E C O N O M I C S B U L L E T I N

A New Variant of RESET for Distributed Lag Models

Dimitris Hatzinikolaou
University of Ioannina, Department of Economics

Athanassios Stavrakoudis
University of Ioannina, Department of Economics

Abstract

We propose a new variant of RESET that is appropriate for distributed lag models. Monte Carlo evidence on size and power strongly supports the use of the new variant instead of the traditional RESET.

Citation: Hatzinikolaou, Dimitris and Athanassios Stavrakoudis, (2005) "A New Variant of RESET for Distributed Lag Models." *Economics Bulletin*, Vol. 3, No. 56 pp. 1−4
Submitted: December 16, 2005. **Accepted:** December 19, 2005.
URL: http://www.economicsbulletin.com/2005/volume3/EB−05C10023A.pdf

1. Introduction

Ramsey's (1969) regression specification error test (RESET) and its variants are known to have high power against certain alternatives, e.g., incorrect functional form, but low against others, e.g., omitted variables or omitted lags; see, e.g., Thursby (1989, Tables 5, 7, 11-13). In this paper, we propose a new variant of RESET, which has high power against omitted lags. Considering such a variant of RESET is important, because in empirical economics we often encounter distributed lag models, e.g., trade balance equations incorporating the J-curve effect, inflation equations, money market equations, etc., and the erroneous omission of lags from these models will in general lead to invalid statistical inference. Using both ordinary least squares (OLS) and the Cochrane-Orcutt (C-O) method, we produce Monte Carlo evidence on the size and power of the proposed variant as well as of Ramsey's (1969) traditional RESET. We consider several true and null models and find that the proposed variant performs much better than the traditional RESET. After describing the test (Section 2) and our Monte Carlo setup (Section 3), we report our results and offer some possible explanations for the reported patterns (Section 4). Section 5 concludes the paper.

2. The tests

Consider the standard linear regression model, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{v}$, and assume that the data on **y** and **X** are stationary time-series. The RESET tests the hypothesis that this (null) model is specified correctly. Choose a $T \times M$ matrix **Z** of "test variables," apply OLS to the equation

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{u}\,, \tag{1}$$

and test the hypothesis H_0: $\boldsymbol{\gamma} = \mathbf{0}$ using a standard F test. Ramsey's (1969) choice of test variables is $\mathbf{z}_t = (\hat{Y}_t^2,\ \hat{Y}_t^3,\ ...,\ \hat{Y}_t^j)$, where $\hat{Y}_t = \mathbf{x}_t'\hat{\boldsymbol{\beta}}$ is the OLS fitted value from the null model. Let this test be denoted as POY(j). Our choice of test variables is $\mathbf{z}_t = (\hat{Y}_{t-1}^2,\ \hat{Y}_{t-2}^2, \ldots, \hat{Y}_{t-m}^2)$, which is appropriate when testing a null distributed lag model for omitted lags, e.g.,

$$Y_t = \alpha + \Sigma_{i=0}^{k} \beta_i X_{t-i} + v_t \tag{2}$$

against the alternative

$$Y_t = \alpha + \Sigma_{i=0}^{l} \beta_i X_{t-i} + u_t\,, \tag{3}$$

where $l > k$, $m > k$, and m may be greater than, equal to, or less than l. Let this variant of RESET be denoted as LOY(m). For example, if the researcher contemplates testing the null hypothesis of no lagged values of X_t ($k = 0$) against the alternative of $l = 1$ lag, then he/she can use LOY(1), which uses the test variable $z_t = \hat{Y}_{t-1}^2$; if he/she contemplates testing $k = 0$ or $k = 1$ against $l = 2$ lags, then LOY(2) is appropriate, where $\mathbf{z}_t = (\hat{Y}_{t-1}^2,\ \hat{Y}_{t-2}^2)$; and so on.

3. Monte Carlo design

The data for X_t and u_t are generated as follows:

$$X_t = \varphi X_{t-1} + \varepsilon_t, \quad \varphi = 0.0,\ 0.5,\ 0.95, \tag{4}$$

$$u_t = \rho u_{t-1} + w_t, \quad \rho = 0.0,\ 0.5,\ 0.95, \tag{5}$$

$$\varepsilon_t \sim i.i.d.\ N(5, 10)\,,\ w_t \sim i.i.d.\ N(0, 1)\,, \tag{6}$$

$$X_0 \sim\ N(5/(1-\varphi), 10/(1-\varphi^2)), \quad u_0 \sim\ N(0, 1/(1-\rho^2))\,. \tag{7}$$

Then, using Equation (3), we generate data for Y_t by assigning specific values to the parameters $l, \alpha, \beta_0, \beta_1, ..., \beta_l$. We consider six models, which emerge by assuming a maximum value of $l = 3$ and $\alpha = 10.0, \beta_0 = 0.6, \beta_1 = \beta_2 = 0.5,$ and $\beta_3 = 0.4$ in Equation (3).

For each sample size, $T = 50$ and $T = 200$, and each combination of the parameters φ and ρ, we generate *one* set of T "observations" for the ε's and 5000 sets for the w's, all from normal distributions, as indicated.[1] Then, we construct one set of "observations" for X_t, which we keep fixed in the 5000 replications; use Equation (3) to generate 5000 sets of T "observations" for Y_t; and use them to estimate 5000 times the null model, Equation (2). In each experiment, we apply POY(j) and LOY(m) and calculate the proportion of rejections using a 5% level of significance. This proportion estimates the power of the test. We also estimate 5000 times the *true* model, Equation (3), apply POY(j) and LOY(m), and calculate the proportion of rejections, thus estimating the *size* of these tests. Note the following observations.

First, in the presence of positive disturbance autocorrelation, POY(j) and LOY(m) tend to be oversized, especially when X_t is also positively autocorrelated, because the conventional standard errors are likely to underestimate the true ones, thus rendering the test variables spuriously significant; see, e.g., Johnston (1972, pp. 248-249) and Porter and Kashyap (1984). This can also occur because of a "spurious correlation" problem; see Leung and Yu (2001). Second, since we use a 5% nominal level of significance and the number of replications is 5000, the 95% confidence interval for the true percentage of rejections is (4.40, 5.60). (This is a standard confidence interval for the population proportion, using a sample size of 5000 observations and a sample proportion of $\hat{p} = 0.05$.) Estimated sizes that fall outside this interval are regarded as significantly different from the nominal size. Third, we have generated POY(2), POY(3), and POY(4), but report the results for POY(2) only, because the three variants behave similarly, and because LOY(m) naturally compares to POY(2), since it uses *second* powers of lagged values of $\hat{Y}_t$.

4. Results

Model 1: $k = 0$, $l = 1$. Using OLS and $T = 50$, we generate the rejection frequencies of POY(2) and LOY(1) and report them in Table 1:

Table 1. Size and power of POY(2) and LOY(1) generated by OLS with $T = 50$ for Model 1

		Size			Power		
φ \ ρ		0	0.50	0.95	0	0.50	0.95
0	POY(2)	5.46	5.24	5.44	1.76	0.90	1.02
	LOY(1)	4.84	4.62	1.50*	100.00	100.00	99.58
0.50	POY(2)	5.22	9.14*	9.60*	4.96	7.00	4.30
	LOY(1)	5.08	7.16*	23.88*	100.00	100.00	95.82
0.95	POY(2)	5.26	19.02*	31.58*	0.08	2.86	15.32
	LOY(1)	4.72	6.60*	20.52*	100.00	100.00	95.42

Notes: (a) The rejection frequencies are given in percentages; (b) size estimates that fall outside the interval (4.40, 5.60) are regarded as significantly different from the nominal size (5%) and are marked by a star (*).

[1] We use the LINUX version of RATS v. 5.01 to carry out the simulations. Random numbers were generated using the function *%RAN*(x) and the starting seed 317811. Note also that before we started drawing the values of ε and w, we let the process run for 500 "periods."

Although the results on size are mixed, those on power clearly support the use of LOY(1). More precisely, the power of POY(2) is less than its size in most of our experiments, which suggests that POY(2) is a biased test for this application, whereas the power of LOY(1) always exceeds 95%. This is not surprising. In this case, Equation (1) is $Y_t = \alpha + \beta_0 X_t + \gamma Z_t + u_t$, where $Z_t = \hat{Y}_t^2$ for POY(2) and $Z_t = \hat{Y}_{t-1}^2$ for LOY(1), and where $\hat{Y}_t$ is the fitted value from the null model, $Y_t = \alpha + \beta_0 X_t + v_t$. The variance estimator of the OLS coefficient $\hat{\gamma}$ is given by

$$S_{\hat{\gamma}}^2 = \frac{S^2}{(1 - r_{XZ}^2)\Sigma(Z_t - \bar{Z})^2}, \tag{8}$$

where S^2 is the residual variance from Equation (1) and r_{XZ} is the correlation coefficient between X_t and Z_t. Consider how the choice of Z_t affects $S_{\hat{\gamma}}^2$. First, S^2 is expected to be larger when $Z_t = \hat{Y}_t^2$ than when $Z_t = \hat{Y}_{t-1}^2$, since $\hat{Y}_{t-1}^2$ is a function of the omitted variable X_{t-1}, and should have more explanatory power in Equation (1) than does $\hat{Y}_t^2$, which is a function of the already included variable X_t. The average values of S^2 from the 5000 replications confirm this expectation. Second, it seems hard to say *a priori* how the choice of Z_t will affect $\Sigma(Z_t - \bar{Z})^2$, but the average values of this sum from the 5000 replications are smaller when $Z_t = \hat{Y}_t^2$ than when $Z_t = \hat{Y}_{t-1}^2$. Third, r_{XZ}^2 is expected to be higher when $Z_t = \hat{Y}_t^2$ than when $Z_t = \hat{Y}_{t-1}^2$, since $\hat{Y}_t^2 = (\hat{\alpha} + \hat{\beta}_0 X_t)^2$ is highly correlated with X_t, whereas $\hat{Y}_{t-1}^2$ may not be correlated with X_t, unless φ takes on a high value, e.g., $\varphi = 0.95$. The average values of r_{XZ}^2 are at least 0.9861 when $Z_t = \hat{Y}_t^2$, but range from 0.0155 to 0.9015 (increasing with φ) when $Z_t = \hat{Y}_{t-1}^2$. For these three reasons, the standard error of $\hat{\gamma}$ when $Z_t = \hat{Y}_t^2$ is at least ten times greater than its counterpart when $Z_t = \hat{Y}_{t-1}^2$, hence the big difference in empirical power.

As was expected from our earlier discussion, the two tests are oversized when *both* $\varphi \geq 0.50$ and $\rho \geq 0.50$. Thus, following Pagan and Hall (1983, pp. 206-209), we use the C-O method, which reduces the size distortion problem drastically: in only four (out of nine) experiments for each test the estimated size now falls outside the interval (4.40, 5.60), and then it only ranges from 4.14 to 7.28 for POY(2) and from 5.64 to 5.90 for LOY(1). Compare these ranges with those from OLS, 9.14 to 31.58 and 1.50 to 23.88, respectively. The C-O method also improves power: the power of POY(2) improves only at $\varphi = 0$, in which case it ranges from 36.66% to 55.28%, whereas that of LOY(1) is now 100% in every case.

For space considerations, we will not report the results for $T = 200$. They differ from the case of $T = 50$ only in that (1) the power of LOY(1) is now 100% in every case, and (2) when the C-O method is used, only two size estimates fall outside the interval (4.40, 5.60), namely 6.24 and 6.58 for POY(2) and 5.68 and 5.70 for LOY(1).

Note that we have also applied LOY(2) and LOY(3). Compared to LOY(1), the only notable difference is that their size is generally greater, except when the C-O method is used with $T = 200$, in which case *all* size estimates fall within the interval (4.40, 5.60).

Model 2: $k = 0$, $l = 2$. We use again POY(2) and LOY(1). Overall, considering both size and power, POY(2) behaves worse and LOY(1) behaves better than in Model 1.

Model 3: $k = 1$, $l = 2$. In this case, we compare POY(2) with LOY(2). The conclusions are similar to those obtained in Models 1 and 2, when we compared POY(2) with LOY(1). It is worth noting that LOY(1) behaves better than LOY(2) in this case.

Model 4: $k = 0$, $l = 3$. Here, we compare POY(2) with LOY(1). The only notable difference from the previous cases is that when $T = 50$ and $\varphi \leq 0.50$, the power of LOY(1) generated by the C-O method ranges only from 71% to 89%. Note that LOY(2) is better in this case.

Model 5: $k = 1$, $l = 3$. We use POY(2) and LOY(2). The results resemble those of Model 3.

Model 6: $k = 2$, $l = 3$. We use POY(2) and LOY(3). The pattern of the results we have seen so far does not change. Note, however, that when $T = 50$, the power of LOY(3) generated by OLS is only 83% in two (out of nine) experiments. LOY(1) and LOY(2) are better.

5. Conclusion

This paper proposes a new variant of RESET that is appropriate for distributed lag models. Monte Carlo evidence on size and power suggests that the traditional RESET is a biased test in the present setup, whereas the new variant has good size properties, provided that it is generated by an autocorrelation robust method, and high power to detect the erroneous omission of lagged values of an explanatory variable.

REFERENCES

Johnston, J. (1972) *Econometric methods*, 2nd Ed., McGraw-Hill: Tokyo.

Leung, S.F., and S. Yu (2001) "The sensitivity of the RESET tests to disturbance autocorrelation in regression residuals" *Empirical Economics* **26**, 721-726.

Pagan, A.R., and A.D. Hall (1983) "Diagnostic tests as residual analysis" *Econometric Reviews* **2**, 159-218.

Porter, R.D., and A.K. Kashyap (1984) "Autocorrelation and the sensitivity of RESET" *Economics Letters* **14**, 229-233.

Ramsey, J.B. (1969) "Tests for specification errors in classical linear least-squares regression analysis" *Journal of the Royal Statistical Society,* **Series B 31**, 350-371.

Thursby, J.G. (1989) "A comparison of several specification error tests for a general alternative" *International Economic Review* **30**, 217-230.