CONDITIONAL MOMENT TESTS WITH EXPLICIT ALTERNATIVES

A. Colin Cameron
and
Pravin K. Trivedi

Working Paper Series
#366

Department of Economics
University of California
Davis, California 95616-8578

CONDITIONAL MOMENT TESTS WITH EXPLICIT ALTERNATIVES

by

A. Colin Cameron
Department of Economics
University of California
Davis, California 95616-8578

Pravin K. Trivedi
Department of Economics
Indiana University
Bloomington, Indiana 47405

Working Paper Series No. 366
September 1990

ABSTRACT

The conditional moment (CM) tests of Newey (1985) and Tauchen (1985) were developed as diagnostic tests of whether or not a moment condition holds. In this paper CM tests against an explicit alternative are presented. In many testing situations this is the natural way to proceed in constructing tests, mirroring the procedure used for classical statistical tests. Furthermore, many diagnostic tests such as the information matrix test, apparently without an alternative, can be interpreted as tests against an explicit alternative. The CM tests considered here are motivated by a regression, and are accordingly called regression-based CM tests. Testing can be based directly on this regression, when the moment fundamental to the test satisfies an additional moment condition. An example where this condition holds yields new tests for misspecification of the central moments. New CM tests based on orthogonal polynomials are also presented.

Some Key Words: conditional moment specification tests; LM tests; score tests; information matrix tests; orthogonal polynomials; generalized linear models; heteroskedasticity; symmetry; kurtosis.

1. INTRODUCTION

Conditional moment (CM) tests, introduced by Newey (1985) and Tauchen (1985), offer a unifying framework for tests of parametric model misspecification, and are often simple to implement.

The richness of this approach is demonstrated in a recent book by White (1990). He draws in to this framework the classical tests - the Wald, Likelihood Ratio and Lagrange Multiplier (or Score) test principles - and non-classical tests - Hausman, Information Matrix, and Encompassing Tests. He obtains the distributions of moment test statistics under minimal assumptions about the data generating process, and proposes auxiliary regressions to compute the test statistics. White advocates a methodology of specification analysis based on this theory. Pagan and Vella (1989) emphasize the simplicity of CM testing and very strongly urge its adoption.

Nonetheless, specification testing based on general moment conditions has been slow to take hold. There are several reasons for this. First, there is a perception that the newer moment tests, i.e. those not based on the three classical test procedures, do not involve the formulation of an alternative hypothesis model beyond stating that the null does not hold, and therefore are more difficult to intuitively interpret. Second, the wide range of moment conditions on which to base CM tests leads to aversion of seemingly ad-hoc choice of moment conditions and reversion to moment tests based on classical test principles whose properties are well known. Finally, in most applications of CM tests, a density is actually specified under the null (e.g. any information matrix test requires this), and a generalization of the information matrix equality is used to justify computation of test statistics by an auxiliary regression, called the outer-product-of-the-gradient regression by Godfrey (1988) and Davidson and MacKinnon (1990), that has very poor size properties.

In this paper we propose an approach to CM testing, the regression-based CM test, that overcomes these objections. A specified conditional moment, called the "fundamental" moment, equals zero under the null hypothesis. The fundamental moment is then embedded in a more general alternative: under the alternative hypothesis the fundamental moment equals a specified function of parameters and exogenous variables. The null hypothesis is tested by performing a significance test on the coefficients from the regression of the fundamental moment on the specified function of parameters and exogenous variables. But this significance test is itself a test of a moment condition, called a "regression-based" CM test. It generalizes the work of Cameron and Trivedi (1990a), who used this test in the context of testing variance-mean equality, a property of the Poisson model.

Like any CM test, implementation of a regression-based CM test requires replacing unobserved parameters by estimates of the parameters of the null hypothesis model. In a number of leading examples, asymptotic results are unchanged by this replacement, so that the significance tests mentioned in the previous paragraph can be performed without modification. This permits direct tests for the significance of subcomponents of the alternative hypothesis. Thus the regression is more than just an auxiliary regression.

In other examples the replacement of parameters by estimates affects the asymptotic distribution. We can then appeal to the general theory of CM tests. Different estimators and different moment conditions lead to different limit distributions of the CM test statistic. The corresponding chi-square test statistics can then be computed by one or more different auxiliary regressions. White (1990) gives a very general treatment, that offers alternatives to using the outer-product-of-the-gradient regression in many common testing situations. Wooldridge (1990a) has developed a procedure to transform the CM test statistic to one, not necessarily equivalent, for which asymptotic theory is valid under relatively weak distributional assumptions.

The regression-based CM test approach decomposes CM tests into two components: the fundamental moment (zero under the null hypothesis) and its parameterization under the alternative hypothesis. The choice of fundamental moment is considered according to whether or not the conditional distribution of the dependent variable is fully parameterized.

If distributional assumptions are to be minimized, it seems natural to let the fundamental moment be the expectation of a function of the central moments of the dependent variable, or equivalently of the error term. In particular, we propose new tests of misspecification of the conditional central moments of the dependent variable, i.e. heteroskedasticity, skewness, and kurtosis.

When the conditional density under the null hypothesis is specified, standard specification tests are the lagrange multiplier (LM), Hausman (H) and information matrix (IM) tests. These tests can be interpreted as regression-based CM tests. For the LM and H tests, the fundamental moment condition is the expectation of a subcomponent of the score vector of the null hypothesis model, while for the IM test the fundamental moment is the expectation of a subcomponent of the sum of the outer product and the derivative of the score vector.

For the LM test, interpretation as a regression-based CM test should not be surprising, since there is usually a moment restriction whose imposition gives the null hypothesis density as a special case of the alternative hypothesis density. However, the regression-based CM test approach is more direct, less parametric, and often considerably easier analytically than the LM test approach. Furthermore, the two approaches frequently give asymptotically equivalent test statistics. To the extent that this happens, the regression based CM test is asymptotically equivalent not only to the LM test, but also to the Wald and Likelihood Ratio tests.

For the IM and H tests, equivalence to a regression-based CM test permits interpretation of the IM and H tests as being a test of a conditional moment against a specific alternative. This re-interpretation of these tests overcomes one of their perceived weaknesses. It also suggests a wider class of IM tests than generally used.

When the conditional density is specified, we also present a new choice of fundamental moment condition, that based on orthogonal polynomial systems. This new theory is related to LM and IM tests in the case of quasi-maximum likelihood estimation for the linear exponential family with quadratic variance function.

The general theory of regression-based CM tests is given in section 2, and contrasted with the standard formulation of CM tests. Direct regressions and auxiliary regressions to implement regression-based CM tests are presented in section 3. Examples are given in section 4. These cover many common testing situations, are easily motivated and implemented, and include some new tests. Distributional assumptions are minimized in sections 2-4. In section 5, the regression-based CM test approach is compared with other principles used to obtain CM tests, including Lagrange multiplier, Hausman and information matrix tests, for models where the conditional density under the null hypothesis is specified. A new class of CM tests, that based on orthogonal polynomial systems is also presented. An application to the linear exponential family with quadratic variance function is given in section 6. Some concluding remarks are made in section 7.

2. REGRESSION-BASED TESTS FOR MODEL SPECIFICATION

2.1 *Conditional Moment Tests*

In regression analysis, we are interested in explaining dependent variables, a vector y_t, conditional on explanatory variables, a vector X_t. For simplicity, this paper focuses on cross-section data, where the data $\{(y_t, X_t), t = 1,\ldots,T\}$ are independent across t. The approach can be extended to dynamic models, where X_t is the vector of current and lagged values of the exogenous variables and lagged values of the dependent variables, and (y_t,X_t) are no longer independent across t.

The true data generating process (d.g.p.) for y given X is unknown. Instead, statistical inference is based on an assumed parameterized density function (quasi-ML estimation) or an assumed parameterization of moments (GMM estimation, e.g. least squares). Conditional moment tests are tests of the validity of moment conditions implied by these assumed parameterizations.

Specifically, a conditional moment test is any test based on an $s\times1$ vector of functions $m(y,X,\theta)$, where θ is a $q\times1$ vector of parameters, that satisfies the moment condition:

$$\text{(2.1)} \qquad \mathbb{E}_0[m(y_t,X_t,\theta) \mid X_t] = 0 \quad ,$$

where the subscript 0 denotes expectation with respect to an assumed model, not necessarily the true d.g.p.

Tests based on a moment condition of the form (2.1), henceforth called CM tests, were introduced by Newey (1985) and Tauchen (1985). Further results are given by Pagan and Vella (1989), White (1987, 1990) and Wooldridge (1990a). The simplest version of a CM test based on (2.1) uses the corresponding sample moment:

$$(2.2) \qquad m_T(\theta) = T^{-1}\sum_{t=1}^{T} m(y_t,X_t,\theta) \quad .$$

To operationalize a CM test, the parameter θ in (2.2) is replaced by an estimator $\hat{\theta}_T$, consistent under the maintained model. CM specification tests are statistical tests of the departure of $m_T(\hat{\theta}_T)$ from zero.

The concern of this paper is CM testing. This has two dimensions: the choice of moment to use in (2.2), and the effect of replacing θ in (2.2) by an estimator. We begin with the first of these.

2.2 *Regression-Based Conditional Moment Tests*

It is assumed that the process generating the data is such that the $n\text{x}1$ vector of functions $r(y,X,\theta)$, satisfies the "fundamental moment condition":

$$(2.3) \qquad H_0\text{: } \mathbb{E}_0[r(y_t,X_t,\theta) \mid X_t] = 0 \quad ,$$

where the subscript 0 denotes expectation with respect to the null hypothesis model. This moment condition is "fundamental" in the sense that the dimension of $r(\cdot)$ in (2.3) is generally considerably less than the dimension of $m(\cdot)$ in (2.1), as should be clear from the ensuing discussion. In fact, $n = 1$ in most commonly used tests, but for completeness we give the theory for the more general case. Sections 4 and 5 will discuss at some length the choice of fundamental moment condition.

Suppose that were the expectation in (2.3) to be taken with respect to the true distribution, (2.3) would no longer hold. In particular, we wish to test against the alternative that for the j-th moment condition:

$$H_1\text{: } \mathbb{E}_1[r_j(y_t,X_t,\theta) \mid X_t] = g_j(X_t,\theta)\cdot\alpha_j \quad , \qquad j = 1,\ldots,n,$$

where the subscript 1 denotes expectation with respect to the true d.g.p., $g_j(X_t,\theta)$ = is a specified $1xp_j$ vector function, and α_j is a p_jx1 vector of additional unknown parameters. Continuing the earlier example, the alternative moment condition may be that the error has non-zero mean due to omitted variables $g(X_t,\theta)$. Combining all n moment conditions we test H_0 against the alternative that:

$$(2.4) \qquad H_1: \mathbb{E}_1[r(y_t,X_t,\theta) \mid X_t] = G(X_t,\theta)\cdot\alpha \quad ,$$

where $G(X_t,\theta)$ is a nxp matrix whose j-th row has g_j in columns $(p_1+ .. + p_{j-1}+ 1)$ to $(p_1+ .. + p_j)$ and zeroes elsewhere, $p = (p_1+ .. + p_n)$, and $\alpha = (\alpha_1', .. ,\alpha_n')'$ is a px1 parameter vector.

Tests of the moment condition under H_0 against that under H_1 are tests of whether $\alpha = 0$. The obvious basis for such a test is the estimated coefficient of α in the following multivariate regression:

$$(2.5) \qquad r(y_t,X_t,\theta) = G(X_t,\theta)\cdot\alpha + \varepsilon_t,$$

where the sx1 heteroscedastic error term ε_t is defined by

$$(2.6) \qquad \varepsilon_t = r(y_t,X_t,\theta) - E[r(y_t,X_t,\theta) \mid X_t] \quad .$$

Weighted least squares estimation of (2.5) with sxs symmetric weighting matrix $W(X_t,\theta)$ yields the usual weighted least squares estimator:

$$(2.7) \quad \hat{\alpha} = \{ \sum_{t=1}^{T} G(X_t,\theta)'\cdot W(X_t,\theta)\cdot G(X_t,\theta) \}^{-1}\cdot \sum_{t=1}^{T} G(X_t,\theta)'\cdot W(X_t,\theta)\cdot r(y_t,X_t,\theta) \ .$$

For the purposes of statistical inference, tests based on $\hat{\alpha}$ are determined by the distribution of:

$$(2.8) \qquad m_{\alpha T}(\theta) = \sum_{t=1}^{T} G(X_t,\theta)' \cdot W(X_t,\theta) \cdot r(y_t,X_t,\theta) \quad .$$

From section 2.1, this is a CM test based on the moment condition:

$$(2.9) \qquad \mathbb{E}_0[G(X_t,\theta)' \cdot W(X_t,\theta) \cdot r(y_t,X_t,\theta) \mid X_t] = 0 \quad .$$

This test is called a "regression-based" CM test, or RBCM test, since this form of the CM test is motivated by the regression of the fundamental moment on to its parameterization under the alternative. This should not be confused with the implementation of a CM test by an auxiliary regression. It is in this latter sense that other authors have used, or perhaps misused, the term "regression-based" in the context of testing. In section 3 we show that sometimes the RBCM test can be implemented directly from the regression (2.5), while at other times an auxiliary regression may be necessary.

RBCM tests were introduced by Cameron and Trivedi (1990a), in the context of testing variance-mean equality. Then in (2.5), y_t is scalar, $r(y_t,X_t,\theta) = \{(y_t - \mu_t(X_t,\theta))^2 - y_t\}$, where $\mu_t(X_t,\theta) = \mathbb{E}_0[y_t \mid X_t]$, and $G(X_t,\theta) = g(\mu_t(X_t,\theta))$ for specified scalar function $g(\cdot)$.

2.3 *Optimal Regression-Based CM Test*

Different choices of the weighting matrix lead to different test statistics. The optimal test within the class of RBCM tests will be that based on the most efficient estimator of α. This is the generalized least squares estimator:

$$(2.10) \qquad m_{\alpha T,\mathrm{opt}}(\theta) \;=\; \sum_{t=1}^{T} G(X_t,\theta)'\cdot\Sigma(X_t,\theta)^{-1}\cdot r(y_t,X_t,\theta) \quad ,$$

where we make the additional assumption:

$$\begin{aligned} (2.11) \qquad & \mathbb{E}_0[\varepsilon_t\varepsilon_t' \mid X_t] \\ & = \mathbb{E}_0[r(y_t,X_t,\theta)\cdot r(y_t,X_t,\theta)' \mid X_t] \\ & = \Sigma(X_t,\theta) \quad , \end{aligned}$$

for specified variance function $\Sigma_t = \Sigma(X_t,\theta)$. So the optimal RBCM test for (2.3) against alternatives of the form (2.4) is a CM test based on:

$$(2.12) \qquad \mathbb{E}_0[G(X_t,\theta)'\cdot\Sigma(X_t,\theta)^{-1}\cdot r(y_t,X_t,\theta) \mid X_t] \;=\; 0.$$

Specification of Σ_t usually requires distributional assumptions for the null hypothesis model additional to the minimal assumptions needed for (2.3). For example, in tests of omitted regressors in the classical regression model, the usual additional assumption is constancy of the error variance. Such additional distributional assumptions are not critical, since it is possible to construct tests that are asymptotically valid even if Σ_t is not fully specified..

2.4 *Discussion*

The RBCM tests differ from conventional CM specification tests in that they are tests against an explicit alternative, given in (2.4). In this subsection the two approaches are compared.

It follows immediately from section 2.3 that any CM test based on the moment condition

$$(2.13) \qquad \mathbb{E}_0[G^*(X_t,\theta)' \cdot r^*(y_t,X_t,\theta) \mid X_t] = 0 \quad ,$$

where G^* is a $n^* x p^*$ matrix and r^* is a $n^* x1$ vector, is a RBCM test of:

$$(2.14) \qquad H_0\text{:}\ \mathbb{E}_0[r^*(y_t,X_t,\theta) \mid X_t] = 0 \quad ,$$

against the alternative hypothesis

$$(2.15) \qquad H_1\text{:}\ \mathbb{E}_1[r^*(y_t,X_t,\theta) \mid X_t] = (W^*(X_t,\theta))^{-1} \cdot G^*(X_t,\theta) \cdot \alpha^* \quad ,$$

where $W^*(X_t,\theta)$ is the $n^* x n^*$ weighting matrix used in the regression, and α^* is a $n^* x1$ parameter vector. (2.13) corresponds to the optimal RBCM test of (2.14) against (2.15) when $W^*(X_t,\theta) = \mathbb{E}_0[r_t^* \cdot r_t^{*\prime} \mid X_t]$.

Note that (2.13) is exactly of the form of the CM test of Newey (1985). In Newey's framework $G^*(X_t,\theta)$ is an arbitrarily chosen matrix of functions, and $r^*(y_t,X_t,\theta)$ satisfies (2.14). The motivation is that if (2.14) holds, then by the law of iterated expectations, (2.13) holds for a wide range of choices of $G^*(X_t,\theta)$. Little guidance is given for the choice of $G^*(X_t,\theta)$, except in the case where the distribution of y_t is specified under both H_0 and H_1, and the score under H_0 is multiplicative in $r^*(y_t,X_t,\theta)$. In this case, Newey gives the optimal choice of $G^*(X_t,\theta)$. By contrast, the optimal RBCM test interpretation imposes less structure, and provides a direct approach to choosing the optimal $G^*(X_t,\theta)$, i.e. specify the variance of $r^*(y_t,X_t,\theta)$ under H_0 and the mean of $r^*(y_t,X_t,\theta)$ under H_1.

Pagan and Vella (1989) also use the form (2.13). Tauchen (1985) considers the simplest case where $G^*(X_t,\theta) = 1$. White (1987, 1990) and Wooldridge (1990a) call r_t^* "generalized residuals" and G_t^* "misspecification indicators". The terminology generalized residual is used following Cox and Snell (1968), though the examples below suggest that a better terminology might be functions of the residual. The RBCM framework provides a natural way to select, or to interpret, the misspecification indicators. The choice of generalized residual, or equivalently of fundamental moment, is still an open issue.

Wooldridge (1990a) factorizes (2.13) further to:

$$(2.16) \qquad \mathbb{E}_0[G^*(X_t,\theta)'\cdot W^*(X_t,\theta)\cdot r^*(y_t,X_t,\theta)] \;=\; 0 \quad .$$

Wooldridge specializes to this form because many specification tests are of this form. From section 2.3, (2.16) is the optimal RBCM test of (2.14) against

$$(2.17) \qquad H_1\colon\ \mathbb{E}_1[r^*(y_t,X_t,\theta)' \mid X_t] \;=\; G^*(X_t,\theta)\cdot\alpha^* \quad ,$$

if $W^*(X_t,\theta) \;=\; \mathrm{Var}(r^*(y_t,X_t,\theta) \mid X_t)$. So the RBCM test approach additionally provides an interpretation of the weighting function in (2.16).

Standard CM tests can be interpreted as RBCM tests. Strictly speaking, hypothesis tests merely reject or do not reject the null hypothesis. However, specifying an alternative hypothesis provides a very direct way to construct CM tests for particular forms of misspecification of a fundamental moment, as illustrated in section 4. The relationship between (2.13) and (2.15) or (2.14) and (2.17) can be used to interpret and contrast many standard tests, as done in section 5. And for a narrow but widely-used class of models, CM tests can be implemented directly by running the regression (2.5) of the RBCM test.

2.5 *Locally Equivalent Alternatives*

In principle, under the alternative hypothesis the right hand side of (2.4) may be nonlinear in α and g. For simplicity, consider the standard case of one fundamental moment condition:

$$(2.18) \qquad H_1\colon \mathbb{E}_1[r(y_t,X_t,\theta) \mid X_t] = h(X_t,\theta,\beta) \quad ,$$

where $h(X_t,\theta,\beta=\beta^*) = 0$. Then by first-order Taylor series expansion about $\beta = \beta^*$, $h(X_t,\theta,\beta^*) = 0 + \nabla_\beta h(X_t,\theta,\beta^*)\cdot(\beta - \beta^*)$, where $\nabla_\beta h(X_t,\theta,\beta^*)$ denotes the derivative of $h(X_t,\theta,\beta)$ w.r.t. β, evaluated at $\beta = \beta^*$. But this is of the form (2.4). The remainder term in the Taylor series expansion will disappear asymptotically for $\beta - \beta^* = O_p(T^{-1/2})$. Thus at least to local alternatives the linear form (2.4) nests nonlinear alternatives of the form (2.18). We use (2.4) as it is simpler, but could use (2.18) and estimate the corresponding regression by nonlinear, rather than linear, least squares.

Locally equivalent tests are discussed in Godfrey (1988). An example given in section 4 is the use of locally equivalent alternatives to transform specification tests for the conditional mean to an omitted variables problem.

3. IMPLEMENTATION OF REGRESSION-BASED CONDITIONAL MOMENT TESTS

The RBCM test approach can always be used to motivate and/or interpret the moment condition tested in a CM test. To implement RBCM tests, the parameter vector θ in (2.5) needs to be replaced by an estimate. The RBCM test is simplest to use when this replacement does not effect the asymptotic distribution of the test.

3.1 Tests when the RBCM Test Regression can be directly implemented

The estimator $\hat{\alpha}$ in (2.7) can be viewed as maximizing, with respect to α, the quasi-likelihood function:

$$(3.1) \qquad -Q_T(\alpha,\theta) = \sum_{t=1}^{T} W(X_t,\theta)\cdot\{r(y_t,X_t,\theta) - G(X_t,\theta)\cdot\alpha\}^2 \quad .$$

To implement RBCM tests, we need to replace θ by an estimate. It is known, e.g. Pagan (1986) or White (1990), that minimization of $Q_T(\alpha,\hat{\theta})$, where $\hat{\theta}$ is consistent for θ_0, the true value of θ, yields an estimator $\hat{\alpha}$ which has the same asymptotic properties as the estimator $\tilde{\alpha}$ which minimizes $Q_T(\alpha,\theta_0)$, if

$$(3.2) \qquad E_0\left[\frac{\partial^2 Q_T(\alpha_0,\theta_0)}{\partial\alpha\cdot\partial\theta} \mid X_t\right] = 0 \quad .$$

For the quasi-likelihood function (3.1), this condition will be satisfied under H_0 (or local alternatives to H_0), if in addition to the fundamental moment condition (2.3), the following moment condition is satisfied:

$$(3.3) \qquad \mathbb{E}_0[\ \nabla_\theta r(y_t,X_t,\theta) \mid X_t] = 0 \quad .$$

This condition implies zero asymptotic covariance between the moment criterion function $T^{1/2}r(y_t,X_t,\hat{\theta})$ and $T^{1/2}(\hat{\theta} - \theta_0)$ which are assumed to be asymptotically jointly normally distributed. (Pierce (1982, p.478). Models for which this condition holds are presented in sections 4 and 6.

When (3.3) holds, the asymptotic theory of section 2 is unchanged by replacing θ_0 by $\hat{\theta}_T$. Therefore we can test H_0 in (2.3) against H_1 in (2.4) by testing the significance of the weighted least squares estimator of α in the regression:

(3.4) $$r(y_t,X_t,\hat{\theta}_T) \;=\; G(X_t,\hat{\theta}_T)\cdot\alpha \;+\; u_t,$$

with weighting matrix $W_t(X_t,\hat{\theta}_T)$, and apply the usual theory of weighted least squares treating $r(y_t,X_t,\hat{\theta}_T)$ as a regular vector dependent variable and $G(X_t,\hat{\theta}_T)$ as a regular matrix of regressors. We have:

(3.5) $$\hat{\alpha} = \{ \sum_{t=1}^{T} \hat{G}_t{}'\hat{W}_t\hat{G}_t \}^{-1}\cdot \sum_{t=1}^{T} \hat{G}_t{}'\hat{W}_t\hat{r}_t \quad .$$

where $\hat{G}_t = G(X_t,\hat{\theta}_T)$, $\hat{W}_t \;=\; W(X_t,\hat{\theta}_T)$, and $\hat{r}_t \;=\; r(y_t,X_t,\hat{\theta}_T)$. We consider the limit distribution of $T^{1/2}\cdot\,\hat{\alpha}_W$ under local alternatives H_L: $\alpha = T^{-1/2}\gamma$, or more formally

(3.6) $$H_L\text{: } E[r(y_t,X_t,\theta) \mid X_t] \;=\; G(X_t,\theta)\cdot(T^{-1/2}\gamma),$$

where γ is a finite vector. Then under H_L:

(3.7) $$T^{1/2}\cdot\hat{\alpha}_W \xrightarrow{d} N(\gamma,\ [\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} G_t{}'W_tG_t]^{-1}\cdot [\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} G_t{}'W_t\Omega_tW_tG_t] \cdot [\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} G_t{}'W_tG_t]^{-1} \quad ,$$

where $G_t = G(X_t,\theta_0)$, $W_t \;=\; W(X_t,\theta_0)$, $r_t \;=\; r(y_t,X_t,\theta_0)$, and $\Omega_t \;=\; \mathbb{E}[r_tr_t{}' \mid X_t]$ is the unspecified conditional variance of r_t.

Specializing to H_0, (3.7) yields t-tests for individual components of α equalling zero. A consistent estimate of the variance-covariance matrix in (3.7) replaces G_t by $\hat{G}_t$, W_t by $\hat{W}_t$, and Ω_t by $\hat{\Omega}_t$, where $\hat{\Omega}_t$ is such that $\operatorname{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} \hat{G}_t{}'\hat{W}_t\hat{\Omega}_t\hat{W}_t\hat{G}_t \;=\; \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} G_t{}'{}_t\Omega_tW_tG_t$. We use the obvious estimator $\hat{\Omega}_t = \sum_{t=1}^{T} \hat{r}_t\hat{r}_t{}'$.

A joint test of whether all components of α equal zero is given by the chi-square test statistic:

$$(3.8) \quad d_W = [\sum_{t=1}^{T} \hat{r}_t' \hat{W}_t \hat{G}_t] \cdot [\sum_{t=1}^{T} \hat{G}_t' \hat{W}_t \hat{\Omega}_t \hat{W}_t \hat{G}_t]^{-1} \cdot [\sum_{t=1}^{T} \hat{G}_t' \hat{W}_t \hat{r}_t] \quad .$$

Under H_0, d_W is asymptotically chi-square distributed with p degrees of freedom.

Implementation of these tests in principle requires a multivariate regression package and matrix multiplication routines. When $n = 1$, however, we need only use an instrumental variables package along the lines suggested by Domowitz (1983) for inference based on heteroskedastic consistent estimators of the variance-covariance matrix of the least squares estimator.

Under H_L, d_W is asymptotically distributed as non-central chi-square with noncentrality parameter:

$$(3.9) \quad \lambda_W = \gamma^2 \cdot V_{G*G*} V_{G**G**} V_{G*G*}$$

where $V_{G*G*} = [\lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} G_t' W_t G_t]^{-1}$, $V_{G**G**} = [\lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} G_t' W_t \Omega_t W_t G_t]$.

The optimal RBCM test based on (2.5) uses weights W_t that maximize the local power by maximizing λ_W in (3.9). As expected, this is $W_t = \Omega_t^{-1}$, the generalized least squares estimator. Thus, we assume that under H_0, $\Omega_t = \Sigma_t$, defined in (2.11), which entails stochastic assumptions in addition to (2.3) and (3.3). We obtain the test statistic:

$$(3.10) \quad d_W^{opt} = [\sum_{t=1}^{T} \hat{r}_t' \hat{\Sigma}_t^{-1} \hat{G}_t] \cdot [\sum_{t=1}^{T} \hat{G}_t' \hat{\Sigma}_t^{-1} \hat{G}_t]^{-1} \cdot [\sum_{t=1}^{T} \hat{G}_t' \hat{\Sigma}_t^{-1} \hat{r}_t] \quad .$$

This optimal test is easily computed as the square root of the explained sum of squares from the OLS regression

$$(3.11) \quad \hat{\Sigma}_t^{-1/2}\hat{r}_t = \hat{\Sigma}_t^{-1/2}\hat{G}_t \cdot \alpha + u_t \ .$$

This regression can also be used to test whether individual components of α equal 0, using the usual t-tests. It is the regression that is directly suggested by the theory of section 2, and is viewed as the regression that is the basis for the test, rather than an auxiliary regression used to compute a test statistic.

The Pitman relative efficiency of d_w relative to d_w^{opt} is given by

$$\mathcal{E}\!f\!f \left[\frac{d_w}{d_w^{opt}} \right] = \frac{V_{G^*G^*} V_{G^{**}G^{**}} V_{G^*G^*}}{V_{G^+G^+}}$$

where $V_{G^+G^+} = [\lim T^{-1} \sum_{t=1}^{T} G_t{'}\Omega_t^{-1}G_t]^{-1}$. This shows the reduction in Pitman relative effciency due to the use of the robust rather than the optimal version of the test.

The optimal RBCM test does not require specification of the conditional distribution of y under H_0. However, it does require sufficient assumptions on the distribution of y to determine the second moment of $r(y_t,X_t,\theta)$, while the less powerful tests based on the weighted least squares regression (3.4) can be implemented using only (3.3) and the initial assumption that $r(y_t,X_t,\theta)$ has zero first moment.

To guard against the possibility of misspecification of Σ_t, we can of course compute asymptotically valid versions of the individual t-tests , using the result (3.7) with $\hat{W}_t = (\hat{\Sigma}_t)^{-1}$.

Furthermore it can be shown that (3.8) is chi-square distributed if $\hat{W}_t$

$= \hat{\Omega}_t$ (a non-trivial result since $\hat{W}_t$ then depends on r_t directly as well as indirectly via $\hat{\theta}_T$). Therefore, T times the uncentered R^2 from the univariate regression of the Tx1 vector of ones on $\hat{G}_t'\hat{W}_t\hat{r}_t$ is asymptotically distributed as chi-square under H_0. This regression, unlike regressions (3.4) or (3.11), is an auxiliary regression.

3.2 Tests when the RBCM Test Regression cannot be directly implemented

Most of the tests in this paper deliberately choose $r(y_t,X_t,\theta)$ so that (3.3) is satisfied. This is a departure from previous studies, aside from that of White (1990), who also exploits the simplification that arises when a condition similar to (3.3) is satisfied.

For tests of misspecification of the conditional mean of y_t, given in section 4.1, (3.3) will not hold. A separate theory for RBCM tests of the conditional mean can be developed, permitting inference on the estimated coefficients from the regression of $r(y_t,X_t,\hat{\theta}_T)$ on $G(X_t,\hat{\theta}_T)$, assuming that $\hat{\theta}_T$ is a weighted least squares estimator or a maximum likelihood estimator. Instead, we consider the joint test of significance of all coefficients, and appeal to the general theory of CM tests.

The general CM test (2.2) is implemented by using an estimator $\hat{\theta}_T$ of θ, consistent under H_0, to form $m_T(\hat{\theta}_T)$. Under appropriate assumptions, $T^{1/2}m_T(\hat{\theta}_T)$ has a limit distribution that is multivariate normal, which is the basis for a chi-square test of the null hypothesis moment condition. This test statistic can often be computed by an auxiliary regression.

The relevant theory is given in great detail in White (1990). A nice exposition is given in Pagan and Vella (1989). For any moment condition, under suitable conditions, by a first-order Taylor series expansion:

$$(3.12) \qquad T^{1/2}m_T(\hat{\theta}_T) = T^{1/2}m_T(\theta_0) + B_0 \cdot T^{1/2}(\hat{\theta}_T - \theta_0) + o_p(1) .$$