Backcasting U.S. Oil Demand over a Turbulent Decade

Hillard G. Huntington

forthcoming in *Energy Policy*
EMF OP 69
April 2011

Energy Modeling Forum
Stanford University
Huang Engineering Center
475 Via Ortega
Stanford, CA 94305-4121

Voice: 650-723-1050
Telefax 650-723-5362

Email: **hillh@stanford.edu**

March 2011 (Revised)

Backcasting U.S. Oil Demand over a Turbulent Decade

Abstract

This paper compares 10-year backcast projections of U.S. petroleum consumption for the 2000-2009 period, based upon a range of functional forms. Although a univariate specification based upon only lagged consumption performed better than most approaches, it provides little value to the policymaker who wants to understand the factors that influence oil consumption and by how much. One structural approach performed considerably better than all other approaches including the univariate model. This specification is the autoregressive distributed lag (ADL) model that allowed oil demand to respond differently (asymmetrically) to price increases and decreases. Certain forms of this equation, however, do well while others perform poorly. Evaluations of backcast projections over different periods and updated frequently can be important byproducts of energy modeling.

Keywords: Energy demand, energy projections

Backcasting U.S. Oil Demand over a Turbulent Decade

1. Introduction

World crude oil prices have varied from the mid $30's to the mid-$140's since 2001. Oil demand growth and retrenchment have played a critical role in this oil-market "roller coaster" over these years. Any evaluation of recent oil market trends (e.g., see Smith, 2009) must begin with understanding the factors contributing to oil demand developments.

This paper explores what is known about projecting future oil demand conditions. It assembles several leading econometric approaches for evaluating demand conditions and compares them in their ability to understand recent trends. Emphasis is placed on the United States, where an extensive and reliable data set on oil demand and contributing factors is available.

This study uses an out-of-sample backcasting approach for evaluating recent historical trends where fitted values from the model are compared with actual levels. The model's parameters are estimated over an initial period and are then used to project energy consumption over a second (out-of-sample) period. Comparing fitted with actual values of consumption levels is very similar to the approach used by the U.S. Energy Information Administration (2008) or by other researchers (Winebrake and Sakva, 2006, and O'Neill and Desai, 2005) when they evaluate reference-case projections of fuel consumption from the National Energy Modeling System (NEMS), an engineering-economic system approach used for developing energy market forecasts since 1982. Unlike these earlier studies, however, this study controls for the actual historical values of the independent variables of oil prices and GDP and hence is more similar to the approach adopted by Fischer, Herrnstadt and Morgenstern (2008). It asks essentially how well can one project oil consumption if you had reliable information about price and economic growth.

The next section describes the objectives for projecting oil demand. Section 3 explains the data sources and key data properties. The main econometric approaches are categorized in section 4. Section 5 explains the procedure for evaluating recent trends through a backcasting technique, before comparing the key results in section 6. Various aspects of the goodness-of-fit of the major equation types are evaluated in section 7, while the main conclusions are summarized in section 8.

2. Objectives

There exists an extensive literature on oil product demand estimation. Dahl and Sterner (1991) compared the price and income elasticities of demand from more than 100 studies on gasoline; more than 10 years later, Goodwin, Dargay and Hanly (2004) updated these estimates by comparing the results from 69 newer studies. In a companion piece, Graham and Glaister (2004) reported results from more than 100 studies. Certainly, this is an issue that has attracted considerable attention from energy economists over the last several decades.

Any individual study included in these surveys will frequently develop its preferred approach, estimate the equation coefficients and conduct tests supporting or refuting various hypotheses. If projections are developed from a single equation, they are not routinely compared with competing specifications to understand important differences in their long-run forecasting properties. As a result, policymakers do not know whether the elasticities reported in any one study come from an approach that explains the recent past well or poorly. While the various surveys often weight each study's elasticity equally when computing mean or median estimates, one could argue for a higher priority for estimates that can track recent history successfully.

In fact, the range of estimated elasticities is very broad across studies. Goodwin, Dargay and Hanly (2004, Table 3), for example, report a range of 0 to -1.81 for the long-run price elasticity of gasoline demand from equations very similar to those estimated in this study. These equations involve dynamic estimation with time series data that allows long-run to be greater than short-run responses. Policy studies of cartel behavior or domestic energy taxes that are based upon estimates at the higher end will have very different market impacts than those at the lower end.

An issue of common interest among policymakers is the long-run trend in energy demand and its relationship to global climate change policy. This paper adopts the perspectives of organizations like the U.S. Energy Information Administration (EIA) and the International Energy Agency (IEA) that have the responsibility for long-run projections, where long-run adjustments in the capital stock may take years rather than months or quarters. Accordingly, the estimates and projections are based upon annual rather than higher-frequency quarterly or even monthly data that might be appropriate for short-run conditions.

Moreover, these groups do much more than develop projections with their models. They also support policymakers who are also interested in addressing how different policy instruments influence oil markets. They may also be keenly interested in explaining how oil prices and other market outcomes may respond under a range of different energy and economic conditions. They can meet these additional issues only by maintaining structural equations that can identify the separate roles of key factors like price and income. For this reason, the paper focuses on structural equations rather than on *univariate* techniques that explain current oil consumption based solely upon past oil consumption. These latter techniques might be more suitable if these groups were interested solely in projections.

The success of different approaches will also depend upon the type of market being evaluated. Some approaches may be particularly appropriate when long-term trends dominate and oil prices are not particularly volatile over the short run. Other approaches may be more relevant if oil prices should suddenly change and exert new pressures on the market. Oil prices may well continue their volatility, so it is important to include periods before as well as after the price escalations since 2003. In all likelihood, analysts are interested in approaches that do well in both situations.

3. Data Sources and Properties

Although many explanatory variables can be included, oil demand functions for many countries, at their most fundamental level, often express the quantity of consumed oil as a function of prices and income. An extensive data set exists for the United States that covers these variables for different products over a very long period (at least since 1949). This data allows total consumption to be segmented into different products and separate price variables to be constructed for different product types.

The U.S. data sources allow considerable flexibility in designing equations that meet a variety of purposes. In this study, oil consumption is defined as total petroleum products excluding residual fuel oil use, as proposed by Ryan and Plourde (2002) and Dargay and Gately (2010). This definition is superior for the U.S. experience, because residual and non-residual consumption appear to have different cointegration patterns with prices and income (Huntington, 2010a). The excluded residual fuel oil has many different properties from other petroleum products. It can be replaced very quickly, at relatively low cost and without major reconfiguration of the capital stock. In contrast, other uses of petroleum products often are much more closely connected to their capital stock that requires the use of petroleum. Examples include gasoline in internal combustion engines of today's automobile fleet, jet fuel in airplanes, and raw material feedstocks in industries producing a range of chemicals and plastics.

The U.S. Energy Information Administration (2010, Table 5.11) reports series on petroleum product consumption, the U.S. Bureau of Labor Statistics (2011) provides data on petroleum product prices, the U.S. Bureau of Economic Analysis (2011) collects information on real Gross Domestic Product (GDP), and the U.S. Bureau of the Census (2011) reports total U.S. population. Both total non-residual petroleum consumption and real GDP (2002 dollars) were expressed in per-capita terms by dividing each by population. The price variable is the producer price index for total refined petroleum products that has been benchmarked to equal \$8.40 per million Btu in 1982, as reported by the EIA (2010, Table 3.3). It is deflated by the producer price index for all commodities to form real prices in 2002 dollars. All variables are converted to logarithms.

Recent studies (e.g., Hughes, Knittel and Sperling, 2008) have emphasized that consumption responds less to price than in the past. One approach for incorporating this instability would be to estimate separate responses over different periods or perhaps truncating the estimation period. Another approach that exploits the full data set has been to decompose the single oil price variable into separate series that are mutually exclusive but that equal the original oil price series when aggregated. Although "asymmetric responses" are often ascribed to this approach, it principal advantage lies in its ability to separate the truly striking price trends of the 1970s and early 1980s from other price movements. As defined by Gately and Huntington (2002), these maximum price, price cut, and price recovery series are:

- The maximum historical price, $P_{max,t}$, which equals the highest logarithmic price between the initial year (1949) and the current year, t;
- The cumulating series of price cuts, $P_{cut,t} \leq 0$, which is monotonically non-increasing in logarithms;
- The cumulating series of sub-maximum price recoveries, $P_{rec,\ t} \geq 0$, which is monotonically non-decreasing in logarithms.

Both price cuts and price recoveries can be considered sub-maximum price changes, while any prices that exceed the previous historical peak can be considered maximum prices.

Specifying the relationship between consumption, prices and economic output will depend upon the data properties of these variables. Each series is initially tested to determine whether the variable is stationary, with its mean and variance remaining stable over time, e.g., as in rolling dice or flipping a coin. A variable will be stationary if its Dickey-Fuller statistic significantly rejects the test's assumption that the variable is non-stationary. These tests also include the constant term and a time trend as deterministic regressors. None of the statistics in Table 1 are significant for logarithmic levels but they all are significant for their first differences. Thus, each of the levels is not stationary, but each variable becomes stationary after conversion to first differences.

Table 1. Augmented Dickey-Fuller Tests for Unit Roots, 1951-2009

	Test Statistic	Lags		Test Statistic	
Consumption	-1.33	2		-5.37	**
GDP	-2.10	0		-6.48	**
Price	-2.41	1		-6.01	**
Maximum Price	-2.07	1		-4.69	**
Price Recovery	-0.03	2		-5.95	**
Price Cut	-1.36	1		-7.09	**
Stochastic regressors incude trend.					
Tests for differences include one less lag.					
** Significantly rejects unit root at 1 percent level.					

Even if two or more variables do not meet this requirement for a stationary process, their levels may still be correlated (or cointegrated) with each other. Table 2 compares Johansen trace cointegration tests for three different long-run relationships: (1) per capita consumption, price and per capita GDP, (2) per capita consumption, the three price components (recoveries, cuts and maximum) and per capita GDP, and (3) per capita consumption, maximum price, sub-maximum price and per capita GDP. Sub-maximum prices consider price recoveries and price cuts as a single variable.[1] The first row of the table shows whether one can reject the hypothesis of not even a single cointegrating relationship. A significant statistic (indicated by an asterisk) indicates that there is a long-run relationship between the variables and they are cointegrated.[2]

Table 2. Johansen Tests for Cointegration, 1949-2009

	(1)	(2)	(3)	
	trace	trace	trace	
rank	statistic	statistic	statistic	
0	25.11	67.42	48.03	*
1	8.2317	38.10	27.20	
2	0.267	21.91	11.76	
3		9.32	1.48	
4		0.83		
* Significantly rejects no cointegration at 5% level.				
(1) Consumption, Price, and GDP.				
(2) Consumption, Maximum Price, Price Recoveries, Price Cuts and GDP.				
(3) Consumption, Maximum Price, Submaximum Price and GDP.				

The tests imply that there is no long-run relationship between these variables when the single price variable is included. The tests fall just short of being significant at the 5 percent level when the three price components replace the single price variable. This failure to reject could be due to the considerably lower power of the test when it is applied to five distinct variables. When price recoveries and cuts are coupled into one variable, the variables are cointegrated at the 5 percent level.

The results from these cointegration tests will be used to interpret the out-of-sample forecasting performance of different equations but will not be used to restrict which forms are tested a priori. The lack of a cointegrating relationship suggests that percent changes may be superior to levels in the backcasting exercise when the effects of oil price increases and decreases have symmetric effects on oil consumption.[3] The presence of a cointegrating relationship when maximum oil prices are separated from other oil price movements suggests that the asymmetric specification may be superior to the symmetric version. Although cointegration can influence long-run projections (e.g., see Thoma and Duy, 1998), it does not necessarily improve the forecasts.

4. Alternative Specifications

Oil economists have estimated a very large number of different specifications over the years. Equations relate the variable levels of consumption, prices and GDP (Jones 1991), the variable differences (Bentzen and Engsted, 1993), or an error-corrections adjustment like the flexible autoregressive distributed lag, or ADL (Hunt, Judge and Ninomiya, 2003; Huntington 2009). The demand responses to price (and maybe income) can be symmetric (Griffin and Schulman 2005) or asymmetric (Dargay, 1992, Dargay and Gately, 1995a, 1995b, 1997). Income effects can change with the income level and decline as the economy becomes richer (Galli, 1998). Technical change can be included as a constant trend (Beenstock and Willcocks, 1981, 1983), a stochastic variable (Hunt and Ninomiya, 2003; Adeyemi and Hunt, 2007), or a set of dummy variables (Griffin and Schulman 2005). The possibilities are almost unlimited.

Estimation can also differ due to the econometric technique applied for relating the variables. For tractability and ease of presentation, our tests are restricted to single-equation specifications based upon ordinary-least-squares (OLS) estimation for policy making. Many energy and transportation agencies are most familiar with this technique, as evidenced by the previously cited surveys of demand studies. Although price elasticity estimates may be biased, the OLS technique does not necessarily produce inferior projections than those based upon simultaneous-equation techniques (Kennedy 2003, p. 187).[4]

Rather than being exhaustive, the analysis selects a few major OLS approaches to represent the possible range of interest.[5] Equations are grouped into two major groups: *levels* (where per capita consumption is a function of price and per capita GDP levels) and *deviations* (where the annual change in per capita consumption is a function of annual changes in prices and per capita GDP).

The *levels* specification expresses oil consumption in period *t* as a function of the current and past price and income levels and previous consumption levels. A trend factor (*t*) may or may not be included to incorporate the possibility of a significant decreasing effect due to exogenous technical change over time.

$$Z_t = \gamma + \sum_{i=0}^{n} \alpha_i X_{t-i} + \sum_{j=1}^{n} \beta_j Z_{t-j} + \delta t \qquad (1)$$

where Z represents per capita oil consumption, X refers to real petroleum price or per capita real GDP, *t* denotes the year or time indicator, and γ, α_i, β_j and δ are the estimated parameters. An attractive feature of this specification is that it can be rewritten as an error-corrections equation where the change in per-capita oil consumption can be determined by the change in oil prices and per-capita real GDP as well as the lagged level for all three variables.[6]

The levels model can be estimated both with a symmetric price response (using a single price variable) and an asymmetric price response (using the three decomposed price series). Sometimes the lagged levels are excluded for the price and GDP variables. When both lagged variables are excluded, the specification becomes the popular Koyck distributed lag model, where oil demand adjusts gradually to oil price movements. The

Koyck model is estimated by setting $\alpha_i=0$ for $i>0$ and $\beta_j=0$ for $j>1$ in equation (1), resulting in the following equation

$$Z_t = \gamma + \alpha_0 X_t + \beta_1 Z_{t-1} + \delta t \qquad (2)$$

When lagged income is included but lagged prices are excluded, the equation allows short-run adjustments to income that may differ from short-run adjustments to prices. This equation looks like equation (2) except there is also a lagged income variable. Gately and Huntington (2002) have used this specification and Dargay and Gately (2010) have developed long-run projections for the OECD nations using this basic approach.[7] Both short cuts reveal serious deficiencies for out-of-sample backcasting of U.S. consumption since 2000, as will be observed below.

The backcast evaluations will compare the structural equations to a univariate approach where consumption is a function of past consumption levels but without any time trend or price and income terms. The univariate function is equation (1) where $\delta=0$ and all $\alpha_i=0$.

The *deviations* specification expresses the change in oil consumption in period *t* as a function of current and past changes in price and income and previous changes in consumption. It is derived directly from equation (1) by solving for Z_t and Z_{t-1} and computing their difference.

$$\Delta Z_t = \sum_{i=0}^{n-1} \alpha_i (\Delta X_{t-i}) + \sum_{j=1}^{n-1} \beta_j (\Delta Z_{t-j}) + \delta \qquad (3)$$

There will be one fewer lagged term in the deviations than in the levels specification. A significant constant term in this equation demonstrates a significant trend factor for exogenous technical progress over time. Backcasts from these equations will be compared with those from univariate equations where the current oil consumption change is governed by its past changes.

All equations can include additional levels (or deviations) from previous years, referred to as lags. The backcasts in section 6 will evaluate the effect of the number of lags by allowing as many as four lagged levels of all variables in the levels specification (as suggested by Jones, 1993) and three lagged differences of all variables in the first-difference specification (as suggested by Bentsen and Engsted, 1996).

5. Backcast Procedure

The projection capability of each specification is based on its ability to forecast accurately over a recent historical period of ten years. Out-of-sample backcasting measures how well the equation tracks historical values of consumption, provided that one has perfect information about the explanatory variables (prices and income) over this same period.[8] It combines the best available information about the equation's parameters at some point, t_e, with perfect information about the explanatory variables after the same point, t_e. Any errors in these price- and income-corrected backcasts can be attributed to the parameter estimates or structural relationships rather than to incorrect price and income assumptions. This approach differs from what some studies (U.S. Energy Information Administration, 2008, and Winebrake and Sakva, 2006) do when they compare projected values directly with actual ones that are observed after the forecast. These latter backcasts include errors in both the projection equation (its structure and parameters) and the explanatory variables.

When the equation includes lagged terms of the dependent variable (oil consumption), it is most useful to allow the equation to compute these values successively each year rather than using the actual historical levels. These projections are sometimes called dynamic or endogenous backcasts. They are completely different from static backcasts where actual consumption values are substituted exogenously for the lagged terms to determine each year's current consumption level. Good dynamic backcasts are much more challenging than good historical fits of the data, because errors in projecting consumption in previous periods will influence errors in projecting current consumption through the lagged effects. These problems can be an important source of error in any of these equations.

Backcasts are prepared for 2000-2009 on the basis of the previously estimated equation for 1949-1999. A widely accepted measure of the error size is the root-mean-square errors (RMSE) associated with the demand level for each specification. It penalizes larger errors proportionately more than smaller errors. The error for any year shows the difference between the logarithms of the projected and actual values, or the percent difference between the two when converted to actual levels. The RMSE measure sums the square of each annual error over the entire backcast period before taking its square root. Hence, the RMS errors are computed as:

$$RMSE = \sqrt{\sum_i (Q_{bi} - Q_{ai})^2 / n}$$

where Q_{bi} and Q_{ai} represent the backcasted and actual logarithmic levels and n denotes the number of backcasted periods. When an equation projects demand deviation rather than level, its estimated deviation is added to the projected demand level from the previous year to construct a new demand level in the current year.

There are also many other possible ways to measure the accuracy of a projection or backcast. Although emphasis will be placed on the RMSE, our table will also report several other measures. The mean absolute deviation (MAD) indicates the average size of the error without distinguishing whether the equation over-predicts or under-predicts the variable. The percent turning-point measure (% tp) shows the number of years where the equation correctly predicts whether oil consumption is increasing or decreasing, but ignores the size of the error. The correlation index (corr) expresses the correlation between actual and predicted changes in oil consumption. The RMSE for the first six years is also reported separately from its counterpart for the full ten years. This alternative measure provides a perspective on how the specification performs over the short-to-mid term, when the errors generally should be smaller than those for the long run. Since there was also more stability in oil markets over the 2000-2005 than the 2006-2009 period, this alternative measure may also be revealing how well the specification performs over more stable market conditions. It may be important to understand that some equations may provide better estimates when prices and income are moving steadily along their trends, while other equations may be better at incorporating large swings in the underlying explanatory variables.

6. Backcast Results

It is important to emphasize that the best functional form for projecting U.S. oil demand over this period may not be the best one for other countries. Nevertheless, the

U.S. results may provide insights into some promising approaches when considering the functional form more globally. It is also highly unlikely that a particular equation will perform well for all years in explaining U.S. oil consumption patterns. Even so, these initial results should be instructive. It would be highly valuable to extend this analysis to other countries and to update these U.S. results occasionally.

Table 3 compares the backcasting accuracy of 20 different demand specifications. Each column reports the result for a different measure of accuracy as discussed in the preceding section. Minimum values for RMSE and MAD and maximum values for the turning point and correlation coefficients are shown in bold type. Thus, the full asymmetric function with lags for prices, per capita GDP and per capita oil consumption has the lowest RMSE (0.038) for the ten-year backcast period.

At the top of Table 3 are the levels specification differentiated by the number of lagged levels in parentheses. Additional levels specification include an equation with one lag and a time trend, "Levels{1}, Trend", and a Koyck model that excludes lags for the price and per capital GDP variables. Next are reported asymmetric specifications that include separate price components for price maximums, recoveries and cuts. The levels specifications are completed with four univariate functions where only lagged per capita oil consumption levels are included. At the bottom of the table are shown difference equations with varying lag lengths, including univariate functions where only lagged percent changes in oil consumption are included.

Table 3. Out-of-Sample Forecasting Accuracy for Different Oil Demand Specifications, 2000-2009

Specification	rmse	rmse(6)	mad	%tp	corr
Symmetric Levels (1)	0.286	0.128	0.076	70%	0.845
Symmetric Levels (1) w Trend	0.310	0.138	0.082	70%	0.843
Symmetric Levels (2)	0.427	0.193	0.117	70%	0.706
Symmetric Levels (3)	0.443	0.196	0.121	70%	0.693
Symmetric Level (4)	0.395	0.167	0.107	80%	0.526
Sym. Levels (1) w/o Lagged Price, GDP	0.095	0.044	0.025	0.700	0.472
Asym Levels (1)	**0.038**	**0.022**	**0.011**	80%	0.948
Asym Levels (1) w/o Lagged Price	0.248	0.122	0.067	70%	0.882
Asym Levels (1) w/o Lagged Price, GDP	0.207	0.113	0.056	70%	0.808
Past Levels (1)	0.142	0.025	0.026	70%	-0.594
Past Levels (2)	0.158	0.041	0.032	60%	0.167
Past Levels (3)	0.169	0.042	0.035	10%	0.078
Past Levels (4)	0.160	0.041	0.033	50%	0.068
Deviations	0.170	0.040	0.038	50%	**0.961**
Deviations w Trend	0.425	0.231	0.120	70%	0.892
Deviations (1)	0.123	0.069	0.032	**90%**	0.442
Deviations (2)	0.246	0.151	0.071	70%	0.348
Deviations (3)	0.309	0.171	0.088	**90%**	0.284
Past Changes (1)	0.405	0.152	0.106	30%	0.265
Past Changes (2)	0.403	0.153	0.106	30%	0.119
Past Changes (3)	0.395	0.148	0.103	30%	0.185

Notes
Minimum or maximum (%tp, corr) are indicated by bold type.
rmse = root mean squared error.
rmse(6) = root mean squared error for first six years (2000-2005).
mad = mean absolute deviation.
%tp = % turning point = % years where direction was predicted correctly.
corr = correlation between actual and fitted percent changes.

The results in this table emphasize four major conclusions:

1) Deviations performed better than levels when symmetric functions are assumed *a priori*. This result may appear surprising, because it implies that prices influence demand very quickly within the first few years rather than through a long delayed adjustment due to the slow turnover in the capital stock. The superiority of deviations appears to be partly due to the earlier finding that there is no long-run relationship

between the *levels* of per capita consumption, price and per capita GDP when expressed as a symmetric function. These results are consistent with the argument advanced by Bentzen and Engsted (1996), who found that the variables of consumption, price and GDP *levels* were not cointegrated. They suggested the use of first differences when non-stationary variables were not cointegrated.

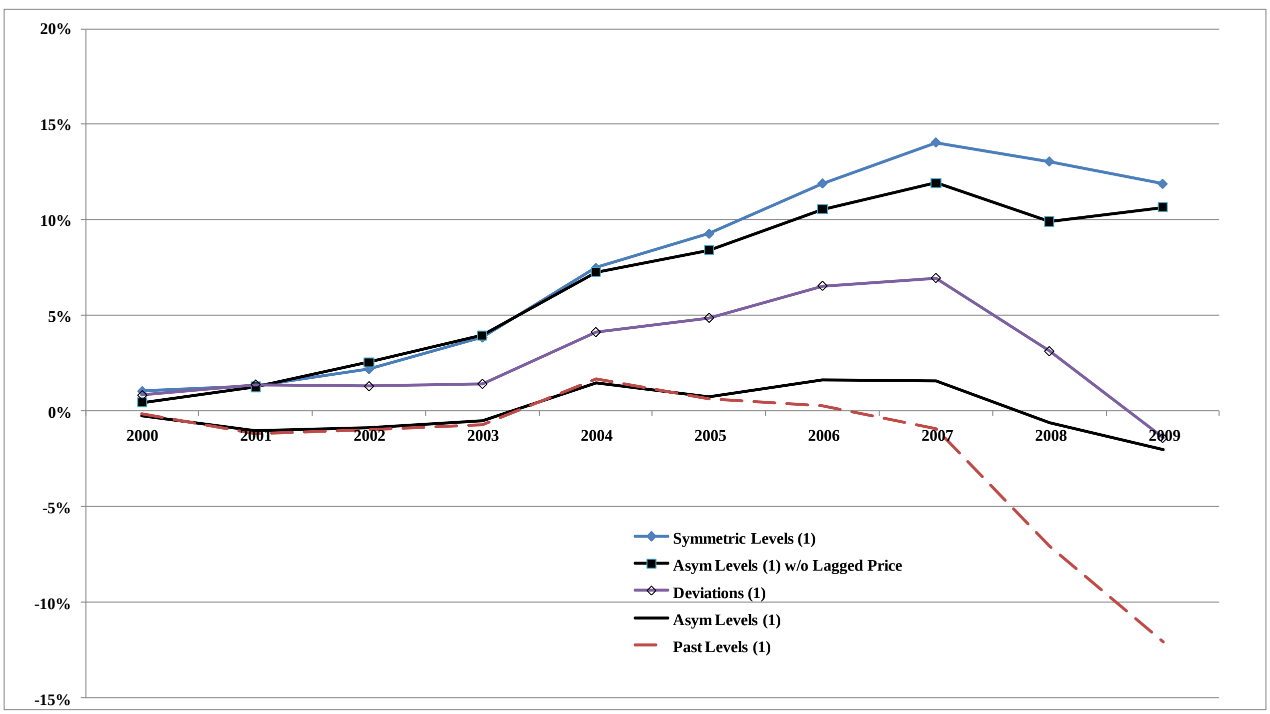
2) Asymmetric functions with one lagged level uniformly perform better than their symmetric counterpart, but ignoring lagged prices causes a severe deterioration in forecasting accuracy. Both the Koyck-lag specification (the third asymmetric equation in the table) and the one immediately above it sacrifice considerable backcasting accuracy and are strongly inferior to an approach that also includes lagged price terms. The RMSE declines from 0.248 and 0.207 to 0.038 in the row labeled "Asym Levels (1)".

3) The improved accuracy of the full-asymmetric function is striking relative to the other asymmetric specifications. Except for this function and perhaps the symmetric Koyck model, all symmetric and asymmetric functions with prices (or price components) and per capita GDP as explanatory variables perform considerably worse than a simple equation including only lagged per capita consumption.

4) The reverse pattern holds for equations using deviations rather than levels. Here, the univariate functions including only lagged consumption changes performs worse than the other specifications.

Although the RMSE is a convenient measure for backcasting accuracy, it fails to provide an intuitive understanding of the errors and how they may change over time. The residual term for each equation is the amount that must be added to (or subtracted from, if a negative number) the fitted value. Each year's error is plotted in Figure 1 for a sample of the different equations, each with one set of lagged variables. The top line in the figure is the symmetric levels specification, which worsens over time until 2007. Immediately below is the asymmetric function without any lagged price terms. It does about as poorly until 2005, when oil prices began to rise sharply. Higher prices presumably contribute to slower oil consumption growth after 2007 according to this model. Neither of these equations, however, do as well as the symmetric equation expressed in terms of deviations, the third line in this figure. The final line for a structural model refers to the full-asymmetric model with lags for the price components and per capita GDP as well as per capita consumption. Through 2005, it closely tracks

Figure 1. Errors for Sample Specifications, One Lag, 2000-2009



both actual demand and the backcasts based purely upon lagged consumption. The latter's backcast tends to deteriorate over the 2005-2009 period.

7. Goodness-of-Fit of the Estimated Equations

Overall, the estimated equations upon which the backcasts are based appear to pass many of the tests frequently used for evaluating empirical estimates.[9] The levels equations have much higher goodness-of-fit over the estimation period but did not perform better than the deviations specifications for out-of-sample backcasting. For example, with one year's lag, the adjusted-R-squares show that the first-difference equations explain about 64% of the variation in the change in per capita consumption. These estimates are less than those for the levels equations that explain more than 99% of the variation. The significant F-tests for both levels and deviations clearly indicate that the coefficients of the explanatory variables are not zero. Additionally, insignificant Jarque-Bera test statistics do not reject that the errors are normally distributed in either equation. Insignificant Breusch-Godfrey test statistics, applied to equations with lagged dependent variables, do not raise any concern about the errors being serially correlated. Autocorrelation would influence the projections, unless an autocorrelation adjustment is applied.

The tests for lag length for estimation and hypothesis testing compare the explanatory power of the equation with the equation's parsimony, as defined by the number of equation constraints or estimated parameters. The Schwarz Information Criteria is the most widely used test for the number of lags, but the Akaike Information Criteria is also reported sometimes. Both tests confirm that one year's lagged value should be included in both the levels and deviations specifications. These equations with one year's lagged values also performed better than others with other lagged structures in the backcasting comparison shown in Table 3 above.

8. Conclusions

There is no blueprint for successful oil demand projections. No one equation for developing petroleum demand projections will outperform all others for all countries and all time periods. However, the practitioner can improve his chances by routinely evaluating different specifications and avoid locking into his favorite functional form.

This paper compares 10-year, backcast projections of U.S. petroleum consumption that began in 2000. These out-of-sample backcasts were based upon coefficients that were estimated for the 1950-1999 period. One should expect significant errors in these backcasts because the decade that began in 2000 covered a very turbulent period.

Symmetric models are specifications that consider the effects of price increases and decreases to have equal but opposite effects on oil consumption. Within this group of models, equations based upon year-to-year deviations in the variables perform noticeably better than equations based upon levels. Equations based upon levels also performed less well than those univariate approaches where only lagged levels were used to explain current consumption levels.

The comparison shows that one structural approach performed significantly better than the other tested specifications. It was also clearly superior to the univariate approach where current consumption was explained as a function of its lagged values only. This specification is the autoregressive distributed lag (ADL) model that allowed oil demand levels to respond asymmetrically to price maximums, recoveries and cuts. It performed considerably better than the comparable model where the responses were assumed to be symmetric. However, not all "asymmetric" formulations performed equally as well. Ignoring lagged values of prices can be particularly problematic. These results suggest caution when using the popular the Koyck-distributed-lag specification or the approach used by Gately and Huntington (2002) and applied by Dargay and Gately (2010) for forecasting future demand.

The preferred asymmetric equation shows considerably higher responses to prices that reach new maximum levels than to prices that retrace previous levels. The long-run price elasticity of oil demand is -0.48 for new maximum prices, while it is considerably lower at -0.18 for other price changes below these maximum levels. If consumers respond to future price movements the way that they have to past ones, oil demand may become much more responsive to prices than they have in the previous few decades. In short, they appear to respond much more to major structural price shocks than to price volatility. The average annual inflation-adjusted U.S. price for all petroleum products in 2010 is virtually at the 2008 annual level, its current maximum level. Sustained price movements beyond that level may provide sufficient incentives for transforming the capital stock to use considerably less oil in the future.

When working with levels, differences or responses that are either symmetric or asymmetric, there are opportunities to choose excellent equations with relatively low errors as well as far inferior ones with relatively high errors. Choosing the basic form is only half the battle. Considerable testing of an equation's forecasting capacity remains an important challenge that will help improve oil demand and market projections.

References

Adeyemi, Olutomi and Lester C Hunt (2007). “Modeling OECD Industrial Energy Demand: Asymmetric Price Responses and Energy-Saving Technical Change”, *Energy Economics* 29(2): 693-709.

Beenstock, M. and P. Willcocks (1981). “Energy Consumption and Economic Activity in Industrialised Countries,” *Energy Economics*, 3, 225-232.

Beenstock, M. and P. Willcocks (1983). “Energy and Economic Activity: A Reply to Kouris,” *Energy Economics*, 5, 212

Bentzen, Jan and Tom Engsted (1996). “On the Estimation of Short- and Long-Run Elasticities in U.S. Petroleum Consumption: Comment,” *Southern Economic Journal* 62(3): 783-787.

Breunig, Robert V. (2008). “Should Single-equation Dynamic Gasoline Models Include Moving Average Terms?” Social Science Research Network Discussion Paper.

Dahl, Carol A. and Thomas Sterner (1991). "Analyzing Gasoline Demand Elasticities: A Survey," *Energy Economics* 13 (July): 203-210.

Dargay, J. M. (1992). “The Irreversible Effects of High Oil Prices: Empirical Evidence for the Demand for Motor Fuels in France, Germany and the UK,” in Hawdon, D. (ed.), *Energy Demand: Evidence and Expectations*, Surrey University Press, Guildford, UK.

Dargay, Joyce M. and Dermot Gately (1995a). “The Imperfect Price-Reversibility of Non-Transport Oil Demand in the OECD,” *Energy Economics*, 17, 59-71.

Dargay, Joyce M. and Dermot Gately (1995b). “The Response of World Energy and Oil Demand to Income Growth and Changes in Oil Prices,” *Annual Review of Energy*, 20, 145-178.

Dargay, Joyce M. and Dermot Gately (1997). The Demand for Transportation Fuels: Imperfect Price Reversibility? *Transportation Research* B, 31, 71-82.

Dargay, Joyce M. and Dermot Gately, (1994). “Oil Demand in the Industrialized Countries”, *Energy Journal*, Vol. 15, Special Issue, pp. 39-67.

Dargay, Joyce M., and Dermot Gately. "World Oil Demand's Shift Toward Faster Growing And Less Price-Responsive Products And Regions." Energy Policy 38, no. 10 (2010): 6261-6277.

Dees, Stephane, Pavlos Karadeloglou, Robert K. Kaufmann, and Marcelo Sanchez (2007). “Modelling the World Oil Market: Assessment of a Quarterly Econometric Model,” *Energy Policy* 35, pp. 178–191.

Fischer, Carolyn, Evan Herrnstadt, and Richard Morgenstern (2008). “Understanding Errors in EIA Projections of Energy Demand,” Resources for the Future Discussion Paper 07-54, November.

Galli, Rosanna (1998). “The Relationship Between Energy and Income Levels: Forecasting Long Term Energy Demand in Asian Emerging Countries,” *The Energy Journal* 19(4): 85-105.

Gately, Dermot and Hillard G. Huntington (2002). “The Asymmetric Effects of Changes in Price and Income on Energy and Oil Demand.” *The Energy Journal* 23(1): 19-55.

Goodwin P., J. Dargay, and M. Hanly (2004). “Elasticities of Road Traffic and Fuel Consumption with Respect to Price and Income: A Review,” *Transport Reviews*, 24 (3): 275-292.

Graham, D. and S. Glaister (2004). “Road Traffic Demand Elasticity Estimates: A Review,” *Transport Reviews*, 24 (3): 261-274.

Griffin, James M., and Craig T. Schulman (2005). “Price Asymmetry in Energy Demand Models: A Proxy for Energy Saving Technical Change.” *The Energy Journal* 26(2): 1-21.

Hendry, D.F. (1995). Dynamic Econometrics. Oxford; New York: Oxford University Press.

Hughes, Jonathan E., Christopher Knittel, and Daniel Sperling (2008). “Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand,” *The Energy Journal* 29(1).

Hunt, Lester C., Guy Judge, and Yasushi Ninomiya (2003). “Underlying Trends and Seasonality in UK Energy Demand: A Sectoral Analysis,” *Energy Economics*, 25, 93- 118.

Hunt, Lester C. and Yasushi Ninomiya (2003). “Unravelling Trends and Seasonality: A Structural Time Series Analysis of Transport Oil Demand in the UK and Japan,” *The Energy Journal*, 24, 63-96.

Huntington, Hillard G. (2010a). “Short- and Long-Run Adjustments in U.S. Petroleum Consumption,” *Energy Economics*, January, 32: 63-72.

Huntington, Hillard G. (2010b). “Oil Demand and Technical Progress,” Applied Economics Letters, 17 (18): 1747 — 1751.

Jones, Clifton T. (1993). “A Single-Equation Study of U.S. Petroleum Consumption: The Role of Model Specification." *Southern Economic Journal*, April, 59(4): 687-700.

Kennedy, Peter (2003). *A Guide to Econometrics*, 5th Edition, Cambride, MA: MIT Press.
O'Neill, Brian C., and Mausami Desai (2005). "Accuracy of Past Projections of U.S. Energy Consumption. Energy Policy 33(8): 979–993.
Ryan, David L. and André Plourde (2002). "Smaller and Smaller? The Price Responsiveness of Non-Transport Oil Demand," *Quarterly Review of Economics and Finance*, 42: 285-317.
Smith, James L. (2009). "World Oil: Market or Mayhem?" *Journal of Economic Perspectives*, 23(3): 145–64.
Smith, James L. (2005). "Inscrutable OPEC? Behavioral Tests of the Cartel Hypothesis," *The Energy Journal*, 26(1): 51-82.
Thoma, Mark and Tim Duy (1998). "Modeling and Forecasting Cointegrating Relationships: Some Practical Experience," *Journal of Economics and Business*, May-June, 50 (3): 291-307.
U.S. Bureau of the Census, *Current Population Reports*, http://www.census.gov/.
U.S. Bureau of Economic Analysis, *National Income and Product Accounts*, http://www.bea.gov/.
U.S. Bureau of Labor Statistics, *Producer Price Indexes*, http://www.bls.gov/ppi/.
U.S. Energy Information Administration (2010). *Annual Energy Review 2009*, Washington, D.C., August, Report No. DOE/EIA-0384(2009).
U.S. Energy Information Administration (2008). *Annual Energy Outlook Retrospective Review: Evaluation of Projections in Past Editions (1982-2008).* DOE/EIA-0640(2008), Washington, D.C.
Winebrake, J.J. and D. Sakva (2006). "An Evaluation of Errors in US Energy Forecasts: 1982–2003," *Energy Policy* 34: 3475–3483.

[1] These price components are also not cumulative summations. Specifically, they are the price variable relative to the maximum price level, as used by Huntington (2010).

[2] The Johansen tests are often used more than the Engle-Granger tests for cointegration when there more than two variables, as discussed by Kremers et al (1991).

[3] For example, see the exchange between Bentsen and Engsted (1996) and Jones (1996).

[4] Moreover, simultaneous equations will produce less bias in the price elasticity only as long as the researcher can find reliable instruments for supply-side factors. Most oil market analysts think that OPEC decisions and actions have dominated the oil supply conditions over the last several decades. Unfortunately, there exists little agreement on a credible and empirically verifiable theory about cartel behavior (e.g., see Smith, 2005).

[5] Although OLS specifications are widely used in policymaking, they do have their shortcomings (Breunig 2008).

[6] The levels specification in Eq. 1 can be transformed into an error-corrections specification by subtracting Z_{t-1} from both sides and adding $\alpha_0 X_{t-1} - \alpha_0 X_{t-1}$ to the right side (see Hendry, 1995, chapter 7).

[7] Technically, these studies constrain the lagged income level by the coefficients on the current income (GDP) and lagged consumption variables. This approach produces very similar responses to lagged income as when the coefficient is left unconstrained, as done in the current study. Dargay and Gately (2010) begin by including all variables as lags but reject the full asymmetric function on goodness-of-fit criterion rather than on the equation's backcasting accuracy. This decision may create problems because their goal appears to be projections rather than hypothesis testing.

[8] Out-of-sample backcasting is a common practice in policy modeling. For example, see the oil price backcasting in the study by Dees *et al* (2007).

[9] A full reporting of the estimated coefficients, the equation's explanatory power and tests on the error term are available from the author upon request.