Inside the Crystal Ball: New Approaches to Predicting the Gasoline Price at the Pump

Christiane Baumeister
Notre Dame

Lutz Kilian
University of Michigan
CEPR

Thomas K. Lee
U.S. EIA

EIA Workshop on Financial and Physical Oil market Linkages

September 29, 2015 / Washington, D.C.

By

Thomas K. Lee

Office of Energy Markets and Financial Analysis

Energy Information Administration

What Are the Existing Gasoline Price Forecasts?

- The U.S. Energy Information Administration (EIA) issues regular forecasts of the retail price of gasoline that are closely monitored and widely discussed by the media.
- The American Automobile Association (AAA) releases its own gas price predictions at irregular intervals.
- There are websites solely devoted to providing daily forecasts of the local price of gasoline for consumers (e.g., www.gaspredictor.com).
- The Michigan Survey of Consumers regularly inquires about consumers' expectations about future retail gasoline prices at selected horizons (see Anderson, Kellogg, and Sallee 2013).

The Academic Literature

- Perhaps surprisingly, the accuracy of retail gasoline price forecasts has not received much attention by academic researchers.
- One recent exception is Anderson, Kellogg, Sallee and Curtin (2011). This study investigates the ability of U.S. consumers to forecast the price of gasoline, as measured by responses in the Michigan Survey of Consumers.
- Another is an evaluation of the accuracy of EIA gasoline price forecasts by Sanders, Manfredo and Boris (2009).
- Neither study, however, addresses the question of how to construct gasoline price forecasts.

Why Has This Question Not Been Investigated?

- There is a perception that forecasting gasoline prices beyond a few days is next to impossible, given publicly available information, and that, for all practical purposes we can think of the current gasoline price as the best predictor of the future price.
- Until recently, this no-change forecast (or random walk forecast) was also considered the best possible forecast of the price of crude oil, but a rapidly expanding literature has overturned this consensus. It is well established now that one can forecast the price of crude oil in real time more accurately than the no-change forecast.

Why the Ability to Forecast Oil Prices Need Not Extend to Gasoline Prices?

- Because gasoline prices are closely tied to the evolution of the price of crude oil with both prices moving together in the long-run.
  - Given it is a product obtained from refining crude oil
- In the short-run, retail gasoline prices also respond to:
  - Changes in gasoline taxes and environmental regulations
  - Refinery shutdowns due to routine maintenance, accidents and hurricanes
  - Important changes in the structure of the refining market in recent years (Borenstein and Kellogg 2014; Kilian 2014)
  - Changes in the market power of refiners (Borenstrin and Shepard 2002; Sweeney 2015)

Objective of this Paper

- This paper provides a comprehensive analysis of the forecastability of the real U.S. price of gasoline, drawing on state-of-the-art regression-based forecasting methods.
- Our objective is to provide a benchmark for future studies and to document the merits of alternative forecasting models.
- The analysis is conducted in real-time, taking account of the delays in the availability of some data and subsequent revisions when data become available.
- We focus on forecasting the average monthly U.S. retail price of gasoline, as defined by the EIA, at horizons up to 24 months.

Key Questions

1. How accurate are retail gasoline price forecasts based only on their own past?
    - AR, ARMA, Exponential smoothing, UC-SV forecasts
2. What is the predictive content of financial market predictors?
    - What gasoline futures prices and spread of the gasoline spot price relative to the spot price of crude oil
3. What is the predictive power of vector autoregressive (VAR) models including both retail gasoline prices and crude oil prices? Such as the role of:
    - The lag structure;
    - The definition of the price of crude oil; and
    - Imposing cointegrationg restrictions.

Key Questions

4. What is the predictive content of models of the gasoline market linking gasoline prices to changes in U.S. gasoline consumption and U.S. real economic activity.
    - Single-equation factor forecasting models
    - Factor augmented VAR (FAVAR) forecasting models
    - VAR for real gasoline consumption and real gasoline price
5. We explore larger-scale VAR models that jointly model the global market for crude oil and the U.S. gasoline market.

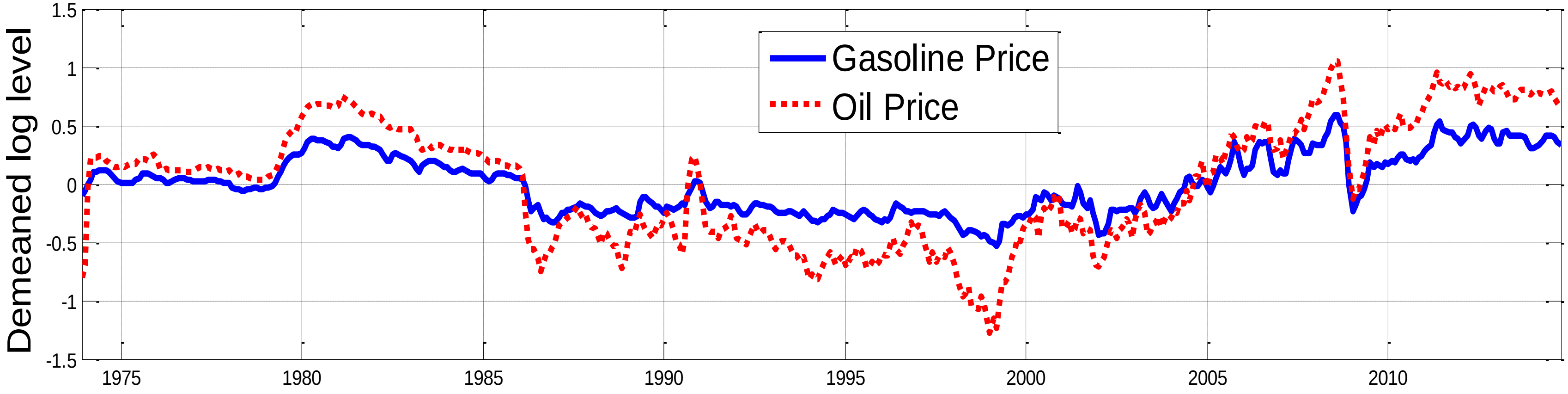
Main Findings

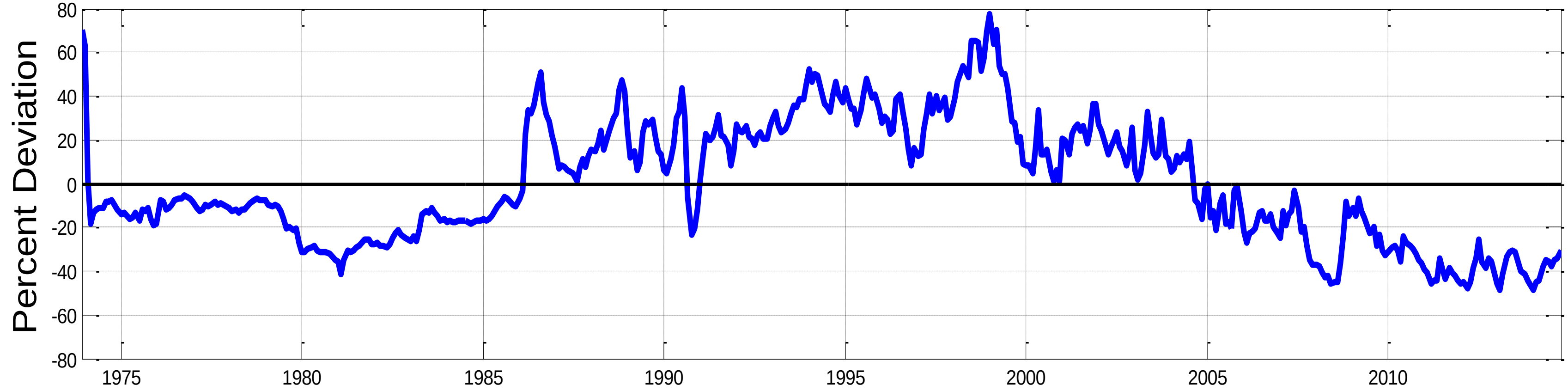
1. Substantial reductions in the mean-squared prediction error (MSPE) of gasoline price forecasts are feasible in real time at horizons up to two years and substantial increases in directional accuracy.

2. The most accurate individual model is a simple bivariate VAR(1) model. This model generates MSPE reductions at all horizons ranging from 10% to 26% and at some horizons has statistically significant directional accuracy as high as 68%.

3. Even more reliable overall is a pooled forecast that assigns equal weight to five of the most successful individual forecasting models. The accuracy of this pooled forecast also is more stable over time than that of the VAR(1) model.

Real U.S. Retail Price of Gasoline and Real Brent Price during 1973.10-2014.9





Q#1: Forecasting the Price based on its own Past

1. Autoregressive and autoregressive-moving average models
   - ARMA(1,1), IMA(1), ARIMA(1,1), AR(12), BAR(12).
   - AR(AIC), BAR(AIC).

2. Exponential smoothing forecasts
   - Forecasts formed as $\hat{r}^{gas}_{T+h|T} = \bar{r}^{gas}_{T}$ and converted to levels by exponentiating.
   - It is designed for series that are not trending over time.

3. Unobserved components-stochastic volatility (UC-SV) model
   - The model treats the log of price as the sum of a permanent component and a serially uncorrelated transitory component.

Real-Time Forecast Accuracy of Univariate Models Evaluated on 1992.1-2014.3

Monthly Horizon	ARMA(1,1)	IMA(1)	ARIMA(1,1)	AR(12)	BAR(12)	AR(AIC)	BAR(AIC)	Exponential Smoothing	UC-SV Model
	(a) MSPE Ratios								
1	1.204	1.202	1.283	1.274	1.226	1.263	1.238	2.704	1.045
3	1.044	1.077	1.159	1.022	1.009	1.037	1.028	1.074	0.987
6	1.032	1.094	1.153	0.936	0.931	0.955	0.950	0.794**	0.950
9	1.077	1.142	1.210	1.022	1.011	1.034	1.025	0.795**	0.921
12	1.088	1.110	1.167	1.155	1.131	1.149	1.129	0.922	0.917
15	1.075	1.099	1.149	1.127	1.103	1.115	1.098	0.883	0.910
18	1.065	1.112	1.158	1.063	1.049	1.051	1.042	0.818*	0.889
21	1.112	1.128	1.179	1.111	1.098	1.095	1.087	0.839*	0.838
24	1.151	1.111	1.161	1.172	1.154	1.154	1.141	0.919	0.776
	(a) Success Ratios								
1	0.644*	0.640*	0.640*	0.659*	0.663*	0.659*	0.655*	0.476	0.375
3	0.525	0.536	0.536	0.620*	0.611*	0.623*	0.623*	0.536	0.509
6	0.538*	0.508	0.508	0.611*	0.607*	0.607*	0.611*	0.519	0.534
9	0.502	0.498	0.494	0.537*	0.533*	0.517**	0.514**	0.494	0.587*
12	0.512**	0.520	0.520	0.488	0.484	0.481	0.461	0.449	0.609
15	0.514	0.510	0.506	0.534*	0.530*	0.526**	0.526**	0.486	0.577
18	0.524	0.496	0.496	0.528	0.536**	0.524	0.528	0.536	0.576**
21	0.490	0.486	0.486	0.490	0.498	0.514	0.518	0.518	0.587
24	0.492	0.508	0.504	0.467	0.492	0.504	0.512	0.488	0.562

Q#2: Forecasting Based on Financial Data

1. Forecasts based on gasoline futures prices

$$\hat{R}^{gas}_{t+h|t} = R^{gas}_{t}\left(1 + f^{h}_{t} - s_{t} - E_{t}(\pi^{h}_{t+h})\right),$$

where the inflation expectation is estimated as in Faust and Wright (Hdbk 2013).

2. Product spread regressions

$$\Delta s^{h,gas}_{t+h|t} = \alpha + \beta\left[s^{gas}_{t} - s^{oil}_{t}\right] + \varepsilon_{t+h}$$

$$\hat{R}^{gas}_{t+h|t} = R^{gas}_{t}\exp\left\{\hat{\alpha} + \hat{\beta}\left[s^{gas}_{t} - s^{oil}_{t}\right] - E_{t}(\pi^{h}_{t+h})\right\},$$

Forecast Accuracy of Models Based on Spot and Futures Market Prices Evaluated on 1992.1-2014.3

Monthly Horizon	Gasoline Futures	Spot Spread Model				TVP Spot Spread Model			
		WTI		Brent		WTI		Brent	
		$\hat{\alpha}_t, \hat{\beta}_t$	$\alpha_t = 0, \hat{\beta}_t$	$\hat{\alpha}_t, \hat{\beta}_t$	$\alpha = 0, \hat{\beta}$	$\hat{\alpha}, \hat{\beta}$	$\alpha = 0, \hat{\beta}$	$\hat{\alpha}, \hat{\beta}$	$\alpha = 0, \hat{\beta}$
	(a) MSPE Ratios								
1	3.831	1.094	1.035	1.055	1.016	0.996	0.972	1.034	0.956
3	1.455	1.054	1.022	1.078	1.013	1.017	1.007	1.098	0.997
6	1.258	1.208	1.026	1.206	1.015	1.127	1.116	1.050	1.037
9	-	1.320	1.042	1.288	1.038	1.157	1.128	1.017	1.005
12	-	1.312	1.035	1.233	0.994	1.107	1.045	0.975	0.941
15	-	1.270	1.056	1.188	1.002	1.110	1.047	1.038	0.932
18	-	1.174	1.105	1.202	1.046	1.141	1.154	1.072	0.983
21	-	1.115	1.130	1.176	1.055	1.319	1.342	1.103	1.036
24	-	1.087	1.084	1.081	0.994	1.604	1.587	1.034	1.032
	(a) Success Ratios								
1	0.487	0.487	0.551	0.494	0.543	0.479	0.491	0.449	0.472
3	0.574*	0.566*	0.434	0.509	0.404	0.551*	0.494	0.521*	0.502
6	0.611*	0.603*	0.424	0.515	0.401	0.527	0.523	0.531	0.527
9	-	0.571**	0.440	0.552	0.421	0.556	0.557	0.548	0.556
12	-	0.492	0.492	0.453	0.473	0.609	0.606	0.609	0.609
15	-	0.466	0.474	0.478	0.451	0.569	0.569	0.569	0.573
18	-	0.484	0.404	0.460	0.400	0.568	0.564	0.568	0.568
21	-	0.518	0.413	0.470	0.445	0.583	0.579	0.583	0.583
24	-	0.537	0.463	0.455	0.557	0.553	0.553	0.570**	0.562

Q#3: Forecasting Based on VAR Models for Real Retail Gasoline and Real Oil Prices

We investigate the role of:

- The lag order (structure);
- The definition of the price of crude oil; and
- Imposing cointegrationg restrictions.

Forecast Accuracy of Models for the Retail Gasoline Price and the Brent Price of Crude Oil Evaluated on 1992.1-2014.3: The Effect of the Lag Order

Monthly Horizon	VAR(12)	BVAR(12)	VAR(6)	BVAR(6)	VAR(1)	BVAR(1)	VAR(AIC)	BVAR(AIC)
	(a) MSPE Ratios							
1	0.908	0.828	0.963	0.927	0.868	0.867	0.928	0.905
3	0.747	0.737	0.732	0.732	0.760	0.760	0.681	0.681
6	0.847	0.949	0.762	0.765	0.738	0.778	0.730	0.728
9	1.013	1.154	0.854	0.858	0.810	0.810	0.888	0.881
12	1.243	1.151	0.914	0.919	0.879	0.880	1.037	1.012
15	1.264	1.069	0.888	0.895	0.856	0.856	0.990	0.962
18	1.167	1.104	0.857	0.864	0.806	0.807	0.920	0.901
21	1.169	1.192	0.908	0.915	0.832	0.833	1.003	0.972
24	1.266	1.024	0.963	0.972	0.899	0.900	1.072	1.035
	(b) Success Ratios							
1	0.678*	0.697*	0.689*	0.693*	0.629*	0.633*	0.693*	0.685*
3	0.679*	0.679*	0.683*	0.676*	0.664*	0.664*	0.732*	0.725*
6	0.618*	0.622*	0.634*	0.634*	0.683*	0.687*	0.679*	0.683*
9	0.517	0.564	0.622*	0.606*	0.602*	0.602*	0.575*	0.579*
12	0.527	0.500	0.563**	0.559**	0.559	0.555	0.500	0.512
15	0.518	0.514	0.577*	0.577**	0.613*	0.613*	0.522	0.538
18	0.536	0.520	0.544	0.544	0.592**	0.592**	0.520	0.524
21	0.522	0.518	0.526	0.514	0.559	0.555	0.518	0.530
24	0.484	0.471	0.516	0.525	0.533	0.528	0.525	0.533

Forecast Accuracy of Models for the Retail Gasoline Price and Alternative Oil Prices Evaluated on 1992.1-2014.3: The Effect of the Oil Price Series

Monthly Horizon	WTI		RAC with WTI Nowcast		RAC with Brent Nowcast		Brent	
	VAR(1)	BVAR(1)	VAR(1)	BVAR(1)	VAR(1)	BVAR(1)	VAR(1)	BVAR(1)
	(a) MSPE Ratios							
1	0.899	0.898	0.939	0.938	0.937	0.936	0.868	0.867
3	0.820	0.820	0.818	0.818	0.819	0.819	0.760	0.760
6	0.819	0.819	0.780	0.780	0.773	0.773	0.738	0.778
9	0.880	0.880	0.836	0.836	0.832	0.832	0.810	0.810
12	0.955	0.955	0.953	0.953	0.940	0.940	0.879	0.880
15	0.915	0.915	0.914	0.914	0.907	0.907	0.856	0.856
18	0.845	0.846	0.822	0.823	0.817	0.818	0.806	0.807
21	0.861	0.862	0.844	0.845	0.839	0.840	0.832	0.833
24	0.919	0.920	0.930	0.931	0.921	0.921	0.899	0.900
	(b) Success Ratios							
1	0.588*	0.588*	0.569*	0.569*	0.581*	0.581*	0.629*	0.633*
3	0.638*	0.642*	0.608*	0.611*	0.611*	0.685*	0.664*	0.664*
6	0.657*	0.657*	0.641*	0.641*	0.641*	0.641*	0.683*	0.687*
9	0.579*	0.583**	0.591*	0.591*	0.583*	0.583*	0.602*	0.602*
12	0.566	0.570	0.543	0.543	0.551	0.551	0.559	0.555
15	0.569	0.569	0.589**	0.585**	0.589*	0.585**	0.613*	0.613*
18	0.588	0.588	0.604*	0.604*	0.600**	0.600**	0.592**	0.592**
21	0.567	0.567	0.591	0.591	0.587	0.587	0.559	0.555
24	0.537	0.537	0.541	0.537	0.545	0.541	0.533	0.528

Forecast Accuracy of Models for the Retail Gasoline Price and Alternative Oil Prices Evaluated on 1992.1-2014.3: The Effect of Imposing Cointegration

Monthly Horizon	Brent		Brent		Brent		Brent	
	VAR(1)	BVAR(1)	VEC(1)	BVEC(1)	VEC(6)	BVEC(6)	VEC(12)	BVEC(12)
	(a) MSPE Ratios							
1	0.868	0.867	1.186	1.182	1.010	0.982	0.892	0.876
3	0.760	0.760	1.247	1.241	0.880	0.878	0.689	0.706
6	0.738	0.778	1.341	1.334	0.929	0.928	0.750	0.769
9	0.810	0.810	1.508	1.500	0.980	0.979	0.870	0.887
12	0.879	0.880	1.519	1.511	1.008	1.005	1.010	1.012
15	0.856	0.856	1.505	1.497	1.027	1.026	0.980	0.988
18	0.806	0.807	1.523	1.515	1.018	1.019	0.916	0.935
21	0.832	0.833	1.581	1.572	1.038	1.037	0.969	0.985
24	0.899	0.900	1.600	1.591	1.059	1.056	1.038	1.044
	(b) Success Ratios							
1	0.629*	0.633*	0.618*	0.618*	0.700*	0.689*	0.697*	0.693*
3	0.664*	0.664*	0.521	0.521	0.626*	0.619*	0.706*	0.687*
6	0.683*	0.687*	0.508	0.508	0.538	0.557*	0.672*	0.691*
9	0.602*	0.602*	0.529	0.533	0.552*	0.552*	0.587*	0.595*
12	0.559	0.555	0.535	0.539	0.543**	0.547**	0.543	0.551**
15	0.613*	0.613*	0.530	0.530	0.557*	0.553*	0.549	0.561**
18	0.592**	0.592**	0.496	0.500	0.592*	0.584*	0.584*	0.560
21	0.559	0.555	0.498	0.498	0.530	0.539**	0.555	0.534
24	0.533	0.528	0.525	0.525	0.537	0.537	0.484	0.471

Q#4: Forecasting Based on Models Containing Information about the Domestic Gasoline Market

1. Factor forecasting model:

$$r_{t+h|t}^{h,gas} = \alpha + \beta\, cfnai_t + \varepsilon_{t+h}$$

   where $r_{t+h|t}^{h,gas}$ denotes the cumulative percent change in the real price of gasoline expressed as a fraction such that $\hat{R}_{t+h|t}^{gas} = \exp\left(r_t^{gas} + \hat{r}_{t+h|t}^{h,gas}\right)$.

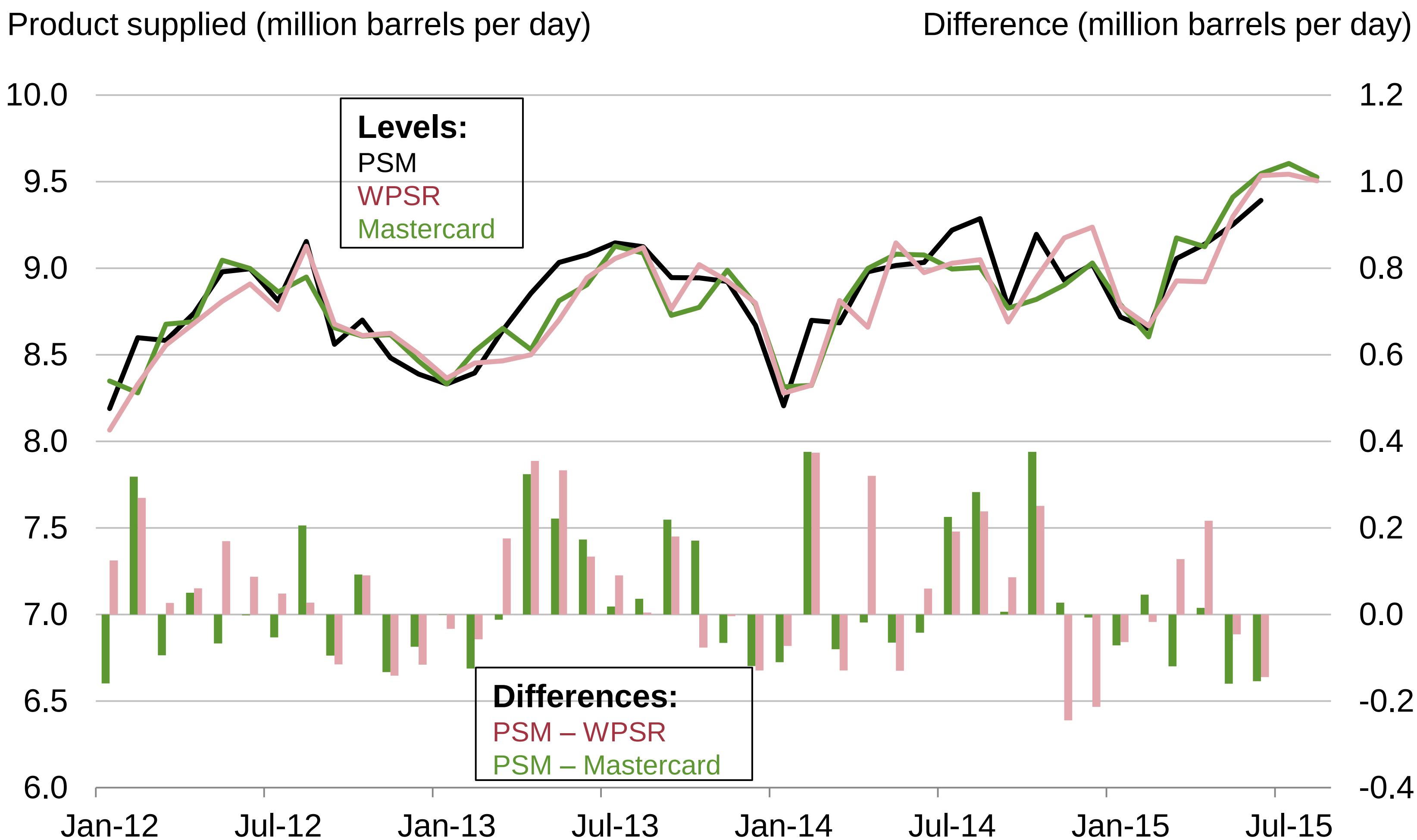
2. Product Factor-augmented VAR (FAVAR) model including the real retail price of gasoline and the CFNAI.

3. VAR model for real U.S. gasoline consumption and real retail gasoline price.

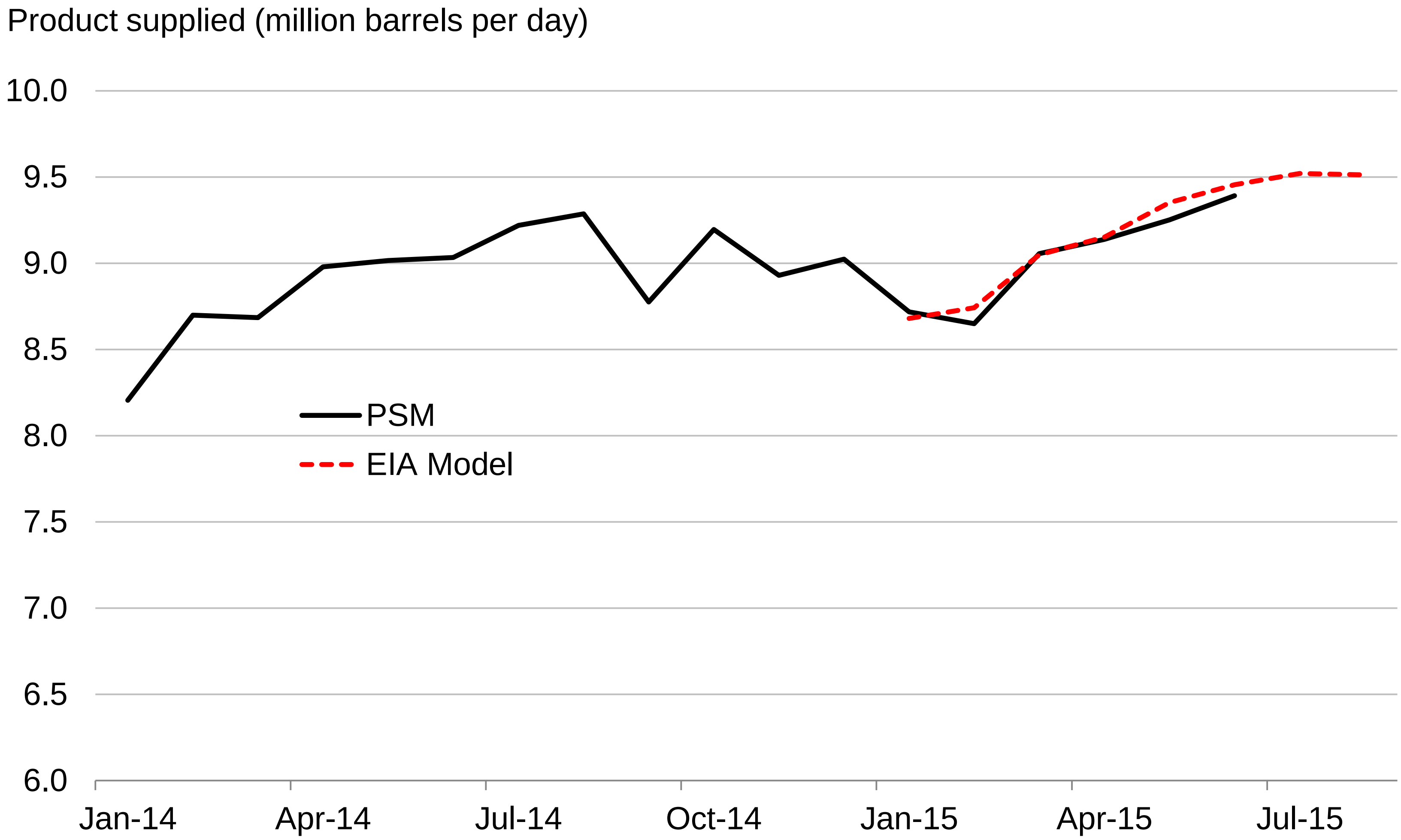
Forecast Accuracy of U.S. Gasoline Market Models Evaluated on 1992.1-2014.3

	Real Gasoline Consumption						U.S. Real Economic Activity (CFNAI)		
Monthly Horizon	VAR(12)	BVAR(12)	VAR(6)	BVAR(6)	VAR(1)	BVAR(1)	Factor Model	FAVAR (12)	BFAVAR (12)
	(b) MSPE Ratios								
1	1.260	1.187	1.267	1.211	0.996	0.996	1.008	1.276	1.189
3	0.997	0.982	1.057	1.050	0.998	0.997	1.027	1.044	1.017
6	0.923	0.912	0.939	0.947	1.002	1.001	1.027	0.969	0.959
9	1.046	1.026	1.016	1.019	1.011	1.010	1.034	1.055	1.035
12	1.194	1.157	1.076	1.071	1.026	1.025	1.039	1.172	1.135
15	1.134	1.104	1.058	1.058	1.025	1.024	1.146	1.140	1.107
18	1.060	1.040	1.029	1.032	1.021	1.019	1.056	1.081	1.063
21	1.120	1.101	1.087	1.086	1.037	1.036	1.079	1.122	1.111
24	1.187	1.164	1.136	1.136	1.058	1.058	1.105	1.172	1.156
	(c) Success Ratios								
1	0.674*	0.678*	0.655*	0.655*	0.532	0.524	0.502	0.663*	0.659*
3	0.642*	0.630*	0.551*	0.566*	0.491	0.494	0.468	0.626*	0.604*
6	0.649*	0.645*	0.576*	0.542*	0.470	0.473	0.489	0.622*	0.618*
9	0.533**	0.521	0.510**	0.510**	0.417	0.421	0.521	0.541*	0.525*
12	0.504*	0.481	0.481	0.496	0.414	0.414	0.512	0.500	0.473
15	0.538*	0.534**	0.526*	0.510**	0.407	0.419	0.490	0.534*	0.518
18	0.536**	0.528**	0.524*	0.520**	0.452	0.448	0.456	0.512	0.500
21	0.526	0.518	0.490	0.494	0.445	0.449	0.453	0.478	0.482
24	0.488	0.471	0.529**	0.525	0.439	0.430	0.463	0.496	0.492

Historical Gasoline Product Supplied Data



Gasoline Product Supplied Nowcast



Q#5: Forecasting Based on Models Containing Both the Domestic Gasoline Market and the Global Oil Market

- Extension of Baumeister and Kilian (JBES 2012, IER 2014) oil price forecasting model with real U.S. retail gasoline price and real U.S. gasoline consumption added.

Forecast Accuracy of Joint Gasoline Market and Oil Market Models Evaluated on 1992.1-2014.3

Monthly Horizon	Kilian and Murphy (2014) oil market model augmented with:				Kilian (2009) oil market model augmented with:			
	Real gasoline price + Real gasoline consumption		Real gasoline price only		Real gasoline price + Real gasoline consumption		Real gasoline price only	
	VAR(12)	BVAR(12)	VAR(12)	BVAR(12)	VAR(12)	BVAR(12)	VAR(12)	BVAR(12)
	(a) MSPE Ratios							
1	0.891	0.803	0.908	0.828	0.876	0.814	0.900	0.843
3	0.732	0.722	0.747	0.737	0.742	0.725	0.766	0.747
6	0.857	0.801	0.847	0.799	0.829	0.795	0.815	0.787
9	1.042	0.968	1.013	0.949	1.008	0.969	0.971	0.942
12	1.277	1.178	1.243	1.154	1.279	1.199	1.226	1.168
15	1.277	1.160	1.264	1.151	1.246	1.173	1.222	1.160
18	1.177	1.075	1.167	1.069	1.129	1.078	1.097	1.061
21	1.178	1.111	1.169	1.104	1.184	1.126	1.156	1.110
24	1.277	1.210	1.266	1.192	1.300	1.233	1.265	1.208
	(b) Success Ratios							
1	0.685*	0.723*	0.678*	0.697*	0.697*	0.715*	0.693*	0.708*
3	0.694*	0.694*	0.679*	0.679*	0.683*	0.691*	0.676*	0.679*
6	0.630*	0.653*	0.618*	0.622*	0.649*	0.645*	0.641*	0.630*
9	0.556	0.556	0.517	0.564	0.583*	0.568**	0.529	0.564
12	0.535	0.520	0.527	0.500	0.516	0.512	0.516	0.508
15	0.530	0.534	0.549	0.514	0.510	0.526	0.518	0.522
18	0.540	0.532	0.518	0.520	0.528	0.540	0.512	0.528
21	0.555	0.506	0.522	0.518	0.498	0.494	0.498	0.514
24	0.475	0.455	0.484	0.471	0.447	0.451	0.455	0.467

Forecast Accuracy of Pooled Real Gasoline Price Forecasts Evaluation period: 1992.1-2014.3

	Equal-weighted combination of five forecasting models	
Monthly horizon	MSPE ratio	Success ratio
1	0.856	0.588*
2	0.768	0.613*
3	0.735	0.642*
4	0.733	0.671*
5	0.733	0.662*
6	0.736	0.649*
7	0.742	0.667*
8	0.763	0.635*
9	0.800	0.606*
10	0.846	0.574
11	0.800	0.572
12	0.895	0.547
13	0.897	0.541
14	0.891	0.534
15	0.879	0.561
16	0.859	0.599*
17	0.839	0.594**
18	0.826	0.596**
19	0.824	0.594**
20	0.825	0.613*
21	0.838	0.583
22	0.857	0.553
23	0.855	0.551
24	0.872	0.541

Pooled forecast models:

(1): Exponential smoothing

(2): UC-SV model

(3): Brent VAR(1)

(4): Brent VEC(12)

(5): KM BVAR(12)

Conclusion

1. Pooled forecasting approaches improves in accuracy compared with the no-change.
   - An equal-weighted average of these five forecasts yields substantial reductions in the MSPE at all horizons and significant directional accuracy.
2. The most accurate individual models are the bivariate Brent-gas price VAR(1) model and the augmented oil price forecasting BVAR(12) model.
3. As expected, consumer survey forecasts are not particularly accurate.
4. In the case of forecasting the real and nominal price of gasoline by quarterly, the accuracy gains are even larger with MSPE reductions as high as 31% and directional accuracy as high as 74%.