

Discussion Paper Series N 2019-09

# Forecasting Energy Commodity Prices: A Large Global Dataset Sparse Approach

Davide Ferrari

Free University of Bozen-Bolzano, Italy

Francesco Ravazzolo

Free University of Bozen-Bolzano, Italy

Joaquin Vespignani

University of Tasmania, Australia

ISBN 978-1-922352-12-5

# Forecasting Energy Commodity Prices: A Large Global Dataset Sparse Approach\*

Davide Ferrari<sup>a,f</sup> Francesco Ravazzolo<sup>a,c,d</sup> Joaquin Vespignani<sup>b,d,e</sup>

<sup>a</sup> Free University of Bozen-Bolzano, Faculty of Economics and Management, Italy

<sup>b</sup> University of Tasmania, Tasmanian School of Business and Economics, Australia

<sup>c</sup> Centre of Applied Macroeconomics and Commodity Prices, BI Norwegian Business School, Norway

<sup>d</sup> Centre for Applied Macroeconomics Analysis, ANU, Australia

<sup>e</sup> Globalization and Monetary Policy Institute, Federal Reserve Bank of Dallas, USA

<sup>f</sup> School of Mathematics and Statistics, University of Melbourne, Australia

December 17, 2019

## Abstract

This paper focuses on forecasting quarterly energy prices of commodities, such as oil, gas and coal, using the Global VAR dataset proposed by Mohaddes and Raissi (2018). This dataset includes a number of potentially informative quarterly macroeconomic variables for the 33 largest economies, overall accounting for more than 80% of the global GDP. To deal with the information on this large database, we apply a dynamic factor model based on a penalized maximum likelihood approach that allows to shrink parameters to zero and to estimate sparse factor loadings. The estimated latent factors show considerable sparsity and heterogeneity in the selected loadings across variables. When the model is extended to predict energy commodity prices up to four periods ahead, results indicate larger predictability relative to the benchmark random walk model for 1-quarter ahead for all energy commodities. In our application, the largest improvement in terms of prediction accuracy is observed when predicting gas prices from 1 to 4 quarters ahead.

**Keywords:** Energy Prices, Forecasting, Dynamic Factor model, Sparse Estimation, Penalized Maximum Likelihood

**JEL Codes:** C1, C5, C8, E3, Q4

---

\*We thank our discussant Shaun Vahey and conference and seminar participants at the CAMA-CAMP-RBA “International Economic Flows: Energy, Finance, Diplomacy and Market Structures” workshop for very useful comments. This paper is part of the research activities at the Centre for Applied Macroeconomics and commodity Prices (CAMP) at BI Norwegian Business School. Corresponding author: Francesco Ravazzolo, Faculty of Economics and Management, Free University of Bozen-Bolzano, Piazza Università 1, 39040 Bolzano, Italy. *Mail:* francesco.ravazzolo@unibz.it.

# 1 Introduction

Energy commodity prices play an important role in individual country-level economies. They are known to affect a number of macroeconomic variables including inflation, gross domestic product, exchange rate, stock market prices and interest rate as well as other commodity prices such as raw materials, metals, minerals and agricultural products whose cost depends on extraction and transportation. Energy commodities (such as oil, gas and coal) are important inputs to many industries, such as transport, agriculture, metal and mineral, construction, manufacturing, and chemical, amongst others. Consequently, a reliable forecast of energy prices is critical for Central Banks, Treasuries, Investors and International Organisations such as the International Monetary Fund (IMF) and the World Bank (WB) to implement policy responses and advice about energy commodity price fluctuations.

Figure 1 shows a strong co-movement of oil, gas and coal, across three distinctive economic periods: the great moderation from the middle 1980s to early 2002, the global financial crisis (GFC) (2008-2009), and the post-GFC period. The Great Moderation period is characterised by stable inflation and economic growth for the global economy. This period is also considered a period of smother business cycle in contrast to previous periods dominated by sharp boom and bust cycles. The second period from 2002 to 2008 is characterised by a rapid increase in commodity prices in line with the fast expansion of large developing economies such as China, India and other highly populated developing economies. These economies have rapidly increased energy consumption. For example, according to the World Bank (WB) China's consumption of energy has increased by 50% in this short period. From January 2002 to July 2008 oil, gas and coal prices increase by a factor of 4.9, 5.2 and 4.4 respectively. In 2008, the GFC took place and commodity prices sharply declined in line with the collapse of the world economy. From August 2008 to February 2009 oil, gas and coal declined by 61.6, 62.9 and 58.3 per cent respectively. From March 2009 to January 2010, oil, gas and coal sharply rebounded, rising by a factor of 1.8, 1.5 and 1.4 respectively. After the initial recovery post-GFC, energy commodity prices experienced a period of higher volatility to the present.

The main focus of this paper is to forecast energy commodity prices (oil, gas and coal) based on a large global database. The literature has applied, often separately, several macroeconomic and financial variables to predict energy and commodity prices. Different measures of economic activity have been largely used to predict energy commodity prices, such as Caldara et al. (2019), Aastveit et al. (2015) and Bjørnland et al. (2018), amongst others. Alquist et al. (2019) develop a factor-based identification strategy to link fluctuations in commodity prices and global economic activity.

The relationship between energy commodity prices and inflation has been intensively studied by academics and central bank researchers, see for example Charles and Fisher (2011). Energy commodity prices and inflation are often understood

in terms of cause-and-effect relationship. For example, Hammoudeh and Reboredo (2018) observed that oil has strong forward and backward linkages with economy and energy price components for the US inflation.

Several authors relate monetary variables such as interest rate and monetary aggregates as important predictors of commodity prices (including energy prices). Cologni and Manera (2008) finds that oil prices and inflation rate shocks are transmitted to the real economy and therefore central banks are forced to increase interest rates, while the findings in Anzuini et al. (2013) suggest that US monetary policy shocks drive both energy and non-energy commodity prices. Ratti and Vespignani (2015) find that monetary aggregates and interest rates in large developed and developing economies are important determinants of energy and non-energy commodity prices. Amongst others, Ravazzolo et al. (2016), Foroni et al. (2015), Chen and Rogoff (2003) show the close connection between exchange rates and commodity prices for countries dependent on commodity prices or commodity currencies.

Finally, the energy and oil prices literature has extensively studied the relationship between oil/commodity prices and stock market returns. For example, Kilian and Park (2009) and Kang et al. (2016) study the relationship between oil prices and the aggregated and industry-disaggregated stock market in the U.S. Abhyankar et al. (2013), related oil price increases with a decline in the Japanese stock market. Diaz et al. (2016) study the impact of oil prices volatility on the G7 economies. They find that the stock market for all G7 economies negatively responded to oil prices volatility from 1970 to 2014. Lombardi and Ravazzolo (2016) extend the analysis to asset allocation.

Building on this evidence we collect and update up to 2018Q4 the global VAR data set of Mohaddes and Raissi (2018). This large dataset contains around 200 series for the 33 largest world economies. The data frequency is quarterly, it spans the sample period from 1979Q1 to 2018Q4, and it includes measures of economic activity (GDP and industrial production), inflation, short- and long-term interest rate, nominal exchange rate and stock markets. Then, we apply a dynamic factor model to it and extract information from all the variables jointly. The large dimensionality of our database relative to the time-series length poses nontrivial challenges to standard dimension reduction techniques such as the factor analysis model. Particularly, when the sample size is small compared to the number of variables, the estimates from the standard factor model are very imprecise. In the most extreme case where the sample size is smaller than the number of variables, the maximum likelihood estimates are meaningless without additional assumptions on the model structure. Motivated by these issues, we assume sparse loadings meaning that more elements of the factor loading matrix become zero and apply a penalized maximum likelihood approach that allows to shrink parameters to zero and to estimate sparse factor loadings. Besides statistical accuracy, another issue concerns the interpretation of the factors in the estimated model. In standard fac-

tor analysis, all the estimated loadings are different from zero, meaning that all the observed variables from all countries in principle contribute to the latent factors. In reality, this may not be the case since certain observed variables may play a role in explaining the latent factors whilst others do not. Finally, after computing the factors, we use a bridge equation based on autoregressive or vector autoregressive models to predict from 1 quarter to 4 quarters ahead energy commodity prices.

The approach to factor construction proposed in this paper is global in nature: it is not constrained within a dominant geographical region and instead relies on exploiting the information from a large number of observable variables representing the entire global economy. Moreover, different factors can be associated with specific variables, such as inflation, equity prices, exchange rates and interest rates, supporting the construction of global factors of macroeconomic variables. Our empirical findings support the presence of pronounced sparsity and heterogeneity in the estimates for the loading matrix, thus showing that each latent factor is likely to depend on a subset of the original variables in the Global VAR data set. Compared to the benchmark random walk (RW) model, our model based on sparse factors improves the accuracy of 1-quarter ahead predictions for all the energy commodity prices. The largest accuracy gains occur when predicting gas prices with the lowest mean square prediction error values for 1- to 4-quarters ahead.

The paper is organized as follows. Section 2 describes the updated VAR global dataset. Section 3 outlines the proposed model and the sparse factor analysis methodology used for estimation. Section 4 summarizes our empirical findings. Section 5 concludes and provides final remarks.

## 2 Data

The data used in this paper is from the global VAR dataset from Mohaddes and Raissi (2018) updated to include 2018Q4 (the most recent available data). We also use quarterly energy commodity prices data from World Bank (WB) for the natural logarithm of the nominal price of oil, gas and coal.<sup>1</sup> The Global VAR dataset includes the most relevant quarterly macroeconomic variables for the 33 largest economies, accounting overall for more than 80% of the global economy as of 2018. The energy commodity prices raw data are shown in Figure 1. The series followed similar trends and experienced strong growth from beginning of 2000 to the US financial crisis when they dramatically dropped; resurged from 2010 up to another drop in 2014; and a final resurgence in 2017-2018. There is, however, large heterogeneity across them with gas for example being more volatile from 2000 to 2009. Therefore, our out-of-sample period from 2000 to 2018 corresponds to very volatile prices and overperforming a simple random walk benchmark might be not

---

<sup>1</sup>We use average prices, see <https://www.worldbank.org/en/research/commodity-markets> for data and definitions.

straightforward.

The country-level variables included in our model are: the natural logarithm of real GDP (RGDP), the rate of inflation calculated by taking the difference of the natural logarithm of the consumer price index (CPI), the natural logarithm of the nominal equity price index deflated by CPI (nomEQ), the natural logarithm of the exchange rate of country  $i$  at time  $t$  expressed in US dollars deflated by country  $i$ 's CPI (Fxdol), the nominal short-term percent interest rate per quarter (Rshort) and the nominal long-term percent interest rate per quarter (Rlong). The main data sources are Haver Analytics, the International Monetary Funds International Financial Statistics (IFS) database, and Bloomberg. When data were not available from these sources, we employed the other national datasets described in Mohaddes and Raissi (2018). Finally, as in the original database from Mohaddes and Raissi (2018), we use final vintage data.

### 3 Model

Let  $x_t = (x_{1,t}, \dots, x_{n,t})'$  be a vector of monthly variables observed at times  $t = 1, \dots, T$  included in our global dataset which have been standardized to have a mean equal to zero and variance equal to one. A dynamic factor model is then given by the following observation equation:

$$x_t = \Lambda f_t + \epsilon_t, \quad \epsilon_t \sim N(0, \Psi), \quad (1)$$

where  $\Lambda$  is a  $(n \times r)$  matrix of factor loadings,  $f_t = (f_{1t}, \dots, f_{rt})'$  is a  $(r \times 1)$  vector of static common factors and  $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{nt})'$  is the idiosyncratic component with zero expectation and  $R$  as covariance matrix. The dynamics of the common factors follows a VAR process:

$$f_t = A f_{t-1} + u_t \quad (2)$$

where  $u_t \sim N(0, Q)$ , and  $Q$  is a  $(r \times r)$  full-rank matrix,  $A$  is a  $(r \times r)$  matrix where all roots of  $\det(I_r - Az)$  lie outside the unit circle. In our empirical analysis, we set the autoregressive lag structure of the VAR model in equal to 1. We investigated the usefulness of higher order lags, but results were qualitatively similar. We also note that factors are demeaned to have zero mean.

The idiosyncratic and VAR residuals are assumed to be independent and identically distributed

$$\begin{bmatrix} \epsilon_t \\ u_t \end{bmatrix} \sim iid N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R & 0 \\ 0 & Q \end{bmatrix} \right) \quad (3)$$

The  $h$ -step ahead predictions of each of the  $S$  commodity prices,  $y_{s,t+h}$ ,  $s = 1, \dots, S$ , are obtained by using a bridge equation where commodity prices ( $y_{s,t}$ ) are expressed

as a linear function of the expected common factors and lags of  $y_{t+h}$  follows:

$$y_{s,t+h} = \alpha_s + \rho_s y_{s,t+h-1} + z_s' f_{t+h} + e_{s,t+h}, \quad e_{s,t+h} \sim N(0, \sigma_{s,e}) \quad (4)$$

where  $z_s$  is an  $r \times 1$  vector of parameters.<sup>2</sup> Accordingly, predictions of commodity prices ( $y_{s,t+h}$ ) are constructed from equation (4), conditional on the estimated parameters and the factor forecasts. The model is labelled as DFM-AR. Note, that equation (4) implies that  $h$ -step ahead forecasts are computed as iterative forecasts. An alternative approach, not considered in this paper, is the direct multi-step ahead forecasting suggested by Marcellino et al. (2006).

We also investigate the use a VAR model to predict jointly the full vector of commodity prices  $y_{t+h} = (y_{1,t+h}, \dots, y_{S,t+h})$  by the multivariate model

$$y_{t+h} = C + \Phi y_{t+h-1} + B' F_{t+h} + e_{t+h}, \quad e_{t+h} \sim N(0, \Sigma_e). \quad (5)$$

The model is labelled as DFM-VAR.

Equations (1) and (2) are commonly estimated following a two-step procedure; e.g., see Giannone et al. (2008). In the first step, principal component analysis or factor analysis are used to estimate the latent factors, while the second step entails finding maximum likelihood estimates for the latent VAR proces. One issue with this approach is that estimation becomes very imprecise when the number of variables  $n$  is relatively large compared to the number of observations. In the most extreme case where the number of observations  $T$  is much smaller than  $n$ , maximum likelihood estimates for the standard factor analysis model do not exist. An additional problem with the traditional approaches is that the estimated matrix of loadings  $\Lambda$  is not sparse, meaning that all its elements are different from zero even when in reality certain elements of  $\Lambda$  might be exactly zero, i.e. some variables may play no role in forming certain latent factors. This means that the additional statistical errors deriving from the estimation of the irrelevant loadings would result in larger forecasting errors.

Motivated by these issues, we propose to use a penalized likelihood approach to extract the latent factors. Penalized likelihood estimation, such as Lasso-type penalization, has been successfully applied in a number of domains in statistics and econometrics to deal with complex likelihood functions with a large number of parameters. In this paper, we follow the penalized factor analysis estimation methodology first considered by Choi et al. (2010). Hirose and Yamamoto (2015) developed algorithms to compute the entire solution path, permitting the application of a wide variety of convex and nonconvex penalties.

---

<sup>2</sup>As for the VAR model in (2) we also investigate to include more lags of  $y_{s,t}$  in (4), but forecast accuracy did not improve.

### 3.1 Penalized maximum likelihood estimation

Assuming fixed factors, the observable random vector  $x_t$  follows a multivariate normal distribution with variance-covariance matrix  $\Sigma = \Lambda\Lambda' + \Psi$ . The log-likelihood function is, up to an additive constant not depending on parameters

$$\ell(\Lambda, \Psi) = -\frac{T}{2} \left[ \log |\Lambda\Lambda' + \Psi| + \text{tr}\{(\Lambda\Lambda' + \Psi)^{-1}\widehat{\Sigma}\} \right], \quad (6)$$

where  $\widehat{\Sigma}$  is the sample variance-covariance matrix. The maximum likelihood estimates of  $\Lambda$  and  $\Psi$  are found by solving  $\partial\ell(\Lambda, \Psi)/\partial\Lambda = 0$  and  $\partial\ell(\Lambda, \Psi)/\partial\Psi = 0$ . Following Hirose and Yamamoto (2015), estimates for the loading,  $\widehat{\Lambda}_{\rho, \gamma}$ , are obtained by maximizing the penalized log-likelihood function

$$\ell_{\rho, \gamma}(\Lambda, \Psi) = \ell(\Lambda, \Psi) - T \sum_{i=1}^n \sum_{j=1}^r \text{pen}_{\rho, \gamma}(|\Lambda_{ij}|), \quad (7)$$

where  $\rho > 0$  and  $\gamma > 0$  are regularization parameter. Here  $\text{pen}(\cdot)$  is a penalty with amount of shrinkage controlled by  $\rho$  and  $\gamma$ . Particularly, for smaller values of  $\rho$ , we have greater shrinkage, meaning that more elements of the loading matrix  $\Lambda$  become zero. In this paper, we focus on the following popular nonconvex penalties:

- The smoothly clipped absolute deviation (SCAD) penalty defined by

$$\text{pen}_{\rho, \gamma}(z) = \begin{cases} \rho z, & z \leq \rho \\ \frac{\gamma \rho z - (z^2 + \rho^2)^2}{2(\gamma - 1)}, & \rho \leq z \leq \gamma \rho \\ \frac{\gamma^2(\gamma^2 - 1)}{2(\gamma - 1)}, & z > \gamma \rho. \end{cases}$$

- The minimax convex penalty (MC+) defined by

$$\text{pen}_{\rho, \gamma}(z) = \begin{cases} \rho z - \frac{z^2}{2\gamma}, & z \leq \rho\gamma, \\ \frac{1}{2}\gamma\rho^2, & z > \rho\gamma. \end{cases}$$

For each value of  $\rho > 0$ ,  $\gamma \rightarrow \infty$  yields the soft threshold operator (i.e., the Lasso penalty), whilst  $\gamma \rightarrow 1^+$  produces hard threshold operator.

The maximization of the penalized likelihood is computed efficiently through a generalized Expectation-Maximization (EM) algorithm. Given number of latent factors  $r \geq 1$ , and a grid of choices for the tuning parameters  $\gamma$  and  $\rho$ , estimated loading matrices are computed in the maximization step through the coordinate-descent approach described in Hirose and Yamamoto (2015). Estimates in this paper are obtained using the R package `fanc` (Hirose et al., 2016).

For our empirical analyses, it is important to select the appropriate value of the regularization parameters  $\rho$  and  $\gamma$ , as well as the unknown number of latent factors

$r$ . The selection of the triple  $(\rho, \gamma, r)$  can be viewed as a model selection. To this end, we use the Bayesian information criterion (BIC)

$$BIC(\rho, \lambda, r) = -2\ell_{\rho, \gamma}(\widehat{\Lambda}_{\rho, \gamma}, \widehat{\Psi}_{\rho, \gamma}) + \log(T) \times \text{df}(r, n), \quad (8)$$

where  $\text{df}(r, n) = rn + r(r + 1)/2$  is an estimate of the degrees of freedom, i.e. the number of effective parameters. Note that our estimate for the degrees of freedom is naive in the sense that we are ignoring the constraints on certain loadings to be zero; thus, we have no guarantee that  $\text{df}(r, n)$  results in model selection consistency. On the other hand, estimation of degrees of freedom in high-dimensional factor analysis is an ongoing research topic and more theory would be needed in this context. In other works involving penalized likelihood approaches, the number of effective parameters is often approximated by taking the number of nonzero elements in  $\widehat{\Lambda}_{\rho, \gamma}$ . By doing this, however, we would under-estimate the effective degrees of freedom of our model since we would treat the selected model as the true model, thus ignoring completely the uncertainty related to the model selection process. Following the Generalized information criterion approach of Konishi and Kitagawa (1996), we also estimated directly the model-selection bias by non-parametric bootstrap for the log-likelihood  $\ell_{\rho, \gamma}(\widehat{\Lambda}_{\rho, \gamma}, \widehat{\Psi}_{\rho, \gamma})$ , finding models with complexity similar to those selected by our naive BIC approach.

## 4 Empirical analysis

This section provides the results of the application of the proposed dynamic factor model to predict energy commodity prices. In the first subsection we focus on in-sample results, in particular evaluating factor loadings, sparsity and factor dynamics. The second subsection provides details of the forecast exercises and results.

### 4.1 In-sample results

Our in-sample analysis focuses on the estimation via penalized maximum likelihood of factors  $f_t$  and factor loadings  $\Lambda$  in equation (1) when applied to the updated global VAR dataset.

The first step is the estimation of the number of factors. We apply the BIC selection in equation (3.1) resulting in  $r = 8$  factors. In particular, the BIC value steadily declines up to 8 factors, then it stabilizes before increasing again when  $r > 10$ . The selected number of factors partly confirms previous findings. Based on US data, Ludvigson and Ng (2009), Stock and Watson (2012) and Casarin et al. (2015) find optimal number of factors equal to 7, 5 and 7, respectively.

The selected tuning parameters are  $\rho = 0.14$  and  $\gamma = 1.01$  essentially corresponding to the hard threshold operator, and indicate a high level of sparsity in

the factor loadings. Indeed, we recall that  $\Lambda$  is a  $(174 \times 8)$  matrix and therefore it requires the estimation of 1392 parameters. Thus, estimation would be imprecise without imposing sparsity. Moreover, it could not be realistic imposing that all variables from all 33 countries contribute to all the factors. Tables 1 and 4 summarize the composition of the estimated factor loadings. In particular, Table 1 reports the proportion of times a certain variable is included in constructing factors. A variable  $i$  has a significant contribution to factor  $j$  if the corresponding factor loading estimate  $\hat{\Lambda}_{ij}$  is different from zero, where  $\hat{\Lambda}$  is the estimated loading matrix. A value of 1 means that all variables associated to a specific group are included; a value of 0 indicates not variables are included. Table 4 provides detailed results. Table 1 organizes the results in two groups: by variables and by geographical regions.

When grouping across all variables (column All), we find that from 30% to 40% of the estimated loadings were exactly zero. This is a substantial sparsity level, indicating not all variables are useful for all factors. On the other hand, we found a large heterogeneity across factors. All 8 factors load on most of the real GDP (RGDP) variables, with a minimum of 7% of the real variables restricted to zero for the eighth factor and a maximum of 21% zero restrictions for the seventh factor. CPI variables are important for the third and, above all, the sixth factors. The seventh factor is an equity price factor (nomEQ) with no sparsity, whether the second factor loads mainly on nominal exchange rate (Fxdol). Factor three loads with no sparsity on short interest rates (Rshort), but this variable is also relevant for the first, the fifth and the sixth factors, supporting the connections of interest rates to many variables, in particular CPI (sixth factors). Finally, all long interest rates (Rlong) variables are included to compute the eighth factor, and most of them in factors one, three, four, five and six.

As a second evidence we classify results in geographical regions. Following Bjørnland et al. (2017), we divide countries in four main continents: North America, South America, Europe and Asia.<sup>3</sup> The heterogeneity is very large, with only the third factor loading on all South American variables, and from 40% to 25% of factor estimates are in most cases restricted to zero. Therefore, our evidence indicates that factors are strongly connected to specific variables and not geographical regions, providing a supportive evidence of the importance of global database.

Figure 2 shows the estimated eight factors. We note specific patterns of the factors that can be explained by the factor loadings. For example, the sixth factor is the one that resembles a global inflation factor. It shows large values in the early part of the sample, then it decreases over the '80s, stabilizes in the '90s and beginning of 2000 before becoming negative starting from the US financial crises. The seventh factor captures equity collapses in October 1987, for the Gulf War (1990-1991), around the Asian financial crisis (1997-1998), associated to the burst

---

<sup>3</sup>Australia, New Zealand and South Africa are grouped in the Asia continent to avoid groups with few variables. We also note that not all variables are available for all countries, in particular equity premium and long interest rates are missing for some of the Asian and South American countries.

of the internet bubble (2000), and to the US financial crisis (2008). Factor three proxies the declines of the short interest rates over the sample. Finally, there is not a clear output factor because most factors upload on RGDP variables and output growth has been quite heterogeneous across countries in our sample. Finally, Table 2 provides the estimate of the matrix  $Q$  in equation (2).

## 4.2 Forecasting results

### 4.2.1 Alternative models

We compare the dynamic factor models DFM-AR in (1)-(2)-(4) and DFM-VAR in (1)-(2)-(5) to three standard models usually applied when investigating energy commodity price predictability, see Alquist et al. (2013).

The benchmark model is the random walk model:

$$y_{t+h,s} = y_{t,s} + e_{t+h,s}. \quad (9)$$

The model assumes that the most accurate prediction for future values of variables  $y_{t+h,s}$  is the last available price  $y_{t,s}$ .

The second model is the (parsimonious) univariate autoregressive (AR) specification, whereas the errors are assumed to be normally distributed with zero mean and  $\sigma_h^2$  variance. The autoregressive model can be written as follows

$$y_{t,s} = \sum_{l=1}^p \phi_l y_{t-l,s} + e_{t,s}. \quad (10)$$

To account for the long autoregressive lag properties of energy commodity prices, see Kilian (2009), we fix  $p=8$  (two years) and select with the AIC criterion the optimal  $p$  lag for each series and for each vintage.

The third model is the vector autoregressive (VAR) model. Let  $\mathbf{y}_{t+h} = (y_{t+h,1}, \dots, y_{t+h,3})'$  denote the  $(3 \times 1)$  vector of energy commodity prices. The VAR model of order  $p$  is formulated as follows:

$$\mathbf{y}_{t+h} = \sum_{l=1}^p \Phi_l \mathbf{y}_{t+h-l} + \mathbf{e}_t, \quad (11)$$

where  $\Phi_l$  is the  $(3 \times 3)$  matrix of autoregressive coefficients. The vector of errors  $\mathbf{e}_t$  is assumed to be serially uncorrelated and normally distributed with zero mean and a full covariance matrix  $\Sigma$ . As for the AR model, we choose optimal  $p$  via AIC for each vintage.

### 4.2.2 Forecast evaluation

We compare the different models in predicting the three energy commodity (logarithmic) prices for four horizons,  $h = 1, 2, 3$  and 4 quarters ahead. We assess the

goodness of our forecasts using the mean square forecasting errors (MSFEs) for each forecast horizon  $h$ . The MSFE for  $h = 1, \dots, 4$ , model  $i = \text{RW, AR, VAR, DFM-AR, DFM-VAR}$ ,  $s = \text{coal, gas and oil}$  is computed as

$$\text{MSFE}_{h,i,s} = \frac{1}{T-R} \sum_{t=R}^{T-1} (\hat{y}_{t+h,i,s} - y_{t+h,s})^2, \quad (12)$$

where  $T$  is the number of observations,  $R$  is the initial in-sample period to compute the first forecasts,  $y_{t+h,s}$  is the realized value of the variable  $s$  at time  $t+h$ , and  $\hat{y}_{t+h,i,s}$  is the model  $i$  forecast of  $y_{t+h,s}$  made at time  $t$ .

In addition, to provide a rough gauge of whether the differences in forecast accuracy based on square forecasting errors are significant, we apply the Model Confidence Set procedure of Hansen et al. (2011) across models for a fixed variable  $s$  and a fixed horizon  $h$  to jointly compare their predictive power. We implemented the MCS procedure with  $T_{max, \mathcal{M}}$  test (Hansen et al., 2011, p. 465) at the  $\alpha = 0.15$  significance level, using the R function `MCSprocedure` within the package `MCS` in Bernardi and Catania (2018).

### 4.2.3 Results

We split the full quarterly sample 1979Q1-2018Q4 into two periods: an initial in-sample period 1979Q1-1999Q4 and an out-of-sample (OOS) period 2000Q1-2018Q4. We use an expanding recursive window to estimate the models and produce 76 forecasts (2000Q1-2018Q4) for each of the energy commodity prices by extending the in-sample period with a new observation at each step. For each of the 76 OOS values, we produce from 1- to 4-step ahead forecasts using several different models described in the previous section. Models are re-estimated for each one of the 76 vintages.

Table 3 reports the OOS forecasting results for the different individual models when predicting the three energy commodity prices at the four horizons we consider. In the second row, the MSFE for the benchmark RW model is reported; on the contrary ratios between the MSFEs of the AR, VAR, FDM-AR and DFM-VAR models and the MSFE of the RW model are reported in the other rows. Values less than 1 indicates that forecasts from the current model are more accurate than forecasts of the RW model.

We find three clear results. The first one is that the benchmark model RW is difficult to beat, as documented in many studies on oil price predictability, see e.g. Alquist et al. (2013) and Baumesteir and Kilian (2015). We generalize the finding to other energy commodity prices as coal and gas. Second, the model that offers larger predictability relative to the RW is the DFM-AR model. The DFM-AR provides the lowest MSFE at 1-quarter ahead horizon for all variables, and at all horizons when predicting gas prices. Moreover, the DFM-AR is the only model always included in the model confidence set in all twelve forecasting exercises we

consider. The DFM-VAR model also does very well when predicting gas prices and it gives the lowest MSFE in all four horizons with reductions up to 15%. But the DFM-VAR does not provide any economic gain when predicting coal and oil. In the latter case, it is not included in the model confidence set at any horizon. This results hints to a different behaviour of gas prices that respond both to gas-market fundamentals and to other commodity energy markets. Other energy commodity prices are mainly driven by their own fundamentals. Finally, the VAR model gives the largest errors. The model has a larger MSFE than that of the RW model in all cases and it is never included in the best set of models.

## 5 Conclusion

This paper proposes the use of a dynamic factor model based on a penalized maximum likelihood approach to forecast oil, gas and coal commodity prices. The nature of our approach is global, since our model is constructed using an update version of the Global VAR dataset proposed by Mohaddes and Raissi (2018), which includes important information from the largest 33 economies globally accounting for more than 80% of the global economy. Specifically, we consider measures of output, inflation, interest rates, equity markets and exchange rates of these countries as well as data for other non-energy commodities (metal and agricultural commodity prices). We estimate a factor model using a penalized likelihood approach that enables shrinking loading parameters to zero, thus resulting in sparse factor loadings. The sparse estimates are shown to improve forecasts compared to standard methods and enhance the interpretation of the role played by observable variables in forming latent factors. Interestingly, each factor mostly captures different economic information and we identify global inflation, interest rates and equity premium factors.

For forecasting, we introduce a bridge equation relating the latent factors to energy commodity prices. We carried out out-of-sample testing of the accuracy of our forecasting by splitting the original sample into an initial sample period (1979Q1-199Q4) and out-of-sample period (2000Q1-2018Q4). Using a moving forecasting window approach, we estimated the models and produced 76 forecasts (1 to 4 periods ahead) for each energy commodity prices (oil, gas and coal). Main results indicate that the model which provides the largest predictability gains relative to the RW is our dynamic factor model. Precisely, our model provides more accurate forecasts than the benchmark RW model for 1-quarter ahead for all energy commodities. When predicting gas prices, the model gives the lowest MSFE values for 1 to 4 quarters ahead.

## References

- K. A. Aastveit, H. C. Bjørnland, and L. A. Thorsrud. What drives oil prices? emerging versus developed economies. *Journal of Applied Econometrics*, 30: 1013–1028, 2015.
- A. Abhyankar, B. Xu, J. Wang, et al. Oil price shocks and the stock market: evidence from japan. *Energy Journal*, 34:199–222, 2013.
- R. Alquist, L. Kilian, and R. J. Vigfusson. Forecasting the price of oil. In G. Elliott and A. Timmermann, editors, *Handbook of Economic Forecasting, Vol. 3*. Elsevier, 2013.
- R. Alquist, S. Bhattarai, and O. Coibion. Commodity-price comovement and global economic activity. *Journal of Monetary Economics*, forthcoming, 2019.
- A. Anzuini, M. Lombardi, and P. Pagano. The impact of monetary policy shocks on commodity prices. *International Journal of Central Banking*, 9:125–150, 2013.
- C. Baumesteir and L. Kilian. Forecasting the real price of oil in a changing world: A forecast combination approach. *Journal of Business and Economic Statistics*, 33(3):338–351, 2015.
- M. Bernardi and L. Catania. The model confidence set package for r. *International Journal of Computational Economics and Econometrics*, 8(2):144–158, 2018.
- H. C. Bjørnland, V. H. Larsen, and J. Maih. Oil and macroeconomic (in) stability. *American Economic Journal: Macroeconomics*, 2018.
- H. Bjørnland, F. Ravazzolo, and L. Thorusrud. Forecasting gdp with global components. this time is different. *International Journal of Forecasting*, 33(1):153–173, 2017.
- D. Caldara, M. Cavallo, and M. Iacoviello. Oil price elasticities and oil price fluctuations. *Journal of Monetary Economics*, 103:1–20, 2019.
- R. Casarin, S. Grassi, F. Ravazzolo, and H. van Dijk. Dynamic predictive density combinations for large data sets in economics and finance. Technical Report 15-084/III, Tinbergen Institute Discussion Paper, 2015.
- E. Charles and J. Fisher. What are the implications of rising commodity prices for inflation and monetary policy? Technical Report 286, Chicago Fed Letter, 2011.
- Y. C. Chen and K. Rogoff. Commodity currencies. *Journal of international Economics*, 60:133–160, 2003.
- J. Choi, G. Oehlert, and H. Zou. A penalized maximum likelihood approach to sparse factor analysis. *Statistics and its Interface*, 3(4):429–436, 2010.

- A. Cologni and M. Manera. Oil prices, inflation and interest rates in a structural cointegrated var model for the g-7 countries. *Energy economics*, 30:856–888, 2008.
- E. M. Diaz, J. C. Molero, and F. P. de Gracia. Oil price volatility and stock returns in the g7 economies. *Energy Economics*, 54:417–430, 2016.
- C. Foroni, F. Ravazzolo, and P. J. Ribeiro. Forecasting commodity currencies: The role of fundamentals with short-lived predictive content. 2015.
- D. Giannone, L. Reichlin, and D. Small. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676, 2008.
- S. Hammoudeh and J. C. Reboredo. Oil price dynamics and market-based inflation expectations. *Energy Economics*, 75:484–491, 2018.
- P. R. Hansen, A. Lunde, and J. M. Nason. The Model Confidence Set. *Econometrica*, 79:453–497, 2011.
- K. Hirose and M. Yamamoto. Sparse estimation via nonconcave penalized likelihood in factor analysis model. *Statistics and Computing*, 25(5):863–875, 2015.
- K. Hirose, M. Yamamoto, and H. Nagata. *fanc: Penalized Likelihood Factor Analysis via Nonconvex Penalty*, 2016. URL <https://CRAN.R-project.org/package=fanc>. R package version 2.2.
- W. Kang, R. A. Ratti, and J. Vespignani. The impact of oil price shocks on the us stock market: A note on the roles of us and non-us oil production. *Economics Letters*, 145:176–181, 2016.
- L. Kilian. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–1069, 2009.
- L. Kilian and C. Park. The impact of oil price shocks on the us stock market. *International Economic Review*, 50:1267–1287, 2009.
- S. Konishi and G. Kitagawa. Generalised information criteria in model selection. *Biometrika*, 83(4):875–890, 1996.
- M. Lombardi and F. Ravazzolo. On the correlation between commodity and equity returns: Implication for portfolio allocation. *Journal of Commodity Markets*, 2(1):45–57, 2016.
- S. Ludvigson and S. Ng. Macro factors in bond risk premia. *Review of Financial Studies*, 22(12):5027–5067, 2009.

- M. Marcellino, J. H. Stock, and M. W. Watson. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135:499–526, 2006.
- K. Mohaddes and M. Raissi. Global var (gvar) database, 1979q2-2016q4. Technical report, University of Cambridge, 2018.
- R. A. Ratti and J. L. Vespignani. Commodity prices and bric and g3 liquidity: A sfavec approach. *Journal of Banking & Finance*, 53:18–33, 2015.
- F. Ravazzolo, T. Sveen, and S. K. Zahiri. Commodity futures and forecasting commodity currencies. 2016.
- J. H. Stock and W. M. Watson. Disentangling the channels of the 2007-09 recession. *Brookings Papers on Economic Activity*, pages 81–156, Spring, 2012.

Table 1: Factor loadings summary

Factors	1	2	3	4	5	6	7	8
Variable groups								
All	0.644	0.592	0.713	0.632	0.707	0.678	0.672	0.609
RGDP	0.909	0.818	0.848	0.909	0.879	0.848	0.788	0.939
CPI	0.606	0.364	0.727	0.424	0.606	0.848	0.424	0.485
nomEQ	0.192	0.269	0.192	0.423	0.462	0.385	1.000	0.154
Fxdol	0.406	0.875	0.563	0.531	0.594	0.250	0.656	0.531
Rshort	0.875	0.531	1.000	0.688	0.844	0.844	0.625	0.625
Rlong	0.889	0.667	0.944	0.889	0.889	0.944	0.556	1.000
Geographical regions								
North America	0.727	0.455	0.636	0.636	0.546	0.546	0.818	0.727
South America	0.810	0.524	1.000	0.524	0.714	0.381	0.524	0.619
Europe	0.662	0.662	0.676	0.662	0.732	0.775	0.747	0.648
Asia	0.563	0.563	0.676	0.634	0.704	0.690	0.620	0.549

*Notes:* Percentage of significant variables included in constructing factors. A variable (RDGP, CPI, etc.) has a significant contribution to a given factor  $j$ , with  $j = 1, \dots, 8$ , if  $\hat{\Lambda}_{ij} \neq 0$  where  $\hat{\Lambda}_{ij}$  is the  $(ij)$ th element of the estimated loading matrix  $\hat{\Lambda}$ . A value of 1 indicates all variables associated to a specific group are included; a value of 0 indicates not variables are included.

Table 2: Factor residual correlations

Factors	1	2	3	4	5	6	7	8
1	1.140	-0.051	-0.153	-0.052	0.330	-0.051	0.126	0.184
2	-0.051	1.312	-0.121	-0.159	0.164	0.288	-0.283	0.059
3	-0.153	-0.121	1.071	0.016	-0.037	0.128	0.099	0.207
4	-0.052	-0.159	0.016	1.330	-0.618	0.157	-0.175	-0.228
5	0.330	0.164	-0.037	-0.618	1.524	0.025	0.127	-0.181
6	-0.051	0.288	0.128	0.157	0.025	1.169	-0.019	0.277
7	0.126	-0.283	0.099	-0.175	0.127	-0.019	1.257	0.163
8	0.184	0.059	0.207	-0.228	-0.181	0.277	0.163	1.354

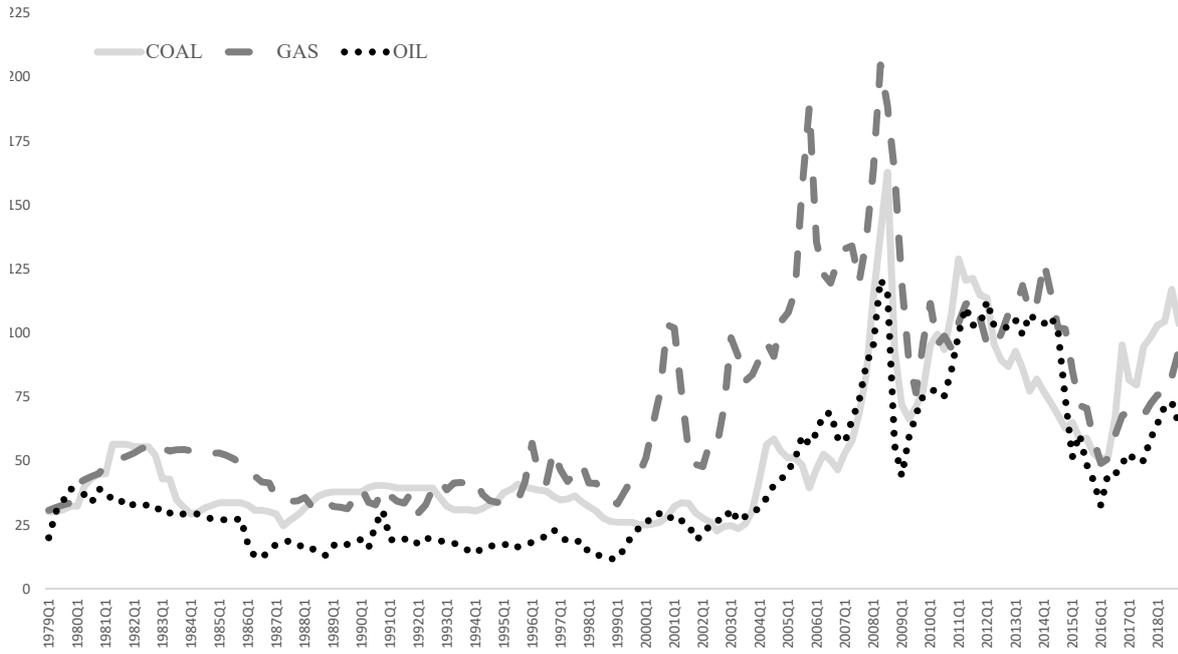
*Notes:* Estimates of the covariance matrix  $Q$  in equation (2).

Table 3: MSFE

	h=1	h=2	h=3	h=4
COAL				
RW	0.020*	0.056*	0.093*	0.127*
AR	1.208*	1.245*	1.299*	1.345*
VAR	1.200	1.441	1.608	1.676
DFM-AR	0.966*	1.036*	1.091*	1.118*
DFM-VAR	1.013*	1.104*	1.165*	1.185*
GAS				
RW	0.022*	0.061	0.097*	0.123*
AR	1.239*	1.270	1.261*	1.346*
VAR	3.022	2.789	2.750	3.106
DFM-AR	0.964*	0.925*	0.912*	0.921*
DFM-VAR	0.913*	0.842*	0.860*	0.952*
OIL				
RW	0.024*	0.057*	0.084*	0.106*
AR	1.355	1.473	1.557	1.665
VAR	1.322	1.781	2.024	2.039
DFM-AR	0.974*	1.077*	1.140*	1.150*
DFM-VAR	1.024	1.162	1.235	1.243

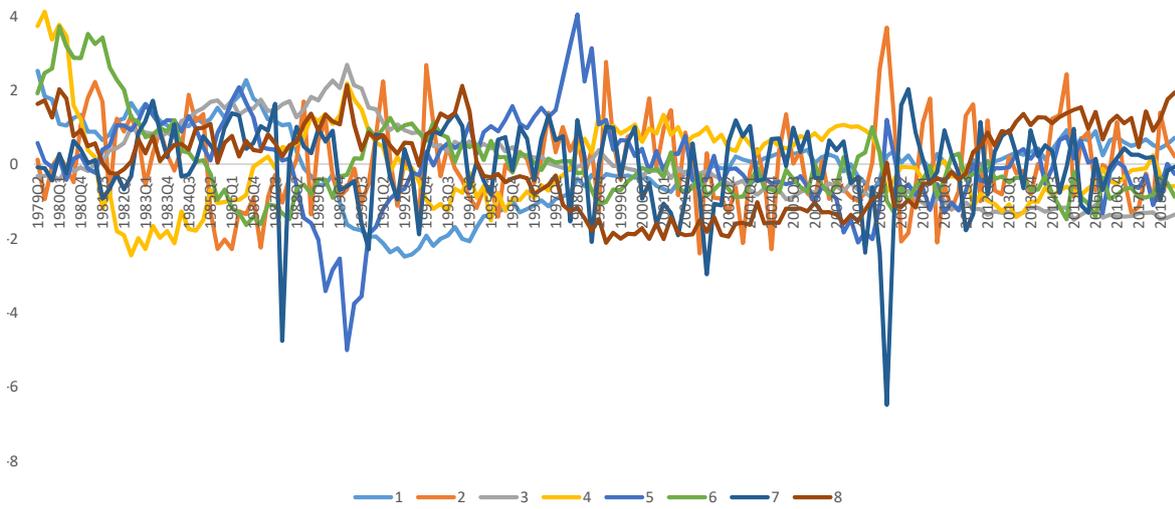
*Notes:* *Notes:* The Table reports mean square forecast error (MSFE) for different models and four different forecasting horizons, h=1, 2, 3 and 4 quarters ahead. For the RW baseline models, the Table reports MSFEs; for all other models, table reports ratio between MSFE of current model and MSFE of RW benchmark. Entries less than 1 indicate that forecasts from current model are more accurate than forecasts from corresponding baseline model. A \* indicates the models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure for each horizon h at confidence level 15%.

Figure 1: Energy Commodity Prices



Notes: The Figure shows coal, gas and oil prices.

Figure 2: Factors



Notes: The Figure shows the eight estimated factors.

# Appendix

This appendix reports details results for factor loadings, displaying results for each variable and each country.

Table 4: Factor loadings

Series	Countries	Factors							
		1	2	3	4	5	6	7	8
RGDP	'Argentina'	0.648	0.012	0.482		0.053	0.522	0.006	0.003
	'Australia'	0.734		-0.160	0.148	0.034	0.528	-0.026	0.000
	'Austria'	0.310			0.484	-0.079	0.599		-0.266
	'Belgium'	0.317	0.011	0.228	0.423	0.085	0.700		-0.151
	'Brazil'	0.553	0.013	0.639	-0.012	-0.029	0.414	-0.002	0.114
	'Canada'	0.711		-0.246	0.359	-0.027	0.345	-0.047	0.015
	'China'	0.035	-0.020	-0.704	-0.217	-0.332	-0.446	-0.058	0.194
	'Chile'	0.624	-0.002	0.293	0.027		0.625	0.001	0.083
	'Finland'	0.843	0.015			-0.489			-0.501
	'France'	0.467	-0.012	-0.086	0.395	0.052	0.675	-0.034	0.000
	'Germany'		0.018	0.387	0.617	0.039	0.329	-0.077	0.284
	'India'	0.605	-0.008	0.507	-0.006	-0.177	0.500	0.009	0.078
	'Indonesia'			0.711	0.088	0.252	0.382	0.051	0.185
	'Italy'	0.380		0.551	0.222	0.127	0.589	0.007	-0.001
	'Japan'	-0.375	-0.011	-0.655	0.001	-0.209	-0.495	-0.007	0.013
	'Korea'	-0.322	-0.030	-0.707	0.173	-0.101	-0.517	-0.011	0.172
	'Malaysia'	-0.308	-0.010	-0.816	-0.057	-0.109	-0.384	-0.014	0.055
	'Mexico'	0.503	-0.017	0.475	-0.011	0.079	0.610		0.059
	'Netherlands'	-0.281	0.044	-0.425	0.612			-0.058	-0.232
	'Norway'	0.617	-0.026		0.258	0.183	0.598		-0.138
	NewZealand'	0.733	-0.017		0.047	0.040	0.578		-0.003
	'Peru'	0.683	0.011	0.554	-0.026		0.396	-0.003	0.072
	Philippines'	0.543	0.002	0.370	0.004	0.007	0.592	-0.001	0.157
	SouthAfrica'	0.472	0.007	0.630	0.024	0.037	0.530	0.015	0.012
	SaudiArabia'	-0.075	-0.016	-0.880	-0.061	-0.274	-0.187	-0.038	0.055
	'Singapore'	-0.256		-0.766	-0.071	-0.172	-0.419	-0.014	-0.031
	'Spain'	0.516	0.002	0.375	0.219	0.018	0.647	-0.009	-0.013
	'Sweden'	0.820	0.001		0.116	-0.168	0.414	-0.047	0.126
	Switzerland'	0.529	0.008	-0.427		-0.387		-0.097	0.453
	'Thailand'		-0.189	-0.451	-0.466	-0.384		0.047	0.158
'Turkey'	0.272	0.004	0.725	0.041	0.109	0.410	0.020	0.159	
'UK'	0.676	-0.005	0.329	0.341	-0.020	0.440	0.001	0.007	
'USA'	0.510	-0.032	-0.333	0.619				-0.028	
CPI	'Argentina'	0.258	0.045	0.546		-0.293			0.160
	'Australia'	0.520	-0.029	0.448		-0.152	0.333		-0.024
	'Austria'					-0.152	0.474	0.162	
	'Belgium'	0.337		0.145	-0.238	-0.254	0.672	0.040	-0.152

Notes: The Table reports factor loading estimates  $\Lambda$  in equation (1). Empty cells indicate the corresponding  $\lambda_{ij}$  are shrink to zero.

Series	Countries	Factors							
		1	2	3	4	5	6	7	8
CPI	'Brazil'	-0.259		0.549	0.022	-0.041			0.331
	'Canada'	0.408		0.257	0.006	-0.212	0.582	0.059	-0.030
	'China'	-0.295		0.222			-0.147	0.014	0.163
	'Chile'	0.221		0.545	0.152		0.275	-0.027	0.261
	'Finland'	0.422	0.046	0.360		-0.222	0.564	0.018	
	'France'	0.447		0.222		-0.068	0.742	0.044	-0.040
	'Germany'					-0.190	0.619	0.092	
	'India'		-0.150				0.239		
	'Indonesia'				0.314	0.354			
	'Italy'	0.370	-0.015	0.345	0.012		0.733	0.008	
	'Japan'				0.193		0.383		0.282
	'Korea'			0.040	0.411	0.045	0.529	0.088	0.143
	'Malaysia'						0.462	0.105	
	'Mexico'	0.438		0.722	-0.171	0.074			
	'Netherlands'						0.603		-0.060
	'Norway'	0.457	0.067	0.353			0.378		
	'New Zealand'	0.547	-0.055	0.418	0.074		0.367		
	'Peru'			0.558		-0.372			0.155
	'Philippines'	0.165	0.080	0.343			0.289		
	'South Africa'	0.158	-0.039	0.596			0.208	-0.101	0.192
'Saudi Arabia'			-0.245		-0.294	0.148			
'Singapore'					-0.228	0.515			
'Spain'	0.308	-0.051	0.348			0.555			
'Sweden'	0.298		0.409		-0.249	0.503		0.037	
'Switzerland'			0.327		-0.217	0.463			
'Thailand'		-0.047		0.316		0.512	0.192		
'Turkey'	-0.414	0.008	0.433	0.420	0.548	0.096		0.199	
'UK'	0.204		0.184	0.298	-0.002	0.495		0.186	
'USA'	0.211		0.144	0.309	-0.080	0.532	0.162		
nomEQ	'Argentina'	0.199		0.386		-0.240		0.235	
	'Australia'				0.201			0.721	
	'Austria'		-0.094			-0.152		0.606	
	'Belgium'		-0.078		-0.002	0.049		0.711	
	'Canada'				0.108			0.610	-0.018
	'Chile'		-0.034	0.257	0.326			0.259	0.161
	'Finland'		0.030		0.000	0.190		0.543	
	'France'				0.096	0.013		0.775	
	'Germany'							0.510	
	'India'					-0.168		0.448	
	'Italy'				0.146	0.119	0.095	0.604	
	'Japan'	0.194						0.567	
	'Korea'						-0.151	0.402	
	'Malaysia'				0.206	-0.023		0.470	
	'Netherlands'	-0.035	0.079		-0.005		0.030	0.825	
	'Norway'				0.133	-0.025		0.724	
'New Zealand'	0.057					0.112	0.521		
'Philippines'			0.207			-0.250	0.409		

Notes: Continue Table 4.

Series	Countries	Factors							
		1	2	3	4	5	6	7	8
nomEQ	'South Africa'				0.157			0.584	
	'Singapore'				0.219	-0.032	0.005	0.680	
	'Spain'			0.076		0.016	-0.020	0.682	-0.045
	'Sweden'		0.095				0.016	0.749	-0.014
	'Switzerland'	-0.143						0.782	
	'Thailand'					-0.179	-0.147	0.453	
	'UK'		0.010	0.007			0.075	0.774	
	'USA'							0.550	
Fxdol	'Argentina'	0.329	0.107	0.504	-0.163	-0.327			
	'Australia'		0.271			0.193		-0.344	0.188
	'Austria'	-0.012	0.941	-0.032	0.043	-0.003	-0.073	0.327	-0.006
	'Belgium'	0.040	0.921	-0.016	0.010	-0.001		0.307	-0.023
	'Brazil'	-0.301	0.013	0.530				-0.011	0.356
	'Canada'	-0.053	0.256			0.250		-0.365	0.252
	'China'			0.240	-0.029				0.156
	'Chile'	0.261	0.202	0.368	-0.221			-0.287	0.000
	'Finland'	-0.038	0.854				-0.029	0.260	0.009
	'France'	0.055	0.920	0.001				0.300	-0.019
	'Germany'	-0.007	0.939	-0.034	0.054		-0.062	0.324	-0.001
	'India'	-0.146	0.236	0.177	-0.119			-0.180	0.158
	'Indonesia'		0.069			0.171			
	'Italy'		0.872	0.018	-0.003		0.026	0.262	0.008
	'Japan'		0.341	-0.148	0.000			0.228	
	'Korea'		0.000			0.113	0.204	-0.247	
	'Malaysia'		0.258			0.189		0.000	
	'Mexico'	0.348	-0.029	0.475	-0.287	0.106		-0.255	
	'Netherlands'		0.938	-0.033	0.059	0.005	-0.056	0.323	
	'Norway'		0.596			0.112			0.107
	'New Zealand'		0.388			0.026		-0.169	0.100
	'Peru'		0.009	0.476		-0.308		0.000	0.112
	'Philippines'			0.183	-0.004	0.234		-0.154	
	'South Africa'		0.347	0.057	-0.165			-0.162	
	'Saudi Arabia'	0.326		0.311	-0.218				
	'Singapore'		0.493			0.232	-0.150		
	'Spain'		0.848		-0.025	0.043	0.048	0.282	0.003
	'Sweden'		0.601		-0.053	0.030			0.054
	'Switzerland'		0.672					0.269	
	'Thailand'		0.235			0.155			
'Turkey'	-0.263	0.237	0.223	0.294	0.462		-0.033	0.278	
'UK'		0.535		-0.132					
Rshort	'Argentina'			0.154		-0.144			
	'Australia'	0.338		0.872	-0.054	-0.224	0.169		-0.051
	'Austria'	-0.043	0.007	0.639	0.037	-0.053	0.585	-0.067	-0.036
	'Belgium'	0.093	-0.001	0.634	-0.013		0.599	-0.051	0.008
	'Brazil'	-0.010		0.234		-0.229			

Notes: Continue Table 4.

Series	Countries	Factors							
		1	2	3	4	5	6	7	8
Rshort	'Canada'	0.326	0.019	0.600		-0.051	0.594	-0.031	0.000
	'China '	-0.371	-0.002	0.654		0.157	0.101	0.072	0.232
	'Chile'	0.234		0.548			0.516	-0.054	0.122
	'Finland'	0.187		0.812	-0.086	-0.114	0.378	-0.040	0.019
	'France'	0.296		0.640	-0.142	0.073	0.621	-0.026	-0.125
	'Germany'	-0.044	0.017	0.585	0.029	-0.065	0.653	-0.083	-0.071
	'India'	-0.412	-0.004	0.501	-0.029	0.420	0.138	0.061	0.386
	'Indonesia'	-0.166	0.047	0.253	0.284	0.512			
	'Italy'	0.214		0.710	-0.169	0.007	0.542	-0.016	0.001
	'Japan'	0.258		0.539	-0.019	-0.116	0.584	-0.005	0.170
	'Korea'	0.136	0.001	0.419	0.393	0.150	0.556	0.002	0.158
	'Malaysia'	0.239		0.691	-0.045	0.304	0.408	0.018	-0.028
	'Mexico'	0.379		0.845	-0.100	0.108	-0.090	-0.004	0.052
	'Netherlands'	0.000	-0.011	0.620	0.136		0.588	-0.077	
	'Norway'	0.248	0.036	0.867	0.000		0.235	-0.110	
	'New Zealand'	0.331	-0.029	0.868	0.077		0.065		
	'Peru'			0.313		-0.362			
	'Philippines'		0.030	0.798		0.067	0.172		
	'South Africa'	-0.248	0.060	0.699	-0.130	0.007	0.000		-0.055
	'Singapore'	0.225	0.020	0.495	0.000	0.155	0.660		
'Spain'	0.146	0.019	0.758	-0.022	0.009	0.428		0.066	
'Sweden'	0.077		0.810	-0.107	-0.047	0.408	-0.023	0.000	
'Switzerland'	-0.395	-0.002	0.643	0.019	-0.197	0.186	-0.003	0.116	
'Thailand'	-0.010		0.587	0.000	0.277	0.518	0.020	0.040	
'Turkey'	-0.552		0.494	0.031	0.463	-0.013	0.000	-0.165	
'UK'	0.170		0.747	0.212	-0.009	0.421			
'USA'	0.324	0.063	0.596	0.068	0.006	0.572	0.018	-0.102	
Rlong	'Australia'	0.308	-0.006	0.860	-0.148	-0.102	0.373	0.006	-0.096
	'Austria'	0.057	-0.017	0.748	-0.032	0.033	0.571	-0.018	-0.191
	'Belgium'	0.222	-0.009	0.725	-0.145	-0.012	0.601	-0.016	-0.149
	'Canada'								0.143
	'France'	0.297		0.702	-0.158	-0.030	0.616	-0.021	-0.150
	'Germany'			0.732	-0.009	0.033	0.579	-0.002	-0.210
	'Italy'	0.167	-0.018	0.652	-0.209	-0.012	0.610	-0.005	-0.001
	'Japan'	0.262		0.621	-0.035	-0.087	0.618		0.040
	'Korea'	0.058		0.520	0.260	0.206	0.605		0.096
	'Netherlands'	0.071	-0.005	0.686	-0.012	0.018	0.644	-0.014	-0.185
	'Norway'	0.293		0.903	-0.076	-0.011	0.288	-0.021	-0.085
	'New Zealand'	0.395	-0.015	0.874	0.043	-0.011	0.203		-0.039
	'South Africa'	-0.233	0.040	0.831	0.002	0.233	-0.027		0.007
	'Spain'	0.184	-0.003	0.730	-0.129	-0.014	0.502		0.032
	'Sweden'	0.121	-0.021	0.828	-0.051	0.025	0.423	0.001	-0.052
	'Switzerland'	-0.152	-0.024	0.784		-0.024	0.436		-0.208
'UK'	0.188	-0.007	0.726	0.036	0.027	0.549		-0.047	
'USA'	0.324	0.047	0.704	-0.092		0.562	0.007	-0.138	

Notes: Continue Table 4.