WORKING PAPER SERIES: WPS/0035

Forecasting GDP with energy series: ADL-MIDAS vs. Linear Time Series Models

Afees A. Salisu and Ahamuefula E. Ogbonna

Cite as:
Salisu A. A and Ogbonna A. E (2017): Forecasting GDP with energy series: ADL-MIDAS vs. Linear Time Series Models - *Centre for Econometric and Allied Research, University of Ibadan Working Papers Series, CWPS 0035*

Forecasting GDP with energy series: ADL-MIDAS vs. Linear Time Series Models

Afees A. Salisu[a,c] and Ahamuefula E. Ogbonna[a,b]

[a] Centre for Econometric & Allied Research, University of Ibadan, Nigeria.
[b] Department of Statistics, University of Ibadan, Nigeria.
[c] Corresponding author:
Email: adebare1@yahoo.com; aa.salisu@cear.org.ng
Phone: +2348034711769

Abstract

In this paper, we offer the following contributions to the extant literature on the energy-growth nexus. First, we test the predictability of the energy predictors in the predictive growth model using the autoregressive distributed lag mixed data sample (ADL-MIDAS) approach. Second, we compare the in-sample and out-of-sample forecast performance of the ADL-MIDAS model with the linear time series models involving the first order autoregressive [AR(1)] model and the autoregressive distributed lag (ARDL) model. Third, we consider an array of energy proxies ranging from aggregate data to sectoral data of energy consumption (residential, commercial, industrial and transportation) and those defined by energy sources (petroleum, natural gas, coal, electricity, nuclear electricity and renewable energy). Fourth, we test whether accounting for asymmetries matters in the ADL-MIDAS regression model for the energy-growth nexus. The results support the significant predictability of energy for growth regardless of the measures of energy. In addition, the in-sample and out-of-sample forecast results overwhelmingly favour the ADL-MIDAS over the conventional linear time series models including the restrictive AR model. Thus, allowing for high frequency data for energy in the low frequency growth model will enhance the forecast accuracy of the model. However, we find accounting for asymmetries may not significantly improve the forecast accuracy of the ADL-MIDAS model in the energy-growth nexus since, forecasts of the positive and negative asymmetric models do not differ significantly.

Keywords: Energy consumption; Growth, ADL-MIDAS; Linear time series models; Forecast evaluation

JEL Classification: C12, C22, Q42, Q43, Q47

Forecasting GDP with energy series: ADL-MIDAS vs. Linear Time Series Models

1.0 Introduction

While classical economists see energy merely as an intermediary factor supporting key growth determinants (capital and labour), ecological economists emphasize the role of energy and its availability in the production process (Stern and Cleveland 2004). Energy has assumed such important role in modern societies that energy consumption is implied in every economic activity, be it production or consumption; on the part of consumers, energy demand assists them to maximize their utility, while on the side of producers, energy is an essential factor of production in addition to capital and labour (Chontanawat et al. 2006; Pinzon 2017). More pertinently, the level of energy consumption has been noted as an indicator of economic growth level achieved by a country (Ucan et al. 2014; Rafindadi 2015; Rafindadi and Ozturk 2015). Thus, economic growth and energy consumption are closely linked such that the process of economic expansion and keeping up with production levels require greater energy demand (Dhungel 2003; Alam et al. 2016; Rafindadi and Ozturk 2017). Since the process of combining labour and capital into final goods requires energy; then the availability of energy will ultimately determine economic growth (Chiou-Wei et al. 2016).

Studies in this line of inquiry have been conducted on the basis of four distinct hypotheses; the growth hypothesis, the conservation hypothesis, the feedback hypothesis, and the neutrality hypothesis (*see* Kraft and Kraft 1978; Berndt 1978; Akarca and Long 1979, 1980; Yu and Hwang 1984; Yu and Choi 1985; Erol and Yu 1987, 1989; Yu et al. 1988 for pioneering ideas on these theories). The growth hypothesis of the economic growth – energy consumption nexus recognizes energy as an input that complements labour and capital in the production process (Narayan 2016). This idea opines that energy consumption triggers economic growth such that policies that attempt to limit/conserve energy constitute barrier to limit economic growth (Alper and Oguz 2016; Brinia et al. 2017). This suggests causation running from energy consumption to economic growth. On the other hand, the idea contained in the conservation hypothesis gives room for energy conservation policy where one-way causation runs from economic growth to energy consumption increase (Naseria et al. 2016).

The neutrality hypothesis suggests no nexus between energy consumption and economic growth, where changes in energy consumption do not affect economic growth, and vice versa (Belke et al. 2011). This view also favours energy conservation policies in that, such policies will not harm economic growth (Narayan 2016). On the part of the feedback hypothesis, it suggests complementarity between energy usage and economic growth where energy consumption and growth enhancing policies are likely to be both beneficial to achieve efficient energy usage as well as stimulate economic activity. On the other hand, negative changes to energy consumption, say energy reducing policy may be detrimental to economic growth because energy consumption represents a fundamental factor of economic growth and not just its result (*see* Mehmet and Alper 2017, for more).

Empirically, the seminal work of Kraft and Kraft (1978) evoked researchers' interests to study the relationship between energy consumption and economic growth in line with the previously discussed four theoretical expositions (*see* Akinlo 2008; Azam et al. 2015; Ahmed and Azam 2016; Adams et al. 2016; Alper and Oguz 2016; Dogan 2016; Saidi et al. 2017; Narayan and Doytch 2017 for detailed empirical literature on the earlier studies based on these hypotheses). However, contemporary studies in this area of research focus on the same goal, which is to provide evidence in support of either of the four theories (*see for example*, Adams et al. 2016; Ahmed and Azam 2016; Alper and Oguz 2016; Chen et al. 2016; Chiou-Wei et al. 2016; Dogan 2016; Narayan 2016; Naseria et al. 2016; Rodríguez-Caballero and Ventosa-Santaulària 2016; Bildirici and Ozaksoy 2017; Mehmet and Alper 2017; Menegaki and Tugcu 2017; Narayan and Doytch 2017;

Rafindadi and Ozturk 2017; Saidi et al. 2017). These studies are however divisive in terms of coverage, methodology and the form of energy considered.

The empirical appraisal reveals a strand of the empirical literature that is divided along methodology and the focus on either renewable (geothermal, wind, solar and biomass) or non-renewable (coal, oil and natural gas) energies. In a study of 16 sub-Saharan African countries and employing a panel vector autoregressive model (PVAR) in a generalized method of moments (GMM) framework, Adams et al. (2016) show support for the feedback hypothesis. Naseria et al. (2016) with an ARDL model find that renewable energy consumption leads to economic growth increase (growth hypothesis) in OECD countries. Supporting evidence from Bildirici (2016) also favours the growth hypothesis in OECD countries. However, results from Rafindadi and Ozturk (2017) show that renewable energy consumption consolidates the Germany's economic growth prospects and also shows feedback effect from economic growth to renewable energy consumption. Specifically focusing on electricity consumption in 17 Latin American countries, Canada and the USA and adopting a VEC specification that allows for structural breaks, Rodríguez-Caballero, and Ventosa-Santaulària (2016) find evidence for the growth hypothesis in eight countries including Canada, conservation hypothesis in three countries including the US, and neutrality hypothesis in three other countries.

More aptly, Narayan and Doytch (2017) consider both renewables and non-renewable energies to show that renewable energies support the neutrality hypothesis, while the feedback, growth and conservative hypotheses strongly feature with non-renewables. More elaborately, Mehmet and Alper (2017) also study both renewable and non-renewable energy consumption. The results reveal growth hypothesis for Peru, the conservation hypothesis for Colombia and Thailand, the feedback hypothesis for Greece and South Korea, and the neutrality hypothesis for the remaining 12 emerging economies in the case of renewable energy consumption. In the case of non-renewable energy consumption however, the study finds evidence to support the growth hypothesis in China, Colombia, Mexico and Philippines; the conservation hypothesis in Egypt, Peru and Portugal; the feedback hypothesis in Turkey and the neutrality hypothesis in the remaining emerging economies. On the basis of time frame, Dogan (2016) finds evidence to support conservation hypothesis and feedback hypothesis for renewable energy consumption in the short run and the long run, respectively, and feedback hypothesis for non-renewable energy in the short run and the long run.

The present study departs markedly from the literature on energy-growth nexus in the following ways. First, to the best of our knowledge, our paper is the first to evaluate the nexus using both the uniform and mixed data frequencies with an array of energy proxies by sectors (residential, commercial, industrial and transportation) and by sources (petroleum products supplied, natural gas consumption, coal consumption, electricity end use, nuclear electricity net generation and renewable energy consumption). Since energy series are usually available at a high frequency, it will be an interesting exercise to examine how the information inherent in such series can be exploited to enhance the forecast performance of the energy – growth nexus. The only closely related study that employs the MIDAS approach is the Valadkhani and Smyth (2017), however, the paper is limited to oil price – output nexus and without any consideration for out-of-sample forecast evaluation. Thus, for comprehensive analyses and robustness checks, we consider energy proxies defined by sectoral distribution and sources of energy (*see* also Narayan and Doytch 2016). However, the Narayan and Doytch (2016) paper involves uniform data frequency and is also restricted to in-sample analyses. The consideration of the different proxies for energy further strengthens the literature in terms of whether the nexus is sensitive to the choice of energy proxy.

Secondly, in addition to the fact that the consideration of mixed data frequencies in the energy-growth literature is scarce (where the Valadkhani and Smyth (2017) work is the only paper known to us), the comparative evaluation of the forecast performance of the energy predictor using both uniform and mixed data frequencies is non-existent. In the search of accurate growth forecasts, such comparison is necessary to test whether having more observations based on data frequency will improve the growth forecast. The forecast evaluation is conducted for both in-sample and out-of-sample forecasts, which is not popular in the energy-growth literature. Even the studies involving uniform data frequency is more prominently limited to impact analyses, which may not be sufficient to support such predictability in the out-of-sample period. Evaluating both in-sample and out-of-sample forecast accuracy is particularly important for policy decisions, where accurate growth forecast is required in the budgeting process and in setting macroeconomic targets for an economy.

Third, given that the literature is replete with evidence that economic variables respond asymmetrically to changes in energy prices (the idea that was formalized by Kilian (2009) using oil price), we further hypothesize that such asymmetries may matter for energy-growth relationship. The two studies, known to us, that have evaluated how growth responds asymmetrically to changes in energy prices are Nusair (2016) (with uniform frequency) and Valadkhani and Smyth (2017) (with MIDAS); however both are limited to oil prices. Unlike these studies, we consider various energy proxies determined by sources and sectors and the probable asymmetric relationship between energy and growth is evaluated for both in-sample and out-of-sample forecasts.

The rest of the paper is structured as follows: Section 2 provides the methodology including the model set up and forecast performance measures. Thereafter, some preliminary analyses are rendered in Section 3. While Section 4 discusses the results, Section 5 however, concludes the paper.

2.0 Methodology

2.1 The Model

In this paper, the predictive model for energy-growth nexus is theoretically motivated by the growth hypothesis. The studies that establish same (Naseria et al. 2016; Bildiric 2016; Rafindadi and Ozturk 2017; Menegaki and Togcu 2017) and the arguments advanced in the literature (*see* Chontanawat et al. 2006; Ucan et al. 2014; Rafindadi 2015; Rafindadi and Ozturk 2015; Alam et al. 2016; Chiou-Wei et al. 2016; Pinzon 2017; Rafindadi and Ozturk 2017) suggest that energy consumption can actually serve as a good predictor for economic growth. The foregoing provides strong motivation for the present study to explore the possibility of forecasting US GDP with energy consumption. The US economy as a highly industrialized economy and a huge consumer of various energy commodities provides robust data that allows for detailed analyses of the energy-growth nexus.

In tune with the ensuing arguments from the growth thesis of and accommodating the dynamics common to economic variables, the following conventional bivariate predictive model for the US GDP is specified as follows:

$$\ln GDP_t = a + \vartheta \ln engy_{t-1} + \varepsilon_t \qquad (1)$$

where $\ln GDP_t$ is the log of US GDP at a given period *t* and $\ln engy_t$ is the log of energy predictor at a given period *t* while ε_t is the error term. The predictability of US GDP is tested on the basis of energy consumption given that the null hypothesis, $H_0 : \vartheta = 0$ is rejected. Equation (1) is the standard bivariate predictive model for testing the predictability between two variables. This approach however is restrictive as limits the dynamics in the model to the predictor. Meanwhile, most economic time series tend to exhibit persistence and therefore allowing for dynamics in both the predicted and the predictor may enhance the forecast performance. The persistence effect evident in the regressand partly justifies why autoregressive models tend to outperform most economic models particularly those relating to exchange rates (*see* e.g., Moosa and Burns 2012, 2014a, b;

Moosa 2013), inflation (*see* e.g., Marcellino 2008; Stock and Watson 2008, 2009) and GDP (*see* e.g., Stock and Watson 2003; Marcellino 2008). Also, we want to ensure that we consider a model that is reasonably and theoretically comparable to the ADL-MIDAS models such that the only difference lies in the data frequency. Thus, the benchmark model for our study is the Autoregressive Distributed Lag [ARDL(*p,q*)] model of Pesaran and Shin (1999) and Pesaran et al. (2001) with uniform frequency and it is specified as:

$$\Delta \ln GDP_t = c + \sum_{i=1}^{p-1} \phi_i \Delta \ln GDP_{t-i} + \sum_{i=0}^{q-1} \varphi_i \Delta \ln engy_{t-i} + \theta \left[\ln GDP_{t-1} - \left\{ \alpha + \delta \ln engy_{t-1} \right\} \right] + \xi_t \quad (2)$$

The equation (2) is the extended version of equation (1) and it is captured in such a way as to assess both the long run and short run relationships between energy and growth. The regressand ($\ln GDP$) and regressor ($\ln engy$) are as previously defined. The parameters, ϕ_i and φ_i are the short run coefficients of the lagged regressand and regressor respectively. The terms in the "[ ]" parenthesis are those of the long run GDP regression model. While δ is the long run coefficient, θ is the coefficient of the speed of adjustment to the long run equilibrium. The optimal lags for the ARDL model are determined on the basis of the Schwartz Information Criterion (AIC) where the model with the least SIC is considered parsimonious and the corresponding lag structure producing such model is considered the optimal lag and is therefore used for the in-sample and out-of-sample forecasts.

Alternatively, to substantiate the possibility of using mixed data frequencies when forecasting the US energy-growth nexus,[1] we express the dynamic model in equation (1) in ADL-MIDAS form.[2] There are different variants of the MIDAS regression models

[1] See (Forsberg and Ghysels 2006; Alper et al. 2008; Bai et al. 2009; Barsoum and Stankiewicz 2015; Jung 2017) for arguments in support of such modelling structure that allows the predictor and predicted variables to be sampled at different data frequencies.

[2] Studies such as Ghysels et al. (2009); Andreou et al. (2013); Albu et al. (2015); Ghysels (2016); Salisu and Ogbonna (2017) also document evidence in support of the ADL-MIDAS framework.

based on how the high frequency regressors are handled in a predictive model with a low frequency regressand. There is the Flat weight aggregation approach (*see* e.g., Asimakopoulos et al. 2013) which involves equal weights for the aggregation of the high frequency data. In our case where the energy variables are monthly series while the GDP is quarterly, the Flat weight approach implies equal weights on each month. However, one of the shortcomings of this approach is that the estimators may be biased if the true weighting scheme is not that of equal weights and this affects the forecasting accuracy of the model (Asimakopoulos et al. 2013) (see appendix 1). Another variant of the MIDAS regression is the Unrestricted MIDAS (U-MIDAS) (*see* e.g. Foroni et al. 2011) which does not require the aggregation of high frequency observations in order to convert to low frequency. A typical representation for energy-growth nexus using the U-MIDAS framework can be expressed as (see also Asimakopoulos et al. 2013):

$$\ln GDP_{t+1}^{Q} = \gamma_0 + \sum_{i=0}^{N_M - 1} \gamma_i \ln engy_{N_M - i,t}^{M} + \varepsilon_{t+1}^{Q} \tag{3}$$

where N_M denotes the number of months in a quarter while Q and M are the quarterly and monthly data frequencies respectively. Unlike the Flat weight aggregation approach, the U-MIDAS does not require any assumption on the weights attached to each month and are therefore considered to be are unrestricted. Therefore, the OLS method can be used to estimate equation (3) directly. One of the shortcomings of the U-MIDAS however is the problem of parameter proliferation. For instance, there are four parameters to be estimated for quarterly/monthly data in which one coefficient is estimated for each month and one for the constant. The parameter proliferation becomes severe if the gap between the low and high frequencies widens, or if the number of lags of each month increases. We thus opt for the ADL-MIDAS model proposed by Ghysels et al. (2006) which does not require any assumption on the weights for aggregation of high frequency variables (as in the Flat weight aggregation MIDAS variant) and it also helps to circumvent the problem of parameter proliferation inherent in the U-MIDAS. Thus, we consider the exponential Almon lag polynomial proposed by Ghysels et al. (2007) which most general weighting scheme compared to others as it is very flexible and can take many shapes. The estimated ADL-MIDAS model for energy-growth nexus is specified as follows with the lag structure $(p_{GDP}^{Q}, q_{engy}^{M})$:[3]

$$\ln GDP_{t+1}^{Q} = \lambda + \sum_{i=0}^{p_{GDP}^{Q}-1} \alpha_i \ln GDP_{t-i}^{Q} + \beta \sum_{i=0}^{q_{engy}^{M}-1} \cdot \sum_{j=0}^{N_M-1} w_{i+j*N_M}(\phi^M) \ln engy_{N_M-j,t-i}^{M} + \zeta_{t+1} \quad (4)$$

where p_{GDP}^{Q} and q_{engy}^{M} denote the number of lags of the quarterly (low) and monthly (high) frequency variables, respectively. $w_i(\phi^H)$ is a weighting structure of a two parameter exponential Almon lag polynomial expressed as:

$$w_i(\phi^H) = w_i(\phi_1\phi_2) = \frac{e^{(\phi_1 i + \phi_2 i^2)}}{\sum_{i=0}^{k} e^{(\phi_1 i + \phi_2 i^2)}} \quad (5)$$

Since the ARDL and ADL-MIDAS are similar except for the data frequency of the regressor(s) and also for consistency, we maintain the same lag structure for both models where the optimal lag is determined using the ARDL.

As previously mentioned, we also test for asymmetry in the ADL-MIDAS predictive model following the Shin et al. (2014) approach which involves partial sum decomposition of energy proxies into positive and negative changes. The asymmetric components of the energy proxies are computed as follows:[4]

$$engy_t^{+} = \sum_{k=1}^{t} \Delta engy_{ik}^{+} = \sum_{k=1}^{t} \max(\Delta engy_{ik}, 0) \quad (6)$$

$$engy_t^{-} = \sum_{k=1}^{t} \Delta engy_{ik}^{-} = \sum_{k=1}^{t} \min(\Delta engy_{ik}, 0) \quad (7)$$

where equations (6) and (7) are the positive and negative partial sum decompositions of changes in energy consumption respectively. By taking account of equations (6) and (7) singly in

[3] The specification is also in line with Asimakopoulos et al. (2013) and Albu et al. (2015).
[4] This approach has also been employed in the literature to decompose energy related variables. Examples of studies in this regard are Hoang et al. (2016); Nusair (2016); Huang et al. (2017); Salisu and Kazeem (2017) and Salisu et al. (2017).

equation (4), we can express the corresponding ADL-MIDAS equations for both $engy_t^+$ and $engy_t^-$ respectively as follows:

$$\ln GDP_{t+1}^{Q} = \lambda^{+} + \sum_{i=0}^{p_{GDP}^{Q}-1} \alpha_i^{+} \ln GDP_{t-i}^{Q} + \beta^{+} \sum_{i=0}^{q_{engy^+}^{M}-1} \cdot \sum_{j=0}^{N_M-1} w_{i+j*N_M}^{+}(\phi^{M}) \ln engy^{+M}_{N_M-j,t-i} + \zeta_{t+1}^{+} \qquad (8)$$

$$\ln GDP_{t+1}^{Q} = \lambda^{-} + \sum_{i=0}^{p_{GDP}^{Q}-1} \alpha_i^{-} \ln GDP_{t-i}^{Q} + \beta^{-} \sum_{i=0}^{q_{engy^-}^{M}-1} \cdot \sum_{j=0}^{N_M-1} w_{i+j*N_M}^{-}(\phi^{M}) \ln engy^{-M}_{N_M-j,t-i} + \zeta_{t+1}^{-} \qquad (9)$$

In other words, equation (4) is the predictive model of growth model with a symmetric energy consumption proxy while equations (8) and (9) represent the asymmetric variants. There is evidence of asymmetry, if the coefficients of β^+ and β^- are statistically different from each other; otherwise their effect on growth is considered identical.

For completeness, we also compare the forecast performance of both models with the first order autoregressive model specified as follows:

$$\ln GDP_t = \alpha + \rho \ln GDP_{t-1} + \varepsilon_t; \qquad (10)$$

where ρ is the first order autogressive coefficient and is expected to satisfy the stationarity condition of $|\rho| < 1$.

2.2 Forecast Evaluation Measures

We test the validity of our forecast models for both in-sample and out-of-sample predictability on the basis of various forecast evaluation criteria. We use the 50 percent and 75 percent of the total observations for the forecast evaluation and the rolling window approach is employed to produce the forecast results for the energy-growth nexus. As previously noted, we compare results for the ARDL and the ADL-MIDAS models with the AR model. We employ two forecast measures: the Campbell-Thompson (C-T) test and the Diebold and Mariano (D-M) test. The C-T test is computed as $1-\left(M\widehat{S}E_1 / M\widehat{S}E_0\right)$. The $M\widehat{S}E_1$ and $M\widehat{S}E_0$ are the mean square error (MSE) of the prediction from the unrestricted and restricted models, respectively. The restricted models in this case are the AR(1) and ARDL models, while the unrestricted version is the ADL-MIDAS. A positive value of the statistic implies that the unrestricted model outperforms the restricted model; otherwise, it does not.

Also, the D-M test is used to test for the equality of forecast accuracy of two competing models at a time. The test statistic is specified as:

$$\text{D-M stat} = \frac{\bar{d}}{\sqrt{\frac{1}{T}V\left(d\right)}} \sim N\left(0,1\right) \tag{11}$$

where $\bar{d} = \frac{1}{T}\sum_{t=1}^{T}\left[g\left(\xi_{it}\right)-g\left(\xi_{jt}\right)\right]$ is the sample mean loss differential and $V\left(d\right)$ is the unconditional variance of d. The $\left\{\xi_{it}\right\}_{t=1}^{T}$ and $\left\{\xi_{jt}\right\}_{t=1}^{T}$ are the forecast errors associated with the two forecasts say $\left\{\hat{y}_{it}\right\}_{t=1}^{T}$ and $\left\{\hat{y}_{jt}\right\}_{t=1}^{T}$ respectively. The $g\left(\xi_{it}\right)$ and $g\left(\xi_{jt}\right)$ are the loss functions associated with these two forecasts while $d_t \equiv g\left(\xi_{it}\right)-g\left(\xi_{jt}\right)$ is the loss differential. The null hypothesis of equal forecast accuracy for the two forecasts is that $E\left[d_t\right]=0$. Thus, the forecast accuracy of two competing models is considered relatively equal if the null hypothesis of the D-M test is not rejected; otherwise, it is not.

3.0 Data and Preliminary Analyses

3.1 Data

In this paper, we utilize the US gross domestic product (GDP) and energy consumption sourced respectively from FRED database at the Federal Reserve Bank of St Louis: https://fred.stlouisfed.org/ and U.S. Energy Information Administration (EIA). The data, which span a period of forty-four (44) years from January 1973 to June 2017, comprise a mixture of frequencies (low and high) with GDP obtained quarterly, while energy consumption is on a monthly basis. Consequently, there are about 178 observations for the quarterly GDP series and 534 observations for the monthly energy consumption. For the purpose of robustness of estimates, the paper considers 11 different proxies for energy consumption with aggregate energy consumption data inclusive. These proxies are subdivided by sectors (residential, commercial, industrial and transportation) and by energy sources (petroleum products supplied, natural gas consumption, coal consumption, electricity end use, nuclear electricity net generation and renewable energy consumption). The detailed description of each variable and the corresponding unit of measurement and frequency are presented in Table 1.

Table 1: Data Description Table

Variable	**Acronym**	**Unit of Measurement**	**Frequency**	**No. of Obs.**
Gross Domestic Product	gdp	Billion USD	Quarterly	178
Energy Consumption				
By Sector:				
Residential	$energy_{residence}$	Trillion Btu	Monthly	534
Commercial	$energy_{commercial}$	Trillion Btu	Monthly	534
Industrial	$energy_{industrial}$	Trillion Btu	Monthly	534
Transportation	$energy_{transportation}$	Trillion Btu	Monthly	534
Aggregate Energy Consumption	$energy_{aggregate}$	Trillion Btu	Monthly	534
By Source:				
Petroleum Products Supplied	$energy_{petroleum}$	Thousand Barrels per Day	Monthly	534
Natural Gas Consumption	$energy_{nat.gas}$	Billion Cubic Feet	Monthly	534
Coal Consumption	$energy_{coal}$	Thousand Short Tons	Monthly	534
Electricity End Use	$energy_{electricity}$	Billion Kilowatthours	Monthly	534
Nuclear Electricity Net Generation	$energy_{nucl.elect}$	Million Kilowatthours	Monthly	534
Renewable Energy Consumption	$energy_{renewable}$	Trillion Btu	Monthly	534

Note: The start and end periods for all the variables are January 1973 and June 2017, respectively. The unit of measurement "Btu" represents British thermal unit. The renewable energy consumption data is obtained by aggregating its components, which include hydroelectric power consumption, geothermal energy consumption, biomass energy consumption (wood and waste energy consumption). The data were sourced from FRED database at the Federal Reserve Bank of St Louis: https://fred.stlouisfed.org/ and U.S. Energy Information Administration (EIA).

3.2 Summary Statistics

The summary statistics presented in Table 2, which includes the mean, standard deviation, skewness and kurtosis, gives a brief description of the nature of the data to be employed in this paper. The average gdp for the 44-year period considered in this study was approximately 8,646.03billion USD, while the average energy consumption (aggregate) was approximately 7,353.75trillion Btu with standard deviation 941.27 trillion Btu. The sector level energy consumption ranged averagely between 1,203.97trillion Btu and 2,643.71trillion Btu, with the least and highest energy consuming sector being the commercial sector and the industrial sector, respectively. On the standard deviation of estimates across sectors, the least deviation was observed in the industrial sector. This indicates that the amount of energy used by the industrial sector over the study period are not as far apart from the obtained sector average, compared to the deviations in other sectors. The other energy consumption proxies are different forms of energy that are used independently or in combined form in one or more sectors. The citizenry of US utilize an average of 18,180,340Barrels of petroleum products supplied, 1,791.30billion cubic feet of natural gas, 73,699,090short Tons of coal, 250.66billion kilowatthours of electricity, 47,371.95million kilowatt-hours of nuclear electricity and 540.98trillion Btu of renewable energy, monthly. All the variables except gdp, $energy_{residence}$, $energy_{nat.gas}$ and $energy_{renewable}$ are negatively skewed, while only $energy_{renewable}$ is leptokurtic.

Table 2: Summary Statistics of the Data (January 1973 to June 2017)

Variables	**Mean**	**Standard Deviation**	**Skewness**	**Kurtosis**
gdp	8,646.03	5,342.00	0.3566	1.8376
$energy_{residence}$	1,532.63	391.99	0.6553	2.8839
$energy_{commercial}$	1,203.97	289.17	-0.1222	1.9581
$energy_{industrial}$	2,643.71	190.84	-0.2198	2.6590
$energy_{transportation}$	1,973.27	294.31	-0.0604	1.7158
$energy_{aggregate}$	7,353.75	941.27	-0.0544	2.1246

$energy_{petroleum}$	18,180.34	1,649.14	-0.1462	2.0468
$energy_{nat.gas}$	1,791.30	429.42	0.5286	2.8731
$energy_{coal}$	73,699.09	15,708.46	-0.1446	2.0906
$energy_{electricity}$	250.66	67.01	-0.0489	1.8865
$energy_{nucl.elect}$	47,371.95	19,880.99	-0.5139	1.8905
$energy_{renewable}$	540.98	130.03	1.0761	3.9272

Note: Summaries were compiled by authors using the data extract from FRED database at the Federal Reserve Bank of St Louis: https://fred.stlouisfed.org/ and U.S. Energy Information Administration (EIA). All the variables have been log-transformed. Statistical significance at 1% level is represented by ***.

We also allow for some plots to highlight the behavior of the series captured over time, which are not usually visible in the summary statistics. Figure 1 shows the plot of the sectoral energy consumption against the Gross Domestic Product (GDP) of the US. As expected, GDP clearly trended upwards, while energy consumption across the sectors fluctuated noticeably and relatively seem to trend upwards as well. The industrial sector could be seen to utilize the largest proportion of energy in US. This commemorates with the US being highly industrialized. For the period 2009m03, when the GDP dropped, a similar drop is observed in the industrial sector energy consumption, which suggests a plausibility of an energy-growth nexus.

Figure 1: Energy Consumption (Sector-Level) and Gross Domestic Product (GDP)

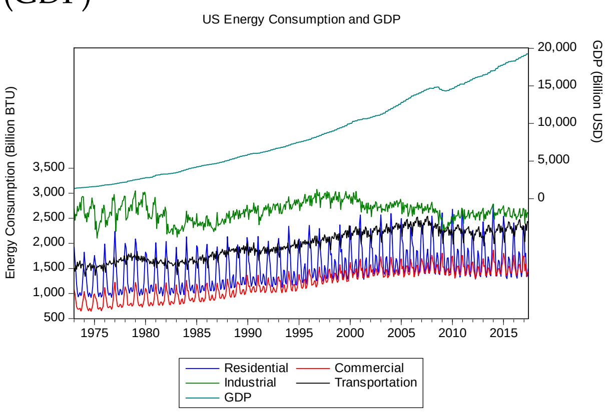


Figure 2: Energy Consumption (Aggregate) and Gross Domestic Product (GDP)

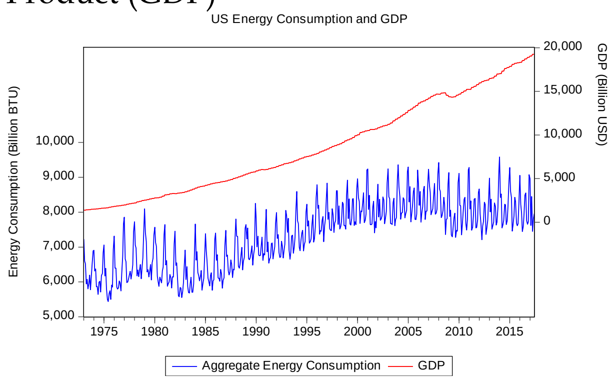
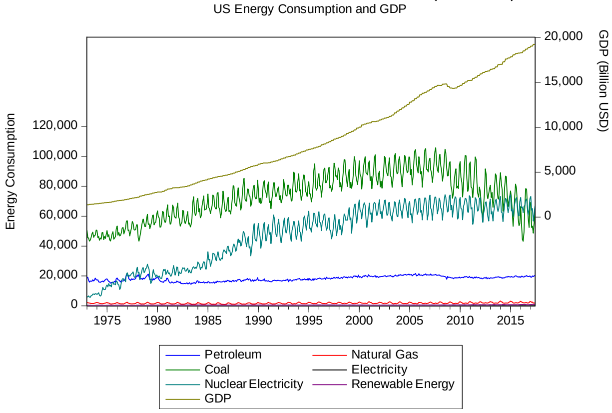


Figure 3: Energy Consumption (Component-Level) and Gross Domestic Product (GDP)



For the plot of the aggregate energy consumption and the gross domestic product displayed in Figure 2, a similar pattern as that of the industrial sector energy consumption is observed, since it accounts for the largest proportion of energy consumption. Although, the aggregate energy consumption showed some kind of seasonal variation, it generally trended upward.

Figure 3 is a display of the energy consumption by energy sources against the gross domestic product for the US. Coal, which is presumably industrially-related, seems to be the most utilized energy source for the period under consideration. Although, the graph shows the steady increase in the use of nuclear electricity, it might replace the use of coal in the nearest future.

4.0 Discussion of Results

The results of the data analysis discussed in this section are based on two different data samples (50% and 75% of full sample) and three competing models - ADL-MIDAS, ARDL and AR(1), evaluated for both in-sample and out-of-sample forecasts with two forecast horizons (*h*) involving four quarters ((*h*=4) and 8 quarters (*h*=8). These results are discussed under three subsections, which include the in-sample predictability, in-sample forecast evaluation and out-of-sample evaluation.

4.1 In-Sample Predictability

This evaluates the statistical significance of the predictor (energy) in the growth model. Thus, our analyses are restricted to the ARDL and ADL-MIDAS. However, we focus essentially on how well the energy proxy will behave in the growth model involving the ADL-MIDAS. The results for the predictability of ARDL are suppressed as several papers have largely reported the significant relationship between energy (using different proxies) and growth (*see* for example, Ocal and Aslan 2013; Cherni and Jouini 2017; Rafindadi and Ozturk 2017). What is relatively scarce however is whether ADL-MIDAS can also be used in the impact assessment of energy on growth. Notwithstanding, all the competing models are considered for the forecast evaluation since that involves determining the model with the best forecast accuracy. This is particularly important in a situation where both ARDL and ADL-MIDAS models validate the inclusion of energy in the predictive regression model for growth.

Table 3: Estimates of the ADL-MIDAS using energy proxies by sectors and sources

	Sectors				Aggregate Energy	Sources					
	Residential	Commercial	Industrial	Transportation		Petroleum	Natural Gas	Coal	Electricity	Nuclear Electricity	Renewable Energy
50% Sample											
$lgdp_{t-1}$	1.2663***	1.2375***	1.1767***	1.1809***	1.1916***	1.1794***	1.2634***	1.1803***	1.2033***	1.2346***	1.2403***
	(0.0926)	(0.0939)	(0.0962)	(0.1035)	(0.1008)	(0.0897)	(0.0909)	(0.0932)	(0.1040)	(0.1223)	(0.0984)
$lgdp_{t-2}$	-0.2716***	-0.2458***	-0.1839*	-0.1898*	-0.1993**	-0.1865**	-0.2677**	-0.1905**	-0.2178**	-0.2574**	-0.2485**
	(0.0921)	(0.0931)	(0.0955)	(0.1027)	(0.1001)	(0.0891)	(0.0905)	(0.0924)	(0.1019)	(0.1179)	(0.0975)
$lenergy_t$	0.0079***	0.0120***	0.0095***	0.0118***	0.0089***	0.0076***	0.0067***	0.0090***	0.0255***	0.0196***	0.0133***
	(0.0019)	(0.0027)	(0.0019)	(0.0023)	(0.0018)	(0.0014)	(0.0016)	(0.0017)	(0.0058)	(0.0053)	(0.0032)
$lenergy_{t-1}$	0.0012	-0.0006	-0.0012	-0.0013	-0.0005	-0.0033	0.0053	-0.0029	-0.3480	0.0005	0.0055
	(2.0928)	(2.2420)	(12.1298)	(4.7396)	(5.4612)	(7.5744)	(4.1037)	(4.1892)	(2.7578)	(1.6568)	(3.7648)
75% Sample											
$lgdp_{t-1}$	1.27***	1.2396***	1.1931***	1.1877***	1.1998***	1.1882***	1.2726***	1.2043***	1.2386***	1.2745***	1.2505***
	(0.0790)	(0.0802)	(0.0840)	(0.0897)	(0.0848)	(0.0781)	(0.0783)	(0.0823)	(0.0926)	(0.0992)	(0.0861)
$lgdp_{t-2}$	-0.275***	-0.2478***	-0.1989**	-0.1953**	-0.2063**	-0.1943**	-0.2769***	-0.2119**	-0.2514***	-0.2884***	-0.2564***
	(0.0786)	(0.0795)	(0.0835)	(0.0890)	(0.0843)	(0.0776)	(0.0780)	(0.0816)	(0.0908)	(0.0969)	(0.0854)
$lenergy_t$	0.0077***	0.01186***	0.0081***	0.0105***	0.0078***	0.0067***	0.0066***	0.0070***	0.0226***	0.0124***	0.0102***
	(0.0015)	(0.0022)	(0.0014)	(0.0017)	(0.0013)	(0.0011)	(0.0013)	(0.0012)	(0.0047)	(0.0029)	(0.0021)
$lenergy_{t-1}$	0.00004	-0.0009	-0.0008	-0.0031	-0.0006	-0.0025	0.0020	-0.0067	-0.0119	-0.0010	0.0547
	(1.49)	(1.7639)	(9.3104)	(3.8924)	(4.2434)	(7.4266)	(2.6730)	(4.2481)	(2.1494)	(2.0975)	(3.2802)

Note: The estimates are obtained from ADL-MIDAS regression models with suppressed intercepts. Statistical significance of the test values at 1%, 5% and 10% are represented by ***, ** and *, respectively. Estimates of the ADL-MIDAS regressions were obtained using the Normalized Exponential Almon lag polynomial $(nealmon)$ weighting scheme in R package. The variables are expressed in logs to enable us produce the estimates in elasticities. In the model employed, the intercept is suppressed, since our interest lies majorly on the significant impact of energy (in different sectors and sources) on growth.

As previously stated, while the interest here is to check the in-sample predictability of our preferred model, we also try to examine the robustness of our preferred model, by using different energy consumption proxies. We further examine three very important features (sign, size and significance) of the resulting estimates across the proxies and the sample structures. A similar pattern is observed regardless of the energy consumption and sample structure used, with $lenergy_t$ having positive significant impact on growth. The significant estimates obtained validate the argument of the growth thesis (*see* Chontanawat et al. 2006; Ucan et al. 2014; Rafindadi 2015; Rafindadi and Ozturk 2015; Chiou-Wei et al. 2016; Alam et al. 2016; Rafindadi and Ozturk 2017; Pinzon 2017, etc.) on the energy-growth nexus and further reveal the importance of any form of energy commodity as a growth determinant, especially in a highly industrialized economy like the US economy, which is a huge consumer of various energy commodities.

4.2 In-Sample Forecast Evaluation

Here, we formally compare the in-sample forecast performance of the ADL-MIDAS with the ARDL and AR(1) models using both the C-T test and the D-M test (see results in Tables 4 and 5). Starting with the C-T test, as shown in Table 4, the statistics are consistently positive for all the considered proxies for energy and data samples implying that the ADL-MIDAS is superior to both ARDL and AR(1) models in terms of in-sample forecast performance. In other words, allowing for high frequency energy data in the low frequency growth model will enhance the accuracy of the growth forecasts. Notwithstanding the evidence from the C-T test, it is also imperative to ascertain if the difference in the forecast performance between the ADL-MIDAS and other competing models is statistically significant. One of the prominent tests used in this regard is the DM test and is adopted in this study. The results of the test are reported in Table 5. For the DM test, significance implies a rejection of the null hypothesis of equality in the forecast accuracy between two competing models. As demonstrated with the C-T test, there are two pairs of competing models, ADL-MIDAS and ARDL as a pair and ADL-MIDAS and AR(1) as the second pair.

As depicted in Table 5, the DM test statistics are consistently significant at the 1 percent level regardless of the energy proxy and data sample. This is an indication that the difference in the forecast results of ADL-MIDAS and other models is statistically significant and since the C-T statistics are positive, it then implies that the ADL-MIDAS offers better in-sample forecast performance than other considered models. Consequently, it suffices to say that the ADL-MIDAS is an improvement on the linear time series models with uniform data frequency particularly in relation to energy-growth nexus. Technically, the findings suggest that the additional information provided by incorporating a high frequency energy predictor in the forecast of a lower frequency growth variable lowers the bias proportion in the in-sample growth forecasts.

Table 4: In-Sample Forecast Evaluation using Campbell-Thompson test Statistic

Energy Consumption Proxies	**50% Sample**		**75% Sample**	
	$ARDL$	$AR(1)$	$ARDL$	$AR(1)$
$lenergy_{residence}$	0.9134	0.9978	0.9370	0.9993
$lenergy_{commercial}$	0.9408	0.9979	0.9368	0.9993
$lenergy_{industrial}$	0.9201	0.9980	0.9528	0.9994
$lenergy_{transportation}$	0.9108	0.9980	0.9329	0.9994
$lenergy_{aggregate}$	0.9108	0.9980	0.9411	0.9994
$lenergy_{petroleum}$	0.9007	0.9980	0.9280	0.9994
$lenergy_{nat.gas}$	0.9589	0.9979	0.9693	0.9993
$lenergy_{coal}$	0.8982	0.9980	0.9528	0.9994
$lenergy_{electricity}$	0.9378	0.9979	0.9603	0.9993
$lenergy_{nucl.elect}$	0.9181	0.9979	0.9808	0.9993
$lenergy_{renewable}$	0.8998	0.9979	0.9680	0.9993

Note: ADL-MIDAS model is unrestricted model in the Campbell-Thompson test reported in the table, while ARDL and AR(1) are the restricted model. A positive statistic implies that the unrestricted model outperforms the competing models, while a negative value implies the reverse.

Table 5: In-Sample Forecast Evaluation using Diebold and Mariano test Statistic

Energy Consumption Proxies	**50% Sample**		**75% Sample**	
	$ARDL$	$AR(1)$	$ARDL$	$AR(1)$
$lenergy_{residence}$	8.995***	14.318***	11.113***	19.240***
$lenergy_{commercial}$	9.524***	14.319***	11.092***	19.240***
$lenergy_{industrial}$	8.936***	14.322***	13.851***	19.242***
$lenergy_{transportation}$	8.045***	14.323***	11.881***	19.242***
$lenergy_{aggregate}$	8.506***	14.321***	12.674***	19.242***
$lenergy_{petroleum}$	7.713***	14.322***	12.093***	19.242***
$lenergy_{nat.gas}$	10.258***	14.319***	9.9256***	19.240***
$lenergy_{coal}$	7.741***	14.322***	13.525***	19.242***
$lenergy_{electricity}$	9.533***	14.320***	11.347***	19.241***
$lenergy_{nucl.elect}$	6.185***	14.321***	14.620***	19.242***
$lenergy_{renewable}$	7.164***	14.321***	12.256***	19.241***

Note: Each cell contains the Diebold and Mariano (DM) test value, which compares ADL-MIDAS with the other models in each column using the squared error as the loss function and its significance implies preference in favour of the row element. Statistical significance of the test values at 1%, 5% and 10% are represented by ***, ** and *, respectively.

4.3 Out-of-Sample Forecast Evaluation

Like the in-sample scenario, we also allow for multiple data samples (50 percent and 75 percent of the full data sample) as well as multiple forecast horizons for robustness. Consequently, both short and long out-of-sample forecasts are produced for four quarters $(h = 4)$ and eight quarters $(h = 8)$, respectively. In the same vein as depicted for the case of the in-sample forecast evaluation, we employ both the C-T test and the DM test to ascertain which model outperforms the other for the out-of-sample case. Again, both tests reveal the superiority of the ADL-MIDAS over ARDL and AR(1), which is not only consistent in both forecast horizons considered, but also in all the energy proxies and sample size. The positive C-T test statistics (see Table 6) and the significant DM test statistics (see Table 7) both confirm this stance, as earlier observed in the in-sample case.

Table 6: Out-of-Sample Forecast Evaluation using Campbell-Thompson test statistics

Energy Consumption Proxies	50%				75%			
	ARDL		$AR(1)$		*ARDL*		$AR(1)$	
	$h=4$	$h=8$	$h=4$	$h=8$	$h=4$	$h=8$	$h=4$	$h=8$
$lenergy_{residence}$	0.9133	0.9120	0.9979	0.9978	0.9375	0.9370	0.9993	0.9993
$lenergy_{commercial}$	0.9407	0.9401	0.9979	0.9979	0.9373	0.9367	0.9993	0.9993
$lenergy_{industrial}$	0.9199	0.9190	0.9980	0.9980	0.9530	0.9522	0.9994	0.9994
$lenergy_{transportation}$	0.9107	0.9097	0.9980	0.9980	0.9329	0.9320	0.9994	0.9994
$lenergy_{aggregate}$	0.9106	0.9096	0.9980	0.9980	0.9413	0.9405	0.9994	0.9994
$lenergy_{petroleum}$	0.9005	0.8993	0.9980	0.9980	0.9282	0.9272	0.9994	0.9994
$lenergy_{nat.gas}$	0.9587	0.9581	0.9979	0.9979	0.9695	0.9693	0.9993	0.9993
$lenergy_{coal}$	0.8983	0.8971	0.9980	0.9980	0.9527	0.9520	0.9994	0.9993
$lenergy_{electricity}$	0.9376	0.9371	0.9979	0.9979	0.9606	0.9602	0.9993	0.9993
$lenergy_{nucl.elect}$	0.9222	0.9233	0.9980	0.9979	0.9807	0.9803	0.9993	0.9993
$lenergy_{renewable}$	0.9003	0.8992	0.9979	0.9979	0.9679	0.9674	0.9993	0.9993

Note: The ADL-MIDAS model is the unrestricted model in the Campbell-Thompson test reported in the table, while ARDL and AR(1) are the restricted model. A positive statistic implies that our preferred model outperforms the competing models, while a negative value implies the reverse.

Table 7: Out-of-Sample Forecast Evaluation using Diebold and Mariano Test statistics

Energy Consumption Proxies	50%				75%			
	ARDL		$AR(1)$		*ARDL*		$AR(1)$	
	$h=4$	$h=8$	$h=4$	$h=8$	$h=4$	$h=8$	$h=4$	$h=8$
$lenergy_{residence}$	4.0960***	3.6286***	5.2960***	3.5372***	5.0183***	4.4343***	7.1146***	4.7603***
$lenergy_{commercial}$	4.0804***	3.3515***	5.2965***	3.5376***	5.0061***	4.4541***	7.1148***	4.7604***
$lenergy_{industrial}$	3.8171***	3.0208***	5.2977***	3.5384***	6.1283***	5.3004***	7.1154***	4.7607***
$lenergy_{transportation}$	3.5105***	2.8334***	5.2981***	3.5386***	5.4864***	5.1645***	7.1155***	4.7608***
$lenergy_{aggregate}$	3.7544***	3.0812***	5.2974***	3.5382***	5.8691***	5.5720***	7.1153***	4.7606***
$lenergy_{petroleum}$	3.3423***	2.6155***	5.2978***	3.5384***	5.6535***	5.3236***	7.1154***	4.7607***
$lenergy_{nat.gas}$	4.1271***	3.1202***	5.2964***	3.5374***	3.9490***	2.8709***	7.1147***	4.7603***
$lenergy_{coal}$	3.5552***	2.9612***	5.2975***	3.5382***	6.1777***	5.6103***	7.1153***	4.7606***
$lenergy_{electricity}$	4.0248***	3.2461***	5.2967***	3.5377***	4.7319***	3.7052***	7.1150***	4.7605***
$lenergy_{nucl.elect}$	2.6480***	2.1292***	5.2969***	3.5380***	5.8679***	4.4131***	7.1152***	4.7606***
$lenergy_{renewable}$	3.4764***	3.0652***	5.2972***	3.5380***	5.1797***	4.0457***	7.1151***	4.7605***

Note: Each cell contains the Diebold and Mariano (DM) test value, which compares ADL-MIDAS with the other models in each column using the squared error as the loss function and its significance implies preference in favour of the row element. Statistical significance of the test values at 1%, 5% and 10% are represented by ***, ** and *, respectively.

Summarily, the combination of the performance(s) of the ADL-MIDAS, in the case of the in-sample and out-of-sample (short and long), overwhelmingly outperforms the other competing models and supports the growth hypothesis as the theoretical underpinning for the energy-growth nexus. Our paper has not only confirmed the energy-growth nexus, but also further established the importance of a MIDAS-based model for improved forecast accuracy of the energy predictors in the growth model.

4.4 Positive Asymmetry vs Negative Asymmetry

Here, we consider the necessity or otherwise, of accounting for asymmetry while predicting the energy-growth nexus. We therefore proceed to compare the positive and negative asymmetric models on the basis of in-sample predictability and forecast accuracy using the Campbell-Thompson test and the Diebold & Mariano test. Both positive asymmetric model (see result in Table 8) and negative asymmetric model (see results in Table 9) seem to predict the energy-growth nexus much in the same way as the earlier case, when the energy series had not been disaggregated into positive and negative series. We find positive (negative) energy changes to impact negatively (positively) on growth, with the results being consistent across data samples and energy consumption proxies in terms of sign, size and magnitude. The coefficients of energy consumption do not differ markedly in both asymmetric models considered across data samples and energy consumption proxies except with regards to the direction. We therefore conduct some formal tests to ascertain the true stance of the role of accounting for asymmetries in the energy-growth nexus.

Table 8: In-Sample Predictability of energy-growth nexus: Positive Asymmetry Model

	Sectors				Aggregate Energy	Components					
	Residential	Commercial	Industrial	Transportation		Petroleum	Natural Gas	Coal	Electricity	Nuclear Electricity	Renewable Energy
50% Sample											
$lgdp_{t-1}$	1.2005*** (0.0958)	1.2036*** (0.0943)	1.2102*** (0.0872)	1.2013*** (0.0954)	1.2005*** (0.0972)	1.2117*** (0.0844)	1.1952*** (0.0968)	1.1978*** (0.0937)	1.1976*** (0.0921)	1.2095*** (0.0907)	1.2012*** (0.0972)
$lgdp_{t-2}$	-0.1975** (0.0962)	-0.2007** (0.0947)	-0.2072** (0.0875)	-0.1983** (0.0957)	-0.1975** (0.0975)	-0.2087** (0.0848)	-0.1923* (0.0971)	-0.1949** (0.094)	-0.1947** (0.0925)	-0.2065** (0.091)	-0.1982** (0.0976)
$lenergy_{t}$	-0.0010*** (0.0002)	-0.0016*** (0.0003)	-0.0029*** (0.0006)	-0.0027*** (0.0005)	-0.0024*** (0.0005)	-0.0041*** (0.0008)	-0.0012*** (0.0002)	-0.0021*** (0.0004)	-0.0022*** (0.0004)	-0.0015*** (0.0003)	-0.0023*** (0.0004)
$lenergy_{t-1}$	0.0002 (23.0837)	0.0001 (23.2081)	0.0002 (67.2883)	-0.0002 (33.7682)	0.0002 (26.4643)	0.0009 (27.2963)	0.0000 (28.0001)	0.0004 (33.8891)	0.0005 (43.7252)	-0.0012 (34.2564)	-0.0006 (32.2377)
75% Sample											
$lgdp_{t-1}$	1.2475*** (0.0842)	1.2320*** (0.0832)	1.2370*** (0.0798)	1.2483*** (0.0848)	1.2420*** (0.0848)	1.2354*** (0.0758)	1.2400*** (0.0853)	1.2466*** (0.0837)	1.2549*** (0.0822)	1.2391*** (0.0827)	1.2547*** (0.0852)
$lgdp_{t-2}$	-0.2450*** (0.0845)	-0.2297*** (0.0835)	-0.2341*** (0.0800)	-0.2458*** (0.0851)	-0.2398*** (0.0851)	-0.2327*** (0.0761)	-0.2374*** (0.0856)	-0.2442*** (0.0840)	-0.2525*** (0.0825)	-0.2365*** (0.0830)	-0.2522*** (0.0855)
$lenergy_{t}$	-0.0006*** (0.0001)	-0.0011*** (0.0002)	-0.0021*** (0.0004)	-0.0017*** (0.0003)	-0.0015*** (0.0003)	-0.0029*** (0.0005)	-0.0007*** (0.0001)	-0.0013*** (0.0003)	-0.0013*** (0.0003)	-0.0010*** (0.0002)	-0.0013*** (0.0003)
$lenergy_{t-1}$	0.0001 (27.3019)	0.0000 (27.2800)	0.0000 (66.2600)	-0.0001 (40.0471)	0.0001 (30.8800)	0.0005 (32.5092)	0.0000 (30.1300)	0.0002 (44.1662)	0.0003 (51.6786)	-0.0002 (40.0358)	-0.0003 (40.9971)

Note: The estimates are obtained from ADL-MIDAS regression models with suppressed intercepts. Statistical significance of the test values at 1%, 5% and 10% are represented by ***, ** and *, respectively. Estimates of the ADL-MIDAS regressions were obtained using the Normalized Exponential Almon lag polynomial $(nealmon)$ weighting scheme in R package.

Table 9: In-Sample Predictability of energy-growth nexus: Negative Asymmetry Model

	Sectors				Aggregate Energy	Components					
	Residential	Commercial	Industrial	Transportation		Petroleum	Natural Gas	Coal	Electricity	Nuclear Electricity	Renewable Energy
50% Sample											
$lgdp_{t-1}$	1.2004*** (0.0933)	1.2029*** (0.0954)	1.2010*** (0.0970)	1.1951*** (0.0968)	1.1952*** (0.0986)	1.1983*** (0.0885)	1.1950*** (0.0968)	1.1935*** (0.0958)	1.1930*** (0.0952)	1.1980*** (0.1001)	1.2065*** (0.0943)
$lgdp_{t-2}$	-0.1974** (0.0936)	-0.1999** (0.0957)	-0.1979** (0.0973)	-0.1921* (0.0972)	-0.1922* (0.0989)	-0.1952** (0.0888)	-0.1925* (0.0971)	-0.1906* (0.0961)	-0.1905** (0.0955)	-0.1951* (0.1004)	-0.2035** (0.0946)
$lenergy_{t}$	0.0010*** (0.0002)	0.0017*** (0.0003)	0.0030*** (0.0006)	0.0029*** (0.0006)	0.0025*** (0.0005)	0.0043*** (0.0009)	0.0012*** (0.0002)	0.0022*** (0.0004)	0.0024*** (0.0005)	0.0017*** (0.0003)	0.0023*** (0.0004)
$lenergy_{t-1}$	0.0001 (44.4989)	-0.0002 (44.3871)	-0.0007 (49.2515)	-0.0005 (29.6998)	-0.0002 (39.1635)	-0.0009 (28.3724)	0.0000 (72.5400)	-0.0002 (26.7552)	-0.0001 (47.2800)	0.0000 (37.0000)	0.0007 (46.3670)
75% Sample											
$lgdp_{t-1}$	1.2478*** (0.0829)	1.2322*** (0.0831)	1.2295*** (0.0845)	1.2432*** (0.0855)	1.2388*** (0.0859)	1.2196*** (0.0798)	1.2390*** (0.0850)	1.2474*** (0.0848)	1.2558*** (0.0841)	1.2410*** (0.0867)	1.2589*** (0.0847)
$lgdp_{t-2}$	-0.2454*** (0.0831)	-0.2296*** (0.0833)	-0.2268*** (0.0848)	-0.2407*** (0.0858)	-0.2362*** (0.0862)	-0.2168*** (0.0800)	-0.2368*** (0.0853)	-0.2450*** (0.0851)	-0.2534*** (0.0843)	-0.2385*** (0.0870)	-0.2565*** (0.0849)
$lenergy_{t}$	0.0006*** (0.0001)	0.0012*** (0.0002)	0.0021*** (0.0004)	0.0018*** (0.0004)	0.0016*** (0.0003)	0.0032*** (0.0006)	0.0007*** (0.0001)	0.0013*** (0.0003)	0.0014*** (0.0003)	0.0011*** (0.0002)	0.0013*** (0.0003)
$lenergy_{t-1}$	0.0000 (53.7044)	-0.0002 (45.7006)	-0.0002 (45.8094)	-0.0002 (39.2998)	-0.0001 (42.3718)	-0.0006 (32.5471)	0.0000 (72.9100)	-0.0002 (37.0363)	-0.0001 (58.0453)	-0.0001 (46.7745)	0.0002 (48.3534)

Note: The estimates are obtained from ADL-MIDAS regression models with suppressed intercepts. Statistical significance of the test values at 1%, 5% and 10% are represented by ***, ** and *, respectively. Estimates of the ADL-MIDAS regressions were obtained using the Normalized Exponential Almon lag polynomial $(nealmon)$ weighting scheme in R package.

Table 10: Campbell-Thompson test: Positive asymmetry vs. Negative asymmetry

	50%			75%		
	In-Sample	Out-of-Sample $(h=4)$	Out-of-Sample $(h=8)$	In-Sample	Out-of-Sample $(h=4)$	Out-of-Sample $(h=8)$
$lenergy_{residence}$	0.0027	0.0022	0.0025	0.0022	0.0019	0.0016
$lenergy_{commercial}$	0.0029	0.0026	0.0027	0.0019	0.0018	0.0016
$lenergy_{industrial}$	-0.0060	-0.0063	-0.0065	-0.0066	-0.0061	-0.0053
$lenergy_{transportation}$	-0.0056	-0.0056	-0.0056	-0.0050	-0.0047	-0.0044
$lenergy_{aggregate}$	-0.0015	-0.0019	-0.0019	-0.0018	-0.0018	-0.0016
$lenergy_{petroleum}$	-0.0091	-0.0093	-0.0093	-0.0113	-0.0115	-0.0116
$lenergy_{nat.gas}$	0.0033	0.0024	0.0024	0.0014	0.0011	0.0009
$lenergy_{coal}$	-0.0020	-0.0018	-0.0015	0.0014	0.0016	0.0018
$lenergy_{electricity}$	-0.0016	-0.0013	-0.0010	0.0014	0.0017	0.0019
$lenergy_{nucl.elect}$	-0.0117	-0.0106	-0.0098	-0.0004	0.0004	0.0012
$lenergy_{renewable}$	0.0039	0.0040	0.0042	0.0034	0.0035	0.0037

Note: The positive asymmetry is used as the reference predictor. A positive C-T value implies that the positive shock outperforms the negative, while the reverse holds if the statistic is negative.

Having confirmed the superiority of the ADL-MIDAS approach over the ARDL and AR(1) models, we considered the possibility of accounting for asymmetries in the ADL-MIDAS model, rather than ignore it. Table 10 shows the comparison of the positive and negative asymmetric models using the Campbell-Thompson test. On the aggregate and sectoral levels, a pattern of performance is observed across the data samples and forecast periods. Negative shocks to aggregate energy consumption, energy consumption by the industrial and the transportation sectors relatively outperform positive shocks in the in-sample, and in the 4- and 8-period ahead out-of-sample forecast horizons. For the residential and commercial sectors, positive shocks outperformed negative shocks across the data sample and forecast periods. On energy consumption by sources, negative shocks to petroleum products supplied consistently outperformed positive shocks, while the reverse was the case for natural gas consumption and renewable energy. Negative shock models seem to be dominant for coal, electricity and nuclear electricity when 50% of full sample was used, while positive shock dominated when 75% of the full sample was used. Overall, negative shock tends to outperform positive shocks about 52% of the time. We therefore proceed to confirm formally the statistical equality or otherwise, of the forecast accuracy of the positive and negative asymmetry using the Diebold and Mariano (DM) test.

Table 11: Diebold and Mariano (DM) test: Positive asymmetry vs. Negative asymmetry

	50%			75%		
	In-Sample	Out-of-Sample $(h=4)$	Out-of-Sample $(h=8)$	In-Sample	Out-of-Sample $(h=4)$	Out-of-Sample $(h=8)$
$lenergy_{residence}$	0.5527	0.4440	0.6494	0.8668	0.7439	0.7686
$lenergy_{commercial}$	0.6297	0.5663	0.7542	0.6845	0.6437	0.7569
$lenergy_{industrial}$	-1.0148	-0.9402	-1.0356	-1.6879	-1.4087	-1.1242
$lenergy_{transportation}$	-1.1653	-1.1888	-1.2104	-1.7119	-1.5898	-1.2726
$lenergy_{aggregate}$	-0.4102	-0.5548	-0.5106	-0.8182	-0.8744	-0.7612
$lenergy_{petroleum}$	-1.1039	-1.2650	-1.1913	-1.6431	-1.9939	-1.5736
$lenergy_{nat.gas}$	0.5703	0.4012	0.5058	0.4366	0.3387	0.3586
$lenergy_{coal}$	-0.8098	-0.5948	-0.6604	0.7675	1.0361	0.8967
$lenergy_{electricity}$	-0.5899	-0.5418	-0.3226	0.7533	0.9928	0.8915
$lenergy_{nucl.elect}$	-1.6251	-1.7120	-1.1854	-0.0981	0.1063	0.2374
$lenergy_{renewable}$	0.7150	0.5044	0.4713	1.1045	0.8266	0.7115

Note: The DM test values test whether the forecast accuracy of the positive asymmetry model differs significantly from that of the negative. The underlying null hypothesis assumes equality of forecast accuracy of both models. A rejection of the null would imply that the forecasts from the two models differ statistically significantly.

Table 11 compares the forecast performance of the positive and negative asymmetric models using the Diebold and Mariano (DM) test. While the values of the Campbell-Thompson tests were generally small, suggesting relative differences, the DM test values in Table 11 puts to rest the question of accounting for asymmetry in the growth forecast model with energy. Across the data sample (50% and 75% of full sample) and energy consumption proxies, both positive and negative asymmetry models did not differ significantly as indicated by the DM test statistics. This suggest equality of the forecast performance of both asymmetric models, which negates the initial disaggregation of the energy series into positive and negative components. Both positive and negative asymmetric models tend to influence growth much in the same way. Consequently, accounting for asymmetry in the energy-growth nexus might not improve forecast accuracy significantly.

5.0 Conclusions

Several studies in the literature have tried to examine the energy-growth nexus on the basis of four distinct theses; the growth thesis, the conservation thesis, the feedback thesis, and the neutrality thesis. While majority of the extant literature focused on providing evidence in support of any of the four theories, they are however divisive in terms of coverage, methodology and the form of energy considered. Also, most literature considered the energy-growth nexus using the restrictive uniform frequency models, which conceals some relevant information due to data aggregation and their analyses are, at best, limited to in-sample (impact) analyses which may not translate into improved out-of-sample forecast performance.

This paper departs from the norm methodologically by adopting a MIDAS-based model, ADL-MIDAS approach, to assess the predictability of this nexus. It also evaluates the relative forecast performance of this approach when compared to the linear time series models such as the autoregressive model and the autoregressive distributed lag (ARDL) model. The forecast evaluation is conducted for both in-sample and out-of-sample periods using multiple data samples and multiple forecast horizons. The focus on the US economy is informed by its huge consumption and/or use of the various forms of energy commodities. Consequently, eleven (11) energy proxies subdivided by sector (residential, commercial, industrial and transportation), aggregate energy consumption and by energy components (petroleum, natural gas, coal, electricity, nuclear electricity and renewable energy). The estimates of the ADL-MIDAS were observed to be consistent in terms of the sign, size and significance of model estimates. The forecast performance of the ADL-MIDAS was compared with two other competing models - ARDL and AR(1) across sample structure and energy proxies. Results revealed overwhelming support in favour of ADL-MIDAS both in the in-sample and the out-of-sample (short and long) forecast periods, which supports the stance that the incorporation of a MIDAS-based model would enhance the predictability of the energy-growth nexus. On role of asymmetry in the energy-growth nexus, while the positive and negative asymmetric models predicted growth in a similar pattern as the model with undecomposed energy series, both models do not differ significantly in their forecast accuracy. The consequent fall out of this finding is that accounting for asymmetries do not significantly improve forecast accuracy in the energy-growth nexus.

References

Adams S, Klobodua EK and Opoku EE (2016) Energy consumption, political regime and economic growth in sub- Saharan Africa. Energy Policy, 96, 36–44

Ahmed M and Azam M (2016) Causal nexus between energy consumption and economic growth for high, middle and low-income countries using frequency domain analysis. Renewable and Sustainable Energy Reviews, 60, 653–678

Akarca AT and Long TV (1979) Energy and employment: a time series analysis of the causal relationship. Resources Energy, 2, 151–162.

Akarca AT and Long TV (1980) On the relationship between energy and GNP: a re-examination. Journal of Energy Development, 5, 326–331

Akinlo AE (2008) Energy consumption and economic growth: Evidence from 11 Sub-Sahara African countries. Energy Economics, 30, 2391–2400.

Alam M, Murad W, Nomanc A and Ozturk I (2016) Relationships among carbon emissions, economic growth, energy consumption and population growth: Testing Environmental Kuznets Curve hypothesis for Brazil, China, India and Indonesia. Ecological Indicators, 70, 466–479

Albu LL, Radu Lupu R and Calin AC (2015) Stock market asymmetric volatility and macroeconomic dynamics in Central and Eastern Europe. Procedia Economics and Finance, 22, 560 – 567

Alper A and Oguz O (2016) The role of renewable energy consumption in economic growth: Evidence from asymmetric causality. Renewable and Sustainable Energy Reviews, 60, 953–959

Alper CE, Fendoglu S and Saltoglu B (2008) Forecasting Stock Market Volatilities Using MIDAS Regression: An Application to the Emerging Markets. MPRA Paper No. 7460.

Andreou E, Ghysels E and Kourtellos A (2013) Should macroeconomic forecasters look at daily financial data? Journal of Business and Economic Statistics, 31, 240 - 251.

Asimakopoulos S, Paredes J and Warmedinger T (2013) Forecasting fiscal time series using mixed Frequency data. Working Paper Series no 1550, the European Central Bank (ECB).

Azam M, Khan A Bakhtyar B and Emirullah C (2015) The causal relationship between energy consumption and economic growth in the ASEAN-5 countries. Renewable and Sustainable Energy Reviews, 47, 732–745

Bai J, Ghysels E and Wright J (2009) State space models and MIDAS regression. Working Paper, NY Fed, UNC and John Hopkins.

Barsoum F and Stankiewicz S (2015) Forecasting GDP growth using mixed-frequency models with switching regimes. International Journal of Forecasting, 31, 33–50

Belke A, Dobnik F and Dreger C (2011) Energy consumption and economic growth: new insights into the cointegration relationship. Energy Economics, 33:5, 782–789.

Berndt E (1978) The Demand for Electricity: Comment and Further Results. MIT Energy Laboratory, WP. No. MIT-EL78-021WP.

Bildirici M (2016) The Relationship Between Hydropower Energy Consumption and Economic Growth. Procedia Economics and Finance, 38, 264 – 270

Bildirici M and Ozaksoy F (2017) The relationship between woody biomass consumption and economic growth: Nonlinear ARDL and causality. Journal of Forest Economics, 27, 60–69

Brinia R, Amara M and Jemmali H (2017) Renewable energy consumption, International trade, oil price and economic growth inter-linkages: The case of Tunisia. Renewable and Sustainable Energy Reviews, 76, 620–627

Campbell JY and Thompson SB (2008) Predicting excess stock returns out of sample: can anything beat the historical average? Review of Financial Studies 21, 1509–1531.

Chen P, Chen S, Hsu C and Chen C (2016) Modeling the global relationships among economic growth, energy consumption and CO2 emissions. Renewable and Sustainable Energy Reviews, 65, 420–431.

Cherni A and Jouini ES (2017) An ARDL approach to the CO2 emissions, renewable energy and growth nexus: Tunisian evidence. International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.08.072

Chiou-Wei S, Zhu Z, Chen S and Hsueh S (2016) Controlling for relevant variables: Energy consumption and economic growth nexus revisited in an EGARCH-M (Exponential GARCH-in-Mean) model. Energy Economics, 109, 391-399

Chontanawat J, Hunt LC and Pierse R (2006) Causality between energy consumption and GDP: evidence from 30 OECD and 78 non-OECD countries. Surrey Energy Economics Centre (SEEC), School of Economics Discussion Papers (SEEDS) 113, University of Surrey

Dhungel KR (2003) Income and price elasticity of the demand for energy: a macrolevel empirical analysis. Pacific and Asian Journal of Energy, 13:2, 73–84.

Diebold F and Mariano R (1995) Comparing Predictive Accuracy. Journal of Business and Economic Statistics, 13, 134-144.

Dogan E (2016) Analyzing the linkage between renewable and non-renewable energy consumption and economic growth by considering structural break in time-series data. Renewable Energy, 99, 1126-1136

Erol U and Yu ES (1987) On the relationship between electricity and income for industrialized countries. Journal of Electricity and Employment, 13, 113-122.

Erol U and Yu ES (1989) Spectral analysis of the relationship between energy and income for industrialized countries. Journal of Energy Development, 13, 113–122.

Foroni C, Marcellino M and Schumacher C (2011) U-MIDAS: MIDAS regressions with unrestricted lag polynomials. Discussion Paper Series 1: Economic Studies 2011, 35, Deutsche Bundesbank, Research Centre.

Ghysels E (2016) MIDAS Matlab Toolbox.

Ghysels EP, Santa-Clara and Valkanov R (2006) Predicting volatility: Getting the most of return data sampled at different frequencies. Journal of Econometrics, 131, 59-95.

Ghysels E, Sinko A and Valkanov R (2007) "MIDAS Regressions: Further Results and New Directions, Econometric Reviews, 26, 53-90.

Ghysels E, Sinko A and Valkanov R (2009) Granger Causality Tests with Mixed Data Frequencies. UNC Discussion Paper.

Hoang TH, Lahiani A and Heller D (2016) Is gold a hedge against inflation? New evidence from a nonlinear ARDL approach. Economic Modelling, 54, 54–66.

Huang S, An H, Gao X and Sun X (2017) Do oil price asymmetric effects on the stock market persist in multiple time horizons? Applied Energy, 185 (2), 1799–1808.

Jung A (2017) Forecasting broad money velocity. North American Journal of Economics and Finance, 42, 421–432

Kilian L (2009) Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. American Economic Review, 19, 1053-1069.

Kraft J and Kraft A (1978) On the relationship between energy and GNP. Journal of Energy Development, 3, 401–403

Marcellino M (2008) A linear benchmark for forecasting GDP growth and inflation? Journal of Forecasting, 27(4), 305–340.

Mehmet AD and Alper A (2017) Renewable and non-renewable energy consumption and economic growth in emerging economies Evidence from bootstrap panel causality, Renewable Energy, 111, 757-763

Menegaki AN and Tugcu CT (2017) Energy consumption and Sustainable Economic Welfare in G7 countries; A comparison with the conventional nexus. Renewable and Sustainable Energy Reviews 69, 892–901

Moosa IA and Burns K (2012) Can exchange rate models outperform the random walk? Magnitude, direction and profitability as criteria. Economia Internazionale, 65, 473–90.

Moosa I (2013) Why is it so difficult to outperform the random walk in exchange rate forecasting? Applied Economics, 45:23, 3340-3346.

Moosa I and Burns K (2014a) The unbeatable random walk in exchange rate forecasting: Reality or myth? Journal of Macroeconomics, 40, 69–81.

Moosa I and Burns K (2014b) Error correction modelling and dynamic specifications as a conduit to outperforming the random walk in exchange rate forecasting. Applied Economics, 46:25, 3107-3118.

Narayan PK and Bannigidadmath D (2015) Are Indian Stock Returns Predictable? Journal of Banking & Finance, 58, 506-531.

Narayan PK and Gupta R (2014) Has oil price predicted stock returns for over a century? Energy Economics. 48, 18–23.

Narayan S (2016) Predictability within the energy consumption–economic growth nexus: Some evidence from income and regional groups. Economic Modelling, 54, 515–521

Narayan S and Doytch N (2017) An investigation of Renewable and Non-renewable Energy Consumption and Economic Growth Nexus using Industrial and Residential Energy Consumption. Energy Economics, In press

Naseria SF, Motamedia S and Ahmadian M (2016) Study of mediated consumption effect of Renewable Energy on Economic Growth of OECD countries. Procedia Economics and Finance, 36, 502 – 509.

Nusair SA (2016) The effects of oil price shocks on the economies of the Gulf Co-operation Council countries: Nonlinear analysis. Energy Policy 91 (2016) 256–267.

Ocal O and Aslan A (2013) Renewable energy consumption–economic growth nexus in Turkey. Renewable and Sustainable Energy Reviews 28, 494–499.

Pesaran MH and Shin Y (1999) An autoregressive distributed lag modelling approach to cointegration analysis, The Ragnar Frisch Centennial Symposium, Cambridge: Cambridge University Press

Pesaran MH, Shin Y, and Smith RJ (2001) Bounds Testing Approaches to the Analysis of Level Relationships. Journal of Applied Econometrics, 16, 289 – 326

Pinzon K (2017) Dynamics between energy consumption and economic growth in Ecuador: A granger causality analysis. Economic Analysis and Policy, In press

Rafindadi AA (2015) Econometric prediction on the effects of financial development and trade openness on the German energy consumption: a startling revelation from the data set. International Journal of Energy & Economic Policy; 5:1, 182–196

Rafindadi AA and Ozturk I (2015) Natural gas consumption and economic growth nexus: is the 10th Malaysian plan attainable within the limits of its resource? Renewable & Sustainable Energy Review, 49, 1221–1232.

Rafindadi AA and Ozturk I (2017) Impacts of renewable energy consumption on the German economic growth: Evidence from combined cointegration test. Renewable and Sustainable Energy Reviews, 75, 1130-1141

Rodríguez-Caballeroa CV and Ventosa-Santaulària D (2016) Energy-growth long-term relationship under structural breaks. Evidence from Canada, 17 Latin American economies and the USA. Energy Economics, 61, 121–134

Saidi K, Rahman MM and Amamri M (2017) The causal nexus between economic growth and energy consumption: new evidence from global panel of 53 countries. Sustainable Cities and Society, 33, 45-56.

Salisu AA. and Isah KO (2017) Revisiting the oil price and stock market nexus: A nonlinear Panel ARDL approach. Economic Modelling, 66, 258-271.

Salisu AA, Isah KO, Oyewole OJ and Akanni LO (2017) Modelling oil price-inflation nexus: The role of asymmetries. Energy, 125, 97-106.

Salisu AA and Ogbonna AE (2017) Improving the predictive ability of oil for inflation: An ADL-MIDAS Approach. Centre for Econometric and Allied Research, University of Ibadan Working Papers Series, CWPS 0025. DOI: 10.13140/RG.2.2.16335.48808

Stern DI and Cleveland CJ (2004) Energy and economic growth. Encyclopaedia of Energy, 2, 35–51.

Stock JH and Watson MW (2003) Forecasting output and inflation: The role of asset prices. Journal of Economic Literature 41, 788-829.

Stock JH and Watson MW (2007) Why has U.S. inflation become harder to forecast? Journal of Money, Credit, and Banking 39, 3-34.

Stock JH and Watson MW (2008) Phillips curve inflation forecast. NBER Working Paper 14322.

Ucan O, Aricioglu E and Yucel F (2014) Energy consumption and economic growth nexus: Evidence from developed countries in Europe. International Journal of Energy Economics and Policy, 45:3, 411–419.

Valadkhani A and Smyth R (2017) How do daily changes in oil prices affect US monthly industrial output? Energy Economics (2017), doi:10.1016/j.eneco.2017.08.009

Yu ES and Choi JY (1985) Causal relationship between energy and GNP: an international comparison. Journal of Energy Development, 10:2, 249–272.

Yu ES and Hwang B (1984) The relationship between energy and GNP: Further results. Energy Economics, 6, 186 -190.

Yu ES, Choi PC and Choi JY (1988) The relationship between energy and employment: a re-examination. Energy Systems Policy, 11, 287–295.

Appendix 1: MIDAS Regressions

A: Flat weight aggregation approach to MIDAS regression

A typical MIDAS regression with Flat weight aggregation can be expressed as (see Asimakopoulos et al., 2013 for further expositions on the subject):

$$\ln GDP_{t+1}^{Q} = \gamma_0 + \gamma_1 \ln engy_t^{Q} + \varepsilon_{t+1}^{Q} \tag{A1}$$

where $\ln engy_t^{Q} = \sum_{i=1}^{N_M} \frac{1}{N_M} \ln engy_{i,t}^{M} = \left(\ln engy_{N_M,t}^{M} + \ln engy_{N_M-1,t}^{M} + \ln engy_{N_M-2,t}^{M} \right) / N_M$, is the quarterly time series for the energy predictor obtained from the monthly data with N_M denoting the number of months in a quarter. Note that Q and M denote quarterly and monthly data frequencies respectively. The model can also be written alternatively as:

$$\ln GDP_{t+1}^{Q} = \gamma_0 + \gamma_1 \sum_{i=1}^{N_M} w_i L_M^i \ln engy_{i,t}^{M} + \varepsilon_{t+1}^{Q} \tag{A2}$$

where w_i are the weights assigned to each month (*i*) while L_M^i is the monthly lag operator. Comparing the two equations and also accounting for the aggregation scheme of the months give:

$$\ln GDP_{t+1}^{Q} = \gamma_0 + \gamma_1 \ln engy_t^{Q} + \gamma_1 \sum_{i=1}^{N_M} \left(w_i - \frac{1}{N_M} \right) L_M^i \ln engy_{i,t}^{M} + \varepsilon_{t+1}^{Q} \tag{A3}$$

If $w_i = 1/N_M$, it suggests that the true weighting scheme is the average/equal weighting scheme and as a consequence, the equation (A3) reduces to (A1) and estimating the latter with OLS will produce unbiased estimates. However, if $w_i \neq 1/N_M$, the OLS regression will have biased estimates due to the omission of the third term in equation (A1).