西南交通大学学报

第 55 卷 第 6 期
2020 年 12 月

JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY

Vol. 55 No. 6
Dec. 2020

ISSN: 0258-2724

DOI：10.35741/issn.0258-2724.55.6.9

Research article

Mathematics

CAUSAL MODELING AND DECOMPOSITION OF CORRELATION STUDY: INDICATORS OF AN INDONESIAN ELECTRICITY COMPANY

相關研究的因果建模與分解：印度尼西亞電力公司的指標

Mustofa Usman [a,*], Edwin Russel [b], Warsono [a], Wamiliana [a], Amanto [a], Fitriani [a], Faiz A.M. Elfaki [c]

[a] Department of Mathematics, Faculty of Mathematics and Sciences, Universitas Lampung, Indonesia, M_usman_alfha@yahoo.com

[b] Department of Management, Faculty of Economics and Business, Universitas Lampung, Indonesia.

[c] Department of Mathematics, Statistics and Physics, College of Art and Sciences, Qatar University, Qatar

Received: April 16, 2020 ▪ Review: July 8, 2020 ▪ Accepted: October 19, 2020

Abstract

The Causal Modeling or Path analysis has become a widely-used statistical method in many areas of study, such as in genetics, biology, social science, environmental science, economic, business, finance and sociology. In this study, causal modeling is used to analyze the relationship among variables of important indicators of the electricity company of the Republic of Indonesia. The variables discussed are Total number of Staff or Employees, Electric Power Installed in megawatts, Operating Cost in million rupees, Electric Power Product in megawatts, Electric Power Sales in megawatts, and Sales Revenue in million rupees. The results of analysis show that the proposed causal modeling indicates the relationships among those six important indicators of the electricity company are very significant and meaningful.

Keywords: Causal Modeling, Decomposition Correlation, Direct Effect, Indirect Effect, Total Effect

摘要 因果模型或路徑分析已成為許多研究領域中廣泛使用的統計方法，例如遺傳學，生物學，社會科學，環境科學，經濟，商業，金融和社會學。在這項研究中，使用因果模型來分析印度尼西亞共和國電力公司重要指標的變量之間的關係。討論的變量是員工總數，以兆瓦為單位安裝的電力，以百萬盧比為單位的運營成本，以兆瓦為單位的電力產品，以兆瓦為單位的電力銷售以及以百萬盧比為單位的銷售收入。分析結果表明，提出的因果模型表明電力公司的六個重要指標之間的關係非常重要和有意義。

关键词: 因果建模，分解相關，直接效應，間接效應，總效應

I. INTRODUCTION

Path analysis or Causal Modeling has a very long history in statistical application. It was introduced by Wright [1],[2], and it was first applied in the field of genetics. The application in genetics can be found in Vogler [3]. It was one of the methods used by many researchers. Causal modeling has become popular and is an analytical method of analysis in social sciences [4]. This method is applied in many field of studies such as education [5], [6], [7], sociology and social sciences [8], [9], [10], [11], [12], and business and management [13], [14], [15]. This study is perhaps the first to use path analysis in an effort to distinguish and measure the effects of medical, anthropometric, behavioral, and sociodemographic factors on the risk of premature birth [16]. Path analysis is not a method to find a model, but this is a method that can be used for testing causal models which have been proposed by a scientist [15]. Therefore, path analysis or causal modeling is a method to test a proposed model offered by researchers [15]. The causal modeling or path analysis does not allow one to determine the direction of causality between two variables; if there is a causal relationship between two variables, then the researchers have to state at the outset what the direction of that relationship is; a decision must be made on the basis of theoretical and substantive grounds [4],[17]. Many books [17] explained one of the approaches to estimate and calculate the direct effect using the concept of linear algebra, system of equations. But this approach for estimating the parameters in path analysis is quite complicated for some researchers. Some statisticians offer a means of calculating the estimation and testing the direct effect and indirect effect by using standardized simple or multiple linear regressions [4], [18],[19]. This approach is simpler and easier to interpret s. And even though Blalock [20] was clear that regression coefficients are estimated quantities (and more fundamentally that the causal models that give regression equations their specifications are subject to simplifying assumptions that may be unrealistic), he still wrote about the resulting coefficients and equations in ways that would surely have excited readers interested in powerful new ways to gain insight from observational data: It is the regression coefficients which give us the laws of science [20]). In causal analyses our aim is to focus on causal laws as represented by regression equations and their coefficients [20]. Even though the path analysis used the standardized simple or multiple linear regression, there are differences in the analysis. In multiple linear regression analysis, each independent (predictor) variable has a direct effect on the dependent variable (or response variable). In the path analysis model, the independent variable (predictor) not only has a direct effect on the response variable, it also has an indirect effect through one or more intervening variables [21]. One of the advantages of path analysis or causal modeling is the ability to explain the direct and indirect effects between variables. Path diagrams are useful as a simple descriptive tool to describe the direct and indirect effects of variables in the model. The coefficient p in the path analysis model is meant to quantify the causal impact of one variable on the other variable as connected by an arrow [22]. Thinking causally about a problem and using an arrow diagram that indicates the causal processes may often give a clearer statement of hypotheses and the interpretation of the topic at hand [23]. In path analysis model, it was assumed that all variables used in a regression model are in standard form, that is, with mean zero and variance one. Therefore, the interpretation of the path coefficients is in standard deviation units [17],[24],[25].

The aims of this study are to answer these questions: (1) Are there direct and indirect effects of NS to EPI? (2) Are there direct and indirect effects of NS and EPI to OC? (3) Are there direct and indirect effects of EPI and OC to EPP? (4) Are there direct and indirect effects of OC and EPP to EPS? (5) Are there direct and indirect effects of OC and EPS to SR, where Total number of Staff or Employees (NS), Electric Power Installed in MW (EPI), Operating Cost (Million Rp) (OC), Electric Power Product in MWh (EPP), Electric Power Sales in MWh (EPS), and Sales Revenue in Million Rp (SR)?

II. CAUSAL MODEL ANALYSIS AND DECOMPOSITION OF CORRELATION

A. Statistical Modeling

The causal relationships of the important indicators of the public electric company of Indonesia are studied. The indicators are: Total number of Staff or Employees (NS), Electric Power Installed in MW (EPI), Operating Cost (Million Rp) (OC), Electric Power Product in MWh (EPP), Electric Power Sales in MWh (EPS), and Sales Revenue in Million Rp (SR). The hypothetical causal model of the causal relationships among those variables is depicted as follows:

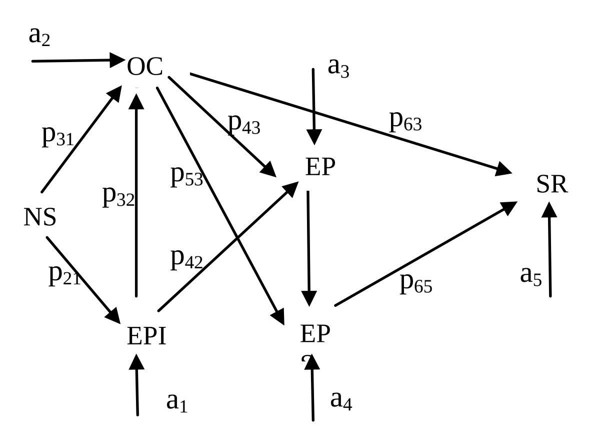

Figure 1. Causal model of the relationships among the important indicators of the public electric company of Indonesia. The indicator variables are Total number of Staff or Employees (NS), Electric Power Installed in MW (EPI), Operating Cost (Million Rp) (OC), Electric Power Product in MWh (EPP), Electric Power Sales in MWh (EPS), and Sales Revenue in Million Rp (SR)

Based on Figure 1, the mathematical causal modeling can be written as follows [26]:

Model 1: $EPI = p_{21}\, NS + p_1\, a_1$

Model 2: $OC = p_{31}\, NS + p_{32}\, EPI + p_2\, a_2$,

Model 3: $EPP = p_{42} EPI + p_{43}\, OC + p_3\, a_3$,

Model 4: $EPS = p_{53}\, OC + p_{54}\, EPP + p_4\, a_4$,

Model 5: $SR = p_{63}\, OC + p_{65}\, EPS + p_5\, a_5$ (1)

where a_1, a_2, a_3 and a_4 are error terms. The main objectives of these models are to test the null hypotheses for respective models. From models (1), (2), (3) and (4) there are four null hypotheses to be tested, namely:

H_{01}: There is no direct effect of NS to EPI;

H_{02}: There are no direct effect of NS and EPI to OC;

H_{03}: There are no direct effect of EPI and OC to EPP;

H_{04}: There are no direct effect of OC and EPP to EPS;

H_{05}: There are no direct effect of OC and EPS to SR.

The parameters of p_1, p_2, p_3, p_4 and p_5 in the error terms can be calculated after [4],[17],[26] as follows:

$$p_i = \sqrt{1 - RSquares_i}\,, \text{ where } i = 1,2,3,4,5 \qquad (2)$$

where R-squares$_i$ are the degrees of determination of Models 1,2,3,4 and 5 above, respectively. Furthermore, from Model 1,2,3,4 and 5, besides analysis of direct and indirect effects, also we will discuss the total effects from one variable to the others variables. The method how to calculate the total effects can be found in [15], [17], and [22].

B. Decomposition of Correlations

One of the advantages of causal modeling analysis, or path analysis, is that it offers a method of explaining the decomposition of correlation among variables in studies by enhancing the interpretation of correlation [15]. One of the interesting features of causal model analysis, or path analysis, is that we can explore the correlation between components. In a given path analysis, we can determine the aspect of correlation between two variables and decompose it into its direct effects and indirect effects [15], [17]. The data of several factors, including Total Staff or Number of Employees (NS), Electric Power Installed in MW (EPI), Operating Cost in million Indonesian Rupiah (Rp) (OC), Electric Power Product in MWh (EPP), Electric Power Sales in MWh (EPS), and Sales Revenue in Million Rp (SR), are transformed into standardized data with mean = 0 and variances = 1. From this standardized data, it is easy to calculate the expected values between two variables, so that E(NS.NS) = 1, E(OC.OC) = 1, E(EPI.EPI) = 1, E(EPP.EPP) = 1, E(EPS.EPS) = 1, and E(SR.SR) = 1, as well as the expected values between two different variables, such that E(NS.EPI) = r_{12}, E(NS.OC) = r_{13}, E(NS.EPP) = r_{14}, E(NS.EPS) = r_{15}, E(NS.SR) = r_{16}, E(EPI.OC) = r_{23}, E(EPI.EPP) = r_{24}, E(EPI.EPS) = r_{25}, E(EPI.SR) = r_{26}, E(OC.EPP) = r_{34}, E(OC.EPS) = r_{35}, E(OC.SR) = r_{36}, E(EPP.EPS) = r_{45}, E(EPP.SR) = r_{46}, and E(EPS.SR) = r_{56}. Here, r_{12}, r_{13}, r_{14}, r_{15}, r_{16}, r_{23}, r_{24}, r_{25}, r_{26}, r_{34}, r_{35}, r_{36}, r_{45}, r_{46}, and r_{56} are correlations between the variables NS and EPI, NS and OC, NS and EPP, NS and EPS, NS and SR, EPI and OC, EPI and EPP, EPI and EPS, EPI and SR, OC and EPP, OC and EPS, OC and SR, EPP and EPS, EPP and SR, and EPS and SR, respectively. From model 1, algebra and tracing rules can be used to determine the decomposition of correlation. Here, both sides are multiplied by NS and then the expected value is taken as given below:

$$E(NS.EPI) = p_{21}\; E(NS.NS) + E(a_1.EPI),$$

So that

$$r_{12} = p_{21}\;,$$

In Model 2, both sides is multiplied by NS and then the expected value is taken as given below:

$$E(OC.NS) = p_{31} E(NS.NS) + p_{32} E(EPI.NS) + p_2 E(a_2.NS), \quad (3)$$

So that

$$r_{13} = p_{31} + p_{32} p_{21} \quad (4)$$

And multiplied both sides by EPI and then the expected value is taken as given below:

$$E(OC.EPI) = p_{31} E(NS.EPI) + p_{32} E(EPI.EPI) + p_2 E(a_2.EPI), \quad (5)$$

So that

$$r_{23} = p_{31} p_{21} + p_{32} \quad (6)$$

In Model 3, both sides is multiplied by NS and then the expected value is taken as given below:

$$E(EPP.NS) = p_{42} E(EPI.NS) + p_{43} E(OC.NS) + p_3 E(a_3.NS), \quad (7)$$

So that

$$r_{14} = p_{42} p_{21} + p_{43} r_{13},$$

$$r_{14} = p_{42} p_{21} + p_{43} (p_{31} + p_{32} p_{21}),$$

or

$$r_{14} = p_{42} p_{21} + p_{43} p_{31} + p_{43} p_{32} p_{21} \quad (8)$$

In Model 3, both sides is multiplied by EPI and then the expected value is taken as given below:

$$E(EPP.EPI) = p_{42} E(EPI.EPI) + p_{43} E(OC.EPI) + p_3 E(a_3.EPI), \quad (9)$$

So that

$$r_{24} = p_{42} + p_{43} r_{23},$$

$$r_{24} = p_{42} + p_{43} (p_{31} r_{12} + p_{32}),$$

or

$$r_{24} = p_{42} + p_{43} p_{31} r_{12} + p_{43} p_{32}. \quad (10)$$

In Model 3, both sides is multiplied by OC and then the expected value is taken as given below:

$$E(EPP.OC) = p_{42} E(EPI.OC) + p_{43} E(OC.OC) + p_3 E(a_3.OC), \quad (11)$$

So that

$$r_{34} = p_{42} r_{23} + p_{43},,$$

$$r_{34} = p_{42} (p_{31} p_{21} + p_{32}) + p_{43},$$

or

$$r_{34} = p_{42} p_{31} p_{21} + p_{42} p_{32} + p_{43} \quad (12)$$

In Model 4, both sides is multiplied by OC and then the expected value is taken as given below:

$$E(EPS.OC) = p_{53} E(OC.OC) + p_{54} E(EPP.OC) + p_4 E(a_4.OC), \quad (13)$$

So that

$$r_{35} = p_{53} + p_{54} r_{34},$$

$$r_{35} = p_{53} + p_{54} (p_{42} p_{31} p_{21} + p_{42} p_{32} + p_{43}),$$

or

$$r_{35} = p_{53} + p_{54} p_{42} p_{31} p_{21} + p_{54} p_{42} p_{32} + p_{54} p_{43} \quad (14)$$

In Model 4, both sides are multiplied by EPP and then the expected value is taken as given below:

$$E(EPS.EPP) = p_{53} E(OC.EPP) + p_{54} E(EPP.EPP) + p_4 E(a_4.EPP), \quad (15)$$

So that

$$r_{45} = p_{53} r_{34} + p_{54},$$

$$r_{45} = p_{53} (p_{42} p_{31} p_{21} + p_{42} p_{32} + p_{43}) + p_{54},$$

or

$$r_{45} = p_{53} p_{42} p_{31} p_{21} + p_{53} p_{42} p_{32} + p_{53} p_{43} + p_{54} \quad (16)$$

In Model 5, both sides is multiplied by OC and then the expected value is taken as given below:

$$E(SR.OC) = p_{63} E(OC.OC) + p_{65} E(EPS.OC) + p_5 E(a_5.OC), \quad (17)$$

So that

$$r_{36} = p_{63} + p_{65} r_{35}$$

$$r_{36} = p_{63} + p_{65} (p_{53} + p_{54} p_{42} p_{31} p_{21} + p_{54} p_{42} p_{32} + p_{54} p_{43}),$$

or

$$r_{36} = p_{63} + p_{65} p_{53} + p_{65} p_{54} p_{42} p_{31} p_{21} + p_{65} p_{54} p_{42} p_{32} + p_{65} p_{54} p_{43}. \quad (18)$$

In Model 5, both sides is multiplied by EPS and then the expected value is taken as given below:

$$E(SR.EPS) = p_{63} E(OC.EPS) + p_{65} E(EPS.EPS) + p_5 E(a_5.EPS), \quad (19)$$

So that

$$r_{56} = p_{63} r_{35} + p_{65}$$

$$r_{56} = p_{63} (p_{53} + p_{54} p_{42} p_{31} p_{21} + p_{54} p_{42} p_{32} + p_{54} p_{43}) + p_{65},$$

or

$$r_{56} = p_{63} p_{53} + p_{63} p_{54} p_{42} p_{31} p_{21} + p_{63} p_{54} p_{42} p_{32} + p_{63} p_{54} p_{43} + p_{65} \quad (20)$$

III. RESULTS AND DISCUSSION

The important indicator data of the public electric company Republic of Indonesia are from the Central Bureau of Statistics Indonesia (BPS, [27]). These important indicators are: Total number of Staff or Employees (NS), Electric Power Installed in MW (EPI), Operating Cost (Million Rp) (OC), Electric Power Product in MWh (EPP), Electric Power Sales in MWh (EPS), and Sales Revenue in Million Rp (SR). Before analyzing the data using path analysis, the data are transformed into standardized form with mean = 0 and variance = 1. The results of analysis for model 1 are presented in Table 1 and Table 2.

Table 1.
Analysis of variance for testing Model 1

Source	DF	Sum of Squares	Mean Square	F Value	P-Value
Model	1	13.108	13.108	53.90	<.0001
Error	16	3.891	0.243		
Corrected Total	17	17.000			

R-squares= 0.7711

Table 2.
Parameter estimated and testing for partial parameter of Model 1

Parameter	Estimate	Standard Error	t Value	P-Value
NS	0.8781	0.1196	7.34	<.0001

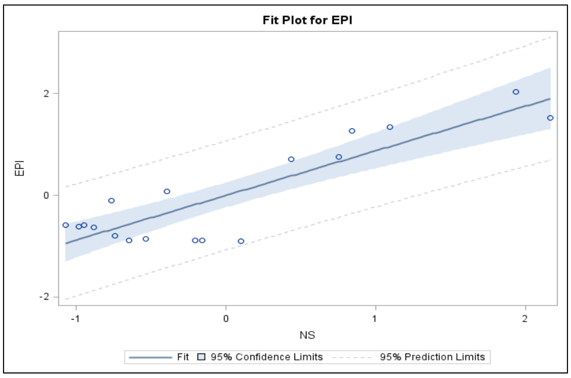

Figure 2. Contour plot of Model 1

Table 1 presents the analysis of variance for testing the parameters in model 1, with the null hypothesis that there is no direct effect of NS to EPI. The results are: F-test=53.90 with P <0.0001. Therefore, the null hypothesis is rejected and we concluded that there is direct effect of NS to EPI. The R-Square=0.7711, meaning that 77.11% of the variation of EPI can be accounted for by NS or by the model. Table 2 shows the results of parameter $p_{21} = 0.8781$ estimation and testing of Model 1 with P< 0.0001. Therefore, the null hypothesis $Ho: p_{21} = 0$ is rejected.

Figure 2 shows the contour plot of model 1, which also indicates a positive correlation: if the value of NS increases, the value of EPI moves increase. The estimation of model 1 is:

EPI = 0.8781 OC

The unexplained variation is

$$p_1 = \sqrt{1 - 0.9606} = 0.1985$$

The results of analysis for model 2 are presented in Table 3 and Table 4.

Table 3.
Analysis of variance for testing Model 2

Source	DF	Sum of Squares	Mean Square	F Value	P-Value
Model	2	16.3297	8.1648	182.72	<.0001
Error	15	0.6703	0.0447		
Corrected Total	17	17.0000			

R-squares=0.9606

Table 4.
Parameter estimated and testing for partial parameter of Model 2

Parameter	DF	Estimate	Standard Error	t Value	P-value
NS	1	-0.1872	0.1072	-1.75	0.1011
EPI	1	1.1403	0.1072	10.64	<.0001

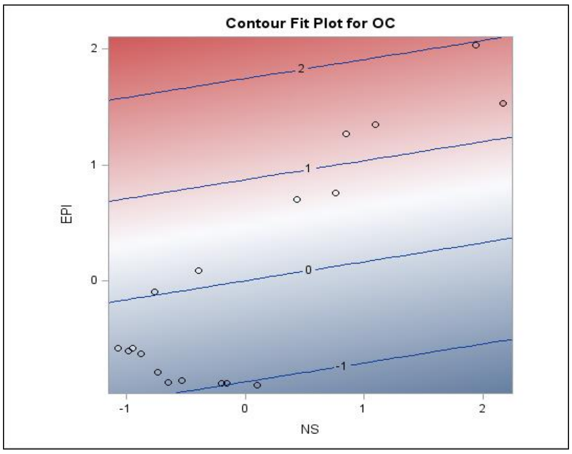

Figure 3. Contour plot of Model 2

Table 3 presents the analysis of variance for testing Model 2, with the null hypothesis that there is no direct effect of NS and EPI to OC. The results are: F-test=182.72 with P <0.0001. Therefore, the null hypothesis is rejected and we concluded that there is direct effect of NS and EPI to OC. The R-Square=0.9606, meaning that 96.06% of the variation of OC can be accounted for by NS and EPI, or by the model. Table 4 shows the results of parameter estimation and partial testing of parameters $p_{31} = -0.1872$ and $p_{32} = 1.14035$ in Model 2 with P=0.1011 and P< 0.0001, respectively. Therefore, the null hypothesis $Ho : p_{31} = 0$ is not rejected, but the null hypothesis $Ho : p_{32} = 0$ is rejected. Even though the parameter p_{31} is not rejected, the absolute value of p_{31} = -0.1872, which is greater than 0.05. According to Pedhazur [17] and Heisse [28], this is meaningful. Figure 3 shows the contour plot of model 2, also indicating the negative correlation if the value of NS increases, the value of OC moves to the red area. The response for OC decreases as NS increases, and the other variable is kept constant. But the trend is positive when the value of EPI increases, the value of OC moves increases, indicated by the increasing straight line as EPI increases and the other variable is kept constant. The estimation of Model 2 is:

OC = -0.1872 NS + 1.1403 EPI

The unexplained variation is

$$p_2 = \sqrt{1 - 0.9606} = 0.1985$$

The results of analysis for Model 3 are presented in Tables 5 and 6.

Table 5.
Analysis of variance for testing Model 3

Source	DF	Sum of Squares	Mean Square	F Value	P-value
Model	2	16.7392	8.3696	481.41	<.0001
Error	15	0.2607	0.0173		
Corrected Total	17	17.0000			

R-squares=0.9847

Table 6.
Parameter estimated and testing for partial parameter of Model 3

Parameter	DF	Estimate	Standard Error	t Value	P-value
EPI	1	0.2976	0.1468	2.03	0.0608
OC	1	0.6997	0.1468	4.77	0.0003

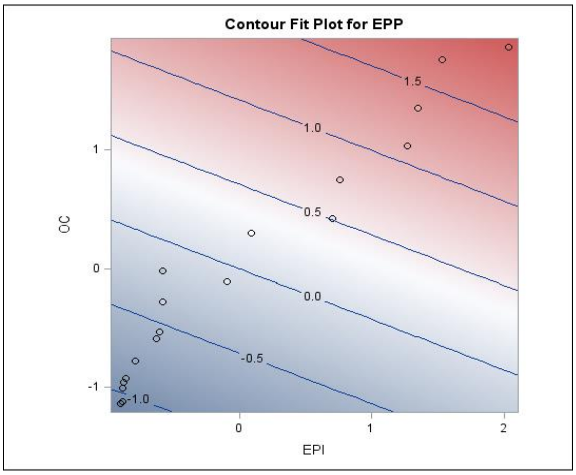

Figure 4. Contour plot of Model 3

Table 5 presented the analysis of variance for testing the model 3, with the null hypothesis there is no direct effect of EPI and OC to EPP, The results F-test=481.41 with P <0.0001, therefore the null hypothesis is rejected and we concluded that there is direct effect of EPI and OC to EPP. The R-Square=0.9847, this mean that 98.47% of the variation of EPP can be accounted for by EPI and OC or by the model. Table 6 shows the results of parameters estimation and testing partial parameter in model 2 are $p_{42} = 0.2976$ and $p_{43} = 0.6997$ with P=0.0608 and P = 0.0003, respectively. Therefore the null hypothesis $Ho : p_{42} = 0$ is not rejected, but the null hypothesis $Ho : p_{43} = 0$ is rejected. Even though the parameter p_{42} is not rejected, but the value of $p_{42} = 0.2976$ is greater than 0.05 in absolute value which according to Pedhazur[17] and Heisse[28] is meaningfulness.

Figure 4 shows the contour plot of model 3 also indicates positive trend if the value of EPI increase, the value of EPP increase, other variable is kept constant. The trend indicates positive if the value of OC increase, the value of EPP moves increase indicated by the increase of straight line as OC increase and the other variable is kept constant. The estimation of Model 3 is:

EPP = 0.2976 EPI + 0.6997 OC.

The unexplained variation is

$$p_3 = \sqrt{1 - 0.9847} = 0.1237$$

The results of analysis for Model 4 are presented in Tables 7 and 8.

Table 7.
Analysis of variance for testing Model 4

Source	DF	Sum of Squares	Mean Square	F Value	P-value
Model	2	16.7788	8.3894	569.05	<.0001

Source	DF	Sum of Squares	Mean Square	F Value	P-value
Error	15	0.2211	0.0147		
Corrected Total	17	17.0000			

R-square=0.9869

Table 8.
Parameter estimated and testing for partial parameter of Model 4

Parameter	DF	Estimate	Standard Error	t Value	P-value
OC	1	0.1202	0.2106	0.57	0.5765
EPP	1	0.8742	0.2106	4.15	0.0009

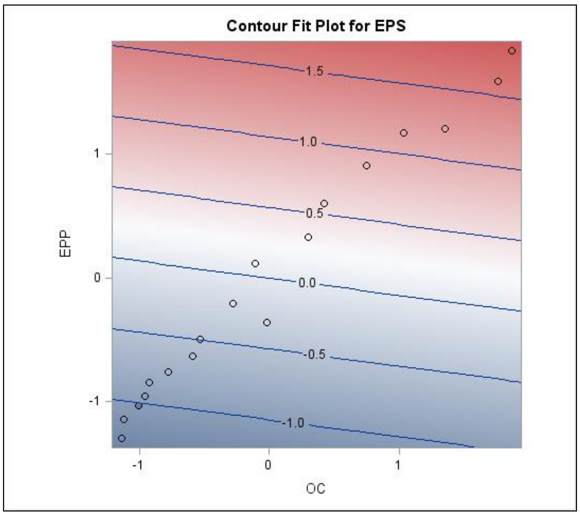

Figure 5. Contour plot of Model 4

Table 7 presented the analysis of variance for testing the model 4, with the null hypothesis there is no direct effect of OC and EPP to EPS, The results F-test=569.05 with P <0.0001, therefore the null hypothesis is rejected and we concluded that there is direct effect of OC and EPP to EPS. The R-Square=0.9869, this mean that 98.69% of the variation of EPS can be accounted for by OC and EPP or by the model 4. Table 8 shows that the results of the parameters estimation and the partial test of parameters of Model 4 are P_{53}=0.1202 and P_{54}=0.1202, with P=0.5765 and P=0.000 respectively. Therefore, the null hypothesis HO:P_{53}=0 is not rejected, but the null hypothesis HO:P_{54}=0 is rejected. Even though the parameter P_{53} is not rejected, the value of P_{53}=0.1202 is greater than 0.05 in absolute value, which is meaningful according to Pedhazur [17] and Heisse [28]. Figure 5 shows the contour plot of Model 4 and indicates a positive trend in that if the value of OC increases, the value of EPS also increases while the other variable is kept constant. The trend is positive if as the value of EPP increases, the value of EPS also increases as indicated by the ascending straight line as EPP increases and the other variable is kept constant.

The estimation of Model 4 is:

S = 0.1202 OC + 0.8742 EPP

The unexplained variation is:

$$p_4 = \sqrt{1-0.9869} = 0.1144$$

The results of the analysis of Model 5 are presented in Tables 9 and 10.

Table 9.
Analysis of variance for testing Model 5

Source	DF	Sum of squares	Mean square	F value	P-value
Model	2	15.927	7.963	111.35	<.0001
Error	15	1.072	0.071		
Corrected Total	17	17.000			

R-square=0.9369

Table 10.
Parameter estimated and testing for partial parameter of Model 5

Parameter	DF	Estimate	Standard error	t Value	P-value
OC	1	1.2908	0.3879	3.33	0.0046
EPS	1	-0.3291	0.3879	-0.85	0.4096

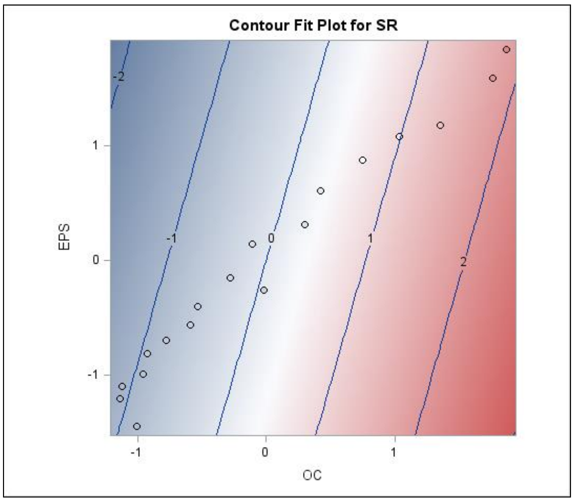

Figure 6. Contour plot of Model 5

Table 9 presents the analysis of variance for testing Model 5, with the null hypothesis that there is no direct effect of OC and EPS on SR. The results are F-test=111.35 with P <0.0001. Therefore, the null hypothesis is rejected, and we conclude that there is a direct effect of OC and EPS on SR. The R-square=0.9369, which means that 93.69% of the variation of SR can be accounted for by OC and EPS or by Model 5.

Table 8 shows that the results of the parameters estimation and the partial test of the parameters of Model 5 are P_{63}=1.2908 and P_{65}=-0.3291 with P=0.0046 and P=4096, respectively. Therefore, the null hypothesis HO:P_{63}=0 is rejected, but the null hypothesis HO:P_{65}=0 is not rejected. Even though the parameter P_{65} is not rejected, the value of P_{65}=-0.3291 is greater than 0.05 in absolute value, which is meaningful according to Pedhazur [17] and Heisse [28]. Figure 6 shows the contour plot of Model 5 and indicates a positive trend in that if the value of OC increases, the value of SR also increases while the other variable is kept constant. The trend is negative if as the value of EPS increases, the value of SR decreases as indicated by the decline of the straight line as EPS increases and the other variable is kept constant.

The estimation of Model 5 is:

$$SR = 1.2908\ OC - 0.3291\ EPS$$

The unexplained variation is

$$p_5 = \sqrt{1-0.9369} = 0.2512$$

Table 11.
Pearson correlation coefficients, N = 18 Prob > |r| under Ho: Rho=0

	NS	EPI	OC	EPP	EPS	SR
NS	1.000	0.8781	0.8142	0.7945	0.7698	0.8654
		<.0001	<.0001	<.0001	0.0002	<.0001
EPI		1.0000	0.9759	0.9805	0.9691	0.9563
			<.0001	<.0001	<.0001	<.0001
OC			1.0000	0.9901	0.9859	0.9663
				<.0001	<.0001	<.0001
EPp				1.0000	0.9933	0.9489
					<.0001	<.0001
EPS					1.0000	0.9435
						<.0001

A. Decomposition of Correlation into Direct and Indirect Effects

The decomposition of the correlation between NS and OC (r_{13}), EPI and OC (r_{23}), EPI and EPP (r_{24}), OC and EPP (r_{34}), OC and EPS (r_{35}), EPP and EPS (r_{45}), OC and SR (r_{36}), and between EPS and SR (r_{56}) and its numerical quantity and meaning are given in Tables 12, 13, 14, 15, 16, 17, 18, and Table 19 respectively.

Table 12.
Decomposition of correlation between NS and OC,
$r_{13} = p_{31} + p_{32}\, p_{21}$

Component	Numerical quantity	Meaning
p_{31}	-0.1872	Because NS has direct effect to OC.
$p_{32} \cdot p_{21}$	1.0013	Because NS has direct effect to EPI and EPI has direct effect to OC.
Total (r_{13})	0.8142	Correlation between NS and OC.

Table 13.
Decomposition of correlation between EPI and OC, $r_{23} = p_{31}\, p_{21} + p_{32}$

Component	Numerical quantity	Meaning
$p_{31} \cdot p_{21}$	-0.1643	Because NS has direct effect to EPI and OC.
p_{32}	1.1403	Because EPI has direct effects to OC
Total (r_{23})	0.9760	Correlation between EPI and OC

Table 14.
Decomposition of correlation between EPI and EPP, $r_{24} = p_{42} + p_{43}\, p_{31}\, p_{21} + p_{43}\, p_{32}$.

Component	Numerical quantity	Meaning
p_{42}	0.2977	Because EPI has direct effect to EPP.
$p_{43} \cdot p_{31} p_{21}$	-0.1150	Because NS has direct effects to EPI and OC, and OC has direct effects to EPP.
$p_{43} \cdot p_{32}$	0.7978	Because EPI has direct effects to OC, and OC has directs effects to EPP.
Total (r_{24})	0.9805	Correlation between EPI to EPP.

Table 15.
Decomposition of correlation between OC and EPP, $r_{34} = p_{42}\, p_{31}\, p_{21} + p_{42}\, p_{32} + p_{43}$

Component	Numerical quantity	Meaning
$p_{42} \cdot p_{21} \cdot p_{31}$	-0.0489	Because NS has direct effects to OC and EPI, and EPI has direct effects to EPP.
$p_{42} \cdot p_{32}$	0.3394	Because EPI has direct effects to OC and EPP.
p_{43}	0.6997	Because OC has direct effects to EPP.
Total (r_{34})	0.9902	Correlation between OC to EPP

Table 16.
Decomposition of correlation between OC and EPS,
$r_{35} = p_{53} + p_{54}\, p_{42}\, p_{31}\, p_{21} + p_{54}\, p_{42} p_{32} + p_{54}\, p_{43}$

Component	Numerical quantity	Meaning
p_{53}	0.1202	Because OC has direct effects to EPS.

$p_{54} \cdot p_{42} \cdot p_{31} \cdot p_{21}$	-0.0427	Because NS has direct effects to EPI and OC, and EPI has direct effects to EPP, and EPP has direct effects to EPS.
$p_{54} \cdot p_{42} p_{32}$	0.2967	Because EPI has direct effect to OC and EPP, and EPP has direct effects to EPS.
$p_{54} \cdot p_{43}$	0.6117	Because OC has direct effects to EPP, and EPP has direct effects to EPS.
Total (r_{35})	0.9859	Correlation between OC and EPS.

Table 17.
Decomposition of correlation between EPP and EPS,
$r_{45} = p_{53}\, p_{42}\, p_{31}\, p_{21} + p_{53}\, p_{42} p_{32} + p_{53}\, p_{43} + p_{54}$

Component	Numerical quantity	Meaning
$p_{53} \cdot p_{42} \cdot p_{53} \cdot p_{42} \cdot p_{31} \cdot p_{21}$	-0.0058	Because NS have direct effect to EPI and OC, and EPI has direct effects to EPP, and OC has direct effects to EPS.
$p_{53} \cdot p_{42} \cdot p_{32}$	0.0408	Because EPI have direct effect to OC and EPP, and OC has direct effect to EPS
$p_{53} \cdot p_{43}$	0.0841	Because OC have direct effect to EPP and EPS.
p_{54}	0.8742	Because EPP has direct effects to EPS.
Total (r_{45})	0.9933	Correlation between EPP and EPS.

Table 18.
Decomposition of correlation between OC and SR,
$r_{36} = p_{63} + p_{65} \cdot p_{53} + p_{65}\, p_{54}\, p_{42}\, p_{31}\, p_{21} + p_{65}\, p_{54}\, p_{42} p_{32} + p_{65}\, p_{54}\, p_{43}\,.$

Component	Numerical quantity	Meaning
p_{63}	1.2908	Because OC has direct effects to SR
$p_{65} \cdot p_{53}$	-0.0395	Because OC has direct effects to EPS, and EPS has direct effects to SR.
$p_{65} \cdot p_{54} \cdot p_{42} \cdot p_{31} \cdot p_{21}$	0.0140	Because NS have direct effect to EPI and OC, EPI has direct effects to EPP, EPP has direct effects to EPS, and EPS has direct effects to SR.
$p_{65} \cdot p_{54} \cdot p_{42} \cdot p_{32}$	-0.0976	Because EPI have direct effect to OC and EPP, EPP has direct effects to EPS, and EPS has direct effects to SR.
$p_{65} \cdot p_{54} \cdot p_{43}$	-0.2013	Because OC has direct effects to EPP, EPP has direct effects to EPS, and EPS has direct effects to SR.
Total (r_{36})	0.9664	Correlation between OC and SR.

Table 19.
Decomposition of correlation between EPS and SR,
$r_{56} = p_{63} p_{53} + p_{63}\, p_{54}\, p_{42}\, p_{31}\, p_{21} + p_{63}\, p_{54}\, p_{42} p_{32} + p_{63}\, p_{54}\, p_{43} + p_{65}\cdot$

Component	Numerical quantity	Meaning
$p_{63} \cdot p_{53}$	0.1552	Because OC have direct effect to EPS and SR.
$p_{63} \cdot p_{54} \cdot p_{42} \cdot p_{31} \cdot p_{21}$	-0.0552	Because NS have direct effect to EPI and OC, EPI has direct effects to EPP, EPP has direct effects to EPS, and OC has direct effects to SR.
$p_{63} \cdot p_{54} \cdot p_{42} \cdot p_{32}$	0.3829	Because EPI have direct effect to OC and EPP, EPP has direct effects to EPS, and OC has direct effects to SR.
$p_{63} \cdot p_{54} \cdot p_{43}$	0.7896	Because OC has direct effects to EPP, EPP has direct effects to EPS, and OC has direct effects to SR.
p_{65}	-0.3291	Because EPS has direct effects to SR
Total (r_{56})	0.9435	Correlation between EPS and SR

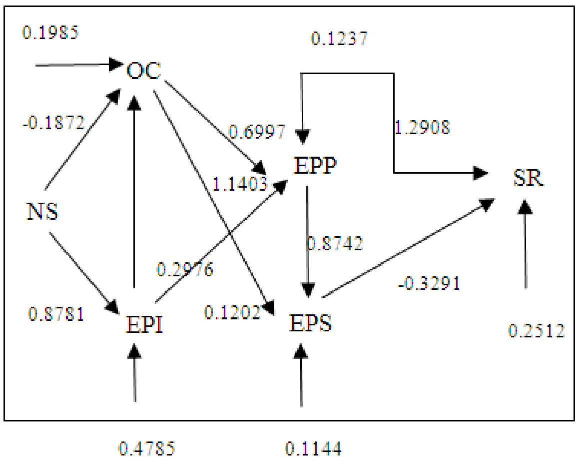

Figure 7. The estimation of the parameters of the model of causal relationships among the important variables of the Public Electric Company of Indonesia. The variables are Total number of Staff or Employees (NS), Electric Power Installed in MW (EPI), Operating Cost (Million Rp) (OC), Electric Power Product in MWh (EPP), Electric Power Sales in MWh (EPS), and Sales Revenue in Million Rp (SR)

B. Direct Effect, Indirect Effect and Total Effect

From Figure 7, Table 2, and Model 1, it can be seen that NS has a direct effect (DE) on EPI of as much as P_{21}=0.8781. It can be concluded that NS has a positive direct effect on EPI. From Figure 7, Table 4, and Model 2 it is clear that NS has a direct effect on OC of as much as P_{31}=-0.1872 and has an indirect effect (IDE) on OC through EPI of as much as P_{21}-

P_{32}=0.8781x1.1404=1.0013. The total effect (TE) of NS on OC is $P_{31}+P_{21}.P_{32}$=0.8141. The direct effect of NS on OC is negative, but the total effect of NS on OC is positive. From Figure 7, Table 6, and Model 3, it is clear that EPI has direct effect on EPP of as much as P_{42}=0.2976 and an indirect effect (IDE) on EPP through OC of P_{32}-P_{43}=1.1403x0.6997=0.7978. The total effect (TE) of EPI on EPP is $P_{42}+P_{32}.P_{43}$=1.0955. The direct effect and total effect of EPI on EPP is positive. While OC has a direct effect on EPP of as much as P_{43}=0.6997 and is positive. From Figure 7, Table 8, and Model 4, it is clear that OC has a direct effect on EPS of as much as P_{53}=0.1202 and has an indirect effect (IDE) on EPS through EPP of as much as $P_{43}.P_{54}$=0.6997x0.8742=0.6117. The total effect (TE) of OC on EPS is $P_{53}+P_{43}.P_{54}$=0.7318. The direct effect and total effect of OC on EPS is positive. EPP has a direct effect on EPS of as much as P_{54}=0.8742 and is positive. From Figure 7, Table 10, and Model 5, it is clear that OC has a direct effect on SR of as much as P_{63}=1.2908 and has an indirect effect (IDE) on SR through EPS of as much as $P_{53}.P_{65}$=0.1202(-0.3291)=-0.0395, and through EPP and EPS, it is as much as $P_{43}.P_{54}.P_{65}$=0.6997x0.8742x(-0.3291)=-0.2833. Therefore, the total effect (TE) of OC on SR is $P_{63}+P_{53}$-$P_{65}+P_{43}.P_{54}.P_{65}$=0.9679. The direct effect and total effect of OC on SR is positive, while EPS has a direct effect on SR of as much as P65=-0.3291, and it is positive.

IV. CONCLUSION

This study investigated the causal relationship among variables that are important indicators of the success of the Public Electric Company of Indonesia. These variables included: Total number of Staff or Employees (NS), Electric Power Installed (EPI), Operating Cost (OC), Electric Power Product (EPP), Electric Power Sales (EPS), and Sales Revenue (SR). From the proposed causal model in this study and the results of the analysis, we can conclude that there are direct effects from the Total number of Staff or Employees (NS) on the Electric Power Installed (EPI); there is a direct effect of the Total number of Staff or Employees (NS) and Electric Power Installed (EPI) on the Operating Cost (OC); There is a direct effect of the Electric Power Installed in MW (EPI) and Operating Cost (OC) on the Electric Power Product (EPP); there a is direct effect of Operating Cost (OC) and Electric Power Product (EPP) on the Electric Power Sales (EPS); and there is a direct effect of Operating Cost (Million Rp) (OC) and Electric Power Sales (EPS) on the Sales Revenue (SR).

ACKNOWLEDGEMENT

The authors would like to thank BPS (Central Bureau of Statistics) Indonesia for providing data in this study. The authors also would like to thank the anonymous referees which have given many suggestions in improving this paper.

REFERENCES

[1] WRIGHT, S. (1921) Correlation and causation. *Journal of Agricultural Research*, 20, pp. 557-585.

[2] WRIGHT, S. (1934) The method of path coefficients. *Annals of Mathematical Statistics*, 5, pp. 161-215.

[3] VOGLER, G. P. (1985) Multivariate path analysis of familial resemblance. *Genetic Epidemiology*, 2, pp. 35-53.

[4] ASHER, H. B., (1983) *Causal modeling*. 2nd ed. Newbury Park, California: SAGE Publication.

[5] BRYANT, L.T. (1977) A path analysis model for secondary physics enrollments. *Journal of Research in Science Teaching*, 14, pp. 177-189.

[6] STANLEY, J.S., and DORAN, R.L. (1996) Replication of a path analysis model of secondary physics enrollments: 20 years later. *Journal of Research in Science Teaching*, 33 (5), pp. 501-512.

[7] BERINSKY, A.J. and LENZ, G.S. (2011) Education and political participation: Exploring the causal link. *Political Behavior*, 33, pp. 357-373.

[8] ALKER, H.R. (1969) Statistics and Politics: the need for causal data analysis. In: LIPSET, S.M. (ed.) *Politics and the social sciences*. New York: Oxford University Press, pp. 244-313.

[9] BLALOCK, H.M. (1971) *Causal models in the social sciences*. Chicago, Illinois: Aldine-Atherton.

[10] ALWIN, D.F. and HAUSER, R.M. (1975) The decomposition of effects in path analysis. *American Sociological Review*, 40, pp. 37-47.

[11] BERK, R.A. (1988) Causal inference for sociological data. In: SMELSER, N.J. (ed.) *Handbook of Sociology*. Newbury Park, California: Sage, pp. 155-172.

[12] MORGAN, S.L. and WINSHIP, C. (2015) *Counterfactuals and causal inference methods and principles for social research.* 2nd ed. New York: Cambridge University Press.
[13] BAGOZZI, R.P. (1980) *Causal models in marketing.* New York: John Wiley.
[14] GAMMERMAN, A. (1999) *Causal models and intelligent data management.* Berlin, Heidelberg: Springer-Verlag.
[15] WARSONO, E.R., WAMILIANA, M.U., WIDIARTI, F.F., and ELFAKI, A. (2020) Causal modeling of the effect of foreign direct investment, industry growth and energy use to carbon dioxide emissions. *International Journal of Energy Economics and Policy*, 10 (3), pp. 348-354.
[16] FIELDS, S.J., LIVSHITS, G., SIROTTA, I., and MERLOB, P. (1996) Path analysis of risk factors leading to premature birth. *American Journal of Human Biology*, 8, pp. 433-443.
[17] PEDHAZUR, E.J. (1997) *Multiple regression in behavioral research: explanation and prediction.* 3rd ed. Belmont, California: Wadsworth Publishing.
[18] WONNACOTT, T.H. and WONNACOTT, R.J. (1990) *Introductory Statistics.* 5th ed. New York: John Wiley and Sons.
[19] LOETHER, H.J. and MCTAVISH, D.G. (1980) *Descriptive and inferential statistics: An introduction.* 2nd ed. Boston, Massachusetts: Allyn and Bacon, Inc.
[20] BLALOCK, H.M. (1964) *Causal inferences in nonexperimental research.* Chapel Hill, North Carolina: University of North Carolina Press.
[21] RETHERFORD, R.D. and CHOE, M.K. (1993) *Statistical models for causal analysis.* New York: John Wiley and Sons, Inc.
[22] RUSSO, F. (2009) *Causality and Causal Modelling in the Social Sciences: Measuring Variations.* New York: Springer.
[23] VAN METER, D.S. and ASHER, H.B. (1973) Causal Analysis: its promise for policy studies. *Policy Studies Journal*, 2, pp. 103-109.
[24] LOEHLIN, J.C. (2004) *Latent variable models: An introduction to factor, path, and structural equation analysis.* 4th ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates Publishers.
[25] WRIGHT, S. (1960) Path coefficients and path regressions: Alternative or complementary concepts. *Biometrics*, 16, pp. 189-202.
[26] WONNACOTT, T.H. and WONNACOTT, R.J. (1981) *Regression: Second course in statistics.* New York: John Wiley and Sons.
[27] BADAN PUSAT STATISTIK (2020) *Selected Indicators of Public Electricity Company 2000-2018.* [Online] Available from: https://www.bps.go.id/dynamictable/2019/03/13/1590/indikator-penting-perusahaan-listrik-2000-2017.html [Accessed 12/10/20].
[28] HEISSE, D.R. (1969) Problems in path analysis and causal inference. In: BORGATTA, E.F. (ed.) *Sociological methodology*. San Francisco, California: Jossey-Bass, pp. 56-67.

参考文:

[1] WRIGHT，S.（1921）相关和因果关系。农业研究杂志，20，第 557-585 页。
[2] WRIGHT，S.（1934）路径系数的方法。数学统计年鉴，5，第 161-215 页。
[3] VOGLER，G.P.（1985）家族相似性的多路径分析。遗传流行病学，2，第 35-53 页。
[4] ASHER，H. B.（1983）因果模型。第二版。加利福尼亚州纽伯里公园市：贤者出版物.
[5] BRYANT，L.T。（1977）中学物理招生的路径分析模型。科学教学研究杂志，14，第 177-189 页。
[6] STANLEY，J.S. 和 DORAN，R.L.（1996）复制中学物理入学人数的路径分析模型：20 年后。科学教学研究杂志，33（5），第 501-512 页。
[7] BERINSKY，A.J。和 LENZ，G.S.（2011）教育和政治参与：探讨因果关系。政治行为，33，第 357-373 页。
[8] ALKER，H.R.（1969）统计与政治：因果数据分析的必要性。在：LIPSET，

S.M. （编。）政治与社会科学。纽约：牛津大学出版社，第 244-313 页。
[9] BLALOCK，H.M。（1971）社会科学中的因果模型。伊利诺伊州芝加哥：阿尔丁· 阿瑟顿。
[10] D.F. ALWIN 和 HAUSER，R.M.（1975）路径分析中的影响分解。美国社会学评论，40，第 37-47 页。
[11] BERK，R.A。（1988）社会学数据的因果推论。在：SMELSER，N.J。（编辑）的社会学手册中。加利福尼亚州纽伯里公园市：贤哲，第 155-172 页。
[12] MORGAN，S.L. 和 WINSHIP，C.（2015）社会研究的反事实和因果推理方法和原则。第二版。纽约：剑桥大学出版社。
[13] BAGOZZI，R.P.（1980）市场营销中的因果模型。纽约：约翰·威利（约翰· 威利）。
[14] GAMMERMAN，A。（1999）因果模型和智能数据管理。柏林，海德堡：施普林格出版社。
[15] E.R. WARSONO，M.U。WAMILIANA，F.F。WIDIARTI 和 A. ELFAKI（2020）外商直接投资，产业增长和能源使用对二氧化碳排放的影响的因果模型。国际能源经济与政策杂志，10（3），第 348-354 页。
[16] FIELDS，S.J.，LIVSHITS，G.，SIROTTA，I., 和 MERLOB，P.（1996）导致早产的危险因素的路径分析。美国人类生物学杂志，8，第 433-443 页。
[17] PEDHAZUR，E.J。（1997）行为研究中的多元回归：解释和预测。第三版。加利福尼亚贝尔蒙特：沃兹沃思出版。
[18] WONNACOTT，T.H。和 R.J. WONNACOTT（1990）介绍性统计。第五版。纽约：约翰·威利父子。
[19] LOETHER，H.J. 和 MCTAVISH，D.G.（1980）描述性和推理性统计：简介。第二版。马萨诸塞州波士顿：艾琳和培根公司.
[20] BLALOCK，H.M。（1964）非实验研究中的因果推论。北卡罗来纳州教堂山：北卡罗来纳大学出版社。
[21] R.D. RETHERFORD 和 M.K. CHOE（1993）因果分析的统计模型。纽约：约翰·威利父子公司
[22] RUSSO，F.（2009）社会科学中的因果关系和因果模型：衡量变异。纽约：施普林格。
[23] D.S. VAN METER 和 H.B. ASHER。（1973）因果分析：它对政策研究的承诺。政策研究期刊，2，第 103-109 页。
[24] LOEHLIN，J.C.（2004）潜变量模型：因子，路径和结构方程分析的简介。第四版。新泽西州希尔斯代尔：劳伦斯·埃尔鲍姆协会出版社。
[25] WRIGHT，S。（1960）路径系数和路径回归：替代或补充概念。生物测定学，16，第 189-202 页。
[26] WONNACOTT，T.H。和 R.J. WONNACOTT（1981）回归：统计学第二门课程。纽约：约翰·威利父子。
[27] 中央统计局（2020）2000-2018 年公共电力公司的某些指标。[在线]可从以下网站获得：HTTPS://WWW.BPS.GO.ID/DYNAMICTABLE/2019/03/13/1590/INDIKATOR-PENTING-PERUSAHAAN-LISTRIK-2000-2017.HTML [访问时间 12/10/20]。
[28] HEISSE，D.R.（1969）路径分析和因果推理中的问题。在：博尔加塔，E.F。（编辑）社会学方法论。加利福尼亚州旧金山：乔西· 贝斯，第 56-67 页。