

Time-Specific Effects and a Bias in the Granger-Causality Test

Chung-ki Min^{*}

This study points out that not only contemporaneous but also lagged time-specific effects can bias the causality test. Treating the lagged time-specific effects as measurement errors in explanatory variables, this study suggests use of the instrumental variables from the literature on measurement errors. Using simulated data, this study shows that the Granger-causality test based on the OLS estimates is biased. This bias can be corrected if appropriately-chosen instrumental variables are used.

Field of Research: Econometrics

1. Introduction

The Granger-causality test is designed to test if lagged values of a variable have significant effects on the current value of another variable. All variables are subject to time-specific changes which are independently and unexpectedly determined in each period. As these time-specific effects are not associated with the causality, we have to separate them from the causality test. One widely-employed approach to control for the time-specific effects is to use dummy variables. However, this dummy-variable approach can be applied only when there are multiple observations in each period, i.e., panel data.

This study examines how time-specific effects can distort the Granger-causality test. In particular, it points out that not only contemporaneous but also lagged time-specific effects can bias the causality test. Treating the lagged time-specific effects as measurement errors in explanatory variables, this study suggests use of the instrumental variables from the literature on measurement errors. The instrumental variables include appropriately lagged variables (Griliches & Hausman 1984) and higher-moment variables (Dagenais & Danenais 1997).

Using simulated data this study shows that the Granger-causality test based on the OLS estimates is biased. This bias, however, is corrected when appropriately-chosen instrumental variables are used.

The discussion in this study can be applied to the tourism research. Tourism is highly sensitive to economic fluctuations, in particular during global economic crisis. In the literature of tourism research there are many studies which examine the causality

^{*} Department of Economics, Hankuk University of Foreign Studies, Seoul, Korea. Email: cmin@hufs.ac.kr. This work was supported by the Hankuk University of Foreign Studies Research Fund.

Min

between tourism development and economic growth. One group of the research uses single time series and examines whether variables are stationary and any cointegrating relation exists, but they do not account for possible bias caused by time-specific effects (Balaguer & Cantavella-Jorda 2002; Lee & Kwon 1995; Oh 2005; Shan & Wilson 2001). Another group of research uses panel data models in estimating tourism demand and the contribution of tourism to economic growth. While they control for unit-specific effects in dynamic panel data models, they do not address the issue of time-specific effects (Fayissa, Nsiah & Tadasse 2008; Garin-Munoz 2006; Ledesma-Rodriguez, Navarro-Ibanez & Perez-Rodriguez 2001). The estimation methods suggested in this study are therefore expected to produce unbiased test results about the Granger causality in the tourism research.

The following section illustrates how time-specific effects can bias the Granger-causality test. Section 3 discusses instrumental variables and estimation methods which can control for time-specific effects and section 4 presents empirical results using simulated data. Section 5 concludes this paper.

2. The Bias in the Granger–causality Test

The Granger-causality is defined as a relationship between a variable and the lagged values of other variables. Thus, the causality test uses VAR (vector autoregressive) models which are reduced-form equations of structural equations. For a case of two variables x^* and y^* , the VAR model is expressed as

$$\begin{aligned}x_t^* &= \alpha_0 + \sum_{k=1}^K \alpha_k^x x_{t-k}^* + \sum_{k=1}^K \alpha_k^y y_{t-k}^* + u_t^x \\y_t^* &= \beta_0 + \sum_{k=1}^K \beta_k^x x_{t-k}^* + \sum_{k=1}^K \beta_k^y y_{t-k}^* + u_t^y\end{aligned}\tag{1}$$

A null hypothesis of $\alpha_k^y = 0$ (for all k) implies that y^* does not Granger-cause x^* , and a null hypothesis of $\beta_k^x = 0$ (for all k) implies that x^* does not Granger-cause y^* . While the OLS method is widely used in estimating the VAR models, it is well known that the OLS estimators become biased if any of the lagged variables on the right-hand side are correlated with the disturbances, called the endogeneity problem.

Every time period the economy experiences unexpected changes and almost all economic variables are influenced by the changes. The effects on economic variables of the changes in each period, called time-specific effects, are merged into the variables x^* and y^* in Eq.(1). A problem arises in estimating Eq.(1) since neither these time-specific effects nor the underlying variables x^* and y^* are observed. Letting δ_t^x and δ_t^y denote the time-specific effects for x and y , respectively, we can observe only the sum of the latent variables x^* and y^* and their corresponding time-specific effects.

$$\begin{aligned}x_t &= x_t^* + \delta_t^x \\y_t &= y_t^* + \delta_t^y\end{aligned}\tag{2}$$

Min

As changes in each period are unexpected, their time-specific effects are independent between periods. Therefore, they are not associated with the causal relationship. As expressed in Eq.(1), the causal relationship holds only for the latent variables. Since x^* and y^* are not observable, however, we need to transform Eq.(1) into equations expressed in observable variables. Substituting Eq.(2) for x^* and y^* in Eq.(1), we obtain

$$\begin{aligned} x_t &= \sum_{k=1}^K \alpha_k^x x_{t-k} + \sum_{k=1}^K \alpha_k^y y_{t-k} + \theta_t^x + u_t^x \\ y_t &= \sum_{k=1}^K \beta_k^x x_{t-k} + \sum_{k=1}^K \beta_k^y y_{t-k} + \theta_t^y + u_t^y \end{aligned} \quad (3)$$

where $\theta_t^x = \delta_t^x - \sum_{k=1}^K \alpha_k^x \delta_{t-k}^x - \sum_{k=1}^K \alpha_k^y \delta_{t-k}^y + \alpha_0$ and $\theta_t^y = \delta_t^y - \sum_{k=1}^K \beta_k^x \delta_{t-k}^x - \sum_{k=1}^K \beta_k^y \delta_{t-k}^y + \beta_0$; these are functions of current and lagged time-specific effects. Notice that the time-specific effects are included not only in θ_t^x and θ_t^y but also in $x_{i,t-k}$'s and $y_{i,t-k}$'s. Thus, ignoring θ_t^x and θ_t^y biases the OLS estimation due to the endogeneity. This resulting bias belongs to the problem of measurement errors in explanatory variables. It indicates that not only contemporaneous but also lagged time-specific effects need to be controlled for in the Granger-causality test.

To illustrate the bias in the Granger-causality test, we generate data using the following two-variable VAR model in which x^* Granger-causes y^* but y^* does not for x^* .

$$\begin{aligned} x_t^* &= 0.5x_{t-1}^* + u_t^x \\ y_t^* &= 0.5x_{t-1}^* + 0.5y_{t-1}^* + u_t^y \end{aligned} \quad (4)$$

where u_t^x and u_t^y are drawn independently from a normal distribution with mean 0 and variance 0.1^2 (for $t=1, \dots, 20$). Figure 1 clearly shows the causal relationship from x^* to y^* .

Min

Figure 1: Causal relationship when there are no time-specific effects

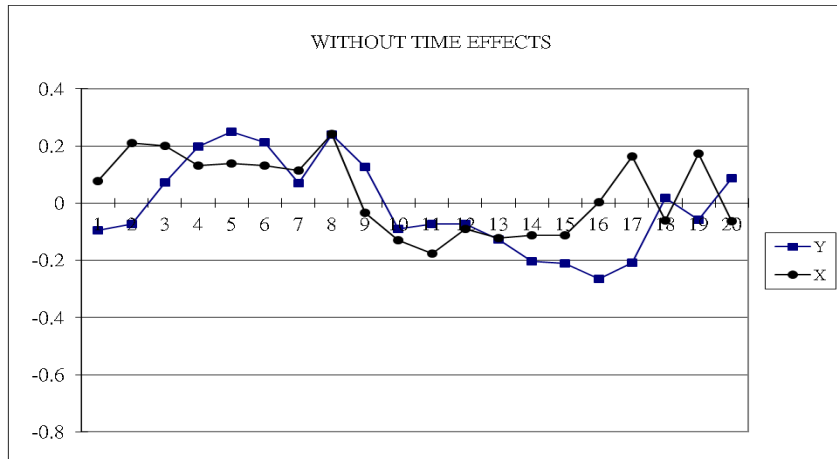


Figure 2: When time-specific effects are added every period

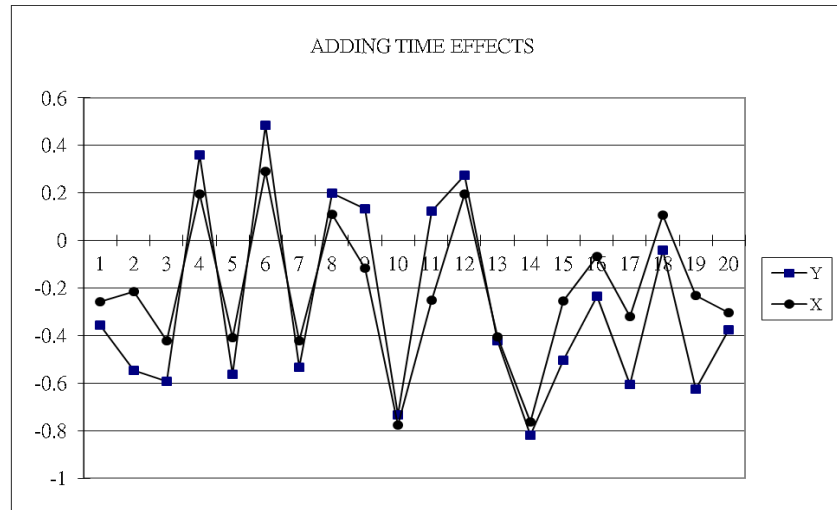
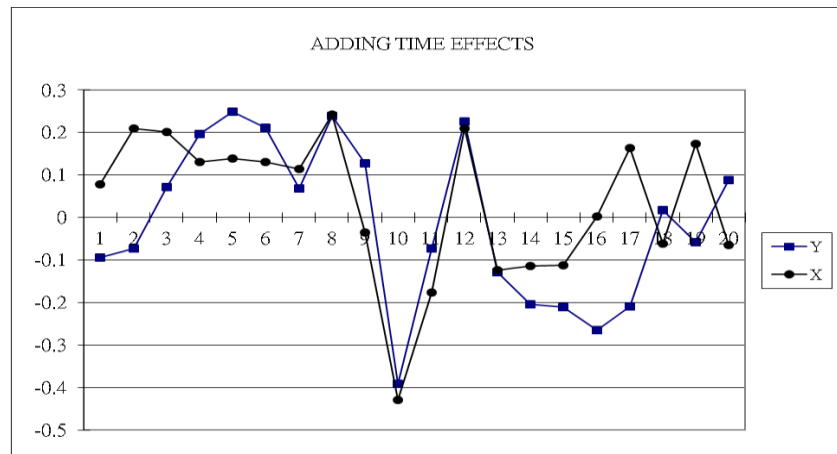


Figure 3: When there are two big time-specific effects in $t=10$ and 12



Min

However, the fundamental variables themselves are not observable but can be observed only with time-specific effects added, i.e., $x_t = x_t^* + \delta_t^x$ and $y_t = y_t^* + \delta_t^y$. As economic variables are affected by common unexpected changes in each period, the time-specific effects δ_t^x and δ_t^y are often highly correlated. In Figure 2, δ_t^x and δ_t^y are drawn from a uniform distribution with a range between -0.7 and 0.3 in a way that they are correlated within each period (the correlation coefficient being 0.92) but independent between periods. In Figure 2, it is hard to see the causal relation between x and y since the contemporaneous time-specific effects have dominated the lagged causal relation.

In Figure 3, we consider a small number of big shocks such as the recent global financial crisis. In generating data we set δ_t^x and δ_t^y to zero (no time-specific effects) for all periods except two: $\delta_{10}^x = \delta_{10}^y = -0.3$ for $t=10$ and $\delta_{12}^x = \delta_{12}^y = 0.3$ for $t=12$. As shown in Figure 3, although the causal relation is present for most of the periods, the Granger-causality test fails to conclude for the existence of causality because of the two big time-specific effects.

3. Methodology: Instrumental Variable Estimation

The above bias in the Granger-causality test is caused by the endogeneity of the explanatory variables which contain lagged time-specific effects. Thus, we need to address the endogeneity problem to correct the bias. Widely-employed estimation methods for the endogeneity problem are the instrumental variable estimation methods, such as the GMM (generalized method of moments) and 2SLS (two stage least squares). For these estimation methods to perform well, the instrumental variables should satisfy two conditions that they be relevant to the endogenous explanatory variables and be exogenous to cut the endogeneity. This study uses two sets of instrumental variables suggested by Griliches and Hausman (1986) and Dagenais and Dagenais (1997). These instrumental variables are expected to satisfy the two conditions.

Griliches and Hausman (1986) point out that the within transformation which controls for the individual effects in panel data often yields unsatisfactory results. It is because measurement errors in explanatory variables exert magnified influence in the within-transformed model. Assuming that the measurement errors are serially uncorrelated, they suggest use of appropriately lagged values of the relevant variables as instruments. This approach does not require any extraneous information.

Dagenais and Dagenais (1997) also propose instrumental variable estimators for linear regression models with errors in explanatory variables which require no extraneous information. These estimators use instrumental variables obtained from sample moments of order higher than two. In contrast to the lagged variables suggested by Griliches and Hausman (1986), these use contemporaneous variables. Consider the following regression model for time-series data.

$$Y = \alpha + X\beta + u_t \quad (5)$$

Min

where X is a $T \times K$ matrix of explanatory variables measured without error. Since \tilde{X} is not observable, we use observable variables X which include measurement errors V .

$$Y = \alpha 1_T + X\beta + (u - V\beta) = \alpha 1_T + X\beta + \eta \quad (6)$$

where $X = \tilde{X} + V$. The higher moment estimator for $\theta = (\alpha, \beta)'$ is derived using the following instrumental variables.

$$\begin{aligned} Z &= (1_T, z_1, \dots, z_7), \\ z_1 &= x * x, \quad z_2 = x * y, \quad z_3 = y * y, \\ z_4 &= x * x * x - 3x[E(x'x/T) * I_K], \\ z_5 &= x * x * y - 2x[E(x'y/T) * I_K] - y\{1_K[E(x'x/T) * I_K]\}, \\ z_6 &= x * y * y - x[E(y'y/T) * I_K] - 2y[E(y'x/T)], \\ z_7 &= y * y * y - 3y[E(y'y/T)] \end{aligned} \quad (7)$$

where the symbol $*$ designates the Hadamard element-by-element matrix multiplication operator and $(x \ y)$ corresponds to $(X \ Y)$ with the variables expressed in mean deviation form. As Z is orthogonal to η as proven in Dagenais and Dagenais (1997), we use Z as instrument variables in the next section.

4. Numerical Illustrations

For an empirical application of the above estimation methods, we generate time-series observations on (x_t, y_t) as follows. For $t=1, 2, \dots, 4000$,

$$\begin{aligned} x_t^* &= 0.8x_{t-1}^* + u_t^x, & u_t^x &\sim \text{uniform}(-0.25, 0.25) \\ y_t^* &= 0.5x_{t-1}^* + 0.8y_{t-1}^* + u_t^y, & u_t^y &\sim \text{uniform}(-0.25, 0.25) \end{aligned}$$

$$\begin{aligned} x_t &= x_t^* + v_t^x, & v_t^x &= \delta_t + \varepsilon_t^x \\ y_t &= y_t^* + v_t^y, & v_t^y &= \delta_t + \varepsilon_t^y \\ \delta_t &\sim \text{Normal}(0, 0.5^2), & \varepsilon_t^x &\sim \text{Normal}(0, 0.5^2), & \varepsilon_t^y &\sim \text{Normal}(0, 0.5^2) \end{aligned}$$

where v_t^x and v_t^y are time-specific effects. According to this model, x Granger-causes y but y does not Granger-cause x . Using observable variables x_t and y_t , we estimate a VAR(2) model:

$$\begin{aligned} x_t &= \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + u_t^x \\ y_t &= \beta_0 + \beta_1 x_{t-1} + \beta_2 y_{t-1} + u_t^y \end{aligned} \quad (8)$$

The estimation results are summarized in Table 1. As expected, the OLS estimation yields biased estimates because it does not control for the time-specific effects. Thus,

Min

the Granger-causality test is also biased; it concluded that y Granger-causes x and x does not Granger-cause y .

In contrast, the GMM estimation controls for the time-specific effects as it uses instrumental variables which are orthogonal to the time-specific effects, as discussed in the previous section. GMM1 used a subset of instrumental variables in Eq.(7) and lagged values, i.e., $\{z_1, z_4, x_{t-2}, y_{t-2}\}$, while GMM2 used all instrumental variables $\{z_1, z_2, z_3, z_4, z_5, z_6, z_7, x_{t-2}, y_{t-2}\}$. GMM1 produced accurate estimates and correctly concluded about the Granger causality. However, GMM2 produced incorrect results about the causality. Further investigation about the instrumental variables of higher moments is needed and planned for future work.

Table 1: Estimation results for simulated data

Explanatory variables	Coefficient values	Estimates (standard errors)		
		OLS ^a	GMM1 ^b	GMM2 ^c
Dep. var= x_t				
$x_{t-1} (\alpha_1)$	0.8	0.08 ^{***} (0.02)	0.60 ^{**} (0.25)	0.34 [*] (0.18)
$y_{t-1} (\alpha_2)$	0	0.04 ^{**} (0.02)	0.09 (0.07)	0.14 ^{**} (0.06)
Dep. var= y_t				
$x_{t-1} (\beta_1)$	0.5	-0.02 (0.02)	0.54 [*] (0.32)	0.35 (0.26)
$y_{t-1} (\beta_2)$	0.8	0.31 ^{***} (0.02)	0.81 ^{***} (0.10)	0.84 ^{***} (0.09)

^a This OLS estimation did not control for the time-specific effects.

^b GMM1 used instrumental variables $\{z_1, z_4, x_{t-2}, y_{t-2}\}$.

^c GMM2 used instrumental variables $\{z_1, z_2, z_3, z_4, z_5, z_6, z_7, x_{t-2}, y_{t-2}\}$.

5. Summary and Conclusions

This study points out a possibility of bias due to time-specific effects in the Granger-causality test. Treating the lagged time-specific effects as measurement errors in explanatory variables, this study suggests use of the instrumental variables from the literature on measurement errors. The instrumental variables include appropriately lagged variables and higher-moment variables.

The discussion in this study can be applied to the tourism research as tourism is highly sensitive to economic fluctuations, in particular during global economic crisis. An application of the estimation methods suggested in this study is planned for future work.

References

- Balaguer, J and Cantavella-Jorda, M 2002, 'Tourism as a long-run economic growth factor: The Spanish case', *Applied Economics*, vol. 34, pp. 877-844.
- Dagenais, M and Dagenais, DL 1997, 'Higher moment estimators for linear regression models with errors in the variables', *Journal of Econometrics*, vol. 76, pp. 193-221.
- Fayissa, B, Nsiah, C and Tadasse, B 2008, 'Impact of tourism on economic growth and development in Africa', *Tourism Economics*, vol. 14, pp. 807-818.
- Garin-Munoz, T 2006, 'Inbound international tourism to Canary Islands: A dynamic panel data model', *Tourism Management*, vol. 27, pp. 281-291.
- Griliches, Z and Hausman, JA 1986, 'Errors in variables in panel data', *Journal of Econometrics*, vol. 31, pp. 93-118.
- Ledesma-Rodriguez, F, Navarro-Ibanez, M and Perez-Rodriguez, JV 2001, 'Panel data and tourism: a case study of Tenerife', *Tourism Economics*, vol.7, pp. 75-88.
- Lee, C and Kwon, K 1995, 'Importance of secondary impact of foreign tourism receipts on the South Korean economy', *Journal of Travel Research*, vol. 34, pp. 50-54.
- Oh, CO 2005, 'The contribution of tourism development to economic growth in the Korean economy', *Tourism Management*, vol. 26, pp. 39-44.
- Shan, J., and K. Wilson, 2001, "Causality between trade and tourism: Empirical evidence from China," *Applied Economics Letter*, 8, 279-283.