Bocconi

IEFE
Centre for Research on Energy and Environmental Economics and Policy

Working Paper Series - ISSN 1973-0381

Forecasting the Oil-Gasoline Price Relationship: Should We Care About the Rockets and the Feathers?

Andrea Bastianin, Marzio Galeotti and Matteo Manera

Working Paper n. 62

February 2014

IEFE - The Center for Research on Energy and Environmental Economics and Policy at Bocconi University
via Guglielmo Röntgen, 1 - 20136 Milano
tel. 02.5836.3820 - fax 02.5836.3890
www.iefe.unibocconi.it – iefe@unibocconi.it

This paper can be downloaded at www.iefe.unibocconi.it
The opinions expressed herein do not necessarily reflect the position of IEFE-Bocconi.

Forecasting the Oil-Gasoline Price Relationship: Should We Care About the Rockets and the Feathers?

Andrea Bastianin
University of Milan and FEEM

Marzio Galeotti
University of Milan and IEFE-Bocconi

Matteo Manera
University of Milan-Bicocca and FEEM

February 2014

Abstract. According to the Rockets and Feathers hypothesis (RFH), the transmission mechanism of positive and negative changes in the price of crude oil to the price of gasoline is asymmetric. Although there have been many contributions documenting that downstream prices are more reactive to increases than to decreases in upstream prices, little is known about the forecasting performance of econometric models incorporating asymmetric price transmission from crude oil to gasoline. In this paper we fill this gap by comparing point, sign and probability forecasts from a variety of Asymmetric-ECM (A-ECM) and Threshold Autoregressive ECM (TAR-ECM) specifications against a standard ECM. Forecasts from A-ECM and TAR-ECM subsume the RFH, while the ECM implies symmetric price transmission from crude oil to gasoline. We quantify the forecast accuracy gains due to incorporating the RFH in predictive models for the prices of gasoline and diesel. We show that the RFH is useless for point forecasting, while it can be exploited to produce more accurate sign and probability forecasts. Finally, we highlight that the forecasting performance of the estimated models is time-varying.

Keywords: Asymmetries, Forecast Evaluation, Gasoline, Crude Oil, Rockets and Feathers.

JEL Codes: C22, C32, C53, Q40, Q47.

Acknowledgments: Previous versions of this paper were presented at the "International Workshop on Oil Price Forecasts and Trends", Fondazione Eni Enrico Mattei, Milan, May 2013, and at the Second Annual Conference of the Italian Association of Environmental and Resource Economists, University of Milan, February 2014.

Corresponding author: Andrea Bastianin, DEMM - Dipartimento di Economia Management e Metodi Quantitativi, University of Milan,Via Conservatorio 7, I-20122 Milano, Italy.
E-mail: andrea.bastianin@unimi.it

Forecasting the Oil-Gasoline Price Relationship: Should We Care About the Rockets and the Feathers?

1. Introduction

Empirical evidence suggests that in many markets the adjustment process of an output price differs depending on the sign of the corresponding input price variations. For instance, Peltzman (2000) reports that output prices tend to respond faster to input increases than to decreases in 160 out of 242 markets.

This tendency, known as Asymmetric Price Transmission (APT), has been widely studied also by energy economists. According to the so-called Rockets and Feathers hypothesis (RFH), the transmission mechanism of positive and negative changes in the price of oil to the price of gasoline is asymmetric. Surveys of the APT literature are provided by Frey and Manera (2007) and Meyer and von Cramon-Taubadel (2004), while Geweke (2004) focuses on the RFH.

Although, starting from Bacon (1991), there have been many contributions addressing how downstream prices respond to increases in upstream prices (see, among others, Al-Gudhea, Kenc and Dibooglu, 2007; Balke, Brown and Yücel, 1998; Borenstein, Cameron, and Gilbert, 1997; Brown and Yücel, 2000; Douglas, 2010; Galeotti, Lanza and Manera, 2003; Godby, Lintner, Stengos, and Wandschneider, 2000; Grasso and Manera, 2007), little is known about the forecasting performance of reduced-form econometric models incorporating APT from crude oil to gasoline. As pointed out by Bachmeier and Griffin (2003), if gasoline prices respond asymmetrically to crude oil price variations, asymmetric cointegration models should produce more accurate forecasts than the symmetric Error Correction Model (ECM). These authors perform a small scale out-of-sample exercise, with the aim of comparing the forecasting accuracy of asymmetric and symmetric ECM for the wholesale price of gasoline.

Our work fills this gap. We focus on U.S. fuel markets and model the oil-gasoline price relation consistently with the RFH. Specifically, we compare point, sign and probability forecasts from a variety of Asymmetric-ECM (A-ECM) and Threshold Autoregressive ECM (TAR-ECM) against a standard ECM. Forecasts from A-ECM and TAR-ECM subsume the RFH, while the ECM implies symmetric price transmission from crude oil to gasoline prices. The aim of our paper is to quantify the forecast accuracy gains due to introducing the RFH in predictive models for the prices of gasoline and diesel. In particular, we provide answers the following research questions:

1. Is the RFH useful when forecasting gasoline price changes (point forecasts)?
2. Is the RFH useful when forecasting the sign of gasoline price movements (direction-of-change or sign forecasts)?

3. Is the RFH useful when forecasting the probability of gasoline price movements (probability forecasts)?
4. Are asymmetries constant through time or time-varying (time-varying forecast accuracy)?
5. At which sampling frequency (daily, weekly or monthly) are forecasts based on the RFH useful?
6. At which stage of the transmission mechanism (i.e. either spot or retail, or both) are the forecasts based on the RFH more accurate than the forecasts obtained from symmetric models?

Our answer to the first question is negative, while questions 2 and 3 have a positive answer. Asymmetries are useful for sign and probability forecasting, but they do not lead to more accurate point forecasts than the symmetric ECM specification. We also show that the forecasting performance of models changes through time: in some periods A-ECM produce more accurate forecasts than the ECM, while in other time periods the ECM dominates the asymmetric specifications. Empirical evidence also highlights that accuracy gains can be achieved mostly at daily or monthly sampling frequency for both spot and retail prices.

Our findings are of great value for a number of economic agents, whose activities involve decisions that are inherently forward-looking. For instance, gasoline producers need accurate point forecasts for hedging activities and portfolio allocation. On the other hand, policy makers exploit point and probability forecasts for stockpiling decisions (e.g. management of inventories and strategic reserves). Moreover, investors rely on direction-of-change forecasts to design technical trading rules and on probability forecasts for risk management (e.g. Value-at-Risk).

The plan of the paper is as follows. Section 2 describes the data. The empirical methods are introduced in Section 3. Section 4 describes the results and Section 5 concludes.

2. Data

Our analysis focuses on the U.S. fuel markets. We consider the relations between the spot price of West Texas Intermediate (WTI) light crude oil and the following petroleum products:

1. spot price of New York Harbour Conventional Gasoline (NY);
2. spot price of U.S. Gulf Coast Conventional Gasoline (GC);
3. spot price of Los Angeles Reformulated RBOB Regular Gasoline (LA);
4. retail price (excluding taxes) of U.S. Regular All Formulations Gasoline (G);
5. retail price (excluding taxes) U.S. No 2 Diesel (D).

Table 1. Data Description

(1)	(2)	(3)	(4)	(5)	(6)
Series	id	Price	Frequency	Sample period	No. Obs. (daily, weekly, monthly)
Cushing, OK WTI Crude Oil	WTI	Spot	Daily	02/06/1986 - 31/01/2013	6712, 1392, 320
New York Harbor Conventional Gasoline Regular	NY	Spot	Daily	02/06/1986 - 31/01/2013	6712, 1392, 320
U.S. Gulf Coast Conventional Gasoline Regular	GC	Spot	Daily	02/06/1986 - 31/01/2013	6712, 1392, 320
Los Angeles Reformulated RBOB Regular Gasoline	LA	Spot	Daily	01/04/2003 - 31/01/2013	2471, 514, 118
U.S. Regular All Formulations Gasoline	G	Retail	Weekly	20/08/1990 - 01/28/2013	-, 1113, 270
U.S. No 2 Diesel	D	Retail	Weekly	06/01/1997 - 01/28/2013	-, 796, 193

Notes: Columns (1) and (2) report a brief description of the series and the short-cut (id) used in the paper. Column (3) describes the type of price series. Column (4) illustrates the highest frequency at which the data are available. The sample period is shown in column (5), while the number of observations for daily, weekly and monthly data are reported in column (6). Retail prices excluding taxes have been calculated from prices including taxes retrieved from the EIA database, as detailed in the Appendix.

We have obtained all price series from the U.S. Energy Information Administration website. Crude oil and gasoline spot prices have been collected at daily sampling frequency, while retail gasoline and diesel prices are available only at weekly frequency.

The spot and retail prices of petroleum products do not include taxes and are denominated in dollars per gallon, while the spot price of oil is expressed in dollars per barrel.

Weekly and monthly spot prices have been calculated by averaging daily prices. Monthly retail prices have been computed by averaging data at weekly frequency. In all cases, in order to have synchronous prices, we preliminary dropped those observations for which it was not to possible to match gasoline or diesel prices with crude oil prices. A description of the dataset is presented in Table 1.[1]

3. Models and Methods

Let O_t be the spot price of WTI crude oil and let P_{kt} denote the price of the k-th petroleum product at time t, k = NY, GC, LA, G, D, $t = 1,\ldots,T$. We use the following notation: $p_{kt} \equiv 100\times\ln(P_{kt})$, $o_t \equiv 100\times\ln(O_t)$, $\Delta p_{kt} \equiv p_{kt} - p_{kt-1}$, and $\Delta o_t \equiv o_t - o_{t-1}$, with ln(.) indicating the natural logarithmic

[1] Retail prices excluding taxes are used in the analysis. To save space a more detailed description of the dataset, including construction of the price series and their plots, is presented in the Appendix.

transformation. From now on we drop the subscript k for ease of notation. Moreover, in this section we will use the generic expression "petroleum product" (PP) to indicate any of the petroleum products considered in the study.

Following previous research on the RFH, we assume that the price of crude oil (o), being oil the main production input, is the only driver of the PP price (p):

$$p_t = \omega_0 + \omega_1\, o_t + z_t \tag{1}$$

where z_t denotes the error term at time t. As highlighted by Bachemeier and Griffin (2003), equation (1) should not be given a structural interpretation. Actually, there are many other factors affecting the price of gasoline (e.g. inventory levels, refinery outages, changes in regulations, refining capacity utilization). If both the price of oil and the PP price are integrated of order one, while their liner combination is stationary, they are said to be co-integrated[2] (Engle and Granger, 1987), and the forecasts for the PP price should be produced with the following Error Correction Model (ECM):

$$\Delta p_t = \alpha + \sum_{i=0}^{p} \beta_i\, \Delta o_{t-i} + \sum_{j=1}^{q} \gamma_j\, \Delta p_{t-j} + \theta\, z_{t-1} + \varepsilon_t \tag{2}$$

where $z_{t-1} \equiv p_{t-1} - \omega_0 - \omega_1\, o_{t-1}$ represents the stationary linear combination (or long-run equilibrium relationship) between the PP price and the price of crude oil. Coefficients β_i and γ_j measure the short-run impact of (lagged) crude oil and PP prices on the current PP price, while θ describes the speed of adjustment to long-run equilibrium. Clearly, the ECM entails a symmetric adjustment process, in that the response of the PP price does not depend on the sign of the disequilibrium between the PP price and the price of crude oil.

A simple way to introduce an asymmetric adjustment mechanism in the ECM is to consider the Asymmetric ECM (A-ECM) of Granger and Lee (1989):

$$\begin{aligned} \Delta p_t = \quad & \alpha + \sum_{i=0}^{p} [\beta_i^{(+)}\, \Delta o_{t-i} \times \mathrm{I}(\Delta o_{t-i} > 0) + \beta_i^{(-)}\, \Delta o_{t-i} \times \mathrm{I}(\Delta o_{t-i} \le 0)] + \ldots \\ & \ldots + \sum_{j=1}^{q} [\gamma_j^{(-)}\, \Delta p_{t-j} \times \mathrm{I}(\Delta p_{t-j} > 0) + \gamma_j^{(-)}\, \Delta p_{t-j} \times \mathrm{I}(\Delta p_{t-j} \le 0)] + \ldots \\ & \ldots + \theta^{(+)} z_{t-1} \times \mathrm{I}(z_{t-1} > 0) + \theta^{(-)} z_{t-1} \times \mathrm{I}(z_{t-1} \le 0) + \varepsilon_t \end{aligned} \tag{3}$$

[2] Results available from the authors show that all price series are integrated of order one and that gasoline and diesel prices are co-integrated with the price of crude oil.

where the indicator function I(.) is used to decompose co-integration residuals and lagged differences of crude oil and gasoline prices into positive and negative values. Notice that the A-ECM reduces to the ECM when the following restrictions hold: $\beta_i^{(+)} = \beta_i^{(-)}$; $\gamma_j^{(+)} = \gamma_j^{(-)}$; $\theta^{(+)} = \theta^{(-)}$, for all i and j.

In order to disentangle the forecast gains deriving from short-run and long-run asymmetries, we also consider two variations of the previous model. An A-ECM with only short-run asymmetries (SR-A-ECM) is obtained by imposing $\theta^{(+)}=\theta^{(-)}$ in equation (3), while restrictions $\beta_i^{(+)} = \beta_i^{(-)}$ and $\gamma_j^{(+)} = \gamma_j^{(-)}$ yield an A-ECM with only long-run asymmetries (LR-A-ECM).

A simple and popular alternative for introducing asymmetries in the ECM is to consider a two-regimes Threshold Autoregressive ECM, TAR-ECM:

$$\Delta p_t = [\alpha^{(+)} + \sum_{i=0}^{p}\beta_i^{(+)}\Delta o_{t-i} + \sum_{j=1}^{q}\gamma_j^{(+)}\Delta p_{t-j} + \theta^{(+)}z_{t-1}] \times \mathrm{I}(q_t > 0) + \ldots$$
$$\ldots + [\alpha^{(-)} + \sum_{i=0}^{p}\beta_i^{(-)}\Delta o_{t-i} + \sum_{j=1}^{q}\gamma_j^{(-)}\Delta p_{t-j} + \theta^{(-)}z_{t-1}] \times \mathrm{I}(q_t \leq 0) + \varepsilon_t \qquad (4)$$

where q_t is a threshold variable. The TAR-ECM reduces to the ECM when $\alpha^{(+)} = \alpha^{(-)}$, $\beta_i^{(+)} = \beta_i^{(-)}$, $\gamma_j^{(+)} = \gamma_j^{(-)}$ and $\theta^{(+)} = \theta^{(-)}$, for all i and j. We consider two versions of the TAR-ECM: TAR1, for $q_t = \Delta o_{t-1}$; TAR2, for $q_t = s^{-1}\sum_{i=1}^{s}\Delta o_{t-i}$, where s = 5, 4, 3 for daily, weekly and monthly data. Previous works based on A-ECM and TAR-ECM include Al-Gudhea et al. (2007), Balke et al. (1998), Douglas (2010), Galeotti et al. (2003), Godby et al. (2000), Grasso and Manera (2007) and Fosten (2012).

All models have been estimated with OLS using a two-step procedure. First, we estimate the equilibrium relation (1) and obtain an estimate of z_t. Second, we estimate the ECM, A-ECM, SR-A-ECM, LR-A-ECM, TAR1 and TAR2 specifications and produce one step ahead forecasts. The optimal lag length (p, q) of each model has been selected by minimizing the Schwarz Information Criterion at each instant of time a new forecast is generated.

3.1 Forecast Evaluation

We denote with $f_{t+1} \equiv f^{(m)}_{kt+1|t}$ the one step ahead point forecast for the k-th oil product (k = NY, GC, LA, GR, DR), issued at time t using model m, m = ECM, A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2. For each model, a total of H forecasts have been obtained with a rolling window estimation scheme.[3] The evaluation of point forecasts relies on the mean squared forecast error (MSFE):

[3] More details are provided in the Appendix. An explanation of why forecasts are based on the rolling window scheme is provided in Footnote 5.

$$MSFE = H^{-1}\sum_{h=1}^{H}(\Delta p_{t+h} - f_{t+h})^2 \tag{5}$$

Direction-of-change (or sign) forecasts are evaluated by comparing the sign of the PP price forecast f_{t+1} with the realized sign of the PP price change Δp_{t+1}. Sign forecasts are particularly relevant for investors aiming to design market timing strategies. We employ two metrics of accuracy, namely the mean forecast trading return (MFTR) and the Success Ratio (SR):

$$MFTR = H^{-1}\sum_{h=1}^{H}\text{sign}(f_{t+h})\Delta p_{t+h} \tag{6}$$

$$\text{SR} = 100 \times [T^{(+,+)} + T^{(-,-)}]/\text{H} \tag{7}$$

where $\text{sign}(x) \equiv I(x > 0) - I(x < 0)$, while $T^{(+,+)}$ and $T^{(-,-)}$ denote the number of correctly predicted price increases and decreases.

The MFTR evaluates the average rate of return from a forecast (Hong and Lee, 2003), while the SR measures the percentage of forecasts that correctly predict the sign of price movements.

For each model and product, we form the corresponding probability forecast $pr_{t+1} = \Pr(\Delta p_{kt+1} < 0 | \Omega_t)$ as $F(-f_{t+1}/\sigma_{t+1})$, where $F(.)$ is the Normal cumulative density function[4] and σ_{t+1} is a (rolling window) volatility forecast obtained by fitting a GARCH(1,1) model to Δp_{kt}.

Probability forecasts are evaluated using the quadratic probability score (QPS):

$$QPS = H^{-1}\sum_{h=1}^{H}2[p_{t+h} - I(\Delta p_{t+h} < 0)\ ]^2 \tag{8}$$

The *QPS*, also known as the Brier score, ranges from 0 to 2, with 0 indicating perfect accuracy (see Diebold and Rudebusch, 1989).

For each of the previous performance measures (PM) we compute the percentage ratio, Δ(PM), as $100\times[(PM_U - PM_{ECM})/ PM_{ECM}]$, where PM = MSFE, MFTR, SR, QPS and subscript "U" denoting forecasts from A-ECM or TAR-ECM.

More accurate forecasts are associated to lower MSFE or QPS and higher MFTR or SR. Therefore, in the case of point and probability forecasts, the ECM is outperformed by an asymmetric model when Δ(PM)<0, while for sign forecasts this happens when Δ(PM)>0.

[4] Experiments with a Logit distribution confirm our results.

3.2 Testing the out-of-sample usefulness of asymmetries

The out-of-sample usefulness of the RFH for point and probability forecasts is investigated with the test put forth by Carriero and Giacomini (2011), CG henceforth.[5]

Since each of the asymmetric models considered in this paper reduces to the symmetric ECM if an appropriate set of parameter restrictions is imposed, we label the forecasts from A-ECM and TAR-ECM as "unrestricted forecasts", f_t^U. Similarly, forecasts from the ECM, f_t^{ECM}, are referred to as "restricted forecasts".

The CG statistic can be thought as an out-of-sample forecast combination test. Actually, it is possible to write the combination of the unrestricted/asymmetric and restricted/ECM forecasts, f_t^C, as:

$$f_t^C = f_t^{ECM} + (1\text{-}\lambda)(f_t^U - f_t^{ECM}) \tag{9}$$

where λ is the weight associated to the restricted ECM forecast. Therefore, asymmetric forecasts are useless if λ=1, that is: $f_t^C = f_t^{ECM}$. Conversely, if $\lambda = 0$, then $f_t^C = f_t^U$ or, equivalently, asymmetric forecasts are useful, with ECM receiving zero weight in the forecast combination.

The following null hypotheses are separately tested: H_0: $\lambda = 1$ and H_0: $\lambda = 0$. The RFH is useful out-of-sample when H_0: $\lambda = 1$ is rejected, while H_0: $\lambda = 0$ is not rejected.

The implementation of the CG test requires the estimation of the combining weight λ. For point forecasts, λ can be estimated with OLS, while for probability forecasts it can be estimated with non-linear least squares (see Kamstra and Kennedy, 1998 for details).

CG propose two versions of the test. Since the procedure discussed above is based on a single estimate of λ over the entire forecast evaluation sample, the CG test under the assumption that λ is constant through time can be thought as a test of "global usefulness" of the RFH. The second version of the CG test relies on time-varying estimates of the combining weight, denoted as λ_t, t = 1,...,H, and provides information about the "local usefulness" of the RFH. Within this alternative setting, if H_0: $\lambda_t = 1$ is rejected (for any t), while H_0: $\lambda_t = 0$ is never rejected, then the RFH is "locally" useful. Moreover, a plot of λ_t can be used to assess whether and how the usefulness of forecasts obtained from symmetric and asymmetric models has changed through time.

[5] The test cannot be applied to sign forecasts, because neither MFTR, nor SR comply with its underlying assumptions. Actually, in addition to some primitive conditions necessary for the law of large numbers and the central limit theorem to apply, the test is derived under the assumptions of convexity and differentiability of the loss function and using a fixed estimation sample. More precisely, the asymptotic distributions of the test is obtained by assuming $H \rightarrow \infty$, where H is the size of the forecast evaluation sample, while the size of the estimation sample and the forecast horizon must be finite. This also explains why forecasts have been obtained with a rolling window estimation scheme.

4. Results

In this section we offer a detailed discussion of the results for the spot price of NY gasoline, a synthetic presentation of the results for the other petroleum products, as well as a summary of the main findings for all the fuel markets analysed in the paper.

Table 2. Accuracy of point forecasts: the N.Y. Gasoline spot price

Model	MSFE	Δ(MSFE)%	λ	t(λ = 0)	t(λ = 1)
Panel (a): Daily data					
ECM	4.894	-	-	-	-
A-ECM	4.922	0.565	1.475	2.249**	0.724
SR-A-ECM	4.916	0.435	1.843	2.178**	0.996
LR-A-ECM	4.904	0.197	1.349	1.612	0.417
TAR1	4.980	1.748	1.455	3.186***	0.997
TAR2	5.004	2.244	1.421	2.910***	0.862
Avg. Asy.	4.926	0.653	-	-	-
Panel (b): Weekly data					
ECM	15.073	-	-	-	-
A-ECM	15.117	0.293	0.587	1.260	-0.885
SR-A-ECM	15.045	-0.185	0.436	0.813	-1.050
LR-A-ECM	15.158	0.568	1.624	2.003**	0.770
TAR1	15.435	2.399	1.002	2.905***	0.007
TAR2	15.548	3.151	0.990	3.190***	-0.032
Avg. Asy.	15.096	0.152	-	-	-
Panel (c): Monthly data					
ECM	35.175	-	-	-	-
A-ECM	33.714	-4.154	0.166	0.687	-3.443***
SR-A-ECM	33.622	-4.416	0.139	0.619	-3.839***
LR-A-ECM	35.636	1.311	1.589	1.457	0.540
TAR1	38.377	9.103	0.923	3.959***	-0.330
TAR2	44.897	27.639	0.887	9.376***	-1.195
Avg. Asy.	34.232	-2.681	-	-	-

Notes: Models are listed in column 1. Columns 2 and 3 show the Mean Squared Forecast Error (MSFE, see Eq. 5) and the percentage MSFE ratio defined as $\Delta(MSFE)\% = 100*[(MSFE_U-MSFE_{ECM})/MSFE_{ECM}]$, for U= A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2, Avg. Asy. "Avg. Asy." denotes the combined forecast from asymmetric/unrestricted models (i.e. A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2). A negative Δ(MSFE)% indicates that point forecasts from model U are on average more accurate than ECM forecasts. Results of the test of Carriero and Giacomini (2011) are shown in columns 4-5. The value of λ in column 4 is the estimated combination weight associated to the restricted/ECM forecast, see equation (9). The statistic t(λ=0)is used to test the null hypothesis that the restricted/ECM forecasts are not useful. The statistic t(λ=1)is used to test the null hypothesis that the unrestricted/asymmetric forecasts are useless. A non-rejection of H_0: λ = 0, coupled with a rejection of H_0: λ = 1, provides evidence that the unrestricted/asymmetric forecasts from model U, are useful. On the contrary, a rejection of H_0: λ = 0, coupled with a non-rejection H_0: λ = 1, provides evidence that asymmetric forecasts are useless. Asterisks *, **, *** denote rejection of the null hypothesis at the 10%, 5%, and 1% significance levels.

4.1 Forecasting the Spot Price of New York Harbour Conventional Gasoline

Results for daily, weekly and monthly point forecasts are reported in Panels a), b), and c) of Table 2, where MSFE and percentage MSFE ratios (Δ(MSFE)%), are shown in columns 2 and 3.

Since the exact nature of the asymmetry is unknown, we also calculate the sample average of all A-ECM and TAR-ECM predictions to form an equally weighted combination of the RFH forecasts (Avg. Asy.).

The MSFE ratios in Panel a) are always positive, thus for daily data asymmetric models are outperformed by the symmetric ECM. The usefulness of ECM forecasts compared to the RFH forecasts is confirmed by the estimates of the combining weight, λ (column 4 of Table 2). These estimates are always close to one, suggesting that a combined forecast should assign a unit weight to the ECM, while forecasts from the models incorporating the RFH should receive zero weight. Furthermore, the CG tests in column 5 and 6 of Table 2 lead to a rejection of $\lambda = 0$, coupled with a non-rejection of $\lambda = 1$ in 4 cases out of 5.

A look at Panel b) of Table 2 shows that in general the results for weekly point forecasts are quite similar. A notable exception is represented by the negative MSFE ratio associated with the SR-A-ECM specification, which is on average slightly more accurate than the ECM. Interestingly, the CG tests indicate that the RFH is useless only for TAR and LR-A-ECM. Although the tests are inconclusive in the remaining cases, for both A-ECM and SR-A-ECM the estimated optimal combining weight is close to 0.5, which supports the interpretation that the introduction of some form of asymmetric price adjustment in the ECM could result in more accurate point forecasts.

Panel c) of Table 2 shows that A-ECM and SR-A-ECM forecasts are the best option for monthly price data. If compared with the ECM, these specifications yield a 4% MSFE reduction (see column 3 of Table 2). The superior performance of these forecasts is confirmed by the output of the CG tests in the last three columns of Table 2. The estimates of λ are close to zero and statistically insignificant. Moreover, since the null hypothesis that λ is unity is rejected, an optimal combination of forecasts should assign zero weight to the symmetric ECM forecasts. On the contrary, results for the remaining models suggest that the RFH is useless for forecasting the NY spot price of gasoline.

A joint inspection of all panels of Table 2 shows that the least accurate models, as measured by the MSFE, are the TAR-ECM, while, among the asymmetric forecasts, the best choice is either the A-ECM or the SR-A-ECM. Lastly, combined forecasts are always associated to quite low MSFE.

The accuracy of direction-of-change or sign forecasts is analyzed in Table 3. On the whole, the specifications based on the RFH can be fruitfully used to improve sign forecasts.

Looking at Success Ratios (column 4 of Table 3), in most of the cases forecasts from asymmetric ECM specifications are to be preferred to symmetric ECM sign predictions. The last column of Table 3 shows that the increase in directional accuracy associated to asymmetric ECM ranges from 0.1% for daily data to 2.1% for monthly data. From a joint inspection of all directional accuracy metrics in Panel b) of Table 3, at weekly frequency the ECM outperforms most of the asymmetric models.

At daily horizon, the overall ranking of asymmetric models is broadly consistent with the results for point forecasts. Asymmetric ECM specifications are better than TAR-ECM, although in this case the former class of models outperforms the symmetric ECM.

The importance of asymmetries for sign forecasting appears to be particularly relevant at monthly horizon. Interestingly, the highest average rate of return, as measured by the MFTR reported in column 2 of Table 3, is associated to the TAR1 model and is equal to 7.5%.

Table 3. Accuracy of direction-of-change forecasts: the N.Y. Gasoline spot price

Panel (a): Daily data				
Model	MFTR	Δ(MFTR)%	SR	Δ(SR)
ECM	1.512	-	76.896	-
A-ECM	1.514	0.148	77.059	0.211
SR-A-ECM	1.517	0.308	77.004	0.141
LR-A-ECM	1.516	0.231	77.221	0.423
TAR1	1.488	-1.609	76.327	-0.740
TAR2	1.508	-0.250	76.490	-0.528
Avg. Asy.	1.512	-0.023	76.842	-0.070
Panel (b): Weekly data				
Model	MFTR	Δ(MFTR)%	SR	Δ(SR)
ECM	2.926	-	76.893	-
A-ECM	2.868	-1.989	76.240	-0.849
SR-A-ECM	2.910	-0.526	77.154	0.340
LR-A-ECM	2.886	-1.362	77.024	0.170
TAR1	2.901	-0.866	76.762	-0.170
TAR2	2.913	-0.437	76.762	-0.170
Avg. Asy.	2.924	-0.078	77.154	0.340
Panel (c): Monthly data				
Model	MFTR	Δ(MFTR)%	SR	Δ(SR)
ECM	7.376	-	81.818	-
A-ECM	7.487	1.511	83.523	2.083
SR-A-ECM	7.483	1.450	83.523	2.083
LR-A-ECM	7.455	1.071	82.386	0.694
TAR1	7.519	1.940	82.386	0.694
TAR2	6.914	-6.263	80.114	-2.083
Avg. Asy.	7.460	1.136	82.955	1.389

Notes: Models are listed in column 1. Columns 2 and 3 show the Mean Forecast Trading Return (MFTR, see Eq. 6) and the percentage MFTR ratio defined as $\Delta(MFTR)\% = 100*[(MFTR_U\text{-}MFTR_{ECM})/MFTR_{ECM}]$, U= A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2, Avg. Asy. "Avg. Asy." denotes the combined forecast from asymmetric/unrestricted models (i.e. A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2). Columns 2 and 3 show the Success Ratio (MFTR, see equation 7) and the percentage SR ratio defined as $\Delta(SR)\% = 100*[(SR_U\text{-}SR_{ECM})/SR_{ECM}]$. A positive Δ(MFTR)% or Δ(SR)% indicates that direction-of-change forecasts from model U are on average more accurate than ECM forecasts.

The highest SR, about 85% of correctly predicted signs, are obtained with A-ECM and SR-A-ECM. Probability forecasts are evaluated in Table 4. At all sampling frequencies, the most accurate probability forecasts are associated to A-ECM and SR-A-ECM specifications, which always lead to reductions of the QPS. For these models, the CG procedure leads to estimates of the combining weights, λ, that are not statistically different from zero, while they are always statistically different from unity.

Hence, irrespective of the sampling frequency, probability forecasts from the ECM are useless. Actually, an optimal combination of forecasts would attach full weight to either A-ECM or SR-A-ECM forecasts.

Conversely, when comparing TAR-ECM against ECM, the latter model is often preferred. Overall, there is evidence that A-ECM and SR-A-ECM yield the most accurate probability forecasts.

Table 4. Accuracy of Probability Forecasts: N.Y. Gasoline Prices

Model	QPS	Δ(QPS)%	λ	t(λ = 0)	t(λ = 1)
Panel (a): Daily data					
ECM	0.348	-	-	-	-
A-ECM	0.347	-0.167	-0.313	-0.587	-2.463**
SR-A-ECM	0.348	-0.086	-0.281	-0.394	-1.796*
LR-A-ECM	0.348	-0.061	-0.088	-0.114	-1.409
TAR1	0.349	0.219	0.876	2.656***	-0.377
TAR2	0.350	0.642	1.415	4.566***	1.339
Avg. Asy.	0.348	-0.015	-	-	-
Panel (b): Weekly data					
Model	QPS	Δ(QPS)%	λ	t(λ = 0)	t(λ = 1)
ECM	0.335	-	-	-	-
A-ECM	0.333	-0.505	-0.092	-0.159	-1.893*
SR-A-ECM	0.333	-0.626	-0.348	-0.564	-2.182**
LR-A-ECM	0.336	0.243	1.728	1.415	0.596
TAR1	0.338	0.929	1.085	2.442**	0.192
TAR2	0.333	-0.602	0.221	0.622	-2.191**
Avg. Asy.	0.333	-0.476	-	-	-
Panel (c): Monthly data					
Model	QPS	Δ(QPS)%	λ	t(λ = 0)	t(λ = 1)
ECM	0.256	-	-	-	-
A-ECM	0.247	-3.488	-1.026	-1.442	-2.847***
SR-A-ECM	0.246	-3.676	-1.217	-1.681*	-3.061***
LR-A-ECM	0.257	0.607	2.370	1.173	0.678
TAR1	0.259	1.193	0.806	1.382	-0.334
TAR2	0.291	14.035	1.421	3.302***	0.979
Avg. Asy.	0.254	-0.504	-	-	-

Notes: Models are listed in column 1. Columns 2 and 3 show the Quadratic Probability Score (QPS, see Eq. 8) and the percentage QPS ratio defined as Δ(QPS)% = 100*[(QPS_U-QPS_{ECM})/QPS_{ECM}], U= A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2, Avg. Asy. "Avg. Asy." denotes the combined forecast from asymmetric/unrestricted models (i.e. A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2). A negative Δ(QPS)% indicates that probability forecasts from model U are on average more accurate than ECM forecasts. Results of the test of Carriero and Giacomini (2011) are shown in columns 4-5. The value of λ in column 4 is the estimated combination weight associated to the restricted/ECM forecast, see equation (9). The statistic t(λ=0)is used to test the null hypothesis that the restricted/ECM forecasts are not useful. The statistic t(λ=1)is used to test the null that the unrestricted/asymmetric forecasts are useless. A non-rejection of H_0: λ = 0, coupled with a rejection of H_0: λ = 1, provides evidence that the unrestricted/asymmetric forecasts from model U are useful. On the contrary, a rejection of H_0: λ = 0, coupled with a non-rejection H_0: λ = 1, provides evidence that asymmetric forecasts are useless. Asterisks *, **, *** denote rejection of the null hypothesis at the 10%, 5%, and 1% significance levels.

The greater accuracy of some specifications based on the RFH is corroborated also by the result that combined forecast are always associated with lower QPS than the ECM.

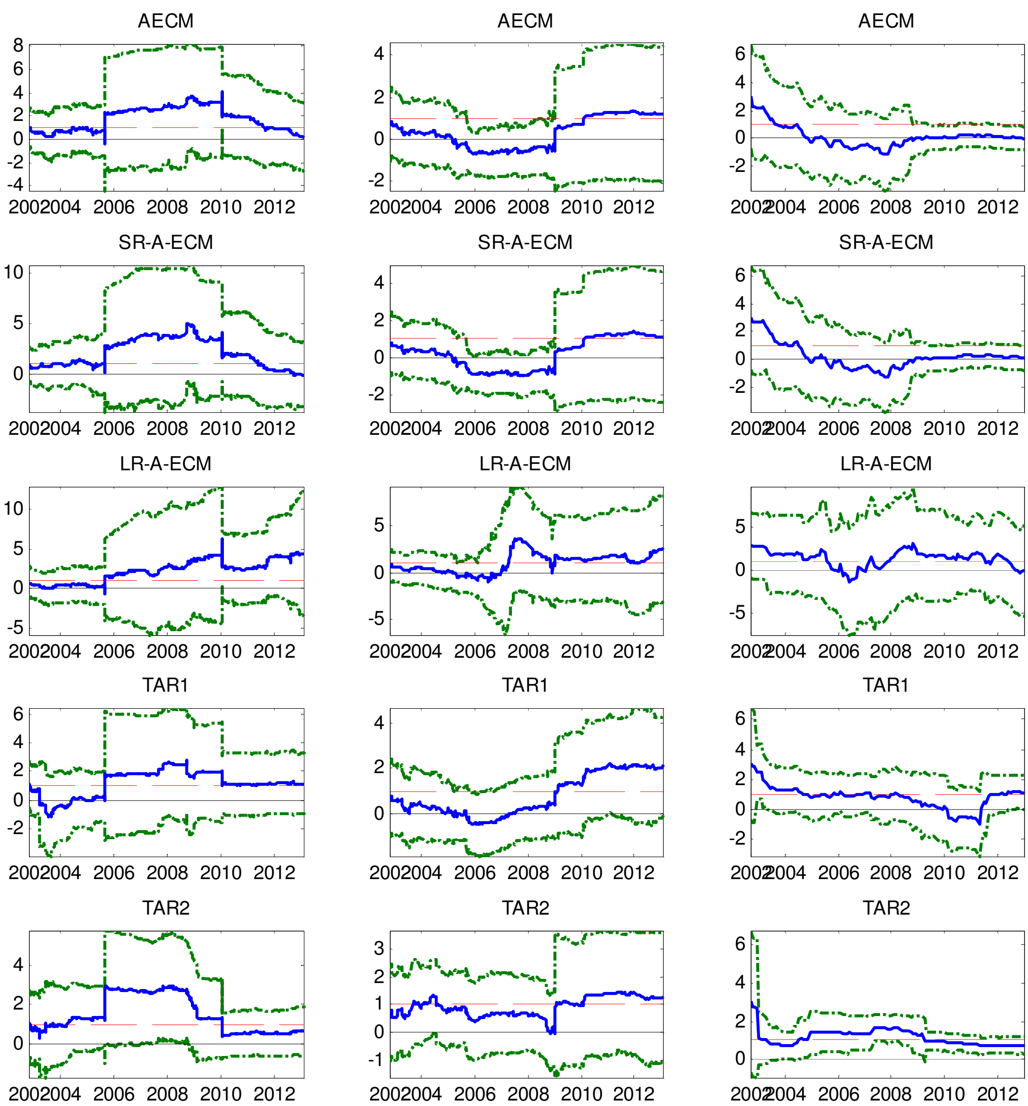
This suggests that some form of asymmetric price transmission from the price of crude oil to the price of gasoline should be incorporated when the aim is probability forecasting.

For both point and probability forecasts, the results of the CG tests rest on the implicit assumption that the weight associated to the symmetric ECM forecasts, λ, is constant through time.

However, given the turmoil that has characterized energy, financial markets and, more generally, the world economy in recent years, the assumption of time-invariant forecast accuracy could turn out to be very restrictive.

Figure 1. Local Usefulness of Asymmetries for Point Forecasts of the NY gasoline Spot Price

(a) daily data (b) weekly data (c) monthly data

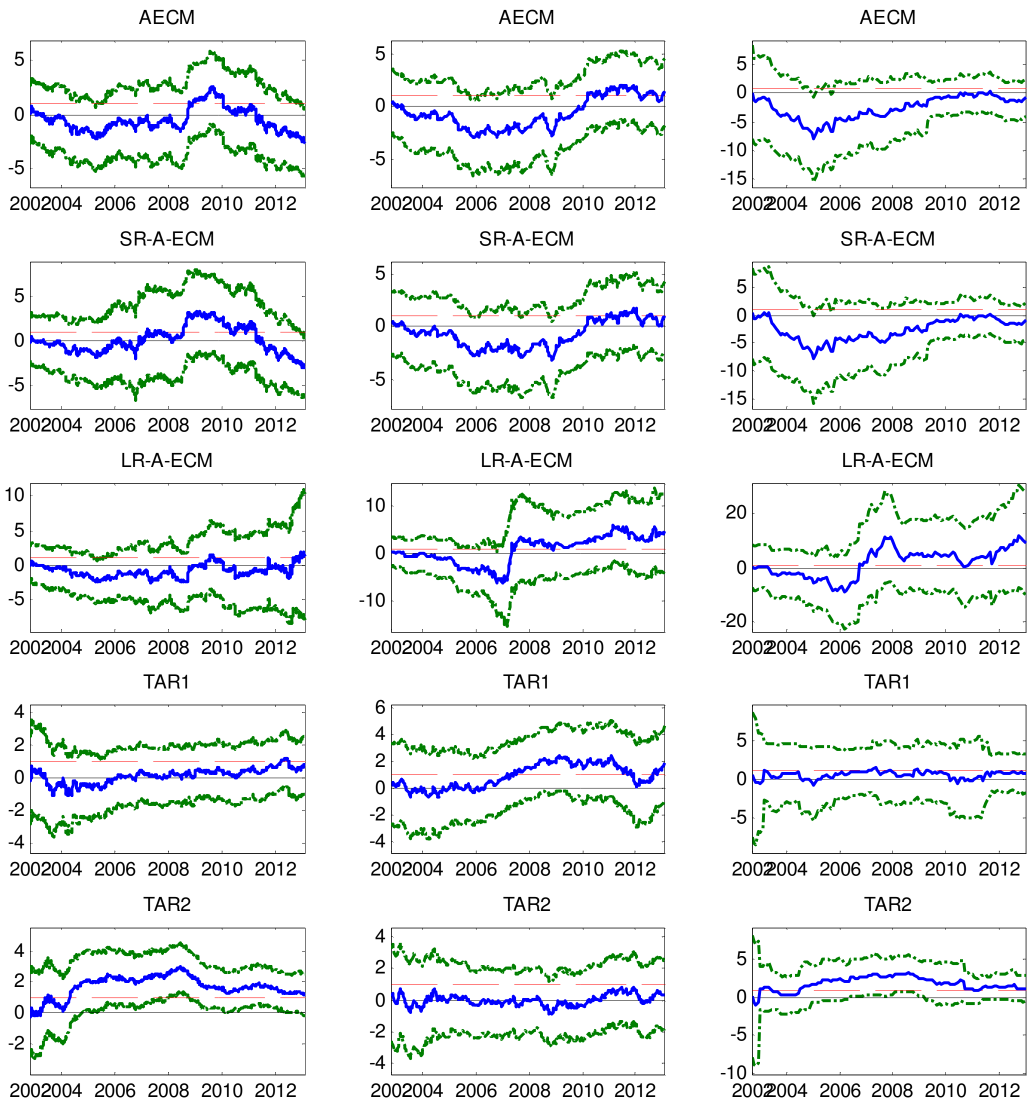


Notes: this figure shows the results of the CG test for the local usefulness of restrictions for point forecasts at daily (first column), weekly (second column) and monthly (third column) frequencies.. The restricted model is the Error Correction Model (ECM), while the model that labels the graph produces the unrestricted forecasts. The blue solid line is the estimated optimal weight. The red dashed and the black continuous horizontal lines are drawn in correspondence of $\lambda_t = 1$ and $\lambda_t = 0$, for all t, respectively. The green dotted lines represent the 95% confidence bands. The null hypothesis $\lambda_t = 0$ ($\lambda_t = 1$) for all t is rejected if the value 0 (the value 1) falls outside the 95% confidence bands for at least one t. The null hypothesis $\lambda_t = 1$ is used to test that the unrestricted forecast is useless. The null hypothesis $\lambda_t = 0$ is used to test that the restricted forecast is useless. Thus, rejection of $\lambda_t = 1$ and non-rejection $\lambda_t = 0$ imply that the asymmetric model is useful.

As already anticipated in the methodological section of this paper, this issue can be addressed with a test of local usefulness of forecasts.

Figure 2. Local Usefulness of Asymmetries for Probability Forecasts of the NY Gasoline Spot Price

(a) daily data (b) weekly data (c) monthly data



Notes: see Figure 1.

Results for point forecasts at daily, weekly and monthly horizons are shown in columns 1, 2 and 3 of Figure 1. A joint inspection of all graphs reveals that the combining weight associated to the ECM forecasts is not constant through time.

Focusing on daily forecasts (first column in Figure 1), in most cases at the beginning (i.e. 2002-late 2005) and at the end (i.e. 2010-2012) of the forecast evaluation sample, the estimates of λ range between 0 and 1, that is some form asymmetric price transmission might be helpful for point forecasting. However, from 2005 to 2009, the period that includes the oil price bubble that originated in March 2008 (Phillips and Yu, 2011), the estimated combining weight associated to the ECM model is much larger, ruling out the possibility to improve forecasts with asymmetric models. This result is confirmed by inspecting the 95% confidence intervals. In many cases, the null hypothesis that $\lambda_t = 0$ is rejected for at least one time period, while the null hypothesis that the ECM should receive a combining weight equal to unity is never rejected.

Looking at the graphs in the second column of Figure 1, we see that the estimates of λ_t are generally close to one starting from late 2008, meaning that after the burst of the oil price bubble weekly point forecasts cannot be improved by incorporating the RFH in the models. On the contrary, since the estimates of λ_t during the period 2002-early 2008 are quite close to zero, asymmetric models might have been more accurate than the ECM. In particular, in the time period spanning 2006 through 2009, the GC test leads to a rejection of the null hypothesis that $\lambda_t = 1$ for A-ECM and SR-A-ECM. Moreover, since the null hypothesis that ECM should receive a zero combination weight is never rejected for those specifications, we can conclude that during the oil price bubble the RFH would have led to more accurate point forecasts.

The tests of local usefulness for monthly data, shown in the last column of Figure 1, display a very similar pattern, suggesting that both A-ECM and SR-A-ECM have outperformed the ECM.

To sum up, results in Figure 1 highlights that the forecasting performance of the models is not constant through time and that A-ECM and SR-A-ECM might outperform the ECM at weekly and monthly horizon. Lastly, we notice that when the RFH is captured by means of LR-A-ECM and TAR-ECM specifications, the ECM always yields more accurate forecasts, irrespective of the sampling frequency of the data.

The local usefulness of asymmetric models for probability forecasting can be assessed by looking at the graphs reported in Figure 2. The estimates of λ_t are more often closer to zero than to one, that is asymmetries are useful for probability forecasting. Confidence intervals for the estimates of λ_t show that, irrespective of the sampling frequency of the data, there are several cases where the value zero always lies within the interval, while the value one is outside the interval. For A-ECM and SR-A-ECM point forecasts, this finding is confirmed in the majority of cases.

Our analysis on the NY gasoline spot price allows to draw some interesting conclusions. First, some asymmetric models yield more accurate probability and sign forecasts than the ECM. Second, there is evidence that the relative performance of the models changes through time.

Third, even if the improvements in forecast accuracy obtained by embedding the RFH are quite low, our results point out that at weekly and monthly horizons asymmetric ECM specifications outperform the standard ECM.

Table 5. Accuracy of Point, Direction-of-change and Probability Forecasts: Gulf Coast and Los Angeles Gasoline Spot Prices, Diesel and Gasoline Retail Prices

Panel (a): Point Forecasts

	Spot prices				Retail prices			
	Gulf Coast (GC)		Los Angeles (LA)		Regular (G)		Diesel (D)	
Frequency	Min MSFE	Median λ	Min MSFE	Median λ	Min MSFE	Median λ	Min MSFE	Median λ
Daily	ECM	1.548	Avg. Asy.	0.662	-	-	-	-
Weekly	LR-A-ECM	1.002	ECM	1.587	SR-A-ECM	0.901	SR-A-ECM	0.587
Monthly	Avg. Asy.	0.745	LR-A-ECM	0.989	ECM	1.365	Avg. Asy.	1.750

Panel (b): Sign Forecasts

	Spot prices				Retail prices			
	Gulf Coast (GC)		Los Angeles (LA)		Regular (G)		Diesel (D)	
Frequency	Max MFTR	Max SR	Max MFTR	Max SR	Max MFTR	Max SR	Max MFTR	Max SR
Daily	A-ECM	SR-A-ECM	Avg. Asy.	Avg. Asy.	-	-	-	-
Weekly	ECM	A-ECM	TAR2	LR-A-ECM	SR-A-ECM	Avg. Asy.	TAR2	A-ECM
Monthly	SR-A-ECM	SR-A-ECM	SR-A-ECM	SR-A-ECM	SR-A-ECM	TAR2	TAR1	Avg. Asy.

Panel (c): Probability Forecasts

	Spot prices				Retail prices			
	Gulf Coast (GC)		Los Angeles (LA)		Regular (G)		Diesel (D)	
Frequency	Min QPS	Median λ	Min QPS	Median λ	Min QPS	Median λ	Min QPS	Median λ
Daily	A-ECM	0.015	Avg. Asy.	0.682	-	-	-	-
Weekly	SR-A-ECM	0.552	ECM	1.201	SR-A-ECM	0.796	A-ECM	0.648
Monthly	SR-A-ECM	0.944	A-ECM	-0.195	TAR2	1.141	TAR1	0.483

Notes: in Panels a), b) and c), entries headed "min MSFE", "max MFTR", "max SR" and "min QPS" indicate the most accurate forecasting model. In Panels a) and c), "Median λ" is the median of the estimates of λ from the CG test. If λ=0, an optimal combination of forecasts should assign unit weight to one of the asymmetric models (i.e. A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2) and zero weight to the ECM; vice versa, when λ=1 the optimal combined forecast coincides with the symmetric ECM. "Avg. Asy." denotes a combined forecasts obtained as the simple average of the forecasts from asymmetric models.

Fourth, in most cases the type of asymmetric adjustment captured by the TAR-ECM is useless for forecasting. Fifth, A-ECM and SR-A-ECM appear to be the best alternatives to incorporate the RFH in forecasting models.

4.2 Forecasting the Gulf Coast and Los Angeles Gasoline Spot Prices, and the Diesel and Gasoline Retail Prices

The evaluation of forecasts of the Gulf Coast and Los Angeles spot prices, and of the retail prices of regular gasoline and diesel is reported in Table 5.

Panel a) of Table 5 shows that A-ECM specifications are often associated to the minimum MSFE, suggesting that the RFH might lead to more accurate point forecasts. However, when taking into account the value of the combining weight, λ, it emerges that, irrespective of the type of price, whether spot or retail, the median of the estimates of λ is close to one in most cases. This means that the ECM specification actually receives full weight in a forecast combination and the RFH does not improve the forecast accuracy of the estimated models. Although these results are not supportive of the usefulness of asymmetries, they can be helpful to draw some indications on the relative validity of the different asymmetric specifications. For instance, in coherence with the results described in the previous section, TAR-ECM models are never selected as the best option.

Notice that none of the asymmetric ECM clearly dominates, since the relative accuracy of A-ECM, SR-A-ECM and LR-A-ECM depends on the series under analysis and on the sampling frequency of the data. Moreover, the combination of forecasts from the asymmetric models often leads to the highest MSFE reduction, thus reinforcing our conclusion that the empirical evidence of superior forecasting performance due to the RFH is scarce and that none of the asymmetric specifications taken into account can fully describe the price transmission mechanism from crude oil to petroleum products.

In the case of sign forecasts results are quite different. From Panel b) of Table 5 we see that the RFH yields more accurate direction-of-change forecasts in 19 comparisons out of 20. The asymmetric ECM specifications are more often associated to higher MFTR or SR than the TAR-ECM models. Consistently with the results for the NY gasoline spot price, the SR-A-ECM appears to be a good choice also for the Gulf Coast and Los Angeles gasoline spot prices. The adequacy of the SR-A-ECM indicates that only short-run asymmetries matter for sign forecasting.

Panel c) of Table 5 confirms these results when probability forecasts are considered. Moreover, as opposed to what has been observed for point forecasts, the median value of the combining weight, λ, is more often closer to zero than to one, that is the RFH can be exploited in most cases to produce more accurate probability forecasts.

Notice that for probability forecasts the ECM is selected as the most accurate specification only once, while in 5 comparisons out of 10 either A-ECM or SR-A-ECM minimize the QPS.

4.3. Summary of Results

Table 6 provides a summary of the main results of the paper. Entries of this table represent the number and percentage of comparisons according to which forecasts embedding the RFH outperform the standard ECM.

Panel a) of Table 6 shows that accuracy gains due to the RFH are rare when point forecasts are considered: asymmetric models yield more accurate predictions than the ECM only 18.5% and 20.8% of the cases for spot and retail prices, respectively. The presence of accuracy gains due to the RFH is influenced by the sampling frequency of the data. The largest number of cases in which asymmetric models outperform the ECM is recorded at monthly sampling frequency for spot prices, while it is at weekly sampling frequency for retail prices.

Table 6. Summary of Results

Panel (a): Point Forecasts (MSFE reductions due to the RFH)

	Spot						Retail									
	New York		Gulf Coast		Los Angeles		Regular		Diesel		Spot		Retail		Spot & Retail	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Daily	0 / 6	0.0	0 / 6	0.0	2 / 6	33.3	- / -	-	- / -	-	2 / 18	11.1	- / -	-	- / -	-
Weekly	1 / 6	16.7	1 / 6	16.7	0 / 6	0.0	2 / 6	33.3	2 / 6	33.3	2 / 18	11.1	4 / 12	33.3	6 / 30	20.0
Monthly	3 / 6	50.0	2 / 6	33.3	1 / 6	16.7	0 / 6	0.0	1 / 6	16.7	6 / 18	33.3	1 / 12	8.3	7 / 30	23.3
Total	4 / 18	22.2	3 / 18	16.7	3 / 18	16.7	2 / 12	16.7	3 / 12	25.0	10 / 54	18.5	5 / 24	20.8	13 / 78	16.7

Panel (b): Directional accuracy (SR increases due to the RFH)

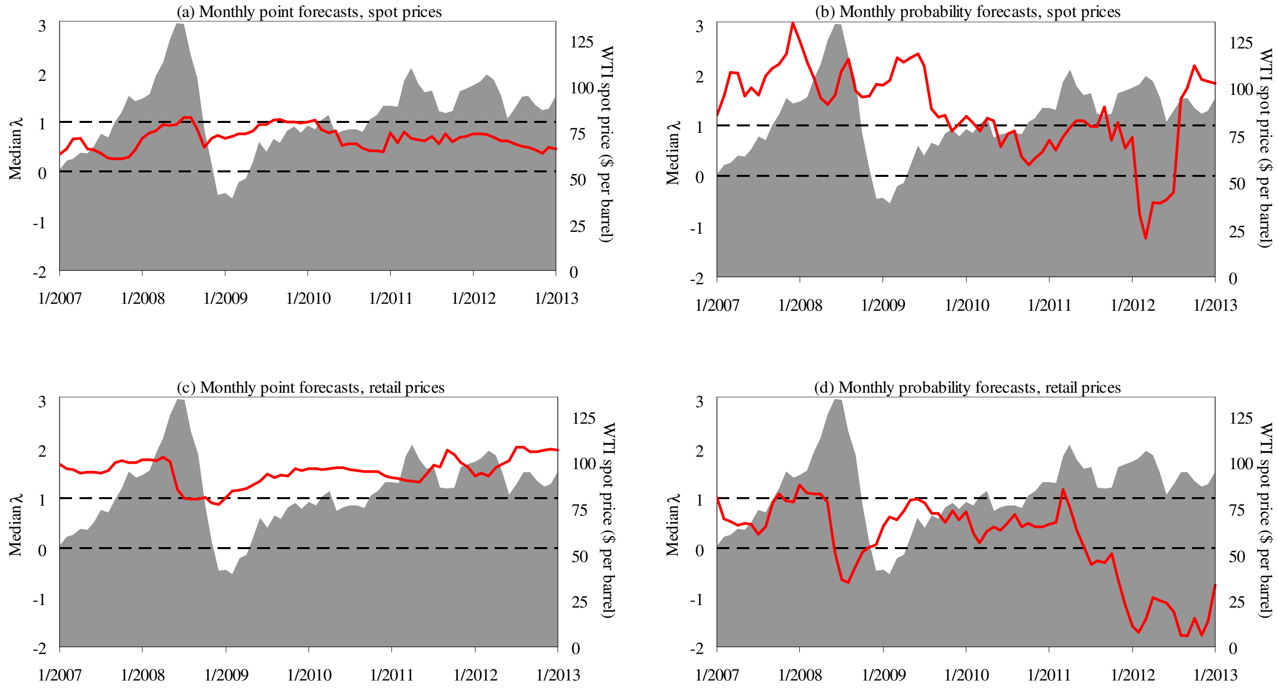
	Spot						Retail									
	New York		Gulf Coast		Los Angeles		Regular		Diesel		Spot		Retail		Spot & Retail	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Daily	3 / 6	50.0	5 / 6	83.3	5 / 6	83.3	- / -	-	- / -	-	13 / 18	72.2	- / -	-	- / -	-
Weekly	3 / 6	50.0	3 / 6	50.0	3 / 6	50.0	3 / 6	50.0	6 / 6	100.0	9 / 18	50.0	9 / 12	75.0	18 / 30	60.0
Monthly	5 / 6	83.3	3 / 6	50.0	5 / 6	83.3	5 / 6	83.3	4 / 6	66.7	13 / 18	72.2	9 / 12	75.0	22 / 30	73.3
Total	11 / 18	61.1	11 / 18	61.1	13 / 18	72.2	8 / 12	66.7	10 / 12	83.3	35 / 54	64.8	18 / 24	75.0	40 / 78	51.3

Panel (c): Probability Forecasts (QPS reductions due to the RFH)

	Spot						Retail									
	New York		Gulf Coast		Los Angeles		Regular		Diesel		Spot		Retail		Spot & Retail	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Daily	4 / 6	66.7	4 / 6	66.7	1 / 6	16.7	- / -	-	- / -	-	9 / 18	50.0	- / -	-	- / -	-
Weekly	4 / 6	66.7	3 / 6	50.0	0 / 6	0.0	2 / 6	33.3	2 / 6	33.3	7 / 18	38.9	4 / 12	33.3	11 / 30	36.7
Monthly	3 / 6	50.0	3 / 6	50.0	4 / 6	66.7	3 / 6	50.0	3 / 6	50.0	10 / 18	55.6	6 / 12	50.0	16 / 30	53.3
Total	11 / 18	61.1	10 / 18	55.6	5 / 18	27.8	5 / 12	41.7	5 / 12	41.7	26 / 54	48.1	10 / 24	41.7	27 / 78	34.6

Notes: entries in this table represent the number and percentage of comparisons for which models incorporating the RFH lead to more accurate point (Panel a), sign (Panel b) and probability forecasts (Panel c). Columns headed "#" can be read as follows: "no. of cases / no. of comparisons", where "no. of cases" corresponds to the number of comparisons in which asymmetric models outperform the ECM. Rows headed "Total" sum over the "#" columns. Percentages are based on those figures. For each series and sampling frequency the figures in the table represent the forecasting comparison of the ECM against the following asymmetric models: A-ECM, SR-A-ECM, LR-A-ECM, TAR1, TAR2, Avg. Asy..

Figure 3. Local Usefulness of Asymmetries for Point and Probability of Spot and Retail Prices



Notes: this figure summarizes the CG test for the local usefulness of restrictions for point forecasts (first column) and probability forecasts (second column) at monthly frequency, and for spot prices (first row) and retail prices (second row). The grey area is the spot price of WTI crude oil, the red line is the median of the estimates of the optimal combining weight λ_t (i.e. the median of the estimates of λ_t from the comparison of all asymmetric models against the ECM, for both spot and retail prices), the black dashed lines are drawn in correspondence of $\lambda_t = 1$ and $\lambda_t = 0$. When $\lambda_t = 1$, forecasts from asymmetric models are useless. Conversely, if $\lambda_t = 0$, ECM forecasts is useless. Median estimates of λ_t lower than one and larger than zero indicate that combining asymmetric and ECM forecasts leads to more accurate predictions.

These results are confirmed by the graphs reported in Panels a) and c) of Figure 3, which illustrate the dynamics of the median value of the estimates of the combining weight, λ, for point forecasts of spot and retail prices. Some interesting considerations can be made. First, Panel a) shows that the median estimate of λ for spot prices is between zero and one most of the time.

This evidence implies that forecast combination might lead to more accurate predictions. Second, after the collapse of WTI price in 2008, the median estimate of λ is close to one, in which is equivalent to say that in this period the RFH was useless. Therefore, the ability of asymmetric models to improve the accuracy of monthly point forecasts of the spot price of gasoline is time-varying, and it is higher when the level of oil price volatility is low. Third, the median estimate of λ for retail prices reported in Panel c) is always close to one, meaning that the RFH is almost useless.

Results for sign forecasts are completely different. Panel b) of Table 6 illustrates that asymmetric models lead to more accurate forecasts than the ECM in most of the comparisons. As for spot prices, this happens 64% of the cases, and more often for monthly and daily forecasts than for weekly forecasts. When retail prices are considered, this percentage increases to 75%, and does not change with the sampling frequency of the data. Lastly, we notice that the directional accuracy is higher for the price of diesel than for the price of regular gasoline.

The results for probability forecasts are similar to the findings obtained for sign forecasts (Panel c of Table 6). However, the number of cases favourable to asymmetric models is slightly lower. Accuracy gains due to the RFH are more frequent for daily and monthly data. Moreover, Panels b) and d) of Figure 3 illustrate that the relative accuracy of models changes through time.

5. Conclusions

Should we care about the Rockets and the Feathers when forecasting the price of petroleum products? While we believe that more work in this area is necessary to confirm our findings, our concise answer to this question is mixed: "yes" for sign and probability forecasts, "no" for point forecasts. Our results have also highlighted that the forecasting performance of the estimated models is time-varying and depends on the sampling frequency of the data.

Consistently with Bachmeier and Griffin (2003), we have shown that models based on the RFH have limited value if the aim is to produce accurate point forecasts. More precisely, according to our results, A-ECM forecasts are at most as accurate as the benchmark forecasts obtained by a standard ECM, while TAR-ECM have always been ranked last in terms of forecast accuracy.

However, we have documented that direction-of-change forecasts from asymmetric ECM often outperform those from the benchmark ECM. These results hold for spot price data at daily and monthly sampling frequencies and for retail price series. As shown by Leitch and Tanner (1991), directional accuracy is highly correlated with the profits that economic agents can make by relying on a given model. Therefore, the accuracy metrics used in this paper, namely MFTR and SR, can be interpreted as economic measures of performance. On the basis of our findings, investors should rely on direction-of-change forecasts from A-ECM specification in order to design profitable market-timing trading strategies.

In the case of probability forecasts, asymmetric ECM perform significantly better than the ECM, at all sampling frequencies for retail data, while at daily and monthly frequencies for spot data. This result suggests that the design of forecast scenarios for gasoline and diesel prices might be improved if the RFH is taken into account.

Moreover, the Carriero-Giacomini test shows that the relative forecasting performance of symmetric and asymmetric price transmission models is not constant through time at all sampling frequencies and for all price series. This implies that the usefulness of models has to be evaluated conditionally on the state of the crude oil market, and that the forecasting performance of each specification changes through the business cycle.

The result that TAR-ECM specifications are often outperformed by asymmetric ECM reveals that the way the RFH is described is important for model building. This finding is reinforced by the result that forecast combination from different models often leads to forecast accuracy gains. Lastly, we have shown that the RFH is more useful at daily and monthly sampling frequency than for weekly data.

References

Al-Gudhea, S., Kenc, T., Dibooglu, S., 2007. Do retail gasoline prices rise more readily than they fall? A threshold cointegration approach. Journal of Economics and Business, 59, 560-574.

Bachmeier, L.J., Griffin, J.M., 2003. New evidence on asymmetric gasoline price responses. Review of Economics and Statistics, 85, 772-776.

Bacon, R.W., 1991. Rockets and feathers: the asymmetric speed of adjustment of UK retail gasoline prices to cost changes. Energy Economics, 13, 211-218.

Balke, N.S., Brown, S.P.A., Yücel, M., 1998. Crude oil and gasoline prices: an asymmetric relationship? Economic and Financial Policy Review, Q1, 2-11.

Borenstein, S., Cameron, A.C., Gilbert, R., 1997. Do gasoline prices respond asymmetrically to crude oil price changes? Quarterly Journal of Economics, 112, 305-339.

Brown, S.P.A., Yücel, M., 2000. Gasoline and crude oil prices: why the asymmetry? Economic and Financial Policy Review, Q3, 23-29.

Carriero, A., Giacomini, R., 2011. How useful are no-arbitrage restrictions for forecasting the term structure of interest rates? Journal of Econometrics, 164, 21-34.

Diebold, F.X., Rudebusch, G.D., 1989. Scoring the leading indicators. Journal of Business, 65, 369-391.

Douglas, C., 2010. Do gasoline prices exhibit asymmetry? Not usually! Energy Economics, 32(4), 918-925.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica, 55, 251-276.

Frey, G., Manera M., 2007. Econometric models of asymmetric price transmission. Journal of Economic Surveys, 21, 349-415.

Fosten, J., 2012. Rising household diesel consumption in the United States: a cause for concern? Evidence on asymmetric pricing. Energy Economics, 34, 1514-1522.

Galeotti, M., Lanza A., Manera, M., 2003. Rockets and feathers revisited: an international comparison on European gasoline markets. Energy Economics, 25, 175-190.

Godby, R., Lintner, A.M., Stengos, T., Wandschneider, B., 2000. Testing for asymmetric pricing in the Canadian retail gasoline market. Energy Economics, 22, pages 349-368.

Granger, C.W.J., Lee, T.H., 1989. Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models. Journal of Applied Econometrics, 4, S145-S159.

Grasso, M., Manera, M., 2007. Asymmetric error correction models for the oil-gasoline price relationship. Energy Policy, 35, 156-177.

Geweke, J., 2004. Issues in the rockets and feathers gasoline price literature. Report to Federal Trade Commission. Available on-line at (last accessed: 04/03/2013): http://www.ftc.gov/bc/gasconf/comments2/gewecke2.pdf

Hong, Y., Lee, T.H., 2003. Inference on predictability of foreign exchange rates via generalized spectrum and nonlinear time series models. Review of Economics and Statistics, 85(4), 1048-1062.

Kamstra, M., Kennedy, P., 1998. Combining qualitative forecasts using logit. International Journal of Forecasting, 14, 83-93.

Leitch, G., Tanner, J.E., 1991. Economic forecast evaluation: profits versus the conventional error measures. American Economic Review, 81, 580-590.

Meyer, J., von Cramon-Taubadel, S., 2004. Asymmetric price transmission: a survey. Journal of Agricultural Economics, 55, 581-611.

Peltzman, S., 2000. Prices rise faster than they fall. Journal of Political Economy, 108, 466-502.

Phillips, P.C.B., Yu, J., 2011. Dating the timeline of financial bubbles during the subprime crisis. Quantitative Economics 2, 455-491.