JETEM

Journal of Economic Theory and Econometrics

The Journal of the Korean Econometric Society

(계 량 경 제 학 보)

September 2009

Volume 20. Issue 3. pp. 56-70

IV Estimation in the Presence of Serially Correlated Regressors and Disturbance Terms*

Chang-Jin Kim †· Donggeun Kim ‡· Geun-Hye Yang§

Abstract

We present a unified framework to solve the endogeneity problem in time-series regression models with serially a correlated disturbance term with ARMA(p,q) dynamics. Our focus is on the case in which lagged regressors are used as instrumental variables. The control function approach provides us with the solution to the problem. Besides, it provides us with an easy method for testing for endogeneity. Our Monte Carlo experiments confirm that the proposed two-step estimation procedure and the proposed test of endogeneity perform well for a sample size as small as 250.

Keywords Endogeneity, IV Estimation, Autoregressive Disturbance term, Alternative Two-Step Procedure, Generated Regressors.

JEL Classification C13, C32

* This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2006-321-B00275). We appreciate Myoung-Jae Lee and two anonymous referees for useful comments.

† corresponding author: Department of Economics, Korea University and Department of Economics, University of Washington (e-mail: changjin@u.washington.edu);

‡ Department of Economics, Ajou University (e-mail: kimdongg@ajou.ac.kr);

§ Kenan-Flagler Business School, University of North Carolina (e-mail: Geunhye_Yang@unc.edu)

1. Introduction

In most time-series regression models with endogenous regressors, lagged regressors are commonly used as instrumental variables. Recent and important examples in empirical macroeconomics include estimation of the hybrid new Keynesian Phillips curve (Gali and Gertler (1999), new Keynesian IS curve (Fuhrer and Rudebusch (2004), and forward-looking monetary policy rule (Clarida, Gali, and Gertler (2000), etc. However, in the presence of serially correlated disturbance term, the lagged regressors may not always serve as valid instruments. When the disturbance term has an AR (p) dynamics, the conventional two-stage nonlinear least squares method or the GMM method (using lagged regressors as instruments) provide consistent estimators. If the disturbance term is serially correlated with ARMA (autoregressive moving average) dynamics, however, the estimators from these methods are not consistent because all the lagged regressors are invalid instruments.

To solve the problem of endogeneity under this situation, we present a two-step estimation procedure based on the control function approach. The idea is to decompose the disturbance term in the regression equation of interest into the two orthogonal portions: one that is correlated with the explanatory variables and the other that is not correlated with the explanatory variables. In the first step, we regress our endogenous explanatory variables on a set of instrumental variables including lagged regressors and get standardized residuals. In the second step, we incorporate the standardized residuals obtained from the first step as additional regressors in the regression equation to correct for the endogeneity problem and the serially correlated disturbances. These additional regressors, which are the portion of the disturbance term correlated with the existing regressors, act as bias correction terms and consistent estimates for the regression coefficients can be obtained. The second-step regression equation with these bias correction terms, modified to deal with serially correlated disturbances, can be estimated by either nonlinear least squares (NLS) or maximum likelihood estimation.

The remainder of the paper is organized as follows. In section 2, we present a two-step procedure for the case in which lagged regressors are used as instrumental variables and the disturbance term follows an AR (p) process. In Section 3, we present how the methodology introduced in Section 2 can be extended to the case of general ARMA (p, q) disturbance term. Section 4 discusses a test for endogeneity in the presence of serially correlated disturbances. In section 5, we report results from Monte Carlo experiments, from which we evaluate the performance of the proposed two-step estimation procedure and the test of endogeneity in finite samples. Section 6 provides some concluding remarks.

2. A Two-Step Procedure in the Presence of an AR (p) Disturbance Term

Let us consider the following model with endogeneity in which both the regressors and the disturbance term are serially correlated:

$$y_t = x_t'\beta + \eta_t \tag{1}$$

$$\Phi(L)\eta_t = \varepsilon_t, \quad \varepsilon_t \sim i.i.d.N(0,\sigma_\varepsilon^2) \tag{2}$$

$$\alpha(L)x_t = \alpha_0 + v_t, \quad v_t \sim i.i.d.N(0,\Sigma_v), \tag{3}$$

where $\Phi(L) = (1-\phi_1 L - \phi_2 L^2 - ... - \phi_p L^p)$ and $\alpha(L) = (1-\alpha_1 L - \alpha_2 L^2 - ... - \alpha_m L^m)$ are pth and mth order lag operators, respectively, and all the roots of $\Phi(L) = 0$ and $\alpha(L) = 0$ are outside the complex unit circle. In equation (3), we assume for simplicity that x_t is correlated with no other exogenous or predetermined variables except its own lagged values. For simplicity of exposition, we assume that all the regressors in the $k \times 1$ vector of x_t are endogenous, and endogeneity in the x_t vector can be specified as the following correlation structure between ε_t and the standardized v_t term:

$$\begin{bmatrix} v_t^* \\ \varepsilon_t \end{bmatrix} \sim i.i.d\, N(0,\, \Omega), \quad \Omega = \begin{bmatrix} I_k & \rho\sigma_\varepsilon \\ \rho'\sigma_\varepsilon & \sigma_\varepsilon^2 \end{bmatrix}, \tag{4}$$

where $v_t^* = \Sigma_v^{-\frac{1}{2}} v_t$ and ρ is a $k \times 1$ vector of the correlations between v_t and ε_t. In this setting, the conventional 2SLS estimator is inconsistent because all the lagged values of the explanatory variables are correlated with the error term η_t making them invalid instrumental variables. However, we note in this paper that the problem can be easily solved by a Heckman-type (1979) two-step procedure, in which we estimate a version of equation (1) modified to incorporate a bias-correction term that allows us to account for the serially correlated disturbances.

In order to deal with the problem of serial correlation in the disturbance term in equation (1), we first multiply both sides of equation (1) by $\Phi(L)$ to obtain:

$$\Phi(L) y_t = \Phi(L) x_t ' \beta + \varepsilon_t \quad . \tag{5}$$

Again, as shown in equation (4), note that the endogeneity problem arises from correlation between x_t and ε_t, which results from correlation between the disturbance terms v_t and ε_t. In order to take care of the endogeneity problem, we next take the Cholesky decomposition of the variance–covariance matrix $\Omega = AA'$ and divide the error term ε_t in equation (5) into two orthogonal components:

$$\begin{aligned} \varepsilon_t &= v_t^{*}{}' \gamma + e_t \\ e_t &\sim i.i.d\,(0,\ (1-\rho'\rho)\sigma_\varepsilon^2) \end{aligned} \tag{6}$$

where $v_t^* = \Sigma_v^{-\frac{1}{2}} v_t$, $A = \begin{pmatrix} I_k & 0 \\ \rho'\sigma_\varepsilon & \sqrt{1-\rho'\rho}\sigma_\varepsilon \end{pmatrix}$, and $\gamma = \rho\sigma_\varepsilon$. Equation (6) leads us to rewrite equation (5) as:

$$\Phi(L) y_t = \Phi(L) x_t ' \beta + v_t^{*}{}' \gamma + e_t \tag{7}$$

where the new disturbance term e_t is uncorrelated with either $\Phi(L) x_t$ or v_t^*. The $v_t^{*}{}'\gamma$ term in equation (7) can be considered as a bias correction term. The nature of the disturbance term in equation (7), along with equation (3), leads us to the following two-step procedure for consistent estimation of equation (7):[1]

[1] Note that when $\Phi(L) = \alpha(L)$ in equations (2) and (3), the model is not identified. Besides, the conventional two-stage nonlinear least squares (based on $\Phi(L) y_t = \Phi(L) \hat{x}_t ' \beta + e_t$) also provides us with a consistent estimator, which is the same as that based on the control function approach discussed in this section.

Step 1: Estimate equation (3) using OLS and obtain a vector of standardized residuals: $\hat{v}_t^* = \hat{\Sigma}_v^{-\frac{1}{2}} \hat{v}_t$.

Step 2: Estimate the following equation using nonlinear least-squares (NLS) by replacing v_t^* with $\hat{v}_t^*$ in equation (7):

$$\Phi(L)y_t = \Phi(L)x_t'\beta + \hat{v}_t^*'\gamma + u_t \quad . \tag{7'}$$

As discussed by Pagan (1984), the variances of the estimators in equation (7') differ from those in equation (7) due to the presence of the generated regressors $\hat{v}_t^*$ in equation (7'). However, we can derive the variances of the estimators relatively easily under some regularity conditions. By re-arranging equation (7'), we have

$$y_t = f(z_t;\theta) + \hat{v}_t^*'\gamma + u_t, \qquad u_t = e_t + (v_t^* - \hat{v}_t^*)'\gamma \tag{8}$$

where $f(z_t;\theta) = \phi(L)y_t + \Phi(L)x_t'\beta$ with $\phi(L) = \phi_1 L + \phi_2 L^2 + ... + \phi_p L^p$ and $\Phi(L) = 1 - \phi(L)$; $z_t = [y_{t-1}\, y_{t-2} \dots y_{t-p}\, x_t'x_{t-1}' \dots x_{t-p}']'$; and $\theta = [\beta', \phi_1, \phi_2, \cdots, \phi_p]'$. Writing equation (8) in matrix notations, we have:

$$Y = F(Z;\Theta) + \hat{v}^*\gamma + u, \qquad u = e + (v^* - \hat{v}^*)\gamma \tag{9}$$

The first-order conditions for the nonlinear least squares problem for equation (9) are given by:

$$\frac{\partial F(Z;\hat{\Theta})}{\partial \Theta}'(Y - F(Z;\hat{\Theta}) - \hat{v}^*\hat{\gamma}) = 0, \tag{10}$$

$$\hat{v}^{*'}(Y - F(Z;\hat{\Theta}) - \hat{v}^*\hat{\gamma}) = 0, \tag{11}$$

where $\hat{\Theta}$ and $\hat{\gamma}$ are the nonlinear least squares estimators. By noting that $\frac{\partial^2 F(Z;\Theta)}{\partial \Theta \partial \Theta'} = 0$, if we take the mean value expansion of equations (10) and (11) around $\hat{\Theta} = \Theta$ and $\hat{\gamma} = \gamma$, we have:

$$\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' \frac{\partial F(Z;\Theta^*)}{\partial \Theta}(\widehat{\Theta}-\Theta)+\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' \hat{v}^*(\hat{\gamma}-\gamma)=\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' u, \tag{12}$$

$$\hat{v}^{*'} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}(\widehat{\Theta}-\Theta)+\hat{v}^{*'}\hat{v}^*(\hat{\gamma}-\gamma)=\hat{v}^{*'}u, \tag{13}$$

where Θ^* lies between $\widehat{\Theta}$ and Θ. The system of equations given by (12) and (13) can be solved for $(\widehat{\Theta}-\Theta)$, and we have the following results:

$$\begin{aligned}
\widehat{\Theta}-\Theta &= \left[\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}\right]^{-1} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} u \\
&= \left[\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}\right]^{-1} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} (e+(v^*-\hat{v}^*)\gamma) \\
&= \left[\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}\right]^{-1} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} (e+v^*\gamma) \\
&\xrightarrow{p} \left[p\lim \frac{1}{T}\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}\right]^{-1} p\lim \frac{1}{T}\left(\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} (e+v^*\gamma)\right) = 0,
\end{aligned}$$

(14)

where $\widehat{M}_{\hat{v}^*} = I_T - \hat{v}^*(\hat{v}^{*'}\hat{v}^*)^{-1}\hat{v}^{*'}$ is an idempotent matrix, and we have

$$\begin{aligned}
\sqrt{T}(\widehat{\Theta}-\Theta) &= \left[\frac{1}{T}\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}\right]^{-1} \frac{1}{\sqrt{T}}\left(\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} (e+v^*\gamma)\right) \\
&\xrightarrow{d} N\left(0,(\sigma_e^2+\gamma'\gamma)\left[p\lim \frac{1}{T}\frac{\partial F(Z;\Theta^*)}{\partial \Theta}' M_{\hat{v}^*} \frac{\partial F(Z;\Theta^*)}{\partial \Theta}\right]^{-1}\right)
\end{aligned} \tag{15}$$

Furthermore, it can be shown that

$$\widehat{\sigma}_u^2 = \frac{1}{T}(Y-F(Z;\widehat{\Theta})-\hat{v}^*\hat{\gamma})'(Y-F(Z;\widehat{\Theta})-\hat{v}^*\hat{\gamma}) \xrightarrow{p} \sigma_e^2 \tag{16}$$

Based on these results, the asymptotic variance of $\widehat{\Theta}$ can be estimated by:

$$\widehat{Var}(\widehat{\Theta}) = (\widehat{\sigma}_u^2+\hat{\gamma}'\hat{\gamma})\left[\frac{\partial F(Z;\widehat{\Theta})}{\partial \Theta}' M_{\hat{v}^*} \frac{\partial F(Z;\widehat{\Theta})}{\partial \Theta}\right]^{-1} \tag{17}$$

When the disturbance terms in (4) are normally distributed, we can estimate the second-step regression equation (7') with the maximum likelihood estimation method. Let $\psi = \left[\Theta' \ \gamma' \ \sigma_u^2\right]'$ be the vector of the parameters and let $\ln L(\psi)$ be the log likelihood function for equation (7'). Then, the maximum likelihood estimator $\widehat{\Theta}_{ML}$ the same as the nonlinear least squares estimator $\widehat{\Theta}$. Furthermore, by noting that the information matrix is block diagonal, one can show that the estimator for the variance-covariance matrix of $\widehat{\Theta}_{ML}$ or $\widehat{\Theta}$ given in equation (17) can alternatively be obtained from the first $k \times k$ block of

$$\widehat{Var}(\widehat{\theta}) = \left[-\frac{\partial^2 \ln L(\psi)}{\partial \psi \partial \psi '} | \psi = \widehat{\psi}^* \right]^{-1} \tag{18}$$

where k is the dimension of Θ and $\widehat{\psi}^* = \left[\widehat{\Theta}' \ \widehat{\gamma}' \ (\widehat{\sigma}_u^2 + \widehat{\gamma}' \ \widehat{\gamma}) \right]'$.

3. A Two-Step Procedure in the Case of an ARMA (p, q) Disturbance Term

In this Section, we consider how the proposed two-step procedure can be modified to deal with the case in which we have the following ARMA (p, q) disturbance term in equation (1):

$$\Phi(L)\eta_t = \theta(L)\varepsilon_t . \tag{19}$$

We assume that all roots of $\Phi(L) = 0$ and $\theta(L) = 0$ lie outside the complex unit circle. Furthermore, we assume that the disturbance terms are normally distributed, in order to employ the maximum likelihood estimation. To deal with the serial correlation in equation (19), we first multiply both sides of equation (1) by $\Phi(L)$. Then, as in the case of the AR (p) disturbance term, we use the Cholesky decomposition of the variance-covariance matrix Ω in equation (4) to divide the error term ε_t into two orthogonal components as in equation (6). This results in the following modification to equation (19):

$$\begin{aligned} \Phi(L)y_t &= \Phi(L)x_t '\beta + \theta(L)\varepsilon_t \\ &= \Phi(L)x_t '\beta + \theta(L)(v_t^* '\gamma + e_t) \\ &= \Phi(L)x_t '\beta + \theta(L)v_t^* '\gamma + \theta(L)e_t . \end{aligned} \tag{20}$$

Note that the $\theta(L)e_t$ term is uncorrelated with either $\Phi(L)x_t '$ or $\theta(L)v_t '$, providing the justification for the two-step procedure given below:

Step 1: Estimate equation (3) using OLS and obtain the standardized residual $\hat{v}_t^*$.

Step 2: Estimate the following equation using MLE by replacing v_t^* in equation (20) with the estimated standardized residual $\hat{v}_t^*$:

$$\Phi(L)y_t = \Phi(L)x_t'\beta + \theta(L)\hat{v}_t^{*}{}'\gamma + \theta(L)u_t, \tag{20'}$$

where $u_t = e_t + (v_t^* - \hat{v}_t^*)'\gamma$.

In order to estimate equation (20') with an MA (q) error term, we need to consider the following state-space representation, which consists of a measurement equation and a transition equation:[2]

Measurement equation: $$\Phi(L)y_t = \Phi(L)x_t'\beta + \theta(L)\hat{v}_t^{*}{}'\gamma + [1\ 0\ 0 \ldots 0]\begin{bmatrix} a_{1t} \\ a_{2t} \\ \vdots \\ a_{mt} \end{bmatrix} \tag{21}$$

$$(\Leftrightarrow \quad y_t = \phi(L)y_t + \theta(L)\hat{v}_t^{*}{}'\gamma + Ha_t)$$

Transition equation: $$\begin{bmatrix} a_{1t} \\ a_{2t} \\ \vdots \\ a_{mt} \end{bmatrix} = \begin{bmatrix} 0\ 1\ 0 \cdots 0 \\ 0\ 0\ 1 \cdots 0 \\ \vdots\ \vdots\ \vdots \quad \vdots \\ 0\ 0\ 0 \quad 0 \end{bmatrix} \begin{bmatrix} a_{1t-1} \\ a_{2t-1} \\ \vdots \\ a_{mt-1} \end{bmatrix} + \begin{bmatrix} 1 \\ \theta_1 \\ \vdots \\ \theta_{m-1} \end{bmatrix} u_t \tag{22}$$

$$(\Leftrightarrow \quad a_t = Fa_{t-1} + \mathrm{R}\,u_t)$$

where $m = \max(p, q+1)$. Given the above state-space representation, the Kalman filter is readily available for estimation via MLE, based on the prediction error decomposition.

Defining $a_{t|t-1} = E(a_t \mid I_{t-1})$, $P_{t|t-1} = Cov(a_t \mid I_{t-1})$, $a_{t|t} = E(a_t \mid I_t)$, $P_{t|t} = Cov(a_t \mid I_t)$, $\xi_{t|t-1} = E(y_t \mid I_{t-1})$, and $f_{t|t-1} = \mathrm{var}(y_t \mid I_{t-1})$, where I_s refers to information up to time s, the Kalman filter equations for the state-space model given by equations (21) and (22) can be derived as:

[2] For a state-space representation of a model that involves an MA(q) disturbance term, refer to Harvey (1981) and Hamilton (1994).

$$a_{t|t-1} = Fa_{t-1|t-1} \tag{23}$$

$$P_{t|t-1} = FP_{t-1|t-1}F' + \sigma_u^2 RR' \tag{24}$$

$$\xi_{t|t-1} = y_t - \phi(L)y_t - \theta(L)\hat{v}_t^{*}{}'\gamma \tag{25}$$

$$f_{t|t-1} = HP_{t|t-1}H' \tag{26}$$

$$a_{t|t} = a_{t|t-1} + P_{t|t-1}H' f_{t|t-1}^{-1}\eta_{t|t-1} \tag{27}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H' f_{t|t-1}^{-1}HP_{t|t-1} \tag{28}$$

Defining $\Theta = \left[\phi_1 .. \phi_p \ \theta_1 .. \theta_q \ \beta'\right]'$ and $\psi = \left[\Theta' \ \gamma' \ \sigma_u^2\right]'$, the above Kalman filter equations provides us with the following log likelihood function based on the prediction error decomposition:

$$\ln L(\psi) = \sum_{t=1}^{T} \ln(f(y_t \mid I_{t-1})), \tag{29}$$

where $f(y_t \mid I_{t-1}) = \frac{1}{\sqrt{2\pi f_{t|t-1}}} \exp(-\frac{\eta_{t|t-1}^2}{2f_{t|t-1}})$.

In light of the discussion in Section 2, the variance-covariance matrix of $\widehat{\Theta}$ that accounts for the generated regressors in the second-step regression can be obtained from the first $k \times k$ block of

$$\widehat{Var}(\widehat{\theta}) = \left[-\frac{\partial^2 \ln L(\psi)}{\partial\psi\partial\psi'} | \psi = \hat{\psi}^{*}\right]^{-1} \tag{30}$$

where k is the dimension of Θ and $\hat{\psi}^{*} = \left[\widehat{\Theta}' \ \hat{\gamma}' \ (\hat{\sigma}_u^2 + \hat{\gamma}'\hat{\gamma})\right]'$.

4. Testing for Endogeneity

An important advantage of the control function approach is that it provides us with an easy method for testing for endogeneity. In order to test for endogeneity, we can check for the significance of the γ parameter in equation (7'). As we have $\gamma = \rho\sigma_\varepsilon$, an estimate of $\gamma = 0$ means there is no correlation between x_t and ε_t, suggesting an absence of endogeneity. In checking the significance of γ, we can employ the conventional Wald test statistic. Note that, under the null hypothesis of $\gamma = 0$, equation (7') is not subject to the 'generated regressor' problem raised by Pagan (1984). Therefore, the conventional Wald test statistic given below has the usual asymptotic chi-square distribution:

$$Wald = \hat{\gamma}' Cov(\hat{\gamma})^{-1} \hat{\gamma} \quad \sim \quad \chi^2(J), \tag{31}$$

where J is the dimension of $\hat{\gamma}$.

5. Monte Carlo Experiment

In this section, we discuss the use of Monte Carlo experiments in order to investigate the finite sample properties of the proposed two-step procedure. We generated 10,000 sets of data (for each sample size T = 250, 500, and 1,000), according to the following data generating process:

$$y_t = \beta_0 + \beta_1 x_t + \eta_t \tag{32}$$

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + v_t \tag{33}$$

[Case 1] for AR (1) disturbances: $\eta_t = \phi\eta_{t-1} + \varepsilon_t$ (34)

[Case 2] for ARMA (1, 1) disturbances: $\eta_t = \phi\eta_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$ (35)

$$\begin{bmatrix} v_t \\ \varepsilon_t \end{bmatrix} \sim i.i.d\ N(0,\ \Omega_0),\ \ \Omega_0 = \begin{bmatrix} \sigma_v^2 & \rho\sigma_v\sigma_\varepsilon \\ \rho'\sigma_v\sigma_\varepsilon & \sigma_\varepsilon^2 \end{bmatrix}, \tag{36}$$

where we set $\beta_0 = 1$; $\beta_1 = 1$; $\alpha_0 = 0.8$; $\alpha_1 = 0.8$; $\sigma_v = 0.5$; $\rho = 0.5$; and $\sigma_\varepsilon = 0.5$. We set $\phi = 0.6$ and $\theta = 0.5$.

We estimate the model using the proposed two-step procedure, and the results for the second-step regression are reported in Tables 1 and 2. Table 1 reports the results when the disturbance term in equation (18) follows an AR (1) process. For estimation of the second-step regression equation, we employ nonlinear least squares (NLS). Regardless of the sample size and the true value of ϕ, ignoring the endogeneity problem results in large and statistically significant bias in the parameter estimates. However, when the proposed two-step procedure is employed, the mean and the standard deviation of the 10,000 estimates for each parameter suggest that the bias almost disappears. Table 2 reports the results when the disturbance term in equation (32) follows an ARMA (1, 1) process. We estimate the second-step regression using the Kalman filter and maximum likelihood estimation based on the state-space representation of the model. As in the simulation results from the AR (1) disturbance process, the bias almost disappears and is statistically insignificant when the proposed two-step procedure is employed. The Monte Carlo experiment results in this section show that the proposed two-step procedure and the proposed test of endogeneity work well for a sample size as small as 250.

In order to gauge the finite-sample properties of the proposed endogeneity tests, we also perform a Monte Carlo experiment based on the DGP above. For a sample size of 250, we consider the empirical size of the Wald test under the Null hypothesis. We thus generate 10,000 sets of data under the Null hypothesis of $\rho = 0$. For each generated data set, we perform the Wald test, using the critical values obtained from the $\chi^2(1)$ distribution. For the case of AR (1) disturbance term, at the 1%, 5%, and 10% significance levels, the percentage rejections of the null hypothesis were 0.8%, 4.0%, and 8.5%, respectively. For the case of ARMA (1, 1) disturbance term, at the 1%, 5%, and 10% significance levels, the percentage rejections of the null hypothesis were 0.8%, 4.9%, and 10.2%, respectively. Thus, the distribution of the Wald test statistic in equation (31) is reasonably well approximated by the χ^2 distribution at a sample size as small as 250.

6. Concluding Remarks

If there exist autoregressive disturbance term in time series models with endogenous regressors, the conventional two-stage nonlinear least squares or the GMM (generalized methods of moments) using lagged regressors as instruments may not always result in consistent estimates. This is particularly the case when the disturbance term follows an ARMA process. In this paper, we show that a two-step estimation procedure based on the control function approach provides us with a clue the solution of the problem under such situation.

Recently, Kim, Osborn, and Zhang (2006) report statistically significant serial correlation in the disturbance terms of the stylized hybrid new Keynesian Phillips curves, casting doubt on the validity of these commonly used instrumental variables. Thus, an alternative two-step procedure based on the control function approach would shed new light on IV estimation of many important empirical macroeconomic time series models, in the presence of ARMA disturbance term. Besides, it provides us with an easy method for testing for endogeneity. Our Monte Carlo experiments confirm that the proposed two-step estimation procedure and the proposed test of endogeneity perform well for a sample size as small as 250.

A nice thing about the proposed methodology based on the control function approach is its expandability. For example, under the normality assumption, it can easily be extended to deal with the endogeneity problem in the presence of heteroscedasticity of known form in the disturbance term.

REFERENCES

Clarida, R., J. Gali, and M. Gertler (2000) "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," Quarterly Journal of Economics, 115(1) 147-180.

Fuhrer, J.C. and G.D. Rudebusch (2004) "Estimating the Euler Equation for Output," Journal of Monetary Economics, 51, 1133-1153.

Gali, J. and M. Gertler (1999) "Inflation Dynamics: a Structural Econometric Analysis," Journal of Monetary Economics, 44, 195-222.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press.

Harvey, A. C. (1981), *Time Series Models*, Oxford: Phillip Allan and Humanities Press.

Heckman, J. (1979), "Sample selection bias as a specification error," Econometrica, 47, pp. 153–162.

Kim, D.H, D.R. Osborn, and C. Zhang (2006) "The New Keynesian Phillips Curve of Rational Expectations," Journal of Money, Credit and Banking 40 (4), June 2008, 667 - 699.

Pagan, A. (1984), "Econometric issues in the analysis of regressions with the generated regressors," International Economic Review, Vol. 25, No. 1, pp. 221–247.

Table 1. Results of the Monte Carlo Experiment: AR (1) Disturbance Term [$\phi = 0.6$]

$$\text{Model: } y_t = \beta_0 + \beta_1 x_t + \eta_t,\ \eta_t = \phi\eta_{t-1} + \varepsilon_t$$

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + v_t$$

$$\begin{bmatrix} v_t \\ \varepsilon_t \end{bmatrix} \sim i.i.d\ N(0, \Omega_0),\quad \Omega_0 = \begin{bmatrix} \sigma_v^2 & \rho\sigma_v\sigma_\varepsilon \\ \rho'\sigma_v\sigma_\varepsilon & \sigma_\varepsilon^2 \end{bmatrix}$$

$$\beta_0 = 1;\ \beta_1 = 1;\ \alpha_0 = 0.8;\ \alpha_1 = 0.8;\ \sigma_v = 0.5;\ \rho = 0.5;\ \sigma_\varepsilon = 0.5$$

$$\text{Transformed model: } y_t = \phi y_{t-1} + \beta_0(1-\phi) + \beta_1(x_t - \phi x_{t-1}) + \gamma\hat{v}_t^* + e_t$$

N = 250					
		Ignoring endogeneity		Proposed 2-step Procedure	
Parameters	True	Mean	SD	Mean	SD
β_0	1.000	0.084	0.137	1.041	0.443
β_1	1.000	1.229	0.028	0.990	0.109
ϕ	0.600	0.608	0.053	0.582	0.050
σ_v	0.500	0.437	0.020	--	--
$\gamma = \rho\sigma_\varepsilon$	0.250	--	--	0.260	0.111
$\sigma_e = \sqrt{(1-\rho^2)}\sigma_\varepsilon$	0.433	--	--	0.429	0.019
N = 500					
		Ignoring endogeneity		Proposed 2-step Procedure	
Parameters	True	Mean	SD	Mean	SD
β_0	1.000	0.086	0.096	1.049	0.338
β_1	1.000	1.228	0.020	0.988	0.083
ϕ	0.600	0.615	0.037	0.592	0.035
σ_v	0.500	0.438	0.014	--	--
$\gamma = \rho\sigma_\varepsilon$	0.250	--	--	0.263	0.085
$\sigma_e = \sqrt{(1-\rho^2)}\sigma_\varepsilon$	0.433	--	--	0.431	0.014
N = 1000					
		Ignoring endogeneity		Proposed 2-step Procedure	
Parameters	True	Mean	SD	Mean	SD
β_0	1.000	0.089	0.067	1.023	0.227
β_1	1.000	1.228	0.014	0.994	0.056
ϕ	0.600	0.619	0.026	0.596	0.025
σ_v	0.500	0.439	0.010	--	--
$\gamma = \rho\sigma_\varepsilon$	0.250	--	--	0.256	0.057
$\sigma_e = \sqrt{(1-\rho^2)}\sigma_\varepsilon$	0.433	--	--	0.432	0.010

Table 2. Results of the Monte Carlo Experiment: ARMA (1, 1) Disturbance Term
[$\phi = 0.6$, θ=0.5]

Model $y_t = \beta_0 + \beta_1 x_t + \eta_t$, $\eta_t = \phi\eta_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$

Transformed model $y_t = \phi y_{t-1} + \beta_0(1-\phi) + \beta_1(x_t - \phi x_{t-1}) + \gamma(\hat{v}_t^* + \theta\hat{v}_{t-1}^*) + \theta e_{t-1} + e_t$

$\beta_0 = 1$; $\beta_1 = 1$; $\alpha_0 = 0.8$; $\alpha_1 = 0.8$; $\sigma_v = 0.5$; $\rho = 0.5$; $\sigma_\varepsilon = 0.5$

		N = 250			
		Ignoring endogeneity		Proposed 2-step Procedure	
Parameters	True	Mean	SD	Mean	SD
β_0	1.000	0.165	0.149	0.993	0.317
β_1	1.000	1.209	0.027	1.002	0.073
ϕ	0.600	0.581	0.062	0.587	0.026
θ	0.500	0.504	0.069	0.510	0.066
σ_v	0.500	0.441	0.020	--	--
$\gamma = \rho\sigma_\varepsilon$	0.250	--	--	0.248	0.079
$\sigma_e = \sqrt{(1-\rho^2)}\sigma_\varepsilon$	0.433	--	--	0.428	0.019
		N = 500			
		Ignoring endogeneity		Proposed 2-step Procedure	
Parameters	True	Mean	SD	Mean	SD
β_0	1.000	0.170	0.105	1.001	0.221
β_1	1.000	1.208	0.019	1.000	0.051
ϕ	0.600	0.588	0.043	0.593	0.039
θ	0.500	0.501	0.048	0.504	0.046
σ_v	0.500	0.443	0.014	--	--
$\gamma = \rho\sigma_\varepsilon$	0.250	--	--	0.250	0.055
$\sigma_e = \sqrt{(1-\rho^2)}\sigma_\varepsilon$	0.433	--	--	0.431	0.014
		N = 1000			
		Ignoring endogeneity		Proposed 2-step Procedure	
Parameters	True	Mean	SD	Mean	SD
β_0	1.000	0.168	0.074	0.999	0.154
β_1	1.000	1.208	0.013	1.000	0.035
ϕ	0.600	0.590	0.031	0.596	0.028
θ	0.500	0.500	0.034	0.503	0.033
σ_v	0.500	0.444	0.010	--	--
$\gamma = \rho\sigma_\varepsilon$	0.250	--	--	0.249	0.038
$\sigma_e = \sqrt{(1-\rho^2)}\sigma_\varepsilon$	0.433	--	--	0.432	0.010