Spurious Regression in Time-Wise Autocorrelated and Cross-Sectionally Heteroskedastic Procedures

M. Fukushige[a]

[a] *Graduate School of Economics, Osaka University, 7-1, Machikaneyama-cho, Toyonaka-shi, Osaka 560-0043 Japan (mfuku@econ.osaka-u.ac.jp)*

Keywords: Spurious regression, Unit root, Time-Wise Autocorrelated and Cross-Sectionally Heteroskedastic Procedures

EXTENDED ABSTRACT

Spurious regression is a serious problem in empirical research when time series has unit roots. This problem might exist when we apply a regression procedure for pooling time series and cross-section data. There are several models to pool time series and cross-section data in regression context, *e.g.* fixed effect or random effect model. Recently, Entorf [1997] and Kao [1999] also pointed out that the spurious regression is still unsolved problem when we apply the least squares with dummy variable (LSDV) procedure for pooling time series and cross-section data.

We have another procedure to pool time series and cross-section data especially for the case when their error terms are autocorrelated and cross-sectionally heteroskedastic. It was proposed by Kmenta [1986]. This procedure estimates autocorrelation coefficients for each time series and admit a heteroskedastic error terms for each series. In the present paper, I investigate whether the spurious regression is a serious for this procedure or not.

I conduct a Monte Carlo simulation to investigate the spurious regression in the time-wise autocorrelated and heteroskedastic procedure proposed by Kmenta [1986].

We set the data generating process (DGP): $y_{is} = y_{i,s-1} + v_{is}$ and $x_{is} = x_{i,s-1} + w_{is}$ for $i = 1,\dots,N$ and $s = 1,\dots,T$, where the error terms (v_{is}, w_{is}) were generated from a bivariate normal distribution with independence across both individual and time period:

$$\begin{bmatrix} v_{is} \\ w_{is} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right).$$

The results show that the spurious regression problem is serious in this procedure too. In comparison to the LSDV procedure, the results are summarized as follows,

1st: Kmenta's procedure has larger probabilities when the estimated coefficient is statistically significant than the LSDV procedure.

2nd: While the spurious regression problem in the LSDV procedure becomes more serious in accordance with the increase of the time length (T), this problem in Kmenta's procedure becomes less serious except the case when the sample size is very small ($N \times T < 900$).

This result suggests that we should pay attention to the spurious regression problem even when we apply the Kmenta's procedure for pooling time series and cross-section data.

1. INTRODUCTION

Spurious regression is a serious problem in empirical research when time series has unit roots. Granger and Newbold [1974] consider the estimation of a simple regression using Monte Carlo simulation and Phillips [1986] investigated the asymptotic distributions of the Durbin=Watson ratio and t-value.

This problem might exist when we apply a regression procedure for pooling time series and cross-section data. There are several models to pool time series and cross-section data in regression context, *e.g.* fixed effect or random effect model (See also Baltagi, 1998). Recently, Entorf [1997] and Kao [1999] also pointed out that the spurious regression is still unsolved problem when we apply the least squares with dummy variable (LSDV) procedure for pooling time series and cross-section data.

We have another procedure to pool time series and cross-section data especially for the case when their error terms are autocorrelated and cross-sectionally heteroskedastic. It was proposed by Kmenta [1986] (See also Baltagi, 1998). This procedure estimates autocorrelation coefficients for each time series and admit a heteroskedastic error terms for each series. In the present paper, I investigate whether the spurious regression is a serious for this procedure or not.

The paper is consisted as follows. In section 2, the phenomenon of the spurious regression is summarized. Time-wise autocorrelated and heteroskedastic procedure proposed by Kmenta [1986] is introduced in section 3. In section 4, the results of the Monte Carlo simulation are reported. Finally, in section 5, I provide further discussion.

2. SPURIOUS EFFECTS IN NONSENSE REGRESSIONS

Granger and Newbold [1974] consider the estimation of a simple regression:

$$y_s = \alpha + \beta x_s + \varepsilon_s \qquad (1)$$

where s=1,…, T, and y_s and x_s are assumed to be generated as two independent random walks:

$$y_{is} = y_{i,s-1} + v_{is} \text{ and } x_{is} = x_{i,s-1} + w_{is}.$$

They found that the calculated t-value for β in (1) is significant with relatively high frequencies, and that the Durbin=Watson (DW) ratio is low. This regression is called a "spurious regression."

Phillips [1986] investigated the asymptotic distribution of the DW. ratio and t-value and obtained the following results:

$$\text{DW.} = O_p(T^{-1}),$$

$$t = O_p(T).$$

These results mean that the calculated t-values diverge, while the DW ratio converges to zero.

Some of the asymptotic results of these nonsense regressions are intuitive and others are obtained theoretically. However, their small sample properties are obscure in general cases. In the present paper, we investigate the spurious regression in estimation of autocorrelated and cross-sectionally heteroskedastic procedure proposed by Kmenta [1986] for pooling time series and cross-section data.

3. TIME-WISE AUTOCORRELATED AND CROSS-SECTIONALLY HETEROSKEDASTIC PROCEDURE

Kmenta [1986] assumes the regression model:

$$y_{is} = x_{is}'\beta + \varepsilon_{is} \quad i = 1,\dots, N, \quad s = 1,\dots, T.$$

$$\begin{aligned}
E(\varepsilon_{is}^2) &= \sigma_i^2 \\
E(\varepsilon_{is}\varepsilon_{js}) &= 0 \qquad i \neq j \\
\varepsilon_{is} &= \rho_i \varepsilon_{i,s-1} + u_{is} \\
\text{where} \quad u_{is} &\sim N(0, \sigma_{ui}^2) \\
\varepsilon_{is} &\sim N\left(0, \frac{\sigma_{ui}^2}{1-\rho_i^2}\right) \\
\text{and} \quad E(\varepsilon_{i,s-1} u_{js}) &= 0 \quad \text{for all} \quad i, j.
\end{aligned}$$

When we set vector y and matrix X as

$$y = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1T} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{NT} \end{bmatrix},$$

$$X = \begin{bmatrix} x_{11,1} & x_{11,2} & \Lambda & x_{11,K} \\ x_{12,1} & x_{12,2} & & x_{12,K} \\ M & M & & M \\ x_{1T,1} & x_{1T,2} & \Lambda & x_{1T,K} \\ x_{21,1} & x_{21,2} & \Lambda & x_{1T,K} \\ x_{22,1} & x_{22,2} & \Lambda & x_{1T,K} \\ M & M & & M \\ x_{NT,1} & x_{NT,2} & \Lambda & x_{NT,K} \end{bmatrix},$$

and the error tem and coefficients as

$$\varepsilon = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ M \\ \varepsilon_{1T} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ M \\ \varepsilon_{NT} \end{bmatrix} \text{ and } \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ M \\ \beta_K \end{bmatrix},$$

we can rewrite the model

$$y = X\beta + \varepsilon \text{ and } E(\varepsilon\varepsilon') = \Omega:$$

where

$$\Omega = \begin{bmatrix} \sigma_1^2 V_1 & 0 & \Lambda & 0 \\ 0 & \sigma_2^2 V_2 & \Lambda & 0 \\ M & M & & M \\ 0 & 0 & \Lambda & \sigma_N^2 V_N \end{bmatrix}$$

and

$$V_i = \begin{bmatrix} 1 & \rho_i & \rho_i^2 & \Lambda & \rho_i^{T-1} \\ \rho_i & 1 & \rho_i & \Lambda & \rho_i^{T-2} \\ M & M & M & & M \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \Lambda & 1 \end{bmatrix}.$$

To estimate this model, we take the following steps. First we apply the ordinary least squares method to all $N \times T$ observations. We obtain unbiased and consistent coefficient, so we can use the residuals e_{is} for estimating parameters in the variance-covariance matrix (Ω).

In the present paper, we estimate ρ_i as follows,

$$\hat{\rho}_i = \frac{\sum e_{is} e_{i,s-1}}{\sqrt{\sum e_{is}^2}\sqrt{e_{i,s-1}^2}}$$

Next, we transform the observations in accordance with the estimated correlation coefficients,

$$y_{is}^* = \beta_1 x_{is,1}^* + \beta_2 x_{is,2}^* + \Lambda + \beta_K x_{is,K}^* + u_{is}^*$$

$$\begin{array}{ccclccl} y_{is}^* & = & \sqrt{1-\hat{\rho}_i^2}\, y_{is} & \text{for} & s & = & 1 \\ y_{is}^* & = & y_{it} - \hat{\rho}_i y_{i,s-1} & \text{for} & s & = & 2,3,\ldots,T \\ x_{is,k}^* & = & \sqrt{1-\hat{\rho}_i^2}\, x_{is,k} & \text{for} & s & = & \\ x_{is,k}^* & = & x_{is} - \hat{\rho}_i x_{i,s-1,k} & \text{for} & s & = & 2,3,\ldots,T \\ k & = & 1,2,\ldots,K & & & & \\ i & = & 1,2,\ldots,N & & & & \end{array}$$

Applying the ordinary least squares to the transformed data, we obtain a consistent estimator and the residuals. Using estimated residuals, we estimate the variance of u_{is}:

$$s_{us}^2 = \frac{1}{T-K}\sum_{s=1}^{T} u_{is}^{*2}\ .$$

Consistent estimator of σ_i^2 can be estimated as

$$s_i^2 = \frac{s_{us}^2}{1-\hat{\rho}_i^2}\,.$$

Using these estimates, we transform the observations:

$$y^{**}_{is} = \beta_1 x^{**}_{is,1} + \beta_2 x^{**}_{is,2} + \Lambda + \beta_K x^{**}_{is,K} + u^{**}_{is}$$

$$y^{*}_{is} = \frac{y^{*}_{is}}{s_{ui}}$$

$$x^{*}_{is,k} = \frac{x^{*}_{is,k}}{s_{ui}} \quad k = 1,2,...,K$$

$$u^{**}_{is} = \frac{u^{*}_{is}}{s_{ui}}$$

$$k = 1,2,...,K$$

$$i = 1,2,...,N$$

Applying the ordinary least squares to the transformed data, we obtain a consistent, asymptotically efficient and asymptotically normal estimator estimator ($\hat{\beta}$) when the explanatory variables are exogenous. See Kmenta [1986]. We call this procedure the Kmenta's procedure afterwards.

4. MONTE CARLO SIMULATION

In the present paper, to investigate the small sample properties of nonsense regression for pooling time series and cross-section data, we conduct a Monte Carlo simulation. We adopt the data generating process (DGP): $y_{is} = y_{i,s-1} + v_{is}$ and $x_{is} = x_{i,s-1} + w_{is}$ for $i = 1,...,N$ and $s = 1,...,T$, where the error terms (v_{is}, w_{is}) were generated from a bivariate normal distribution with independence across both individual and time period:

$$\begin{bmatrix} v_{is} \\ w_{is} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right).$$

This assumes strong exogenous and no serial correlation. Random numbers (v_{is}, w_{is}) were generated by the GAUSS procedure RNDNS. The data were generated by creating $T + 1000$ observations and discarding first 1000 observations to remove the effect of the initial conditions.

Using generated data, we estimate the model:

$$y_{is} = \alpha + \beta x_{is} + \varepsilon_{is}$$

by the Kmenta's procedure. In the Monte Carlo simulation, 5000 replication was made.

The results of the cases when $N = T$ are reported below. Table 1 reports the mean and standard deviation of the estimated coefficient ($\hat{\beta}$ and SSD($\hat{\beta}$)), the mean of the standard error of the coefficient (ESE($\hat{\beta}$)) and the mean of the DW ratio.

Table 1 Coefficients $\hat{\beta}$ and DW ratio

N(T)	$\hat{\beta}$	SSD($\hat{\beta}$)	ESE($\hat{\beta}$)	DW
10	0.00232	0.11809	0.01103	1.199
20	0.00299	0.07960	0.00440	1.319
30	0.00230	0.06538	0.00295	1.367
40	0.00152	0.05608	0.00241	1.409
50	-0.00035	0.04782	0.00213	1.442
60	0.00122	0.04114	0.00193	1.471
70	0.00037	0.03595	0.00178	1.498
80	-0.00011	0.03241	0.00167	1.523

The mean and standard deviation of the t-value (t_β and SSD(t_β)) and the probability when the estimated coefficient is significant (P($|t_\beta| > 1.96$)) are reported in Table 2.

Table 2 t-value t_β

N(T)	t_β	SSD(t_β)	P($\lvert t_\beta \rvert > 1.96$)
10	0.219	20.449	0.844
20	1.094	24.936	0.916
30	1.058	26.340	0.924
40	0.963	26.453	0.932
50	-0.367	24.597	0.927
60	0.593	22.675	0.929
70	0.105	21.193	0.914
80	-0.076	20.341	0.918

To compare these results to Kao's results for the LSDV (least squares with dummy variables) procedure, the corresponding results in Kao [1999] are reported in Table 3 and 4.

Table 3 Coefficients $\hat{\beta}$ and DW ratio by Kao[1999]

N(T)	$\hat{\beta}$	SSD($\hat{\beta}$)	ESE($\hat{\beta}$)	DW
10	-0.0028	0.2865	0.1495	0.6184
20	0.0024	0.2039	0.0727	0.3057
30	0.0012	0.1652	0.0479	0.2033
40	0.0006	0.1423	0.0358	0.1519
50	0.0001	0.1265	0.0286	0.1212
60	0.0004	0.1160	0.0237	0.1009

70	-0.0006	0.1072	0.0203	0.0864
80	-0.0001	0.1002	0.0178	0.0755

Table 4 t-value t_β by Kao[1999]

N(T)	t_β	SSD(t_β)	P($\|t_\beta\| > 1.96$)
10	-0.0185	1.9603	0.3150
20	0.0339	2.8375	0.4889
30	0.0248	3.4693	0.5723
40	0.0160	4.0065	0.6166
50	0.0060	4.4392	0.6598
60	0.0207	4.8987	0.6894
70	-0.0365	5.3014	0.7096
80	-0.0057	5.6369	0.7264

Table 1 and 2 show that the spurious regression still unsolved problem in this procedure too. As the number of observations ($N \times T$) becomes large, the average of the estimated coefficient and its standard deviation converges to zero, which means consistency of the estimator. The average of the estimated standard error of the coefficient also converges to zero. In consequence, the average of the estimated t-value converges to zero but its standard deviation remains relatively large. These results lead the high frequency of the case when the estimated coefficient becomes statistically significant. This phenomenon is used be called "spurious regression."

This problem is more serious than those in the LSDV procedure. The reason why such a result arises is that the estimated standard errors are relatively small in Kmenta's procedure while the estimated coefficients are similar in both cases. Furthermore, the average of the DW ratio is relatively large and becomes larger and larger in accordance with the number of the observations becomes large. This result indicates that detecting the spurious regression from the DW ratio becomes hard. This is an additional bothersome problem from an empirical researcher's point of view.

Another finding is that the spurious regression problem becomes less serious in Kmenta's procedure in accordance with the increase of the observations except the case when the sample size is very small ($N \times T$ < 900) while that problem in the LSDV procedure becomes more serious in accordance with the increase of the observations.

The results of the cases when we fix $N = 30$ are reported in Table 5 and 6, and Kao's corresponding results in Table 7 and 8.

Table 5 Coefficients $\hat{\beta}$ and DW ratio

T	$\hat{\beta}$	SSD($\hat{\beta}$)	ESE($\hat{\beta}$)	DW
10	0.00100	0.06838	0.00349	1.254
20	0.00204	0.06690	0.00309	1.310
30	0.00230	0.06538	0.00295	1.367
40	0.00112	0.06139	0.00297	1.407
50	0.00028	0.05797	0.00298	1.443
60	0.00352	0.05379	0.00300	1.468
70	0.00046	0.50404	0.00297	1.495
100	-0.00013	0.04130	0.00291	1.557
150	-0.00021	0.02910	0.00279	1.631

Table 6 t-value t_β

T	t_β	SSD(t_β)	P($\|t_\beta\| > 1.96$)
10	0.876	30.415	0.913
20	1.155	28.220	0.926
30	1.058	26.340	0.924
40	0.339	23.567	0.924
50	-0.069	22.183	0.916
60	1.117	20.142	0.912
70	0.166	18.877	0.899
100	-0.191	15.269	0.890
150	-0.097	11.129	0.837

Table 7 Coefficients $\hat{\beta}$ and DW ratio by Kao[1999]

T	$\hat{\beta}$	SSD($\hat{\beta}$)	ESE($\hat{\beta}$)	DW
10	0.0006	0.1669	0.0861	0.5686
20	0.0022	0.1660	0.0593	0.2991
30	0.0012	0.1652	0.0479	0.2033
40	0.0004	0.1639	0.0413	0.1538
50	-0.0001	0.1651	0.0368	0.1238
60	0.0013	0.1643	0.0336	0.1033
70	-0.0012	0.1647	0.0310	0.0890
100	0.0026	0.1634	0.0259	0.0626
150	0.0002	0.1647	0.0211	0.0418

Table 8 t-value t_β by Kao[1999]

T	t_β	SSD(t_β)	P($\|t_\beta\| > 1.96$)
10	0.0104	1.9546	0.3130
20	0.0397	2.8192	0.4858
30	0.0248	3.4693	0.5723
40	0.0131	3.9855	0.6218
50	-0.0074	4.5078	0.6651
60	0.0381	4.9167	0.6889
70	-0.0419	5.3318	0.7103
100	0.1072	6.3427	0.7548
150	0.0802	7.8426	0.8022

Additional fact findings are as follows. The effects of the time length (T) change are similar to those of the number of observations ($N \times T$) change, *e.g.*, the effects on the average of the estimated coefficient, its standard deviation, and the estimated standard error of the coefficient.
In comparison to the results of the LSDV procedure, the effects of the time length (T) are similar to those of the number of observations ($N \times T$), *e.g.*, the effects on the probability when the estimated coefficient is statistically significant.

The results of the cases when we fix $T = 30$ are reported in Table 9 and 10, and Kao's corresponding results in Table 11 and 12.

Table 9 Coefficients $\hat{\beta}$ and DW ratio

N	$\hat{\beta}$	SSD($\hat{\beta}$)	ESE($\hat{\beta}$)	DW
10	0.00508	0.10191	0.00710	1.364
20	0.00385	0.07581	0.00411	1.319
30	0.00230	0.06538	0.00295	1.364
40	0.00206	0.05849	0.00237	1.366
50	0.00187	0.05270	0.00202	1.365
60	0.00088	0.04844	0.00178	1.365
70	0.00004	0.04476	0.00160	1.362

Table 10 t-value t_β

N	t_β	SSD(t_β)	P($\lvert t_\beta \rvert > 1.96$)
10	1.003	19.754	0.882
20	0.909	23.250	0.909
30	1.058	26.340	0.924
40	1.168	28.680	0.936
50	1.235	29.393	0.946
60	0.553	30.341	0.942
70	0.306	30.686	0.943

Similar to the effects of the number of observations ($N \times T$) and the time length (T), the average of the estimated coefficient and its standard deviation converge to zero, and the average of the estimated standard error of the coefficient also converges to zero as the number of cross-section observations become large.

Table 11 Coefficients $\hat{\beta}$ and DW ratio by Kao[1999]

N	$\hat{\beta}$	SSD($\hat{\beta}$)	ESE($\hat{\beta}$)	DW
10	0.0029	0.2848	0.0832	0.2232
20	0.0037	0.2005	0.0588	0.2080
30	0.0012	0.1652	0.0479	0.2033
40	0.0009	0.1425	0.0415	0.2007
50	0.0001	0.1265	0.0371	0.1993
60	0.0017	0.1163	0.0339	0.1983
70	0.0020	0.1080	0.0313	0.1976

Table 12 t-value t_β by Kao[1999]

N	t_β	SSD(t_β)	P($\lvert t_\beta \rvert > 1.96$)
10	0.0261	3.5000	0.5732
20	0.0567	3.4577	0.5788
30	0.0248	3.4693	0.5723
40	0.0195	3.4532	0.5692
50	0.0259	3.4306	0.5699
60	0.0444	3.4478	0.5635
70	0.0588	3.4613	0.5745

In comparison to the results of the LSDV procedure, the changes of the number of cross-section observations (N) have similar effects on the LSDV and Kmenta's procedures, except the levels of some estimated statistics.

5. CONCLUSION

I conduct a Monte Carlo simulation to investigate the spurious regression in the time-wise autocorrelated and heteroskedastic procedure proposed by Kmenta [1986]. The results show that the spurious regression problem is serious in this procedure too. In comparison to the LSDV procedure, the results are summarized as follows,

1st: Kmenta's procedure has larger probabilities when the estimated coefficient is statistically significant than the LSDV procedure.

2nd: While the spurious regression problem in the LSDV procedure becomes more serious in accordance with the increase of the time length (T), this problem in Kmenta's procedure becomes less serious except the case when the sample size is very small ($N \times T$ < 900).

This result suggests that we should pay attention to the spurious regression problem even when we apply the Kmenta's procedure for pooling time series and cross-section data.

Finally, we should mention about the remaining problems. First one is that in the present paper we did not investigate the effects of the alternative methods to estimate the serial correlation coefficient. We adopt the correlation coefficient between e_{is} and $e_{i,s-1}$. For this estimation, we can take other methods:

The ratio of autocorrelation coefficients:

$$\hat{\rho}_i = \frac{\frac{1}{T-1}\sum_{s=2}^{T} e_{is} e_{i,s-1}}{\frac{1}{T}\sum_{s=1}^{T} e_{is}^2} \text{ or}$$

the regression coefficient e_{is} on $e_{i,s-1}$:

$$\hat{\rho}_i = \frac{\sum e_{is} e_{i,s-1}}{\sum e_{i,s-1}^2}.$$

These alternative methods might affect the small sample properties. We should investigate this remaining problem in further investigation.

6. REFERENCES

Baltagi, B.H., *Econometrics*, Berlin: Springer-Verlag, 1998.

Choi, I., Spurious regressions and residual based tests for cointegration when regressors are cointegrated, *Journal of Econometrics*, 60, 313-320, 1994.

Entorf, H., Random walks with drifts: nonsense regression and spurious fixed-effect estimation, *Journal of Econometrics*, 80, 287-296, 1997.

Granger, C.W.J., and P. Newbold, Spurious regressions in econometrics, *Journal of Econometrics*, 2, 111-120, 1974.

Kao, C., Spurious regression and residual-based tests for cointegration in panel data, *Journal of Econometrics*, 90, 1-44, 1999.

Kmenta, J., Elements of Econometrics, MacMillan, 1986.

Phillips, P.C.B., Understanding spurious regressions in econometrics, *Journal of Econometrics*, 33, 311-340, 1986.