Redundancy of lagged regressors revisited

Stanislav Anatolyev

New Economic School, Moscow

Abstract

In a recent *Econometric Theory* problem, it was demonstrated that in a conditionally heteroskedastic time series regression with martingale difference errors the use of lagged values of regressors as instruments may not increase the efficiency of estimation relative to ordinary least squares. We provide an example of an analogous phenomenon in a model with serially correlated errors, where the optimal instrumental variables estimator is asymptotically as efficient as the instrumental variables estimator constructed as optimal when ignoring the presence of conditional heteroskedasticity.

Address

New Economic School, Nakhimovsky prospect, 47, room 1721, Moscow, 117418, Russia. E-mail: sanatoly@nes.ru

Acknowledgements

I thank a referee for providing useful comments which improved the presentation, and the coeditor Paolo Paruolo for his patience.

Problem and motivation

In a recent *Econometric Theory* problem proposed in Anatolyev (2003), it is demonstrated that under conditional heteroskedasticity (CH) the use of lagged values of regressors as instruments may not increase the efficiency of generalized method of moments (GMM) estimation relative to ordinary least squares (OLS). Carrasco (2004) proves this result using the Hansen (1985) optimal instrumental variables (OIV) approach, while Dhaene (2004) derives it from the instrument redundancy criteria of Breusch, Qian, Schmidt and Wyhowski (1999).

It appears that an analogous phenomenon can occur in models with serially correlated errors, with the OLS estimator being replaced by the GMM estimator that would be OIV in the absence of CH. Consider the following stationary time series regression similar to that in Anatolyev (2003):

$$y_t = \beta x_t + e_t, \quad E\left[e_t | x_t, e_{t-2}, x_{t-1}, e_{t-3}, \cdots\right] = 0,$$

where all variables are scalars. Note that e_{t-1} is missing in the conditioning information which implies that the error e_t is serially correlated, in contrast to Anatolyev (2003) where the error had the martingale difference property. Suppose that the error e_t is conditionally heteroskedastic with

$$\begin{aligned} E\left[e_t^2 | x_t, e_{t-2}, x_{t-1}, e_{t-3}, \cdots\right] &= \omega_0 + \lambda_0 \left(x_t - \mu_0\right)^2, \quad \omega_0 > 0, \quad \lambda_0 \geq 0, \\ E\left[e_t e_{t-1} | x_t, e_{t-2}, x_{t-1}, e_{t-3}, \cdots\right] &= \omega_1 + \lambda_1 \left(x_t - \mu_1\right)^2, \quad \omega_1 > 0, \quad \lambda_1 \geq 0. \end{aligned}$$

The object of estimation is β, and the class of GMM estimators we consider contains elements that use the regressor x_t and its lags as instruments, possibly an infinite number thereof. Let us compare two estimators from this class. The first one is the OIV estimator that acknowledges the presence of CH in the form described above. It is an analog of the OIV estimator in the Anatolyev (2003) problem. The second one is the GMM estimator that is constructed to be OIV in the absence of CH, i.e. in exactly the same problem as above except that the conditional variance and first-order autocovariance are thought to be constant. This is a legitimate GMM estimator, but, of course, it need not be asymptotically as efficient as the OIV in general. It is an analog of the OLS estimator in the Anatolyev (2003) problem; indeed, OLS is OIV when martingale difference errors are conditionally homoskedastic.

Assume that x_t can be represented as

$$x_t = \sum_{i=0}^{\infty} \varphi_i \eta_{t-i},$$

where η_t's are IID standard normal. Determine under what restrictions placed on parameters the two estimators are asymptotically equally efficient.

Solution and discussion

Let the representation of the optimal instrument that acknowledges the presence of CH be

$$z_t = \sum_{i=0}^{\infty} g_i \eta_{t-i},$$

and the representation of the instrument that would be optimal in the absence of CH be

$$\mathring{z}_t = \sum_{i=0}^{\infty} \mathring{g}_i \eta_{t-i}.$$

We are looking for optimal instruments in these forms because the space of present and past x_t is exactly spanned by the present and past innovations in x_t. In other words, instead of the system of moment conditions $E\left[x_{t-k}e_t\right] = 0$ for $k \geq 0$ we exploit the equivalent system $E\left[\eta_{t-k}e_t\right] = 0$ for $k \geq 0$. Such reformulation is more convenient because the innovations are mutually orthogonal, and as a result the system of optimality conditions simplifies greatly as in West (2001, section 3).

The optimality condition for z_t is (Hansen, 1985, equation 4.9; West, 2001, proposition 1)

$$\forall k \geq 0 \qquad E\left[\eta_{t-k}x_t\right] = E\left[\eta_{t-k}z_t e_t^2\right] + E\left[\eta_{t-k-1}z_t e_t e_{t-1}\right] + E\left[\eta_{t-k}z_{t-1}e_t e_{t-1}\right]. \qquad (1)$$

Note that exactly three terms figure in the right hand side because the moment functions $\eta_{t-k}e_t$ are conditionally serially correlated of order one. Indeed, by the Law of Iterated Expectations, for $k \geq 0$ and $j \geq 1$,

$$\begin{aligned} E\left[\eta_{t-k-j}z_t e_t e_{t-j}\right] &= E\left[E\left[\eta_{t-k-j}z_t e_t e_{t-j}|x_t, e_{t-2}, x_{t-1}, e_{t-3}, \cdots\right]\right] \\ &= E\left[\eta_{t-k-j}z_t E\left[e_t|x_t, e_{t-2}, x_{t-1}, e_{t-3}, \cdots\right] e_{t-j}\right] \\ &= 0. \end{aligned}$$

Denote $\sigma_x^2 \equiv \sum_{j=0}^{\infty} \varphi_j^2$ and $\tau \equiv E[\eta_t^4] - 1$ (in our case, $\tau = 2$). The left hand side in (1) equals φ_k. The components of the right hand side in (1) are equal to

$$\begin{aligned} E\left[\eta_{t-k}z_t e_t^2\right] &= E\left[\eta_{t-k}\left(\sum_{i=0}^{\infty} g_i\eta_{t-i}\right)\left(\omega_0 + \lambda_0\left(\sum_{j=0}^{\infty}\varphi_j\eta_{t-j} - \mu_0\right)^2\right)\right] \\ &= \left(\omega_0 + \lambda_0\left(\mu_0^2 + \sigma_x^2 + \tau\varphi_k^2\right)\right) g_k + 2\lambda_0\varphi_k \sum_{i=0,\, i\neq k}^{\infty} \varphi_i g_i, \end{aligned}$$

$$\begin{aligned} E\left[\eta_{t-k-1}z_t e_t e_{t-1}\right] &= E\left[\eta_{t-k-1}\left(\sum_{i=0}^{\infty} g_i\eta_{t-i}\right)\left(\omega_1 + \lambda_1\left(\sum_{j=0}^{\infty}\varphi_j\eta_{t-j} - \mu_1\right)^2\right)\right] \\ &= \left(\omega_1 + \lambda_1\left(\mu_1^2 + \sigma_x^2 + \tau\varphi_{k+1}^2\right)\right) g_{k+1} + 2\lambda_1\varphi_{k+1} \sum_{i=0,\, i\neq k+1}^{\infty} \varphi_i g_i, \end{aligned}$$

and

$$E\left[\eta_{t-k}z_{t-1}e_te_{t-1}\right] = E\left[\eta_{t-k}\left(\sum_{i=1}^{\infty}g_{i-1}\eta_{t-i}\right)\left(\omega_1+\lambda_1\left(\sum_{j=0}^{\infty}\varphi_j\eta_{t-j}-\mu_1\right)^2\right)\right]$$

$$= \begin{cases} \left(\omega_1+\lambda_1\left(\mu_1^2+\sigma_x^2+\tau\varphi_k^2\right)\right)g_{k-1}+2\lambda_1\varphi_k\sum\limits_{i=0,\, i\neq k-1}^{\infty}\varphi_{i+1}g_i & \text{if } k>0, \\ 2\lambda_1\varphi_0\sum\limits_{i=0}^{\infty}\varphi_{i+1}g_i \quad \text{if } k=0. \end{cases}$$

Therefore, the system (1) can be written in matrix form as follows:

$$\Phi = \mathrm{SG}, \tag{2}$$

where $\Phi \equiv \begin{bmatrix}\varphi_0\\ \varphi_1\\ \vdots\\ \varphi_k\\ \vdots\end{bmatrix}$, $\mathrm{G} \equiv \begin{bmatrix}g_0\\ g_1\\ \vdots\\ g_k\\ \vdots\end{bmatrix}$, $\mathrm{S} \equiv \begin{bmatrix} S_{0,0} & S_{0,1} & \cdots & S_{0,k} & \cdots \\ S_{1,0} & S_{1,1} & \cdots & S_{1,k} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ S_{k,0} & S_{k,1} & \cdots & S_{k,k} & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{bmatrix}$, and

$S_{k,k} = \omega_0 + \lambda_0\left(\mu_0^2+\sigma_x^2+\tau\varphi_k^2\right) + 4\lambda_1\varphi_k\varphi_{k+1},\ k \geq 0,$
$S_{k,k-1} = \omega_1 + \lambda_1\left(\mu_1^2+\sigma_x^2+\tau\varphi_k^2\right) + 2\lambda_0\varphi_k\varphi_{k-1} + 2\lambda_1\varphi_{k+1}\varphi_{k-1},\ k \geq 1,$
$S_{k,k+1} = \omega_1 + \lambda_1\left(\mu_1^2+\sigma_x^2+\tau\varphi_{k+1}^2\right) + 2\lambda_0\varphi_k\varphi_{k+1} + 2\lambda_1\varphi_k\varphi_{k+2},\ k \geq 0,$
$S_{k,m} = 2\lambda_0\varphi_k\varphi_m + 2\lambda_1\varphi_{k+1}\varphi_m + 2\lambda_1\varphi_k\varphi_{m+1},\ k \geq 0,\ m < k-1 \text{ or } m > k+1.$

Similarly, the optimality condition for $\mathring{z}_t$ is (recall that it is optimal *under conditional homoskedasticity*)

$$\forall k \geq 0 \qquad E\left[\eta_{t-k}x_t\right] = E\left[\eta_{t-k}\mathring{z}_t\right]E\left[e_t^2\right] + E\left[\eta_{t-k-1}\mathring{z}_t\right]E\left[e_te_{t-1}\right] + E\left[\eta_{t-k}\mathring{z}_{t-1}\right]E\left[e_te_{t-1}\right]. \tag{3}$$

Again, the left hand side in (3) equals φ_k. In the right hand side in (3),

$$E\left[e_t^2\right] = \omega_0 + \lambda_0\left(\mu_0^2+\sigma_x^2\right),$$

$$E\left[e_te_{t-1}\right] = \omega_1 + \lambda_1\left(\mu_1^2+\sigma_x^2\right),$$

$$E\left[\eta_{t-k}\mathring{z}_t\right] = E\left[\eta_{t-k}\left(\sum_{i=0}^{\infty}\mathring{g}_i\eta_{t-i}\right)\right] = \mathring{g}_k,$$

$$E\left[\eta_{t-k-1}\mathring{z}_t\right] = E\left[\eta_{t-k-1}\left(\sum_{i=0}^{\infty}\mathring{g}_i\eta_{t-i}\right)\right] = \mathring{g}_{k+1},$$

$$E\left[\eta_{t-k}\mathring{z}_{t-1}\right] = E\left[\eta_{t-k}\left(\sum_{i=1}^{\infty}\mathring{g}_{i-1}\eta_{t-i}\right)\right] = \begin{cases}\mathring{g}_{k-1} & \text{if } k>0,\\ 0 & \text{if } k=0.\end{cases}$$

Therefore, the vector of weights $\mathring{\mathrm{G}}$ implied by $\mathring{z}_t$ is a solution of the system

$$\Phi = \mathring{\mathrm{S}}\mathring{\mathrm{G}}, \tag{4}$$

where $\mathring{S}$ is a Toeplitz matrix with $\omega_0 + \lambda_0 (\mu_0^2 + \sigma_x^2)$ on the main diagonal, $\omega_1 + \lambda_1 (\mu_1^2 + \sigma_x^2)$ on the two diagonals next to the main diagonal, and zeros elsewhere.

Now when $\tau = 2$, some algebra shows that

$$\mathrm{S} = \mathring{\mathrm{S}} + 2\lambda_0 \Phi\Phi' + 2\lambda_1 \Phi_+ \Phi' + 2\lambda_1 \Phi\Phi'_+,$$

where Φ_+ is Φ with elements shifted one position up (i.e., $(\Phi_+)_k = \Phi_{k+1}$, $k \geq 0$). Then equation (2) may be written as

$$\Phi = \left(\mathring{\mathrm{S}} + 2\lambda_0 \Phi\Phi' + 2\lambda_1 \Phi_+ \Phi' + 2\lambda_1 \Phi\Phi'_+ \right) \mathrm{G}.$$

Suppose that $\Phi_+ = \phi\Phi$ for some scalar ϕ, then

$$\Phi \left(1 - 2 \left(\lambda_0 + 2\lambda_1 \phi\right) \Phi' \mathrm{G}\right) = \mathring{\mathrm{S}}\mathrm{G}.$$

This is the same as (4) up to the multiplicative scalar factor $1 - 2(\lambda_0 + 2\lambda_1\phi)\Phi'\mathrm{G}$. In this case, the solution G of (2) differs from the solution $\mathring{\mathrm{G}}$ of (4) only by a multiplicative factor, which implies that z_t and $\mathring{z}_t$ are essentially the same instrument.

The condition $\Phi_+ = \phi\Phi$ implies the AR(1) process for the regressor x_t, including the case of a serially independent regressor. Note that the derivation takes advantage of the fact that the distribution of η_t is symmetric with $\tau = 2$, although it may not be normal.

The author doubts that it is possible to solve the problem via the redundancy of moment conditions criteria as in Dhaene (2004).

References

Anatolyev, S. (2003) Redundancy of lagged regressors in a conditionally heteroskedastic time series regression: problem 03.1.2. *Econometric Theory* 19, 225–226.

Breusch, T., H. Qian, P. Schmidt & D. Wyhowski (1999) Redundancy of moment conditions. *Journal of Econometrics* 91, 89–111.

Carrasco, M. (2004) Redundancy of lagged regressors in a conditionally heteroskedastic time series regression: solution 03.1.2. *Econometric Theory* 20, 228–229.

Dhaene, G. (2004) Redundancy of lagged regressors in a conditionally heteroskedastic time series regression: solution 03.1.2. *Econometric Theory* 20, 227.

Hansen, L.P. (1985) A method for calculating bounds on the asymptotic variance-covariance matrices of generalized method of moments estimators. *Journal of Econometrics* 30, 203–228.

West, K.D. (2001) On optimal instrumental variables estimation of stationary time series models. *International Economic Review* 42, 1043–1050.