THE EFFECT OF OMITTED VARIABLE BIAS ON REGRESSION ANALYSIS TO EMPIRICAL ACCOUNTING RESEARCH

KAZUHIRO MANANBE

School of Contemporary International Studies, Nagoya University of Foreign Studies
E-mail: kmana@nufs.ac.jp

Abstract - This study examines the industry effect on financial variables and models the relation between financial variable and industry characteristic variables by econometrics approach. The results of this research show the following. Theoretical evidence gives economic meaning to the regression coefficients of industry dummy variables and suggests effectiveness of the industrial dummy variable as the control variable. It suggests that intercept coefficient estimates industry effect on financial variable for base group, and that each slope coefficient estimates additional industry effect on financial variable for each group. Furthermore, to control industry effect, it is only necessary to insert industry dummy variables to the regression model.In other words, we do not need to identify and measure industry characteristic variables.

Keywords - Value Relevance Research; Omitted Variable Bias; Industry Dummy Variable; Capital Markets-Based Accounting Research

I. INTRODUCTION

In recent year, significant advances have been made in the factor of corporate profitability and these impacts.This research issues have received scholarly attention of management, accounting, and industrial organization.In particular, Misangyi et al. (2006) investigate the effect of industry, corporate, segment on corporate profitability in a quantitative way.Their evidence indicated that industry could explain about 20 percent of the variability of corporate profitability.

Many research point out that there are the homogeneity of corporate fundamentals in each industry(Amit and Livant, 1990; Guenther and Rosman, 1994; Krishnan and Press, 2003; Kimura,2009; Aratani,2010).These research focus on industrial classifications which are defined by government department, stock exchange, database vendor, and investigate the reliability of these classifications by testing the homogeneity of financial ratio within industry.Based on the consequence of Kimura (2009), although there are the homogeneity of company size, liquidity, way of financing, there are not the homogeneity of corporate profitability, growth, ratios related to stock market.

These relationship between financial variable and industry characteristics might request a review of prior statistical method and erode trust in empirical evidence.If industry characteristics have an effect on variables which were add to a regression model, and these effects are not perfectly controlled, we could not establish causal relation between variables.This phenomenon called as endogenous variable problemoften occur not only in economics and finance research but in case where we conduct study for non-experimental data. In a more severe case, it has been suggestedthatthis phenomenon produces may inconsequence result. These results were pervasively found in empirical accounting research and, in particular value relevance research.Easton and Sommers (2003) estimates the regression model of market capitalization on the book value of common of equity and net income. This regression is conducted for each year from 1963 to 1999.According to their results, the coefficients of book value are negative in 11 of the 36 years.Iatridis and Dimitras (2013) estimates the regression model of stock price on book value of equity per share and the net profit per share. The empirical analysis examines companies of five European nations (i.e., Portugal, Ireland, Italia, Greece, and Spain) for each year from 2005 to 2011.According to their results, the coefficients of book value are negative for Italian companies in 3 years period from 2005 to 2007.And Usui (2005) estimates the regression model of deflated market capitalization on the deflated book value of common of equity and deflated net income. The empirical analysis examines Japanese companies for each year from 1965 to 2012.According to their results, the coefficients of book value are negative in 10 years within 48 years period.

For such a problem, previous research suggests empirical model which controls industry effects on financial numbers.Obinata(2003) emphasizes need to control industrial effect, because intensity of intra industry competition adversely affects the estimated relevance of accounting information.The suggested specific solution is to introduce industry dummy variable in regression model.Industry model variable is defined to equal one if a company is involved particular industry and zero otherwise.In recent years, previous research introduces industry dummy variables in regression model as control for industry effect.

The results of this research show the following. Theoretical evidence gives economic meaning to the regression coefficients of industry dummy variables and suggests effectiveness of the industrial dummy

variable as the control variable.It suggests that intercept coefficient estimates industry effect on financial variable for base group, and that each slope coefficient estimates additional industry effect on financial variable for each group.Furthermore, to control industry effect, it is only necessary to insert industry dummy variables to the regression model.
In other words, we do not need to identify and measure industry characteristic variables.

II. THEORITICAL ANALYSIS

This section considers what the industrial dummy variables control in a regression model, in other word, economic meaning of the regression coefficients.For the sake of brevity, We suppose that a particular financial variable, X, is related to two industry characteristic variables, F_1,F_2, which thus evoke industry effect, and the error, u, as

$$X = \alpha_1 F_1 + \alpha_2 F_2 + u,$$

where α_1 and α_2 are the population slope parameters, respectively.The population error uis independent of explanatory variable F_1, F_2 and is normally distributed with zero mean and variance σ^2:

$$u \sim N(0, \sigma^2)$$

We also suppose that the population consists companies in three industries (i.e., industry 1, industry 2, industry 3), as a result F_1 and F_2 take only three values $F_{1,1}, F_{1,2}, F_{1,3}$ and $F_{2,1}, F_{2,2}, F_{2,3}$, respectively.
Next, we choose industry 1 to be the base group, that is, the group against which comparisons are made.Using $F_{1,1}$ and $F_{2,1}$, which is values of independet variables on base group, above equation become

$$\begin{aligned} X &= \alpha_1 F_1 + (\alpha_1 F_{1,1} - \alpha_1 F_{1,1}) \\ &\quad + \alpha_2 F_2 + (\alpha_2 F_{2,1} - \alpha_2 F_{2,1}) + u \\ &= \alpha_1 F_{1,1} + \alpha_1 (F_1 - F_{1,1}) \\ &\quad + \alpha_2 F_{2,1} + \alpha_2 (F_2 - F_{2,1}) + u \\ &= \alpha_1 F_{1,1} + \alpha_2 F_{2,1} + \alpha_1 \Delta F_1 + \alpha_2 \Delta F_2 + u, \end{aligned}$$

Where $\Delta F_{1,2}$, $\Delta F_{2,2}$, $\Delta F_{1,3}$,and $\Delta F_{2,3}$ are constants which represent the deviation from value of F_1 and$F_2$for indutry 1 (i.e., $F_{1,1}$, $F_{2,1}$), respectively.

$$\begin{aligned} \text{Industry 1: } X &= \alpha_1 F_{1,1} + \alpha_2 F_{2,1} \\ &\quad + \alpha_1 \Delta F_{1,1} + \alpha_2 \Delta F_{2,1} + u \\ &= \alpha_1 F_{1,1} + \alpha_2 F_{2,1} + \varepsilon \end{aligned}$$

$$\begin{aligned} \text{Industry 2: } X &= \alpha_1 F_{1,1} + \alpha_2 F_{2,1} \\ &\quad + \alpha_1 \Delta F_{1,2} + \alpha_2 \Delta F_{2,2} + u \end{aligned}$$

$$\begin{aligned} \text{Industry 3: } X &= \alpha_1 F_{1,1} + \alpha_2 F_{2,1} \\ &\quad + \alpha_1 \Delta F_{1,3} + \alpha_2 \Delta F_{2,3} + u, \end{aligned}$$

Where $\Delta F_{1,2}$, $\Delta F_{2,2}$, $\Delta F_{1,3}$,and $\Delta F_{2,3}$ are constants which represent the deviation from value of F_1 and$F_2$for industry 1 (i.e., $F_{1,1}$, $F_{2,1}$), respectively.

Using dummy varibles, ditribution of financial number, X, for each industry can be expressed as single regression models:

$$\begin{aligned} X &= \alpha_1 F_{1,1} + \alpha_2 F_{2,1} + \alpha_1 \Delta F_{1,2} * D_1 \\ &\quad + \alpha_2 \Delta F_{2,2} * D_1 + \alpha_1 \Delta F_{1,3} * D_2 \\ &\quad + \alpha_2 \Delta F_{2,3} * D_2 + u \\ &= \alpha_1 F_{1,1} + \alpha_2 F_{2,1} \\ &\quad + (\alpha_1 \Delta F_{1,2} + \alpha_2 \Delta F_{2,2}) * D_1 \\ &\quad + (\alpha_1 \Delta F_{1,3} + \alpha_2 \Delta F_{2,3}) * D_2 + u, \end{aligned}$$

Where D_1 is defined to equal one if a company is involved in industry 2 and zero otherwise. $D_2$is defined to equal one if a company is involved in industry 3 and zero otherwise.
While we have restricted the number of industry and industry characteristic variable for purpose of simplication, we can expand to L industries and M industry characteristic variables.In this situation, distribution of financial number, X, can be expressed as following regression models:

$$\begin{aligned} X &= \alpha_1 F_{1,1} + \alpha_2 F_{2,1} + \cdots + \alpha_M F_{M,1} \\ &\quad + (\alpha_1 \Delta F_{1,2} + \alpha_2 \Delta F_{2,2} + \cdots + \alpha_M \Delta F_{M,2}) * D_1 \\ &\quad + (\alpha_1 \Delta F_{1,3} + \alpha_2 \Delta F_{2,3} + \cdots + \alpha_M \Delta F_{M,3}) * D_2 \\ &\quad + \cdots \\ &\quad + (\alpha_1 \Delta F_{1,L} + \alpha_2 \Delta F_{2,L} + \cdots + \alpha_M \Delta F_{M,L}) * D_{L-1} \\ &\quad + \varepsilon \end{aligned}$$

$$\begin{aligned} X &= \sum_{m=1}^{M} \alpha_m F_{m,1} + \left(\sum_{m=1}^{M} \alpha_m \Delta F_{m,2} \right) \times D_1 \\ &\quad + \left(\sum_{m=1}^{M} \alpha_m \Delta F_{m,3} \right) \times D_2 + \cdots \\ &\quad + \left(\sum_{m=1}^{M} \alpha_m \Delta F_{m,L} \right) \times D_{L-1} \\ &\quad + u. \end{aligned}$$

If we estimate the regression model which regress the financial variable X on industry dummy variables as following:

$$X = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \cdots + \beta_{L-1} D_{L-1} + u,$$

then expected values of each regression coefficient go as below

$$E[\hat{\beta}_0] = \sum_{m=1}^{M} \alpha_m F_{m,1},$$

$$E[\hat{\beta}_1] = \sum_{m=1}^{M} \alpha_m \Delta F_{m,2},$$

$$\cdots,$$

$$E[\hat{\beta}_{L-1}] = \sum_{m=1}^{M} \alpha_m \Delta F_{m,L}$$

where $\beta_0, \beta_1, \cdots, \beta_{L-1}$ are regression coefficient, or parameter and u is the error term.

This theoretical evidence gives economic meaning to the regression coefficients of industry dummy variables and suggests effectiveness of the industrial dummy variable as the control variable.It suggests that intercept coefficient estimates industry effect on financial variable for base group, and that each slope coefficient estimates additional industry effect on financial variable for each group.Furthermore, to control industry effect, it is only necessary to insert industry dummy variables to the regression model.

In other words, we do not need to identify and measure industry characteristic variables.

III. CONCLUSION

In recent year, significant advances have been made in the factor of corporate profitability and these impacts.This research issues have received scholarly attention of management, accounting, and industrial organization.Many research point out that there are the homogeneity of corporate fundamentals in each industry(Amit and Livant, 1990; Guenther and Rosman, 1994; Krishnan and Press, 2003; Kimura,2009; Aratani,2010).

These relationship between financial variable and industry characteristics might request a review of prior statistical method and erode trust in empirical evidence.If industry characteristics have an effect on variables which were add to a regression model, and these effects are not perfectly controlled, we could not establish causal relation between variables.These results were pervasively found in empirical accounting research and, in particular value relevance research.For such a problem, previous research suggests empirical model which controls industry effects on financial numbers.Obinata(2003) emphasizes need to control industrial effect, because intensity of intra industry competition adversely affects the estimated relevance of accounting information.The suggested specific solution is to introduce industry dummy variable in regression model.

The results of this research show the following.Theoretical evidence gives economic meaning to the regression coefficients of industry dummy variables and suggests effectiveness of the industrial dummy variable as the control variable.It suggests that intercept coefficient estimates industry effect on financial variable for base group, and that each slope coefficient estimates additional industry effect on financial variable for each group.Furthermore, to control industry effect, it is only necessary to insert industry dummy variables to the regression model.

In other words, we do not need to identify and measure industry characteristic variables.

REFERENCES

[1] Amit, P., and Livant, J. (1990). Grouping of conglomerates by their segement's economic attributes: towards a more meaningful ratio analysis. Journal of Business Finance & Accounting, 17(1), 85-99.

[2] Ball, R., and Brown, P. R. (2014). Ball and Brown (1968): A retrospective. The Accounting Review, 89(1), 1–26.

[3] Barth, M. E., Beaver, M. H., and Landsman, W. R. (2001). The relevance of the value relevance literature for financial accounting standard setting: Another view. Journal of Accounting and Economics, 31, 77–104

[4] Barth, M. E., and Clinch, G. (2009). Scale effects in capital markets-based accounting research. Journal of Business Finance & Accounting, 36(3–4), 253–288.

[5] Barth, R. E., and Kallapur, S. (1996). The effects of cross-sectional scale differences on regression results in empirical accounting research. Contemporary Accounting Research, 13, 527–567.

[6] Brown, P. (1994). Capital Markets-based Research in Accounting: An Introduction. Melbourne: Coopers & Lybrand.

[7] Brown, S., Lo, K., and Lys, T. (1999). Use of R2 in accounting research: Measuring changes in value relevance over the last four decades. Journal of Accounting and Economics, 28, 83–115.

[8] Collins, D. W., Maydew, E. L., and Weiss, I, S. (1997). Changes in the value-relevance of earnings and book values over the past forty years. Journal of Accounting and Economics, 24, 39–67.

[9] Easton, P.D., and Harris, T. S. (1991). Earnings as an explanatory variable for returns. Journal of Accounting Research, 29, 19–36.

[10] Easton, P. D., and Sommers, G. A. (2003). Scale and the scale effect in market-based accounting research. Journal of Business Finance & Accounting, 30, 25–55.

[11] Feltham, G. A., and Ohlson, J. A. (1995). Valuation and Clean Surplus Accounting for Operating and Financial Activities. Contemporary Accounting Research, 11(2), 689-731.

[12] Freeman, R. N., and Tse, S. Y. (1992). A nonlinear model of security price responses to unexpected earnings. Journal of Accounting Research, 30, 185–209.

[13] Guenther, D. A., and Rosman, A. J. (1994). Differences between COMPUSTAT and CRSP SIC Codes and related effects on research. Journal of Accounting and Economics, 18(1), 115-128.

[14] Hao, S., Jin, Q., and Zhang, G. (2011). Investment Growth and the Relation between Equity Value, Earnings, and Equity Book. The Accounting Review, 86(2), 605-635.
[15] Holthausen, R. W., and Watts, R. L. (2001). The relevance of the value relevance literature for financial accounting standard setting. Journal of Accounting and Economics, 31, 3–75.
[16] Hyan, C. (1995). The information content of losses. Journal of Accounting and Economics, 20, 125–153.
[17] Kimura, F. (2009). A comparison of Reliability of IndustrialClassifications in Japan.Contemporary Disclosure Research, 9,33-42
[18] King. G., Keohane, R.O., and Verba. S. (1994). Designing Social Inquiry: Scientific Inference in Qualitative Research, Princeton Univ.
[19] Kothari, S. P. (2001). Capital market research in accounting. Journal of Accounting and Economics, 31, 105–231.
[20] Kothari, S. P., and Zimmerman, J. L. (1995) Price and return models. Journal of Accounting and Economics, 20, 155-192.
[21] Kothari, S. P., Ramanna, K., and Skinner. D. J. (2010). Implication for GAAP from an analysis of positive research in accounting. Journal of Accounting and Economics, 50, 246–286.
[22] Krishnan, J., and Press. E. (2003). The North American Industry Classification System and its implications for accounting research. Contemporary Accounting Research, 20(4), 685-717.Lev, B. (1989). On the usefulness of earnings and earnings research: Lessons and directions from two decades of empirical research. Journal of Accounting Research, 27, 153–192.
[23] Lev, B., Li, S., and Sougiannis, T. (2010). The usefulness of accounting estimates for predicting cash flows and earnings. Review of Accounting Studies, 15, 779–807.
[24] Lev, B., and Zarowin, P. (1999). The boundaries of financial reporting and how to extend them. Journal of Accounting Research, 37, 353–385.
[25] Ohlson, J. A. (1995). Earnings, Book values, and Dividends in Equity Valuation. Contemporary Accounting Research, 11(2), 661-687.
[26] Stock, J. H., and Watson, M. W. (2011). Introduction to Econometrics (third edition). Pearson Education Limited.
[27] Zhang, X. J. (2000). Conservative accounting and equity valuation. Journal of Accounting and Economics, 29, 125-149.

★★★