Forecasting Long-run Coal Price in China: A Shifting Trend Time Series Approach

Boqiang Lin, Baomin Dong[†] and Xuefeng Li[*]

Abstract. The paper studies the behavior of mid- to long-run real energy prices, especially coal price in China. The problem is of great importance because the coal takes 70% share in China's energy mix and China is the world's second largest carbon emitter. An accurate forecast in coal price is crucial in predicting China's future energy consumption mix as well as private sector's energy-type-related investment decisions. In estimation and forecasting, the shifting trend time series model suggested by Pindyck's (1999) is used to capture the technological progress etc. that are unobservable to the econometrician. It is found that the shifting trend model with continuous and random changes in price level and trend outperforms plain vanilla ARIMA models. It is argued that the model postulated by Pindyck is robust even in a transition economy where energy prices are subject to relatively rigid regulatory control. Out-of-sample forecasts are provided.

Keywords: Shifting Trend, Energy Price, Coal

JEL Classification Numbers: Q30, Q40, C53

1. Introduction

It has been recognized that fluctuations in energy prices have important and lasting effects on the economies of industrialized countries (Hamilton (2003)). From a microeconomic or managerial perspective, forecasting energy prices can also be of very pragmatic[1]. The literature has extensively studied the impact of oil price changes on the economy. Energy plays a fundamental role in China's economy and continues to support the rapid economic growth and growing living standards. Because coal is such a dominant component of the energy structure in China, the coal industry has been important in this context. Coal is used in all sectors of China's economy and by households. It accounts for over two-thirds of primary energy consumption and more than three-quarters of electricity generation[2].

[†] Contacting author, Department of Economics, University of International Business and Economics, baomindong@uibe.edu.cn

[*] Department of Economics, University of International Business and Economics, Beijing, China 100029

[1] In private sectors especially energy producers and consumers regularly attempt to forecast prices of oil, coal, and other resources over time horizons as long as twenty or thirty years. Producers make these forecasts for general purposes of strategic planning, and for specific purposes of evaluating investment decisions, e.g., related to resource exploration, reserve development, and production. Industrial consumers, such as petrochemical companies or electric utilities, make these forecasts for the same kinds of reasons – oil, coal, and natural gas are important input costs that can affect investment decisions (e.g., an oil- versus coal-fired power plant for an electric utility), or even the choice of products to produce (e.g., a set of chemicals or the processes used to produce those chemicals).

[2] In 2004, the raw coal production is 1956 MT in China and the consumption demand in 2005 will exceed 2.1 BT, surpassing the demand in 2004 by 0.15 BT, where the increase is largest in power coal which accounts for 0.1 BT.

Forecasting the behavior and patterns of energy prices, however, can be quite challenging. In the case of China, the rate of growth in China's energy consumption has been a little more than moderate since 2000, in comparison to the notable decline in domestic coal consumption in late 1990s. The 2005 overall coal price increases by more than 50 Yuan/ton compared with 2004 level. Contributing to this have been newly operated coal-fired power generators, fast growing economy and decrease in the marginal benefits of efficiency improvements in electricity generation and in industry. This view is shared by a number of economists and practitioners. Thus, to characterize the price behaviors of coal and other energy, apart from domestic and international supply and demand conditions, one also needs to take into account market regulations, technological advances, and geopolitical considerations. These non-market-related aspects present the essential challenges for the econometrician, since they are largely unpredictable. For that reason, Pindyck (1999) suggests that, rather than fully articulated structural model, it is preferable to adopt simple models for our long-run forecasting needs where prices grow in real terms and at a fixed rate.

Despite being simple, the shifting trend models is flexible, allowing prices to grow from their current level (i.e., prices follow a random-walk process with drift) and/or from a changing trend line (i.e., prices revert to a possibly moving mean). Such differences can be thought of as reflecting differing assumptions regarding resource depletion and technological change. Indeed, using a simple Hotelling model on depletable resources, Pindyck shows that long-run energy prices should revert to an unobservable trending long-run marginal cost, with continuous random changes in their level and in the slope of their trend.

The coal price in China grew very slowly under the centrally-planned regime during 1950's to 1980's. The average mine-mouth sales price of raw and washed coal in major state-owned mines increased very slowly from 1953 to 1980 (see appendix A). The story is changed after major reform measures took place. The reform of coal prices in China has experienced five stages since the opening door policy enacted in late 1970s, from completely centrally planned one into a market oriented one gradually. To eliminate the possible biases caused by rigidity of the regulatory price control before 1994, our model uses price data ranging from 1994 to 2005.

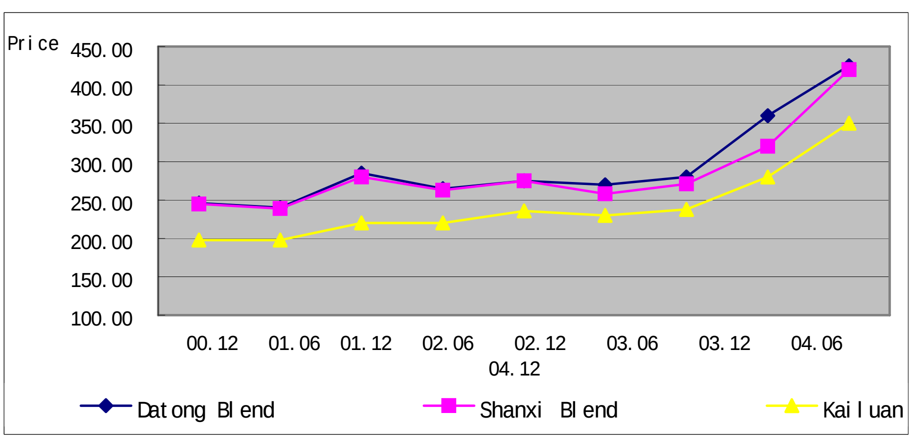


Figure 1.1 Prices in Qinghuangdao Coal Market

Table 1－1　Key State Coalmine Price 1953～2004 (in Yuan/ton)

Pricing Scheme		Year	Average Price
Coal Price determined by the State	Single Track System	1953	11.00
		1960	
		1965	17.68
		1970	
		1975	16.48
		1980	21.33
Deregulation of Coal Prices	Dual Track System	1985	26.05
		1990	43.85
		1991	58.45
		1992	90.67
Gradual Liberalization of Coal Prices		1993	105.42
		1994	108.94
		1995	115.00
		1996	125.00
		1997	166.60
		1998	160.20
		1999	143.98
		2000	140.19
		2001	150.99
Completely Market Determination		2002	167.81
		2003	173.81
		2004	206.43

Source: *The Electric Power and Coal in China*, Coal Industry Press

With rapid economic growth in China since the economic reform in 1979, the nation recently experienced overall energy shortage. In particular, the energy shortage is largely constituted by severe bottlenecks such as restricted electricity transmission capacity from western regions with abundant power resources to the eastern regions with relatively poor energy resources and fast growing energy demand. On the other hand, economists claim that most energy input sources are underpriced, causing resource distortions in the economy, possibility of unsustainability of continuous growth, and environmental problems. With abundant endowment of coal and heavy reliance on coal as main energy input source across the country, price of coal is essential to the nation's economic growth.

Historically, there is a gap between commercial coal price and power coal price in China because of the inflexible electricity tariffs and priority of electricity generation among all industrial sectors[3].

[3] In 2004, the average price of commercial coal by key state coal firms is 206.54 Yuan/ton, with a 47.3% increase. Power coal price was 161.55 Y/t in average in the same year, who sees a 33.6% increase. The difference between average commercial coal price and average power coal price was widened to 44.99 Y/t, or 14% difference in corresponding increases.

Table 1-2. Commercial Coal Price vs. Power Coal Price during 1997～2004 (Y/t)

	1997	1998	1999	2000	2001	2002	2003	2004
Average Commercial Coal Price	166.34	160.20	142.74	140.19	150.99	167.88	178.61	206.54
Average Power Coal Price	137.33	133.27	121.48	120.93	123.94	137.25	145.25	161.55
Difference	29.01	26.93	21.26	19.26	27.05	30.56	33.36	44.99

Coal-fried generation takes a share as high as 80% in China. Power coal consumption is around half of the total coal consumption in China. The growing demand for electricity continuously drives the demand for coal to increase. Power coal price increase is a main component of coal price increase. But the price gap is about to an end, since the NDRC has lifted up the last restriction on power coal pricing – the suggested price – since January 2006 Coal Ordering Conference. To resolve potential conflict between coal producers and power generators, the regulator has prepared and indeed activated a price linkage program[4] which allows power prices to float in line with the change in coal prices.

The environmental valuation of powerplant emissions may put a limit for the coal price if the prices for alternative fuels remain constant or with moderate increase compared with coal price. Already, China has begun to charge powerplants for emissions of sulphur dioxide and nitrogen oxides. The tax on these pollutants may increase over the years, as a means of internalizing their environmental and health costs to China's citizens. In addition, as concerns grow over global warming, a market value may increasingly be attached to greenhouse gas emissions.

Therefore, in our estimation, we only focus on the price behavior of commercial coal because of the fact that power coal price is now also completely subject to market determination.

In this research we examine the mid- to long-run behavior of coal prices in China without considering structural modeling. Following Pindyck (1999), we first test whether prices are mean-reverting. The results support the hypothesis that coal price follow a mean-reverting trend, but the rate of mean revision is very slow so that a decision maker can treat the price movement as a geometric Brownian motion, or other random-walk process. In addition, similarly to the Pindyck's result on US energy price long run data, the trends to which prices revert are themselves fluctuating over time.

2. Brief Literature Review

The idea of shifting trend, as pointed out earlier, is not new. Perron (1989) developed a stochastic switching model that allows for discrete shifts in the slope or level of the trend line. Applying the model to data on real GNP, he found two events that seem to represent permanent changes in the underlying process, the Great Crash (which shifted the trend line downward) and

[4] Under the scheme, power companies can transfer 70 per cent of the rise in fuel costs to end users, should coal prices increase by more than 5 per cent within a six-month period. Power companies will have to absorb the remaining 30 per cent themselves.

the 1973 oil price shock (which changed the slope of the trend line). Other econometricians, such as Videgaray (1998), examined the possibility of such a structural change in the context of the price of crude oil and found a structural change around 1973 which supports Perron's argument.

The present research will focus on time-varying trend or sometimes called shifting-trend models. In energy price forecast model literature, Huntingon (1994) shows that the forecasting performance of ten structural models was problematic since the errors in structural models were due to factors such as exogenous GNP assumptions, resource supply conditions outside OPEC, and demand adjustments to price changes. Lynch (2002) concludes in similar fashion in a comparison between the theory and empirics of oil supply forecasting. Koomey et al (2003) point out that factors like technological innovation and inaccuracy of oil reserve forecasts may also contribute to the forecast errors. Tang and Hammoudeh (2002) show that omission of market participants' expectations account for forecast errors too. Pindyck (1999) points out that structural models may not be always accurate in long run forecast, but they are better suited at providing understandings of the causes of short or intermediate run fluctuations of prices and other variables. Pindyck 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) and statistically significant instabilities for coal and natural gas prices were found.

Using the model of depletable resource production, Pindyck (1999) argues that the forecast of energy prices in the model is based on the long run total marginal cost. To increase the prediction accuracy, Radchenko (2005) relaxes some assumptions on model parameters and white noise terms in Pindyck (1999), and proposes a Bayesian approach to estimate the model with autocorrelation. Radchenko (2005) also combines shifting trend model, random walk model and autoregression model and finds that the combined model substantially decreases the mean squared forecast error (MSFE).

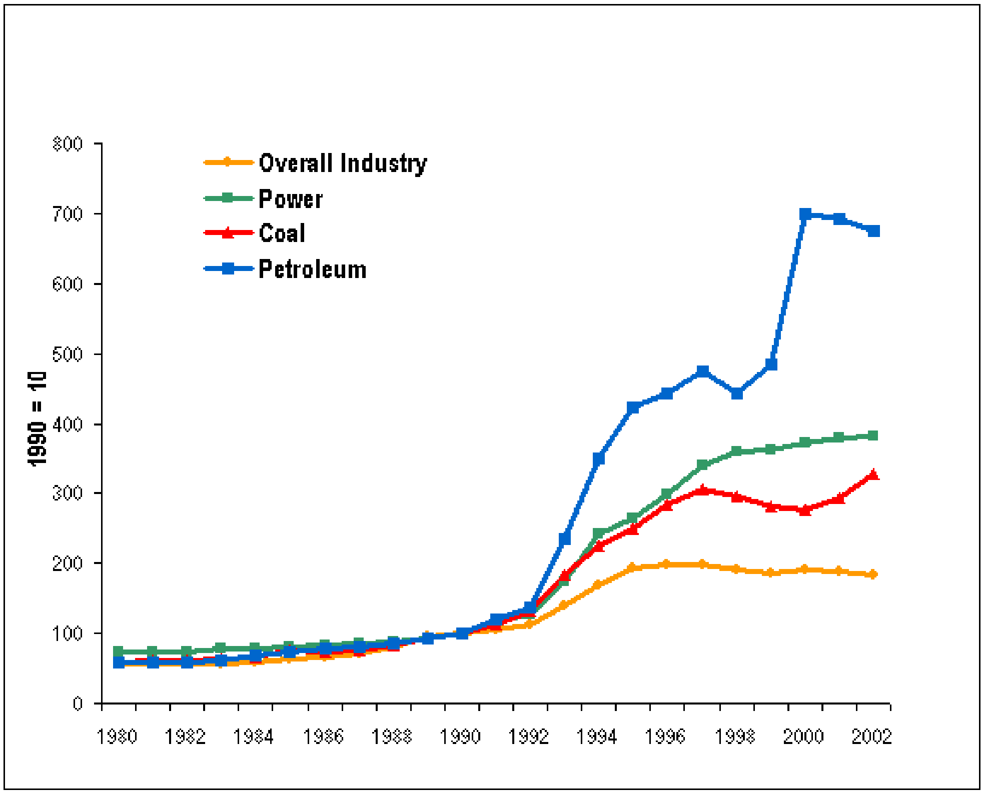
The current research presents the shifting trend model and compares it with a benchmark time series ARIMA model. Ye *et al* (2002) propose a short run monthly autoregressive distributed lag forecast model of WTI crude oil spot price using OECD inventory levels. Zeng and Swanson (1997) examine the predictive accuracy of various econometric models for the crude oil price using daily futures prices. Chacra (2002) uses a quarterly forecast model to examine the relationship between world oil prices and components of CRP-energy and builds short-sighted forecasts. Sun and Peng (2000) fit the historical prices (1994~1999 time series data) of high quality Shanxi coal (5,500 kCal/kg) in a time series model and conclude the data suits the ARIMA (1,1,1) model, using AIC (Akike Information Criterion).

Our estimation shows that the coal price follows reversion to trend lines with slopes and levels that are both shifting continuously and unpredictably over time, so that each price follows a multivariate stochastic process. The shifts themselves are mean-reverting, but ignoring the stochastic components here is misleading, and can lead to sub-optimal forecasts[5].

---

[5] Pindyck (1999) has shown that a multivariate model with continuous fluctuations in the trend line slope and level is consistent with basic models of exhaustible resource production.

Figure2-1. Ex-Factory Price Indices for Industrial Products by Sector (1990=100)



3. The Model

In the existing literature of Econometric models and forecasts on energy demand, single equation time series model is a frequently used choice. On one hand, the naïve time series approach has the merit of avoiding misspecification errors compared with panel data model. On the other hand, it is obvious that most economic data commonly used in demand analysis are normally nonstationary and Econometric studies that overlook this particular characteristic may lead to a 'purious result', that is, obtaining a high R^2 although there is no meaningful relationship existing between the two sides of the equation. This problem occurs because when both the dependent and independent variables involved exhibit strong trends, the high R^2 observed is due to the presence of the trends, not to a true relationship between them.

To overcome the spurious problem, one of the best solutions is to apply cointegration and error correction models (ECMs)[6]. We use the traditional ECM time series model as a benchmark to be compared with.

The model of interest here is the one developed by Pindyck (1999) who proposes a reduced

[6] See appendix B for details.

form model with time-varying parameters for the analysis of stochastic properties of energy prices. 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).

The univariate shifting trend time series model is as follows:

$$p_t = \gamma p_{t-1} + b_1 + b_2 t + \phi_{1t} + \phi_{2t} t + u_t \,, \tag{3.1}$$

$$\phi_{1t} = c\phi_{1,t-1} + v_{1t} \,, \tag{3.2}$$

$$\phi_{2t} = s\phi_{2,t-1} + u_{2t} \,, \tag{3.3}$$

$$u_t = \psi u_{t-1} + e_t \,, \tag{3.4}$$

where ϕ_{1t} and ϕ_{2t} are unobservable state variables, p_{1t} is a real price of 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, the model converges to the one in 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 is proved difficult to estimate this model 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 (3.3) to 1 for coal model, sets a parameter c to zero for natural gas model, and exclude the unobserved state from the estimation of gas model. We do not restrict parameter *s* to 1 or *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, we compare mean forecast squared errors for shifting trend models with the random walk model and univariate autoregressive models. Therefore, we 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. We construct linear combinations of forecasts to check how well they perform relative to single forecasts. Combination forecasts are constructed using the following formula:

$$\hat{y}^{c}_{t+h} = \sum_{m=1}^{M} k_{m,h,t}\,\hat{y}_{t+h,m} \tag{3.5}$$

$$k_{m,h,t}y = \frac{(1/MFSE_{m,h,t})^{\omega}}{\sum_{m=1}^{M}(1/MFSE_{j,h,t})^{\omega}} \qquad (3.6)$$

where $\hat{y}^{c}_{t+h}$ denotes a constructed combination forecast, $\hat{y}_{t+h,m}$ denotes the forecasts of the considered model (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 computing the weights, we set the coefficient ω=1 for simplicity, which implies that the weight of model is chosen as inversely proportional to its MFSE. We also check the performance of combination forecast when the coefficient ω for larger value which implies that the best performing model receives more weight.

The Gibbs sampling algorithm, one of the simplest Markov chain Monte Carlo algorithms, is used in estimation in univariate models with shifting trends. Gelfand and Smith (1990) discussed the value of the Gibbs algorithm for a range of problems in Bayesian analysis.

To define the Gibbs sampling algorithm, let the set of full conditional distributions be

$$\{\pi(\psi_1|\psi_2,\dots,\psi_p)\};\{\pi(\psi_2|\psi_1,\psi_3\dots,\psi_p)\};\dots,\{\pi(\psi_p|\psi_1,\dots,\psi_{d-1})\}, \qquad (3.7)$$

Now one cycle of the Gibbs sampling algorithm is completed by simulating $\{\psi_k\}_{k=1}^{p}$ from these distributions, recursively refreshing the conditioning variables. The Gibbs sampler in which each block is revised in fixed order is defined as follows.

Gibbs Sampling:

1. Specify an initial value $\psi^{(0)}=(\psi_1^{(0)},\dots,\psi_p^{(0)})$
2. Repeat for *j*=1,2,...,M

   Generate $\psi_1^{(j+1)}$ from $\pi(\psi_1|\psi_2^{(j)},\psi_3^{(j)},\dots,\psi_p^{(j)})$

   Generate $\psi_1^{(j+1)}$ from $\pi(\psi_2|\psi_1^{(j+1)},\psi_3^{(j)}\dots,\psi_p^{(j)})$

   ... ...

   Generate $\psi_1^{(j+1)}$ from $\pi(\psi_p|\psi_1^{(j+1)},\dots,\psi_{p-1}^{(j+1)})$
3. Return the values Generate $\{\psi^{(1)},\psi^{(2)},\dots,\psi^{(M)}\}$

Since when the *k*-th block is reached, the previous (*k*-1) blocks have been updated. Thus, the transition density of the chain, under the maintained assumption that *π* is absolutely continuous, is given by the product of transition kernels for each block:

$$K(\psi,\psi') = \prod_{K=1}^{P}\pi(\psi_k\,|\,\psi_1^{(j+1)},\dots,\psi_{k-1}^{j+1},\psi_{k+1}^{(j)},\dots,\psi_p^{(j)}). \qquad (3.8)$$

In our case, let Θ^i denote the *i*th 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 *i*th 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 Gibbs sampling algorithm is used to generate the draw of parameters Θ^{1+i} and states $\Phi_1^{(i+1)}$ and $\Phi_2^{(i+1)}$.

We use conjugate prior distributions to simplify computations in the Gibbs sampling algorithm. The choice of hyperparameters in our univariate model should be based on the Econometrician's available information and trials.

For the distribution specification of hyperparameters for variance, we follow the existing literature, i.e., assume Inverted-Wishart distribution for the variance parameters. The potential difference in the prior for variance across different energy input prices is consistent with the estimates of unobserved states and their variance in Pindyck (1999). A simple way to incorporate the Gibbs sampling procedure is through a Kalman filter model. In Eviews or TSP, the Kalman filter treats the trend line slope and level as state variables that evolve stochastically, and that cannot be observed directly. The values are estimated recursively over the sample horizon, along with any fixed parameters. Thus the Kalman filter is a type of time-varying parameter estimator.

4. Data and Estimation Results

We examine the coal prices over the 11-year period from 1994 to 2005. All coal prices are Qinghuangdao coal market average trading prices, available from Development Research Center of State Council (PRC) and *Coal Information* Magazine. Natural logarithm of each deflated series is taken.

Firstly, an AR(1) with fixed trend time series model is estimated as follows.

$$\log p_t = \alpha + \beta_1 \cdot t + \beta_2 \cdot t^2 + \beta_3 \cdot \log p_{t-1}$$

The estimation results show that

$$\begin{array}{c} \log p_t = 0.307 - 0.0008t + 8\mathrm{E}(-6) \cdot t^2 + 0.947 \log p_{t-1} \\ (0.114) \quad (0.0003) \quad 2.2\mathrm{E}(-6) \qquad (0.0207) \end{array}$$

where numbers in parentheses are standard errors of the parameter estimates.

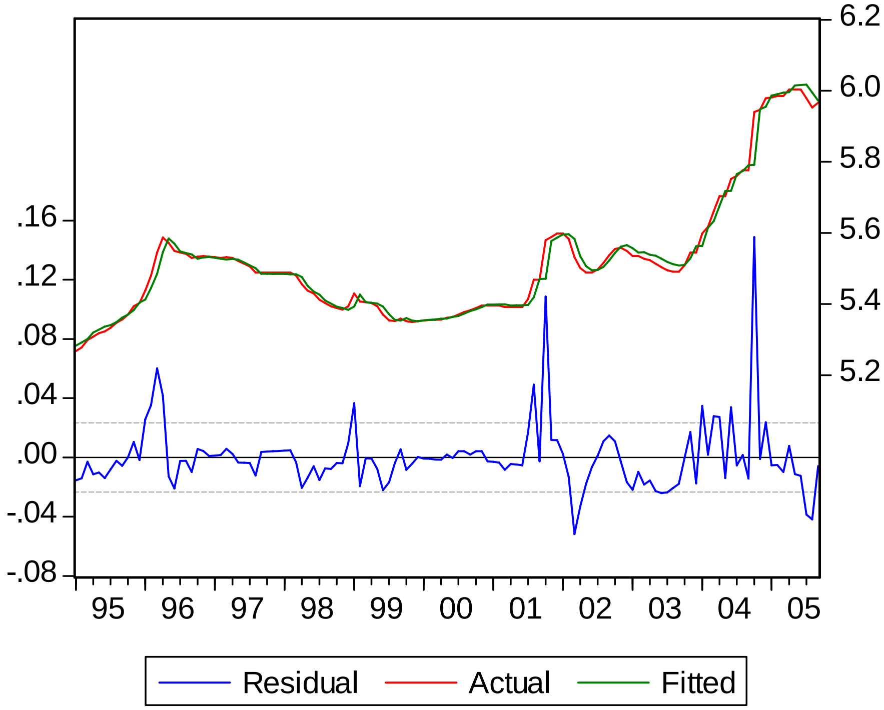
Regression 1: AR(1)

Sample (adjusted): 1995M01 2005M09

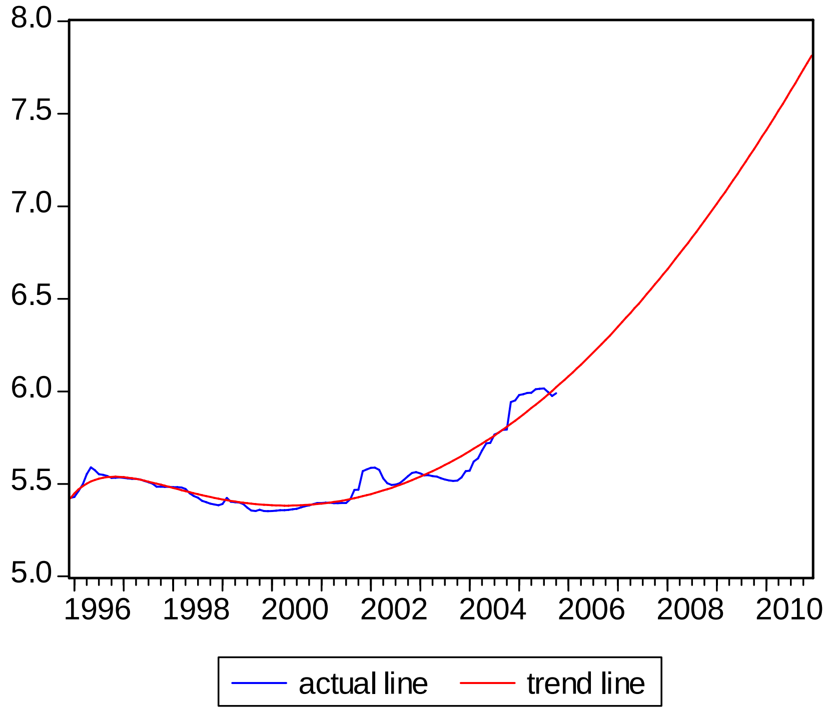
Variable	Coefficient	Std. Error	t-Statistic	Prob.
α	0.306938	0.114407	2.682857	0.0083
β_1	-0.000822	0.000260	-3.163818	0.0020
β_3	0.946818	0.020704	45.73091	0.0000
β_2	8.00E-06	2.21E-06	3.623324	0.0004

R-squared	0.983396	Mean dependent var	5.517869
Adjusted R-squared	0.982998	S.D. dependent var	0.179204
S.E. of regression	0.023367	Akaike info criterion	-4.644470
Sum squared resid	0.068252	Schwarz criterion	-4.555794
Log likelihood	303.5683	F-statistic	2467.794
Durbin-Watson stat	1.569310	Prob(F-statistic)	0.000000

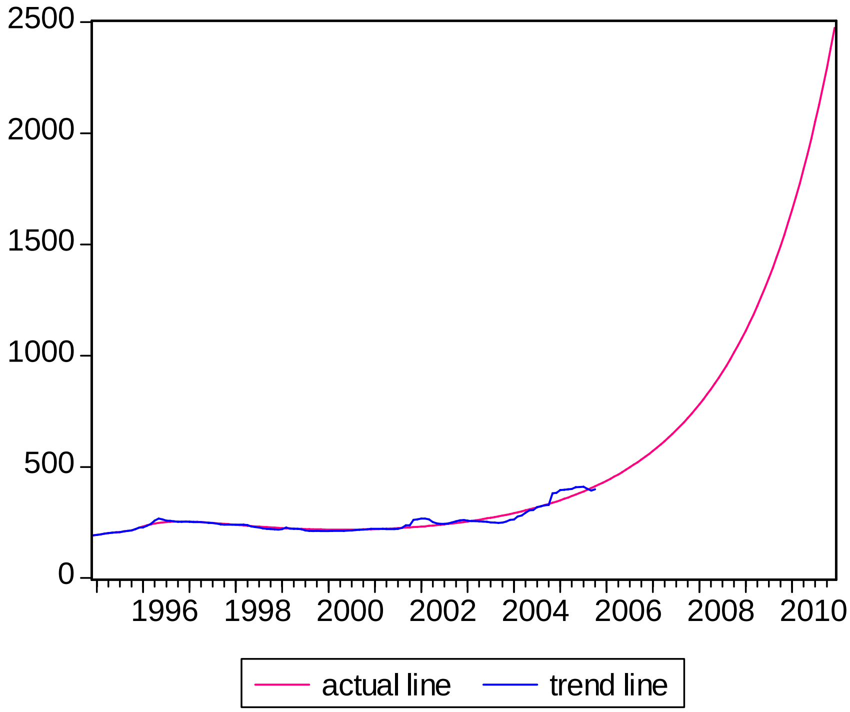
The actual, residual and fitted graph is as follows.



The out of sample forecast is done as follows.



The following graph is the one that transformed into prices (not logarithm).



The second model estimates the same model but with a smaller sample, i.e.,

$$\log p_t = C_1 + C_2 \cdot t + C_3 \cdot t^2 + C_4 \cdot \log p_{t-1}$$

$$\log p_t = -0.326 - 0.0289t + 0.00031 \cdot t^2 + 0.945 \log p_{t-1} - 0.997 u_{t-1}$$
$$(0.19) \quad (0.0038) \quad 3.44\text{E}(-5) \quad (0.0456) \quad (0.119)$$

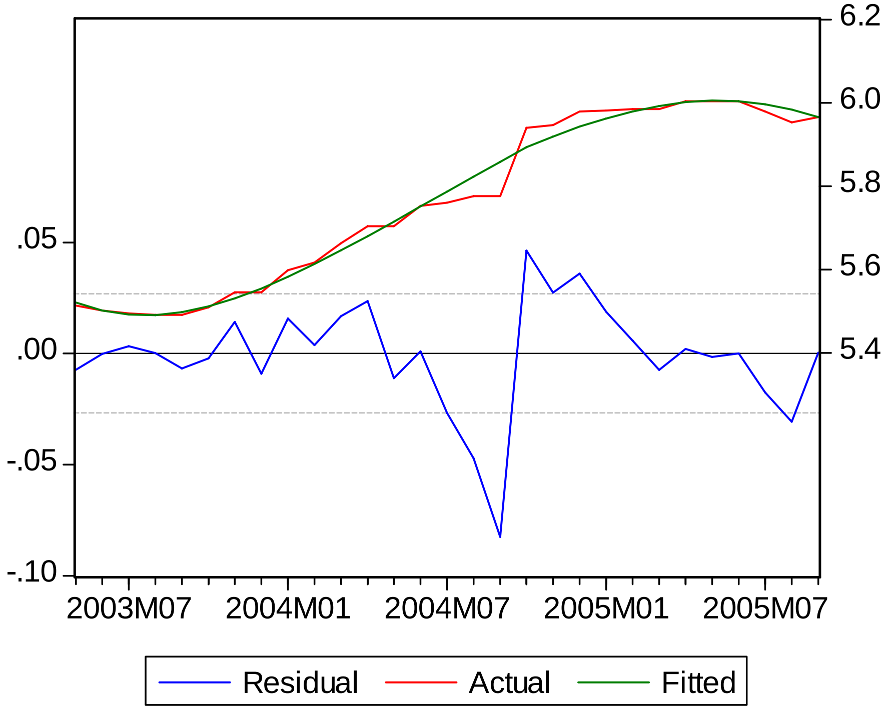
where numbers in parentheses are standard errors of the parameter estimates.

Regression 2

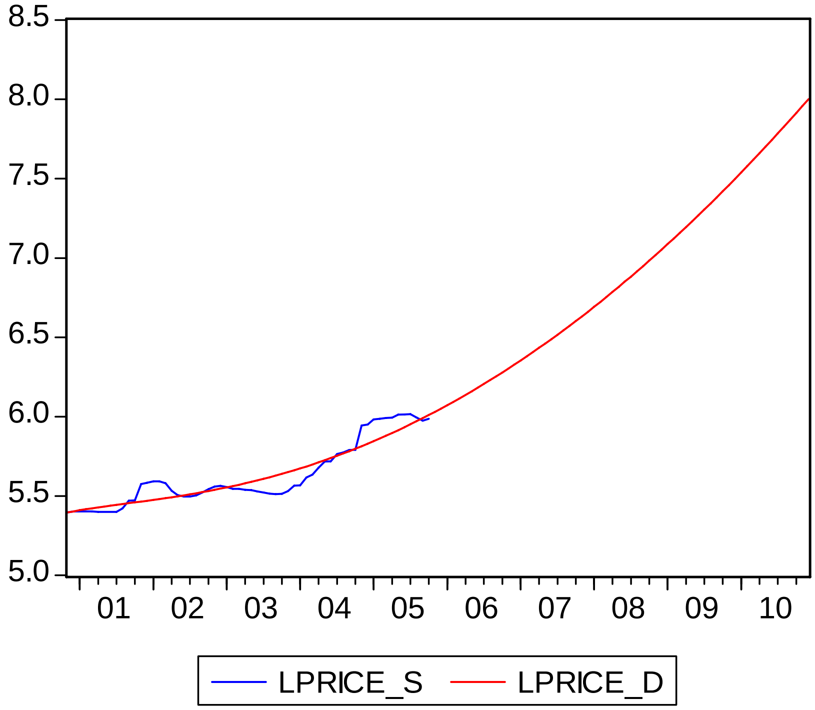
Sample: 2003M05 2005M09

Variable	Coefficient	Std. Error	t-Statistic	Prob.
T	0.028824	0.003817	7.552282	0.0000
T2	-0.000306	3.44E-05	-8.912848	0.0000
C	-0.326043	0.190096	-1.715148	0.0992
LPRICE(-1)	0.945278	0.045642	20.71060	0.0000
MA(1)	-0.997406	0.118892	-8.389143	0.0000

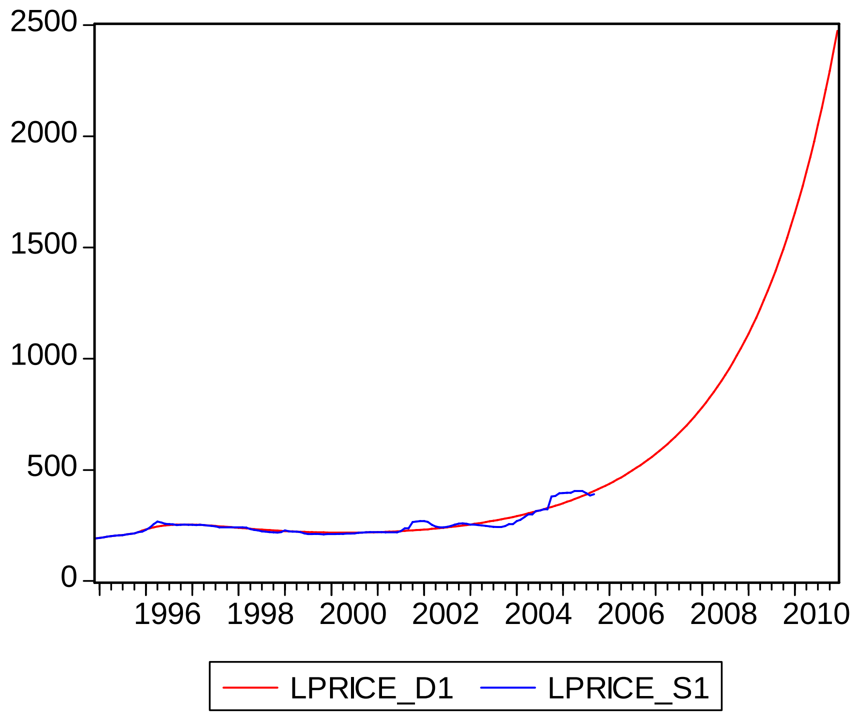
R-squared	0.984735	Mean dependent var	5.764485
Adjusted R-squared	0.982190	S.D. dependent var	0.201065
S.E. of regression	0.026833	Akaike info criterion	-4.242810
Sum squared resid	0.017280	Schwarz criterion	-4.007069
Log likelihood	66.52074	F-statistic	387.0486



Forecast from this model shows that price will increase substantially in five years.



As one can see clearly that the second model also gives unreasonable price forecasts.



The third model, shifting trend model, outperforms all other estimations we have performed. The model we estimate is

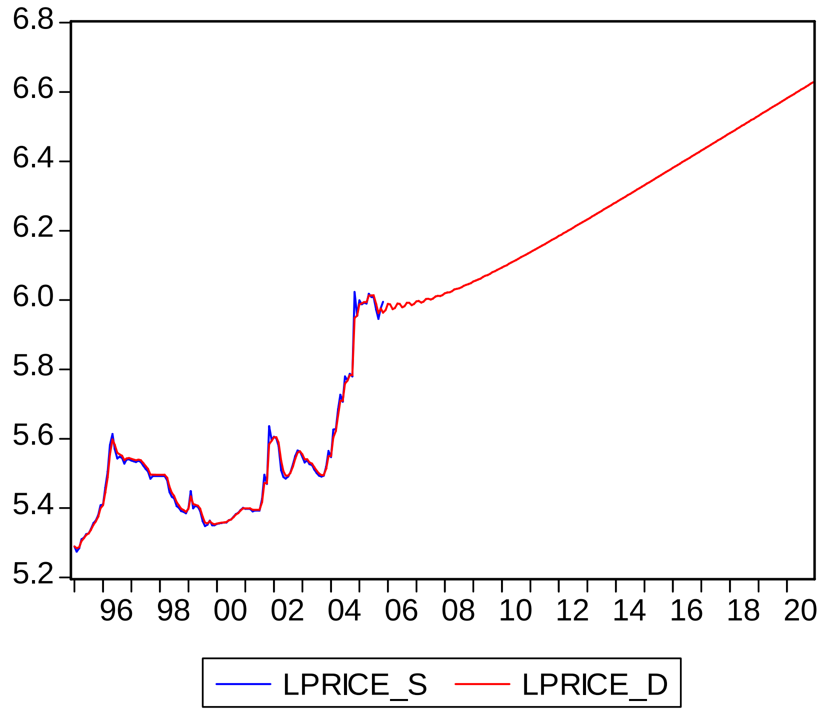
$$\log p_t = c_1 + c_2 p_{t-1} + \phi_{1t} + \phi_{2t} t + \varepsilon_t$$
$$\phi_{1t} = c_3 \phi_{1,t-1} + \upsilon_{1t},$$
$$\phi_{2t} = c_4 \phi_{2,t-1} + \upsilon_{2t},$$

Regression 3, Shifting Trend Model

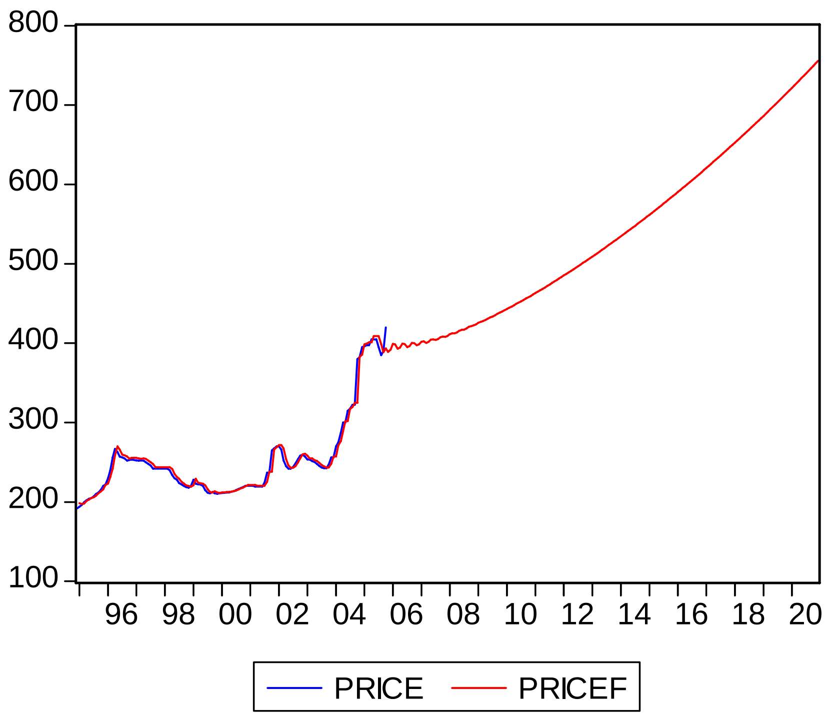
	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	2.834709	0.351328	8.068556	0.0000
C(2)	0.467773	0.070751	6.611539	0.0000
C(3)	0.961771	0.143505	6.701984	0.0000
C(4)	0.994278	0.006032	164.8336	0.0000
C(5)	-8.894320	0.207795	-42.80339	0.0000
C(6)	-16.51694	0.092315	-178.9197	0.0000
	Final State	Root MSE	z-Statistic	Prob.
SV1	0.008447	0.042130	0.200491	0.8411
SV2	0.002584	0.000413	6.255229	0.0000

Log likelihood	317.0173	Akaike info criterion	-4.821974
Parameters	6	Schwarz criterion	-4.688959
Diffuse priors	0	Hannan-Quinn criter.	-4.767927

The following graph shows the estimated line and original line.



The following graph is the one that transformed into prices (not logarithm). One can see that the shifting trend model forecast is much more reasonable than that of the simple ECM model.



References

Baker, Malcolm P., E. Scott Mayfield, and John E. Parsons, "Alternative Models of Uncertain Commodity Prices for Use with Modern Asset Pricing Methods," *The Energy Journal*, 1998, **19**, 115–48.

Campbell, John, and N. Gregory Mankiw, "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, 1987, **102**, 857–80.

Chang, Youngho and Jiang Chan (2003), Oil price fluctuations and China economy, Energy policy, 31(11):1151-1165.

Im, K., Pesaran, M. and Shin, Y. (2002), Testing for unit roots in heterogeneous panels. Mimeo.

Kao, C. (1999), Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics* 90, 1-44.

MacKinnon, James, "Critical Values for Cointegration Tests," in R. F. Engle and C. W. J. Granger, eds., *Long-Run Economic Relationships*, (Oxford: Oxford University Press), 1991.

Manthy, Robert S., *Natural Resource Commodities: A Century of Statistics*, (Baltimore, Johns Hopkins Press), 1978.

McCoskey, S. and Kao, C. (1998), A residual-based test of the null cointegration in panel data. *Econometric Reviews* 17, 57-84.

Perron, Pierre, "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Journal of Political Economy*, December 1989, **57**, 1361–1401.

Phillps, P. C. B. and Hansen. B. E. (1990), Statistical inference in instrumental variables regression with I(1) processes, *Review of Economic Studies* 57.99-125.

Phillps, P. C. B.(1986), Understanding spurious regressions in econometrics, *Journal of Econometrics*, 33.311-340.

Pindyck, Robert S., "The Optimal Exploration and Production of Nonrenewable Resources," *Journal of Political Economy*, October 1978, **86**, 841–61.

Pindyck, Robert S., "Uncertainty and Exhaustible Resource Markets," *Journal of Political Economy*, 1980, **88**, 1203–25.

Pindyck R. S. 1999. "The Long Run Evolution of Energy Prices." *The Energy Journal* 20: 1-27.

Pindyck Robert S., "Inventories and the Short-Run Dynamics of Commodity Prices," *The RAND Journal of Economics*, Spring 1994, **25**,141–59.

Appendix A. Coal Resources in China by Type of Coal (Gt)

Coal Classification	Estimated Resource	Share
brown coal	190.31	4.2%
bituminous coal	2,421.51	53.2%
gas coal	939.24	20.6%
rich coal	103.21	2.3%
coking coal	195.73	4.3%
lean coal	80.38	1.8%
meager coal	146.89	3.2%
anthracite coal	474.24	10.4%
total	4,551.50	100.0%

This is the Chinese category of " Geologically Identified Reserves"

Average Minemouth Sales Price of Raw and Washed Coal, Major State-Owned Mines (yuan/t)

Year	Average Price of Raw and Washed Coal	Datong Bureau	Kailuan Bureau	Huaibei Bureau
1953	11.00			
1957	11.46			
1965	17.68			
1975	16.48			
1980	21.33			
1982	21.58			
1984	22.73			
1985	26.05			
1986	26.51	27.93	26.84	
1987	26.28	28.32		27.04
1988	27.72			
1990	43.85			
1991	47.89	49.14	53.10	55.45
1992	54.73	58.04	61.65	58.28

Source: Energy Research Institute.

Appendix B. Cointegration and Error Correction Model (ECM)

In ECM, let us start from the following simple equilibrium equation:

$$p_t = \alpha + \beta x_t \, , \tag{A.1}$$

When the residual terms are not zero, this quantity measures the extent of disequilibrium between p_t and x_t, one typical form of which is

$$p_t = \gamma + \delta_0 x_t + \delta_1 x_{t-1} + \tau p_{t-1} + u_t \qquad \text{(A.2)},$$

where u_t is the noise term. Such formulation obviously yields

$$\Delta p_t = \delta_0 \Delta x_t - \mu(p_{t-1} - \alpha - \beta x_{t-1}) + u_t, \qquad \text{(A.3)}$$

where μ=1-τ. When higher order lagged variables are introduced, the above equation is modified into:

$$\Delta p_t = \sum_{i-1}^{k-1} \Psi_i \Delta y_{t-1} + \sum_{i=0}^{k-1} \delta_i \Delta x_t - \mu(p_{t-k} - \alpha - \beta x_{t-k}) + u_t, \qquad \text{(A.4)}$$

The popularity of ECM is due to the works of Granger and Engle on cointegration. The importance of cointegration stems from the fact that statistical inference from conventional regression is only valid when the variables in a model are stationary. Most economic time series are not, however, stationary. The relation of the ECM to cointegration analysis derives from a representation theorem proved by Engle and Granger (1987). A maximum likelihood procedure is developed by Johansen and Juselius (1990) and Johansen (1991) began to replace the simple regression approach to estimate the long run coefficients. Chan and Lee (1997) estimate China's demand for coal using ECM and cointegration based on annual data from 1953 to 1994.

We perform the augmented Dickey-Fuller unit root tests, as is often done. Suppose that the detrended logarithm of price around its mean follows an AR(1) process:

$$p_t = \rho p_{t-1} + \varepsilon_t,$$

where 0 ≤ρ≤1, and εt is a white noise process. Then the asymptotic standard deviation of ρ is given by:

$$\text{s.d.}(\rho) = \left(\frac{1-\rho^2}{T} \right)^{1/2}$$

where T is the number of observations. Now, suppose we run a Dickey-Fuller unit root test to determine whether $\rho \leq 1$. To reject the hypothesis that $1 - \rho = 0$ at the 5 percent level, we would need a t-statistic on $1 - \rho$ of at least 2.89. Thus, we would need

$$(2.89)^2 \leq \frac{T(1-\rho)^2}{1-\rho^2}$$

or equivalently, T ≥ (8.352)(1– ρ ^{2})(1– ρ) ^{2}.

We perform augmented Dickey-Fuller unit root tests on the full sample of data. The test involves regressions of the form:

$$\Delta p_t = (\rho - 1)\, p_t - 1 + \sum_{k-1}^{N} \alpha_k \Delta p_{t-k} = c_0$$

where *p* is the log real price, and *N* is the number of lags. The test is based on the MacKinnon (1991) critical values for the *t*-statistic on *ρ* – 1.

Appendix C. The Trend Line for a Depletable Resource

For a depletable resource such as coal, we would expect both the level of the log price trajectory and its slope to fluctuate over time in response to fluctuations in demand, extraction costs, and reserves. To see this, consider the basic Hotelling model of a depletable resource produced in a competitive market with a constant marginal cost of extraction, *c*. In this model the price trajectory is $dP/dt = r(P - c)$. Hence the price level itself is given by

$$p_t = p'_0 e^{rt} + c,$$

where $P'_0 = P_0 - c$ is the net price. If the demand function is isoelastic, i.e., is of the form $Q_t = AP_t^{-n}$, the trajectory for the rate of production will be given by

$$Q_t = A\left(c + p_0^{'} e^{rt}\right)^{-\eta}$$

In this case, we can find the initial net price $\boldsymbol{P}_0$ by making use of the fact that cumulative production over the life of the resource must equal the initial reserve level, R_0:

$$R_0 = \int_0^\infty Q_t dt = \int_0^\infty A\left(c + p_0^{'} e^{rt}\right)^{-\eta} dt$$

For arbitrary values of the elasticity of demand, *η*, this equation can be solved numerically for $p_0^{'}$.

For a unitary elasticity of demand (*η* = 1), we can solve it analytically:

$$R_0 = \frac{A}{rc} \log \frac{c + p_0^{'}}{p_0^{'}}$$

or

$$p_0^{'} = \frac{c}{e^{rcR_0/A} - 1}$$

Hence the price level at any time t is given by

$$p_t = c + \frac{ce^{rt}}{e^{rcR_0/A} - 1}$$

and the slopes of the price trajectory and log price trajectory are given by

$$\frac{dp_t}{dt} = \frac{rce^{rt}}{e^{rcR_0/A} - 1}$$

and

$$\frac{d \log p_t}{dt} = \frac{rc}{\left(e^{rcR_0/A} - 1\right)ce^{-rt} + c}$$

From these equations it is easy to see that an upward shift in the demand curve, i.e., an increase in A, leads to an increase in the price level P_t, and an increase in the slopes of both the price trajectory and the log price trajectory. An increase in the level of extraction cost, c, leads to an increase in price, but a decrease in the slopes of the price and log price trajectories. Finally, if new discoveries result in an unexpected increase in the reserve level, R_0, this will cause a decrease in price, and will also lead to decreases in the slopes of the price and log price trajectories. For most depletable resources, one would expect demand, extraction costs, and reserves all to fluctuate continuously and unpredictably over time. Whether or not the processes that these variables follow are stationary is an open matter. But in either case, we would expect price to revert to a trend line with a level and slope that likewise fluctuate over time.

If demand, extraction costs, and reserves change very infrequently but by large, discrete amounts, then a switching model of the sort estimated by Perron (1989) is appropriate as a description of price. It is important to note that this trend line to which price reverts, and which represents long-run total marginal cost, is itself unobservable. We might estimate the "parameters" of the trend line (and hence marginal cost itself) at any point in time using data up to that point, but those parameters (and hence the corresponding level of marginal cost) will change over time.

Appendix D. The Derivation of Shifting Trend Model of Price

A model of long-run commodity price evolution should incorporate two key characteristics: (i) reversion to an unobservable long-run total marginal cost, which follows a trend; and (ii) continuous random fluctuations in both the level and slope of that trend. A continuous time model that has these characteristics is a version of the multivariate Ornstein-Uhlenbeck process.

Suppose, first, that the log price follows a simple trending Ornstein-Uhlenbeck (OU) process. If the trend is quadratic, that process can be written in continuous time as:

$$\mathrm{d}p = -\gamma p \mathrm{d}t + \sigma \mathrm{d}z$$

where $\overline{p} = p - \alpha_0 - \alpha_1 t - \alpha_2 t^2$ is the detrended price. In terms of the price level itself, this is equivalent to:

$$dp = \left[-\gamma\left(p - \alpha_0 - \alpha_1 t - \alpha_2 t^2\right) + \alpha_1 + 2\alpha_2 t\right]dt + \sigma dz$$

Note that the parameter γ describes the rate of reversion to the (fixed) trend line. If $\gamma = 0$, the log price follows an arithmetic Brownian motion (so price is a geometric Brownian motion), and the variance ratio would approach 1.

In the bivariate case, we could write the multivariate Ornstein-Uhlenbeck process as:

$$d\overline{p} = \left(-\gamma \overline{p} + \lambda x\right)dt + \sigma dz_p$$

where x is itself an Ornstein-Uhlenbeck process:

$$dx = -\delta x dt + \sigma dz_x$$

where dzp and dzx may be correlated.

In discrete time, this process would be given by the following equations:

$$p_t = \alpha_p p_{t-1} + \lambda x_{t-1} + \varepsilon_{p,t}$$

$$x_t = \alpha_x x_{t-1} + \varepsilon_{x,t}$$

where ε_{pt} and ε_{xt} are normally distributed with mean 0, and with some covariance. As discussed below, if x were unobservable, this process could be estimated using the Kalman filter.

Pindyck (1999) considers a slightly more general multivariate version of this process that allows for fluctuations in both the level and slope of the trend. In particular, suppose that in continuous time the process for the log price is:

$$d\overline{p} = \left(-\gamma \overline{p} + \lambda_1 x + \lambda_2 yt\right)dt + \sigma dz_p$$

with

$$dx = -\delta_1 x dt + \sigma_x dz_x$$
$$dy = -\delta_2 y dt + \sigma_y dz_y$$

Rewriting yields

$$dp = [-\gamma\,(p - \alpha_0 - \alpha_1 t - \alpha_2 t\,) + \alpha_1 + 2\alpha_2 t + \lambda_1 x + \lambda_2 yt]dt + \sigma dz_p\,.$$

Combining terms, we can write this as:

$$dp = (-\gamma p - \alpha'_0 - \alpha'_1 t - \alpha'_2 t^2 + \lambda_1 x + \lambda_2 yt)dt + \sigma dz_p\,.$$

These equations describe a process in which the log price reverts to a trend line with level and slope that fluctuate stochastically, and which may or may not be observable

These equations imply the following discrete-time model, which will be the basis of the empirical work that follows:

$$p_t = \rho p_{t-1} + b_1 + b_2 t + b_3 t^2 + \phi_{1t} + \phi_{2t} t + \varepsilon_t$$

$$\phi_{2t} = c_2 \phi_{2,t-1} + v_{2t}$$

ϕ_{1t} and ϕ_{2t} are treated as unobservable state variables. This is appropriate, since marginal cost at any point in time, the resource reserve base, and the demand parameters are all unobservable.

If we make the further assumption that the distribution of the error terms, ε_t, v_{1t}, and v_{2t} is multivariate normal and that ε_t is uncorrelated with v_{1t} and v_{2t}, then a natural estimator of this system of equations is the Kalman filter. To simplify matters, we assume that the error in the state equations, v_{1t}, and v_{2t}, are uncorrelated.