Selçuk Ün. Sos. Bil. Ens. Der. 2018; (39): 191-200
- İstatistik / Araştırma -

Modeling The Causality Relationships Between Gdp/Gni and Electricity Consumption According to Income Levels of Countries By Generalized Estimating Equations*

Harun YONAR**
Neslihan İYİT***

ABSTRACT

Gross domestic product (GDP) and energy consumption in the economic evaluations of countries are seen as two basic concepts of development. The need for energy resources in recent years has brought countries closer to technology, but in some cases, it causes problems such as wars. It is also important to determine the economic feasibility of energy consumption as well as the feasibility of many aspects such as the origin, usage, and necessity of energy. When we look at the crises that have taken place in the last 20 years, it is once again seen that energy is the dynamism and indispensable necessity of the countries. If we look at the effect of the consumed energy on the country's economy, the first economic variable will be GDP. Interpretation and evaluation of GDP, which reveals steady growth, will give effective results on economic indicators of the country. A lot of research has been done in the literature between the amount of energy consumption (according to the sectors, type of energy used, supply, and etc.) and the GDP which is the most important indicator of the country's economy. The final relationship between these two variables has been examined in details for different countries and energy concepts. In previous studies, it is sometimes observed that energy consumption is a cause of GDP or vice versa, and sometimes a two-way causality between them is determined. On the other hand, a causality relationship can not be always determined between the variables. In this case a suitable regression model can be established without looking for causality.

In this study, the causality relationship between the GDP values, categorized by five income levels, and the energy consumptions of the countries between 1980 and 2014 is determined by using the Granger causality test. When we look at the results of the causality test, we find that only one causality relationship exists between high income level countries by GDP and the energy consumption of them. According to the causality test result, dependent and independent variable are determined before generalized estimating equations (GEE) method is used for modelling the data. In GEE method, the smallest values of QIC and QICC information criteria are found in the direction of causality relationships. The same causality assessment is done between gross national incomes (GNI) of countries categorized by income levels and energy consumptions, and it is concluded that the GEE models established according to the causality relationship direction are much better fit to the data. These findings obtained from this study suggests that causality test is a guide for us when we have insufficient knowledge in determining dependent and independent variables before fitting regression models to the data.

Keywords: Gross Domestic Product, Gross National Income, Energy Consumption, Granger Causality Test, Generalized Estimating Equations.

Ülkelerin Gelir Düzeylerine Göre Gayri Safi Yurtiçi Hasılaları ile Elektrik Tüketimleri Arasındaki Nedensellik İlişkilerinin Genelleştirilmiş Tahmin Denklemleri ile Modellenmesi

ÖZ

Ülkelerinin ekonomik değerlendirmelerinde gayri safi yurtiçi hasıla (GSYİH) ve enerji tüketimi kalkınmanın iki temel unsuru olarak görülmektedir. Son yıllarda enerji kaynaklarına duyulan ihtiyaç, ülkeleri teknolojiye yaklaştırdığı gibi bazende onları yok edebilecek savaşlara sebep olabilmektedir. Enerjinin temini, kullanımı ve ihtiyaç nedeni gibi bir çok açıdan ele alınabilirliği kadar enerji tüketiminin ekonomik karşılığının tespit edilebilmesi de son derece önemlidir. Son 20 yılda enerji kaynaklı yaşanan krizler incelendiğinde, enerjinin ülkelerin vazgeçilmez bir ihtiyacı ve dinamiği olduğu gerçeği bir kez daha görülmektedir. Tüketilen enerjinin ülke ekonomisine etkisininin ne olduğuna bakılacak olursa, ilk karşılaşılacak ekonomik değişken GSYİH olacaktır. Bir ülkenin GSYİH'nın ekonomik göstergelerdeki yorumlanması ve değerlendirilmesi son derece etkin sonuçlar vermektedir. Literatürde farklı ülkeler için enerji tüketimi miktarları (sektörlere göre, kullanılan enerji türlerine göre, teminine göre ve benzeri) ile ülke ekonomisinin en önemli göstergelerinden biri olan GSYİH değerleri arasındaki nedensellik ilişkilerini incelemek üzere yapılan bir çok çalışma mevcuttur. Yapılan çalışmalarda farklı ülkeler için kimi zaman enerji tüketiminin GSYİH'nın bir nedeni

*This study is a part of Harun Yonar's Ph.D. Thesis entitled "Aktüeryal Risk Değerlendirilmesinde Genelleştirilmiş Lineer Modeller" supervised by Assist.Prof.Dr.Neslihan İyit continuing in Statistics Department, Graduate School of Natural Sciences, Selçuk University. Neslihan İyit is the major author of this study and also Harun Yonar's Ph.D. supervisor. An earlier version of this study is presented by the same authors at II.International Academic Research Congress (INES 2017), 18-21 October 2017, Antalya.

**Res.Ass. Selçuk University, orcid no: 0000-0003-1574-3993, hyonar@selcuk.edu.tr
***Assist.Prof.Dr. Selçuk University, corresponding author, orcid no: 0000-0002-5727-6441, niyit@selcuk.edu.tr

Makalenin Gönderim Tarihi: 28.12.2017; Makalenin Kabul Tarihi: 25.01.2018

olduğu, kimi zamansa GSYİH'nın enerji tüketiminin bir nedeni olduğu gösterilmiştir. Bu iki değişken arasında çift yönlü bir nedensellik ilişkisinin saptandığı çalışmalar da mevcuttur. Öte yandan, değişkenler arasında her zaman bir nedensellik ilişkisi tespit edilemese de, uygun bir regresyon modeli kurularak değişkenler arasındaki açıklayıcılık incelenebilir. Bununla birlikte, kurulan her regresyon modelinde de bir nedensellik ilişkisi aramak doğru değildir.

Bu çalışmada öncelikle 1980-2014 yılları arasında 5 farklı gelir düzeyine göre kategorize edilmiş ülkelerin GSYİH değerleri ile enerji tüketimleri arasındaki nedensellik ilişkileri Granger nedensellik testi kullanılarak tespit edilmiştir. Nedensellik testi sonuçlarına bakıldığında, sadece gelişmiş ülkeler için GSYİH'ları ile enerji tüketimleri arasında bir nedensellik ilişkisi olduğu tespit edilmiştir. Yapılan bu nedensellik testinin sonucuna göre, kurulacak genelleştirilmiş tahmin denklemleri (GEE) için hangi değişkenin bağımlı, hangi değişkenin bağımsız olduğuna karar verilmiştir. Bu bağlamda, çalışmada kullanılan QIC ve QICC bilgi kriterlerinin de en düşük değerleri, değişkenler arasında belirlenen nedensellik ilişkisinin yönü doğrultusunda elde edilmiştir. Aynı nedensellik ilişkisi değerlendirmesi gelir düzeylerine göre kategorize edilmiş ülkelerin gayri safi milli hasılaları (GSMH) ile enerji tüketimleri için de yapılmış olup, belirlenen nedensellik yönü doğrultusunda kurulan GEE modellerinin veriyi çok daha iyi modellediği sonucuna varılmıştır. Çalışmadan elde edilen bu sonuç bize istatistiksel olarak veriyi modelleyebilmek için bağımlı ve bağımsız değişken seçiminde yetersiz bilgiye sahip olduğumuz durumlarda nedensellik testinin yönlendirici bir kılavuz olduğunu göstermektedir.

Anahtar Kelimeler: Gayri Safi Yurtiçi Hasıla, Gayri Safi Milli Hasıla, Enerji Tüketimi, Granger Nedensellik Testi, Genelleştirilmiş Tahmin Denklemleri.

1.Introduction

Energy consumption and economic development are two indispensable factors in the stability of countries' development performance. The direction of the relationship between these two factors differs according to the different variables, time intervals, and countries used in the model. Econometric tests are effective methods in determining these relationships which are based on causality. Granger (1969; 424-428) proposed a method known as the "Granger causality test" and many studies are done belonging to determining the causality relationships between energy consumption and economic growth of the countries in the literature given Table 1. For more literature review about the causality analysis, see (Aydın, 2010; Öztürk et al., 2010).

Table 1. Summary of the literature review between energy consumption and various economic indicators based on Granger causality test

Authors (Year)	Period	Country	Causality relationships
Hamilton (1983; 228-248)	1948-1972	USA	$EC \rightarrow ED$
Yu and Choi (1985; 249-272)	1954-1976	Philippines, Korea Republic	$EC \rightarrow ED$(Philippines) $ED \rightarrow EC$ (Korea Republic)
Hwang and Gum (1991; 219-226)	1961-1990	Taiwan	$EC \leftrightarrow ED$
Yang (2000; 309-317)	1954-1997	Taiwan	$EC \leftrightarrow ED$
Glasure (2002; 355-365)	1961-1990	Korea	$EC \leftrightarrow ED$
Chiou-Wei et al. (2008; 3063-3076)	1954-2006	Developing countries	$ED \rightarrow EC$((Philippines, and Singapore) $EC \rightarrow ED$ (Taiwan, Hong Kong, Malaysia, and Indonesia)
Shahbaz et al. (2012;2947-2953)	1971-2009	Pakistan	$ED \rightarrow EC$
Hossain (2014;347-376)	1976-2009	SAARC countries namely Bangladesh, India, and Pakistan	$EC \leftrightarrow ED$
Saidi and Mbarek (2016; 364-374)	1990 -2012	9 developed countries	$EC \leftrightarrow ED$

In Table 1, ED indicates economic development, EC indicates energy consumption, $ED \rightarrow EC$ indicates unidirectional causality (from economic development to energy consumption), $EC \rightarrow ED$ indicates unidirectional causality (from energy consumption to economic development), and $EC \leftrightarrow ED$ indicates bi-directional causality (between economic development and energy consumption).

The purpose of this study is to evaluate the causality relationships between GDP and GNI values and electricity energy consumptions of the countries that are categorized according to their income levels as lower income, lower-middle income, middle income, upper-middle income, and high income) between

1980 and 2014, based on Granger causality test. And then by using the determined causality relationships between GDP/GNI values and electricity energy consumptions of the countries, generalized estimating equations (GEE) approach will be applied to the data.

2. Granger Causality Test

The Granger causality test is a method used to determine the direction of the statistically significant causality relationship between two variables. Stationary, which means that the statistical properties of a stochastic process such as the mean, variance, autocorrelation, and etc. are constant over time in time series, is necessary in order to be able to perform the Granger causality test. In Granger causality test, stationary assumption is checked by using Dickey–Fuller (DF) test improved as unit root test (Dickey and Fuller, 1979; 427-431). Simple causality models for variables X_t^* and Y_t^*, providing stochastic and stationary conditions, are given as follows (Işığıçok, 1994; 90-96);

$$X_t^* = \sum_{j=1}^{m} a_j X_{t-j} + \sum_{j=1}^{m} b_j Y_{t-j} + \varepsilon_t \quad (1)$$

$$Y_t^* = \sum_{j=1}^{m} c_j X_{t-j} + \sum_{j=1}^{m} d_j Y_{t-j} + \eta_t \quad (2)$$

where a_j, b_j, c_j and d_j are prediction coefficients, ε_t and η_t are white-noise series. Assume that these white-noise series are uncorrelated at $s \neq t$ time point as follows;

$$E[\varepsilon_t \varepsilon_s] = E[\eta_t \eta_s] = 0 \quad (3)$$

Simple causality models given in Eq. (1) and Eq.(2) presume that X_t^* and Y_t^* are related to their past values of themselves as X_{t-j} and Y_{t-j}, respectively (Gujarati, 2003; 652-657).

This can be assessed in four cases;

- If the estimated coefficients on the lagged X_t^* in Eq. (1) are statistically different from zero and the set of estimated coefficients on the lagged Y_t^* in Eq. (2) is not statistically different from zero, it shows that there is a unidirectional causality from X_t^* to Y_t^*.
- If the estimated coefficients on the lagged X_t^* in Eq. (1) are not statistically different from zero and the set of estimated coefficients on the lagged Y_t^* in Eq. (2) is statistically different from zero, it shows that there is a unidirectional causality from Y_t^* to X_t^*.
- If both of the above conditions are applied, X_t^* causes Y_t^*, and Y_t^* causes X_t^*, it means that there is a bilateral causality between them.
- Finally, when the sets of estimated coefficients on the lagged X_t^* and Y_t^* variables are not statistically significant, it shows that there is no causality between them. The fact that there is no causality between variables means that the variables are independent of each other (Florens and Mouchart, 1982; 583-591).

The joint hypotheses for testing the statistically significance of overall parameters of simple causality models given in Eq.(1) and Eq.(2), respectively, are given as follows;

$$\text{Model 1} \quad \left\{ \begin{array}{l} H_0 : all\ b_j = 0 \\ H_1 : at\ least\ one\ b_j \neq 0 \end{array} \quad j = 1,2,...,m \; ; \; j \neq j' \right\} \quad (4)$$

$$\text{Model 2} \quad \left\{ \begin{array}{l} H_0 : all\ d_j = 0 \\ H_1 : at\ least\ one\ d_j \neq 0 \end{array} \quad j = 1,2,...,m \; ; \; j \neq j' \right\} \quad (5)$$

F-test statistics, developed by Wald for testing the joint hypothesis given in Eq.(4) and Eq.(5) is given as follows (Işığıçok, 1994; 90-96);

$$F_{(m,n-2m)} = \frac{(\sum \varepsilon_{1t}^2 - \sum e_{1t}^2)/m}{\sum e_{1t}^2/(n-2m)} \quad (6)$$

Much as there is no prior knowledge about the size of m in the specified models, m will be assumed finite and shorter than the given time series (Granger, 1969; Işığıçok, 1994).

3. Generalized Estimating Equations (GEE)

The term 'generalized linear model' (GLM) was first introduced in a landmark paper by Nelder and Baker (1972) . When the assumptions of a normally distributed response variable with constant variance are violated, alternative approaches are proposed as data transformations, weighted least squares and generalized linear models (GLMs) (Montgomery et al., 2012; 421-474). Generalized estimating equations (GEE) extend the GLMs approach when panel data is used in many applied fields such as life, mortality, risk classification, non-life insurance, premium pricing, and etc. modeling. The random component part of GEE involves the distribution of the dependent variable coming from the explonential family including many common discrete and continuous distributions such as normal, binomial, multinomial, Poisson, Gamma, Inverse Gaussian, and etc. as follows (İyit et al., 2016; 397-400);

$$f(y;\theta,\phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right\} \quad (7)$$

where $a(\phi)$, $b(\theta)$, and $c(y,\phi)$ are some specific functions, θ is canonical parameter, and ϕ is dispersion parameter (McCullagh, 1984; 285-292).

Dependent variables used in this study have gamma distribution with the following probability-density function;

$$f_{Y_i}(y_i;\omega,\psi) = \left(\frac{y_i}{\omega}\right)^{\psi-1} \frac{\exp -\frac{y_i}{\omega}}{\omega\Gamma(\psi)} \; ; \; y_i > 0 \; ; \; i = 1,2,...n \quad (8)$$

where $\omega > 0$ is the scale parameter and $\psi > 0$ is the dispersion parameter. The expected value (mean) and the variance for gamma distributed dependent variable are as follows (Fox, 2015; 418-425);

$$\mu_i = E(Y_i) = \omega\psi \text{ and } V(Y_i) = \omega^2\psi \quad (9)$$

The systematic component part of GEE involves the linear predictor (η) obtained by a linear combination of covariates as follows (McCullagh, 1984; Agresti, 2015);

$$\eta = \sum_{j=1}^{p} \beta_j x_{ij} \; ; \quad i = 1,2,...,n \quad (10)$$

The parameter estimates in GEE are obtained by using iterative methods like Newton-Raphson, Fisher's Scoring or Hybrid algorithms based on the maximization of the quasi log-likelihood function (Montgomery, 2012; 421-474).

The link function part of GEE involves the relationship between the function of the mean of the dependent variable given in Eq.(9), and the systematic component part given in Eq.(10) (Agresti, 2015; 3-15) as amonotonic and differentiable function;

$$g(\mu_i) = \sum_{j=1}^{p} \beta_j x_{ij} \quad i = 1,2,...,n \tag{11}$$

$1/g(\mu_i)$ is called the inverse link function and called as canonical (natural) link function for Gamma distribution. For more details, see (Dobson and Barnett, 2008; McCullagh, 1984; Montgomery et al., 2012).

The parameter estimates of β_j ; $j = 1,2,...,p$ and standard errors of these parameter estimates are based on the working correlation matrix (Hardin, 2005; 62-69). There are several choices for the working correlation matrix structures for repeated measurements in GEE such as (Pan, 2001; Agresti, 2015);

- Exchangeable working correlation matrix: $corr\left(y_{ij}, y_{ik}\right) = \alpha$
- Independence working correlation matrix: $corr(y_{ij}, y_{ik}) = 0$
- Unstructured working correlation matrix: $corr(y_{ij}, y_{ik}) = \alpha_{ij}$
- First-order autoregressive working correlation matrix: $corr(y_{ij}, y_{ik}) = \alpha^{|i-j|}$
- M-dependent working correlation matrix: $corr(y_{ij}, y_{ik}) = \alpha_{|i-j|}$

The selection of the most appropriate working correlation matrix structure in longitudinal data analysis makes more reliable statistical inferences (Davis, 2002; Wang and Carey, 2003; Pan, 2001). Quasi-log-likelihood under the independence model information criteria as QIC, and also a corrected version of QIC as QICC are used for the most appropriate working correlation matrix structure selection in GEE. QIC value for the working correlation matrix of interest can be calculated as follows;

$$QIC(\mathrm{R}) = -2Q(\hat{\beta}(\mathrm{R}), \phi) + 2trace(\hat{\Omega}_I \hat{V}_{\mathrm{R}}) \tag{12}$$

where R is the working correlation of interest, $\hat{\beta}(R)$ is the parameter estimates belonging to the working correlation structure of interest, $\hat{V}_{\mathrm{R}}$ is robust variance estimator, and $\hat{\Omega}_I$ is an other variance estimator (Jang, 2011; Pan, 2001).

4. Generalized Estimating Equations (GEE) for the Relationships Between GDP/GNI and Electricity Consumption According to Income Levels of Countries Based on Granger Causality Test

In this section, firstly, the causality relationships and the directions of these relationships between GDP and GNI values of the countries categorized by income levels and their electricity consumptions between 1980 and 2014, are tried to be determined by using Granger causality test. And then by determining the dependent and independent variables as a result of this test, the data are modelled by using generalized estimating equations (GEE) approach. The data used in this study are taken from The World Bank WDI Database Archieves.

After detecting nonstationary in series belonging to electricity consumption, GDP and GNI by using Dickey-Fuller (DF) test, these series are stabilized by taking the first differences. And then the results obtained from Granger casusality test are given in Table 2.

Table 2. Granger causality test results for EC, GDP and GNI

Income levels	$EC \rightarrow GDP$			$EC \rightarrow GNI$		
Lower income	Chi-square	df	p	Chi-square	df	p
	0.276220	2	0.8710	0,761408	2	0,6834
	$GDP \rightarrow EC$			$GNI \rightarrow EC$		
	Chi-square	df	p	Chi-square	df	p
	0.100012	2	0.9512	0,106987	2	0,9479
Lower-middle income	$EC \rightarrow GDP$			$EC \rightarrow GNI$		
	Chi-square	df	p	Chi-square	df	p
	1.165108	2	0.5585	1.479461	2	0.4772
	$GDP \rightarrow EC$			$GNI \rightarrow EC$		
	Chi-square	df	p	Chi-square	df	p
	0.903766	2	0.6364	0.835048	2	0.6587
Middle income	$EC \rightarrow GDP$			$EC \rightarrow GNI$		
	Chi-square	df	p	Chi-square	df	p
	0.792749	2	0.3733	1.460963	2	0.2268
	$GDP \rightarrow EC$			$GNI \rightarrow EC$		
	Chi-square	df	p	Chi-square	df	p
	0.583800	2	0.4448	1.282648	2	0.2574
Upper-middle income	$EC \rightarrow GDP$			$EC \rightarrow GNI$		
	Chi-square	df	p	Chi-square	df	p
	1.967106	2	1.967106	3.069031	2	0.2156
	$GDP \rightarrow EC$			$GNI \rightarrow EC$		
	Chi-square	df	p	Chi-square	df	p
	2.151205	2	2.151205	0.900173	2	0.6376
High income	$EC \rightarrow GDP$			$EC \rightarrow GNI$		
	Chi-square	df	p	Chi-square	df	p
	1.569175	2	0.4563	8.394581	6	0.2106
	$GDP \rightarrow EC$			$GNI \rightarrow EC$		
	Chi-square	df	p	Chi-square	df	p
	6.381348	2	**0.0411***	17.92671	6	**0.0064***

***(p<0,05) indicates statistically significant causality relationships between variables**

According to the Granger causality test results, there are no causality relationships between the GNI/GDP values and electricity consumptions of the countries in lower, lower-middle, middle and upper-middle income levels as shown in Table 2. On the other hand, causality relationships between these variables are only determined in high income level. Also the directions of these statistically significant causality relationships are examined for high income level.

For the relationship between EC and GDP $(EC \rightarrow GDP)$, the first hypothesis set for Model 1 is rejected at the level of 5% significance. But for the relationship between GDP and EC $(GDP \rightarrow EC)$, the second hypothesis set for Model 2 is not rejected at the level of 5% significance. It means that the causality relationship between electricity consumption and GDP is unidirectional causality from GDP to electricity consumption (p=0.0411<0.05).

Furthermore, for $EC \rightarrow GNI$ is obtained likewise $EC \rightarrow GDP$ and the causality is detected between electricity consumption and GNI, which is a unidirectional causality from GNI to electricity consumption EC (p=0.0064<0.05). GEE models are established as a result of the $GDP \rightarrow EC$ and $GNI \rightarrow EC$ causality tests. In both cases GDP and GNI are independent variables and the EC is dependent variable, respectively.

The QIC values of the model, based on causality test, is compared with the QIC values of the model, which without considering $GDP \rightarrow EC$ causality. This comparison reveals that the model, based on the direction of causality is better than the other model, established in the opposite direction. In the same way, this procedure is applied for $GNI \rightarrow EC$.The model of $GDP \rightarrow EC$ is build as; EC dependent variable,which comes from Gamma distribution, the canonical link function of it is inverse, and the working correlation matrix structure is independent. Also, the same choice for modelling are used for the causality of $GNI \rightarrow EC$. Goodness-of-fit test statistics as QIC and QICC belonging to the GEE models based on causality and non-causality relationships determined by Granger causality test are given Table 3.

Table 3. Goodness-of-fit test statistics belonging to the GEE models based on causality and non-causality relationships

Causality	Model	QIC	QICC*
$GDP \rightarrow EC$	$(EC) = 6,546 + 1.791*10^{-13}(GDP)$	**128.615**	**86.046**
$EC \rightarrow GDP$	$(GDP) = 29,014 + 0.000269(EC)$	138.326	106.658
$GNI \rightarrow EC$	$(EC) = 6,554 + 1.797*10^{-13}(GNI)$	**125.854**	**84.735**
$EC \rightarrow GNI$	$(GNI) = 28,962 + .00025(EC)$	134.516	102.433

The smallest values of QIC and QICC in Table 3 indicate the most appropriate relationship by GEE approach. From Table 3, the smallest values of QIC and QICC indicate $GDP \rightarrow EC$ relationship is better than $EC \rightarrow GDP$ relationship. And resulted as **128.615** (QICC=**86.046**) for is based the causality of $GDP \rightarrow EC$ and $GNI \rightarrow EC$ relationship is better than $EC \rightarrow GNI$. So the GEE models constituted in the direction of causality, have smaller QIC and QICC values. It means that the GEE models with causality relationships better fit to the data than the GEE models with non-causality relationships.

5.Results and Discussion

Electricity consumption and GDP are are two major indicators that change together in the country's development evaluations. When this relationship is examined, there may be differences between countries in terms of causality. The income level of the countries is a factor affecting the direction of this relationship so in this study, it is discussed separately for each income level of the countries.The Granger causality test is conducted for 5 different income levels ,but a statistically significant difference is determined only in the high income level. This result suggests that GDP is the cause of electricity consumption. And also the GNI and electricity consumption causality is examined. GNI is determined as the cause of electricity consumption. The causality of electricity consumption in both assessments is from GNI/GDP to EC. Causality is detected between these variables only in the high income level and then the data is modelled by GEE approach. In addition, the data is modelled by GEE for cases where causality is not determined. The results indicate that the GEE models based on causality, better fit to the data than the GEE models established without consideration of causality. This founding shows that the selection of the variables as independent or dependent by causality analysis before constructing GEE models gives better results in statistical modelling.

When the causality relationship $\mathbf{GDP \rightarrow EC}$ is modelled, $(EC) = 6,546 + 1.791*10^{-13}(GDP)$ GEE model is obtained. When the causality relationship $EC \rightarrow GDP$ is modelled, $(GDP) = 29,014 + 0.000269(EC)$ GEE model is obtained. Actual and predicted values for these GEE models in the directions of $\mathbf{GDP \rightarrow EC}$ and $\mathbf{EC \rightarrow GDP}$ causality relationships in high income level between 1980 and 2014 are given in Table 4. When we look at the actual vs predicted line graphs for both models in Figure 1, it is obviously seen that the predictions of the GEE model gives better fit to the data constructed in the causality direction in high income level.

Table 4. Actual and predicted values for GEE models in the directions of $GDP \rightarrow EC$ and $EC \rightarrow GDP$ causality relationships in high income level between 1980 and 2014

Years	$GDP \rightarrow EC$			$EC \rightarrow GDP$		
	Actual	**Predicted**	**Residual**	**Actual**	**Predicted**	**Residual**
1980	3.762173	3.148754	0.613419	12.94932	13.27742	-0.3281
1985	3.806529	3.191138	0.615391	13.00571	13.35018	-0.34447
1990	3.865108	3.49218	0.372928	13.27638	13.45839	-0.18201
1995	3.902116	3.730058	0.172058	13.41198	13.53468	-0.1227
2000	3.947264	3.790163	0.157101	13.44046	13.63699	-0.19653
2005	3.968273	4.133215	-0.16494	13.57471	13.68835	-0.11364
2010	3.970043	4.395274	-0.42523	13.65502	13.69279	-0.03777
2014	3.95842	4.571161	-0.61274	13.70164	13.66395	0.03769

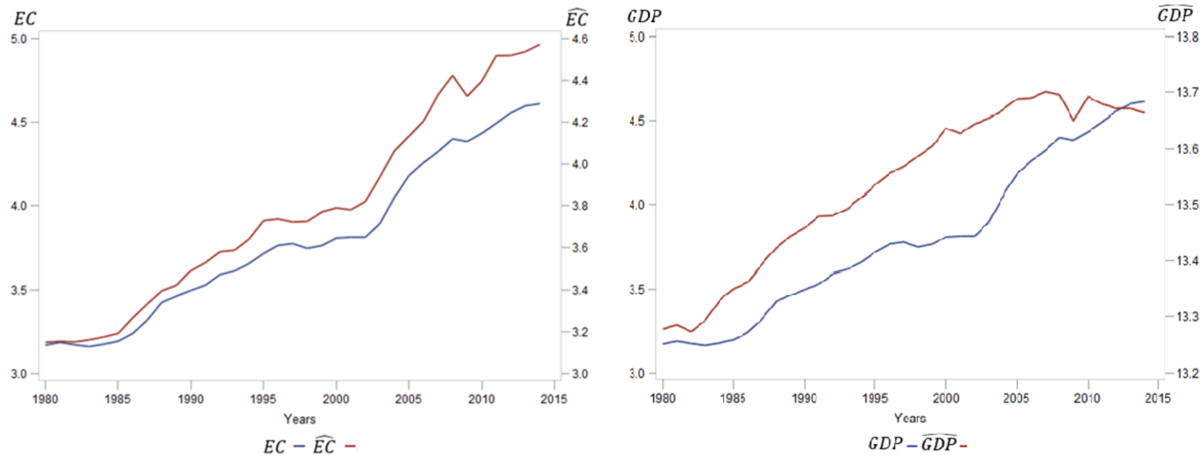

Fig.1. Actual and predicted line plots for GEE models in the directions of $GDP \rightarrow EC$ and $EC \rightarrow GDP$ causality relationships in high income level between 1980 and 2014

When the causality relationship $GNI \rightarrow EC$ is modelled, $(EC) = 6,554 + 1.797*10^{-13}(GNI)$ GEE model is obtained. When the causality relationship $EC \rightarrow GNI$ is modelled, $(GNI) = 28,962 + .00025(EC)$ GEE model is obtained. Actual and predicted values for these GEE models in the directions of $GNI \rightarrow EC$ and $EC \rightarrow GNI$ causality relationships in high income level between 1980 and 2014 are given in Table 5. When we again look at the actual vs predicted line graphs for both models in Figure 2, it is obviously seen that the predictions of the GEE model gives better fit to the data constructed in the causality direction in high income level.

Table 5. Actual and predicted values for GEE models in the directions of $GNI \rightarrow EC$ and $EC \rightarrow GNI$ causality relationships in high income level between 1980 and 2014

Years	$GNI \rightarrow EC$			$EC \rightarrow GNI$		
	Actual	**Predicted**	**Residual**	**Actual**	**Predicted**	**Residual**
1980	3.762173	3.16911	0.593063	12.96967	13.26874	-0.29907
1985	3.806529	3.192587	0.613942	13.00016	13.34303	-0.34287
1990	3.865108	3.495089	0.370019	13.2728	13.45353	-0.18073
1995	3.902116	3.716994	0.185122	13.40055	13.53142	-0.13087
2000	3.947264	3.807735	0.139529	13.4436	13.63588	-0.19228
2005	3.968273	4.181939	-0.21367	13.58637	13.68832	-0.10195
2010	3.970043	4.432876	-0.46283	13.66114	13.69286	-0.03172
2014	3.95842	4.611588	-0.65317	13.70749	13.66341	0.04408

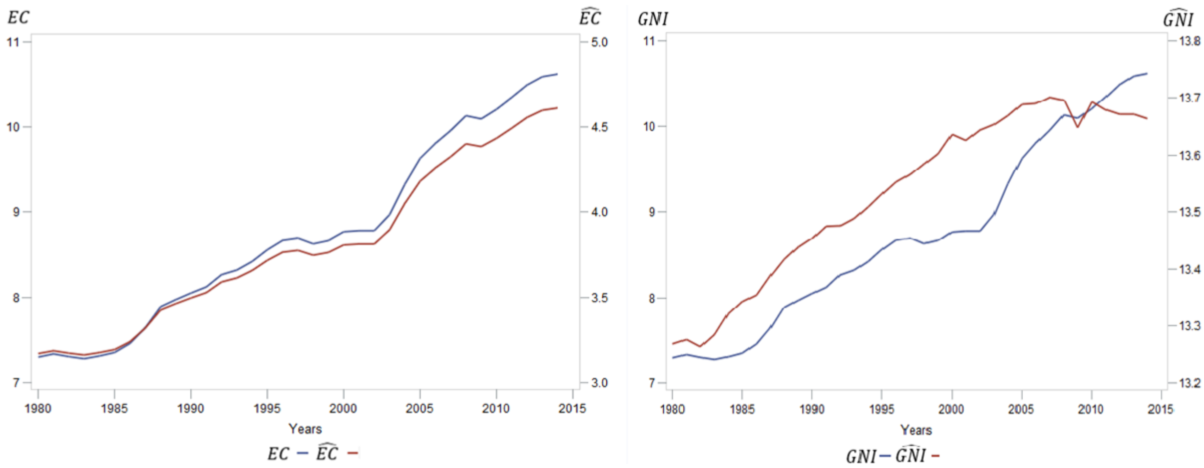

Fig.2. Actual and predicted line plots for GEE models in the directions of $GNI \rightarrow EC$ and $EC \rightarrow GNI$ causality relationships in high income level between 1980 and 2014

6.Conclusion

In this study, it is investigated that causality is considered as a pioneering approach in selecting variables for statistical modelling. The causality test is also useful tool in determining the dependent and independent variables selection especially in economy framework. GEE can of course also be used when there is no causality, but it only reveals which variable in the model is related to the other, and the result is far away from the explanation of the causal relation between variables. The model based on the causality relationship will guide an effective evaluation for statistical modelling especially investigated in economics. The detection of causality direction is a very effective approach in economic studies, especially when there are difficulties in determining dependent variables. In this study, the causality relationships between GDP/GNI and electricity consumption have been determined only for the countries in the high income level. From this study, it can be concluded that the countries with higher income levels are more likely to benefit from the energy they use.

References

Agresti, A. Foundations of Linear and Generalized Linear Models, Cambridge, John Wiley & Sons, 2015.

Aydın, F. F. (2010). "*Enerji Tüketimi ve Ekonomik Büyüme*", Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, (35), p. 317-340.

Chiou-Wei, S. Z., Chen, C.-F., & Zhu, Z. (2008). "Economic Growth and Energy Consumption Revisited—Evidence from Linear and Nonlinear Granger Causality", Energy Economics, 30(6), p. 3063-3076.

Davis, C. S. Statistical Methods for the Analysis of Repeated Measurements, New York, Springer Science & Business Media, 2002.

Dickey, D. A., & Fuller, W. A. (1979). "*Distribution of the Estimators for Autoregressive Time Series with A Unit Root*", Journal of the American Statistical Association, 74(366a), p. 427-431.

Dobson, A. J., & Barnett, A. An Introduction to Generalized Linear Models, ülke adı, CRC Press, 2008.

Florens, J.-P., & Mouchart, M. (1982). "*A Note on Noncausality*", Econometrica: Journal of the Econometric Society, p. 583-591.

Fox, J. Applied Regression Analysis and Generalized Linear Models, Australia, Sage Publications, 2015.

Glasure, Y. U. (2002). "*Energy and National Income in Korea: Further Evidence on the Role of Omitted Variables*", Energy Economics, 24(4), p. 355-365.

Granger, C. W. (1969). "*Investigating Causal Relations by Econometric Models and Cross-Spectral Methods*", Econometrica: Journal of the Econometric Society, p. 424-438.

Gujarati, D. N. Basic Economics". 4th Edition, McGraw Hill, New York, 2003.

Hamilton, J. D. (1983). "*Oil and the Macroeconomy since World War II*" Journal of political economy, 91(2), p. 228-248.

Hardin, J. W. Generalized Estimating Equations (GEE), Boca Raton-FL, Wiley Online Library, 2005.

Hossain, S. (2014). "*Multivariate Granger Causality Between Economic Growth, Electricity Consumption, Exports and Remittance for the panel of three SAARC Countries*", European Scientific Journal, ESJ, 8(1).

Hwang, D. B., & Gum, B. (1991). "*The Causal Relationship between Energy and GNP: the case of Taiwan*" The Journal of Energy and Development, p. 219-226.

Işığıçok, E. Zaman Serilerinde Nedensellik Çözümlemesi: Türkiye'de Para Arzı ve Enflasyon Üzerine Amprik bir Araştırma, Bursa Uludağ Üniversitesi Basımevi, 1994.

İyit, N., Yonar, H., & Genç, A. (2016). "*Generalized Linear Models for European Union Countries Energy Data*" Acta Physica Polonica, A., 130(1), 397-400.

Jang, M. J. (2011). "*Working Correlation Selection in Generalized Estimating Equations*", Phd thesis, USA, The University of Iowa.

McCullagh, P. (1984). "*Generalized Linear Models*" European Journal of Operational Research, 16(3), p. 285-292.

Montgomery, D. C., Peck, E. A., & Vining, G. G. Introduction to Linear Regression Analysis, New Jersey, John Wiley & Sons, 2012.

Nelder, J. A., & Baker, R. J. (1972). "*Generalized linear models* ", John Wiley & Sons, Inc..

Ozturk, I., Aslan, A., & Kalyoncu, H. (2010). "*Energy Consumption and Economic Growth Relationship: Evidence From Panel Data for Low and Middle Income Countries*", Energy Policy, 38(8), p. 4422-4428.

Pan, W. (2001). "*Akaike's Information Criterion in Generalized Estimating Equations*", Biometrics, 57(1), p. 120-125.

Saidi, K., & Mbarek, M. B. (2016). "*Nuclear Energy, Renewable Energy, Co 2 Emissions, and Economic Growth for Nine Developed Countries: Evidence from Panel Granger Causality Tests*" Progress in Nuclear Energy, 88, p. 364-374.

Shahbaz, M., Lean, H. H., & Shabbir, M. S. (2012). "*Environmental Kuznets Curve Hypothesis In Pakistan: Cointegration and Granger Causality*" Renewable and Sustainable Energy Reviews, 16(5), p. 2947-2953.

The World Bank WDI Database Archieves. (Online) http://databank.worldbank.org/data/reports.aspx?source=wdi-database-archives-(beta), (Date of access: 10 October 2017)

Wang, Y. G., & Carey, V. (2003). "*Working Correlation Structure Misspecification, Estimation and Covariate Design: Implications for Generalised Estimating Equations Performance*", Biometrika, 90(1), p. 29-41.

Yang, H.-Y. (2000). "*A Note on the Causal Relationship Between Energy and GDP in Taiwan*" Energy Economics, 22(3), p. 309-317.

Yu, E. S., & Choi, J.Y. (1985). "*The Causal Relationship Between Energy and GNP: An International Comparison*" The Journal of Energy and Development, 10(2), p. 249-272.