The Long-Run Forecasting of Energy Prices Using the Model of Shifting Trend

Stanislav Radchenko*

University of North Carolina at Charlotte

This version: February, 2005

Abstract

This paper constructs long-term forecasts of energy prices using a reduced form model of shifting trend developed by Pindyck (1999). A Gibbs sampling algorithm is developed to estimate models with a shifting trend line which are used to construct 10-period-ahead and 15-period ahead forecasts. An advantage of forecasts from this model is that they are not very influenced by the presence of large, long-lived increases and decreases in energy prices. The forecasts form shifting trends model are combined with forecasts from the random walk model and the autoregressive model to substantially decrease the mean forecast squared error compared to each individual model.

Keywords: energy forecasting, oil price, coal price, natural gas price, shifting trends model

JEL Classification Numbers: C53

*Department of Economics, University of North Carolina at Charlotte, NC 28223. E-mail address: sradchen@email.uncc.edu.

I would like to thank Robert S. Pindyck for kindly sharing his data with me.

1 Introduction

Developing models for accurate long-term energy price forecasting is an important problem because these forecasts should be useful in determining both supply and demand of energy. On the supply side, long-term forecasts determine investment decisions of energy-related companies. A study of trends in the USA natural gas market conducted by Energy Information Administration indicates that one of the reasons for high growth of natural gas prices in 2000 was a decline in investments in exploration and production.[1] It is argued that a decline in investment activity in late 90's was induced by a history of low energy prices. The investment decline of such scale would not occur if there were accurate and reliable long-run forecasts of natural gas prices.[2]

On the demand side, investments in physical capital and durable goods depend on price forecasts of a particular energy type. Forecasting long-run trend movements in energy prices is very important on the macroeconomic level for several developing countries because energy prices have large impacts on their real output, the balance of payments, fiscal policy, etc. In this paper, I use Bayesian methods to estimate the model with a time-varying trend to construct long-run forecasts of energy prices: crude oil price, natural gas price, and bituminous coal price.

The long-term forecasting of energy prices is a challenging problem. The forecasts of energy price (oil, natural gas, and coal) are constructed on regular basis by the Department of Energy,[3] but performance of these forecasts is criticized.[4] There are several structural approaches to forecasting the oil market and energy prices in terms of a supply-demand equilibrium schedule.[5] These approaches proved to be challenging because they require the

[1]The study points out that oil and gas investment in exploration and production from 1990 through 1996 averaged $15 billions annually in real 1999 dollars, as compared with investment in excess of $30 billions before 1986.

[2]Several papers have analyzed the actual and forecast levels of oil exploration and extraction in U.K. Continental Shelf. Kemp and Kasim (2003) point out that the majority of forecast results have been inaccurate because of inaccurate resource price forecasts.

[3]One may check the page http://www.eia.doe.gov/oiaf/petgas.html for the petroleum and natural gas forecasts and http://www.eia.doe.gov/oiaf/coal.html for coal forecasts.

[4]See Lynch (2000).

[5]A recent example of structural approach is the paper by Dees *et al.* (2004) who evaluate a quarterly

modeling of oil supply of a cartel (OPEC) and several non-OPEC producers, the modeling of investment decisions and other variables influencing supply and demand, the reaction of oil price to changes in oil market conditions and OPEC behavior, etc.

Huntington (1994) shows that the forecasting performance of ten structural models existing in 1980-1981 was problematic. The errors in structural models were due to such factors as exogenous GNP assumptions, resource supply conditions outside OPEC, and demand adjustments to price changes. Lynch (2002) arrives to similar conclusions by comparing the theory and practice of oil supply forecasting. In their retrospective study, Koomey et al. (2003) point to factors like technological innovation and human behavior for inaccuracy of oil forecasts. Tang and Hammoudeh (2002) show that omission of market participants' expectations contributes to forecasting errors. Gately (1995) shows that the projection of oil prices in structural models depends crucially on assumptions about the model parameters.

Pindyck (1999) points out that structural models are not always useful for long-run forecasting, but they are better suited at providing understanding of the causes of short- or intermediate-run fluctuations of prices and other variables.[6] The author argues that the dynamics of real energy prices is mean-reverting to trend lines with slopes and levels that are shifting unpredictably over time. The hypothesis of shifting long-term trend lines was statistically tested by Benard *et al.* (2004). The authors find statistically significant instabilities for coal and natural gas prices.[7] Instead of fixing all modeling issues in a structural model to construct reliable long-term forecasts, I continue the research of energy prices in the framework of continuously shifting levels and slopes of trend lines started by Pindyck (1999).

model of oil market that includes a pricing rule and demand and supply schedules based on the forecasting performance. A price rule relates real oil price to measures of OPEC behavior and market indicators of supply/demand balance.

[6]In view of the problems with energy forecasts from structural models, Koomey et al. (2003) argue that the accuracy of long-range energy forecasts is not important, but the ability of forecasts to describe the consequences of police shocks is important.

[7]In a different framework, McAvinchey and Yannopoulos (2003) present evidence for structural changes in energy prices for Germany and UK. In their paper, the parameter constancy tests and diagnostic specification statistics indicate the superiority of the model with structural changes, while results from root mean forecast squared error give mixed results for Germany and UK.

The examined model offers both parsimonious approach and perspective on the developments in energy markets. Using the model of depletable resource production, Pindyck (1999) argued that the forecast of energy prices in the model is based on the long-run total marginal cost. Because the model of a shifting trend is based on the competitive behavior, one may examine deviations of oil producers from the competitive behavior by studying the difference between actual prices and long-term forecasts. To construct the long-run forecasts of energy prices, I modify the univariate shifting trends model of Pindyck (1999). I relax some assumptions on model parameters, the assumption of white noise error term, and propose a new Bayesian approach to estimate the model with autocorrelation. To improve the long-term forecasts, I suggest to combine forecasts from the shifting trend model, random walk model, and an autoregressive model.

The main results may be summarized as follows. The shifting trend model offers an alternative approach to construct long-run forecasts of energy prices which give additional insights about the competitive nature of energy markets. The constructed forecasts seem to be unaffected by the presence of large long-lived increases and decreases in energy prices. Also, these forecasts may be combined with the forecasts from random walk model and univariate autoregressive model to substantially decrease the mean squared forecast error (MSFE) of energy forecasts.

2 Econometric Model for Forecasting Energy Prices

Even though the energy price forecasting received a lot of attention in the private sector, there are only few reduced form models in the literature for forecasting crude oil, natural gas, or bituminous coal. Ye *et al.* (2002) propose a short-term monthly forecasting model of WTI crude oil spot price using OECD inventory levels. To form forecasts, they use an autoregressive distributed lag model. Zeng and Swanson (1997) examine the predictive accuracy of various econometric models for the crude oil price using daily futures prices. Chacra (2002) builds a quarterly forecasting model to examine the relationship between world oil prices and components of CPI-energy (gasoline, heating oil, natural gas), but he constructs only one-quarter and two-quarter ahead forecasts.

Tang and Hammoudeh (2002) examine the dynamics of oil price based on a version of the target zone model. The authors consider a short-term forecasting experiment and argue that the target zone model has good forecasting ability and the inclusion of expectation effect of the market participants decreases forecasting errors. McAvinchery and Yannopoulos (2003) use the ECM with and without structural changes to construct forecast up to a 5-year horizon.

In this paper, I use the model developed by Pindyck (1999) who proposes a reduced form model with time-varying parameters for the analysis of stochastic properties (persistence) of energy prices. The author indicates that the model may be applied to forecasting. This conjecture is supported by Benard et al. (2004) who present evidence in favor of the class of time-varying parameter models suggested by Pindyck (1999).

Following Pindyck (1999), I use the following univariate model with shifting trends for the long-term forecasting of energy prices:

$$p_t = \gamma p_{t-1} + b_1 + b_2 t + \phi_{1t} + \phi_{2t} t + u_t \quad (1)$$

$$\phi_{1t} = c\phi_{1,t-1} + v_{1t} \quad (2)$$

$$\phi_{2t} = s\phi_{2,t-1} + v_{2t} \quad (3)$$

$$u_t = \psi u_{t-1} + e_t \quad (4)$$

where ϕ_{1t} and ϕ_{2t} are unobservable state variables, p_t is a real price of crude oil, natural gas, or coal. The distribution of the error terms e_t, v_{1t} and v_{2t} is multivariate normal, e_t is uncorrelated with v_{1t} and v_{2t}, in particular $e_t \sim N(0, \omega^2)$, $v_{1t} \sim N(0, \sigma_1^2)$, $v_{2t} \sim N(0, \sigma_2^2)$. When the parameter ψ is set to zero, one obtains a model examined by Pindyck (1999). The assumption of autocorrelation in the error term is based on the preliminary analysis of the error terms in white noise model.

Despite a parsimonious form of the model, it proved to be difficult to estimate this mode for energy prices using the maximum likelihood estimation approach. This model has problems with convergence and estimation of long-term trends for coal and natural gas prices. Pindyck (1999) suggests that these problems may be attributed to the nonstationarity of the unobservable states. Thus the Bayesian analysis seems to be a better alternative for model estimation because the results in Bayesian framework are less influenced by whether variables are stationary or not.

To estimate different energy price models, Pindyck (1999) sets a parameter s in equation (3) to one for coal model, sets a parameter c to zero for natural gas model, and exclude the unobserved state ϕ_{1t} from the estimation of gas model. In all models, authors set the parameter b_2 to zero.[8] Similar to previous papers, I impose a zero restriction on the parameter b_2 for all energy models. However, I do not restrict a parameter s to one or a parameter c to zero for any model and do not exclude the unobserved state ϕ_{1t} from estimation.

To check how well shifting trend models perform, I compare mean forecast squared errors for shifting trend models with the random walk model and univariate autoregressive models. Therefore, I consider three forecasting models in total.

It is well known that combining several forecasts can yield a mean square forecast error lower than that of a single forecast.[9] I construct linear combinations of forecasts to check how well they perform relative to single forecasts. Combination forecasts are constructed using the following formula:

$$\begin{aligned} \hat{y}^c_{t+h} &= \sum_{m=1}^{M} k_{m,h,t} \hat{y}_{t+h,m} \\ k_{m,h,t} &= \frac{(1/MFSE_{m,h,t})^w}{\sum_{j=1}^{M}(1/MFSE_{j,h,t})^w} \end{aligned} \tag{5}$$

where $\hat{y}^c_{t+h}$ denotes a constructed combination forecast, $\hat{y}_{t+h,m}$ denotes the forecasts of the

[8] Benard et al. (2004) impose the same restrictions.

[9] One may check papers by Wright (2003) or Stock and Watson (2003) among others.

considered models ($m = 1, 2, 3$), h is the forecast horizon (h=10 years, 15 years), $k_{m,h,t}$ denotes a weight for each model and forecasting horizon. In computation of weights, I set the coefficient w equal to 1 which implies that the weight of model is chosen as inversely proportional to its MFSE. I also check the performance of combination forecast when the coefficient $w = 5$ which implies that the best performing model receives more weight.[10]

2.1 Prior distributions

I use conjugate prior distributions to simplify computations in the Gibbs sampling algorithm. The choice of prior hyperparameters for univariate models is presented in Table 1. Usually, the choice of hyperparameters is not trivial. On one hand, a researcher needs to incorporate as much subjective information as possible, but on the other hand a researcher should be careful not to introduce information that is not available.

Even though the previously obtained estimates for the price autoregressive parameter (γ) are not very high, my prior expectations are that the autoregressive parameters for energy prices should be close to one for all energy commodities.[11] So, I set the prior mean for parameters to one and set prior variance to 0.2 for the oil and natural gas models and 0.1 for the coal model. The chosen values of prior variance are high relative to values of prior mean and distributions for parameters γ are not very informative. Based on the previously reported estimates for an autoregressive parameters of ϕ_{1t} and ϕ_{2t} processes, I impose a persistent and tight prior distribution for the parameters c and s with the prior mean 0.9 for the coal and natural gas models and 0.95 for the crude oil model. The prior variance is 0.05 for crude oil model and 0.1 for the gas and coal models. The prior distribution for the parameter ψ is set to be uninformative reflecting a lack of prior knowledge about this parameter.

The most difficult and important in practical implementation of the Gibbs sampling is the specification of hyperparameters for variances, because it seems that these hyperparameters

[10] I follow Marcellino (2002) in setting the value for w.

[11] The estimated parameters in Pindyck (1999) are not very high: 0.804, 0.014, 0.687 for oil, coal, and natural gas price respectively.

influence the convergence of the Gibbs sampling algorithm. Following the literature, I assume Inverted-Wishart distribution for the variance parameters. The hyperparameters and degrees of freedom are presented in Table 1. I set the prior distribution for the variances of oil and coal models tight, but the prior distribution for natural gas model is not informative. The difference in the prior for variance is consistent with the estimates of unobserved states and their variance in Pindyck (1999).

2.2 The Gibbs sample algorithm

The Gibbs sampling algorithm for estimation of the univariate models with shifting trends is presented in Appendix. The main idea in drawing parameters γ, b_1, ψ and variances is to transform the state space model into liner regression models from which it is easy to derive posterior distributions for the parameters of interest. Once the model is transformed into a linear form, I use a standard approach for linear models with conjugate prior distributions to draw parameters. To handle the presence of autocorrelated error terms, I follow the approach of Kim and Nelson (1998).

A more difficult part is the draws of unobserved states ϕ_{1t} and ϕ_{2t}, continuously changing levels and slopes of linear time trend. The draw of states is done separately. First, I formulate the state space representation to draw unobserved levels ϕ_{1t}. Then I formulate another state space representation to draw unobserved slopes ϕ_{2t}. To understand the derivation of the second state space model in Appendix, note that equation (3) can be written as:

$$\phi_{2t}t = \left(\frac{t}{t-1}\right) s\phi_{2,t-1}(t-1) + tv_{2t}. \tag{6}$$

Next, equation (6) is divided by T to prevent the variance-covariance of the error term from exploding in the implementation of the algorithm:

$$\phi_{2t}\frac{t}{T} = \left(\frac{t}{t-1}\right) s\phi_{2,t-1}\frac{(t-1)}{T} + \frac{t}{T}v_{2t}. \tag{7}$$

Using definitions in Appendix, equation (7) is written as

$$\tilde{\phi}_{2t} = \tilde{s}_t \tilde{\phi}_{2,t-1} + \tilde{v}_{2t} \tag{8}$$

which is the expression used in the second state-space representation in the Appendix. The state space representation that I have applied is not unique. I have considered several alternative state-space representations in the implementation of Gibbs sampling algorithm and I have chosen the model with the best convergence of unknown parameters and unobserved states.

In the implementation of the Gibbs sampling algorithm, I use 17,000 draws and discard first 2,000 draws. To insure the convergence of the algorithm, I restrict parameters γ, c and s to be less than 1.2 in Gibbs sampling. The draws of parameters are used to find estimates of the parameters (posterior means) and their standard deviations.The convergence of the algorithm for some parameters may be slow and depends on the energy commodity.

3 Data and Results

I estimate shifting trend models of real prices of crude oil and bituminous coal using the sample over 127-year period 1870-1996, while the sample for natural gas prices is for the period 1919 - 1996. The data set is the same as used by Pindyck (1999).[12] The time series of prices is deflated to 1967 dollars using the Wholesale Price Index for all commodities through 1970, and the Produce Price Index from 1970 onwards.[13] Even though it is likely that quarterly data are better suited for the analysis of short-term changes, annual data should reflect long run developments in energy sector more accurately which is what I am interested in.[14] Also, using the same sample allows to compare results in this paper with results in Pindyck (1999) and Benard et al. (2004).

[12]I would like to thank Robert S. Pindyck for kindly sharing his data with me.
[13]Check Pindyck (1999) for details on data transformation.
[14]McAvinchey (2002) examines the combination of annual data and quarterly data in the analysis of energy market changes.

To get an idea about the dynamics of energy prices, I present the graphs of the series in Figure 1 for the period 1919 - 1996. One may notice from the figure that oil and natural gas prices exhibit similar dynamics, but the behavior of coal price seems to be different from the dynamics of crude oil and natural gas prices. This is also confirmed by correlation coefficients among the variables: the correlation between crude oil and natural gas prices is 0.82, while correlation between oil and coal prices is 0.65 and coal and natural gas prices is 0.46.

3.1 Individual Model Results

In Figure 2, I present the estimates of the slopes and levels of unobserved trend for the full sample which provide additional information about the long-run developments in energy markets. The estimated slopes of long-term trend of oil prices indicate that real oil prices declined by 1-2% per year before 1895. The trend seems to stabilize around 1900, but it resumes falling after 1905. The real oil price trend declines slightly for the period 1925-1950 and remains virtually unchanged for the period 1950-1975. From 1975 onward, oil prices experience slight upward trend. The trend dynamics for coal prices is also interesting. The real coal price has a downward trend for the period 1870-1920 with the biggest decline reaching 1.5% per year, but it start to increase after 1920. The estimates of trend slope of natural gas price indicate a rapid decline in natural gas prices for the period 1920-1950 (up to 10% in the beginning of the sample), but natural gas price experiences a slight upward trend afterward. Notice also that trend for all energy prices seem to stabilize for the last ten years or so. These results agree and extend the results of Cashin and McDermott (2002) who point out that real commodity prices have declined about 1 percent per year over the last 140 years, but price variability is very large relative to a trend with prices changing as much as 50 percent per year.[15]

The constructed forecasts are depicted in Figures 3 - 5. I construct four long-term

[15] Cashin and McDemott (2002) examine the behavior of the industrial commodity-price index of The Economist.

forecasts for each energy series for the following periods: (i) 1986 - 2011, (ii) 1981-2011, (iii) 1976-2011, (iv) 1971 - 2011. I expect that long-run forecasts of trend lines may be very different from actual prices for some periods of time because of high variability in energy prices. Not surprisingly, the oil price model fails to predict oil price increases in 1973-75 and 1979-80. It predicts the long-term real oil price level in the range of 1974-75 oil prices. An advantage of models with shifting trend line is that they are not very influenced by the presence of large, long-lived increases and decreases in energy prices. Cashin and McDermott (2002) argue that this quality is important for any forecasting model of commodity prices. They point out that it is important not to be misled by "booms" and "slumps" because they are unlikely to indicate major changes in long-run prices. Notice that all the forecasts for the crude oil and coal models converge to approximately the same value in the long-run indicating the robustness of long-term forecasts.

Similar results are obtained for coal forecasts. The model fails to predict the increase in coal prices in 1969-75, but forecasts since 1976 are satisfactory because they all correctly predict the downward decline in coal prices.

A different behavior is observed for forecasts of the natural gas price. While the crude oil and coal forecasts indicate the downward trend in prices, the gas model forecasts indicate the upward trend in price. The natural gas model is the only model that predicts the commodity price increase in 1973-83. Notice that the model does not predict a price decline in late eighties and nineties and predicts the real natural gas price in future at approximately the level of 1983 prices.

The parameter estimates for all energy models obtained using the full sample are presented in Table 2. Even though prior distributions for autoregressive parameters γ are centered at very persistent values, the estimates of γ for oil and coal models are somewhat low, 0.305 and 0.197 respectively. The parameter estimate for the natural gas model is higher with $\hat{\gamma} = 0.761$. The obtained results qualitatively are similar to the estimates of Pindyck (1999). The only difference is the magnitude of the estimate of the autoregressive parameter for the oil model. Pindyck (1999) obtains a more persistent estimate of autoregressive parameter for oil model (0.804), while the estimate that I obtain is not persistent.

The estimates of autoregressive parameters for unobserved states ϕ_{1t} and ϕ_{2t} are not as high as the estimates reported in Pindyck (1999) implying lower persistence of these states. However, I must note that there is high uncertainty about parameter estimates which is reflected in high standard errors. Also, Benard et al. (2004) report estimates of these parameters that are lower than the estimates reported by Pindyck (1999) but higher than the estimates reported here.

The difference in the estimates of variance for different energy models is significant. The parameter ω is estimated to be 4.25 for the natural gas model, 0.04 for the oil model, and only 0.0047 for the coal model. The variance estimates for the coal and natural gas models are different from the prior means, but the variance estimate for the oil model is close to the prior mean. It implies the low information of likelihood function with respect to the variance and the importance of prior information in estimation of the oil model.

3.2 Combination of Forecasts

By construction, the long-term forecasts of energy prices from shifting trend models do not have any mechanism to model and forecast the level of OPEC short-term or long-term cooperation (or any cooperation among producers of energy products) to maintain higher than competitive prices. That is why the model seems to predict a decline in energy prices more often than is practically reasonable. To decrease this problem, I suggest to combine forecasts from shifting trends model with the random walk model.

The combination of forecasts from these two models is likely to produce a superior forecast. First, Wright (2003) points out that it is a part of folk wisdom that combination of forecasts from any models is likely to produce a better forecast. Second, while a shifting trend model tends to predict declines in energy prices to competitive marginal cost levels, the random walk model implies that market conditions will stay the same and price levels in future will be determined by the same demand and supply factors. This implication is not very reasonable and the random walk forecast is not appealing for long-term forecasting if it is used alone. But the random walk forecast may be used as an approximation for the same level of cooperation among produces in short run to keep energy prices at the currently observed level. Therefore, a random walk forecast can be thought of as a proxy for the future cooperation among producers and combination of this forecast with the shifting model forecast should result in the forecast improvement. I also check whether forecasts from univariate autoregressive models are useful in forecast combination. In estimation of AR models, I use the Akaike Information Criteria to select the lag length.

In Table 3, I compare the MSFE of the shifting trend model with the random walk and autoregressive models. The obtained results are mixed and do not indicate a clear "winner". The shifting trend models outperform other models in three cases (natural gas forecasts at 10-period-ahead and 15-period-ahead horizons and crude oil price forecast at 10-period-ahead horizon); the univariate AR model has the lowest MSFE in two experiments (coal price forecasts at 10-period-ahead and 15-period-ahead horizons); the random walk model seems to outperform marginally in only one case (crude oil price forecast at 15-period-ahead horizon).

I form two combination forecasts. The first combination forecast is constructed by setting the parameter $w = 1$ in equation (5). This implies that the weight of each model is chosen inversely proportional to MSFE. The second combination forecast is constructed by setting the parameter $w = 5$ in equation (5). This implies that the best performing model receives the biggest weight.

As expected, combining forecasts of three models results in substantial forecast improvement. The average decline in MSFE of the combination forecast 1 over the best performing model is 25%, while for the combination forecast 2 the average decline in MSFE is 12%.

4 Conclusions

In this paper I propose a new Bayesian framework to estimate and construct the long-term forecasts of energy prices using the shifting trend model of energy prices proposed by Pindyck

(1999). An advantage of forecasts from this model is that they are not very influenced by the presence of large, long-lived increases and decreases in energy prices and produce robust long-term forecasts. Using the annual data for 1870 - 1996, I show that forecasts from this model may be combined with forecasts from a random walk model and an autoregressive model to substantially decrease mean forecast squared error. The estimates of slopes of trend provide additional information about long-run developments in energy markets.

Appendix A: The Gibbs sampling algorithm

Let Θ^i denote the ith draw of all model parameter, $\Theta_i = (\gamma^i, b_1^i, c^i, s^i, \omega^{2i}, \sigma_1^{2i}, \sigma_2^{2i})$, and Φ_1^i be the ith draw of the first unobserved states and Φ_2^i be the ith draw of the second unobserved states. Given the draw of parameters Θ^i, the following Gibbs sampling algorithm can be used to generate the draw of parameters Θ^{i+1} and states $\Phi_1^{(i+1)}$ and $\Phi_2^{(i+1)}$:

1. The draw of parameters $\gamma^{(i+1)}$. The model (1) - (3) can be written:

$$\tilde{P} = P^*_{-1}\Gamma' + E$$

where $\tilde{P} = (1-\psi^i L)(P - Xb_1^i - \Phi_1^i - Z^i)$, $P^*_{-1} = (1-\psi L)P_{-1}$; $P = [\, p_1 \quad p_2 \quad \cdots \quad p_T\,]'$ and $P_{-1} = [\, p_0 \quad p_1 \quad \cdots \quad p_{T-1}\,]'$ are $T \times 1$ vectors, X is a $T \times 1$ vector of ones, $\Phi_1^i = [\, \phi_{11}^i \quad \phi_{12}^i \quad \cdots \quad \phi_{1T}^i\,]'$ is a $T \times 1$ vector of the first unobserved state, $Z^i = [\, z_1^i \quad z_2^i \quad \cdots \quad z_T^i\,]'$ is a $T \times 1$ vector of the second unobserved state multpiled by the time trend, $z_t^i = \phi_{2t}^i t$. The parameters γ are then drawn from:

$$\begin{aligned} \gamma^{(i+1)} &\sim N(\tilde{\gamma}, V_\gamma) \\ \tilde{\gamma} &= V_\gamma \left(V_{\gamma 0}^{-1}\gamma_0 + \left((\omega^{2i})^{-1} P^*_{-1}{}' P^*_{-1} \right) \hat{\gamma} \right) \\ V_\gamma &= \left(V_{\gamma 0}^{-1} + \left((\omega^{2i})^{-1} P^*_{-1}{}' P^*_{-1} \right) \right)^{-1} \end{aligned} \tag{A-1}$$

where $\hat{\gamma} = (P^*_{-1}{}'P^*_{-1})^{-1}P^*_{-1}{}'\tilde{P}$, γ_0 and $V_{\gamma 0}$ are the prior mean and variance of γ.

2. The draw of parameters $b_1^{(i+1)}$. The model (1) - (2) can be written:

$$\bar{P} = X^* b_1 + E$$

where $\bar{P} = (1 - \psi^i L)(P - P_{-1}\gamma^{(i+1)} - \Phi_1^i - Z^i)$, $X^* = (1 - \psi^i)X$. The parameter b_1 is then drawn from the following distribution:

$$
\begin{aligned}
b_1^{i+1} &\sim N(\tilde{\beta}_1, V_\beta) \\
\tilde{\beta}_1 &= V_\beta \left(V_{\beta 0}^{-1}\beta_0 + \left((\omega^{2i})^{-1}(X^{*\prime}X^*) \right) \hat{\beta} \right) \\
V_\beta &= \left(V_{\beta 0}^{-1} + \left((\omega^{2i})^{-1}(X^{*\prime}X^*) \right) \right)^{-1}
\end{aligned}
\tag{A-2}
$$

where $\hat{\beta}_1 = (X^{*\prime}X^*)^{-1}X^{*\prime}\tilde{P}$, β_0 and $V_{\beta 0}$ are the prior mean and variance of β_1.

3. The parameter ψ^{i+1} is drawn using the following model

$$U^{i+1} = U_{-1}^{i+1}\psi + E$$

where $U^{i+1} = P - P_{-1}\gamma^{(i+1)} - Xb_1^{i+1} - \Phi_1^i - Z^i$ and U_{-1}^{i+1} is the matrix consisting of the first lag of U^{i+1}. The parameter ψ is drawn from the following distribution:

$$
\begin{aligned}
\psi^{i+1} &\sim N(\tilde{\psi}, V_\psi) \\
\tilde{\psi} &= V_\psi \left(V_{\psi 0}^{-1}\psi_0 + \left((\omega^{2i})^{-1}(U_{-1}^{i+1\prime}U_{-1}^{i+1}) \right) \hat{\psi} \right) \\
V_\psi &= \left(V_{\psi 0}^{-1} + \left((\omega^{2i})^{-1}(U_{-1}^{i+1\prime}U_{-1}^{i+1}) \right) \right)^{-1} \\
\hat{\psi} &= (U_{-1}^{i+1\prime}U_{-1}^{i+1})^{-1}U_{-1}^{i+1\prime}U^{i+1}
\end{aligned}
\tag{A-3}
$$

4. The draw of $\omega^{2(i+1)}$ is done using the usual Inverted Wishart (IW) formula:

$$\omega^{2(i+1)} \sim IW(SSR, df) \tag{A-4}$$

$$SSR = E^{(i+1)\prime}E^{(i+1)}$$

where $E^{(i+1)} = U^{i+1} - U_{-1}^{i+1}\psi$, $df = T + 3$.

5. The draw of $c^{(i+1)}$ is done from the following distribution:

$$\begin{aligned} c^{(i+1)} &= N(\tilde{c}, V_c) \\ \tilde{c} &= V_c\left(V_{c0}^{-1}c_0 + \left((\sigma_1^{2i})^{-1}(\Phi^{i\prime}_{1(-1)}\Phi^{i}_{1(-1)})\right)\hat{c}\right) \\ V_c &= \left(V_{c0}^{-1} + \left((\sigma_1^{2i})^{-1}(\Phi^{i\prime}_{1(-1)}\Phi^{i}_{1(-1)})\right)\right)^{-1} \end{aligned} \tag{A-5}$$

where $\hat{c} = (\Phi^{i\prime}_{1(-1)}\Phi^{i}_{1(-1)})^{-1}\Phi^{i\prime}_{1(-1)}\Phi^{i}_{1}$, $\Phi^{i}_{1(-1)}$ is a $T \times 1$ vector that denotes the first lag of the matrix Φ^{i}_{1}, $\Phi^{i}_{1(-1)} = [\,\phi^{i}_{10} \quad \phi^{i}_{11} \quad \cdots \quad \phi^{i}_{1T-1}\,]'$; and c_0 and V_{c0} are the prior mean and variance of c.

6. The draw of $s^{(i+1)}$ is done from the following distribution:

$$\begin{aligned} s^{(i+1)} &= N(\hat{s}, V_s) \\ \tilde{s} &= V_s\left(V_{s0}^{-1}s_0 + \left((\sigma_2^{2i})^{-1}(\Phi^{i\prime}_{2(-1)}\Phi^{i}_{2(-1)})\right)\hat{s}\right) \\ V_s &= \left(V_{s0}^{-1} + \left((\sigma_2^{2i})^{-1}(\Phi^{i\prime}_{2(-1)}\Phi^{i}_{2(-1)})\right)\right)^{-1} \end{aligned} \tag{A-6}$$

where $\hat{s} = (\Phi^{i\prime}_{2(-1)}\Phi^{i}_{2(-1)})^{-1}\Phi^{i\prime}_{2(-1)}\Phi^{i}_{2}$, $\Phi^{i}_{2(-1)}$ is a $T \times 1$ vector that denotes the first lag of the vector Φ^{i}_{2}, s_0 and V_{s0} are the prior mean and variance of s.

7. The draw of $\sigma_1^{2(i+1)}$ and $\sigma_2^{2(i+1)}$ is done using Inverted Wishart distribution:

$$\sigma_1^{2(i+1)} \sim IW(SSR_1, df), \quad SSR_1 = V_1^{(i+1)\prime}V_1^{(i+1)} \tag{A-7}$$

$$\sigma_2^{2(i+1)} \sim IW(SSR_2, df), \quad SSR_2 = V_2^{(i+1)\prime} V_2^{(i+1)}$$

where $V_1^{(i+1)} = \Phi_1^i - \Phi_{1(-1)}^i c^{(i+1)}$, $V_2^{(i+1)} = \Phi_2^i - \Phi_{2(-1)}^i s^{(i+1)}$, $df = T + 3$.

8. The draw of $\Phi_1^{(i+1)}$. To draw the states Φ_1, I assume that the draw of the second state is given. In this case I can use the following state-space representation:

$$\tilde{p}_t = H\xi_{1t} + e_t$$
$$\xi_{1t} = F\xi_{1,t-1} + \eta_{1t}$$

where $\tilde{p}_t = (1 - \psi^{i+1}L)(p_t - p_{t-1}\gamma^{i+1} - X_t b_1^{i+1} - \phi_{2t}^i t)$ and

$$H = \begin{bmatrix} 1 & -\psi^{i+1} \end{bmatrix}, \quad F = \begin{bmatrix} c^{i+1} & 0 \\ 1 & 0 \end{bmatrix}, \quad \xi_{1t} = \begin{bmatrix} \phi_{1t} \\ \phi_{1t-1} \end{bmatrix}, \quad \eta_{1t} = \begin{bmatrix} v_{1t} \\ 0 \end{bmatrix}$$

I use the usual Kalman filter and smoother recursions to estimate states Φ_1^{i+1}.

9. The draw of $\Phi_2^{(i+1)}$. Once I draw the states $\Phi_1^{(i+1)}$, I use the following state-space representation to draw the states $\Phi_2^{(i+1)}$

$$\tilde{p}_t = H_t\xi_{2t} + e_t$$
$$\xi_{2t} = F_t\xi_{2t-1} + \tilde{v}_{2t}$$

where $\tilde{p}_t = (1 - \psi^{i+1}L)(p_t - p_{t-1}\gamma^{i+1} - X_t b_1^{i+1} - \phi_{1t}^{i+1})$ and

$$H_t = \begin{bmatrix} T & -T\psi \end{bmatrix}, \quad F_t = \begin{bmatrix} \tilde{s}_t & 0 \\ 1 & 0 \end{bmatrix}, \quad \xi_{2t} = \begin{bmatrix} \tilde{\phi}_{2t} \\ \tilde{\phi}_{2t-1} \end{bmatrix}, \quad \eta_{2t} = \begin{bmatrix} \tilde{v}_{2t} \\ 0 \end{bmatrix}$$

$\tilde{\phi}_{2t} = \frac{1}{T}\phi_{2t}t$, $\tilde{s}_t = s^{i+1}\frac{t}{t-1}$, $\tilde{v}_{2t} = \frac{t}{T}v_{2t}$, $E(\tilde{v}_{2t}\tilde{v}'_{2t}) = \frac{t^2}{T^2}\sigma_2^{2(i+1)}$.

I use the standard Kalman filter and smoother recursions to estimate states Φ_2^{i+1}.

References

[1] Bernard, J.-T., Khalag, L., and M. Kichian, 2004. Structural change and forecasting long-run energy prices. Bank of Canada working paper 2004-5.

[2] Cashin, P., and C. J. McDermott, 2002. The long-run behavior of commodity prices: small trends and big variability. IMF Staff Reports 49(2), 175 - 199.

[3] Chacra, M., 2002. Oil price shocks and retail energy prices in Canada. Bank of Canada working paper 2002-38.

[4] Chan, H. L., and S. K. Lee, 1997. Modelling and forecasting the demand for coal in China. Energy Economics 19, 271 - 287.

[5] Dees, S., Karadeloglou, P., Kaufmann R. K., and M. Sanchez, 2004. Modeling the world oil market assessment of a quarterly econometric model. Working paper.

[6] Energy Information Administration, 2001. U.S. natural gas markets: recent trends and prospects for the future. http://www.eia.doe.gov/oiaf/servicerpt/naturalgas/pdf/oiaf00102.pdf

[7] Gately, D., 1995. Strategies for OPEC's pricing and output decisions. The Energy Journal 16(3), 1-37.

[8] Huntigton, H. G., 1994. Oil price forecasting in the 1980s: what went wrong? The Energy Journal 15(2), 1-22.

[9] Kemp, A. G., and A. S. Kasim, 2003. Forecasting activity levels in the UK continental shelf: the role of perceptions. Energy Economics 25, 713 - 739.

[10] Kim C.-J., and C. R. Nelson, 1998. State-space models with regime switching: classical and Gibbs-sampling approaches with applications. The MIT press.

[11] Koomey, J., Craig, P., Gadgil, A., and D. Lorenzetti, 2003. Improving long-range energy modeling: a plea for historical perspectives. The Energy Journal 24(4), 75 - 92.

[12] Lynch, M. C., 2002. Forecasting oil supply: theory and practice. The Quarterly Review of Economics and Finance 42, 373 - 389.

[13] Marcellino, M., 2002. Forecast pooling for short time series of macroeconomic variables. IEP-Bocconi University working paper n. 212.

[14] McAvinchey, I. D., 2002. Modelling and forecasting in an energy demand system with high and low frequency information. Economic Modelling 20, 207 - 226.

[15] McAvinchey, I. D., and A. Yannopoulos, 2003. Stationarity, structural change and specification in a deman system: the case of energy. Energy Economics 25, 65 - 92.

[16] Pindyck, R. S., 1999. The long-run evolution of energy prices. The Energy Journal 20(2), 1-27.

[17] Stock, J. H., and M. W. Watson, 2003. Forecasting output and inflation: the role of asset prices. Journal of economic literature 41(3), 788 - 829.

[18] Tang, L., and S. Hammoudeh, 2002. An empirical exploration of the world oil price under the target zone model. Energy Economics 24, 577 - 596.

[19] Wright, J. H., 2003. Bayesian model averaging and exchange rate forecasts. Board of Governors of the Federal Reserve System, International Finance Discussion Papers N779.

[20] Ye, M., Zyren, J., and J. Shore, 2002. Forecasting crude oil spot price using UECD petroleum inventory levels. International Advances in Economic Research 8(4), 324-233.

[21] Zeng, T., and N. R. Swanson, 1997. Predictive evaluation of econometric forecasting models in commodity futures markets. Studies in Nonlinear Dynamics & Econometrics 4(2), 159-177.

Table 1: Prior mean and variance for model parameters.

	Oil Model	*Coal Model*	*Gas Model*
γ	N(1.0, 0.2)*	N(1.0, 0.1)	N(1.0, 0.2)
b_1	N(2.0, 0.3)	N(1.0, 0.3)	N(1.0, 0.3)
c_1	N(0.95, 0.05)	N(0.9, 0.1)	N(0.9, 0.1)
s_1	N(1.0, 0.4)	N(1.0, 1.0)	N(1.0, 1.0)
ψ	N(0.2, 0.2)	N(-0.2, 0.2)	N(-0.2, 0.2)
ω	IW(0.2, 6)	IW(0.1, 6)	IW(33, 6)
$E[\omega]$	0.04	0.02	6.6
$Var[\omega]$	0.004	0.001	10.89
σ_1^2	IW(0.02, 6)	IW(0.002, 6)	IW(28, 6)
$E[\sigma_1^2]$	0.004	0.0004	5.6
$Var[\sigma_1^2]$	4×10^{-6}	4×10^{-8}	7.84
σ_2^2	IW(0.02, 6)	IW(0.002, 6)	IW(28, 6)
$E[\sigma_2^2]$	0.004	0.0004	5.6
$Var[\sigma_2^2]$	4×10^{-6}	4×10^{-8}	7.84

* I use the conjugate prior distributions to simplify the implementation of the Gibbs sampling algorithm

Table 2: Estimates of a shifting trend models.

	Oil Model	*Coal Model*	*Gas Model*
γ	0.305	0.197	0.761
	(0.17)	(0.14)	(0.28)
b_1	1.797	1.183	0.983
	(1.01)	(0.21)	(0.53)
c	0.831	0.672	0.640
	(0.17)	(0.21)	(0.20)
s	0.797	0.908	0.585
	(0.16)	(0.07)	(0.37)
ψ	0.162	0.058	0.037
	(0.11)	(0.12)	(0.16)
ω^2	0.040	0.0047	4.25
	(0.009)	(0.001)	(2.01)
σ_1^2	0.0018	0.00018	2.82
	(9.1×10^{-4})	(1.00×10^{-4})	(1.2)
σ_2^2	0.00010	0.000011	0.298
	(1.14×10^{-5})	(1.30×10^{-6})	(0.04)

I use data for the period 1870-1996 in estimation of shifting trend model for oil and coal prices and data for the period 1919-1996 in estimation of the natural gas price model.

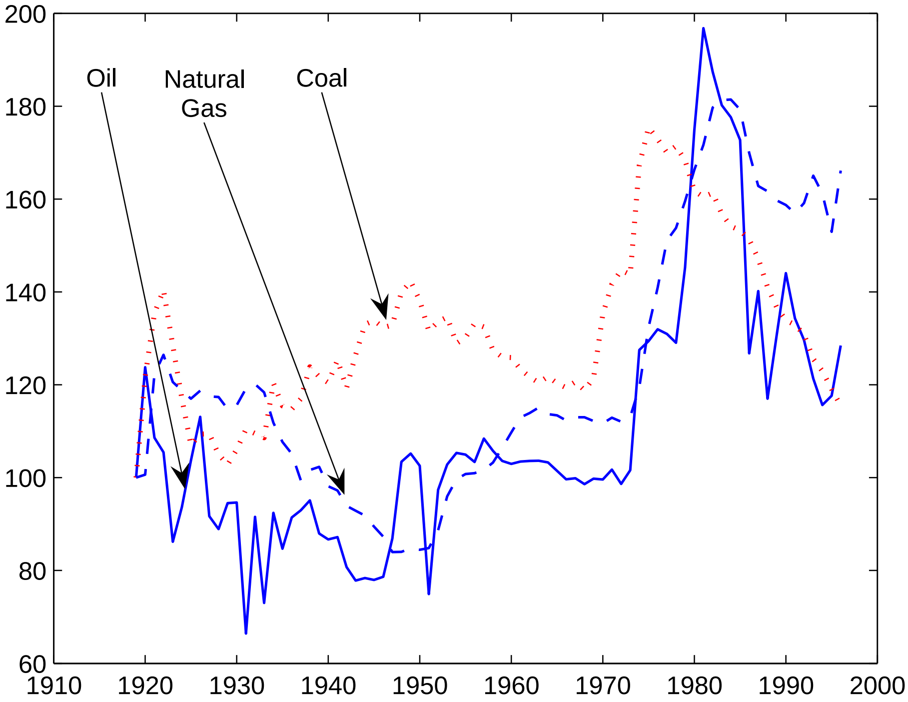
Table 3: Mean forecast squared error of different models

	Oil	*Coal*	*Gas*
	10-period-ahead		
Shifting trends	0.245	0.178	0.180
Univariate AR	0.427	0.083	0.777
Random Walk	0.361	0.142	0.499
*Combination Forecast 1**	0.176	0.063	0.200
*Combination Forecast 2***	0.210	0.082	0.179
	15-period-ahead		
Shifting trends	0.287	0.220	0.196
Univariate AR	0.562	0.175	1.971
Random Walk	0.285	0.206	1.144
Combination Forecast 1	0.165	0.120	0.123
Combination Forecast 2	0.184	0.138	0.196

* The combination forecast 1 is constructed by setting the parameter $w = 1$ in equation (5). This implies that the weight of each model is chosen inversely proportional to MFSE.

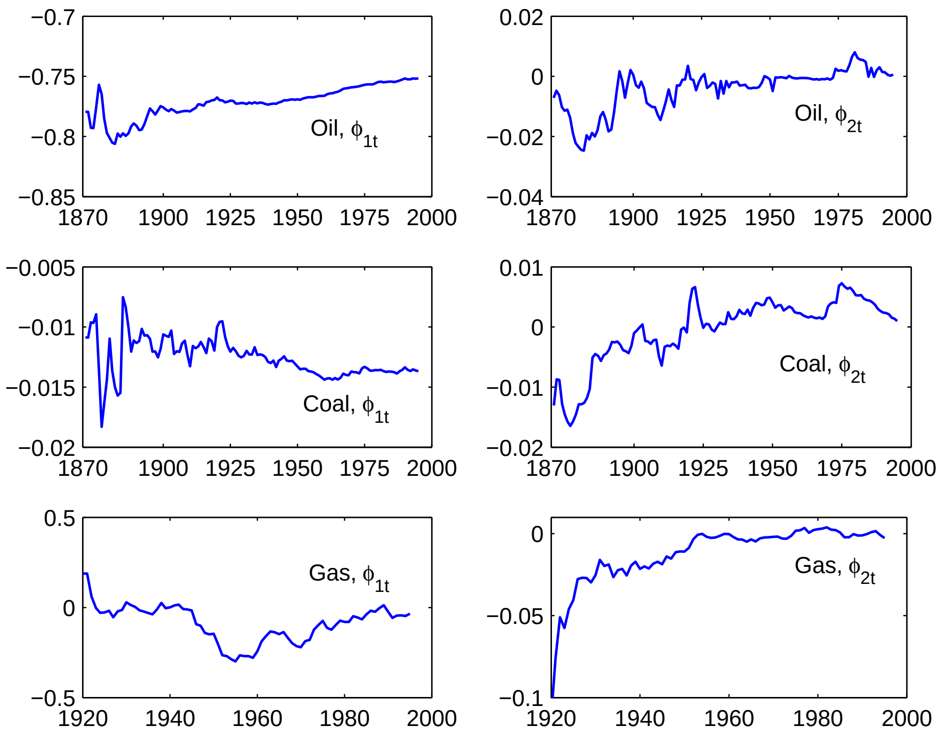
** The combination forecast 2 is constructed by setting the parameter $w = 5$ in equation (5). This implies that the best performing model receives the biggest weight.

Figure 1: The indices of real energy prices.



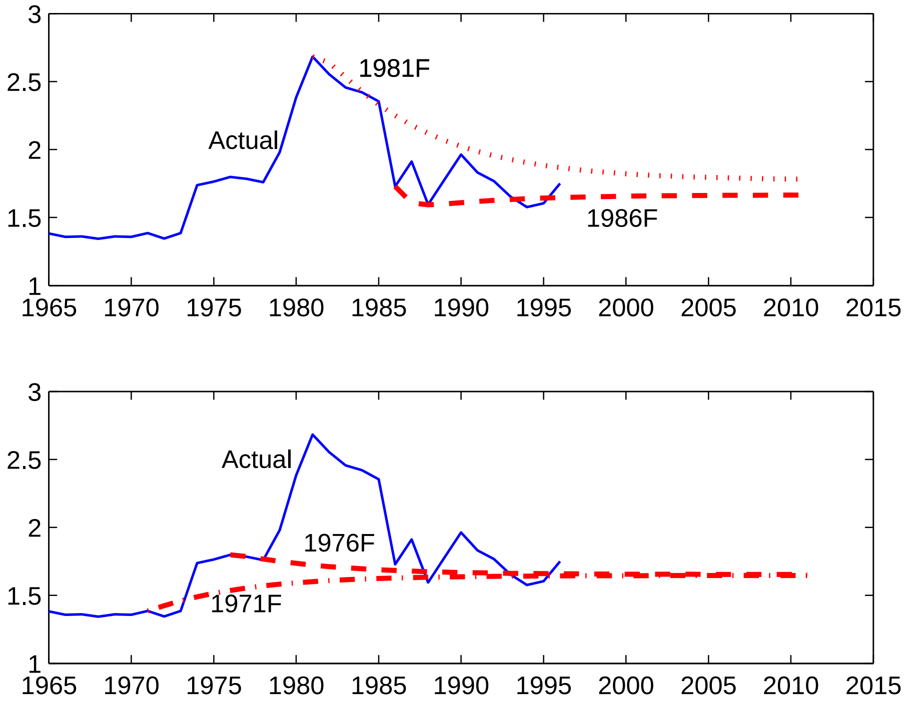
Notes: For comparison of the price dynamics, I normalize all real energy prices to 100 in 1919. Correlation coefficients among the variables: (i) the correlation between crude oil and natural gas prices is 0.82, (ii) the correlation between oil and coal prices is 0.65, (iii) the correlation between coal and natural gas prices is 0.46.

Figure 2: Estimates of unobserved shifting slopes and levels.



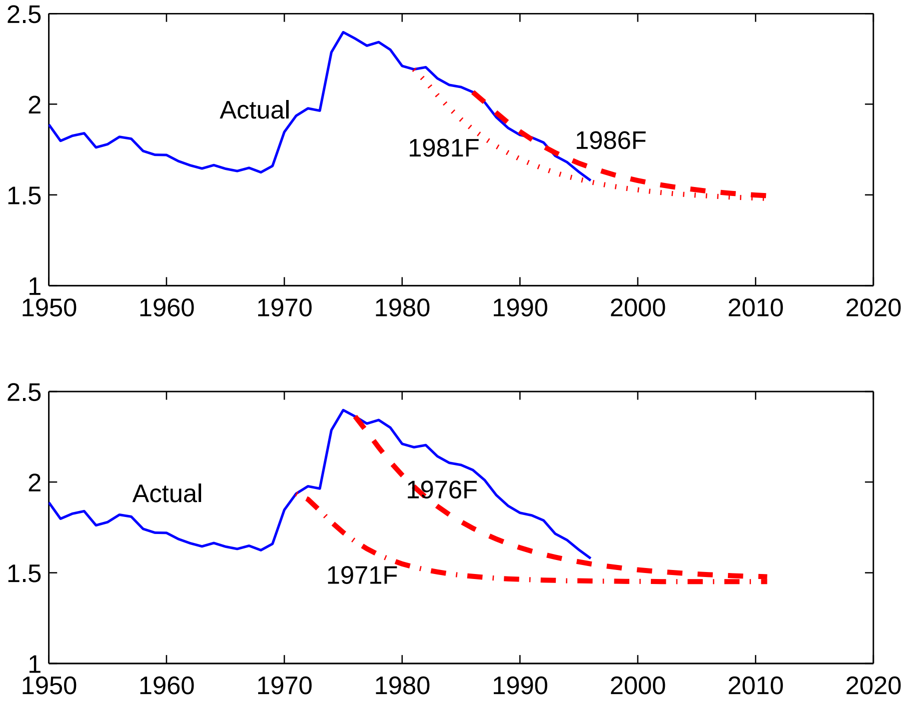
Notes: The shifting trend models for oil and coal prices are estimated using annual data for the period 1870-1996. For the natural gas price, the model is estimated using annual data for the period 1919-1996.

Figure 3: Crude oil forecasts.



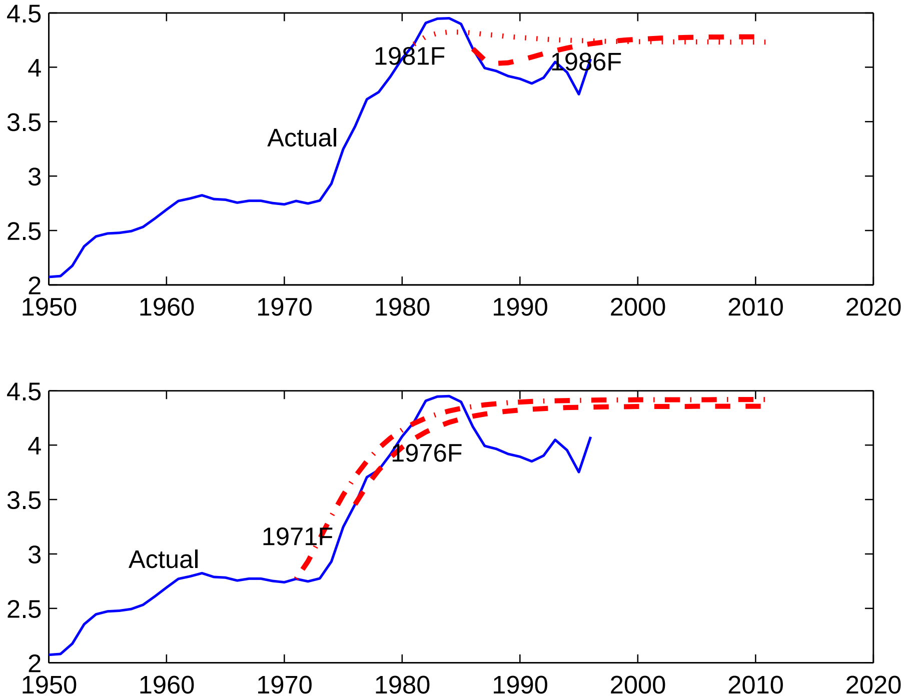
Notes: Crude oil forecasts are constructed for following periods: (i) 1986 - 2011, (ii) 1981-2011, (iii) 1976-2011, (iv) 1971 - 2011. The forecasts are computed based on the model of shifting trend lines proposed by Pindyck (1999). The model assumes depletable resource production and competitive behavior of producers.

Figure 4: Coal forecasts.



Notes: Bituminous coal forecasts are constructed for following periods: (i) 1986 - 2011, (ii) 1981-2011, (iii) 1976-2011, (iv) 1971 - 2011. The forecasts are computed based on the model of shifting trend lines proposed by Pindyck (1999). The model assumes depletable resource production and competitive behavior of producers.

Figure 5: Natural gas forecasts.



Notes: Natural gas forecasts are constructed for following periods: (i) 1986 - 2011, (ii) 1981-2011, (iii) 1976-2011, (iv) 1971 - 2011. The forecasts are computed based on the model of shifting trend lines proposed by Pindyck (1999). The model assumes depletable resource production and competitive behavior of producers.