Gauss-Newton, Milliken-Graybill, and Exact Misspecification Testing Using Artificial Regressions

Kenneth G. Stewart
Assistant Professor

Department of Economics
University of Victoria
Victoria, British Columbia
Canada V8W 2Y2

Telephone: (250) 721-8534
Fax: (250) 721-6214
e-mail: kstewart@uvvm.uvic.ca

November 27, 1998

Abstract

The Gauss-Newton regression (GNR) is widely used to compute Lagrange multiplier statistics. A regression described by Milliken and Graybill yields an exact F test in a certain class of nonlinear models which are linear under the null. This paper shows that the Milliken-Graybill regression is a GNR. Hence one interpretation of Milliken-Graybill is that they identified a class of nonlinear models for which the GNR yields an exact test.

1 Introduction

Artificial regressions provide a convenient means of computing Lagrange multiplier (LM) statistics. The large literature developing the relevant econometric theory has been definitively unified and consolidated by MacKinnon (1992) and Davidson and MacKinnon (1990, 1993). Since LM statistics are based on the estimation of the restricted form of the model they tend to be particularly heavily used for misspecification testing, where the restricted model is well defined and often linear. Hence, as emphasized by MacKinnon (1992), the artificial regressions approach to computing LM tests provides a means of unifying much of the methodology of misspecification testing.

The distribution theory associated with LM tests is, of course, asymptotic, and the nature of many types of misspecification is such that this is the best that can be hoped for. In a few instances, however, test statistics having a known finite-sample distribution have been derived. As one would expect, these obtain under classical circumstances of normality, nonstochastic regressors, and a model linear in its coefficients.

The trivial example is the case of a regression

$$y = X_1\beta_1 + X_2\beta_2 + u, \qquad u \sim \mathrm{NID}(0, \sigma^2 I),$$

in which the role of the variables X_2 is in question. The null $H_0 : \beta_2 = 0$ is the hypothesis that the simpler model $y = X_1\beta_1 + u$ is correctly specified; the alternative is that it is misspecified.

To cite a particular application, consider the problem of testing for structural change. The null hypothesis is of parameter constancy, the alternative that a structural change has taken place. In this case X_2 consists of appropriately defined dummy variable terms, β_2 has the interpretation as a vector of "shift coefficients," and the restriction $\beta_2 = 0$ corresponds to the hypothesis of an absence of misspecification.

Although one could compute the LM test of this restriction, by an artificial regression or some other means, this is obviously neither the simplest nor the preferred procedure. The natural choice is instead an F statistic which, in the context of this testing problem, takes the form

$$\frac{y'M_1X_2(X_2'M_1X_2)^{-1}X_2'M_1y}{y'M_{M_1X_2}M_1y} \cdot \frac{n-k}{r} \sim F_{r,n-k}. \tag{1}$$

The notation is $M_1 \equiv I - X_1(X_1'X_1)^{-1}X_1'$ and similarly for $M_{M_1X_2}$ in terms of M_1X_2; n

is the sample size, k is the number of coefficients in the unrestricted model, and r is the number of restrictions (in this case, the number of coefficients in the vector β_2). In the context of the structural change application the statistic (1) is, of course, the well known Chow test statistic.

More interesting examples of exact distributional results applying to misspecification tests include the Ramsey-Schmidt (1976) version of the RESET test, the J_A test for non-nested models proposed by Fisher and McAleer (1981),[1] and the Andrews (1971) test for linear versus loglinear regression models;[2] McAleer (1987) usefully summarizes the underlying statistical basis for an exact distributional result applying to the latter two. All turn out to be applications of Milliken and Graybill (1970), who showed that an exact F test applies to a particular set of zero-restrictions in a certain class of nonlinear regression models. The Milliken-Graybill F statistic is obtained from a regression which is here termed the Milliken-Graybill regression (MGR), and which is described in Section 4 below.

This paper considers the best-known artificial regression—the Gauss-Newton regression (GNR)—and investigates the availability of exact tests as an alternative to the LM tests that would normally be yielded by it. It is shown that, in terms of the test statistic generated, the Milliken-Graybill regression effectively *is* a Gauss-Newton regression. Hence the circumstances under which the MGR applies are those in which an exact test is available as an alternative to the LR test that would normally be computed from the GNR. Indeed, one interpretation of the Milliken-Graybill result is that they identified a class of models for which the GNR yields an exact test, rather than one having only an asymptotic justification.

Although, as these introductory remarks have suggested, misspecification testing is a particularly fruitful area of application and so serves to motivate the discussion, the results are not limited to this context. The original Milliken-Graybill work was illustrated with applications to testing interaction terms in analysis of variance models, and to nonlinear regression.

[1] See also Fisher (1983), Godfrey (1983), and McAleer (1983).

[2] Related papers are Godfrey and Wickens (1981), Bera and McAleer (1983), and Godfrey, McAleer, and McKenzie (1988).

2 The Gauss-Newton Regression

In developing the Gauss-Newton regression it is convenient to adopt the notation of Davidson and MacKinnon (1993), who specify the nonlinear regression model as

$$y = x(\beta) + u. \tag{2}$$

The most common approach to testing restrictions on β is to apply one of the three classical test citeria: the Wald, likelihood ratio, and Lagrange multiplier principles.

Focusing on the latter, the calculation of the LM statistic is facilitated with an artifical regression. The Gauss-Newton regression is of the form

$$y - x(\beta^*) = X(\beta^*)b + \text{disturbance},$$

where $X(\beta) \equiv Dx(\beta)$ denotes the matrix of derivatives of $x(\beta)$ with respect to β, and β^* denotes the parameter vector at which the GNR is evaluated. Although the GNR may be used for a variety of purposes, depending on the choice of β^*, our interest lies in the computation of LM test statistics. Without loss of generality, the model may be reparameterized so that arbitrary restrictions on β may be expressed as zero restrictions on a subvector β_2. That is, the model (2) becomes

$$y = x(\beta_1, \beta_2) + u. \tag{3}$$

Let us denote the restricted nonlinear least squares estimator by $\tilde{\beta} \equiv [\tilde{\beta}_1; 0]$. Partitioning the matrix of derivatives $X(\beta)$ into $X(\beta_1)$ and $X(\beta_2)$ defined by

$$X(\beta_1) = \frac{\partial x(\beta)}{\partial \beta_1} \tag{4a}$$

$$X(\beta_2) = \frac{\partial x(\beta)}{\partial \beta_2}, \tag{4b}$$

the GNR evaluated at $\tilde{\beta}$ is

$$y - \tilde{x} = \tilde{X}_1 b_1 + \tilde{X}_2 b_2 + \text{disturbance}. \tag{5}$$

In terms of notation, the matrices of explanatory variables are, respectively, (4a) and (4b) evaluated at $\tilde{\beta}$. Note that the dependent variable is the vector of residuals arising from the estimation of (3) under the null. This GNR gives rise to a family of LM statistics described by Davidson and MacKinnon (1993, sec. 6.4). To paraphrase MacKinnon (1992, p. 109), "The fundamental result for tests based on the GNR is that any asymptotically valid test of $b_2 = 0$ in (5) also provides an asymptotically valid test of $\beta_2 = 0$ in (3)."

In terms of the small sample behavior of alternative statistics arising from the GNR, Davidson and MacKinnon recommend as an alternative to an LM statistic the closely related F statistic for $b_2 = 0$ in (5).[3] Applying (1), the formula for this statistic is

$$\frac{(y-\tilde{x})'\tilde{M}_1\tilde{X}_2(\tilde{X}_2'\tilde{M}_1\tilde{X}_2)^{-1}\tilde{X}_2'\tilde{M}_1(y-\tilde{x})}{(y-\tilde{x})'\tilde{M}_{\tilde{M}_1\tilde{X}_2}\tilde{M}_1(y-\tilde{x})}\cdot\frac{n-k}{r}, \tag{6}$$

where $\tilde{M}_1 \equiv I - \tilde{X}_1(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'$ and similarly for $\tilde{M}_{\tilde{M}_1\tilde{X}_2}$ in terms of $\tilde{M}_1\tilde{X}_2$. The available small sample evidence suggests, then, that the distribution of this statistic is well approximated by $F_{r,n-k}$.

3 Application to the Milliken-Graybill Model

The above F statistic is a version of the LM test—one which appears to have appealing small-sample behavior—for arbitrary restrictions on the general nonlinear model (2). As an illustration of its application, let us consider a particular class of nonlinear models studied by Milliken and Graybill (1970). This class is of the form

$$y = X\beta_1 + F(X\beta_1)\beta_2 + u. \tag{7}$$

The matrix $F(X\beta_1)$ is some nonlinear function of $X\beta_1$; this is the sole source of nonlinearity in the coefficients. The restriction of interest is $\beta_2 = 0$. Since under this null the model is linear, the LM test afforded by the GNR (5) is of natural interest. In terms of notation the matrices of derivatives (4) are

$$X(\beta_1) = \frac{\partial[X\beta_1 + F(X\beta_1)\beta_2]}{\partial\beta_1} = X + \frac{\partial F(X\beta_1)\beta_2}{\partial\beta_1}$$
$$X(\beta_2) = \frac{\partial[X\beta_1 + F(X\beta_1)\beta_2]}{\partial\beta_2} = F(X\beta_1)$$

Evaluating at the restricted estimates $\tilde{\beta} = [\tilde{\beta}_1; 0]$, these are

$$\tilde{X}_1 = X \qquad \tilde{X}_2 = F(X\tilde{\beta}_1),$$

yielding the GNR

$$y - X\tilde{\beta}_1 = Xb_1 + \tilde{F}b_2 + \text{disturbance}, \tag{8}$$

[3]Papers considering aspects of the small-sample behavior of alternate versions of the LM test in various contexts include Davidson and MacKinnon (1983), Kiviet (1986), and Bera and McKenzie (1986). In assessing this evidence Davidson and MacKinnon (1993, p. 190) conclude: "Based partly on theory and evidence, then, and partly on the convenience of using the same form of test for Gauss-Newton regressions as would normally be used with genuine regressions, we therefore recommend using the F test ... "

where $\tilde{F} \equiv F(X\tilde{\beta}_1)$. The restricted estimates are, of course, simply

$$\tilde{\beta}_1 = (X'X)^{-1}X'y, \tag{9}$$

and the dependent variable is the associated OLS residual vector

$$y - X\tilde{\beta}_1 = y - X(X'X)^{-1}X'y = [I - X(X'X)^{-1}X']y = M_X y. \tag{10}$$

Applying (6), the preferred test statistic is

$$\frac{(y - X\tilde{\beta}_1)'M_X\tilde{F}(\tilde{F}'M_X\tilde{F})^{-1}\tilde{F}'M_X(y - X\tilde{\beta}_1)}{(y - X\tilde{\beta}_1)'\tilde{M}_{M_X\tilde{F}}M_X(y - X\tilde{\beta}_1)} \cdot \frac{n-k}{r}. \tag{11}$$

As indicated by (10), in these expressions $M_X \equiv I - X(X'X)^{-1}X'$, and similarly for $M_{M_X\tilde{F}}$ in terms of $M_X\tilde{F}$.

4 The Milliken-Graybill Regression

Milliken-Graybill's motivation for studying the class of models (7) was an interest in examining the scope for exact testing in nonlinear models. As a means of testing $\beta_2 = 0$ they proposed the following procedure.

Step 1 Estimate (7) under the null. This yields $\tilde{\beta}_1$ given by (9).

Step 2 Computing $\tilde{F} \equiv F(X\tilde{\beta}_1)$, estimate the original model (7), replacing $F(X\beta_1)$ with $\tilde{F}$. Let us denote this linear regression by

$$y = Xb_1 + \tilde{F}b_2 + \text{disturbance}, \tag{12}$$

which we term the Milliken-Graybill regression (MGR).

Milliken and Graybill's contribution was to show that a standard F test of $b_2 = 0$ in the MGR is an exact test of $\beta_2 = 0$ in the original model (7). That is, applying (1), the statistic

$$\frac{y'M_X\tilde{F}(\tilde{F}'M_X\tilde{F})^{-1}\tilde{F}'M_Xy}{y'\tilde{M}_{M_X\tilde{F}}M_Xy} \cdot \frac{n-k}{r} \tag{13}$$

is distributed as $F_{r,n-k}$ in finite samples. Intuitively, the reason for this is that since $X\tilde{\beta}_1 = X(X'X)^{-1}X'y$ operates on y so as to project it onto the space spanned by the columns of X, the columns of $\tilde{F}$ may be treated as fixed regressors; this is why it is essential that $\tilde{\beta}_1$ enter the argument of $F(\cdot)$ only via $X\tilde{\beta}_1$.

From the above development it is apparent that, although not previously recognized in these terms—and certainly not motivated in this way by Milliken and Graybill—the MGR (12) is tantalizingly similar to the GNR (8). The only difference is the dependent variable, which in the GNR is OLS residual vector under the null, but in the MGR is the original dependent variable. For the purpose of computing the test statistics (11) and (13), however, this difference is irrelevant.

Result. *The test statistics* (11) *and* (13) *are numerically identical.*

Proof. Heuristically, the inclusion of X on the right hand side of the regressions means that the use of the dependent variable $y - X\tilde{\beta}_1 = M_X y$ in the GNR versus simply y in the MGR has no effect on the F statistic for $b_2 = 0$. More rigorously, using (10) the numerator of (11) is, by the symmetry and idempotence of M_X,

$$y' M_X' M_X \tilde{F}(\tilde{F}' M_X \tilde{F})^{-1} \tilde{F}' M_X M_X y = y' M_X \tilde{F}(\tilde{F}' M_X \tilde{F})^{-1} \tilde{F}' M_X y,$$

which is the numerator of (13). Similarly the denominator is

$$\begin{aligned} y' M_X' \tilde{M}_{M_X \tilde{F}} M_X M_X y &= y' M_X' [I - M_X \tilde{F}(\tilde{F}' M_X \tilde{F})^{-1} \tilde{F}' M_X] M_X y \\ &= y' [M_X - M_X \tilde{F}(\tilde{F}' M_X \tilde{F})^{-1} \tilde{F}' M_X] y \\ &= y' \tilde{M}_{M_X \tilde{F}} M_X y. \end{aligned}$$

■

5 Conclusion

This development establishes that one interpretation of Milliken and Graybill's analysis—perhaps the most useful interpretation—is that they identified a class of nonlinear models for which the GNR yields an exact test. This provides, incidentally, an analytical foundation for the empirical finding that the F version of the GNR LM test has the best small sample properties: for a particular class of models—one that has had some prominence in applied work—it is in fact exactly F distributed in finite samples. Finally, whereas the Milliken-Graybill procedure has tended to be seen as something of an ad hoc curiosity, it is now apparent that it is directly linked to the systematic methodology of artifical regressions.

References

Andrews, D.F. (1971) A note on the selection of data transformations. *Biometrika* 58, 249–254.

Bera, A.K. & M. McAleer (1983) Some exact tests for model specification. *Review of Economics and Statistics* 65, 351–354.

Bera, A.K. & C.R. McKenzie (1986) Alternative forms and properties of the score test. *Journal of Applied Statistics* 13, 13–25.

Fisher, G.R. (1983) Tests for two separate regressions. *Journal of Econometrics* 21, 117–132.

Fisher, G.R. & M. McAleer (1981) Alternative procedures and associated tests of significance for non-nested hypotheses. *Journal of Econometrics* 16, 103–119.

Davidson, R. & J.G. MacKinnon (1983) Small sample properties of alternative forms of the Lagrange multiplier test. *Economics Letters* 12, 269–275.

Davidson, R. & J.G. MacKinnon (1990) Specification tests based on artificial regressions. *Journal of the American Statistical Association* 85, 220–227.

Davidson, R. & J.G. MacKinnon (1993) *Estimation and Inference in Econometrics.* New York: Oxford University Press.

Godfrey, L.G. (1983) Testing non-nested models after estimation by instrumental variables or least squares. *Econometrica* 51, 355–365.

Godfrey, L.G., McAleer, M. & C.R. McKenzie (1988) Variable addition and Lagrange multiplier tests for linear and logarithmic regression models. *Review of Economics and Statistics* 70, 492–503.

Godfrey, L.G. & M.R. Wickens (1981) Testing linear and log-linear regressions for functional form. *Review of Economic Studies* 48, 487–496.

Kiviet, J.F. (1986) On the rigour of some misspecification tests for modelling dynamic relationships. *Review of Economic Studies* 53, 241–261.

MacKinnon, J.G. (1992) Model specification tests and artificial regressions. *Journal of Economic Literature* 30, 102–146.

McAleer, M. (1983) Exact tests of a model against nonnested alternatives. *Biometrika* 70, 285–288.

McAleer, M. (1987) Specification tests for separate models: a survey. In M.L. King and D.E.A. Giles (eds.), *Specification Analysis in the Linear Model,* pp. 146–195. London: Routledge and Kegan Paul.

Milliken, G.A. & F.A. Graybill (1970) Extensions of the general linear hypothesis model. *Journal of the American Statistical Association* 65, 797–807.

Ramsey, J.B. & P. Schmidt (1976) Some further results on the use of OLS and BLUS residuals in specification error tests. *Journal of the American Statistical Association* 71, 389–390.