Limiting efficiency of OLS vs. GLS when regressors are fractionally integrated[1]

by

Walter Krämer and Uwe Hassler

Fachbereich Statistik, Universität Dortmund
D-44221 Dortmund, Germany

and

Fachbereich Wirtschaftswissenschaften, Freie Universität Berlin
Boltzmannstraße 20, D-14195 Berlin, Germany

Abstract

We show that previous results on the asymptotic efficiency of OLS versus GLS in the context of trending data carry over to regressors of the fractionally integrated type.

1 Introduction

We consider the standard linear regression model $y_t = \beta' x_t + u_t$ $(t = 1, \ldots, T)$ with stochastic regressors x_t, where the disturbances u_t are possibly autocorrelated but independent of the regressors x_t. There is a huge literature in econometrics and statistics in this context on the efficiency of the Ordinary Least Squares (OLS) estimates of the β's, which ignore the correlation of the errors, relative to the Generalized Least Squares (GLS) estimates, which take this correlation into account. In particular, starting with Grenander (1954), there has been much interest in conditions on the regressors and the error

[1] Research supported by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 475

correlation structure which guarantee that OLS is asymptotically as efficient as GLS (Grenander and Rosenblatt 1957, Chipman 1979, Krämer 1980, 1982, 1986, Phillips and Park 1988). For nonstochastic regressors and stationary errors, Grenander (1954) shows that in the bivariate model with a single nonstochastic regressor, the limiting ratio of the respective estimator variances depends on the spectral distribution functions of the regressor and error sequences, respectively, taking the value of one when the spectral density of the disturbances is constant on the spectrum of the regressor sequence, a condition which is easily seen to be satisfied for arbitrary stationary disturbances with finite spectral density when the regressor sequence is a polynomial or trigonometric function of time. This result has been generalized to multivariate regressors by Grenander and Rosenblatt (1957) and to stochastic regressors of the ARIMA–type by Krämer (1986) and Phillips and Park (1988).

Below we consider the bivariate regression

$$y_t = \alpha + \beta x_t + u_t \qquad (t = 1, \ldots, T), \tag{1}$$

where the regressor is nonstationary fractionally integrated:

$$(1 - L)^d x_t = v_t, \tag{2}$$

where $0,5 < d < 1,5$ and v_t is i.i.d. $(0, \sigma_v^2)$. Variables of this type are entertained with increasing frequency as models for economic time series such as interest rates, exchange rates and other variables which are not exactly integrated but are not stationary either (Robinson 1994, Hassler and Wolters 1995 or Baillie 1996 and the recent special issue of the Journal of Econometrics that is devoted to this topic). If one such variable is used to explain, say, individual investment or savings or some other economic micro variable, one can reasonably assume that the disturbances in this regression are possibly autocorrelated but stochastically independent of the regressor, so the problem of the relative efficiency of OLS immediately suggests itself.

2 Limiting efficiency of GLS

Following Krämer (1986) and Phillips and Park (1988), we focus on the $AR(p)$–case where

$$u_t + \rho_1 u_{t-1} + \ldots + \rho_p u_{t-p} = \varepsilon_t, \tag{3}$$

where the ε_t's are i.i.d. $(0, \sigma_\varepsilon^2)$ and where all roots of the polynomial $\rho(z) = 1 + \rho_1 z + \ldots + \rho_p z^p$ are outside the unit circle. Also we confine ourselves to the estimation of the slope parameter β in (1). We consider both the case where an intercept is included ($\alpha \neq 0$) and where an intercept is excluded ($\alpha = 0$) in (1).

Without an intercept the OLS–estimator for β is given by

$$\hat{\beta} = x'y/x'x, \tag{4}$$

where $x = (x_1, \ldots, x_T)'$ and $y = (y_1, \ldots, y_T)'$. With an intercept, one has to replace x and y by the deviations from the respective means. The GLS estimator $\tilde{\beta}$ for β is obtained by applying OLS to the transformed model

$$\tilde{y}_t = \alpha + \beta \tilde{x}_t + \varepsilon_t \qquad (t = p+1, \ldots, T), \tag{5}$$

where

$$\begin{aligned} \tilde{y}_t &= y_t + \rho_1 y_{t-1} + \ldots + \rho_p y_{t-p} \quad \text{and} \\ \tilde{x}_t &= x_t + \rho_1 x_{t-1} + \ldots + \rho_p x_{t-p}. \end{aligned} \tag{6}$$

i.e.

$$\tilde{\beta} = \frac{\tilde{x}'\tilde{y}}{\tilde{x}'\tilde{x}}. \tag{7}$$

In the model with an intercept, we again take deviations from the respective means (while it is not at all obvious that one can do this in the case of GLS, this procedure is legal here, see Krämer (1986) or Fiebig et al. 1996).

For ease of notation, we have in (5) discarded observations $1, \ldots, p$. As we are only interested in large sample asymptotics, this involves no loss in generality (in our finite sample Monte–Carlo–estimates below, however, all data are used for GLS).

When both the disturbances u_t in (1) and the innovations v_t in the $I(d)$–representation of x_t in (2) are i.i.d., the limiting distribution of $\hat{\beta} = \tilde{\beta}$ in the model without an intercept is derived by Wright (1996), who shows that

$$T^d(\hat{\beta} - \beta) \xrightarrow{d} (\int_0^1 W_d(r)^2 dr)^{-1} \int_0^1 W_d(r) dV(r), \tag{8}$$

where $\frac{1}{\sigma_u} V(r)$ is standard Brownian Motion,

$$W_d(r) = \frac{1}{\Gamma(d)} \int_0^r (r-s)^{d-1} dW(s) \tag{9}$$

is fractional Brownian motion and where $W_d(r)$ and $V(r)$ are the probability limits of

$$T^{\frac{1}{2}-d} x_{[Tr]} \quad \text{and} \quad T^{\frac{1}{2}} \sum_{t=1}^{[Tr]} u_t, \tag{10}$$

respectively.

In particular, we see from (8) that $\hat{\beta}$ converges to the true β faster than in the case of stationary regressors.

The proof of (8) rests on $\hat{\beta} - \beta = x'u/x'x$ and

$$\frac{1}{T^d} x'u \xrightarrow{d} \int_0^1 W_d(r) dV(r), \tag{11}$$

$$\frac{1}{T^{2d}} x'x \xrightarrow{d} \int_0^1 W_d(r)^2 dr. \tag{12}$$

With autocorrelated disturbances, these results do not apply. Neither can they be used to deduct the limiting distribution of GLS in (5), as the innovations $\tilde{v}_t$ in

$$(1-L)^d \tilde{x}_t = \tilde{v}_t \tag{13}$$

are no longer i.i.d. (rather, $\tilde{v}_t = v_t + \rho_1 v_{t-1} + \ldots + \rho_p v_{t-p}$). However, appealing to Phillips and Park (1988, p. 114, proof of theorem 2.2) with $V(r)/\sigma_\varepsilon(\Sigma\rho_i)$ a standard Brownian Motion, it is easily seen that (11) continues to hold. So (8) continues to hold as well, the only difference being that $V(r)$ must now be divided by $\sigma_\varepsilon(\Sigma\rho_i)$ (which is different from σ_u) to obtain a standard Brownian Motion.

THEOREM:

In the model (1) without an intercept, the limiting distributions (as $T \to \infty$) of OLS and GLS are identical. The same applies for the model with an intercept.

The theorem shows that OLS is asymptotically as efficient as GLS despite the presence of autocorrelated disturbances in the regression. As in Grenander (1954), the proof relies heavily on the fact that there is only one regressor. However, we conjecture that similar results obtain if the regressors can be represented as a multivariate $I(d)$–process or as independent $I(d)$ processes with (possibly) different $d's > 0,5$, but this has be shown along different lines, perhaps emulating Phillips and Park (1988) or Grenander and Rosenblatt (1957).

Proof of Theorem:

The key to the proof is the result from Hassler (1994, 1997) that empirical autocorrelations of $I(d)$–processes for $d > 0,5$ tend in probability to one. Using this result, we show that

$$\frac{\tilde{\beta} - \beta}{\hat{\beta} - \beta} \xrightarrow{p} 1 \quad \text{as} \quad T \to \infty, \tag{14}$$

which in turn immediately implies the theorem: From (14) we have

$$\begin{aligned} \tilde{\beta} - \beta &= \hat{\beta} - \beta + (\hat{\beta} - \beta) o_p(1), \quad \text{so} && (15) \\ T(\tilde{\beta} - \beta) &= T(\hat{\beta} - \beta) + o_p(1), \quad \text{so} && (16) \\ T(\tilde{\beta} - \hat{\beta}) &\xrightarrow{p} 0. && (17) \end{aligned}$$

We first prove (14) for a regression without an intercept. We can then write

$$\frac{\tilde{\beta}-\beta}{\hat{\beta}-\beta}=\frac{\tilde{x}'\varepsilon}{x'u}\times\frac{x'x}{\tilde{x}'\tilde{x}}, \tag{18}$$

where

$$\frac{\tilde{x}'\tilde{x}}{x'x}=\frac{\left(\sum(\rho_i x_{-i})\right)^2}{x'x}=\sum_{i,j=0}^{p}\rho_i\rho_j\frac{x'_{-i}x_{-j}}{x'x}\xrightarrow{p}\sum_{i,j=0}^{p}\rho_i\rho_j=\left(\sum\rho_i\right)^2 \tag{19}$$

and where the crucial relationship is

$$x'_{-i}x_j/x'x\xrightarrow{p}1\quad\text{for all}\quad i,j=0,\ldots,p.$$

Similarly, we have

$$\frac{\tilde{x}'\varepsilon}{x'u}=\frac{(\sum\rho_i x_{-i})'(\sum\rho_i u_i)}{x'u}=\sum_{i,j=0}^{p}\rho_i\rho_j\frac{x'_{-i}u_{-j}}{x'u}, \tag{20}$$

where once again, the second factor tends in probability to one. This can be seen along the lines of Krämer (1986), by verifying that

$$x'_{-i}u_{-j}=x'u+O_p(T^{\frac{1}{2}}) \tag{21}$$

and noting that $x'u=O_p(T^d)$.

To prove (15) for regressions with an intercept, one simply replaces all relevant vectors with deviations from their respective means, and verifies that (20) and (21) continue to hold.

Note that our proof shows slightly more than is asserted in the statement of the theorem: Not only have $T(\hat{\beta}-\beta)$ and $T(\tilde{\beta}-\beta)$ identical limiting distributions; in addition, $T(\hat{\beta}-\tilde{\beta})\xrightarrow{p}0$.

3 Relative efficiency of OLS in finite samples

Next we provide some Monte Carlo evidence as to the finite–sample relevance of our theorem. Table 1 gives the empirical mean square errors for various values of d, T and ρ obtained from $N = 1000$ replications when the disturbances are stationary $AR(1)$. It shows that the convergence to one of the relative

Table 1: Var(OLS) / Var(GLS)

T	ρ			
	-0.9	-0.3	0.3	0.9
a) d = 0.6				
25	7.13	1.14	1.10	6.17
50	12.33	1.14	1.07	5.65
100	16.63	1.13	1.09	4.54
250	19.14	1.14	1.07	4.34
400	20.02	1.12	1.04	3.65
b) d = 0.8				
25	6.40	1.14	1.06	4.28
50	9.06	1.13	1.06	3.73
100	10.90	1.08	1.04	2.83
250	10.54	1.04	1.04	2.21
400	8.05	1.04	1.05	1.86
c) d = 1.0				
25	4.14	1.10	1.10	4.62
50	5.74	1.04	1.10	2.40
100	5.74	1.06	1.03	1.86
250	4.28	1.03	1.03	1.45
400	3.06	1.01	1.00	1.18
d) d = 1.2				
25	3.50	1.07	1.06	2.66
50	3.54	1.07	1.01	2.52
100	3.06	1.03	1.03	1.40
250	2.10	1.02	1.00	1.24
400	1.53	0.99	1.00	1.09
e) d = 1.4				
25	2.34	1.07	1.03	2.11
50	2.53	1.03	1.01	1.92
100	1.76	1.02	1.00	1.34
250	1.51	1.00	1.01	1.16
400	1.21	1.00	1.01	1.11

efficiency of OLS is not uniform in ρ and that large inefficiencies persist for sample sizes up to $T = 400$ and disturbance–autocorrelation coefficients close to the border of the parameter space, in particular for $\rho \to -1$, i.e. for large positive correlations among the disturbances. Not surprisingly, the relative efficiency increases ceteris paribus as d increases (i.e. the more pronounced the trend, the closer comes OLS to GLS); however, it is not monotone in sample size: for large disturbance correlations, the relative efficiency of OLS first decreases as T increases, before the large sample asymptotics take effect, confirming Krämer (1986, table 1).

References

Baillie, R.T. (1996): "Long memory processes and fractional integration in econometrics." *Journal of Econometrics* 73, 5 – 59.

Chipman, J.S. (1979): "Efficiency of least–squares estimation of linear trend when residuals are autocorrelated." *Econometrica* 47, 115 – 127.

Fiebig, D.; Krämer, W. and Bartels, R. (1996): "The Frisch–Waugh Theorem and Generalized Least Squares." *Econometric Reviews* 15, 431 – 444.

Grenander, U. (1954): "On the estimation of regression coefficients in the case of an autocorrelated disturbance." *Annals of Mathematical Statistics* 25, 252 – 272.

Grenander, U. and Rosenblatt, M. (1957): *Statistical Analysis of Stationary Time Series*, New York (Wiley).

Hassler, U. (1994): "The sample autocorrelation function of I(1) processes." *Statistical Papers* 35, 1 – 16.

Hassler, U. (1997): "Sample autocorrelations of nonstationary fractionally integrated series." *Statistical Papers* 38, 43 – 62.

Hassler, U. and Wolters, J. (1995): "Long memory in inflation rates: International evidence." *Journal of Business and Economic Statistics* 13, 37 – 47.

Krämer, W. (1980): "Finite sample efficiency of Ordinary Least Squares in the Linear Regression Model with autocorrelated Errors." *Journal of the American Statistical Association* 75, 1005 – 1009.

Krämer, W. (1982): "Note on estimating linear trend when residuals are autocorrelated." *Econometrica* 50, 1065 – 1067.

Krämer, W. (1986): "Least squares regression when the independent variable follows on ARMA process." *Journal of the American Statistical Association* 81, 150 – 154.

Phillips, P.C.B. and Park, J.Y. (1988): "Asymptotic equivalence of Ordinary Least Squares and generalized Least Squares in regressions with integrated regressors." *Journal of the American Statistical Association* 83, 111 – 115.

Robinson, P.M. (1994): "Time series with strong dependence." In C.A. Sims (ed.): *Advances in Econometrics*, Cambridge University Press, 47 – 95.

Wright, J. (1996): "Least squares and fractional cointegration." Mimeo, Harvard University.