Volatility and a Century of Energy Markets Dynamics*

Apostolos Serletis† and Libo Xu
Department of Economics
University of Calgary
Calgary, Alberta T2N 1N4
Canada

Forthcoming in: *Energy Economics*

January 20, 2016

*We thank Adonis Yatchew and Dimitri Dimitropoulos for making their fuels data available to us. We also thank two referees for comments that greatly improved the paper.

†Corresponding author. Phone: (403) 220-4092; Fax: (403) 282-5262; E-mail: Serletis@ucalgary.ca; Web: http://econ.ucalgary.ca/serletis.htm.

Abstract:

How similar is the price behavior of oil, natural gas, and coal? Are there any interactions among these three fuel prices and their volatilities? Using the Yatchew and Dimitropoulos (2015) annual data for the United States, over the period from 1870 to 2014, and state-of-the-art econometric methodology, we explore for spillovers and interactions among the three energy markets. In doing so, we use a range of univariate and multivariate volatility models. The key contribution to the literature is the estimation of a trivariate BEKK model that allows for the interdependence of oil, natural gas, and coal returns and volatilities, using the longest span prices that have ever been studied before.

JEL classification: E32, C32.

Keywords: Oil, Natural gas, Coal, Volatility, VARMA, GARCH-in-Mean model.

1 Introduction

In recent years, multivariate volatility models are becoming standard in economics and finance. These models, first proposed by Bollerslev *et al.* (1988), allow for rich dynamics in the variance-covariance structure of time series, making it possible to model spillovers in both the values and the conditional variances of the series under study. They can be used to investigate a large number of issues in economics and finance. For example, as Bauwens *et al.* (2006, p. 79) put it, "is the volatility of a market leading the volatility of other markets? Is the volatility of an asset transmitted to another asset directly (through its conditional variance) or indirectly (through its conditional covariances)? Does a shock on a market increase the volatility on another market, and by how much? Is the impact the same for negative and positive shocks of the same amplitude?"

In this paper we estimate univariate and multivariate volatility models for the prices of three hydrocarbons — oil, natural gas, and coal — using the Yatchew and Dimitropoulos (2015) annual data for the United States that span over a century, from 1870 to 2014. The key contribution to the literature is the estimation of a trivariate volatility model that explores the interdependence of oil, natural gas, and coal prices and volatilities, using the longest span data that have ever been studied before. Existing studies have used multivariate volatility models to explore the relationships among several electricity markets [see Worthington *et al.* (2005)], between oil and natural gas markets [see Ewing *et al.* (2002) and Serletis and Shahmoradi (2006)], and between oil markets and financial or macroeconomic indicators [see Lee *et al.* (1995), Sadorsky (2012), Elder and Serletis (2010), and Rahman and Serletis (2012)]. However, to the best of our knowledge, no study has used multivariate volatility models to model oil, natural gas, and coal prices in a systems context, although there is similar work by Efimova and Serletis (2014) who model crude oil, natural gas, and electricity prices using daily data for the United States over a short period, from 2001 to 2013.

Our research is distinguished from the current literature in a number of ways. The present paper is significantly different from Efimova and Serletis (2014), and even though our research question is similar to Ewing *et al.* (2002), who discuss the volatility transmission in the oil and natural gas markets, our empirical models have more features. We use a long span, low frequency data set, over the period from 1870 to 2014, and model the returns of oil, natural gas, and coal which are all raw non-renewable energy resources and are the main utilized fuels for energy worldwide. Although temporally aggregated data exhibit smaller GARCH effects than higher frequency data, since the persistence of conditional volatility tends to increase with the sampling frequency, we fit the univariate and multivariate volatility models to the low frequency data. Also, there is no simple method which links the presented ARCH and GARCH effects to the estimation results at higher frequencies — see Drost and Nijman (1993), Hafner (2008), and Zivot (2009) for more details regarding these issues.

This paper is also significantly different from Efimova and Serletis (2014). Their work uses daily data (over the period from January 2, 2001 to April 26, 2013) and focuses on two primary forms of energy (oil and natural gas) and an energy carrier (electricity), produced using these two primary energy sources. In this paper we derive univariate forecasting model specifications for three primary energy sources — oil, natural gas, and coal — based on annual time series data. We also model and investigate the spillovers and interactions among the three primary energy sources using the longest span annual price series (from 1870 to 2014) that have ever been studied before. Forecasts of hydrocarbons prices and estimates of the spillovers and interactions among the hydrocarbons markets affect the economic outlook of the country, guiding the development of natural resources and investments in infrastructure. They also play an important role in firm investment and production. Users of hydrocarbons price forecasts and estimates of spillovers and interactions among the hydrocarbons markets include governments, central banks, international organizations, and a range of industries in the broad areas of manufacturing, mining, and utilities.

We focus on the price of crude oil, although oil is not consumed directly but is used as a factor of production in the refining industry (in the production of gasoline, diesel, heating oil, and jet fuel). We are interested in the interaction between the oil market and the coal market, because even before the emergence of a global oil market in the 1970s, there was a well developed market for coal. In fact, coal was the primary fuel in shipping until the 1920s and in railroading and home heating until the 1950s. We are also interested in the interaction between the oil and coal markets and the natural gas market, because natural gas competes with coal in producing electricity and in the manufacturing of chemicals and metals. In 2011, for example, 35.3% of the U.S. energy was generated by oil, while natural gas contributed 24.8%, and coal contributed 19.7%. These fuels exhibit some degree of substitutability and their prices are closely correlated, suggesting that we model interactions among the different markets in a systems context. See Kilian (2015) for more details regarding a historical perspective of the evolution of the oil, coal, and natural gas markets in the United States.

The outline of the paper is as follows. In Section 2 we present the Yatchew and Dimitropoulos (2015) price data on the three hydrocarbons and investigate their time series properties using unit root and stationarity tests. In Section 3 we present two alternative formulations of univariate volatility models for each fuel return and in Section 4 we estimate a trivariate volatility model that explores the interdependence of oil, natural gas, and coal returns and volatilities. The final section briefly concludes the paper.

2 Data and Basic Facts

We use the Yatchew and Dimitropoulos (2015) annual data on oil and coal for the period from 1870 to 2014 and for natural gas for the period from 1919 to 2014. It is the same data that Yatchew and Dimitropoulos (2015) use and were obtained from Manthy (1978) prior to 1973 and augmented using prices published by the U.S. Energy Information Administration from 1973 to 2014. See Yatchew and Dimitropoulos (2015) for more details regarding the fuels data.

Figures 1 to 3 plot the natural logs of the oil, natural gas, and coal prices, and their returns, respectively. The oil price is dollars per barrel, the natural gas price is cents per 1000 cubic feet, and the coal price is dollars per ton. Table 1 presents summary statistics for the log levels, $\ln o_t$, $\ln g_t$, $\ln c_t$, and the first differences of the logs, $\Delta \ln o_t$, $\Delta \ln g_t$, and $\Delta \ln c_t$, of the three price series. In general, the p-values for skewness and kurtosis point to significant deviations from symmetry and normality with both the logged series and the first differences of the logs. In fact, the Jarque-Bera (1980) test statistic, distributed as a $\chi^2(2)$ under the null hypothesis of normality, rejects the null hypothesis.

One interesting feature of the data is the contemporaneous correlation between the different prices series. These correlations are reported in Table 2 for log levels (in panel A) and for first differences of log levels (in panel B). To determine whether these correlations are statistically significant, Pindyck and Rotemberg (1990) is followed and a likelihood ratio test of the hypotheses that the correlation matrices are equal to the identity matrix is performed. The test statistic is

$$-2 \ln \left(|R|^{N/2} \right)$$

where $|R|$ is the determinant of the correlation matrix and N is the number of observations. This test statistic is distributed as χ^2 with $0.5q(q-1)$ degrees of freedom, where q is the number of series. The test statistic is 724.811 with a p-value of 0.000 for the logged prices and 74.145 with a p-value of 0.000 for the first differences of the logs. Clearly, the hypothesis that these price series are uncorrelated is rejected. Notice, however, that the correlations indicate a weak relationship between the natural gas and coal price series. The correlation patterns documented in Table 2 manifest in the graphical representation of the logged levels of the series in Figure 4.

The first step in volatility modelling and the examination of trends in a set of variables is to test for the presence of a stochastic trend (a unit root) in the autoregressive representation of each individual series. Nelson and Plosser (1982) argue that most macroeconomic and financial time series have a unit root (a stochastic trend), and describe this property as one of being 'difference stationary' so that the first difference of a time series is stationary. An alternative 'trend stationary' model has been found to be less appropriate.

We conduct a battery of unit root and stationary tests in panel A of Table 3 in the natural logs of each price series. In particular, we use the Augmented Dickey-Fuller (ADF) test [see Dickey and Fuller (1981)] and the Dickey-Fuller GLS test [see Elliot, Rothenberg, and Stock (1996)], assuming both a constant and trend, to determine whether the series have a unit root. The optimal lag length is taken to be the order selected by the Bayesian information criterion (BIC), after we assume a maximum lag length of 4 for each series. Moreover, given that unit root tests have low power against trend stationary alternatives, we also use the KPSS test [see Kwiatkowski *et al.* (1992)] to test the null hypothesis of stationarity around a trend. As shown in panel A of Table 3, the null hypothesis of a unit root cannot in general be rejected at conventional significance levels by both the ADF and DF-GLS test statistics. Moreover, the null hypothesis of trend stationarity can be rejected at conventional significance levels by the $\hat{\eta}_\tau$ KPSS test.[1] We thus conclude that each of the three fuel prices is nonstationary, or integrated of order one, $I(1)$. In panel B of Table 3 we repeat the unit root and stationarity tests using the first differences of the logarithms of the series. The null hypotheses of the ADF and DF-GLS tests are in general rejected and the null hypothesis of the KPSS test cannot be rejected, suggesting that the first differences of the logarithms of the series are stationary, or integrated of order zero, $I(0)$.

Due to the presence of unit roots in the logged levels, in the next section we estimate all univariate volatility models using the first differences of the logarithms of the series, $\Delta \ln o_t$, $\Delta \ln g_t$, and $\Delta \ln c_t$. Therefore, we actually model each fuel return in the univariate volatility models.

3 Univariate Volatility Modelling

This section presents a range of univariate GARCH models for oil, natural gas, and coal returns. As Efimova and Serletis (2014, pp. 265) put it, "univariate GARCH models have been neglected by academic research in recent years despite their strong performance. Moreover, univariate models produce accurate forecasts, converge much faster in maximum likelihood estimation, and allow for the inclusion of a significant number of additional parameters whereas multivariate systems quickly become overparameterized."

3.1 Oil

In this section, we estimate two GARCH (1,1) models that differ in their mean equations to model the annual oil return. The first mean equation is a random walk with drift

$$\Delta \ln o_t = \beta_0 + \varepsilon_t \tag{1}$$

where $\Delta \ln o_t$ is the first difference of the logarithm of the oil price. The number of autoregressive (and moving-average) terms in equation (1) was chosen using the Bayesian Information Criterion (BIC). The second mean model is equation (1) augmented with a GARCH-in-Mean term, as follows

$$\Delta \ln o_t = \beta_0 + \beta_1 h_t + \varepsilon_t \tag{2}$$

where h_t is the time-varying variance of the oil return.

[1] It should be noted that the KPSS test has limitations, as there are size distortions from the number of observations and the persistence of the data. For example, some data tend to be persistent even when they are actually stationary, and the KPSS test may reject the null of stationarity under this circumstance. In this regard, Caner and Kilian (2001) suggest that monthly and quarterly data based on small samples mainly suffer from this issue.

We use a GARCH (1,1) specification for the variance equation and also include the GJR asymmetry coefficient of Glosten *et al.* (1993), $\varepsilon_{t-1}^2 \times I_{\varepsilon<0}(\varepsilon_{t-1})$, which captures the disproportionate response of a commodity's variance to unexpected return decreases. The resulting variance equation is

$$h_t = c_0 + a_1\varepsilon_{t-1}^2 + b_1 h_{t-1} + d_1\varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1}). \tag{3}$$

The empirical estimates for both models, equations (1) and (3) and equations (2) and (3), using the annual data from 1870 to 2014, are presented in panels A and B of Table 4. The estimation is performed in RATS 8.3. The GARCH-in-Mean coefficient in the extended oil return equation (2) is large and statistically significant, suggesting that current oil return volatility has a significant impact on the direction or magnitude of the oil return. Moreover, since the two GARCH models (the baseline and extended) are nested, their log-likelihood values are directly comparable. We find that the extended model, equations (2) and (3), has a higher log-likelihood value than the baseline model, equations (1) and (3).

The most striking finding from the variance equation in panel B of Table 4 is the contrast of very high ARCH coefficient on ε_{t-1}^2 and a moderate GARCH coefficient on h_{t-1}. This suggests that volatility reacts intensely to oil return movements, but shocks to the conditional variance die out quickly, thus rendering volatility spiky. Moreover, we find negative and statistically significant asymmetric effects (-0.780 with a p-value of 0.011).

Additionally, in panel C of Table 4 we report the log-likelihood values and diagnostic test statistics for the standardized residuals, $\hat{\varepsilon}_t = \varepsilon_t/\sqrt{h_t}$, including descriptive statistics, the Jarque-Bera statistic, the Ljung-Box Q test for residual autocorrelation, and the McLeod-Li Q^2 test for squared residual autocorrelation. Both tests assume the null hypothesis that the data are independently distributed, and an alternative hypothesis of autocorrelation. The Q and Q^2 statistics are reported for 10 lags, with p-values in parentheses. The Ljung-Box and McLeod-Li tests pass at conventional significance levels, suggesting that there is no significant evidence of autocorrelation in the levels or squares of the standardized residuals.

Finally, in panel D of Table 4 we present two structural break tests for each of the baseline and extended models. In part, this is related to the treatment of the data as having been governed by a single process over the course of one hundred years — a span over which energy markets in the United States (and globally) have undergone very significant structural changes due to technological innovation, revisions in regulatory regimes and market structures, the emergence of new players such as OPEC, and changes in the sources of supply and demand. We perform the Andrews and Ploberger (1994) and the Andrews-Quandt structural break tests for a single structural break at an unknown point within the sample. In particular, we generate a series of LM statistics at each of the points in the middle range of the data set, and use the maximum of the LM statistics as the test statistic in the case of the Andrews-Quandt test and the geometric mean of the LM statistics in the case of the Andrews and Ploberger (1994) test. According to the results in panel D of Table 4 (based on asymptotic p-values computed as in Hansen (1997), the baseline model suggests a structural break in 1972, very close to the first oil crisis in 1973. However, this break is highly insignificant. Moreover, the structural break in the extended crude oil model is not significant as well. Therefore, we conclude that our model specifications fit the series of oil returns very well without a structural break, suggesting that the assumption of a single data generation process over this time-frame is a good one to make.

3.2 Natural Gas

Like in the previous subsection, we estimate two univariate GARCH (1,1) models of natural gas returns which again differ only in their mean equations. The first model uses the following baseline MA (3) mean equation

$$\Delta \ln g_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \beta_3 \varepsilon_{t-3} + \varepsilon_t \tag{4}$$

where $\Delta \ln g_t$ is the first difference of the logarithm of the natural gas price. Again, the number of (AR and) MA terms in (4) was chosen using the Bayesian Information Criterion. The second model for the mean equation is a MA (3) model augmented with a GARCH-in-Mean term

$$\Delta \ln g_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \beta_3 \varepsilon_{t-3} + \beta_4 h_t + \varepsilon_t \tag{5}$$

where h_t now denotes the time-varying variance of the natural gas return.

We use the same GARCH (1,1) variance specification as we did for oil

$$h_t = c_0 + a_1 \varepsilon_{t-1}^2 + b_1 h_{t-1} + d_1 \varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1}). \tag{6}$$

Empirical estimates for models (4) and (6) and (5) and (6), using the annual data from 1919 to 2014, are presented in Table 5, in the same fashion as those for oil in Table 4. Basically we get consistent estimates across the baseline and the extended model. The MA coefficients are all positive and statistically significant. Moreover, the GARCH-in-Mean coefficient in the extended natural gas return equation (5) is positive and statistically significant (1.698 with a p-value of 0.094), suggesting that natural gas return volatility has a large and significant effect on the change of the natural gas return.

Estimates of the variance equation coefficients are reported in panel B of Table 5. We find positive and statistically significant ARCH and GARCH effects, 0.552 with a p-value of 0.013 and 0.651 with a p-value of 0.000, respectively. In particular, we find that natural gas return volatility reacts less intensely to natural gas return movements compared to the case of oil and that shocks to the conditional variance are more persistent, rendering natural gas return volatility less spiky compared to the case of oil return volatility. Moreover, we find no significant asymmetric effects.

Panel C of Table 5 presents the log-likelihood values for models (4) and (6) and (5) and (6), as well a range of statistics and diagnostic tests applied to the standardized residuals, in the same fashion as for oil in Table 4. These statistics and diagnostic tests suggest that the standardized residuals, $\hat{\varepsilon}_t$, are normally distributed and serially uncorrelated. We also report the two structural break tests in panel D of Table 5, and find that there does not exist a structural break in the series of natural gas returns.

3.3 Coal

Following the pattern established in the previous two subsections, we also construct two GARCH models for coal, starting with a baseline AR (1) model

$$\Delta \ln c_t = \beta_0 + \beta_1 \Delta \ln c_{t-1} + \varepsilon_t \tag{7}$$

where $\Delta \ln c_t$ is the first difference of the logarithm of the coal price. Again, the number of AR and MA terms in equation (7) was chosen using the Bayesian Information Criterion. The second mean model for coal is an AR (1) model augmented with a GARCH-in-Mean term

$$\Delta \ln c_t = \beta_0 + \beta_1 \Delta \ln c_{t-1} + \beta_2 h_t + \varepsilon_t. \tag{8}$$

The variance equation for each model is again an asymmetric GARCH (1,1)

$$h_t = c_0 + a_1 \varepsilon_{t-1}^2 + b_1 h_{t-1} + d_1 \varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1}). \tag{9}$$

Empirical estimates of models (7) and (9) and (8) and (9), using annual data from 1870 to 2014, are reported in Table 6, in the same fashion as those for oil and natural gas in Tables 4 and 5, respectively. Again we get consistent estimates across the baseline and the extended model. Moreover, the GARCH-in-Mean coefficient in the extended coal return equation (8) is statistically insignificant and the ARCH and GARCH terms in the variance equation (9) are both positive and statistically significant, 1.057 with a p-value of 0.000 and 0.546 with a p-value of 0.000, respectively. We also find negative but statistically insignificant asymmetric effects (-0.317 with a p-value of 0.345), and there is no evidence of a structural break in the series of coal returns.

4 Multivariate Volatility Modelling

In this section we follow Efimova and Serletis (2014) and estimate a trivariate vector autoregressive moving average (VARMA), GARCH-in-Mean, BEKK model [see Engle and Kroner (1995) for more details], which models natural gas, coal, and oil returns and volatilities as a system. We prefer the VARMA framework because it allows us to capture features of the data generating process in a more parsimonious way without adding a large number of parameters or lagged variables. This formulation allows us to model the transmission of return volatility from one fuel to another, and estimate the effects of volatility in any of the three markets on the return of each fuel.

For the mean equation, we choose a trivariate VARMA (1,1) specification in the first differences of the logarithms of the natural gas, coal, and oil prices (multiplied by 100) forming the dependent variables

$$\boldsymbol{z}_t = \mathbf{a} + \boldsymbol{\Gamma} \boldsymbol{z}_{t-1} + \boldsymbol{\Psi}\sqrt{\boldsymbol{h}_t} + \boldsymbol{\Theta} \boldsymbol{e}_{t-1} + \boldsymbol{\epsilon}_t \tag{10}$$

$$\boldsymbol{\epsilon}_t | \, \Omega_{t-1} \sim (\mathbf{0}, \, \boldsymbol{H}_t) \, , \quad \boldsymbol{H}_t = \begin{bmatrix} h_{gg,t} & h_{gc,t} & h_{go,t} \\ h_{cg,t} & h_{cc,t} & h_{co,t} \\ h_{og,t} & h_{oc,t} & h_{oo,t} \end{bmatrix}$$

where Ω_{t-1} is the information set available in period $t-1$ and

$$\boldsymbol{z}_t = \begin{bmatrix} \Delta \ln g_t \\ \Delta \ln c_t \\ \Delta \ln o_t \end{bmatrix}; \quad \boldsymbol{\epsilon}_t = \begin{bmatrix} \epsilon_{g,t} \\ \epsilon_{c,t} \\ \epsilon_{o,t} \end{bmatrix}; \quad \boldsymbol{h}_t = \begin{bmatrix} h_{gg,t} \\ h_{cc,t} \\ h_{oo,t} \end{bmatrix};$$

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{bmatrix}; \quad \boldsymbol{\Psi} = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix}; \quad \boldsymbol{\Theta} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{bmatrix}.$$

In the trivariate model we use the first differences of the logarithms of the three fuel prices. The reason for using the first differences of the logarithms is that we find no cointegration in the log levels of the three fuel prices after implementing the Engle and Granger (1987) and Johansen (1988) cointegaration tests (these results are available upon request). Also, we do not consider structural breaks, since we have already shown that each fuel return does not have a structural break.

For the variance equation, we use the asymmetric BEKK model introduced by Grier *et al.* (2004). We choose the BEKK(1,1,1) specification, which is a multivariate extension of GARCH(1,1). The resulting variance equation is

$$\boldsymbol{H}_t = \boldsymbol{C}'\boldsymbol{C} + \boldsymbol{B}'\boldsymbol{H}_{t-1}\boldsymbol{B} + \boldsymbol{A}'\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1}\boldsymbol{A} + \boldsymbol{D}'\boldsymbol{u}_{t-1}\boldsymbol{u}'_{t-1}\boldsymbol{D} \tag{11}$$

where $\boldsymbol{C}$, $\boldsymbol{B}$, $\boldsymbol{A}$, and $\boldsymbol{D}$ are 3×3 matrices with $\boldsymbol{C}$ being a triangular matrix to ensure positive definiteness of $\boldsymbol{H}$. This specification allows past volatilities, $\boldsymbol{H}_{t-1}$, as well as lagged values of $\boldsymbol{\epsilon}\boldsymbol{\epsilon}'$ and $\boldsymbol{u}\boldsymbol{u}'$ to show up in estimating current volatilities of natural gas, coal, and oil returns. The asymmetry vector $\boldsymbol{u}_{t-1}$ is defined by $\boldsymbol{u}_{t-1} = \boldsymbol{\epsilon}_{t-1} \circ I_{\boldsymbol{\epsilon}<0}(\boldsymbol{\epsilon}_{t-1})$ where $\circ$ is the Hadamard product. Since $\boldsymbol{H}$ matrix is symmetric, equation (11) produces six unique equations modeling the dynamic variances of natural gas, coal, and oil returns, as well the covariances between them. We refrain from adding additional explanatory variables, since our model already contains 30 mean equation parameters and 33 variance equation parameters, for a total of 63 free parameters.

4.1 Empirical Estimates

The trivariate BEKK model formed by equations (10) and (11) was estimated in Estima RATS (version 8.3) using quasi-Maximum Likelihood. We used the BFGS (Broyden, Fletcher, Goldfarb & Shanno) estimation algorithm, which is recommended for GARCH models, combined with the derivative-free Simplex pre-estimation method. Table 7 reports the coefficients obtained (with p-values in parentheses), as well as key diagnostics for the standardized residuals

$$z_{jt} = \frac{\epsilon_{jt}}{\sqrt{h_{jt}}}$$

for $j =$ natural gas, coal, and oil.

The autoregressive coefficients in the $\mathbf{\Gamma}$ matrix are significant along the main diagonal, suggesting that for each of the fuels, today's return is a good predictor of tomorrow's return. In particular, natural gas returns will be high if they were high in the previous period, whereas coal and crude oil returns will be low if they were high in the previous period. Moreover, natural gas and oil returns experience significant spillover effects from each of the other markets whereas coal returns only receive spillover effect from the natural gas market.

The moving-average coefficients along the diagonal of the $\mathbf{\Theta}$ matrix are large and significant, suggesting that the dynamics of fuel returns are consistent with a typical ARMA process. In particular, only the natural gas return has a negative relationship with shocks originating in its own market. Another interesting result is that there are spillover effects in the MA terms as well. The natural gas and oil returns receive spillover effects form each of the other markets. For example, an unexpected increase of one unit in the coal return is associated with a 0.707 unit increase in the natural gas return in the next period. Similarly, an unexpected increase of one unit in the oil return is associated with a 0.407 unit increase in the natural gas return next period. We find that news in the natural gas and oil markets have very little effect on the coal market, suggesting that coal returns do not respond to surprise developments in each of the natural gas and oil markets.

The estimates of the GARCH-in-Mean coefficients matrix $\mathbf{\Psi}$ suggest that each of the fuel returns is affected by its own volatility. We find that only the natural gas return decreases when its own volatility increases, suggesting that natural gas returns will be low if there is a lot of uncertainty about the natural gas market. On the other hand, the returns of coal and oil increase as their volatilities increase. Also, the volatility in each market not only affects its own return but also has spillover effects on the other markets. For example, the natural gas return will increase if the volatility of coal and oil returns increases. Interestingly, the $\mathbf{\Psi}$ matrix shows that the response of each fuel return to the volatility in its own market and the response of each fuel return to the volatility in the other market are totally opposite. For example, the natural gas return decreases as its own volatility increases. However, the natural gas return increases as the volatilities increase in the other markets. Moreover, this pattern is also valid in the other two markets. Overall, we find that the volatility in each fuel return influences its own return. Moreover, the volatility shows significant spillover effects across markets.

Finally, we note that the spillover effects which we find in the autoregressive coefficients, moving-average coefficients, and the GARCH-in-Mean coefficients, are rather asymmetric in terms of the sign of the coefficients. This asymmetry could be found once we compare the significant coefficients (with p values less than 0.05) which lay on the off-diagonal of each of the $\boldsymbol{\Gamma}$, $\boldsymbol{\Theta}$ and $\boldsymbol{\Psi}$ matrices. For example, ψ_{12} implies that the natural gas return increases as the coal return volatility increases. The spillover effect of natural gas return volatility on the coal return is given by ψ_{21} and suggests the the coal return decreases as the natural gas return volatility increases. Therefore, the spillover effects of volatility on return are asymmetric between the natural gas and coal markets. Moreover, this pattern also exists when we look at γ_{12} and γ_{21}, γ_{13} and γ_{31}, θ_{13} and θ_{31}, and ψ_{13} and ψ_{31}. Obviously, these asymmetries show up between the natural gas and coal markets and exists between the natural gas and oil markets. Another interesting finding is related to the spillover effects between the coal and oil markets. It is found the spillovers are unidirectional. For example, γ_{32} implies that the oil return will be low if the coal return is high in the previous period. However, the oil return in the previous period has no effect on the coal return this period, since γ_{23} is statistically insignificant. This pattern could be found in the Θ matrix as well. Therefore, we mainly observe spillovers from coal to oil.

The asymmetry found in the spillover effects given by the autoregressive coefficients, the moving-average coefficients, and the GARCH-in-Mean coefficients is in general consistent with the economic explanation provided by Hossain and Serletis (2016) who investigate interfuel substitution in the United States using annual price and quantity data over almost the same sample period as in the present paper. They show that the elasticities of substitution among oil, natural gas, and coal are positive and statistically significant, indicating substitutability and significant interactions among the three primary energy sources.

The estimates for the variance equation show high and significant ARCH coefficients along the main diagonal of the $\boldsymbol{A}$ matrix, except for the coal market. It suggests that volatility is persistent in the natural gas and oil markets. In particular, the oil return has the most persistent ARCH effect, $\hat{a}_{33} = 1.085$, implying an ARCH effect of 1.085^2. The off diagonal elements of the $\boldsymbol{A}$ matrix indicate significant spillover ARCH effects as well. For example, an unexpected change in the natural gas and coal returns will increase the volatility of the oil return, since $\hat{a}_{13} = 1.127$ (with a p-value of 0.000) and $\hat{a}_{23} = -1.081$ (with a p-value of 0.000). Also, an unexpected change in the coal return will increase the volatility of the natural gas return ($\hat{a}_{21} = -0.589$ with a p-value of 0.000). A particular finding about ARCH effects is that the volatility of the coal return is affected by shocks in the oil and natural gas returns without responding to shocks in its own return. It implies that natural gas and oil shocks play a dominant role in the coal return volatility.

The main diagonal coefficients of the $\boldsymbol{B}$ matrix indicate that there are statistically significant GARCH effects only in the markets of natural gas and oil. This phenomenon happens in the ARCH effects as well. Moreover, the coal return volatility does not respond to the volatility in natural gas and oil returns. Therefore, the coal return volatility can be mainly explained by shocks in the natural gas and oil markets when we model the interactions among the three fuel markets. Regarding spillover GARCH effects, these effects are only seen in the oil market. It is found that the volatility in natural gas and coal returns will increase the oil return volatility slightly, because the spillover GARCH effect of natural gas (0.076^2) and the spillover GARCH effect of coal (0.106^2) are small and significant.

Finally, the $\boldsymbol{D}$ matrix presents the asymmetric ARCH effects in the fuel markets. The diagonal coefficients of the $\boldsymbol{D}$ matrix suggest that negative shocks (bad news) to fuel returns are associated with more volatility in the natural gas and oil markets, compared with positive shocks (good news). Moreover, some spillover ARCH effects show asymmetries as well. The natural gas return will be more volatile if there is a negative shock to the coal return rather than a positive shock. In the case of coal, negative shocks to the natural gas and oil markets are associated with more volatility in the coal returns, compared with positive shocks. In the oil market, the oil return will be more volatile if there is a negative shock to the natural gas return. Overall, we find that positive and negative shocks play different roles in the volatility transmission across the three fuel markets.

Overall, the trivariate VARMA-BEKK model shows significant interactions among the three fuel returns, including spillovers from surprise return changes in one fuel to the return volatility of another fuel. Thanks to the long span data set and the powerful VARMA-BEKK model structure, we are able to not only detect these spillover effects, but also estimate their magnitude.

5 Conclusion

Fuel price and its return volatility are of great concern to market participants and policymakers. Being able to accurately forecast the volatility and predict its spillover effects carries direct implications for hedging and derivatives trading. Motivated by these considerations, we estimate univariate and multivariate volatility models for the prices of three hydrocarbons — oil, natural gas, and coal — using the longest span data that have ever been studied before. We contribute to the understanding of fuel price and fuel return volatility in energy markets, and suggest several effective models that would be of use to energy market participants, derivatives market participants, large energy consumers interested in hedging strategies, and policymakers. In particular, the use of low frequency data over nearly a 100-year time horizon in this paper help us understand the interactions of energy markets in the long term.

The optimal choice of the model would depend on the research question (e.g. forecasting versus price and volatility spillovers). The advantage of the multivariate volatility model is the potential to investigate the interactions among all three fuel prices, their returns, and their volatilities, which makes it possible for us to discover surprising and significant spillover effects.

References

[1] Andrews, D. and W. Ploberger. "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative." *Econometrica* 64 (1994), 1383-1414.

[2] Bauwens, L., S. Laurent, and J.V.K. Rombouts. "Multivariate GARCH Models: A Survey." *Journal of Applied Econometrics* 21 (2006), 79-109.

[3] Bollerslev T., R.F. Engle, and J. Wooldridge. "A Capital Asset Pricing Model with Time Varying Covariances." *Journal of Political Economy* 96 (1988), 143–172.

[4] Caner, M. and L. Kilian. "Size Distortions of Tests of the Null Hypothesis of Stationarity: Evidence and Implications for the PPP Debate." *Journal of International Money and Finance* 20 (2001), 639-657.

[5] Dickey, D.A. and W.A. Fuller. "Likelihood Ratio Tests for Autoregressive Time Series with a Unit Root." *Econometrica* 49 (1981), 1057-1072.

[6] Drost, F.C. and T.E. Nijman. "Temporal Aggregation of GARCH Process." *Econometrica* 61 (1993), 909-927.

[7] Efimova, O. and A. Serletis. "Energy Markets Volatility Modelling Using GARCH." *Energy Economics* 43 (2014), 264-273.

[8] Elder, J. and A. Serletis. "Oil Price Uncertainty." *Journal of Money, Credit and Banking* 42 (2010), 1137-1159.

[9] Elliot, G., T.J. Rothenberg, and J.H. Stock. "Efficient Tests for an Autoregressive Unit Root." *Econometrica* 64 (1996), 813-836.

[10] Engle, R.F. and C.W. Granger. "Cointegration and Error Correction: Representation, Estimation and Testing." *Econometrica* 55 (1987), 251-276.

[11] Engle, R.F. and K.F. Kroner. "Multivariate Simultaneous Generalized ARCH." *Econometric Theory* 11 (1995), 122-150.

[12] Ewing, B.T., F. Malikb, and O. Ozfidanc. "Volatility Transmission in the Oil and Natural Gas Markets." *Energy Economics* 24 (2002), 525-538.

[13] Glosten, L.R., R. Jaganathan, and D. Runkle. "On the Relation between the Expected Value and the Volatility of the Normal Excess Return on Stocks." *Journal of Finance* 48 (1993), 1779-1801.

[14] Grier, K.B., O.T. Henry, N. Olekalns, and K. Shields. "The asymmetric Effects of Uncertainty on Inflation and Output Growth." *Journal of Applied Econometrics* 19 (2004), 551-565.

[15] Hafner, C.M. "Temporal Aggregation of Multivariate GARCH Processes." *Journal of Econometrics* 142 (2008), 467-483.

[16] Hansen, B.E. "Approximate Asymptotic p Values for Structural-Change Tests." *Journal of Business and Economic Statistics* 15 (1997), 60-67.

[17] Hossain, N. and A. Serletis. "A Century of Interfuel Substitution in the United States." Mimeo. Department of Economics, University of Calgary (2016).

[18] Jarque, C.M. and A.K. Bera. "Efficient Tests for Normality, Homoscedasticity, and Serial Independence of Regression Residuals." *Economics Letters* 6 (1980), 255-259.

[19] Johansen, S. "Statistical Analysis of Cointegration Vectors " *Journal of Economic Dynamics and Control* 12 (1988), 231-254.

[20] Kilian. L. "Energy Price Shocks." In: S.N. Durlauf and L.E. Blume (eds), The New Palgrave Dictionary of Economics (Online edition), March 2015, 1-15.

[21] Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root." *Journal of Econometrics* 54 (1992), 159-178.

[22] Lee, K., S. Ni, and R.A. Ratti. "Oil Shocks and the Macroeconomy: The Role of Price Variability." *The Energy Journal* 16 (1995), 39-56.

[23] Manthy, R.S. "Natural Resource Commodities - A Century of Statistics: Prices, Output, Consumption, Foreign Trade, and Employment in the United States, 1870–1973." Baltimore, MD: Johns Hopkins University Press (1978).

[24] Nelson, C.R. and C.I. Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications." *Journal of Monetary Economics* 10 (1982), 139-162.

[25] Pindyck, R.S. and J.J. Rotemberg. "The Excess Co-movement of Commodity Prices." *The Energy Journal* 100 (1990), 1173-1189.

[26] Rahman, S. and A. Serletis. "Oil Price Uncertainty and the Canadian Economy: Evidence from a VARMA, GARCH-in-Mean, Asymmetric BEKK Model." *Energy Economics* 34 (2012), 603-610.

[27] Sadorsky, P. "Correlations and Volatility Spillovers between Oil Prices and the Stock Prices of Clean Energy and Technology Companies." *Energy Economics* 34 (2012), 248-255.

[28] Serletis, A. and A. Shahmoradi. "Measuring and Testing Natural Gas and Electricity Markets Volatility: Evidence from Alberta's Deregulated Markets." *Studies in Nonlinear Dynamics & Econometrics* 10 (2006), Article 10.

[29] Worthington, A., A. Kay-Spratley, and H. Higgs. "Transmission of Prices and Price Volatility in Australian Electricity Spot Markets: A Multivariate GARCH Analysis." *Energy Economics* 27 (2005), 337–350.

[30] Yatchew, A. and D. Dimitropoulos. "Discerning Trends in Commodity Prices." *Macroeconomic Dynamics* (2015, forthcoming).

[31] Zivot, E. "Practical Issues in the Analysis of Univariate GARCH Models." In T.G. Anderson, R.A. Davis, J.-P. Kreiss, and T. Mikosch (eds), *Handbook of Financial Time Series*. Springer (2009), 113-156.

Figure 1. Logged Oil Price and Its Growth Rate

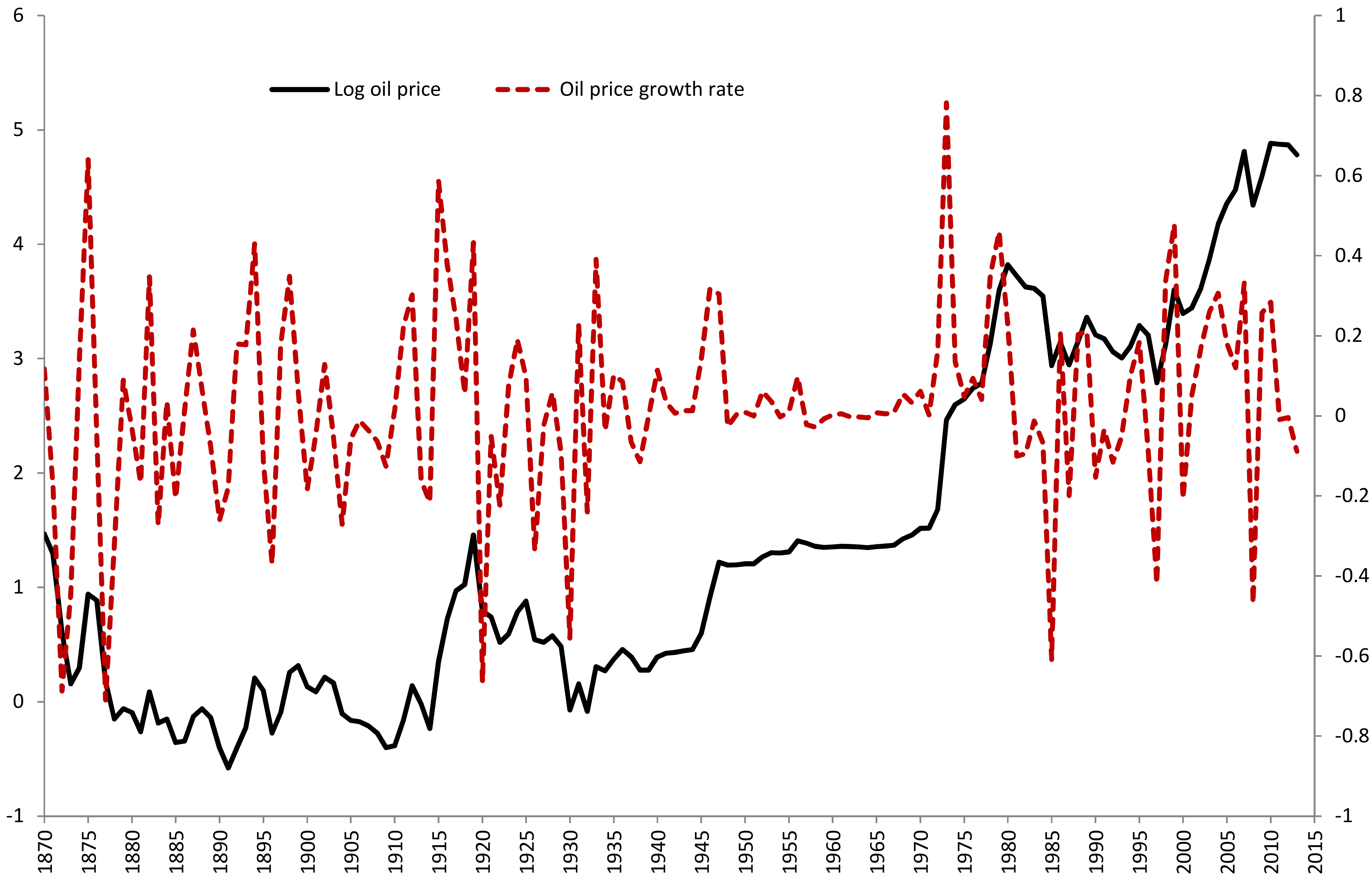

Figure 2. Logged Natural Gas Price and Its Growth Rate

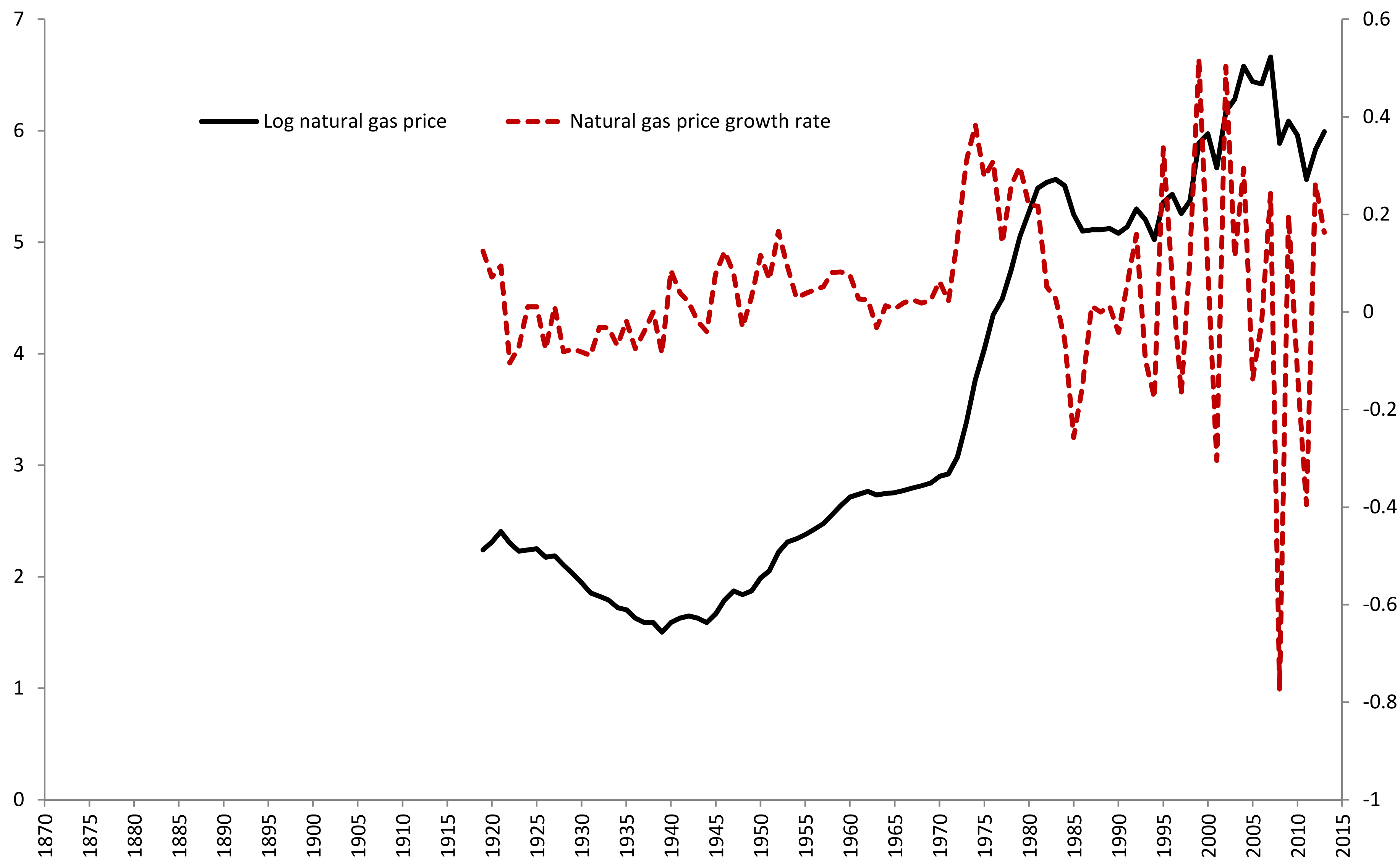

Figure 3. Logged Coal Price and Its Growth Rate

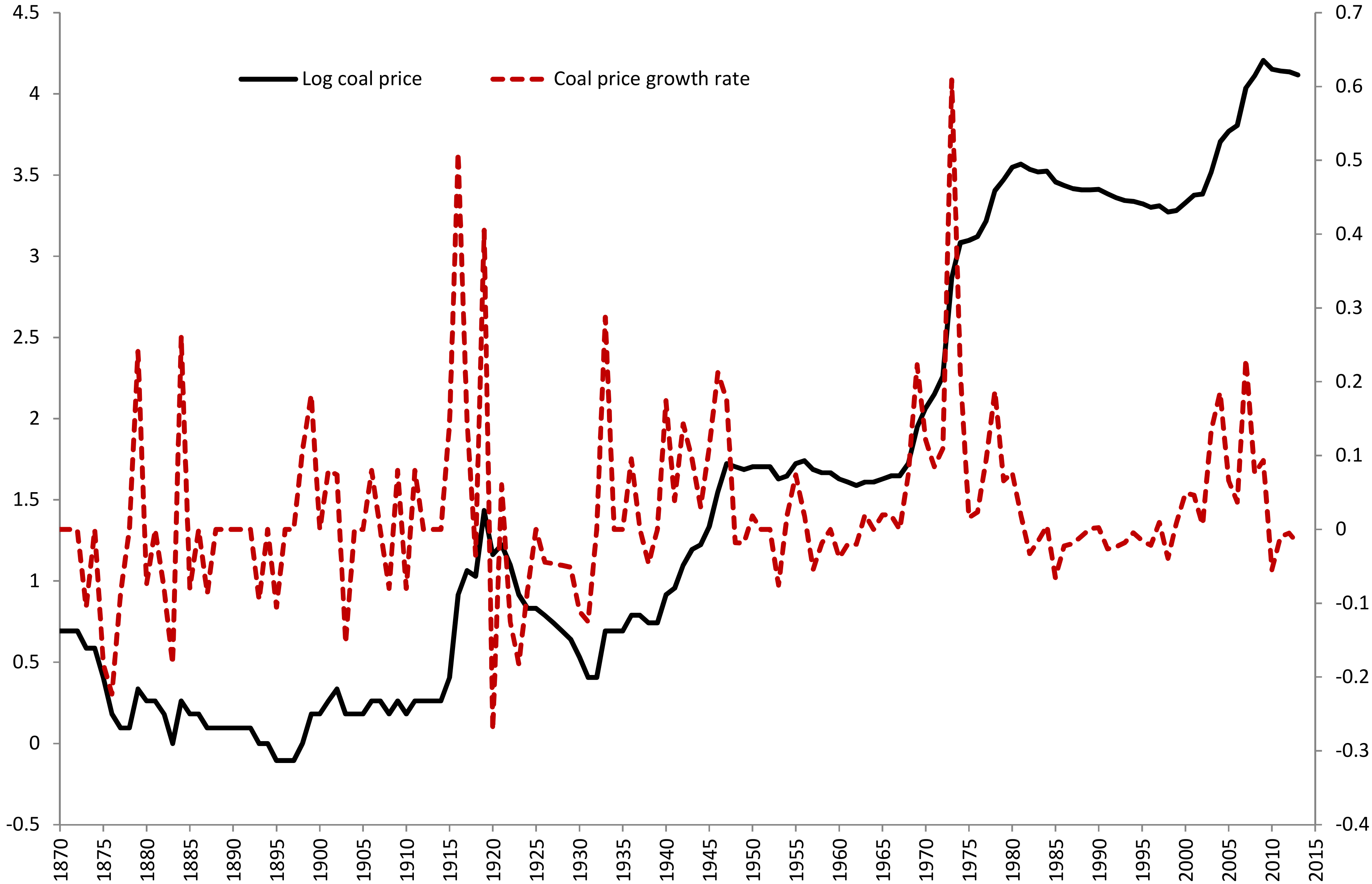

Table 1. Summary statistics

Series	Mean	Variance	*p*-values Skewness	kurtosis	Normality
		A. Log levels			
Oil	1.385	2.295	0.000	0.167	0.000
Natural gas	3.566	2.868	0.131	0.004	0.003
Coal	1.574	1.831	0.008	0.005	0.000
		B. First differences of log levels			
Oil	0.024	0.058	0.165	0.000	0.000
Natural gas	0.041	0.030	0.005	0.000	0.000
Coal	0.024	0.014	0.000	0.000	0.000

Note: Sample period, annual observations, 1870-2014 for oil and coal and 1919-2014 for natural gas.

Table 2. Contemporaneous correlations between prices

	A. Logged levels			B. First differences of log levels		
	Oil	Natural gas	Coal	Oil	Natural gas	Coal
Oil	1	0.978	0.986	1	0.507	0.619
Natural gas	0.978	1	0.990	0.507	1	0.321
Coal	0.986	0.990	1	0.619	0.321	1
	$\chi^2(3) = 724.811$			$\chi^2(3) = 74.145$		

Note: Sample period, annual data, 1919-2014.

Figure 4. Logs of Oil, Natural Gas, and Coal Prices

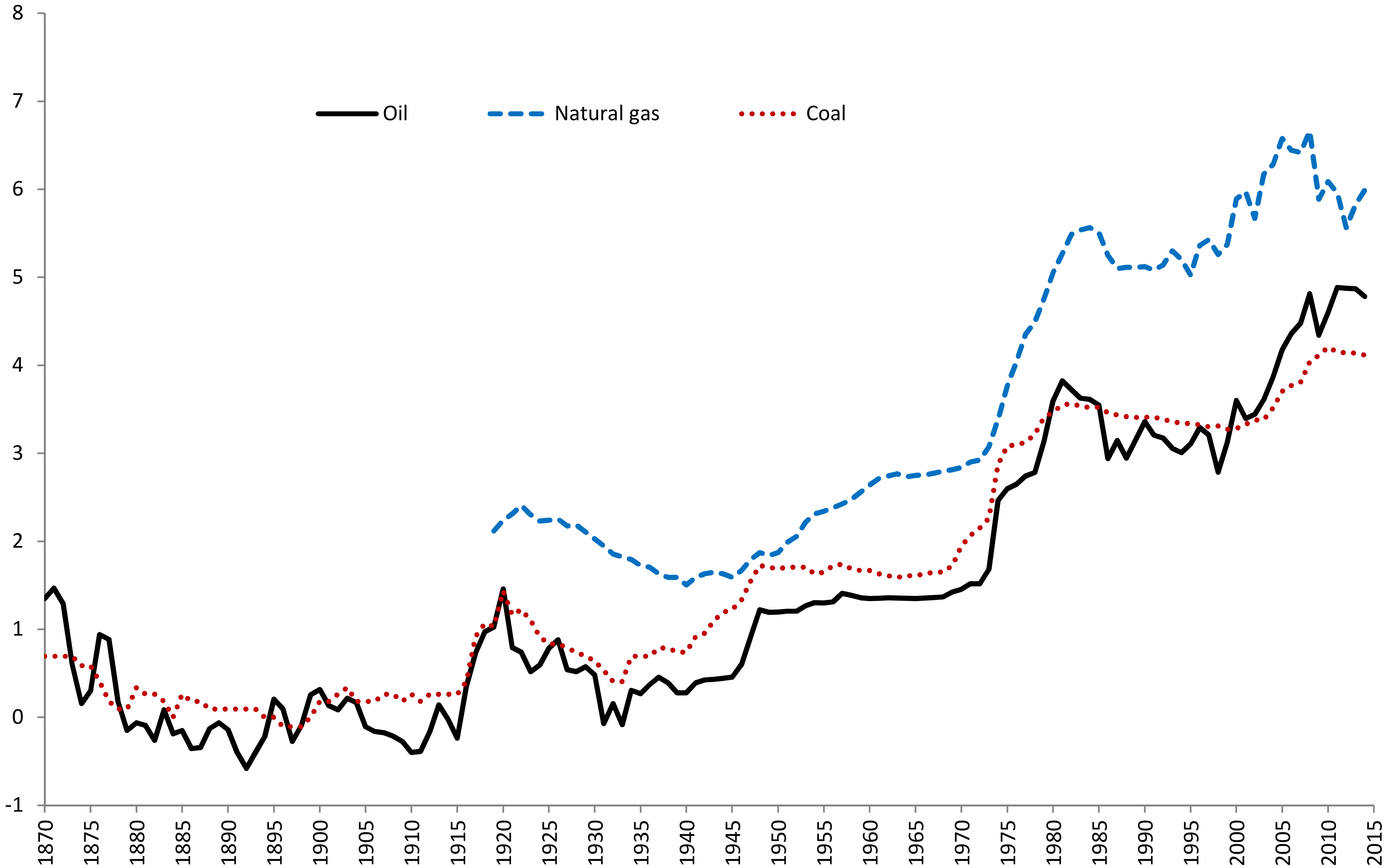

Table 3. Unit Root and Stationary Tests

Series	Test			Decision
	ADF	DF-GLS	KPSS	
		A. Logged levels		
Oil	-2.303	-0.866	0.530	$I(1)$
Natural gas	-3.154	-2.054	0.304	$I(1)$
Coal	-3.196	-1.539	0.458	$I(1)$
		B. Logged first differences		
Oil	-8.723	-8.846	0.049	$I(0)$
Natural gas	-3.377	-3.138	0.116	$I(0)$
Coal	-9.690	-9.710	0.061	$I(0)$

Notes: Sample period, annual observations, 1870-2014 for oil and coal and 1919-2014 for natural gas.
The 1% and 5% critical values are -4.023 and -3.441 for the ADF test, -3.529 and -2.988 for the DF-GLS test, and 0.216 and 0.146 for the KPSS test, respectively.

Table 4. Univariate GARCH oil models

Coefficient	GARCH model: Baseline	GARCH model: Extended
A. Conditional mean equation		
Constant	0.018 (0.364)	0.006 (0.591)
h_t		0.418 (0.062)
B. Conditional variance equation		
Constant	0.010 (0.217)	0.000 (0.280)
ε_{t-1}^2	0.293 (0.262)	1.143 (0.000)
h_{t-1}	0.622 (0.005)	0.612 (0.000)
$\varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1})$	−0.152 (0.463)	−0.780 (0.011)
C. Standardized residual diagnostics		
Mean	0.023	−0.021
Standard error	1.012	1.009
Variance	1.024	1.019
Skeweness	(0.694)	(0.571)
Kurtosis	(0.000)	(0.000)
Jarque-Bera	(0.000)	(0.000)
$Q(10)$	(0.230)	(0.075)
$Q^2(10)$	(0.983)	(0.801)
Log likelihood	5.047	8.583
D. Structural break tests		
Andrews-Quandt	1972 (0.388)	1892 (0.139)
Andrews-Ploberger	1972 (0.572)	1892 (0.126)

Notes: Sample period, annual data, 1870-2014.
Numbers in parentheses are p-values.

Table 5. Univariate GARCH natural gas models

Coefficient	GARCH model Baseline	Extended
A. Conditional mean equation		
Constant	0.034 (0.026)	0.021 (0.238)
ε_{t-1}	0.446 (0.001)	0.433 (0.003)
ε_{t-2}	0.279 (0.013)	0.308 (0.009)
ε_{t-3}	0.319 (0.002)	0.327 (0.000)
h_t		1.698 (0.094)
B. Conditional variance equation		
Constant	0.000 (0.157)	0.000 (0.157)
ε_{t-1}^2	0.525 (0.013)	0.552 (0.013)
h_{t-1}	0.666 (0.000)	0.651 (0.000)
$\varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1})$	−0.212 (0.499)	−0.206 (0.544)
C. Standardized residual diagnostics		
Mean	−0.005	−0.067
Standard error	0.999	0.999
Variance	0.998	0.998
Skeweness	(0.331)	(0.558)
Kurtosis	(0.166)	(0.097)
Jarque-Bera	(0.219)	(0.190)
$Q(10)$	(0.109)	(0.092)
$Q^2(10)$	(0.869)	(0.897)
Log likelihood	75.139	76.439
D. Structural break tests		
Andrews-Quandt	1985 (0.160)	1985 (0.158)
Andrews-Ploberger	1985 (0.152)	1985 (0.107)

Notes: Sample period, annual data, 1919-2014.
Numbers in parentheses are p-values.

Table 6. Univariate GARCH coal models

Coefficient	GARCH model Baseline	Extended
	A. Conditional mean equation	
Constant	−0.009 (0.043)	−0.010 (0.009)
$\Delta \ln c_{t-1}$	0.209 (0.128)	0.143 (0.282)
h_t		0.328 (0.349)
	B. Conditional variance equation	
Constant	0.000 (0.571)	0.000 (0.583)
ε_{t-1}^2	1.075 (0.000)	1.057 (0.000)
h_{t-1}	0.530 (0.000)	0.546 (0.000)
$\varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1})$	−0.280 (0.538)	−0.317 (0.345)
	C. Standardized residual ($\widehat{\varepsilon}$) diagnostics	
Mean	0.249	0.219
Standard error	0.971	0.978
Variance	0.942	0.957
Skeweness	(0.000)	(0.000)
Kurtosis	(0.000)	(0.001)
Jarque-Bera	(0.000)	(0.000)
$Q(10)$	(0.448)	(0.399)
$Q^2(10)$	(0.541)	(0.569)
Log likelihood	120.747	121.215
	D. Structural break tests	
Andrews-Quandt	1969 (0.820)	1978 (0.674)
Andrews-Ploberger	1969 (0.658)	1978 (0.436)

Notes: Sample period, annual data, 1870-2014.
Numbers in parentheses are p-values.

Table 7. The trivariate VARMA, GARCH-in-Mean BEKK model for natural gas, coal, and oil (in that order)

A. Conditional mean equation

$$\mathbf{\Phi} = \begin{bmatrix} -7.351\ (0.000) \\ -4.949\ (0.001) \\ 2.086\ (0.041) \end{bmatrix}; \mathbf{\Gamma} = \begin{bmatrix} 1.097\ (0.000) & -0.664\ (0.000) & -0.205\ (0.000) \\ 0.282\ (0.000) & -0.554\ (0.000) & 0.060\ (0.105) \\ 0.817\ (0.000) & -0.803\ (0.000) & -0.609\ (0.000) \end{bmatrix};$$

$$\mathbf{\Psi} = \begin{bmatrix} -0.239\ (0.000) & 1.263\ (0.000) & 0.066\ (0.011) \\ -0.196\ (0.001) & 1.170\ (0.000) & -0.136\ (0.083) \\ -0.830\ (0.000) & -0.275\ (0.087) & 1.014\ (0.000) \end{bmatrix}; \mathbf{\Theta} = \begin{bmatrix} -0.848\ (0.000) & 0.707\ (0.000) & 0.407\ (0.000) \\ -0.068\ (0.391) & 0.786\ (0.000) & -0.098\ (0.113) \\ -0.757\ (0.000) & 0.209\ (0.004) & 1.231\ (0.000) \end{bmatrix}.$$

B. Residual diagnostics

	Mean	Variance	$Q(5)$	$Q^2(5)$
z_{g_t}	−0.160	1.059	(0.349)	(0.310)
z_{c_t}	0.065	0.941	(0.000)	(0.228)
z_{o_t}	−0.103	0.983	(0.183)	(0.828)

C. Conditional variance equation

$$\boldsymbol{C} = \begin{bmatrix} -0.066\ (0.850) & & \\ 4.471\ (0.000) & 5.493\ (0.000) & \\ 1.287\ (0.001) & 1.617\ (0.000) & 0.000\ (0.999) \end{bmatrix}; \ \boldsymbol{A} = \begin{bmatrix} 0.936\ (0.000) & 0.519\ (0.000) & 1.127\ (0.000) \\ -0.589\ (0.000) & -0.023\ (0.627) & -1.081\ (0.000) \\ 0.046\ (0.183) & 0.162\ (0.000) & 1.085\ (0.000) \end{bmatrix};$$

$$\boldsymbol{B} = \begin{bmatrix} 0.362\ (0.000) & 0.000\ (0.999) & -0.076\ (0.030) \\ 0.029\ (0.563) & -0.139\ (0.297) & -0.106\ (0.048) \\ -0.040\ (0.147) & 0.050\ (0.337) & 0.249\ (0.000) \end{bmatrix}; \boldsymbol{D} = \begin{bmatrix} 1.430\ (0.000) & 0.272\ (0.000) & 0.877\ (0.000) \\ -0.702\ (0.000) & 0.059\ (0.550) & -0.106\ (0.424) \\ 0.008\ (0.744) & -0.247\ (0.000) & -0.633\ (0.000) \end{bmatrix}.$$

Note: Sample period, annual data: 1919-2014. Numbers in parentheses are tail areas of tests.