PREDICTING THE UNPREDICTABLE: COMPARING FORECASTING MODELS OF ENERGY PRICES AND INTERNATIONAL COMOVEMENTS OF ENERGY CONSUMPTIONS

by

MUSTAFA ASLAN

JUNSOO LEE, COMMITTEE CHAIR
XIAOCHUN LIU
BURCU KESKIN
PIYALI BANERJEE

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts
in the Department of Economics, Finance, and Legal Studies
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2022

ABSTRACT

The first chapter examines the performance of various forecasting models on US electricity prices. This study is, in essence, a "meta-analysis," employing nearly 20 different forecasting models. Our model selection acts as a retrospective on the many different generation forecasting. We largely generalize them into two groups: *Old School* (OS) and *New School* (NS) models. This distinction was initially noted in Enders et al. (2009). They suggest that the *Old School* forecasting methods utilize standard linear models or various smoothing functions. The *New School* models involve threshold models, structural change models, regime-switching, and other modern approaches. We add some new forecasting models to them. For performance assessment, we focus on out-of-sample forecasting errors.

The second chapter examines how international comovements of energy consumption create global and regional effects on nations' energy consumption. The existing literature links the energy consumption behaviors of countries to country-specific components and examines the comovement of certain variables with the energy consumption of some specific countries. However, this study presents detailed empirical evidence of exogenous factors on energy consumption. We employ the time-varying Bayesian Dynamic Factor model with loadings and stochastic volatility, which permits us to estimate the global, regional, and country-specific factors affecting the energy consumption of the 52 countries divided by six different regions. We find that the magnitude of the effects is time-dependent, the global factor is generally a more dominant driving source of energy consumption than regional and idiosyncratic components, and

there is a significant regional factor effect in the European region. Moreover, we find clear evidence of international shocks linked to real-world events such as the 1973 OPEC crisis, the 1979 Iranian revolution, the 1997 Asian Financial crisis, the 2008 Great Recession, and renewable energy agreements between Nordic countries. We emulated the approach of the literature by comparing variables labeled as driving factors to our factor model and found that GDP and urbanization variables have pronounced effects on the international comovement of energy consumption.

DEDICATION

I dedicated this master thesis to my family. I am extremely grateful to my parents for their love, prayers, caring, and sacrifices for educating and preparing me for my future. Their emotional support during my master thesis process means a lot to me.

LIST OF ABBREVIATIONS

NS	New school
OS	Old school
RS	Regime Switching
D_1	First differenced AR(p) model
D_2	Second differenced AR(p) model
$AR_{(p)}$	P_{th} Level Regressive Integrated Moving Average Model
ES	Exponential Smoothing Model
Rm1	RALS forecasting with 2nd moment
Rm2	RALS forecasting with 3rd moment
Rm3	RALS forecasting with 2nd, and 3rd moment
ARMA	Autoregressive moving average
AP	Andrews Ploeberger
$AP_{(a)}$	Andrews and Ploeberger methodology all variables included
$AP_{(p)}$	Andrews and Ploeberger methodologies based on post break data
BP	Bai Perron
$BP_{(a)}$	Bai-Perron methodology, all variables included
$BP_{(p)}$	Bai-Perron methodology based on post-break data
TAR	Threshold autoregressive
M-TAR	Momentum threshold autoregressive

RALS	Residual augmented least squares
LAD	Least absolute deviations
LS	Least square
DF	Dickey-Fuller
ADF	Augmented Dickey-Fuller
US	United States
DM	Diebold Mariano
T	Number of observations
SSR	Sum of squared residuals
m	Mean of loss differential
BIC	Bayesian information criteria
MSPE	Mean squared prediction error
NYMEX	New York Mercantile Exchange
BRICS	5 major economic countries; Brazil, Russia, India, China, and South Africa
DFM	Dynamic factor model
EU	Europe
OLS	Ordinary least squares
NARDL	Nonlinear autoregressive distributed lag
VAR	Vector Autoregression
SVAR	Structural VAR
TVP-VAR	Time Varying Parameters VAR
GMM	Generalized Methods of Moments
ARDL	Autoregressive Distributed Lag Bounds

HGCM	Hsiao's Versions of Granger Causality Method,
VCM	Vector Correction Model,
PVECM	Panel Vector Error Correction Model
DPDM	Dynamic Panel Demand Models
IRF	Impulse Response Function
BTA	Bounds Testing Approach
PDA	Panel Data Analysis
ECM	Error Correction Models
PCM	Panel Cointegration Methods
PCT	Panel Causality Test
PP	Philips Perron
SDA	Structural Decomposition Analysis
FDI	Foreign Direct Investment
P	Population or population growth rate
GDP	Gross Domestic Product
EG	Economic Growth
EP	Energy Price
GE	Gas Emissions
CE	CO2 emissions
URN	Urbanization
I	Income
FD	Financial Development
IND	Industrialization

WDI	World development indicator
OPEC	Organization of the petroleum exporting countries
CI	Confidence interval
FE	Fixed effect
FM	Factor model
IFE	Interactive fixed effect

ACKNOWLEDGEMENTS

Firstly, I would like to thank the Republic of Turkey Ministry of Energy and Natural Resources for providing me with the scholarship for my MA program. I would also like to thank my advisor, Professor Junsoo Lee, for his invaluable advice, continuous support, and patience during my study. I would also like to express my gratitude to Professor Xiaochun Liu, Professor Burcu Keskin, and Dr. Piyali Banerjee for agreeing to serve on my committee and providing helpful comments. Last, I express my thanks to Paul Bousquet, a Ph.D. student in the Economics department, for his precious suggestions and valuable help.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: PREDICTING THE UNPREDICTABLE: COMPARING FORECASTING MODELS OF ENERGY PRICES

1.1 Introduction

Due to their connection to the global political and economic climate, which is notoriously difficult to predict, energy prices have a comprehensive and complex price structure. Namely, there is always a potential for a shock somewhere in the world that significantly affects the status quo for energy production or demand. Killian (2015) distills a broad history of volatility in the modern world to focus on a few key types of such occurrences, such as internal political conflicts, wars involving petrostates, and fluctuations in energy production input prices. Examples of such events that had a blatant impact on energy prices include the 1990 invasion of Kuwait, Venezuelan unrest in 2002, the Iraq War, and Arab Spring. Especially after becoming a crude oil importer in the 1970s, US energy prices were adversely affected in all these cases. With the advent of globalization and the corresponding trend of rapid industrialization for developing nations (e.g., China), many countries have become economically dependent on fossil fuels and ramped up demand for raw materials. Such a dependency furthers the propensity for US energy prices to be affected by non-domestic factors. In other words, the shocks occurring in these factors cannot be considered separately from each other.

Given the empirical abundance of unforeseen events and trends fostering stark shifts in price, we would expect to see some level of structural change in a time series of US energy prices. However, it is not always straightforward to determine the size and location of such

breaks. Moreover, the issues of non-stationarity and structural change are extremely relevant to creating good forecasts, and they are used as a beginning motivation for our results. The heightened likelihood of structural breaks in this series signals additional complications in forecasting (i.e., the difficulty of constructing a model that anticipates the break) and determining a unit root, which compounds the effect on forecasting. With these considerations in mind, the broad aim of this paper is to use a sweeping array of models to determine which produces the most accurate assessment of the behavior of monthly US electricity prices and which has the most predictive power. Our analysis incorporates nearly 20 different forecasting specifications. We note that many past studies focus on one or a few models in their results and discussion. Overall, the structure of our approach can be thought of as a "model race" or a meta-analysis of the performance of popular models from different generations and schools of thought.

Since we have so many models at our disposal, we largely generalize them into two groups: New School (NS) and Old School (OS). This distinction of the New School versus Old School model is derived from Enders et al. (2009). They suggest that *Old School* forecasting methods utilize standard linear models or various smoothing functions. As a good reference point for forecast performance, the OS methods include traditional time series models (*ARIMA*), exponential smoothing, and simple differenced models. In contrast, the *New School* view adopts properly estimated break dates which can be used to control for regime shifts when forecasting. The NS models are centered around variations of the Andrews-Ploberger (AP, (1994)) method, which allows a single break in the series, and the Bai Perron (1998) estimation methods, which can control multiple breaks. In this paper, we add the regime-switching *TAR* and *M-TAR* models developed by Tong and Lim (1980), providing the opportunity to change between states but not predict the location of breaks. Moreover, we also consider two recent developments of the NS

models: the Residual Augmented Least Squares (*RALS*) estimation, which improves prediction performance by decreasing the error variance, and a Fourier model, which estimates the location structural breaks using a functional form approximation.

As alluded to earlier, a relevant question for any researcher with a forecasting motivation is whether a unit root is present in the data. A common approach is to adopt a pre-testing procedure. An influential paper by Diebold and Kilian (1999) suggests that unit root test results could motivate the selection of forecast models. That is, one may use differenced data for forecasting if the null of a unit root is not rejected and use level data otherwise. However, there can be two major questions.

First, is it always better to use differenced data if a variable of interest is non-stationary? This paper shows that using level data for a non-stationary variable does not necessarily yield inferior forecasts than a differenced version of the same series. It is possible that the forecasting models using the level data of non-stationary variables can perform better. This ambiguity has not been fully considered in the literature. One prominent reason for using the first differenced data comes from the argument of spurious regression. When the variables are non-stationary, usual statistical inference can be affected, and the possibility of spurious regression remains in usual regression models. But, the forecasting model may not be subject to the problem of spurious regression. For example, an autoregressive regression of the usual *ARMA* model can be viewed as a cointegration model since a non-stationary variable, say y_t, is trivially cointegrated with its lagged variables, y_{t-k}; see Campbell and Shiller (1987). Then, the level regression can have a stable long-run cointegration regression. Also, one practical issue is which model specification is more effective in estimating structural changes. When using differenced data, the usual procedure tends to pick up too many structural changes, although the number of major

structural changes is rather limited, especially when evaluated from the models using level data. We will show such a case in our application to US electricity prices.

Second, there is a question of which unit root tests to use. A popular approach is to use modified versions of the Dickey-Fuller (DF) tests, given the potential for non-linearity trends and structural change. We use such tests, as well as a more comprehensive alternative of the Fourier unit root test from Enders and Lee (2012). However, considering that results from unit root tests can vary depending on the adopted test, this creates additional justification to not solely consider differenced data if faced with a lack of unit root rejections.

We examine the performance of the forecasting models using level data against a differenced series, given initial results pointing to non-stationarity, where we identify for which models a level or differenced data should be preferred (with respect to our given time series). We formally examine this conjecture using a Diebold-Mariano statistic.

Our model race shows a wide disparity in performance. Even some NS specifications had less predictive power than the traditional models. Overall, the Andrew-Perron tests produced the best forecasts according to squared error, with the Fourier, *TAR*, and the *RALS* models also demonstrating low error rates. Additionally, our Fourier tests support non-stationarity, but the Diebold-Mariano tests show that model forecasting with level data was better for most models.

The rest of the paper is organized as follows. Section 1.2 includes information on our data, approach, and models. Section 1.3 gives the unit root and forecasting results. Section 1.4 provides a comparison of forecasts in level and with differenced data. In Section 1.5, we make concluding remarks.

1.2 Methodological Background

1.2.1 Data, Initial Unit Root Testing, and Approach

Our paper focuses on forecasting US electricity prices. Specifically, our time series of choice is monthly average retail prices of electricity data (Cent per Kilowatt hour, including taxes) from 1990-2019 (yielding 359 observations), seasonally adjusted and sourced from the U.S Energy Information Administration.

With a primary research focus on forecasting, the potential for non-stationarity in the time series is a relevant concern. We begin with Dickey-Fuller (DF) tests as a baseline frame of reference. But, the DF tests fail in the presence of structural changes (Perron, (1989)). One may consider using dummy variables for structural changes. To help account for this difficulty, Perron (1998) suggests that if dates of the breaks are known, the standard DF test can be appropriate and changes in trends and levels by adding in dummy variables. While adapting the Perron test can be appealing from a simplicity perspective, it is unlikely that the duration, number, and type of structural breaks will be known. Moreover, multiple breaks can exist. A more flexible and effective method is available. Enders and Lee (2012) suggest approximating unknown forms of structural breaks and nonlinearity using the Fourier function. The Fourier unit root test, in which the deterministic component is expressed with a time-dependent function such as $s(t)$, is developed based on the below equations

$$\Delta y_t = s(t) + \beta y_{t-1} + \sum_{\mathrm{i}=1}^{p} \rho_i \Delta y_{t-i} + \mu_t \tag{1}$$

$$s(t) = c_0 + \sum_{\mathrm{k}=1}^{m} \left(c_k \sin\frac{2kt\pi}{T} + a_k \cos\frac{2kt\pi}{T} \right); \mathrm{m} \leq \frac{T}{2} \tag{2}$$

where, ε_t has constant variance (σ_ε^2), k is the fundamental frequency parameter, T is the number of observations, and m is the number of cumulative frequencies.

Conforming to the standard paradigm, the null of unit root ($\beta = 0$) is tested against the alternative hypothesis $\beta < 0$. In the instances where $c_k = a_k = 0$, for any k, holds, there is no non-linear trend, and the standard DF test works better. One can consider two cases. First, one can use the number of cumulative frequencies (m) to facilitate the selection. For example, if m = 1, we use k = 1. If m = 2, we use the first two frequencies, k = 1 and 2. Also, if m = 3, we employ three frequencies, k = 1, 2, and 3. And so on. Usually, using m = 1, 2, or 3 is recommended, and the difference is not noticeable. Indeed, the first frequency k = 1 captures major parts of a non-linear trend. Second, one can use one particular frequency, k. Assessing the potential of non-linearity can be done using a standard hypothesis test of this property, whereby the following F statistic shows if there is a statistically significant difference between (1) and the series without trigonometric terms

$$F(k) = \frac{\frac{SSR_0 - SSR_1(k)}{2}}{\frac{SSR_1(k)}{T-q}} \qquad (3)$$

where the sum of the squared value of the restricted model is identified by $SSR_1(k)$ and SSR_0 for the unrestricted model. Here, T and q, respectively, represent the number of observations and the number of regressors. The fundamental value frequency (k) of the Fourier cycle is estimated with a grid-search methodology. Then, the value of k that yields the smallest SSR is chosen.

Further, although the Fourier unit root test can be limited to estimate sharp breaks, it shows better performance when the series has smooth breaks. When these breaks are momentary or instant, applying the standard DF test dummy can be more convenient (Enders and Lee,

(2012)). In this study, the breaks that were wanted to control are expected to be not instantaneous; therefore, applying the Fourier Unit Root test can give better results.

Much of the discussion centered around data with a unit root is that its behavior is difficult to predict, or more formally, it does not empirically revert to a mean and that differencing the data can remove this trait. However, as noted above, the forecasting ability is a different issue. The forecasting models using level data, even with non-stationary variables, can be valid and exhibit better performance.

To compare the forecasting performances of models in level and first differenced data, we employ the Diebold Mariano (1995) test. The Diebold Mariano (DM) test statistic can be given by

$$D = \frac{\bar{m}}{\sqrt{\frac{2\pi\hat{z}_m(0)}{T}}} \qquad (4)$$

where T is the number of observations, and $\hat{z}_m(0)$ is a stable estimator of the spectral density of the zero-frequency loss differential function (m). Squared deviation of errors of the forecasting models is defined with $m_i = e_{1i}^2 - e_{2i}^2$ and $\bar{m}$ represents mean of loss differential (m_i). In general, the DM test can be utilized to examine the relative performance of two competing models.

As stated earlier, the unique characteristic of this study is the volume of model specifications used. Thus, we can develop not only a comparison of how well US electricity prices are forecasted but have a wide array of DM statistics to give a more comprehensive picture of our indirect check on the quality of competing models. The procedure can also be used to examine if forecasting power using level data are equal to forecasts with differenced data.

Our interest lies in examining which models perform better. Enders, Lui, and Prodan (2009) define the traditional models as the Old School (OS) Models, which are; *ARIMA (p, d, q)*, first and second differences, and exponential smoothing models against the modern approaches (i.e., the New School Models). The new school models with regime-changing functionality include the model founded by Andrews Ploeberger ((1994), *AP*) and the break models suggested by Bai-Perron ((1998), *BP*). We also include threshold regime-switching models, *RALS*, and Fourier models. We have discussed briefly the Fourier models above. We will discuss more details of these new models below. Table 1.1 provides the list of all the models (and specifications) we plan on using and their corresponding abbreviations.

1.2.2 Traditional Approaches

We begin to discuss *ARIMA* models (p, d, q), also known as the Box-Jenkins method. They can be used to predict the mean in a time series. This process is defined by

$$\Delta^d y_t = c_0 + \sum_{i=1}^{p}(c_i \Delta^d y_{t-i}) + \sum_{i=1}^{q}(a_i \Delta^d y_{t-i}) + \varepsilon_t \qquad (5)$$

In the equation, ε_t is expressed as the error term, while y_t is defined as the predictor variable, and p and q values show the degree of the autoregressive process and the number of lagged forecast errors in the prediction equation orderly. To find the optimum values of p and q, BIC (Bayesian Information Criteria) is used with $ARMA_{(p,\,q)}$ and $AR_{(p)}$ processes with different values of p and q. Also, the *AR* models used only in stationary models are stabilized by taking a serial difference to use non-stationary models. The value of d indicates the degree of this difference, for which different unit root tests are used.

Moreover, the first and second lagged *AR* processes ($AR_{(p,\,1)}$, $AR_{(p,\,2)}$ called as D_1, and D_2 models) are also included in the OS models to increase the forecasting comparison performance.

Here it should be noted that where the Y_t series has no unit root, differencing the Y_t variable in order to reduce the impact of the breaks, create a unit root in the series. Enders et al. (2009) suggest a new method based on Diebold and Killian's (1999)[1] recommendation. The method is to perform a Dickey-Fuller unit root test for the simulated Old School series, and according to pretest results, the variable is estimated in level or indifference. The "*Pre*" model added to forecasting comparison not to select true model between differenced and undifferenced series, but to find out if the pretest improves the prediction performance.

As a final OS model specification, the exponential smoothing method (*ES*), one of the forms of the weighted moving average, is applied. *ES* estimates the model based on the exponentially weighted average of past observations. Even though *ES* models have a simple structure, other versions include seasonality or a trend. One of the most notable advantages of this *ES* forecasting, while it is working, the impact of past breaks would be decreased based on how long the break was in the past. The general form of *ES* can be given by

$$\hat{x}_t = \hat{x}_{t-1} + T_{t-1} + \gamma(y_{t-1} - \hat{x}_{t-1}) \qquad (6)$$

where $\hat{x}_t$ and T_t are the past observations weighted averages and trend components, respectively.

Again, we stress that these traditional models are included to give a reference point for forecast performance. Furthermore, the OS models are known to still produce useful results, even if they do not control for structural breaks. In addition, it has been observed that NS models show low forecasting performance if the break dates and magnitudes are not well adjusted (Enders et al., (2009)). Therefore, considering the models in this section (and the varying specifications) is important for this paper to be as comprehensive as possible.

[1] In cases where the series does not have a unit root, differencing the series will cause an over differencing problem. Diebold and Killian (1999) recommend that to avoid this over-differencing problem, a pretest should be performed for unit root before taking the difference of the series.

1.2.3 The Modern Approaches

As a modern approach, we can first consider the models of Andrews and Ploeberger (1994), in which two different forms of a single break are imposed. The general form follows

$$y_t = \beta_j + \sum_{i=1}^{p} \alpha_i \, y_{t-i} + \varepsilon_t \qquad (7)$$

$$y_t = \beta_j + \sum_{i=1}^{p} \alpha_i^j \, y_{t-i} + \varepsilon_t \qquad (8)$$

$$y_t = \begin{cases} \beta_1 + \sum_{i=1} \alpha_i^1 \, y_{t-i} + \varepsilon_t & \text{if } T_B > t \\ \beta_2 + \sum_{i=1} \alpha_i^2 \, y_{t-i} + \varepsilon_t & \text{if } T_B \leq t \end{cases} \qquad (9)$$

where T_B defines the estimated period at which breaks happen.

There are a few key things to observe about this form. Namely, the partial structural break model (which is called *AP* model in this paper) in equation (7), the break can only happen on the intercept of the equation (β_1, β_2). Unlike equation (7), for the $AP_{(a)}$ model in equation (8), we set the pure structural break model, and all parameters are allowed to change according to T_B. The break location T_B is searched in all sampling periods except the first and last 5% of the observations (i.e., trimming value ε = 0.05 is chosen), and the optimum values of break dates are estimated. Moreover, the supremum F-test is applied for each possible break date to decide whether there is a structural break in the series. According to the test result, if there is a break in the series at the moment at T_B, we can predict the equation using (7) or (8) and obtain out-of-sample forecasts. Otherwise, when the test concludes with no structural break, all data set is estimated using the *AR* process.

As the Andrews Ploberger method allows only one break, this method is insufficient when multiple breaks are present. To solve this problem, Bai and Perron (2003) suggested a new

methodology to allow more than 1 break. In this methodology, for each break, a new T_B value is created (T_{B_1}, T_{B_2},...., T_{B_n}) using (7) and (8), and these equations are redefined based on the value of j. Also, for this methodology, the trimming value is chosen 0,05 again, and a maximum of 4 breaks are allowed. Whether there were any breaks in the series was determined again using the supreme F test.[2]

1.2.4 Regime Switching Models

The threshold autoregressive model (*TAR*) was first suggested by Tong and Lim (1980) and later developed further by Tong (1990). We can determine a certain threshold value and apply it to different regimes with this model. Unlike NS models, the break process is predicted by the other parameters of the model in the RS models. The model allows for several different regimes with a separate autoregressive model for each regime. In other words, although all breaks are considered permanent in NS models and cannot be considered as part of the data generation process, breaks occur suddenly and instantaneously when exceeding the threshold value. Moreover, even though all series show non-linear characteristics, it is accepted that each regime is linear in itself. The general form of 2-regime *TAR* models follows

$$y_t = \phi_t \left[\beta_{10} + \sum_{i=1}^{p} \beta_{1i}\, y_{t-i}\right] + (1-\phi_t)\left[\beta_{20} + \sum_{i=1}^{p} \beta_{2i}\, y_{t-i}\right] + \varepsilon_t \tag{10}$$

$$\phi_t = \begin{cases} 1 & \text{if } y_{t-d} \geq \tau \\ 0 & \text{if } y_{t-d} < \tau \end{cases} \tag{11}$$

[2] In addition to the above methods, Pesaran and Timmerman (2004) proposed a new method that allows the series to be predicted in the *AR* process and uses only the final break if the break in the series is too large. Therefore, using only the $T_{B_n} \leq t \leq T$ time path, the series is predicted by the $AR_{(p)}$ process. This allows us to form a "Post-Break" structure.

In (10), the first regime parameters are defined with β_{10}, β_{1i}, second regime parameters are shown with β_{20}, β_{2i}, the τ value indicates the threshold value obtained by grid research, and d, p, and ϕ define a lag parameter, the degree of $AR_{(p)}$ process, and Heaviside indicator function, respectively. In both regimes, it is assumed that the error terms are independent, identically distributed, and have white noise characteristics. In this *TAR* model, in the "high" case of $\phi = 1$, the value of y_{t-d} exceeds the threshold value. In contrast, if y_{t-d} is below the threshold value, this state is called "low," and the autoregressive process is created according to the $\phi = 0$ state.

Another regime-switching model to capture breaks in the series is the momentum threshold autoregressive (*M-TAR*) model. *M-TAR* model, which allows regime change according to the first difference of y_t (i.e., Δy_{t-d}), was developed by Enders and Granger (1998). The specification for the *M-TAR* models is defined as

$$\phi_t = \begin{cases} 1 & \text{if } \Delta y_{t-d} \geq \tau \\ 0 & \text{if } \Delta y_{t-d} < \tau \end{cases} \tag{12}$$

where Δy_{t-d} is used as a threshold variable instead of y_{t-d}. In the case of a “high” state, Δy_{t-d} is greater than a threshold value. That is, when $\Delta y_{t-d} > \tau,\ \phi_t = 1,\ and\ y_t = \left[\beta_{10} + \sum_{i=1}^{p} \beta_{1i}\, y_{t-i}\right] + \varepsilon_t$). When $\Delta y_{t-d} < \tau$, *M-TAR* model indicates a low state case (i.e. $\phi_t = 0,\ and\ y_t = \left[\beta_{20} + \sum_{i=1}^{p} \beta_{2i}\, y_{t-i}\right] + \varepsilon_t$). In the case that y_t is assumed stationary, testing for the deepness hypothesis is matching to the testing the no skewness in Δy_t; see Tayyab et al. (2012). The main purpose of these threshold autoregressive models is that when a sufficiently large shock is sent to the series, this shock can be accounted for by triggering a switch in the system between regimes.

1.2.5 Miscellaneous Forecast Methods

It is a well-known fact that standard OLS estimators can be efficient regardless of the error distribution. The efficiency holds under both normal and non-normal distributions. However, their efficiency can be improved further under non-normal distributions. The residual Augmented Least Squares (*RALS*) estimator, developed by Im and Schmidt (2008), aims to increase the efficiency of autoregressive estimates by reducing the error variances by using covariate-augmented auto-regression. They claim that higher moments are not connected to regressors. Moreover, the *RALS* estimator has been shown to be more efficient than the LS estimator under moment constraints compensated by non-normal distributions. Also, Li and Lee (2015) show that the *RALS* estimator performs better than LAD and LS estimators if the data is skewed with lower MSPE values. Another advantage is that it does not require nonlinear estimations since the *RALS* estimator can be calculated by linear regressions extended by functions of LS residuals.

The *RALS* estimator requires a few steps. First, the p-th order AR (p) model is fitted with the LS method, under the assumptions that ($E(e_t|\beta_t) = E(e_t) = 0$, where the error ($e_t$) follows non-normal distributions (when $\mu_s = E(e_t^s)$, there is at least one integer s providing that $\mu_{s+1} \neq s\mu_{s-1}\mu_2$). A time series of interest is y_t, $(t = 1,2,3,..,n)$ and we can consider an *AR* process.

$$y_t = \beta_t'\rho + e_t \tag{13}$$

The $(\rho + 1) \times 1$ vector of regressors is given by $\beta_t = (1, y_{t-1}, y_{t-2}, .., y_{t-p})'$, $\rho = (\rho_0, \rho_1, \ldots, \rho_p)'$ shows the coefficient vector. After estimating (13), we use the residual ($\hat{e}_t = y_t - \beta_t'\hat{\rho}$) to construct a moment condition based on the following covariates.[3]

[3] Under the assumption that p is consistent, the sample covariate ($\hat{\theta}_t$) is asymptotically equal to population covariate ($\hat{\theta}_t$).

$$\hat{\theta}_t = \hat{e}_t^s - \hat{\mu}_s - s\hat{\mu}_{s-1} - \hat{e}_t \tag{14}$$

where $\hat{\mu}_s = n^{-1}\sum \hat{e}_t^s$ shows a sample moment. On the next step, we construct covariates $\gamma_t = (w_{1t}, w_{2t})'$, where $w_{1t} = \hat{e}_t^2 - \hat{\mu}_2$, and $w_{2t} = \hat{e}_t^3 - \hat{\mu}_3 - 3\hat{\mu}_2\hat{e}_t$, following the moment condition based on (13)[4]. Then, the $AR_{(p)}$ model is expanded to these covariates. The *RALS* estimator is obtained from the following autoregressive model by the LS method.

$$y_t = \beta_t'\rho^r + \gamma_t\lambda + \epsilon_t, \tag{15}$$

where we let $e_t = \gamma_t\lambda + \epsilon_t$ with $E(\gamma_t\epsilon_t) = 0$, which show the orthogonal decomposition of the error, and where $\lambda \equiv \frac{E(\gamma_t e_t)}{E(\gamma_t{}^2)}$. Finally, with this new *RALS* estimator ($\hat{\rho}^r$), the j-th step ahead forecast created with;

$$\hat{y}_{t+j}^r = \hat{\rho}_0^r + \sum_{i=1}^{j} \hat{\rho}_i^r\, \hat{y}_{t+k-i}^r. \tag{16}$$

While the *RALS* estimator has previously shown no bias when used in cross-sectional models, it is known to be biased when used in autoregressive time series models such as this study. Lee and Li (2015) made the following assumption.

Assumption 1: the bias converges to zero as the sample size increases,

$$[E(\hat{\rho}^r) - \rho][E(\hat{\rho}^r) - \rho]' = o(n^{-1}) \tag{17}$$

We assess this assumption as relatively non-restrictive, and the *RALS* estimator is $\sqrt{n}$-consistent.

Next, the Fourier approach is added. It captures the gradual structural break process in series using a few trigonometric parameters, as given in (2). The jth-step ahead Fourier forecasts are obtained for $(j = 1,2,..,23)$ using

[4] It was emphasized that the second and third moments are sufficient for the most appropriate practical approach; see Li and Lee (2015).

$$\Delta\hat{y}_{t+j} = \hat{s}_{t+j} + \beta\hat{y}_{t-1} + \sum_{i=1}^{n} \hat{\rho}_i \, \Delta\hat{y}_{t+j-i} + \mu_t. \tag{18}$$

The main advantage here is that the frequency components created by the trigonometric variables mimic the breaks in the series, and by adding them to the *AR (1)* process, the structural breaks can be controlled.

1.3 Application to Energy Price Data

1.3.1 Testing for Stationarity

When we look at the plot of the monthly US electricity price total over time, we can see that the series has a non-linear trend or structural break (Figure 1.1). The plot for the Fourier function and the plot of the energy price data are shown in Figure *1.1*. According to this figure, we can see that the series has 0.7 cycles from 1990:1 to 2019:11, meaning it will take approximately 41 years for the Fourier cycle to repeat itself. This is added confirmation that the series has a break throughout the time path.

To begin with, we first adopt the ADF tests, which do not allow for breaks. The calculated ADF value is -0.248, well below any standard $\tau_{critical}$ value for ADF test, meaning we cannot reject the null hypothesis and (according to this result) the series is non-stationary. Given the empirical evidence for a non-linear trend of energy data in the literature, we also check on the presence of non-linear breaks. Per our previous discussion, the F-statistic in equation (3) is 6.59, so we can reject the null of linearity at the 1% level. As stated earlier and further inferred from the non-linearity result, there is much reason to have lacking confidence in the ADF in this setting, so we then apply the Fourier unit root test. The t statistic for the Fourier unit root test is estimated as -3.196, which means we cannot reject the null hypothesis of the unit root (critical value is -3.75). That result implies that our data is non-stationary, and one may argue that proper

differencing of time series for forecasting comparison will provide greater stationarity. However, it will be helpful to examine if forecasting models using level data can yield better forecasts since the usual *AR* regression can be a cointegration regression model. It should be noted that differencing is applied not to eliminate structural breaks, which the Fourier test and some modeling methods can account for, but to remove the unit root in the series. Still, having in mind these results, we adopt the strategy of using both the level and the difference data and comparing the results.

1.3.2 Forecasting Results in Level

For our forecasting comparison, we apply all the models and specifications listed in Table 1.1 to our data set and evaluate the in-sample and the out-of-sample performance of these models. The in-sample period spans from 2013:1 to 2017:11, and we attempt to forecast the out-of-sample values from 2017:12 to 2019:10. The 23 steps forward estimation results of the models are obtained, and these results are compared with the data of the last 23 observations we separated from the estimation. For each model, the ratio of the Mean Square Prediction Error (MPSE) relative to the $AR_{(p)}$, MPSE is calculated as a comparison metric, with a lower ratio being indicative of superior performance. We first make assessments of the models using level data. The results are given in Table 1.2.

First, we evaluate the in-sample fit using the aforementioned cutoffs using the BIC values. Among the Old School (OS) models, the best performing models are the D_1, *Pre*, and $AR_{(p)}$ models, while the best results among the New School (NS) models and Regime Switching (RS) models such as the $BP_{(a)}$, *BP*, and $AP_{(a)}$ models. Also, these NS models have the best results of all models, although the difference between the Old School models and the New School models is marginal. However, in the previous study on the "Terrorism series" containing

structural breaks, it was observed that the sampling performances of the NS models were better than the OS models, but this result was reversed when out-of-sample forecasting performances were compared; see Enders et al. (2009). As such, we consider the "out of sample" prediction performance as bearing more weight for our model selection.

When the MSPE results obtained for 23 steps forward prediction is analyzed, the best performing model among all models is the *AP* model, followed by the $AP_{(p)}$, and Fourier models, respectively. Among the Old School models, the best result belongs to the *ES* model, while it is seen that the NS and RS models yield significantly better results than the OS models. Moreover, supporting Enders et al.'s (2009) findings, pre-test strategy doesn't increase the predictive accuracy among old school models. Although we see the modest forecasting accuracy increase in the *Pre* model, the difference does not look significantly important.

Overall, NS models outperform OS models. These results may reflect the presence of breaks in the series since NS models can capture breaks. Looking at the performance of OS models, we find that the exponential smoothing model surpasses the efficiency of other traditional models. The exponential smoothing (*ES*) model is specially equipped to reduce the weight of a break, further indicated by its strong performance with our data. We find that *AP* has the lowest MSPE values among all models. This may be partly because the series has a single large permanent structural break in the trend, rather than multiple offsetting structural breaks[5,6]. Also, the superiority of the *AP* and $AP_{(p)}$ models over the others are in line with the finding of

[5] Enders, in his previous forecasting study on series with breaks, found that OS models are superior to the NS models if the series contain offsetting and small break, as the result of the Monte Carlo experiment, otherwise (i.e., if the break in the data is large and relatively permanent), the NS models perform better.

[6] Andrews Ploeberger methodologies allows only one break while Bai Perron can empower more than one.

Casini and Perron (2018) that partial break models predict more precisely than the pure break models.

For the non-OS or NS models, the *RALS (Rm1)* forecasting method, which reduces the error variance by using 2nd order moment, shows good performance when it uses the information on non-normality. Although it does not explicitly incorporate structural breaks, its forecast errors are lower than many of the OS, NS, and RS models, which could be attributed to the kurtosis seen in the distribution of the U.S. energy price series. Finally, regarding the MSPE, we find that, other than the *AP* model, the Fourier model delivers the best performance. This result is encouraging because it shows the Fourier function, even with the low data frequency, is capable of mimicking smooth and gradual breaks, which can be observed from the plot of energy prices time series.

We have shown the forecasting performances of the models in Figures 1.2 and 1.3. We have omitted the results of D_1, and *ARMA* models since they are reasonably similar to the $AR_{(p)}$ model. As we can see from the forecasted values in Figure 1.2, the trend of forecasted values from the *ES* model is smaller than that from the OS models. Also, all the traditional models have an increasing trend in the predicted values, while the NS models show a decreasing trend. The general *AP* model exhibits the least monotonicity. In panel 4 of Figure 1.2, the predictive performances of *RALS*, Fourier, *AP*, *TAR*, and *ES* models are given. The predicted values from the Fourier model follow a slight fluctuating path till 2018:05, and then it shows a decreasing trend till the end of the forecast horizon, while *RALS* shows an increasing trend. Compared to the *AP* model, which gives the best estimation result, the slope of the Fourier forecasting trend is much less steep. The trend of the predicted values from the regime-switching models, *TAR*, and *M-TAR* models is not clear as it is mixed with positive or negative trends.

In Figure 1.3, we examine how closely the 23-step forward prediction results coincide with the real data. The predicted behavior of all models is shown together with the real data in the period 2017:12-2019:10. It is clearly seen here that among the OS models, the *ES* model shows the closest forecasting performance to catch the real data, while the D_2 model exhibits the worst behavior. In Panel 2, you can see how effective the *AP* model is in making 23 steps forward forecasting of energy price data. In addition, while *BP*, and $BP_{(p)}$ models follow a very close prediction path to each other, $AP_{(a)}$ has the worst forecasting characteristics.

In general, we find that the Fourier and *BP* models have the most dynamic forecasting potential. Most of the models output what is essentially a near-perfect linear trend, which is not desirable since the relevant data clearly exhibits non-linear trends. Thus, we prefer a model that can exhibit lots of fluctuations in trend as opposed to largely uniform monotonicity.

Estimated break dates in level: AP and BP models

We now examine the estimated breakpoints. We report the estimated break dates in Table 1.3. The AP models include only one break by construction. The *BP* model gives the result of four breaks, where the BIC value is minimized, while The $BP_{(a)}$ model gives the result of one break.

Notably, the value of a barrel of crude oil on NYMEX (New York Mercantile Exchange), which is one of the main factors in determining energy prices, was below US $25/barrel from 1980 to 2003, while in 2003, this price rose to over $30, and reached the level of $60 in August 2005. Later, showing an increasing trend, it reached $147 around July 2008. The reasons for this subsequent price increase were attributed to several events, such as the rising turmoil in the Middle East, the rise in demand from China, the falling US dollar value, and the decline of US oil reserves; see Killian (2016). As a result, the date of this increasing trend, which caused a

change in the regime, was estimated very closely by the *AP* model. This observation is consistent with the energy price plot shown in Figure 1.1. This result is consistent with the good forecast performance of the *AP* model. The $AP_{(a)}$, and $BP_{(a)}$ models find the break date as 2013:11 date, indicating relevant sticking points such as Libyan turmoil and Iranian sanctions.

1.3.3 Forecasting Results with Differenced Data

In this section, we examine the forecasting performance of various models when the first difference of the series is taken to make price series stationary. We have seen that the Fourier unit root test fails to reject the unit root null hypothesis. Thus, one may argue that it might be natural to use the model based on the differenced data for forecasting. As before, the last 23 observations are saved to examine the out-of-sample performance. A similar process was used as in the above subsection. We report the in-sample and out-of-sample performance in terms of the BIC and MSPE values in Table 1.4.

We begin with the in-sample assessment, with the parameters outlined in the preceding section. Among the Old School models, the best performing models are the *ARMA*, *Pre*, and $AR_{(p)}$ models, while the best results among the New School models and RS models belong to the $BP_{(a)}$, and *BP* models. Also, these NS models have better results, although the difference between the Old School and the New School models is modest. Forecasting using the Fourier models has the lowest in-sample fit performance among all estimation methods. As previously stated, the in-sample performance is premature and less important in evaluating the forecasting performance. Therefore, we look for MSPE values from the out-of-sample forecasts.

When we evaluate the out-of-sample performance using MSPE, we find that the best performing model among all models is the Fourier forecasting model, followed by the *RALS* with 2nd moment *(Rm1)*, and *BP* models, respectively. They have been obtained from 23 steps

forward prediction. Among the Old School models, the best result is shown from the $AR_{(p)}$ model, and we see that the pre-testing increases the forecasting accuracy slightly among old school models. However, the *RALS, RS,* and Fourier models give better results than the OS and the NS models. These results present an immediate quandary: when using the level data, the *AP* model performs the best, but this does not hold for differenced data. One explanation would be that the differenced data could exhibit multiple structural changes. Thus, allowing for one break would not be sufficient. Another explanation could be derived from the fact that the regime-switching models also outperform the OS, RS, and NS models using MSPE. This result suggests that the differenced series has multiple smooth offsetting breaks rather than sharp and permanent structural breaks. Thus, the aforementioned F-test may not be a sufficient break detection tool for the differenced data. The *RALS (Rm1)* forecasting method, which reduces the error variance by using 2nd order moment, performs well when utilizing the information on non-normality. It delivers a lower out-of-sample MSPE than all OS and NS models. Overall, the Fourier model performs the best, in terms of the MSPE selecting criteria, seemingly due to its ability to capture the smooth and gradual breaks for differenced prices, even with a low data frequency.

Next, we present the plots of the series along with the forecasted values in Figures 1.4 and 1.5. We again find that the Fourier model shows the most potential for outputting the desired non-linear trend, whereas even with the differenced data, most other models generate a functionally constant or linear fit. Compared with the old school models in Panel 1 in Figure 1.4, the estimation results of preliminary and *ARMA* models are similar to those of the $AR_{(p)}$ model. Since the data is already differenced, D_2 delivers extremely large MSPE values as a result of over-differencing, and is subsequently excluded from Figure 1.4. We observe that all the OS models show an increasing trend on the forecast horizon, but $AR_{(p)}$ has the smallest mean. On the

other hand, the D_1 model has an increasing trend. For the NS models shown in Panel: 2, *AP* and *BP* models follow an increasing trend path. In Panel: 4, the predictive values from the *RALS*, Fourier, *TAR*, and *BP* models are shown. They give the best forecasting results among all models. We can see that the forecasted values from *RALS* and *TAR* models have smaller variances of the confidence intervals than those from the Fourier models. In contrast, the Fourier model outperforms other models in terms of the MSPE values.

Figure 1.5 shows more precise plots of the 23-step forward prediction values. We present them along with the actual data that we have reserved for examining the out-of-sample performance for the period 2017:12 - 2019:10. It is seen here that among the OS models, the $AR_{(p)}$ model shows the closest forecasting performance to catch the real data, while the D_1 model exhibits the worst behavior. In Panel 4, you can see how effective the Fourier model is in making 23 steps forward forecasting of Energy Price data. In addition, we see that *RALS* and *TAR* models follow a very close prediction path to each other. These results are quite surprising, given the structures of these two different models. Moreover, as shown in Panel 3, the *TAR* model also shows substantially gratifying predictive efficiency. They also show a fluctuating path while their forecasting variances are pretty low.

Estimated break dates with differenced data

We now examine the estimated break dates. We rely on the *BP* model, which gives two estimated dates of 2004:12 and 2008:07. Again, the break dates coincide with the dates when the value of a barrel of crude oil on the New York Mercantile Exchange (NYMEX, (2014)) was below US$25/barrel, which occurred from 1980 to 2003. Later, in 2003 this price rose to over $30 and reached $60 in August 2005. Regarding the second estimated break date, we observe a

rapid decrease in energy prices after July 2008. The reasons for this subsequent price increase between 2003 and 2005 years were discussed above.

1.4 Comparing Level and Differenced Forecasts

In this section, we compare the forecasting performance of the models using the level data with the models using the first differenced data. Earlier, we noted that the unit root tests imply the energy prices are non-stationary, regardless of whether breaks are allowed for or not. Thus, it might be practical to use the first differenced data. But, forecasting could be a different matter, and it is possible that forecasting using the level data could work better even with non-stationary variables. Thus, it will be interesting to evaluate the forecasting performance of two different approaches for each model.

We report the Diebold-Mariano ((1995), DM) test results in Table 1.5. The first three columns of Table 1.5 show the DM statistic and the corresponding p-values. The DM test is frequently used to evaluate the relative accuracy of forecasts derived from two competing models. It is used here to evaluate the forecasting performance of the models using level and differentiated data for each of the same models. A positive value of the DM test statistic implies that the model using first differences outperforms the corresponding model using level data. Negative values of the DM test results indicate that the opposite is true.

Examining the DM test results, we see that the null of equal forecast accuracy can be rejected for 14 out of 17 models at the 5% significance level. In addition, DM test results indicate that the forecasting performances are better for 9 out of those 14 models when forecasted using the level data, namely the *RALS* models (*Rm1*, *Rm2*, and *Rm3*), $AP_{(p)}$, *AP*, *ES*, and Fourier models. On the other hand, the sign of DM statistic is positive for five models among the statistically significant results: *BP* models (*BP*, and $BP_{(a)}$), $AR_{(p)}$, *ARMA*, and *M-TAR*

models. Thus, we cannot say that the forecasts made from the models using level data are universally superior to our energy price forecasts. However, when we look at the models that give the best forecasting results for both cases (level and first difference), there is strong evidence that the models forecasted at the level have better predictive ability. In particular, the Fourier model has the best prediction performance among the models estimated using first difference data. Still, we see that the performance is even better when the Fourier is forecasted with level data. And even though the Fourier model in level is better than the best model for differenced data (itself), its performance with the undifferenced series lags behind *AP* and $AP_{(p)}$, models. Thus, in our analysis of forecasting performance using the electricity price series, the models using level data provide the largest propensity for accuracy.

In addition to the superiority of *AP* and $AP_{(p)}$ models to others using level data, we can say that Fourier and *RALS* models show better predictive ability than other old school and Bai Perron models regardless of whether the series is predicted at level or first difference. However, the *AP* model shows the best performance when forecasted in level, but they perform poorly when differenced data are used. It might be due to the failure of capturing multiple breaks, as the *AP* model allows for only one break. Possibly, more breaks could be observed when using differenced data since the date can exhibit more fluctuations. Here, when taking the difference of the series containing the breaks, it can be said that the prediction performance decreases roughly. Among the remaining models, since D_1 and D_2 models may have over-differencing problems when used with differenced data, it is not surprising that they show lower performance.

For a clearer explanation, the estimation performances of the models selected above by difference and level are compared in Figure 1.6. For the first three panels, based on the DM test results, the models that outperform others converge to recent values of the series. Here, panel 1

and panel 2 are based on the level models, while panel 3 is based on the differenced model. When it comes to the *RALS* model in panel 4, the model estimated at the level has an ascending trend towards the end of the sample period that diverges from the recent values of the series.

Overall, it can be said that forecasting models work better if the difference of the data is not taken and taking the difference of the data in particular to the US Monthly Energy Price index reduces the forecasting performance in the models, which perform better relative to others. Perhaps, this is a new finding that has not been examined in the literature. Our overall results show that the forecasting models using level data can perform better even when the variables are viewed as non-stationary. This finding makes a sharp contrast to the previous view and common practice that the first differenced data need to be used for forecasting models.

1.5 Conclusion

While the primary goal at the end of this model comparison is to find the model that gives the best forecasting accuracy using the univariate US monthly Energy Price index, this race is also effective in assessing the overall behavior of US energy prices.

The issue of non-stationarity was a beginning motivation for the research focus of forecasting. It was confirmed that the energy price series has structural breaks, and the Enders and Lee test gives the result of non-stationarity. However, the Diebold-Mariano test suggests that the estimations from the undifferenced series containing break are more consistent in many cases than the models forecasted with differenced series. Thus, we find evidence that forecasting models using level data can perform better even when the variable of interest is non-stationary. The intuition is clear: the usual *ARMA* models, for instance, can offer a valid cointegration regression. Other models can contain the same features. Thus, the issue of forecasting gives additional insight and a different perspective from the inference of the models with non-

stationary variables. A common conception of spurious regression could linger in the eyes of practitioners when a variable is non-stationary. Which model gives better forecasts is a different issue, and it is an empirical matter. We find that forecasting was more appropriate when using level data in many cases.

The New school (NS) and the RS models appear to be better than the Old School (OS) models for their out-of-sample performance. The main reason for this situation is the series has breaks, it is not linear, and the NS and the RS models are known to give good predictive results if the location and size of this break are well determined by the NS models. Thus, these results tell us that NS models not only increase forecasting performance but also show acceptable performance on predicting break dates and size on the energy price index. Among the NS models, the *AP* models created by Andrews Perron performed better than the *BP* models created by Bai-Perron. This gives us the information that although there is a structural break in the series, the number of breaks that will cause a change in the regime is a single, large and permanent break. Moreover, since the *ES* model outperforms the other OS models, we can conclude that the location of the break in the series is in the relatively past term of the sample period. Finally, since the Fourier model performs well with both level and differenced data and shows the capacity to output non-monotonic and non-linear trends, it could be seen as the default for this data in the future given the high likelihood of fluctuations and the uncertainty over whether to use differencing.

Overall, Enders and Liu reported in their 2009 study that NS models performed relatively well when the DGP included a single break at the beginning or the middle of the sampling period. In addition to the *AP* model, the Fourier model outperforms other OS, NS, RS models, and *RALS* methodology in all forecasting models in level and differences. The Fourier model's

superiority is more valuable when other models are used for series with structural breaks such as *TAR* and *BP* as the competitors. In addition, a Fourier model is a very low-cost model since the structure of the model is not complex like the new school methodologies; as such, this cost reduction may make the Fourier model more preferable, even though it does not have the lowest MSPE value. Our results shed light both on the characteristics of these models and US energy prices, with a unique confidence that we have selected an optimal model from the sheer number of options considered.

Table 1.1 Summary of various forecasting models

Abbreviation	Method	Abbreviation	Method	Abbreviation	Method
AP	AP model, only intercept included	*TAR*	Threshold Autoregression Model	*Rm1*	RALS forecasting with 2nd moment
$AP_{(a)}$	AP model, all variables included	*M-TAR*	Momentum Threshold Autoregression Model	*Rm2*	RALS forecasting with 3rd moment
$AP_{(p)}$	AP model, based on post-break data	$AR_{(p)}$	P_{th} Level Autoregressive Model	*Rm3*	RALS forecasting with the covariate of 2nd and 3rd moment
BP	BP model, only intercept included	D_1	First differenced AR(p) model	*Fourier*	Fourier Forecasting
$BP_{(a)}$	BP model, all variables included	D_2	Second differenced AR(p) model		
$BP_{(p)}$	Bp model, based on post-break data	*ES*	Exponential Smoothing Model		
		ARMA	Autoregressive Moving Average Model		

Table 1.2 BIC and MSPE results of all models in level

	The OS Models			The NS Models				
	BIC	MSPE		BIC	MSPE		BIC	MSPE
$AR_{(p)}$	300.72	1	*AP*	291.8	0.01	*TAR*	298.22	0.07
D_1	294.91	0.98	$AP_{(p)}$		0.02	*MTAR*	304.41	0.59
D_2	335.37	2.29	$AP_{(a)}$	290.78	15.63	*Rm1*		0.07
Pre	294.91	0.98	*BP*	290.45	0.44	*Rm2*		0.12
ES	328.63	0.56	$BP_{(p)}$		0.32	*Rm3*		0.12
ARMA	300.72	1	$BP_{(a)}$	278.69	1.26	*Fourier*	377.57	0.03

Table 1.3 The Estimated Break Dates belongs to AP and the BP methods

	Break Number	Break Dates
AP	1	2004;12
$AP_{(a)}$	1	2013;11
BP	4	1991;08-2000;09-2005;03-2007;09
$BP_{(a)}$	1	2013;11

Table 1.4 BIC and MSPE results of the models using the first difference data

	OS Models			NS Models		RS, RALS, Fourier		
	BIC	MSPE		BIC	MSPE		BIC	MSPE
$AR_{(p)}$	294.91	1	*AP*	294.91	1	*TAR*	300.29	0.07
D_1	335.37	2.32	$AP_{(p)}$		1	*MTAR*	292.03	0.07
D_2	444.85	120.8	$AP_{(a)}$	294.91	1	*Rm1*		0.083
Pre	294.91	1	*BP*	291.51	0.21	*Rm2*		0.083
ES	323.44	1.8	$BP_{(p)}$		0.31	*Rm3*		0.083
ARMA	294.91	1	$BP_{(a)}$	291.51	0.21	*Fourier*	340.37	0.067

Table 1.5 The MSPE values, and DM test results

DM Results	t-stat	p-value	In Level++	MSPE	In Difference++	MSPE
Rm1	-5.35	<.001	*AP*	0.00165	*Fourier*	0.00857
Rm2	-4.53	0.002	$AP_{(p)}$	0.00217	*TAR*	0.00867
Rm3	-4.46	0.003	*Fourier*	0.00362	*M-TAR*	0.00921
D_2	-5.717	0.024	*Rm1*	0.01	*Rm1*	0.0106
D_1	-6.85	0.006	*TAR*	0.009	*Rm2*	0.0106
$BP_{(p)}$	2.821	0.126+	*Rm2*	0.0152	*Rm3*	0.0106
BP	4.11	0.04	*Rm3*	0.0157	*BP*	0.0266
$BP_{(a)}$	6.56	0.001	$BP_{(p)}$	0.041	$BP_{(a)}$	0.0266
$AR_{(p)}$	7.11	0.004	*BP*	0.057	$BP_{(p)}$	0.03918
ARMA	7.11	0.004	*ES*	0.0726	$AR_{(p)}$	0.12743
$AP_{(p)}$	-3.02	<.001	*M-TAR*	0.076	*ARMA*	0.12743
AP	-3.12	<.001	D_1	0.12743	$AP_{(p)}$	0.12743
$AP_{(a)}$	6.62	0.08+	$AR_{(p)}$	0.12944	*AP*	0.12743
ES	-7.28	0.003	*ARMA*	0.12944	$AP_{(a)}$	0.12743
TAR	2.24	0.28+	$BP_{(a)}$	0.162	*ES*	0.2299
M-TAR	3.43	0.01	D_2	0.2962	D_1	0.2962
Fourier	-1.55	<.001	$AP_{(a)}$	2.0235	D_2	15.401

Note: (+) Indicate that the null hypothesis of equal forecasting accuracy cannot be rejected for these models. (++) Forecasting results of the models in level/difference are ordered based on their MSPE values from smallest one to biggest one on the table

Figure 1.1 Spread of monthly energy price index and Fourier

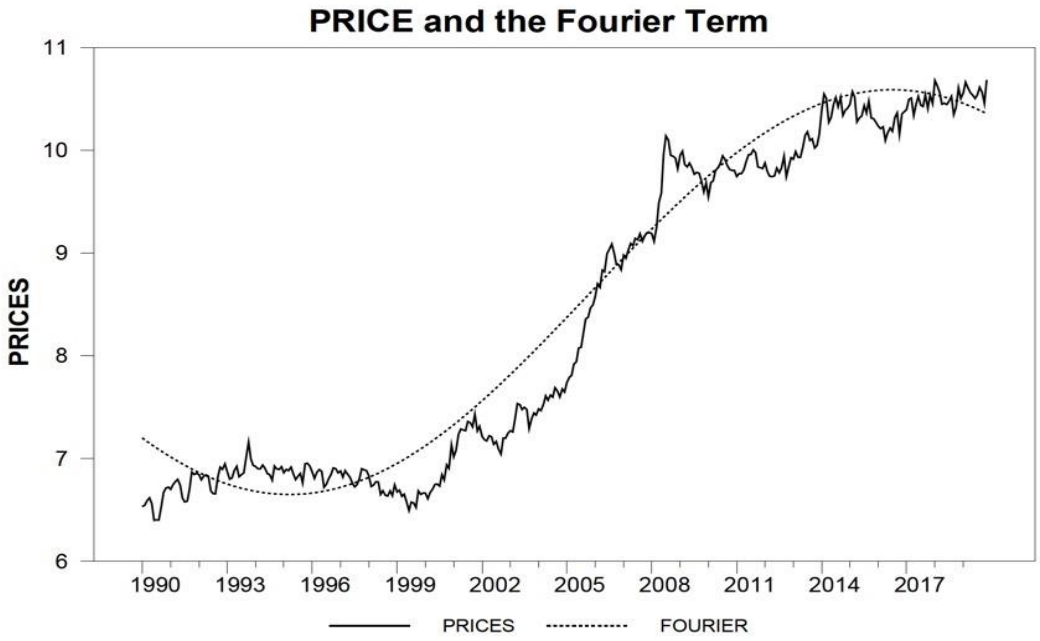


Figure 1.2 Forecasting Performance Comparison of the OS, the NS, and the RS

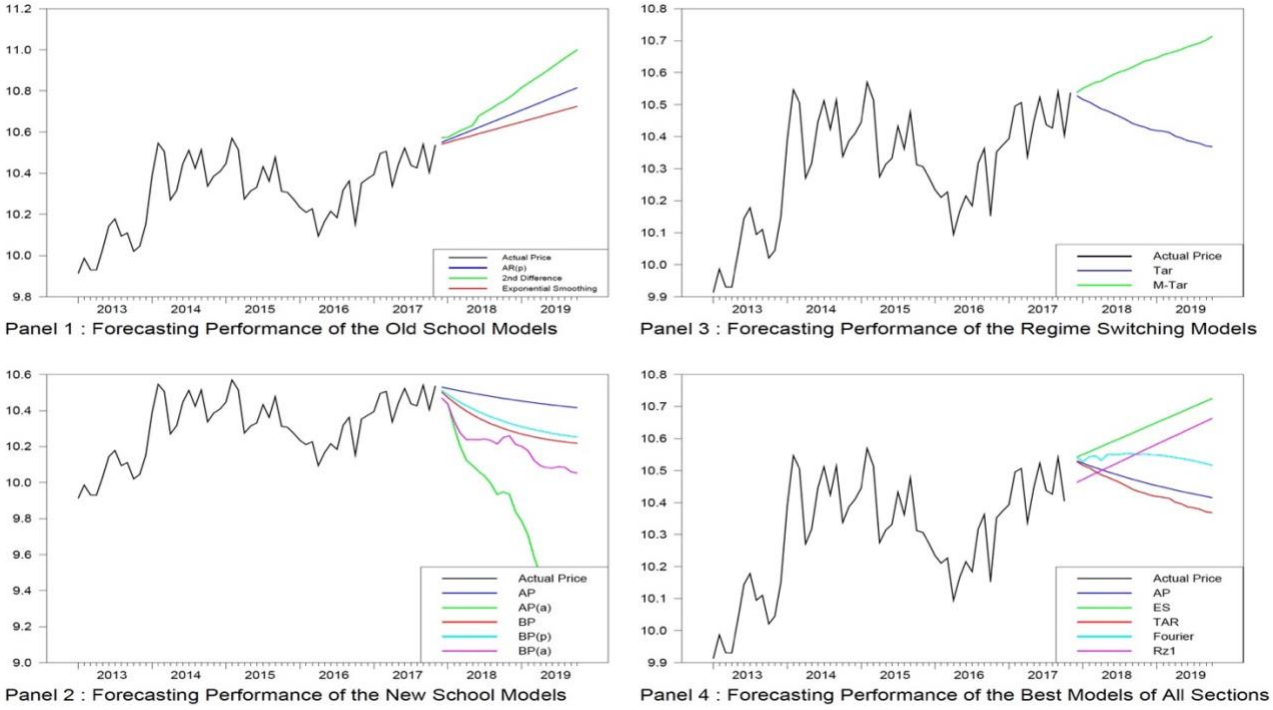


Figure 1.3 Forecasting performance comparison of the OS, the NS, and the RS in short scale

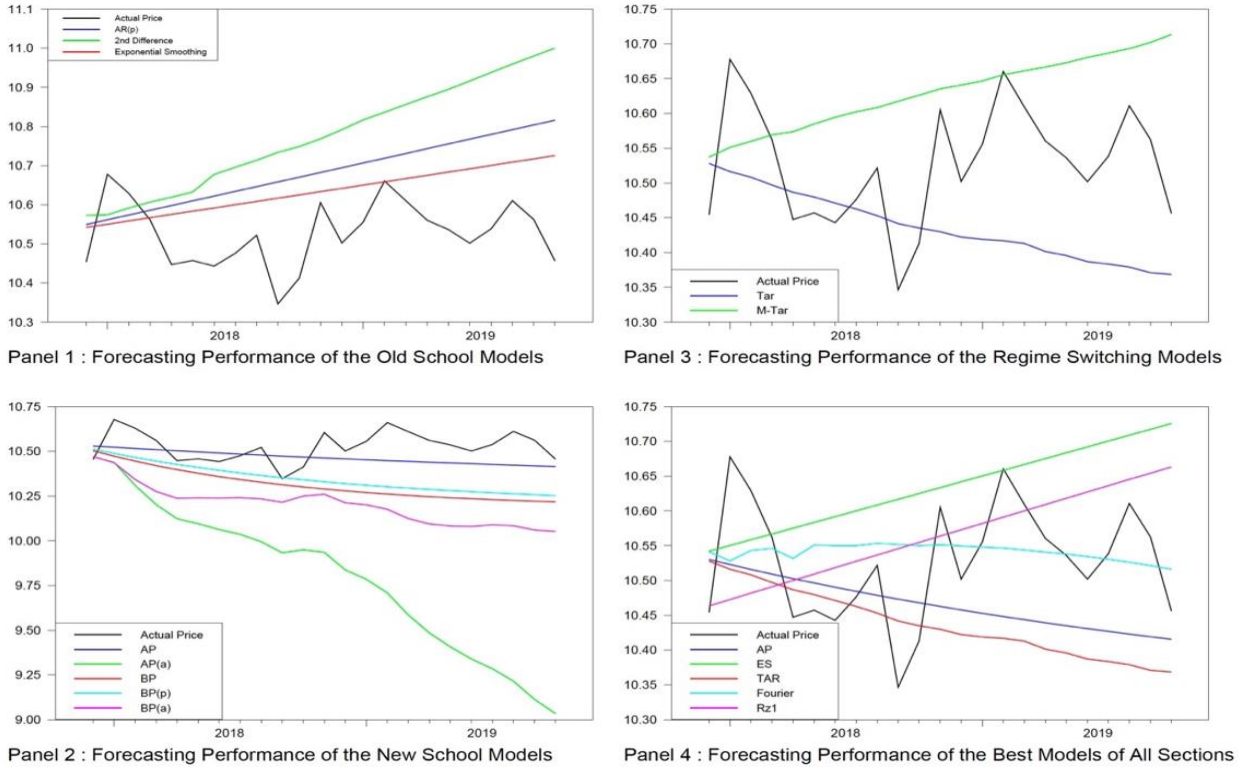


Figure 1.4 Forecasting performance comparison of the OS, the NS, and the RS in long scale

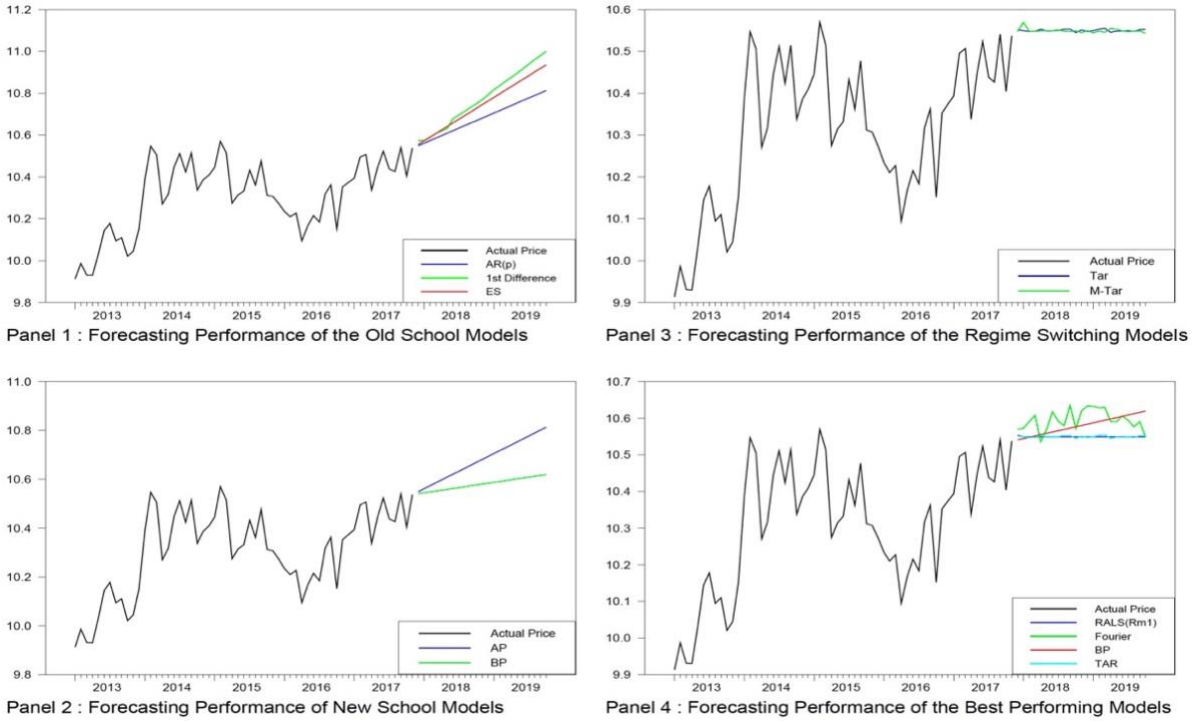


Figure 1.5 Forecasting performance comparison of the OS, the NS, and the RS in short scale

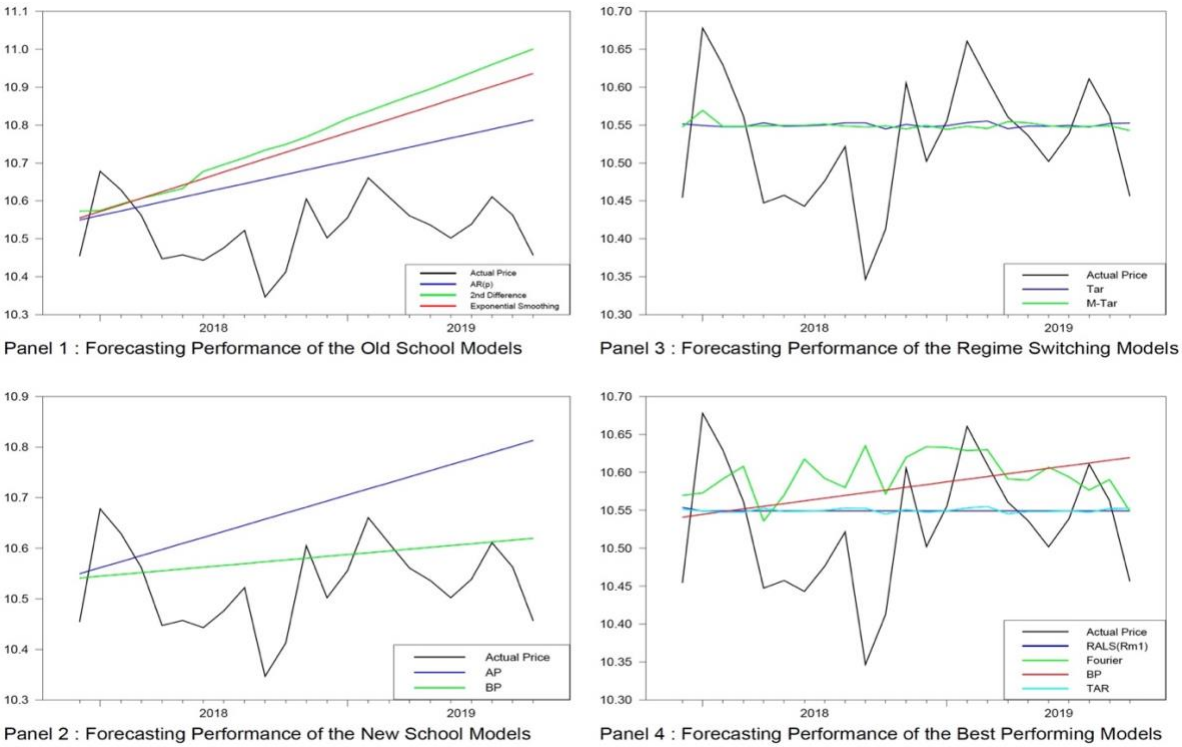
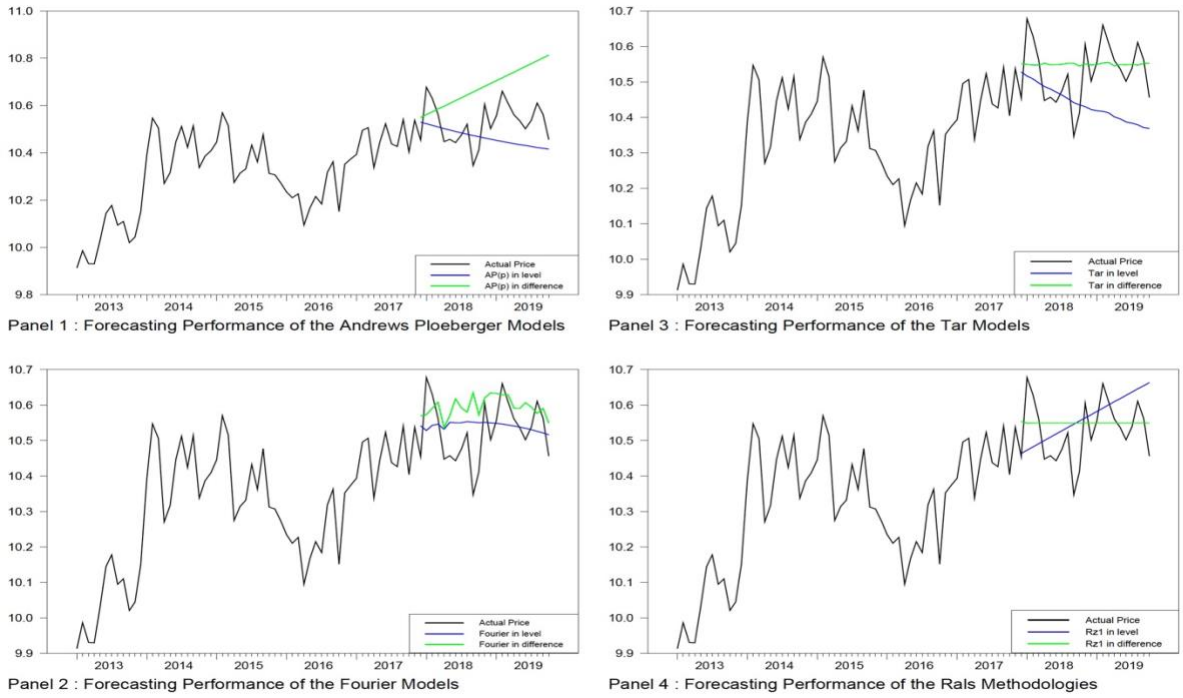


Figure 1.6 Forecasting performance comparison of the best models in level to models in difference



CHAPTER 2: INTERNATIONAL COMOVEMENTS OF ENERGY CONSUMPTIONS

2.1 Introduction

Domestic energy consumption and production habits are intrinsically linked to many important issues faced by policymakers, including national security (e.g., energy independence), infrastructure, and climate policy. Sahira and Qureshi (2007) argue that the impact and prevalence of energy in so many prominent areas of governance has made it one of the most strategic commodities, given that its generation and distribution are especially intentional. Moreover, Zhang et al. (2017) note that economic growth has necessitated increases in energy consumption and dependence across the globe. Often labeled as a contributor to the broad progression of economic growth, globalization has further created linkages by facilitating a more interconnected energy grid across countries and rising demand for international travel, both for consumers and businesses. Saidi and Hammami (2015) argue this creates a confluence of factors virtually guaranteeing energy's continued importance in the global economy in the next century. In addition, the waning abundance of energy, fossil fuel restrictions imposed by international agreements to control carbon emissions, and energy production volatility from wars and political conflicts are current stumbling blocks highlighting the importance of having a well-developed energy policy.

As noted in Pesaran et al. (1998), most studies to date on energy consumption tend to focus on looking into country-specific components whose effects are purportedly unique to that country, such as economic growth, carbon emissions, and energy prices as explanatory variable

of energy consumption. One potential caveat of previous studies is that the effects of exogenous or non-domestic factors are not fully taken into account. Global and regional trends are also likely significant determinants in the energy consumption of many countries. Yet, considering the nested nature of energy consumption, a given nation's energy consumption has a structure that is inherently susceptible to international shocks and affected by global energy policies. Therefore, as explanatory variables, we conjecture that the countries' global and regional comovements should have a highly non-trivial effect on energy consumption. Accordingly, a few recent studies look into exogenous factors. For example, Hu and Li (2021), adopting the global vector autoregressive approach on 41 countries and regions, highlight that a shock occurring in a large economy's energy growth and consumption can directly change other nations' growth/consumption behaviors. In addition, fossil fuel embargos, implemented by actors such as Arab Petroleum Exporting, have empirically caused global crises volatility in energy price and consumption. Also, Valeria Andreoni (2020) points out that non-energy events can cause shocks, like the 1979 oil shock caused by the Iranian revolution, shocks caused by the 2008 great recession, and massive price and demand dips seen in the COVID-19 pandemic. There are also shocks stemming from less dramatic circumstances. Shahbaz et al. (2018) detail that the formation of the BRICS (Brazil, South Africa, Russia, and China) entanglement has created a significant exogenous component in each country's energy consumption. These occurrences imply a likely existence of pervasive linkages in the energy market across all countries, given the reality of strengthening international interdependence.

In addition to strong global ties, regional bonds are also evident and non-negligible. With EC and GDP data from 12 Asian countries, Chang et al. (2013) reported a cross-sectional dependency between countries using four different test methods, implying that a shock in one of

the 12 Asian countries seems to be transmitted to other countries. McCormick and Neij (2009) document that Nordic countries often collaborate on energy policy and efficiency. So clearly, we cannot ignore the potential of regional factors in energy consumption.

This paper investigates how international comovements of energy consumption create global and regional effects on energy consumption. This paper attempts to contribute to the existing literature in different ways. Although most of the literature focuses on domestic components influencing consumption, we try to incorporate exogenous factors and look at a broad cross-section of countries. To formally accomplish this goal, we employ the time-varying dynamic factor model (DFM) with loadings and stochastic volatility pioneered by Del Negro and Otrok (2008). The use of Del Negro and Otrok's DFM has been well-established in several comparable settings, including public debt, international business cycles, and GDP. We apply the model to energy consumption indexes of 52 countries divided into six regions (North America, South Central America, Europe, Middle East, Africa, and the Asia Pacific) from 1965-2019. We interpret how global, regional, and idiosyncratic comovements for each country change over time and how they react to global shocks, along with their time-varying volatility. Overall, the broad trends suggest that the global factor dominates in high-income countries, while country-specific factors dominate in petrostates (fossil fuel producers) and relatively low GDP countries like Turkey, Greece, and Mexico. We also note very strong regional comovement for Nordic countries and no definitive dominating effect throughout the sample for Africa. In general, these trends are extremely time-dependent and sensitive to global shocks.

The rest of the paper structure is organized as follows; Section 2.2 presents the literature review, and Section 2.3 provides the data. Section 2.4 introduces the DFM model, while Section 2.5 provides DFM results and interpretation. The DFM model results and driving sources of the

global and regional factors are shown in Section 2.6. The final discussion and conclusion are provided in section 2.6.

2.2 Literature Review

Energy consumption and its driving factors have been discussed extensively in the literature. Although different studies have relied on various countries, methodologies, sample periods, and instruments for energy consumption (EC), most of these contributions have concentrated on the country-specific energy consumption factors. While some of these studies have examined the effects of more than one variable on energy consumption, a considerable number of them have individually examined the relationship between energy consumption and the variables of Economic growth, CO_2 emissions, GDP, and energy prices using bivariate analyses. First, we will overview the papers that focused on country-specific factors before moving on to analyses incorporating regional and global effects.

The literature seeking to isolate the effect of the domestic factors on energy consumption has identified several contributors, with economic growth (EG) emerging as arguably the most significant. Chang et al. (2013), using a Granger causality approach on 12 Asian countries from 1970-2010, found an influential relationship running through from economic growth to energy consumption. Kasman and Duman (2015), focusing on 38 Union for Mediterranean countries from 1980-2010, also confirm the causality linking economic growth and energy consumption with a short-run unidirectional panel causality running from GDP to EC. Similarly, Soytas and Sari (2003) find that for the top 10 emerging market countries, excluding China and the G-7, there is bi-directional causality between GDP and EC in Argentina and the unidirectional GDP to EC causality for Italy and Korea). By applying a dynamic OLS model in the post-Deng Xiaoping era, Fei et al. (2011), affirmed the long-run cointegrated relationship between EG and EC for

China. There are also additional studies supporting unidirectional and/or bidirectional causality between the EG and EC for India and Pakistan (Aqeel & Butt, 2001; Cheng, 1999; Shaista Alam & Butt, 2002)

Another popular domestic contributor appearing in the current literature is CO_2 emissions. Studies identifying unidirectional causality from CO_2 to energy consumption have confirmed a definitive impact of carbon emissions on EC. At the national level, the causality from CO_2 emissions to energy consumption has been addressed by Hwang and Yoo (2014) and Aviral Kumar (2011) for Indonesia using ECM and VAR methodologies orderly. Based on international contributions, the literature that focuses on CO_2 emissions on EC is controversial according to the methodologies being analyzed. For example, Lean and Symth (2010), using a panel vector error correction model, analyzed the relationships among carbon emissions, electricity consumption, and economic growth for 5 ASEAN countries between 1980-2006 years. They found that there is a statistically significant positive relationship between electricity consumption and emissions. Also, applying the Granger causality test between energy consumption and carbon emissions, they pointed out that unidirectional Granger causality from carbon emissions to electricity consumption in the short term. Moreover, Wang et al. (2016), using the panel cointegration test for 8 ASEAN countries over the period 1980 to 2009, show the long-term equilibrium relationship among urbanization, energy use, and carbon emissions. They also applied the Granger causality test to these variables. However, they figure out the unidirectional Granger causality from energy use to carbon emissions in the long term, contrary to Lean and Symth's (2010) results.

The vast majority of existing literature changes its interest in searching for the causal relationship between a univariate proxy variable and EC to testing the effects of multiple

variables on the nation's EC in the last decade. Jing Zhang et al. (2012) analyze causality between a combination of variables and energy consumption in China from 1985-2007 using a principal component analysis. The findings show that GDP, income, industrial output value, commodity exports, percentage of the rural population, price index, household numbers have been influencing energy consumption positively while the effects of commodity imports and efficiency of electricity cause to reduce EC. Later, Azlina Abd. Aziz (2013) ascertains the driving factors of EC as income, price, economic structure, and CO_2 emissions in the study employing the panel data approach to 16 developing countries over 30 years. Muhammad Azam et al. (2015) find that FDI inflows, economic growth, trade openness, and human development index variables increase EC in Indonesia, Malaysia, and Thailand with a standard log-linear regression. Liu et al. (2018) performed an empirical study for China introducing index decomposition analysis to the production decomposition analysis. This approach decomposes changes that affect energy consumption into eight different factors. The results exhibit that technological improvements related to energy savings and economic developments have a notable impact on reducing energy consumption. Also, they pointed out that regardless of geographical regions, potential economic development is the most important factor affecting energy consumption growth. For the European countries' context, the studies in the literature have been examined driving factors/EC nexus for Eastern Europe Post-Communist countries, and European Union 28 countries applying the stochastic frontier model and panel data approach; see Sinevience et al. (2017) and Zaharia et al. (2019). In these studies, the proxy variables for EC are concluded as gas emissions, GDP, population, labor growth, and healthcare expenditure.

To summarize the above discussion, the reported directions of the driving factors and their significance on EC may vary from country to country. Focusing on this observation, Wang et al. (2019) separated 186 countries into three different groups as high, upper-middle, and low-middle income groups. Then, they investigated the driving factors considering income gaps between countries adopting the Granger causality approach and the impulse response functions. The results demonstrate that the impact of urbanization on EC is only solid for high and low-middle income countries, while the causality is running from GDP to EC is valid for all groups. Furthermore, energy price appears to impact EC in high and lower-middle-income countries negatively, yet it affects EC positively in upper-middle-income countries. The results specify that the impact of urbanization on EC is only solid for high and low-middle income countries, while the causality is running from GDP to EC is valid for all groups.

We wish to replicate the structure of Wang et al. (2019) in that we include a variety of countries from around the world to control for a global effect, but instead, use a methodology that we have more confidence in. There have been a few other studies seeking to isolate non-domestic factors. Shahbaz et al. (2018) investigated the effect of globalization on EC utilizing the NARDL (Nonlinear Autoregressive Distributed Lag) bound approach in the time frame 1970-2015. The outcome shows an important globalization effect among EC countries. The impact of globalization increases EC in Brazil and South Africa while reducing EC for Russia and China. Hu and Li (2021) investigated the indirect impacts of the United States' COVID-19 economic recession on energy consumption in other nations utilizing a global VAR approach over the 1990-2013 period for 41 major countries/regions. Based on the study's results, the shrinkage in the US economy and overall decrease in travel because of the COVID-19 global crisis causes the reduction in the other developed and developing countries' energy consumption.

Additionally, the importance of the spillover effect is higher for developing countries in the short term, while for developed countries is higher in the long term. Chang et al. (2013) provide a notable look at regional effects with their investigation of cross-sectional dependency among 12 Asian countries using Lagrange Multiplier and cross-dependence tests in the study aimed to search for economic growth/energy consumption relationship. They found significant cross-sectional dependency in the group, implying joint behavior or that a shock from one of these countries extended to the others

The studies presented above and more, examining domestic factors affecting energy consumption in the current literature, are summarized in Table 2.1. These country-specific analyses examining the driving factors of energy consumption make up the bulk of the existing literature. Some papers examine the effects of domestic energy shocks on other countries' EC, where the analyses look for impacts of exogenous(global/regional) factors, but these are in the minority.

2.3 Data

As the main variable, the energy consumption data of the countries created by British Petroleum (BP) for the World Energy Statistical Review covers the years 1965 and 2019. Here, primary energy is expressed in *exajoules* (10^{18} joules). All energy sources used to generate energy from non-fossil sources are calculated on an input-equivalent basis (i.e., It is the calculation of the amount of fossil fuel required to obtain the electricity produced in a standard thermal power plant). Our panel includes 52 countries, divided into six different regions: North America, South-Central America, Europe, Middle East, Africa, and the Asia Pacific. The observations of the countries missing up to 2 observations were interpolated by taking the averages of consecutive years through interpolating.

GDP per capita growth rates were obtained from the World Databank were used in annual periods between 1965 and 2019. In addition, the variables used in the Analysis of the global factor obtained from our model, including world-scale annual energy consumption, and crude oil prices (normalized with 2020 USD), were obtained from BP's statistical review 2021 report. Other independent variables, such as gross fixed capital (GFC) formation (with respect to 2010 USD), GDP per capita growth (annual %), CO_2 emissions (metric tons per capita), and urban population growth (% of the total population), are collected in annual scale years from 1970-2019 from the world development indicator (WDI) website.

In Figure 1, the plots of the energy consumption of 52 countries are given in exajoules between 1965 and 2020. As can be seen from the chart, the consumption for every country but China, the USA, India, and Japan fluctuated between 0 and 14 EJ, while in 2019 China reached 141.7 EJ, USA 94.65 EJ, India 34.06 EJ, Japan 18.67 EJ. When we look at the empirical time-variant behavior of the countries' energy consumption, we can say that they were unilaterally affected by the price shocks of early 1973, 1979, and 2000s. They decreased their energy consumption in these periods and showed a common increase in the remaining years. These shocks correspond to notable tumultuous events. The OPEC oil crisis began in October 1973 when the Organization of Arab Petroleum Exporting Countries proclaimed an oil embargo against countries (Canada, Japan, Netherlands, United Kingdom, and the United States) believed to have supported Israel during the Yom Kippur War. Although this oil embargo was applied only against these countries, it created a global crisis because of the importance of the involved countries in the fossil fuel market. The Iran Revolution of 1979 also caused a global shock because Iran is one of the most prominent oil suppliers. For the 2000s, the global financial crisis in 2008 is sufficient to explain the common decline in energy consumption. Supporting that,

Andreoni (2020), using a Multi-scale Integrated Analysis of Societal and Ecosystem Metabolism (MuSIASEM), proved the decrease in energy consumption caused by the 2008 and 2015 financial crisis in detail with countries energy reduction rates in her study on European countries.

2.4 Methodology of The Dynamic Factor Model

We adopt the Dynamic Factor Model (DFM) to examine the global and regional effects on energy consumption. We use the energy consumption series of 52 countries to estimate unobserved comovements of energy consumption in the world. We follow the procedures illustrated in Del Negro and Otrok (2008). It is a flexible model that allows us to examine time-varying loading parameters and stochastic volatility, and it has been used to examine the comovements of some major macro variables.[7] To our knowledge, such an approach has not been adopted for examining the comovement in energy consumption.

We consider the following model specification,

$$y_{it} = \alpha_{it} + \emptyset_{it} F_t^g + \beta_{it} F_t^r + \varepsilon_{it}$$

where y_{it} is the energy consumptions of country *i* at time *t.* F_t^g indicate the global factor affecting countries energy consumption at time t, and $\emptyset_{it}$ indicates the time-varying factor loading to global factor, indicating each country's response to the global factor. F_t^r denotes the regional factors at time t in the format of a vector of dummy variables covering six different regional factors at time t while β_{it} denote time-varying factor loading to those regional factors. Finally, ε_{it} implies country-specific component.

[7] The DFM model has recently been used for some major macro variables. For example, we can find some important studies by Kose, Otrok, and Whiteman (2003) on international business cycles, Bhatt, Kishor, and Ma (2017) on long-term sovereign bond yields, Isomitdinow et al. (2020) on public debt, among others. We are grateful for the authors of Bhatt et al. (2017) who shared the Matlab codes that were initially developed for these procedures.

Here $b_{it} = [\emptyset_{it} + \beta_{it}]'$ indicate vector of time-varying loading parameters, and it is assumed that it follows random walk process without drift;

$$b_{it} = b_{i,t-1} + \sigma_{ni} n_{it}$$

Here, n_{it} $N(0,1)$ and it is independent across i. It is also assumed that global and regional factors follow AR(p) process,

$$F_t^g = \varphi_1^g F_{t-1}^g + \ldots + \varphi_p^g F_{t-p}^g + \varepsilon^{s_t^g} \upsilon_t$$

$$F_t^r = \varphi_{z,1}^r F_{z,t-1}^r + \ldots + \varphi_{z,p}^r F_{z,\ t-p}^r + \varepsilon^{s_{z,t}^r} \upsilon_{z,t}$$

where $z = 1, \ldots\ldots, z_r$ indicate the number of regions, υ_t^g $N\left(0, \sigma_{F_t^g}^2\right)$,$\upsilon_{z,t}^r$ $N\left(0, \sigma_{z,F_t^{r2}}^2\right)$, and $\sigma_{F^g}^2 = \sigma_{z,F^r}^2 = 1$. However, time-varying stochastic volatilities of global and regional factors assumed that they follow random walk process,

$$s_t^{F^g} = s_{t-1}^{F^g} + \sigma_s^{F^g} \omega_t^{F^g}$$

$$s_{z,t}^{F^r} = s_{z,t-1}^{F^r} + \sigma_z^{F^r} \omega_{z,t}^{F^r}$$

where $\omega_t^{F^g}$ $N(0,1)$, $\omega_{z,t}^{F^r}$ $N(0,1)$. For the error term (country-specific component), it follows stationary AR_{q_i} process,

$$\varepsilon_{it} = \varphi_{i,1} \varepsilon_{i,t-1} + \ldots + \varphi_{i,q_i} \varepsilon_{i,t-q_i} + \varepsilon_{\omega_{it}}^{s_{it}}$$

where ω_{it} $N(0,1)$, $s_t = s_{t-1} + \sigma_s \omega_t$. Also, the stochastic volatility of both factor variables and idiosyncratic components are orthogonal to each other.

Then we can utilize the global, regional factors and idiosyncratic components, decomposing the variance of energy consumption,

$$Var(y_{it}) = \emptyset_{it}^2 \, Var\left(F_t^g\right) + \beta_{it}^2 \, Var(F_t^r) + Var(\varepsilon_{it})$$

Then, the variance contributions can be estimated as $\frac{\emptyset_{it}^2 \, Var(F_t^g)}{Var(y_{it})}$ for the global factor, $\frac{\beta_{it}^2 \, Var(F_t^r)}{Var(y_{it})}$ for the regional factor, and $\frac{Var(\varepsilon_{it})}{Var(y_{it})}$ for the idiosyncratic factor, respectively. These are the key results to show. The interpretations of those estimated variance contributions for each country and six different regions are presented in Figures 2.3a-3f and Section 2.5 (variance decomposition results).

2.5 Results from Dynamic Factor Model

Estimated Global and Regional Factors

To estimate the DFM, we use the first differences of the energy consumption variable. We obtain the estimates of the global and regional factors using the differenced data from the median values of the Bayesian draws based on the Markov Chain Monte Carlo (MCMC) procedure. Then, we use the cumulative sum to obtain the estimated factors. The estimated global and regional factors are presented in Figure 2.2-2.3. It shows the plots of estimated global (top) and regional (bottom) factors together with their confidence intervals (5th and 95th percentiles) and their medians for the period from 1965 to 2019. The estimated global factor shows a fluctuating path throughout the sample period, and it follows a downward trend. It seems that three major shocks occurred on the global factor during the sampling period, twice between 1970 and 1980 and once towards the end of the 2000s. These shocks correspond with the 1973 OPEC crisis, the 1979 Iranian revolution, and the great financial crisis of 2008, and they might have caused sharp swings in the magnitude of the global factor. However, the global factor tends to revert its pre-shock levels.

We next examine the estimated regional factors. We observe a sharp decline in the Asia Pacific regional factor. It might have been caused by the Asian financial crisis that started first in the Asia-pacific regions in 1997 until the end of the 1990s. It seems that the regional factors of the South-Central America region were affected by the global crisis in 2008. We also observe that the confidence interval for North America, Europe, and the Asia Pacific regions is very tight, which means that the predictive performance is relatively high in these regions.

Variance Decomposition Results

One of the key results of the DFM is associated with those from variance decomposition analysis. It gives the results on the relative contributions of each of the global, regional, and country-specific factors to the variation of energy consumption of each country. If the share of the global factor is high in a country, it implies that the energy consumption of that country is more explained by the global factor rather than the regional or country-specific factors. Then, the country is more exposed to global shocks regarding energy consumption. The results on the relative shares of global, regional, and country-specific factors are given in Figures 2.4 – 2.9. For convenience, we present the results of each country under the corresponding regions where a country belongs. The blue areas represent the contribution of the global factor to the energy consumption of each country, the green and yellow areas show the contribution of the regional and country-specific factors, respectively. At first glance, the regions have relatively unique characteristics. While the country-specific factor is more dominant for the countries in Asia-pacific and the Middle East, except for Japan, the global factor seems to be the most vital determinant in energy consumption for the European and North American regions. For the South and Central America region, we can say that mostly the regional factor is more responsible than other factors.

The effect of the global factor decreased rapidly at the beginning of the sample period in Asia Pacific countries and decreased to negligible levels by the 2000s. Especially for China, India, Singapore, and Thailand, the country-specific factor has been the most determining factor in energy consumption (increased up to 95 percent in 2019). In general, the economic growth that came with the increase in over-manufacturing in Asian countries since the 1970s has separated Asian countries from global factors. In parallel with these findings, it has been suggested in previous studies that there is a one-way causality relationship between energy consumption and GDP for India, Thailand, and Singapore. China has also implemented many domestic factors such as increasing energy efficiency by taking many energy-saving measures along with economic growth and increasing energy intensity, which shows why the country-specific factor is most effective.

For the countries in Europe, we see that the share of the global factor on countries' energy consumption declined through the end of the 90s, then, the global factor effect, which increased until 2007, decreased slightly again with the great global crisis in 2007. In addition, while Belgium, France, Italy, Netherlands, Portugal, Spain, and the United Kingdom are almost entirely dominated by the global factor (variance contributions over 90 percent), the regional factor is more important in Denmark, Iceland, and Norway. On the other hand, Sweden, Greece, Switzerland, and Turkey are influenced by country-specific factors in energy consumption from the 1965's. The heterogeneity in this region can be explained by noting that domestic factor effects are more pronounced in countries with much lower GDP, while countries with relatively high GDP are more exposed to global factors. Moreover, renewable energy policies agreed upon among Nordic countries and their effect on the energy consumption of these countries can explain why regional factors are high in these countries.

The effect of the global factor in energy consumption in the North America region, where the global factor is the most important, has reached the level of 89%. The fact that the USA and Canada have become energy importing countries since 1979 explains the high global factor effects in these countries. For the same reason, the country-specific factor is the most effective in the energy consumption of oil-producing countries such as Iran, Saudi Arabia, Egypt, and Venezuela. Another important finding is that in countries in the European region, global and regional factors are superior to country-specific factors in energy consumption in many countries. While the global factor is the most important in countries with strong economies such as the United Kingdom, France, Spain, Italy, etc., in Nordic countries such as Denmark, Finland, Iceland, and Norway, the effect of the regional factors is superior to global and country-specific factors. Renewable energy agreements between Nordic countries may be the reason behind regional factor dominance. In addition, it has been observed that country-specific factors dominate energy consumption in Asian Pacific countries (except for Japan). The increased overproduction in Asian countries since the 1970s may have led these countries to move away from global factors and increase the importance of domestic factors in energy consumption.

Stochastic Volatility Results

In Figure 2.10, stochastic volatility of global and regional factors of energy consumption is shown with 90% confidence intervals. The volatility of the global factor experienced a sharp decline between 1970 and 1980 and remained stable until the great financial crisis in 2008, then the modest upward trend we observed in volatility stopped in 2010 and remained stable until 2020. When we examine the volatility of the regional factor of the European region, we see that the behavior in the regional volatility is quite similar to the volatility of the global factor but less aggressive. The interpretation here is that the developments in the global factor have a great

impact on the regional factor for the European region as well. For North America and Middle East regions, volatility in the regional factor remains stable throughout the sample period, while for Africa and South-Central America, volatility decreases slightly over the sample period. For the Asia Pacific region, the volatility, which decreased towards the end of the 1990s, remained stable between 2000 and 2020, showing a slight increase, which is thought to be because of the Asian crisis at the end of the 1990s.

Cross-Sectional Dispersion in Volatility

In Figure 2.11, we present the cross-sectional distribution of volatility of all countries over time. The Dynamic Factor model decomposes volatility into two components, one due to the global factor and the other due to country-specific fluctuations. This figure helps us to understand whether the decrease in volatility stems from global or country-specific factors. Variance explained by global factors experiences a persistent decrease in the first ten years from the 1970s, and then it stabilizes until a modest increase in 2007. However, when the distribution of the variance explained by the country-specific factors is examined, a slight increase is observed, especially after the 1990s. However, this increase due to individual components in the dispersion is rather insignificant when compared to the decrease due to the global factor. Thus, the decrease in cross-sectional dispersion in volatility was driven by global components. This indicates that the impact of global shocks on energy consumption is becoming more and more similar between countries.

Average Cross-Country Correlations

The average cross-country correlations obtained by dividing all countries into both common and six different regions, along with confidence intervals over the sampling period, are

shown in Figure 2.12. Although the cross-country correlations of each region follow different levels and fluctuations, they have caught common increasing and decreasing trends between certain periods. When the cross-country correlations of all regions and all countries together are examined, the highest correlation was seen in the North America region as expected (around 0.6), followed by the correlations of the European region (between 0.3 and 0.5) and the global factor (between 0.1 and 0.15), respectively. For the Middle East, Africa, and Asia Pacific regions, although there were fluctuations throughout the sample period, no statistically significant correlation was observed. Surprisingly, cross-country correlations of the European region and the global factor give very similar responses to shocks that occur in the correlation over time, implying that changes in European countries are mainly related to global comovement. The cross-country correlations, which decreased from the 1970s to the 1990s for all countries in the data set (around 0.05) and the Europe region (around 0.1), remained almost constant from the beginning of the 90s to the crisis in 2008. Then the correlations increased until 2010 after the crisis in 2008, later from 2010 to the end of the sampling period has been gradually decreasing. These results also show that while the inter-country correlations are higher in regions where the global factor is important, the correlations can be neglected in regions where the country-specific factor is dominated by the others. This is in line with the variance decomposition results above, which show the contribution of global and country-specific factors to energy consumption.

Loadings to Common and Regional Factors

In Appendix Figure 1, we present the time-varying loadings to the common factor (top) and regional factor (bottom) for all countries with their confidence intervals. It is difficult to identify a common pattern that changes over time between the loadings to the global factor and loading to

the regional factor of countries. However, we can use these graphs to surmise the degree to which countries and regions are subject to the global factor. While the regions most affected by the global factor are North America and Europe, this rate is almost non-existent in other regional countries, or the rates are relatively low. The results show that the US is the country most exposed to the global factor, both within the North America region and among all countries, followed by Japan, and countries with relatively higher GDP than the others in the European region (France, Italy, United Kingdom, Spain), respectively. Among these countries, exposure rates to the global factor remain constant or fluctuate modestly over the sample period. The loadings to regional factor rates can be said to be quite low when compared to exposure to the global factor, even in most countries, this rate can be assumed to be 0 statistically. But when the countries that are most exposed to the regional factor are examined, these countries are matched with the countries that are most exposed to the global factor. Overall, we infer from these results that the countries that are relatively more exposed to the global energy consumption factor than the others are more exposed to the regional factor.

2.6 Analysis of DFM Result and Driving Sources of Global Factor

Estimated Factors and World Energy Consumption

As we saw in the above sections, fluctuations in the energy consumption of countries can mostly be explained by fluctuations in the global factor. Figure 2.13 presents the time-varying plots of global-scale energy consumption data with the global, regional factors obtained from DFM with comparative graphs. The overall trend of the global factors is similar to the real-world energy consumption index, including the timing of shocks. In addition, the world energy consumption data and regional factors that change over time were compared, and it was seen that

the trends gave very similar reactions at the times of the breakouts for Europe, North America, and South-Central America regions. These results also support the finding from the DFM model that there are important regional factors in these regions.

Moreover, real energy consumption data for each country and time-varying 5-year window rolling correlations of Global and Regional factors are presented in Figure 2.14. Based on these graphs we see a strong correlation around by 0.9 between the global factor and energy consumption of some countries like the USA, United Kingdom, France, Japan. Also, there is a robust correlation around 0.8 between the regional factor and energy consumption of some countries like Colombia, Egypt, Iran, Denmark. Even these correlations are quite volatile and fluctuate throughout the time period, these observations confirm that for some countries, global and regional factors are more correlated with these countries' time-varying energy consumption.

Partial Correlation Coefficients Between the Global Factor and Driving Forces

The question of interest is which variables might be major determinants of the global factor. This is an issue related to the driving forces of the global factor of energy consumption. Thus, we examine the relationship between the global factor and global macro variables. We consider the following variables: the world level GDP, the gross fixed capital formation (GFC), oil prices, CO_2 emission, and urbanization rates. We expect easily that oil prices affect energy consumption; see Sarwar et al. (2017). CO_2 emission and urbanization rates are the two most common variables for determinants of energy consumption variables; see Lean and Smyth (2010) and Wang et al. (2016-2019). The sources of these data are discussed in the previous section.

We first look at the partial correlation of these driving force variables and the global factor of energy consumption. For this, the following regression is considered,

$$y_t^* = \sum_{i=1}^{v} \delta_i n_{it}^* \ + \ \varepsilon_t$$

where y_t^* is the standardized global factor of energy consumption, n_{it}^* denotes the standardized variables of the driving source variables explained above. Here δ_i indicates the partial correlations of each variable with the global factor. Table 2.2 reports these results. We see strong correlations from the GDP growth rate and urbanization growth rate, while the coefficients of CO_2 emissions, GFC, and oil price are not statistically significant. On the other hand, when we use real-world energy consumption as the dependent variable, only the GDP growth rate is statistically significant.

Panel Data Analysis

We now examine the determinants of energy consumption of each country using panel data models. We use the energy consumption data for each country as a dependent variable. We use global macro variables discussed above as independent variables. Moreover, we add the country-level GDP growth rates to the set of regressors. In Table 2.3, we present the estimation results using a battery of different estimation methods: ordinary least squares (OLS), fixed effects estimator (FE), fixed effects estimator with factors (FE-F), and the iterative FE (IFE) models of Bai (2009). The FE-F estimator uses the estimated factors as additional control variables. Since the factors are estimated exogenously from the dynamic factor models, this procedure can be subject to the issue of the generated regressor. However, the IFE estimator is free of this issue since the factors are jointly estimated with the parameters in the model, using principal component analysis (PCA). Instead of using the estimated global and regional factors, IFE includes the principal components to capture cross-correlations. Then, using an iterative procedure, the parameters in panel models and the PCA are jointly estimated.

The results in Table 2.3 support the results from the partial correlation analysis. The coefficients of growth rates of GDP and urbanization indexes are highly significant. These variables continue to be important driving forces in the panel data analysis. The OLS, FE, FM, and IFE estimators identify these two variables as significant determinants. The coefficient of country-level GDP growth rates is also significant. We find that the coefficient of growth rates of GFC and oil price indexes are significant from the IFE estimator, which accounts for cross-correlations using the PCA.

2.7 Discussion and Conclusion

This article has examined the significance of exogenous factors affecting energy consumption globally. We have estimated the dynamic factor model (DFM) using data of 52 countries divided into six regions from 1965 to 2019. The DFM estimates the global, regional, and country-specific factors in determining the energy consumption in the world. We find significant evidence of the existence of comovements in energy consumption. They are the exogenous factors affecting energy consumption, and they change over time. These factors are affected by international shocks over time.

The findings show that the ratio of global factors affecting energy consumption worldwide is at 73%, which is a highly significant figure. The ratio of regional factors is 5%, 6%, 5% in North America, South Central America, and Asia, respectively. The ratio of the regional factor is 18% in the European region. The global factor has a declining trend. It is significantly affected by international shocks such as the 1973 OPEC crisis, the 1979 Iranian revolution, the 1997 Asian financial crisis, and the 2008 Great Recession or the renewable energy agreements signed among Nordic countries. While previous studies focused on domestic factors in examining the energy consumption of countries, this study provides significant

evidence of the importance of global comovements. Perhaps, no previous studies addressed this matter. That is the main point by which the paper attempts to the literature. In the future, more papers should focus on attaining more precise measurement of exogenous factors because of the international network by which energy consumption is dictated for so many countries

Table 2.1 Summary of the Studies Related to Energy Consumption

Reference	Period	Level/Country	Method	Major Drivers
Muhammad Azam et al. (2015)	1980-2021	International (Malaysia, India, Thailand)	LS	FDI, EG, trade openness, etc.
Jiang Zhang et al. (2012)	2006-2015	National/China	Principal Component Analysis	GDP, I, commodity exports, Industrial output value etc.
Xiao Liu et al. (2018)	2007-2012	National/China	Production Decomposition Analysis	EG, technological change effects.
Tsangyao Chang et al. (2013)	1970-2010	International (12 Asian Countries)	Panel Causality Analysis	EG
Alina Zaharia et al. (2019)	1995-2014	International (EU28 Countries)	PDA, and Bibliometrics Analysis	GE, GDP, labor growth, healthcare expenditure, etc.
Lina Sineviciene et al. (2017)	1996-2013	International (EEPCC)	Comparative Analysis, and Stochastic Frontier Approach	GDP growth, EP*
Esmaeili, and Rafei (2021)	1992-2017	National/Iran	SVAR, and TVP-VAR models	Inflation shocks, GDP*, oil revenues*
Azlina Abd. Aziz et al. (2013)	1978-2003	International (16 Developing Countries)	PDA	I, EP, economic structure, CE
Qiang Wang et al. (2019)	1980-2015	International	Causality Test Approach, and IRF Analysis	EP, URN, GDP
Saidi, and Hammami (2015)	1990-2012	International	GMM	CE, EG
Ozturk, and Acaravci (2010)	1968-2005	National/Turkey	ARDL, and BTA	EG, GE, the employment ratio
Wang et al. (2016)	1980-2019	International (ASEAN Countries)	Pedroni Panel Cointegration Test	URN, CE*
Alam, and Butt (2002)	1960-1998	National/ Pakistan	Johansen and Juselius Technique	EG

Benjamin S. Cheng (1999)	1952-1995	National / India	HGCM	EG
Fei et al. (2011)	1985-2007	National / China	Dynamic OLS	GDP
Aqeel, and Butt (2001)	1955-1996	National/Pakistan	HGCM	EG
Chan, and Lee (1996)	1953-1993	National/China	VCM	EP, I, Share of industry output
Lean, and Smyth (2010)	1980-2006	International (ASEAN Countries)	PVECM	CE
Perry Sadorsky (2011)	1996-2006	International (Central, and Eastern European Countries)	DPDM	FD
Shahbaz, and Lean (2011)	1971-2008	National/Tunisia	ARDL	FD, IND
Islam et al. (2013)	1971-2008	National/Malaysia	ARDL	EG, FD
Esseghir, and Khouni (2014)	1980-2010	International/ (38 UFM Countries)	ECM	EG
Kasman and Duman (2014)	1992-2010	International (New UE Member and Candidate Countries)	PCM, PCT	GDP
Soytas, and Sari (2003)	1950-1992	International (Top 10 emerging markets, except for China and G7 countries	DF, ADF, PP	GDP
Hwang, and Yoo (2012)	1965-2006	National/ Indonesia	ECM	CE
Aviral Kumar (2011)	1971-2007	National/ India	VAR	CE
Liu et al. (2014)	1992-2007	International/ (China and USA)	SDA	International trade, domestic investment
Fan, and Xia (2012)	1987-2007	National/ China	RAS	Industry structure, technology improvements
Cui et al. (2018)	1990-2015	National/ China	VAR	P, GDP, URN

[1]Where abbreviations are defined as VAR: Vector Autoregression, SVAR: Structural VAR, TVP-VAR: Time-Varying Parameters VAR, GMM: Generalized Methods of Moments, ARDL: Autoregressive Distributed Lag Bounds, HGCM: Hsiao's Versions of Granger Causality Method, VCM: Vector Correction Model, PVECM: Panel Vector Error Correction Model, DPDM: Dynamic Panel Demand Models, IRF: Impulse Response Function, BTA: Bounds Testing Approach, PDA: Panel Data Analysis, ECM: Error Correction Models, PCM: Panel Cointegration Methods, PCT: Panel Causality Test, PP: Philips Perron, DF: Dickey-Fuller, ADF: Augmented Dickey-Fuller, SDA: Structural Decomposition Analysis

[2]Where driving factors abbreviations are defined as FDI: Foreign Direct Investment, P: Population or population growth rate, GDP: Gross Domestic Product, EG: Economic Growth, EP: Energy Price, GE: Gas Emissions, CE: CO2 emissions, URN: Urbanization, I: Income, FD: Financial Development, IND: Industrialization

[3](*) indicator means the effect of the variables is tested but resulted in no statistically significant relationship between variables and energy consumption.

Table 2.2 Partial Correlations Between the Global Factor/World Energy Consumption

VARIABLES	Standardized values of (Global_e)	Standardized values of (W_EC_gw)
Standardized values of (GDP_gw)	0.642**	0.831***
	(0.27)	(0.18)
Standardized values of (CO2_Em_gw)	-0.058	0.234
	(0.10)	(0.15)
Standardized values of (GFC_gw)	0.030	-0.142
	(0.26)	(0.13)
Standardized values of (URN_gw)	-0.179*	-0.060
	(0.09)	(0.07)
Standardized values of (Oil_P_gw)	-0.066	0.076
	(0.09)	(0.07)
Observations	47	47
R-squared	0.464	0.804

Robust standard errors in parentheses *** p < 0:01, ** p < 0:05, * p < 0:1
Where variables are defined as GDP_gw : GDP per capita growth rate, CO2_Em_gw : Co2 emissions (metric tons per capita) growth rate, GFC_gw : Gross fixed capital formation (with respect to 2010 USD) growth rate, URN_gw : urban population growth (% of the total population), Global_e: Estimated global factor of energy consumption, W_EC_gw : World energy consumption growth rate

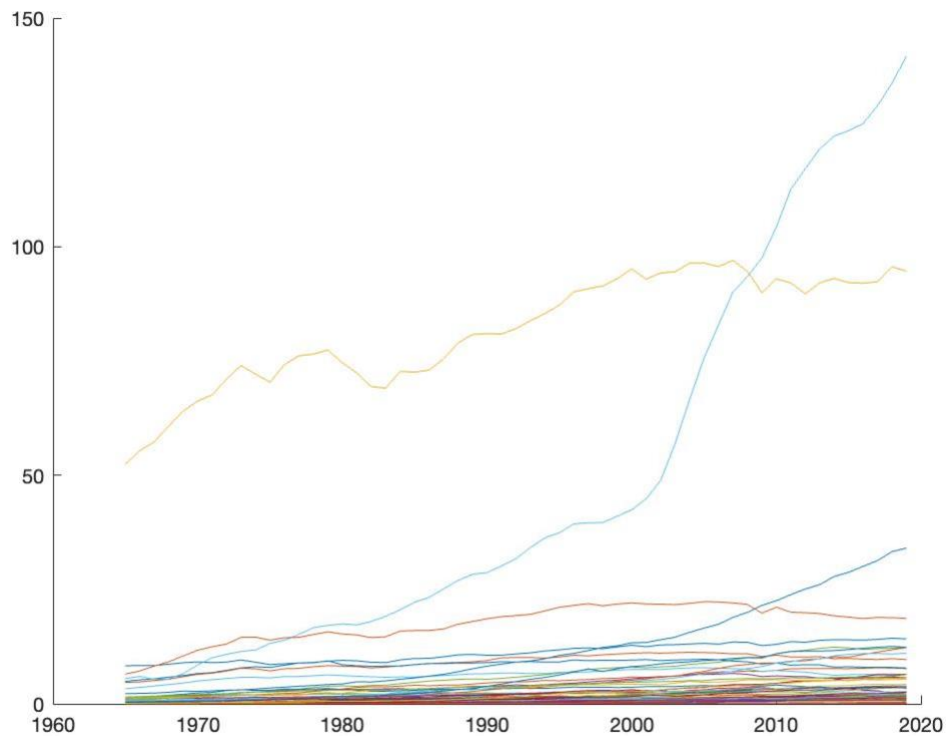
Table 2.3 Panel Regression Results

VARIABLES	OLS	FE	FE_F	IFE
Country-level GDP_gw	0.004***	0.004***	0.004***	0.004***
	(0.00)	(0.00)	(0.00)	(0.00)
GDP_gw	0.007***	0.006**	0.003*	0.018***
	(0.00)	(0.00)	(0.00)	(0.01)
CO2_Em_gw	-0.008	-0.008	0.002	-0.003
	(0.01)	(0.01)	(0.01)	(0.02)
GFC_gw	-0.148	-0.147	0.038	-0.974**
	(0.19)	(0.20)	(0.17)	(0.47)
URN_gw	-5.011**	-4.986*	-7.762***	-16.303***
	(2.35)	(2.57)	(1.93)	(5.37)
Oil_P_gw	-0.006	-0.006	-0.004	-0.058***
	(0.01)	(0.01)	(0.01)	(0.02)
Constant	0.015	0.032***	0.045***	0.019
	(0.01)	(0.01)	(0.01)	(0.02)
Observations	2,350	2,350	2,350	2,3
R-squared	0.166	0.152		
Number of ids		50		

Robust standard errors in parentheses *** p < 0:01, ** p < 0:05, * p < 0:1

Where variables are defined as GDP_gw : GDP per capita growth rate, CO2_Em_gw : Co2 emissions (metric tons per capita) growth rate, GFC_gw : Gross fixed capital formation (with respect to 2010 USD) growth rate, URN_gw *:* urban population growth (% of the total population), Where the methodologies are defined as OLS : Ordinary Least Squares , FE : Fixed Effects Estimator, FE-F : Fixed Effects Estimator with Factors , and IFE : Iterative FE models of Bai (2009)

Figure 2.1 Plots of Energy Consumption in Each Country



The plots of the energy consumption of 52 countries are given in exajoules (10^{18} joules) between 1965 and 2020. The consumption rates for every country but China, the USA, India, and Japan fluctuated between 0 and 14 EJ, while in 2019 China reached 141.7 EJ, USA 94.65 EJ, India 34.06 EJ, Japan 18.67 EJ.

Figure 2.2 Estimated Global Factor

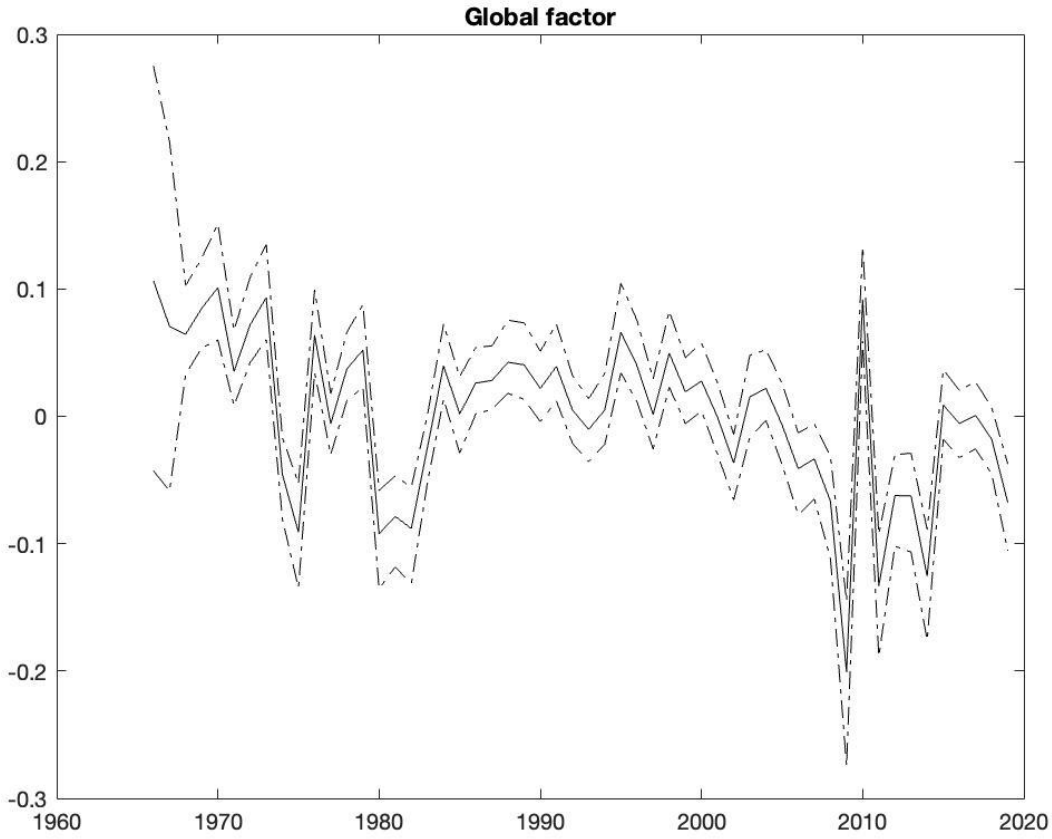


Figure 2.3 Estimated Regional Factors

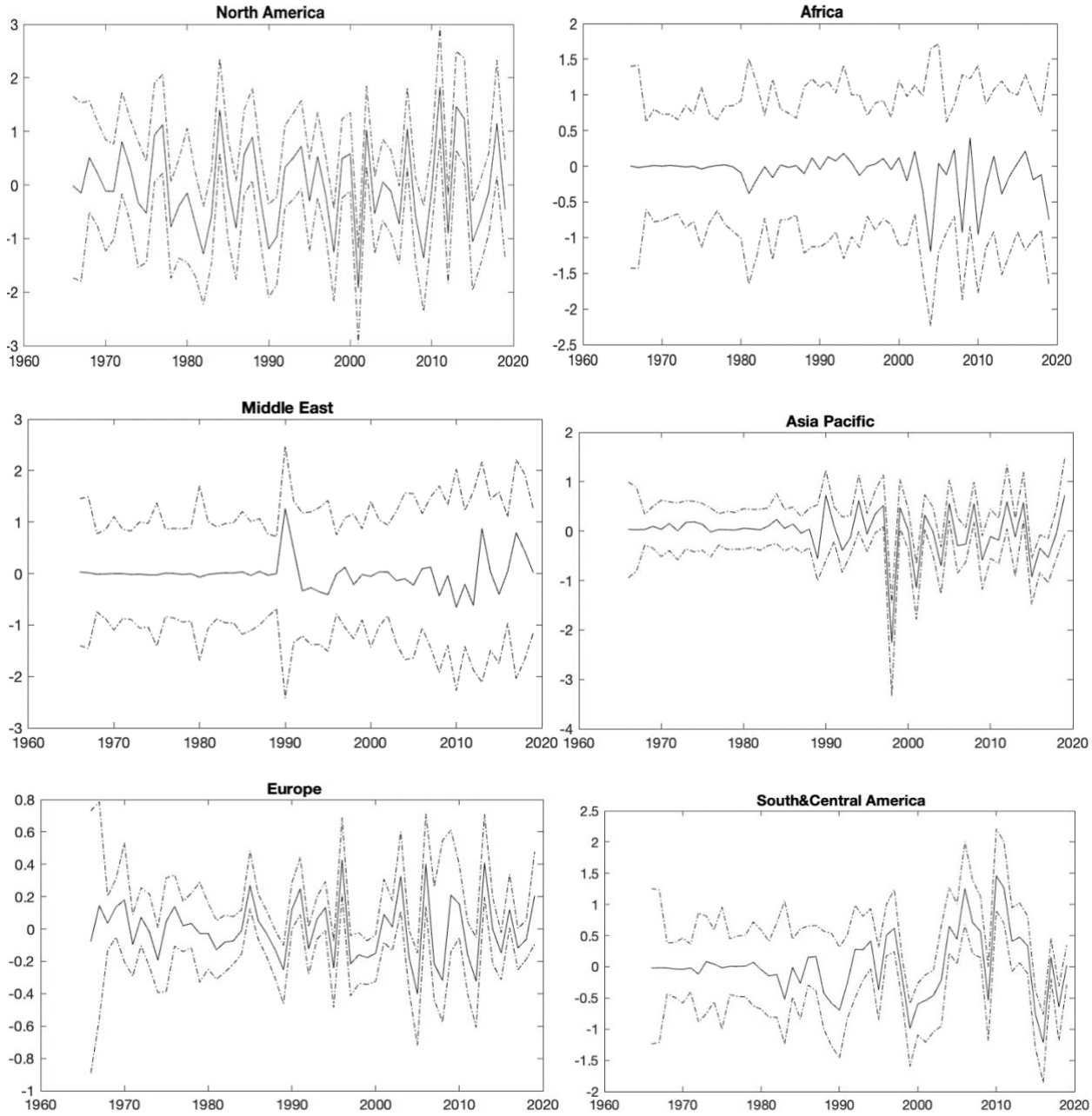


Figure 2.4 Variance decomposition in the Asia Pacific region

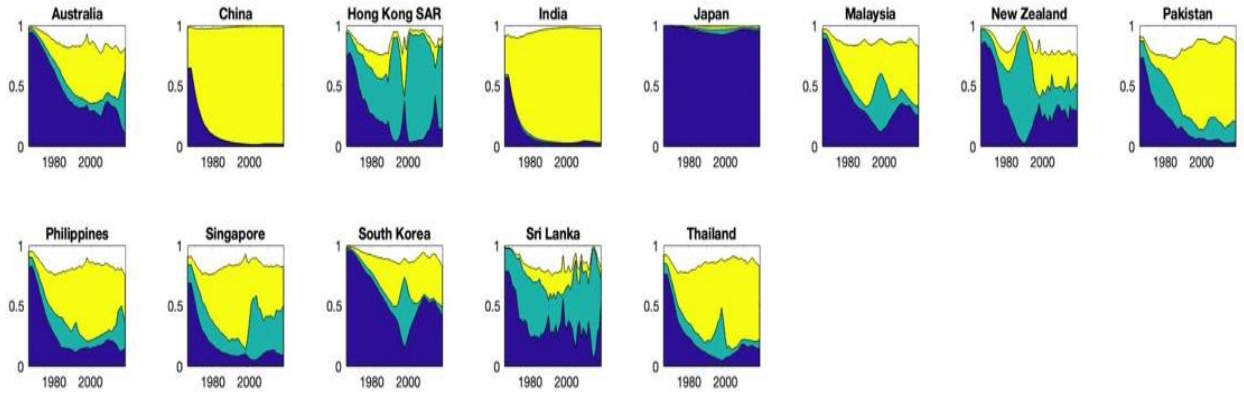


Figure 2.5 Variance decomposition in the Europe region

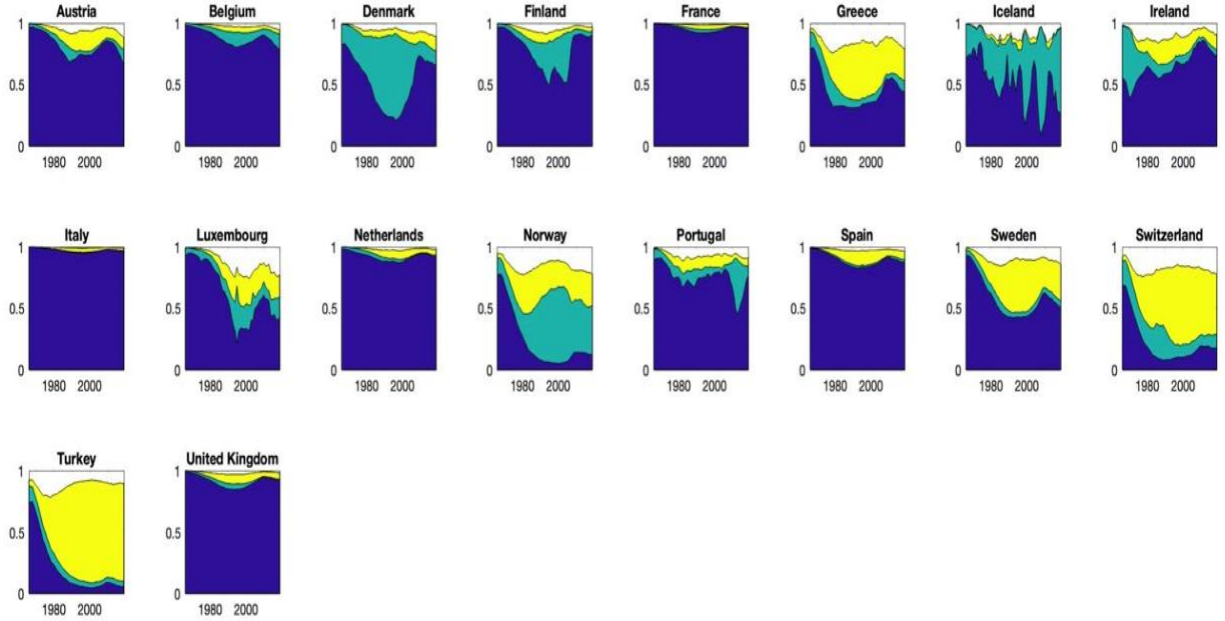


Figure 2.6 Variance decomposition in the Africa region

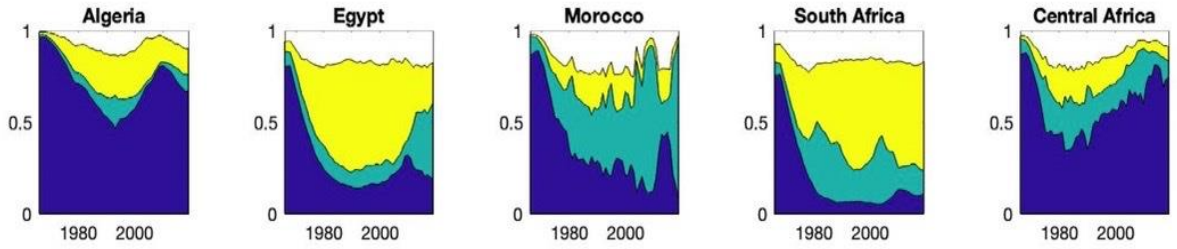


Figure 2.7 Variance decomposition in the Middle East region

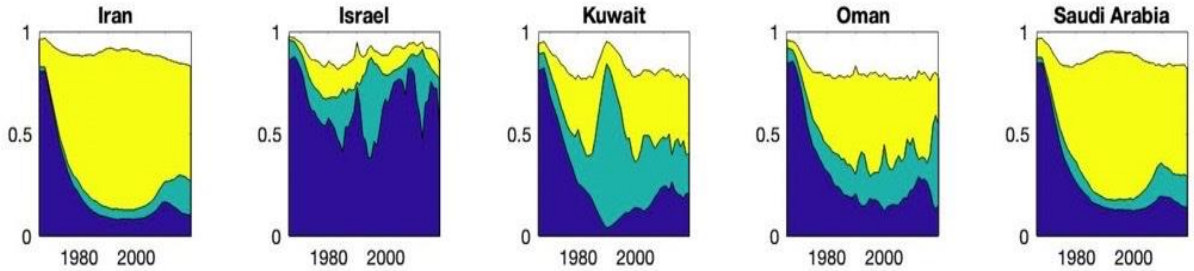


Figure 2.8 Variance decomposition in the North America region

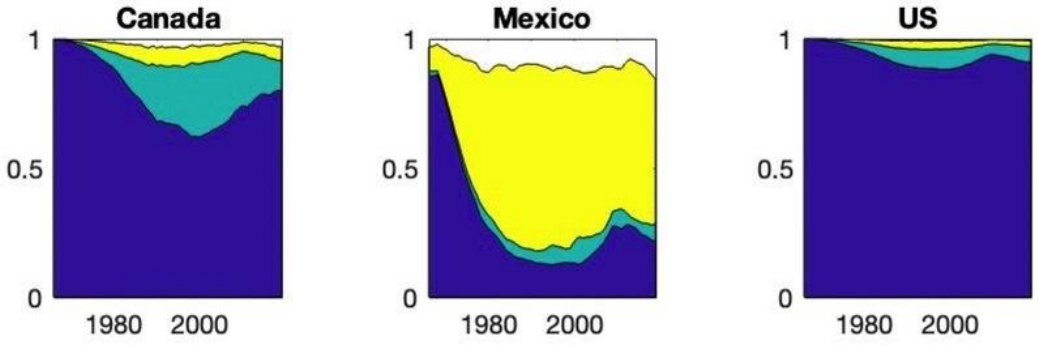


Figure 2.9 Variance decomposition in the South and Central America region

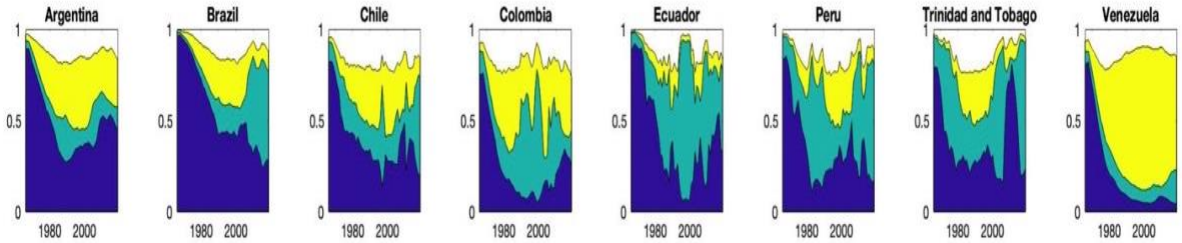


Figure 2.10 Stochastic volatility of Global and Regional factors

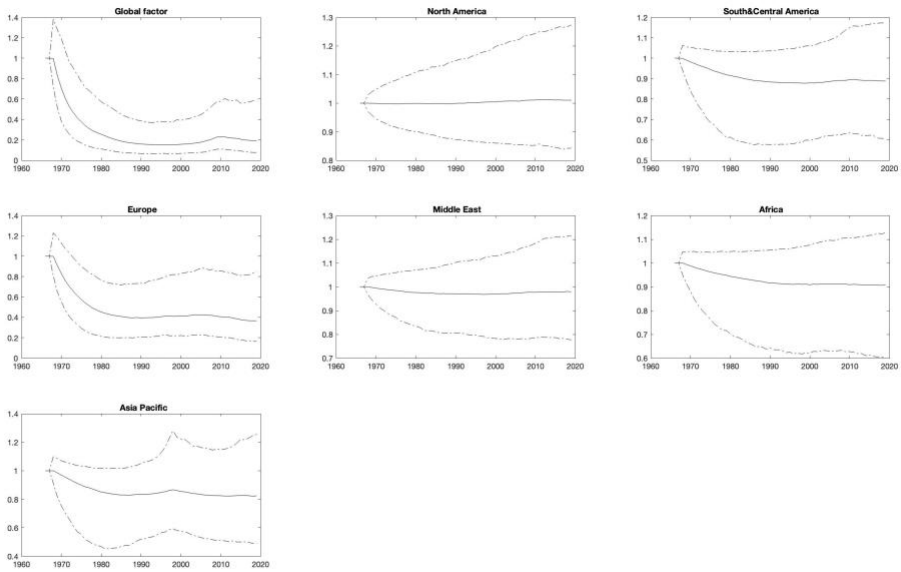


Figure 2.11 Cross-Sectional Dispersion in Volatility

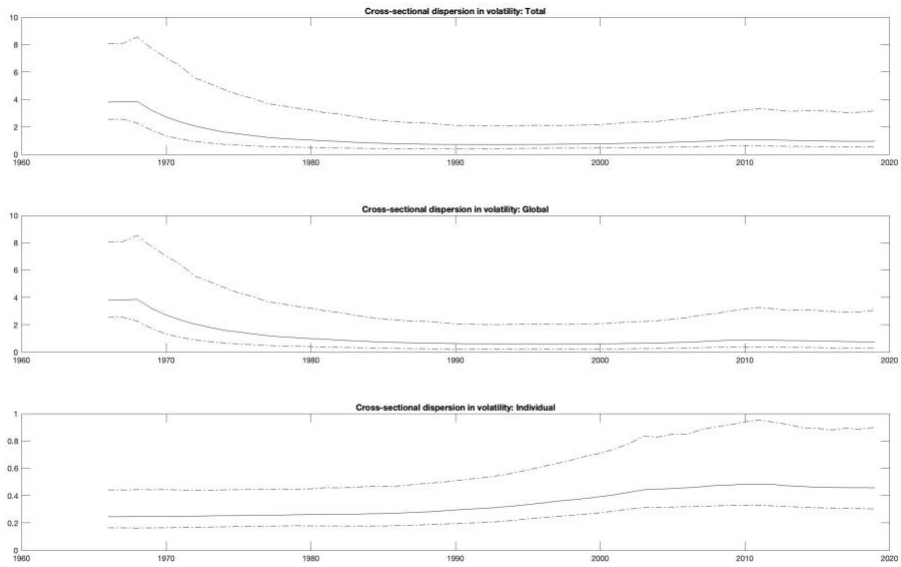


Figure 2.12 Average cross-country correlations

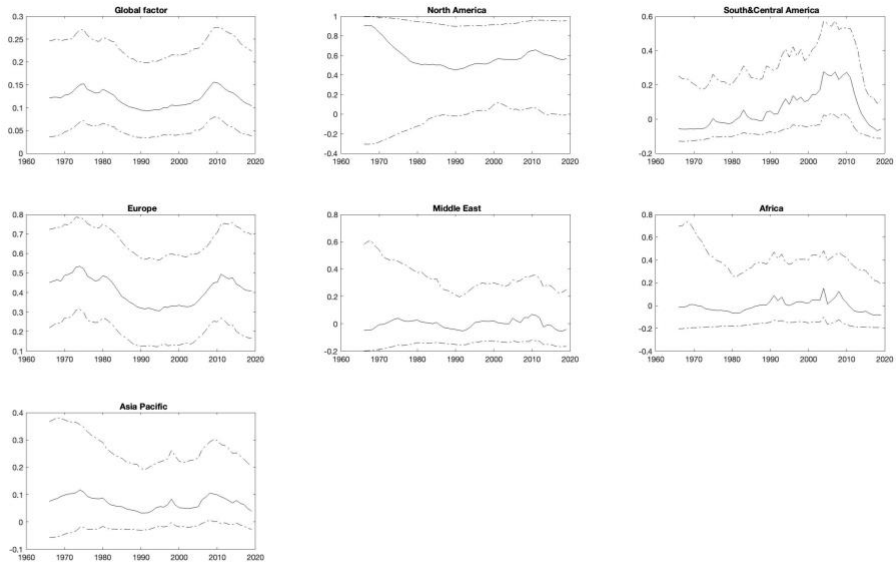


Figure 2.13 Plots of Real World Energy Consumption Index and Global and Regional Factors

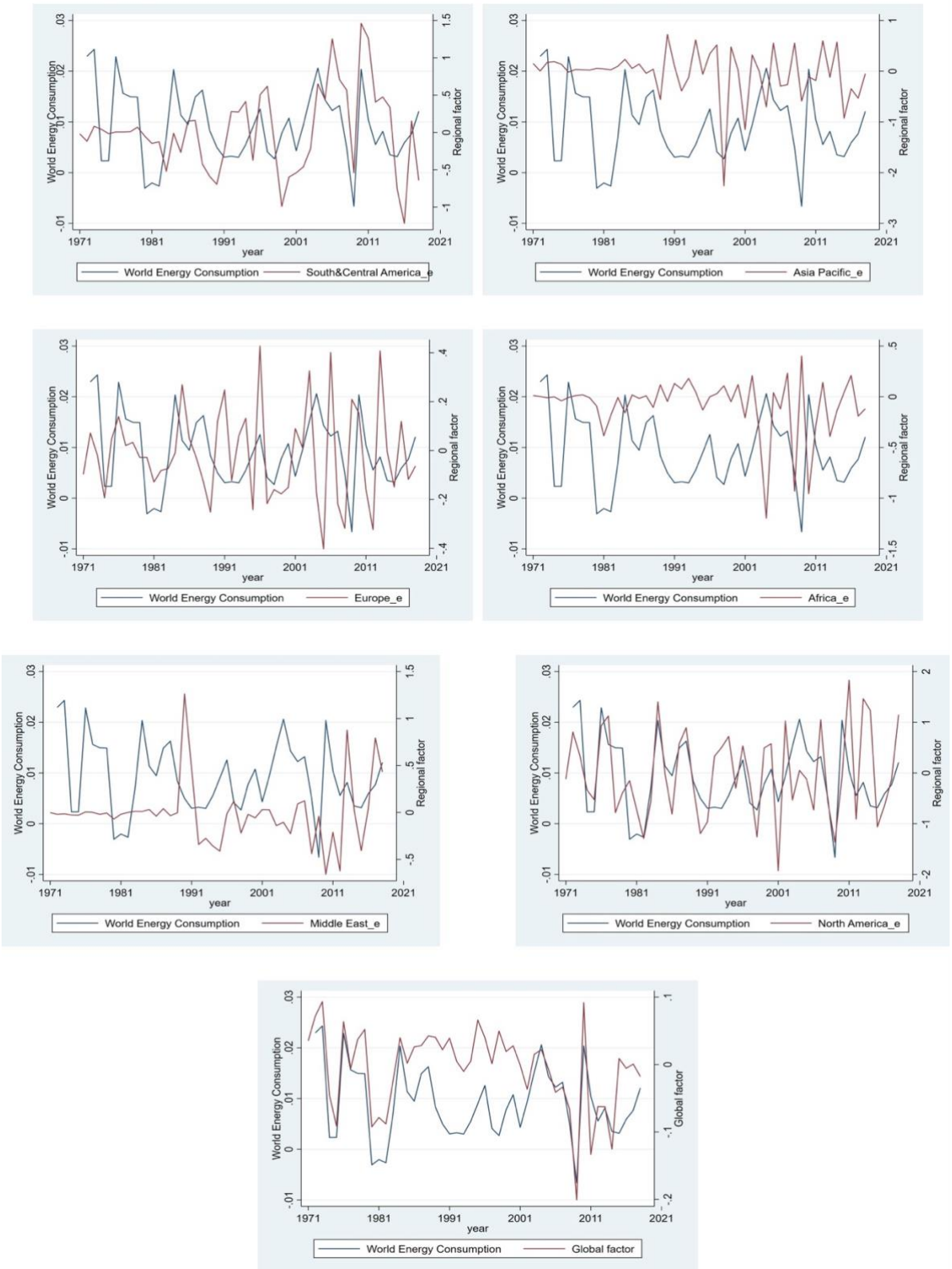


Figure 2.14 Rolling Window Correlations with World Energy Consumption in Each Country

REFERENCES

Andreoni, V. (2020). The energy metabolism of countries: Energy efficiency and use in the period that followed the global financial crisis. *Energy Policy*, *139*, 111304. https://doi.org/10.1016/j.enpol.2020.111304

Andrews, D. W. K., & Ploberger, W. (1994). Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative. *Econometrica*, *62*(6), 1383. https://doi.org/10.2307/2951753

Aqeel, A., & Butt, M. S. (2001). THE RELATIONSHIP BETWEEN ENERGY CONSUMPTION AND ECONOMIC GROWTH IN PAKISTAN. *Asia-Pacific Development Journal*, *8*(2), 10.

Azam, M., Khan, A. Q., Zaman, K., & Ahmad, M. (2015). Factors determining energy consumption: Evidence from Indonesia, Malaysia and Thailand. *Renewable and Sustainable Energy Reviews*, *42*, 1123–1131. https://doi.org/10.1016/j.rser.2014.10.061

Aziz, A. A., Mustapha, N. H. N., & Ismail, R. (2013). *Factors Affecting Energy Demand in Developing Countries: A Dynamic Panel Analysis*. 6.

Bai, J. (2009). Panel Data Models With Interactive Fixed Effects. *Econometrica*, *77*(4), 1229–1279. https://doi.org/10.3982/ECTA6135

Bai, J., & Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, *66*(1), 47. https://doi.org/10.2307/2998540

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22. https://doi.org/10.1002/jae.659

Bhatt, V., Kishor, N. K., & Ma, J. (2017). The impact of EMU on bond yield convergence: Evidence from a time-varying dynamic factor model. *Journal of Economic Dynamics and Control*, *82*, 206–222. https://doi.org/10.1016/j.jedc.2017.06.008

Campbell, J. Y., & Shiller, R. J. (1987). Cointegration and Tests of Present Value Models. *Journal of Political Economy*, *95*(5), 1062–1088. https://doi.org/10.1086/261502

Casini, A., & Perron, P. (2018). Structural Breaks in Time Series. *ArXiv:1805.03807 [Econ, Stat]*. http://arxiv.org/abs/1805.03807

Chan, H. L., & Lee, S. K. (1996). Forecasting the Demand for Energy in China. *The Energy Journal*, *17*(1), 19–30.

Chan, K. S., & Tong, H. (1990). On Likelihood Ratio Tests for Threshold Autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, *52*(3), 469–476.

Chang, T., Chu, H.-P., & Chen, W.-Y. (2013). Energy consumption and economic growth in 12 Asian countries: Panel data analysis. *Applied Economics Letters*, *20*(3), 282–287. https://doi.org/10.1080/13504851.2012.692869

Cheng, B. S. (1999). Causality Between Energy Consumption and Economic Growth in India: An Application of Cointegration and Error-Correction Modeling. *Indian Economic Review, New Series*, *34*(1), 39–49.

Cui, H., Wu, R., & Zhao, T. (2018). Dynamic Decomposition Analysis and Forecasting of Energy Consumption in Shanxi Province Based on VAR and GM (1, 1) Models. *Mathematical Problems in Engineering*, *2018*, 1–11. https://doi.org/10.1155/2018/4234206

Del Negro, M., & Otrok, C. M. (2008). Dynamic Factor Models with Time-Varying Parameters: Measuring Changes in International Business Cycles. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1136163

Diebold, F., & Kilian, L. (1999). *Unit Root Tests Are Useful for Selecting Forecasting Models* (No. w6928; p. w6928). National Bureau of Economic Research. https://doi.org/10.3386/w6928

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253. https://doi.org/10.2307/1392185

Enders, W., & Lee, J. (2012). A Unit Root Test Using a Fourier Series to Approximate Smooth Breaks*: A unit root test using a Fourier series. *Oxford Bulletin of Economics and Statistics*, *74*(4), 574–599. https://doi.org/10.1111/j.1468-0084.2011.00662.x

Enders, W., Liu, Y., & Prodan, R. (2009). FORECASTING SERIES CONTAINING OFFSETTING BREAKS: OLD SCHOOL AND NEW SCHOOL METHODS OF FORECASTING TRANSNATIONAL TERRORISM. *Defence and Peace Economics*, *20*(6), 441–463. https://doi.org/10.1080/10242690802425772

Esmaeili, P., & Rafei, M. (2021). Dynamics analysis of factors affecting electricity consumption fluctuations based on economic conditions: Application of SVAR and TVP-VAR models. *Energy*, *226*, 120340. https://doi.org/10.1016/j.energy.2021.120340

Esseghir, A., & Haouaoui Khouni, L. (2014). Economic growth, energy consumption and sustainable development: The case of the Union for the Mediterranean countries. *Energy*, *71*, 218–225. https://doi.org/10.1016/j.energy.2014.04.050

Fan, Y., & Xia, Y. (2012). Exploring energy consumption and demand in China. *Energy*, *40*(1), 23–30. https://doi.org/10.1016/j.energy.2011.09.049

Fei, L., Dong, S., Xue, L., Liang, Q., & Yang, W. (2011). Energy consumption-economic growth relationship and carbon dioxide emissions in China. *Energy Policy*, *39*(2), 568–574. https://doi.org/10.1016/j.enpol.2010.10.025

Hu, S., & Li, R. (2021). Investigating the Effects of the United States' Economic Slowdown Related to the COVID-19 Pandemic on Energy Consumption in Other Countries—A Global Vector Autoregressive Model. *Energies*, *14*(11), 2984. https://doi.org/10.3390/en14112984

Hwang, J.-H., & Yoo, S.-H. (2014). Energy consumption, CO2 emissions, and economic growth: Evidence from Indonesia. *Quality & Quantity*, *48*(1), 63–73. https://doi.org/10.1007/s11135-012-9749-5

Im, K. S., & Schmidt, P. (2008). More efficient estimation under non-normality when higher moments do not depend on the regressors, using residual augmented least squares. *Journal of Econometrics*, *144*(1), 219–233.

Islam, F., Shahbaz, M., Ahmed, A. U., & Alam, Md. M. (2013). Financial development and energy consumption nexus in Malaysia: A multivariate time series analysis. *Economic Modelling*, *30*, 435–441. https://doi.org/10.1016/j.econmod.2012.09.033

Kasman, A., & Duman, Y. S. (2015). CO2 emissions, economic growth, energy consumption, trade and urbanization in new EU member and candidate countries: A panel data analysis. *Economic Modelling*, *44*, 97–103. https://doi.org/10.1016/j.econmod.2014.10.022

Kilian, L. (2015). Energy Price Shocks. In Palgrave Macmillan (Ed.), *The New Palgrave Dictionary of Economics* (pp. 1–12). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-349-95121-5_3008-1

Kose, M. A., Otrok, C., & Whiteman, C. H. (2003). International Business Cycles: World, Region, and Country-Specific Factors. *American Economic Review*, *93*(4), 1216–1239. https://doi.org/10.1257/000282803769206278

Kumar, A. (2011). ENERGY CONSUMPTION, CO2 EMISSIONS AND ECONOMIC GROWTH: A REVISIT OF THE EVIDENCE FROM INDIA. *Applied Econometrics and International Development*, *11*, 25.

Lean, H. H., & Smyth, R. (2010). CO2 emissions, electricity consumption and output in ASEAN. *Applied Energy*, *87*(6), 1858–1864. https://doi.org/10.1016/j.apenergy.2010.02.003

Li, J., & Lee, J. (2015). Improved autoregressive forecasts in the presence of non-normal errors. *Journal of Statistical Computation and Simulation*, *85*(14), 2936–2952. https://doi.org/10.1080/00949655.2014.945930

Liu, H., Polenske, K. R., Guilhoto, J. J. M., & Xi, Y. (2014). Direct and indirect energy use in China and the United States. *Energy*, *71*, 414–420. https://doi.org/10.1016/j.energy.2014.04.089

Liu, X., Zhou, D., Zhou, P., & Wang, Q. (2018). Factors driving energy consumption in China: A joint decomposition approach. *Journal of Cleaner Production*, *172*, 724–734. https://doi.org/10.1016/j.jclepro.2017.10.238

McCormick, K., & Neij, L. (2009). Experience of Policy Instruments for Energy Efficiency in Buildings in the Nordic Countries. *International Institute for Industrial Environmental Economics (IIIEE)*, 74.

NYMEX. (2014). *Crude Oil EmiNY*. https://web.archive.org/web/20140502013508/http://tfc-charts.com/chart/QM/W

Ozturk, I., & Acaravci, A. (2010). CO2 emissions, energy consumption and economic growth in Turkey. *Renewable and Sustainable Energy Reviews*, *14*(9), 3220–3225. https://doi.org/10.1016/j.rser.2010.07.005

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, *57*(6), 1361. https://doi.org/10.2307/1913712

Pesaran, M. H., Smith, R. P., & Akiyama, T. (1998). *Energy demand in Asian developing economies*. Oxford Institute for Energy Studies, 57 Woodstock Road, Oxford, OX2 6FA, UK. https://www.osti.gov/etdeweb/biblio/675385

Pesaran, M. H., & Timmermann, A. (2004). How costly is it to ignore breaks when forecasting the direction of a time series? *International Journal of Forecasting*, *20*(3), 411–425. https://doi.org/10.1016/S0169-2070(03)00068-2

Sadorsky, P. (2011). Financial development and energy consumption in Central and Eastern European frontier economies. *Energy Policy*, *39*(2), 999–1006. https://doi.org/10.1016/j.enpol.2010.11.034

Sahir, M. H., & Qureshi, A. H. (2007). Specific concerns of Pakistan in the context of energy security issues and geopolitics of the region. *Energy Policy*, *35*(4), 2031–2037. https://doi.org/10.1016/j.enpol.2006.08.010

Saidi, K., & Hammami, S. (2015). The impact of CO2 emissions and economic growth on energy consumption in 58 countries. *Energy Reports*, *1*, 62–70. https://doi.org/10.1016/j.egyr.2015.01.003

Sarwar, S., Chen, W., & Waheed, R. (2017). Electricity consumption, oil price and economic growth: Global perspective. *Renewable and Sustainable Energy Reviews*, *76*, 9–18. https://doi.org/10.1016/j.rser.2017.03.063

Shahbaz, M., Shahzad, S. J. H., Alam, S., & Apergis, N. (2018). Globalisation, economic growth and energy consumption in the BRICS region: The importance of asymmetries. *The Journal of International Trade & Economic Development*, *27*(8), 985–1009. https://doi.org/10.1080/09638199.2018.1481991

Shaista Alam, & Butt, M. S. (2002). Causality Between Energy and Economic Growth in Pakistan: An Application of Co-intergration and Error-Correction Modeling Techniques. *Pacific and Asian Journal of Energy*, *12*(2), 151–165.

Sineviciene, L., Sotnyk, I., & Kubatko, O. (2017). Determinants of energy efficiency and energy consumption of Eastern Europe post-communist economies. *Energy & Environment*, *28*(8), 870–884. https://doi.org/10.1177/0958305X17734386

Soytas, U., & Sari, R. (2003). Energy consumption and GDP: Causality relationship in G-7 countries and emerging markets. *Energy Economics*, *25*(1), 33–37. https://doi.org/10.1016/S0140-9883(02)00009-9

Tayyab, M., Tarar, A., & Riaz, M. (2012). Threshold autoregressive (TAR) &Momentum Threshold autoregressive (MTAR) models Specification. *Research Journal of Finance and Accounting*, *3*(8), 9.

Tong, H., & Lim, K. S. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, *42*(3), 245–292.

Wang, Q., Su, M., Li, R., & Ponce, P. (2019). The effects of energy prices, urbanization and economic growth on energy consumption per capita in 186 countries. *Journal of Cleaner Production*, *225*, 1017–1032. https://doi.org/10.1016/j.jclepro.2019.04.008

Wang, Y., Chen, L., & Kubota, J. (2016). The relationship between urbanization, energy use and carbon emissions: Evidence from a panel of Association of Southeast Asian Nations (ASEAN) countries. *Journal of Cleaner Production*, *112*, 1368–1374. https://doi.org/10.1016/j.jclepro.2015.06.041

Zaharia, A., Diaconeasa, M. C., Brad, L., Lădaru, G.-R., & Ioanăș, C. (2019). Factors Influencing Energy Consumption in the Context of Sustainable Development. *Sustainability*, *11*(15), 4147. https://doi.org/10.3390/su11154147

Zhang, J., Hu, Z., Zheng, Y., Zhou, Y., & Wan, Z. (2017). Sectoral Electricity Consumption and Economic Growth: The Time Difference Case of China, 2006–2015. *Energies*, *10*(2), 249. https://doi.org/10.3390/en10020249

Zhang, J., Yang, X., Shen, F., Li, Y., Xiao, H., Qi, H., Peng, H., & Deng, S. (2012). Principal Component Analysis of Electricity Consumption Factors in China. *Energy Procedia*, *16*, 1913–1918. https://doi.org/10.1016/j.egypro.2012.01.292

APPENDIX

Appendix Figure 1 Factor Loadings to Global and Regional Factors