and Elena Pesavento, "Oil Price Shocks, Systematic Monetary Policy, and the 'Great Moderation,'" *Macroeconomic Dynamics* 13 (2009), 107–137.
- Inoue, Atsushi, and Lutz Kilian, "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?" *Econometric Reviews* 23 (2004), 371–402.
- Kilian, Lutz, "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review* 99 (2009), 1053–1069.
- "Explaining Fluctuations in Gasoline Prices: A Joint Model of the Global Crude Oil Market and the U.S. Retail Gasoline Market," *Energy Journal* 31 (2010), 105–130.
- Kilian, Lutz, and Cheolbeom Park, "The Impact of Oil Price Shocks on the U.S. Stock Market," *International Economic Review* 50 (2009), 1267–1287.
- Leduc, Sylvain, and Keith Sill, "A Quantitative Analysis of Oil-Price Shocks, Systematic Monetary Policy, and Economic Downturns," *Journal of Monetary Economics* 51 (2004), 781–808.
- Lee, Kiseok, and Shawn Ni, "On the Dynamic Effects of Oil Price Shocks: A Study Using Industry Level Data," *Journal of Monetary Economics* 49 (2002), 823–852.
- Rotemberg, Julio J., and Michael Woodford, "Imperfect Competition and the Effects of Energy Price Increases on Economic Activity," *Journal of Money, Credit, and Banking* 28, Part 1 (1996), 550–577.
- Sullivan, Ryan, Allan Timmermann, and Halbert L. White, "Dangers of Data Mining: The Case of Calendar Effects in Stock Returns," *Journal Econometrics* 105 (2001), 249–286.
- White, Halbert L., "A Reality Check for Data Snooping," *Econometrica* 68 (2000), 1097–1126.