

# The Importance of the Effect of Exogenous Interaction Factors on Endogenous Variables in Accounting Modeling

Dodik Siswanto<sup>1,\*</sup>, I Gusti Ngurah Agung<sup>2</sup>

<sup>1</sup>Department of Accounting, Universitas Indonesia, Depok, 16424, Indonesia  
<sup>2</sup>Department of Management, Universitas Indonesia, Depok, 16424, Indonesia

**Abstract** Most researchers in accounting use models having dummy independent variables to divide the research objects into two or more groups of objects or individuals. In other researches, they employ more than one variable in the model. However, the dummy variable(s) that stand alone may implicate improper correlations between the numerical independent variables with the corresponding dependent variable within the groups, namely the numerical independent variables have the same effects on the dependent variables within all groups of individuals. To overcome this problem, interaction factors between the dummies and the numerical variable(s) should be used as additional independent variables. So that the models considered would be heterogeneous regression models. It is found, only very few researchers employ the heterogeneous regression models, or the interaction models in general, in top accounting journal, recently. This paper attempts to show that the absence of interaction models may mislead the causal relationship between the independent and dependent variables, specifically between the numerical variables, of the models. The main objective of this paper is to present the importance of interaction models should be applied as the basic statistical models. It is expected that this paper can contribute for a better understanding to the students and less experience researchers in developing the statistical models, in various field of studies.

**Keywords** Accounting Model, Interaction Models, Causal Relationship

## 1. Introduction

It seems that there is no procedural standard when we use one or two dummy variables in the accounting model for research. Some books of SPSS or Eviews do not treat this specifically when there is a dummy variable used. Actually, when we employ a dummy variable we assume that there is a difference between the groups of the research objects generated by the dummy variable. In fact, some researchers may not consider those differences to present their models. This is actually important tools to see whether the dummy variable is needed or otherwise.

This paper tries to see the importance of the counting model having interaction independent variable(s), in a theoretical sense, since a number of additive counting models have been presented in accounting journal recently. For this paper, we have observed a total of 268 papers in four international journals, namely the International Journal of Accounting (IJA), Journal of Accounting and Economics (JAE), British Accounting Review (BAR) and Advances in Accounting (AA), from 2008 to 2010.

There are a total of 63 papers use dummy variable, but only five are using interaction models, where two of the five

interaction models are full interaction models, as presented in table 1.

Note that this table shows that the only Journal of Accounting and Economics and International Journal of Accounting have been presenting the interaction models. This may be caused the need of categorization to differentiate the inherent characteristics for each journal.

**Table 1.** The total number of models by their classification, presented in four journals, observed in the last three years

	IJA	JAE	BAR	AA	Total
Total	49	95	54	70	268
Non-interaction model	22	24	5	12	63
Interaction model	1	2	0	0	3
Full interaction model	1	1	0	0	2

IJA= International Journal of Accounting  
JAE= Journal of Accounting and Economics  
BAR= British Accounting Review  
AA= Advances in Accounting

An interaction models, including the heterogeneous regression model, should be considered as the very basic statistical model, in general. Since if we have several exogenous (cause, source, upstream or independent) variables, then it can easily be identified, in a theoretical sense, that the effect of an exogenous variable should depend on at least one of the others.

Hence, the main objective of this paper is to present the basic theoretical concepts of heterogeneous regression

\* Corresponding author:

dodik.siswanto@ui.ac.id (Dodik Siswanto)

Published online at <http://journal.sapub.org/ijfa>

Copyright © 2012 Scientific & Academic Publishing. All Rights Reserved

models, the interaction models in general. Furthermore, it is expected that this paper would contribute for a better understanding to the students, as well as less experience researchers, in developing the statistical models in various field of studies.

## 2. Basic Theoretical Concept of the Interaction Models

Corresponding to the interaction models, Agung (2010, 2009, 2008 and 2006) has demonstrated the applications of interaction models, even for the models having two or three independent variables, either categorical or numerical variables, where each categorical should be transformed to dummy variables. In fact, heterogeneous regressions model was presented by Johnson & Neyman in 1936 (quoted by Huitema, 1980, p.270). So that the interaction models have been applied for more than 75 years, hence, the suggestion for using interaction model(s) may be important issue for enriching the research. This section presents the basic theoretical concept of the interaction models, which could easily be extended to higher order interaction models with many independent variables.

### 2.1. Models Having Two Dichotomous Independent Variables

Having a numerical endogenous variable Y, including the zero-one problem indicator, and two dichotomous exogenous variables A and B, then the model should be considered is a two-way ANOVA model. Note that the dichotomous variables A and B, respectively, should be represented using pairs of dummy variables, namely (DA1, DA2), and (DB1, DB2), then the following alternative two-way interaction models can be applied (Agung, 2010).

(a) A pure interaction model:

$$Y = \beta_0 + \beta_1 * DA1 + \beta_2 * DA2 + \beta_3 * DA1 * DA2 + \epsilon \quad (1)$$

Table 2. Difference in differences model 1

	DB1=1, DB2=0	DB1=0, DB2=1	Diff.
DA1=1, DA2=0	$\beta_0 + \beta_1$	$\beta_0 + \beta_3$	$\beta_1 - \beta_3$
DA1=0, DA2=1	$\beta_0 + \beta_2$	$\beta_0$	$\beta_2$
Diff.	$\beta_1 + \beta_2$	$\beta_3$	$DID = \beta_1 - \beta_2 - \beta_3$

By applying pure interaction model would cause  $(\beta_1 - \beta_2 - \beta_3)$  difference in differences (DID). Model 1 gives meaningful result in DID.

(b) A non-hierarchical model:

$$Y = \beta_0 + \beta_1 * DA1 + \beta_2 * DA1 * DB1 + \beta_3 * DA2 * DB1 + \epsilon \quad (2)$$

Model with non-hierarchical would result different DID  $(\beta_2 - \beta_3)$ . Model with dichotomous variable should interact with all other variable to give better DID.

(c) Another non-hierarchical model:

$$Y = \beta_0 + \beta_1 * DB1 + \beta_2 * DA1 * DB1 + \beta_3 * DA1 * DB2 + \epsilon \quad (3)$$

(d) A full-factorial or hierarchical model:

$$Y = \beta_0 + \beta_1 * DA1 + \beta_2 * DB1 + \beta_3 * DA1 * DB1 + \epsilon \quad (4)$$

Table 3. Difference in differences model 2

	DB1=1, DB2=0	DB1=0, DB2=1	Diff.
DA1=1, DA2=0	$\beta_0 + \beta_1 + \beta_2$	$\beta_0 + \beta_1$	$\beta_2$
DA1=0, DA2=1	$\beta_0 + \beta_3$	$\beta_0$	$\beta_3$
Diff.	$\beta_1 + \beta_2 - \beta_3$	$\beta_1$	$DID = \beta_2 - \beta_3$

Note that all models should have four parameters, namely  $\beta_0, \beta_1, \beta_2,$  and  $\beta_3$  which equals to the number of cells generated by the factors A and B, and they are representing the effect of the factor A (or B) on Y depends on the factor B (or A), in a theoretical sense. So that, the following additive model should be considered an improper or worst model.

$$Y = \beta_0 + \beta_1 * DA1 + \beta_2 * DB1 + \epsilon \quad (5)$$

In other words, this model has hid assumptions, in addition to the basic assumption of its error terms, since the model has only three parameters for the four cells generated by the factors A and B. DID would give 0 which means meaningless as dichotomous variables do not have each interaction. Dichotomous variable is the lowest group under ordinal, interval and ratio data. Therefore, it should be grouped with interaction for specific case.

Table 4. Difference in differences model 5

	DB1=1	DB1=0	Diff.
DA1=1	$\beta_0 + \beta_1 + \beta_2$	$\beta_0 + \beta_2$	$\beta_1$
DA1=0	$\beta_0 + \beta_1$	$\beta_0$	$\beta_1$
Diff.	$\beta_2$	$\beta_2$	$DID = 0$

### 2.2. Models Having a Numerical and a Dichotomous Independent Variables

Having a numerical endogenous variable Y, including the zero-one problem indicator, a numerical exogenous variable X, and a dichotomous exogenous variables A, then in general X has different effects on Y between the two groups generated by the factor A. In this case then the simplest heterogeneous regressions can be applied, which has the following alternative forms.

(a) A model using a dummy variable with intercept parameter:

$$Y = \beta_1 + \beta_2 * DA1 + \beta_3 * X + \beta_4 * X * DA1 + \epsilon \quad (6)$$

(b) A model using two dummy variables without intercept parameter.

$$Y = (\beta_1 + \beta_2 X) * DA1 + (\beta_3 + \beta_4 X) * DA2 + \epsilon \quad (7)$$

### 2.3. Models Having Two Numerical Independent Variables

It is recognized that if we have a numerical endogenous variable Y, and two numerical exogenous variables X1 and X2, then in most cases the linear effect of an exogenous variables on Y, depends of the other exogenous variable, in a

theoretical sense. Thence, the interaction factor  $X1*X2$  should be used as an independent variable of any models considered. In this case, it is recommended to apply the following simple alternative two-way interaction models.

(a) A hierarchical or full interaction model:

$$Y = \beta_0 + \beta_1 * X1 + \beta_2 * X2 + \beta_3 * X1 * X2 + \varepsilon \quad (8)$$

(b) Alternative reduced models:

$$Y = \gamma_0 + \gamma_1 * X1 + \gamma_2 * X1 * X2 + \varepsilon \quad (9)$$

$$Y = \delta_0 + \delta_1 * X2 + \delta_2 * X1 * X2 + \varepsilon \quad (10)$$

$$Y = \theta_0 + \theta_3 * X1 * X2 + \varepsilon \quad (11)$$

### 3. Interaction Models Presented in Selected Papers

This section presents selected papers from international journals that use interaction models. It will be explained the significant effect of interaction model, and the need of different test for each categorical variable. This is important to see whether the categorization of dummy variable is needed or otherwise. The complex problem may arise if more than one dummy variable are employed, the matrix table may be useful to see different test.

First, Francis & Martin (2010) present a three-way interaction model (12) having a single dummy variable (D), defined as  $D=1$  if the return is negative ( $R < 0$ ) and  $D=0$  if otherwise, and 24 parameters. In other words, the model has 23 independent variables. Note that the model has two-interaction variables, such as  $D_{i,t-1} * R_{i,t-1}$ ,  $AcqCAR_{i,t} * D_{i,t-1}$  and  $AcqCAR_{i,t} * R_{i,t-1}$ , and three-way interaction variables:  $AcqCAR_{i,t} * D_{i,t-1} * R_{i,t-1}$  and  $Leverage_{i,t-1} * D_{i,t-1} * R_{i,t-1}$ . Thence, it can be concluded that the numerical independent variables have different effects on  $X_{i,t-1}$  between the groups indicated by  $D=1$  and  $D=0$ . But they do not apply different test between the groups.

$$\begin{aligned} X_{i,t-1} = & \beta_1 + \beta_2 D_{i,t-1} + \beta_3 R_{i,t-1} + \beta_4 D_{i,t-1} * R_{i,t-1} + \beta_5 AcqCAR_{i,t} + \beta_6 AcqCAR_{i,t} * D_{i,t-1} + \beta_7 AcqCAR_{i,t} * R_{i,t-1} \\ & + \beta_8 AcqCAR_{i,t} * D_{i,t-1} * R_{i,t-1} + \beta_9 Leverage_{i,t-1} + \beta_{10} Leverage_{i,t-1} * D_{i,t-1} + \beta_{11} Leverage_{i,t-1} * R_{i,t-1} \\ & + \beta_{12} Leverage_{i,t-1} * D_{i,t-1} * R_{i,t-1} + \beta_{13} Log(assets)_{i,t-1} + \beta_{14} Log(assets)_{i,t-1} * D_{i,t-1} + \beta_{15} Log(assets)_{i,t-1} * R_{i,t-1} \\ & + \beta_{16} Log(assets)_{i,t-1} * D_{i,t-1} * R_{i,t-1} + \beta_{17} MB_{i,t-1} + \beta_{18} MB_{i,t-1} * R_{i,t-1} + \beta_{19} MB_{i,t-1} * D_{i,t-1} + \beta_{20} MB_{i,t-1} * R_{i,t-1} * D_{i,t-1} \\ & + \beta_{21} LIT_{i,t-1} + \beta_{22} LIT_{i,t-1} * D_{i,t-1} + \beta_{23} LIT_{i,t-1} * R_{i,t-1} + \beta_{24} LIT_{i,t-1} * D_{i,t-1} * R_{i,t-1} + e_{i,t} \end{aligned} \quad (12)$$

Kann & Watts (2009) also employ a three-way interaction model (13). It is presented that each of the interaction independent variables ( $Ret*Size$ ), ( $Ret*M/B$ ), ( $D*Ret$ ), ( $D*Ret*Size$ ), ( $D*Ret*Lev$ ) and ( $D*Size$ ) has a significant adjusted effects on X. But they do not apply different test between the groups indicated by the dummy variable D.

$$\begin{aligned} X_i = & \beta_1 + \beta_2 D_i + R_i * (\mu_1 + \mu_1 Size_i + \mu_3 M/B_i + \mu_1 Lev_i) + D_i * R_i * (\lambda_1 + \lambda_2 Size_i + \lambda_3 M/B_i + \lambda_4 Lev_i) \\ & + (\delta_1 Size_i + \delta_2 M/B_i + \delta_3 Lev_i + \delta_4 D_i * Size_i + \delta_5 D_i * M/B_i + \delta_6 D_i * Lev_i) + \varepsilon_i \end{aligned} \quad (13)$$

Shah, et.al (2009) employ a two-way interaction model (14) using S as a dummy variable, such that  $S=1$  if the firm classified as large,  $S=0$  otherwise. Note that the two-way interaction between S and each of the other variables, namely A, RD, E, BV, D and CC, indicate that the six variables have different effects on MV between the two groups of the firms indicated by the dummy variable S.

$$\begin{aligned} MV_{it} = & \alpha_0 + \alpha_{01} S_{it} + \alpha_1 A_{it} + \alpha_{11} S_{it} * A_{it} + \alpha_2 RD_{it} + \alpha_{21} S_{it} * RD_{it} + \alpha_3 E_{it} \\ & + \alpha_{31} S_{it} * E_{it} + \alpha_4 BV_{it} + \alpha_{41} S_{it} * BV_{it} + \alpha_5 D_{it} + \alpha_{51} S_{it} * D_{it} + \alpha_6 CC_{it} + \alpha_{61} S_{it} * CC_{it} + \varepsilon_{it} \end{aligned} \quad (14)$$

Jermias & Setiawan (2008) utilize a full interaction model (15) to see the effect of three independent variables to performance (Perform). They use two dummy variables, hierarchy and control system. In their paper, they do not conduct different test for each dummy variable, so we hardly to know whether this categorization is needed or otherwise. But all interaction variables show significant effect to Performance for dichotomous control. In this case, only (Control\*Participation) variable that do not show significant effect.

$$\begin{aligned} Perform_i = & \gamma_0 + \gamma_1 Hier_i + \gamma_2 Control_i + \gamma_3 Participation_i + \gamma_4 Control_i * Hier_i + \gamma_5 Control_i * Participation_i \\ & + \gamma_6 Hier_i * Participation_i + \gamma_7 Hier_i * Control_i * Participation_i + \varepsilon_i \end{aligned} \quad (15)$$

Lastly, Hugon & Muslu (2010) adopt a full interaction model, but the dummy is actually under interaction model for other variable (REV). Only variables (REV\*INS) and (REV\*CONSERV\*INS) show significant effect to CAR.

$$\begin{aligned} CAR_{j,t} = & \lambda_0 + \lambda_1 REV_{j,t}^i + \lambda_2 CONSERV_{t-1}^i + \lambda_3 INS_j^i + \lambda_4 REV_{j,t}^i * CONSERV_{t-1}^i + \lambda_5 REV_{j,t}^i * INS_{j,t-1}^i + \lambda_6 CONSERV_{t-1}^i * INS_{j,t-1}^i \\ & + \lambda_7 REV_{j,t}^i * CONSERV_{t-1}^i * INS_{j,t-1}^i + \sum_m \Pi m Controls + \sum_m \Im m REV_{j,t}^i * Controls + v_{j,t}^i \end{aligned} \quad (16)$$

From journals that use accounting model, we can see interaction variable can give significant effect to the model or dependent variable. Based on the interaction model, it can be shown that the absence of interaction model may mislead the analysis.

## 4. Conclusions

In a theoretical sense, interaction model should be applied if we have at least two cause or source variables, especially if we have categorical independent variable(s) which should be represented using dummy variables. This is because a dummy variable is a tool to see whether there is a significant effect among variables or otherwise. Thus, we should interact it to other variables as it attach to them directly. Mapping or matrix model may be a good explanation to see this.

Research results show that only few papers adopt a full interaction model when they employ a dummy variable/s. Significant effect of interaction variable may show that the interaction model is needed when we use a dummy variable(s).

Even though the interaction models are very importance models, we would never present empirical models using all possible interactions, if the models have three or more numerical independent variables. So that the researchers should select some important interactions, based on their best judgments, supported by their knowledge and experiences.

## ACKNOWLEDGEMENTS

This paper is a revision of the paper, which has been presented at The 3rd International *Accounting Conference* held by Universitas Indonesia, 27 - 28 October 2010. Bali, Indonesia.

---

## REFERENCES

- [1] Agung, I. G. N., Penerapan model rerata-sel multivariat dan model ekonometri dengan SPSS. Jakarta: Yayasan SAD Satria Bhakti, 2006.
- [2] Agung, I. G. N., Manajemen penulisan skripsi, tesis dan disertasi statistika: Kiat-kiat untuk mempersingkat waktu penulisan karya ilmiah yang bermutu. (4th ed.). Jakarta: PT RajaGrafindo Persada, 2008.
- [3] Agung, I. G. N., *Time series data analysis using E-Views*. Singapore: John Wiley & Sons, 2009.
- [4] Agung, I. G. N., Cross section and experimental data analysis using E-Views. Singapore: John Wiley & Sons, 2011.
- [5] Artur Hugon, Volkan Muslu, "Market demand for conservative analysts", *Journal of Accounting and Economics*, vol. 50, pp. 42–57, 2010.
- [6] Jere R. Francis, Xiumin Martin, "Acquisition profitability and timely loss recognition", *Journal of Accounting and Economics*, vol. 49, pp.161–178, 2010.
- [7] Johnny Jermias, Trisnawati Setiawan, "The moderating effects of hierarchy and control systems on the relationship between budgetary participation and performance", *The International Journal of Accounting*, vol. 43, pp. 268–292, 2008.
- [8] Huitema, B. E. *The Analysis of Covariance and Alternatives*. John Wiley & Sons, New York, 1980
- [9] Mozaffar Khan, Ross L. Watts, "Estimation and empirical properties of a firm-year measure of accounting conservatism", *Journal of Accounting and Economics*, vol. 48, pp. 132–150, 2009.
- [10] Syed Zulfiqar Ali Shah, Andrew W. Stark, Saeed Akbar, "The value relevance of major media advertising expenditures: Some U.K. Evidence", *The International Journal of Accounting*, vol. 44, pp.187–206, 2009