Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 January 2020 doi:10.20944/preprints202001.0008.v1

Energy Demand Relationship: Theory and Empirical Application. A Short Note

Fakhri J. Hasanov [a,b], Jeyhun I. Mikayilov [a,c]

[a] King Abdullah Petroleum Studies and Research Center, P.O. Box 88550, Riyadh 11672, Saudi Arabia
[b] Research Program on Forecasting, Economics Department, The George Washington University, 2115 G Street, NW, Washington, DC 20052, USA
[c] Department of Statistics and Econometrics, Azerbaijan State University of Economics (UNEC), Istiqlaliyyat Str. 6, Baku AZ1141, Azerbaijan

Abstract

In this short note, the described step-by-step derivations of the industrial energy demand function from the production function framework and provided researchers with two specifications. Then we applied these theoretical specifications to the time series data as empirical analysis. We concluded that theories should be considered at the beginning of the empirical analyses but the data also should be allowed to speak freely. Hence, the main suggestion of this short note is that it would be a better strategy to consider the combination of theory-driven and data-driven approaches in the empirical analyses.

Key words: Theory of energy demand, empirical application, cointegration, theory-driven approach, data-driven approach

© 2020 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 January 2020 doi:10.20944/preprints202001.0008.v1

1. Background

Undoubtedly, energy is one of the key factors of not the only economic activity but also social life. Even, it is impossible to imagine modern life without the existence of energy. Resultantly, countless studies have investigated different aspects of energy including economic, environmental, and demographics. Supply and demand for energy have been key topics in the research to enhance understanding of the human being about them. It would not be wrong if we state that the demand-side of the energy has been examined by a massively huge number of studies. One of the crucial aspects is to understand how various economic, technical and demographic factors shape our demand for energy and this has been the core task of modeling and forecasting exercises from a variety of domains such as academia, government agencies, and the private sector. Therefore, we believe that any research that helps us to understand or enhance our existing knowledge about theoretical and empirical aspects of the energy demand has a value and is worth considering.

Our short note here has two objectives. The first one is to describe the step-by-step derivation of the industrial energy demand function from the production function framework. The second is to provide empirical researchers with two theoretically grounded specifications of the industrial energy demand. We believe there is a value in both objectives. The first one provides researchers with an understanding of how demand for energy, as an input of the production process, is related to other inputs of the production. This understanding could also help them to augment the standard energy demand framework with other economic or social factors in line with their research questions at hand. The second makes the researchers able to select an energy demand specification that best fits their research objectives given the data availability.

As we showed in the empirical section, it is not always possible and or plausible to stick exactly with what theory implies without considering the properties of a given data set. This is because of many the reason as Hendry (2018) discusses including inconsistencies between data measures, proxies and what theory implies caused mainly by either vagueness of theories or data unavailability, mismeasurement or both among others.

Therefore, the key recommendation of this short note is that it is always better to apply the combination of theory-driven and data-driven approaches in the empirical analyses as suggested by seminal studies such as Hendry (2018); Hoover, Johansen, and Juselius (2008); Pagan (2003) among others.

The section below describes the derivation of the energy demand equation and Section 3 applies the derived specification to time series data.

2. Theoretical derivations of the energy demand equations

Cost minimization approach

In order to derive a functional specification for the energy demand function the set up can be formalized as follow. The Cobb-Douglas production function, which related the output to the production factors is given:

$$Q = AK^{\alpha}L^{\beta}E^{\gamma}M^{\delta} \quad (1)$$

Where: Q, K, L, E, and M are output, capital, labor, energy consumption and materials, respectively.

The target is to minimize the total cost, in other words to define the quantities of Q, K, L, E, and M which gives minimum value to the following total cost function:

$$C = p_kK + p_lL + p_eE + p_mM \quad (2)$$

Where: C is total cost and p_k, p_l, p_e, p_m are capital, labor, energy and material prices, respectively.

Then treating the total cost function as an objective function and the production function as a constraint the exercise can be formulated as a constrained optimization problem:

$$C = p_kK + p_lL + p_eE + p_mM \rightarrow Min \quad (3)$$

Subject to

$$Q = f(K, L, E, M) = AK^{\alpha}L^{\beta}E^{\gamma}M^{\delta} \quad (4)$$

Using the Lagrange multipliers method for constrained optimization we can modify our optimization set up as follow (let's call it G):

$$G = C + \lambda(Q - f(K, L, E, M)) \rightarrow Min \quad (5)$$

Which becomes:

$$G = p_kK + p_lL + p_eE + p_mM + \lambda(Q - AK^{\alpha}L^{\beta}E^{\gamma}M^{\delta}) \rightarrow Min \quad (6)$$

Now we have an unconstrained optimization problem with objective function G, as given above.

Based on Lagrange multipliers method next we should calculate the partial derivatives with respect K,L, E, M and λ. The derivatives are given below:

$$\begin{cases} G'_\lambda = Q - AK^\alpha L^\beta E^\gamma M^\delta \\ G'_K = p_k - \lambda A\alpha K^{\alpha-1} L^\beta E^\gamma M^\delta \\ G'_L = p_l - \lambda A\beta K^\alpha L^{\beta-1} E^\gamma M^\delta \\ G'_E = p_e - \lambda A\gamma K^\alpha L^\beta E^{\gamma-1} M^\delta \\ G'_M = p_m - \lambda A\delta K^\alpha L^\beta E^\gamma M^{\delta-1} \end{cases} \quad (7)$$

As a next step in order to find the extremum point we should equate all the above derivatives to zero.

$$\begin{cases} Q - AK^\alpha L^\beta E^\gamma M^\delta = 0 \\ p_k - \lambda A\alpha K^{\alpha-1} L^\beta E^\gamma M^\delta = 0 \\ p_l - \lambda A\beta K^\alpha L^{\beta-1} E^\gamma M^\delta = 0 \\ p_e - \lambda A\gamma K^\alpha L^\beta E^{\gamma-1} M^\delta = 0 \\ p_m - \lambda A\delta K^\alpha L^\beta E^\gamma M^{\delta-1} = 0 \end{cases} \quad (8)$$

Let's take the second terms of each equation to the right side of the equations and then take logs of both sides, then we will get the following system of equations:

$$\begin{cases} lnQ = lnA + \alpha lnK + \beta LnL + \gamma lnE + \delta lnM \\ lnp_k = ln\lambda\alpha A + (\alpha - 1)lnK + \beta lnL + \gamma lnE + \delta lnM \\ lnp_l = ln\lambda\beta A + \alpha lnK + (\beta - 1)lnL + \gamma lnE + \delta lnM \\ lnp_e = ln\lambda\gamma A + \alpha lnK + \beta lnL + (\gamma - 1)lnE + \delta lnM \\ lnp_m = ln\lambda\delta A + \alpha lnK + \beta lnL + \gamma lnE + (\delta - 1)lnM \end{cases} \quad (9)$$

Since, our purpose is to derive a formula for Energy demand function, we should express all other variables in terms of energy demand, namely E.

Let's express K, L and M in terms of E and then consider this expression in the first equation of the last system of equations. Subtracting the fourth equation of the system from the second one and using the property of logarithmic function we get:

$$ln\frac{p_k}{p_e} = ln\frac{\alpha}{\gamma} - lnK + lnE \quad (10)$$

Modifying this equation, a little we can derive a formula relating K to E:

$$lnK = -ln\frac{p_k}{p_e} + ln\frac{\alpha}{\gamma} + lnE \text{ or } lnK = ln\frac{p_e}{p_k}\frac{\alpha}{\gamma}E \text{ and therefore: } K=\frac{p_e}{p_k}\frac{\alpha}{\gamma}E \quad (11)$$

In a similar way subtracting the fourth equation consequently from the third and fifth equations

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 January 2020 doi:10.20944/preprints202001.0008.v1

we get the formulas relating L and M to E:

$$L=\frac{p_e}{p_l}\frac{\beta}{\gamma}E \text{ (12) and } M=\frac{p_e}{p_m}\frac{\delta}{\gamma}E \quad (13)$$

Considering (11), (12) and (13) in the first equation of the system (9) we end up with:

$$lnQ = lnA + \alpha\ln(\frac{p_e}{p_k}\frac{\alpha}{\gamma}E) + \beta\ln(\frac{p_e}{p_l}\frac{\beta}{\gamma}E) + \gamma lnE + \delta\ln(\frac{p_e}{p_m}\frac{\delta}{\gamma}E) \quad (14)$$

Using the properties of the logarithmic function (14) can be modified as follow:

$$lnQ = lnA + \alpha\ln\left(\frac{p_e}{p_k}\right) + \alpha\ln\left(\frac{\alpha}{\gamma}\right) + \alpha lnE + \beta\ln\left(\frac{p_e}{p_l}\right) + \beta\ln\left(\frac{\beta}{\gamma}\right) + \beta lnE + \gamma lnE + \delta\ln\left(\frac{p_e}{p_m}\right) + \delta\ln\left(\frac{\delta}{\gamma}\right) + \delta lnE \quad (15)$$

Rearranging (15) and combining the constant terms and coefficient of *lnE* we get:

$$lnQ = \left[lnA + \alpha\ln\left(\frac{\alpha}{\gamma}\right) + \beta\ln\left(\frac{\beta}{\gamma}\right) + \delta\ln\left(\frac{\delta}{\gamma}\right)\right] + \alpha\ln\left(\frac{p_e}{p_k}\right) + \beta\ln\left(\frac{p_e}{p_l}\right) + \delta\ln\left(\frac{p_e}{p_m}\right) + [\alpha + \beta + \gamma + \delta]lnE \quad (16)$$

Let's find *lnE* from this expression:

$$[\alpha + \beta + \gamma + \delta]lnE = lnQ - \left[lnA + \alpha\ln\left(\frac{\alpha}{\gamma}\right) + \beta\ln\left(\frac{\beta}{\gamma}\right) + \delta\ln\left(\frac{\delta}{\gamma}\right)\right] - \alpha\ln\left(\frac{p_e}{p_k}\right) - \beta\ln\left(\frac{p_e}{p_l}\right) - \delta\ln\left(\frac{p_e}{p_m}\right) \quad (17)$$

Taking into account the fact that $(-\ln(\frac{x}{y})) = ln\frac{y}{x}$ we can modify (17) as follow:

$$[\alpha + \beta + \gamma + \delta]lnE = lnQ + \left[-lnA + \alpha\ln\left(\frac{\gamma}{\alpha}\right) + \beta\ln\left(\frac{\gamma}{\beta}\right) + \delta\ln\left(\frac{\gamma}{\delta}\right)\right] + \alpha\ln\left(\frac{p_k}{p_e}\right) + \beta\ln\left(\frac{p_l}{p_e}\right) + \delta\ln\left(\frac{p_m}{p_e}\right) \quad (18)$$

Now in order to get expression for *lnE,* let's divide both sides of (18) by the coefficient of *lnE*:

$$lnE = -\frac{1}{[\alpha+\beta+\gamma+\delta]}lnQ + \frac{\left[-lnA+\alpha\ln\left(\frac{\gamma}{\alpha}\right)+\beta\ln\left(\frac{\gamma}{\beta}\right)+\delta\ln\left(\frac{\gamma}{\delta}\right)\right]}{[\alpha+\beta+\gamma+\delta]} + \frac{\alpha}{[\alpha+\beta+\gamma+\delta]}\ln\left(\frac{p_k}{p_e}\right) + \frac{\beta}{[\alpha+\beta+\gamma+\delta]}\ln\left(\frac{p_l}{p_e}\right) + \frac{\delta}{[\alpha+\beta+\gamma+\delta]}\ln\left(\frac{p_m}{p_e}\right) \quad (19)$$

In order to have energy, capital, labor and material prices as individual variables in the right hand side let's modify the right side of (19):

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 January 2020 doi:10.20944/preprints202001.0008.v1

$$lnE = -\frac{1}{[\alpha+\beta+\gamma+\delta]} lnQ + \frac{\left[-lnA + \alpha\ln\left(\frac{\gamma}{\alpha}\right) + \beta\ln\left(\frac{\gamma}{\beta}\right) + \delta\ln\left(\frac{\gamma}{\delta}\right)\right]}{[\alpha+\beta+\gamma+\delta]} +$$

$$\frac{\alpha}{[\alpha+\beta+\gamma+\delta]}\ln p_k + \frac{\beta}{[\alpha+\beta+\gamma+\delta]}\ln p_l + \frac{\delta}{[\alpha+\beta+\gamma+\delta]}\ln p_m + \frac{(-\alpha-\beta-\delta)}{[\alpha+\beta+\gamma+\delta]} p_e \qquad (20)$$

The equation (20) can be expressed as follow:

$$lnE = \alpha_0' + \alpha'\ln p_k + \beta'\ln p_l + \delta'\ln p_m + \gamma' lnp_e + \eta' lnQ \qquad (21)$$

Where: $\alpha_0' = \frac{\left[-lnA+\alpha\ln\left(\frac{\gamma}{\alpha}\right)+\beta\ln\left(\frac{\gamma}{\beta}\right)+\delta\ln\left(\frac{\gamma}{\delta}\right)\right]}{[\alpha+\beta+\gamma+\delta]}$, $\alpha' = \frac{\alpha}{[\alpha+\beta+\gamma+\delta]}$, $\beta' = \frac{\beta}{[\alpha+\beta+\gamma+\delta]}$, $\delta' = \frac{\delta}{[\alpha+\beta+\gamma+\delta]}$, $\gamma' = \frac{(-\alpha-\beta-\delta)}{[\alpha+\beta+\gamma+\delta]}$, $\eta' = -\frac{1}{[\alpha+\beta+\gamma+\delta]}$

As a results we derived a formula for energy demand which expresses it as function of capital, labor, material and energy prices and total output.

As discussed in Nordhaus (1975) the demand functions of the economy for each product can be written as:

$$Q_i = f^i(P_1, P_2, P_3, \dots, P_n, Y), i = 1, \dots, n \qquad (22)$$

Where P_i's are prices and Y is the total income. This function (in logs) can be written explicitly as:

$$lnQ = \theta_0 + \theta_1\ln p_k + \theta_2\ln p_l + \theta_3\ln p_m + \theta_4 lnp_e + \theta_5 lnY \qquad (23)$$

If we substitute lnQ in (21) with its expression in (23) and rearrange according to the coefficients we get:

$$lnE = (\alpha_0' + \eta'\theta_0) + (\alpha' + \eta'\theta_1)\ln p_k + (\beta' + \eta'\theta_2)\ln p_l + (\delta' + \eta'\theta_3)\ln p_m + (\gamma' + \eta'\theta_4) lnp_e + \eta'\theta_5 lnY \qquad (24)$$

(24) can be written is the following simpler form:

$$lnE = \alpha_0'' + \alpha''\ln p_k + \beta''\ln p_l + \delta''\ln p_m + \gamma'' lnp_e + \eta'' lnY \qquad (25)$$

Where

$$\alpha_0'' = \alpha_0' + \eta'\theta_0, \alpha'' = \alpha' + \eta'\theta_1, \beta'' = \beta' + \eta'\theta_2, \delta'' = \delta' + \eta'\theta_3,$$
$$\gamma'' = \gamma' + \eta'\theta_4 \text{ and } \eta'' = \eta'\theta_5$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 January 2020 doi:10.20944/preprints202001.0008.v1

Equation (25) is the derived energy demand equation and it shows that the demand for energy is a function of prices of production factors and the total income.

Apparently, the difference between equations (21) and (25) is that the former links energy demand to output while the latter considers it as a function of income alongside the price measures.

Obviously, the framework linking energy demand to its own price and income, which is the widely-used specification in empirical analysis, is a modification of equation (25), where prices of other inputs are omitted.

Note that depending on the research objective, equations (21) and (25) can be further modified theoretically to include socio-economic variable(s) of interest in them. For example, Beenstock, and Dalziel (1986) modified equation (21) by replacing industrial production with industrial capital stock while Hasanov (2018) modified equation (25) by including demographic variable in it.

3. The energy demand equations in empirical analysis

In this section, as an empirical application, we econometrically estimated equations (21) and (25) using time series data over the 1980-2018 period[1]. Equally, one may apply the equations to the cross-sectional or panel data but given availability and accessibility issues of the required data across the countries or across the industrial branches, especially when it comes to price and cost of capital data, we opted to the time series estimation here. As a country, we selected Saudi Arabia because of the convenience of obtaining the required time-series data. Of course, different countries can be selected as a case study. We considered electricity consumption of the Saudi Arabian non-oil manufacturing sector as a measure of energy and again, different energy products and in different industry branches can be considered.

Table 1 presents variables we used in the empirical analysis and their descriptions/definitions.

[1] The period is dictated by the availability of the time-series data.

Table 1. Variables and their descriptions

Variable	Notation	Description/definition	Source
Electricity consumption	E	Demand for electricity in industrial sector, mtoe.	IEA
Cost of capital in real terms	P_K	This is the United States seven-year Treasury note yield, at constant maturity, adjusted for the US inflation rate.	OEGEM
Average annual wage in real terms	P_L	$P_L = \frac{W_N}{CPI} * 100$ W_N is average annual wage in nominal term, which is calculated as below. $W_N = \frac{ER}{ET}$ ER is total earnings in thousand SAR. ET is total employment in thousand persons. CPI is Consumer Price Index, 2010=100.	Own calculation using GSTAT data
Price of intermediate goods in real terms	P_M	$P_M = \frac{P_{MAN}}{P_{NOIL}} * 100$ P_{NOIL} is deflator of the non-oil value added, which is calculated as below. $P_{NOIL} = \frac{QN_{NOIL}}{QR_{NOIL}} * 100$ QN_{NOIL} is nominal value added in non-oil manufacturing in million SAR. QR_{NOIL} is real value added in non-oil manufacturing in million SAR at 2010 prices.	
Price of electricity consumed in industry in real terms	P_E	$P_E = \frac{P_N}{P_{MAN}} * 100$ P_N is nominal price of the electricity consumed in industry in SAR/toe. P_{MAN} is deflator of the non-oil manufacturing value added, which is calculated as below. $P_{MAN} = \frac{Q_N}{Q} * 100$ Q_N is nominal value added in non-oil manufacturing in million SAR.	Own calculation using GSTAT data
Output in non-oil manufacturing in real terms	Q_O	This is sum of value added and intermediate consumption both in non-oil manufacturing (excluding petroleum manufacturing) in million SAR at 2010 prices.	GSTAT and OEGEM
Value added in non-oil manufacturing in real terms	Q_V	Value added in manufacturing excluding petroleum manufacturing, million SAR at 2010 prices.	GSTAT

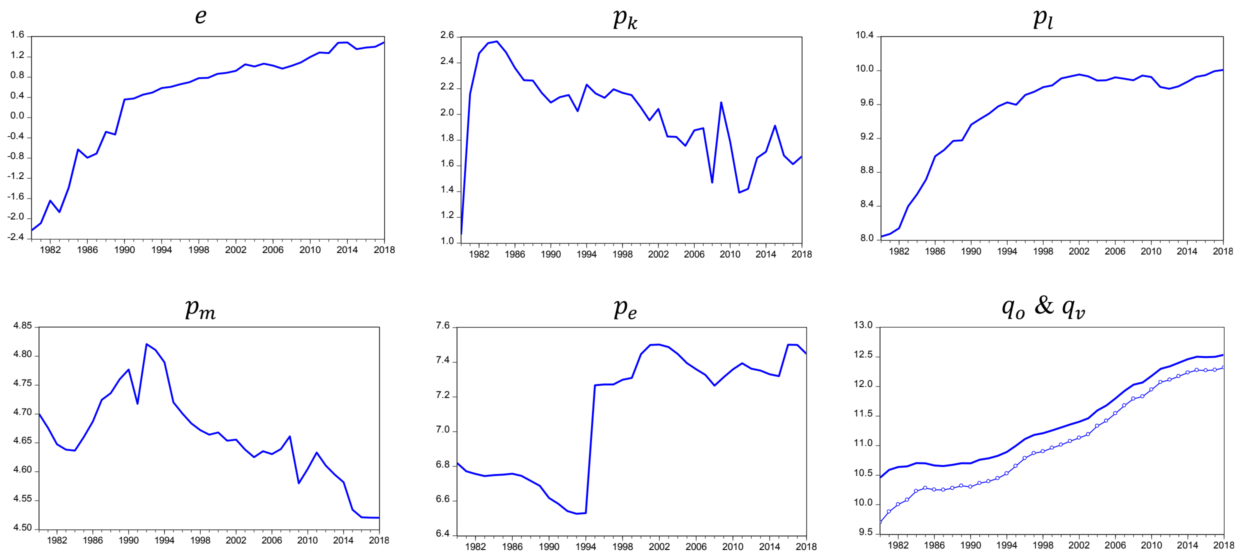
Note: IEA=International Energy Agency; GSTAT=General Authority for Statistics of Saudi Arabia; OEGEM=Oxford Economics Global Economic Model database.

We used the United States seven-year Treasury note yield instead of using Saudi long-run interest rate because of such long-run interest rate data for Saudi Arabia does not available to the best of our knowledge. There are interest rates on Saudi riyal deposits for one, three, six and 12 months maturities in Saudi Arabian Monetary Authority Annual statistics. However, these interest rates are (a) short-run interest rates and (b) available only from 1997 (SAMA, 2019). It is worth noting that earlier studies on Saudi economy also faced the issue of obtaining accurate interest rate data with long sample size and hence ended up with using foreign countries, especially the US, interest rates (see El Mallakh, 1982; Chatah, 1983; Darrat, 1984, 1986; Al-Bassam, 1990; Al Rasasi and Banafea, 2018; Al Rasasi and Qualls, 2019).

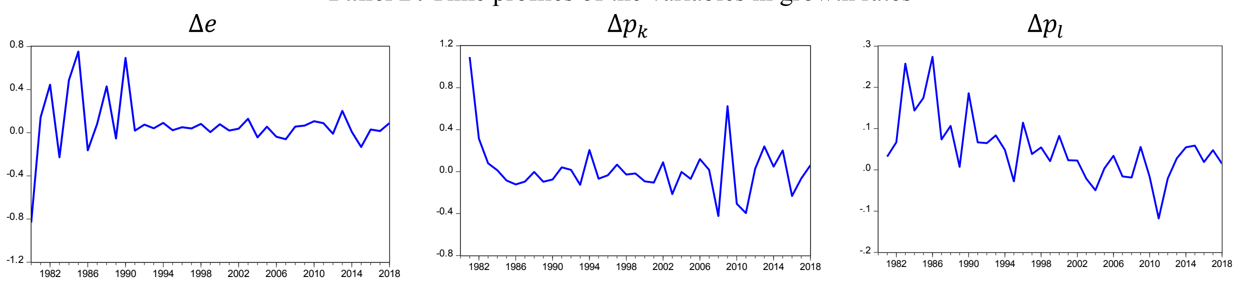
Figure 1 below illustrates the time profiles of the variables in the table expressed in natural logarithm and growth rates. We did not discuss the data here because it is not our main purpose in this note. Also, without providing detailed information about the stochastic properties of the variable, note that they all are unit root processes at their log level and stationary at their growth rate form. In other words, they all follow the integrated order of one, i.e., I (1) process. Details of the unit root test results are available from the authors under request.

Figure 1. Time profiles of the variables

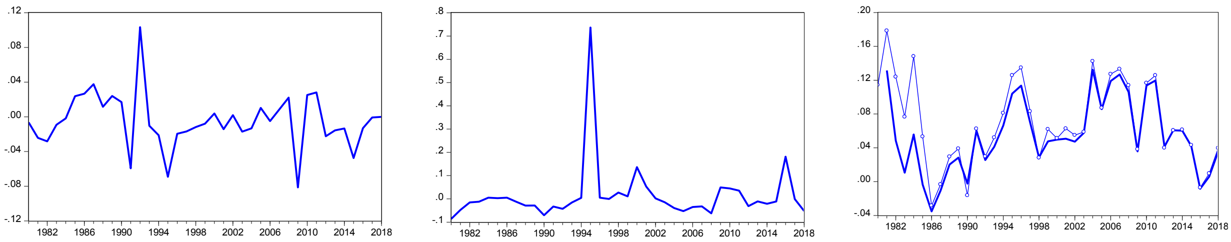
Panel A. Time profiles of the variables in logs



Panel B. Time profiles of the variables in growth rates



Δp_m Δp_e Δq_o & Δq_v



Since the variable are I(1) processes and to be consistent with the theory above, we tested whether the variables are cointegrated and estimated level/long-run relationship among them. For this exercise, we used Fully Modified Ordinary Least Squares (FMOLS) estimator developed by Phillips and Hansen (1990) and Phillips and Loretan (1991) as this method avoids over-parametrization and addresses the endogeneity issue. Tables 2 and 3 documents the estimation and testing results for equations (21) and (25) with different options.

Table 2. FMOLS estimation and test results for equation (21).

	Panel A. Estimation of equation (21)		**Panel B.** Estimation of equation (21) without p_k		**Panel C.** Estimation of equation (21) only with p_e and q_o	
Estimated long-run elasticities						
Regressor	***Coef.***	***P-value***	***Coef.***	***P-value***	***Coef.***	***P-value***
p_k	0.097	0.643	–	–	–	–
p_l	1.453	0.000	1.505	0.000	–	–
p_m	1.980	0.086	1.493	0.087		
p_e	-0.224	0.200	-0.285	0.075	0.630	0.392
q_o	0.608	0.001	0.532	0.000	0.892	0.021
SER	0.171098		0.168028		0.604501	
R^2_{ADJ}	0.969779		0.970854		0.622769	
Post-estimation test results						
Q	0.368	0.544	0.316	0.574	0.809	0.000
JB	0.369	0.832	0.849	0.654	6.330	0.042
F for p_k	–	–	0.200	0.658	0.567	0.456
F for p_l	–	–	–	–	507.355	0.000
F for p_m	–	–	–	–	51.613	0.000
Cointegration test results						
EG^a_τ	-5.492	0.027	-5.545	0.011	-2.623	0.440
EG^a_z	-34.067	0.024	-34.239	0.009	-5.674	0.852

Notes:
e is the dependent variable in the estimations;
SER is standard error of regression.
R^2_{ADJ} is adjusted R-Squared, i.e., Determination coefficient.
Q is the Q-statistic of the first order auto-correlation coefficient with the null hypothesis that the residuals are not correlated.
JB is the Jarque-Bera statistic of Normality test with the null hypothesis that the residuals are normally distributed.
F is the F-statistics of the Omitted Variable test with the null hypothesis of a tested variable can be omitted.
EG^a_τ and EG^a_z are the degree of freedom adjusted Engle-Granger tau- and z-statistics.
Coef. and *P-value.* mean coefficient and its probability value.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 January 2020 doi:10.20944/preprints202001.0008.v1

For simplicity, intercepts are not reported.
Estimation period: 1980-2018.

Table 3. FMOLS estimation and test results for equation (25).

	Panel D. Estimation of equation (25)		**Panel E.** Estimation of equation (25) without p_k		**Panel F.** Estimation of equation (25) only with p_e and q_o	
Estimated long-run elasticities						
Regressor	***Coef.***	***P-value***	***Coef.***	***P-value***	***Coef.***	***P-value***
p_k	0.030	0.885	–	–		
p_l	1.368	0.000	1.439	0.000		
p_m	1.838	0.111	1.621	0.056		
p_e	-0.218	0.217	-0.249	0.105	0.364	0.727
q_v	0.540	0.002	0.499	0.000	1.019	0.029
SER	0.170074		0.169650		0.608458	
R^2_{ADJ}	0.970140		0.970289		0.672940	
Post-estimation test results						
Q	0.594	0.441	0.669	0.413	31.490	0.000
JB	0.164	0.921	0.922	0.631	0.687	0.709
F for p_k	–	–	0.029	0.865	0.808	0.375
F for p_l	–	–	–	–	399.451	0.000
F for p_m	–	–	–	–	67.250	0.000
Cointegration test results						
EG^a_τ	-5.351	0.035	-5.378	0.015	-1.853	0.795
EG^a_z	-33.126	0.031	-33.258	0.013	-5.137	0.881

Notes:
e is the dependent variable in the estimations;
SER is standard error of regression.
R^2_{ADJ} is adjusted R-Squared, i.e., Determination coefficient.
Q is the Q-statistic of the first order auto-correlation coefficient with the null hypothesis that the residuals are not correlated.
JB is the Jarque-Bera statistic of Normality test with the null hypothesis that the residuals are normally distributed.
F is the F-statistics of the Omitted Variable test with the null hypothesis of a tested variable can be omitted.
EG^a_τ and EG^a_z are the degree of freedom adjusted Engle-Granger tau- and z-statistics.
Coef. and *P-value.* mean coefficient and its probability value.
For simplicity, intercepts are not reported.
Estimation period: 1980-2018.

Three useful observations can be extracted from Tables 2 and 3.

- Panels A and D report the results of the estimation of equation (21) and (25). All the explanatory variables have theoretically expected signs. Apparently, the cost of capital is highly statistically insignificant in both estimations. These are the cases, i.e., Panels A and D, where we imposed the theory of energy demand on the data and ignored information coming from data, i.e., the insignificance of the capital cost. Our results were theory-driven and we positioned ourselves at the upper part of Pagan's trade-off curve (Hendry, 2018; Pagan, 2003).

- Panels B and E report the results of the estimation of equation (21) and (25) both without the cost of capital. All the explanatory variables have theoretically expected signs and are statistically significant. Apparently, the cost of capital was excluded from the estimations as it was statistically insignificant. These are the cases, i.e., Panels B and E, where our estimation results were derived from nesting the theory of energy demand with data. In other words, we used the theory but not imposed it on the data and we also considered information coming from data, i.e., the insignificance of the capital cost and the result of the omitted variable test on the variable. Put differently, our results were both theory-driven and data-driven and we positioned ourselves in the middle part of Pagan's trade-off curve (Hendry, 2018; Pagan, 2003).
- Panels C and F report the results of the estimation of equation (21) and (25) both only with energy price and income. It is worth stating that considering energy demand as a function of its own price and income is the widely-used specification in the literature. These are the cases, i.e., Panels C and F, where we ignored theoretically predicted variables, i.e., prices of capital, labor and intermediate consumption in equations (21) and (25), without testing their statistical significance of them to see whether they can contribute to the explanation of the energy demand pattern. Obviously, we missed some useful information, which could come from the prices of the labor and intermediate consumption as the omitted variable tests indicated.

Thus, the question occurs here is that which option above should we prefer to in our energy demand modeling for policy analysis and or forecasting? In our explanations below, we tried to justify them theoretically and statistically.

Undoubtedly, no one would prefer to model energy demand as we did in Panels C and F because such specification of the energy demand yielded theoretically inconsistent (i.e., omissions of the variables predicted by the theory, the wrong sign of energy price) and statistically poor (insignificance of the energy price, higher SER and lower R^2_{ADJ} compared to other options, auto-correlation in the residuals, omitted variable tests results) results.

Perhaps, theory-driven researchers, such as CGE or DSGE modelers would prefer to modeling energy demand as we did in Panels A and D. The researchers, who are in favor of hybrid modeling that is nesting theory-driven and data-driven approaches would prefer modeling energy demand as we did in Panels B and E.

With respect to both type of researchers, we believe that the modeling energy demand like it was done in Panels B and E is the best option to consider. Statistically, in our case, this is because of the fact that the specifications without the cost of capital (Panels B and E) outperform those with the cost of capital (Panels A and D) as the former ones have lower standard errors and a higher level of approximations compared to the latter ones as well as results of the omitted variable test indicate no information loss by excluding the cost of the capital variable. Theoretically, the following points should be recalled briefly.

- Sometimes a variable articulated by the theory cannot be exactly measured in practice due to data availability issue and proxies provide poor estimates and thus they do not help us to approximate Data Generation Process (DGP) of the variable under interest. This is exactly what we faced in our analysis here. The theory above articulates cost of capital as an explanatory variable of the industrial energy demand. However, we could not find exactly the cost of capital data for the non-oil manufacturing[2]. We proxied it following earlier studies but it appeared that the selected proxy did not contribute to the DGP of the industrial energy demand and it was statistically insignificant.
- Sometimes theory is vogue about the variables when it comes to their consideration in the empirical analysis and the considered variable may not contribute to the DGP. For example, money demand theories consider income as a scale variable in explaining the behavior of money balance. However, it is not quite clear which income measure should be considered in the empirical analysis. Therefore, from GDP to retail turnover, consumption, government expenditure, and industrial production index are considered in the empirical analyses of money demand (Sriram, 2001).
- All theories are based on certain assumptions and these assumptions may not be held in a given country for a time period considered.
- Theories do not tell us anything about structural breaks and location shifts which can play a considerable role in explaining a given process at hand.

Details of the above-listed points alongside others are comprehensively discussed in Hendry (2018) and in the references therein.

To conclude, theories should be considered at the beginning of the empirical analyses but the data also should be allowed to speak freely.

[2] It can be argued that it is not the case just for Saudi Arabia and even for many developing and perhaps developed countries, the cost of capita data is not available for the branches of industry.

Thus, the main recommendation of this short note here is that it would be a better strategy to consider the combination of theory-driven and data-driven approaches in the empirical analyses.

Acknowledgments

The views expressed in this paper are the authors' and do not necessarily represent the views of their affiliated institutions.

References

Al-Bassam, K.A. (1990). Money Demand and Supply in Saudi Arabia: An Empirical Analysis. PhD Thesis, University of Leicester.

Al Rasasi, M., Banafea, W.A. Estimating Money Demand function in Saudi Arabia: Evidence from Cash in Advance Model. Saudi Arabian Monetary Agency Working Paper, WP/18/4.

Al Rasasi, M., Qualls J. 2019. Revisiting the Demand for Money in Saudi Arabia. Saudi Arabian Monetary Authority, SAMA Working Paper, WP/19/01.

Beenstock, M. and Dalziel, A. 1986. The demand for energy in the UK: A general equilibrium analysis. Energy Economics 8(2), 90-98.

Chatah, M. 1983. A Monetary Framework for Saudi Arabia', Unpublished Ph.D. Dissertation, University of Texas, Austin.

Cobb, C. W., Douglas, P. H. 1928. A Theory of Production. American Economic Review 18, 139–165.

Darrat, A.F. 1984. The money demand relationship in Saudi Arabia: An empirical investigation. Journal of Economic Studies, 11(3), 43-50.

Darrat, A.F. 1986. The demand for money in some major OPEC members: regression estimates and stability results. Applied Economics, 18(2), 127-142

El-Mallakh, Regaei. 1982. The Saudi Arabia Rush to Development, London: Croom Helm Ltd.

Hasanov, Fakhri. 2018. Theoretical Framework for Industrial Electricity Consumption Revisited: Empirical Analysis and Projections for Saudi Arabia. King Abdullah Petroleum Studies and Research Center, Discussion Papers, Doi: 10.30573/KS--2019-DP66.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 January 2020 doi:10.20944/preprints202001.0008.v1

Hendry, David F. "Deciding between alternative approaches in macroeconomics." International Journal of Forecasting 34, no. 1 (2018): 119-135.

Holtedahl, P., Joutz, F.L. 2004. Residential electricity demand in Taiwan. Energy Economics 26, 201–224.

Hoover, Kevin D., Soren Johansen, and Katarina Juselius. 2008. "Allowing the Data to Speak Freely: The Macroeconometrics of the Cointegrated Vector Autoregression." American Economic Review 98(2): 251-55.

Joutz, F. and Gardner, T. 1996. Economic Growth, Energy Prices and Technological Innovation. Southern Economic Journal 62 (3), 653-666.

Nordhaus, W. D. 1975. The demand for energy: an international perspective. Cowles Foundation Discussion Paper no 405.

Pagan, Adrian (2003). "Report on modelling and forecasting at the Bank of England/Bank's response to the Pagan report." Bank of England. Quarterly Bulletin 43, no.1: 60.

Phillips, P.C.B, Hansen, B.E. 1990. Statistical inference in instrumental variables regression with I(1) processes, Review of Economics Studies, 57 (1), 99-125.

Phillips, Peter CB, and Mico Loretan. "Estimating long-run economic equilibria." The Review of Economic Studies 58, no. 3 (1991): 407-436.

Silk, J., Joutz, F. 1997. Short and long-run elasticities in US residential electricity demand: A co-integration approach. Energy Economics 19, 493–513.

Sriram, S. 2001. A survey of recent empirical money demand studies. IMF Staff papers 47(3), 334-365.