Tests of Independence in Parametric Models

Cameron, A.; Trivedi, P.

Publication date:
1992

entER
– for
omic Research

Discussion paper

TILBURG UNIVERSITY

No. 9237

TESTS OF INDEPENDENCE IN PARAMETRIC MODELS: WITH APPLICATIONS AND ILLUSTRATIONS

by A. Colin Cameron
and Pravin K. Trivedi

October 1992

ISSN 0924-7815

TESTS OF INDEPENDENCE IN PARAMETRIC MODELS: WITH APPLICATIONS AND ILLUSTRATIONS

By

A. Colin Cameron
Department of Economics
University of California
Davis, CA, 95616, U.S.A.

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

August 1991

This version: May 1992

ABSTRACT

Tests of independence between variables in a wide variety of discrete and continuous bivariate and multivariate regression equations are derived using results from the theory of series expansions of joint distributions in terms of marginal distributions and their related orthonormal polynomials. The tests are conditional moment tests based on covariances between pairs of orthonormal polynomials. Examples include tests of serial independence against bilinear and/or ARCH alternatives, tests of dependence in multivariate normal regression model, and dependence in count data models. Monte Carlo simulation based on bivariate count models is used to evaluate the size and power properties of the proposed tests. A multivariate count data model for Australian health care utilization data is used to empirically illustrate the tests.

Some Key Words: *SERIES EXPANSIONS; ORTHOGONAL POLYNOMIALS; SCORE TEST; DYNAMIC INFORMATION MATRIX TEST; ARCH AND BILINEAR MODELS; COUNT DATA.*

1. INTRODUCTION

In this paper we develop and apply a general framework for testing the assumption of zero correlation and, more generally, independence between pairs of random variables. This problem arises routinely in time series work and frequently in multi-equation cross section models. Tests of zero correlation in multivariate Gaussian regression models have been discussed in the econometric literature by Breusch and Pagan (1980) and Shiba and Tsurumi (1988) amongst others. But in non-Gaussian regression models zero correlation and independence are equivalent only for special classes of distributions and, in general, independence rather than zero correlation may be the interesting restriction to test. This is now recognized in nonlinear time series models; see Brock et. al. (1991), Hsieh (1989), and Robinson (1991). But for cross sectional work there is a relative dearth of tests. A general framework which considers both time series and cross section data is desirable.

This paper develops score type tests of independence based on a series expansion of the unknown joint pdf of the observations. This is simpler than the alternative approach of writing down the joint density explicitly and deriving score tests of independence, because in some non-Gaussian situations a flexible specification of the joint density is often not readily available. This also makes the construction of Wald and likelihood ratio tests difficult and partly explains the relative infrequency with which such tests are developed or used. By contrast the approach of this paper requires the specification of the univariate marginals which are then used to form an approximation to the joint distribution. Given correct specification of the marginals, the validity of the resulting independence tests does not depend on the adequacy of this approximation, though the power of the test will.

A general framework for testing dependence must address the following problem: except in special cases, tests of independence involve, in principle, an infinite number of restrictions. So an approach is required which will either test a smaller subset of these restrictions, or test the restrictions through one or more parameters in the joint distribution. How to derive and justify such restrictions is an important issue which is addressed by the general method of testing for independence between random variables considered here. It is based on a characterization of bivariate and multivariate distributions, introduced by Lancaster (1958) and subsequently elaborated and extended in Lancaster (1963, 1969), and Eagleson (1964). Infinite series expansions for the bivariate or multivariate joint distributions are constructed using the univariate marginal distributions and their associated orthonormal polynomials. The tests are conditional moment tests based on low order terms in the series expansion.

A brief comparison of the approach of this paper with other approaches in the literature may provide an improved perspective. In econometrics tests of dependence are most highly developed in the context of time series. Serial correlation tests are the most common, but the literature also considers nonlinear dependence of other types; for example, bilinear and ARCH dependence (Granger and Andersen (1978), Engle (1982), Weiss (1986)), and ARCH-M dependence (Engle et. al. (1987)). Much of this work is restricted by the assumption of conditionally Gaussian or symmetrically distributed errors. More recently, a nonparametric approach to testing for nonlinear time series dependence using the correlation integral has been investigated and applied by Brock with a number of co-authors; see Brock, Hsieh and LeBaron (1991) and Brock and Potter (1991). Robinson (1991) has also proposed a nonparametric test of independence of y_t and y_{t-1} for a stationary process $\{y_t\}$ based on the Kullback-Leibler entropy measure of the difference between the joint distribution and the product of the two marginals. By contrast, tests of dependence in cross sectional work have not received much attention. These pose additional problems because features such as non-normality, truncation and censoring are common in the data. The framework of this paper can address issues of dependence in many time series *and* cross sectional models and in both cases can accomodate non-normal parametric distributions. For example it can be applied to the following: tests of independence in the linear multivariate Gaussian regression model; tests of zero correlation between errors of seemingly unrelated non-Gaussian regression models; tests of serial dependence in time series models, including bilinear and ARCH processes; tests of independence in bivariate survival models; and tests of independence between the conditional mean of one random variable and the conditional variance (or higher moments) of another. The approach has some similarities with Hall's (1990) paper on score tests of normality against seminonparametric alternatives in which he uses a Gallant-Tauchen type nonorthogonal series expansion for the conditional density of $y(t)$, given $y(t-1)$; we use an orthogonal series expansion for the joint pdf.

The remainder of this paper is organized as follows. Section 2 deals with the underlying theory, and section 3 with its application to testing independence. Sections 4 and 5 provide examples. Section 6 reports the results of a simulation to evaluate the operating characteristics of the proposed tests and section 7 provides empirical illustrations. Section 8 concludes.

2. ORTHOGONAL SERIES EXPANSIONS FOR BIVARIATE DENSITIES

A key concept in this paper is the construction of a series expansion of a density using a sequence of orthonormal polynomials. This has been extensively discussed in statistics. Cameron and Trivedi (1990) proposed its use in specification tests for univariate regression models. In the bivariate case, considered here, the density $g(y_1, y_2)$ will be approximated by a series expansion, where the terms in the series expansion are orthonormal polynomials of the univariate marginal densities $f_1(y_1)$ and $f_2(y_2)$. The purpose of this section is to provide a brief but self-contained mathematical and statistical background for the reader's convenience. Proofs of theorems are not provided, but references to the literature are given. We begin with the univariate case.

Let f(y) denote the density of the independently distributed scalar continuous random variable y_t. (After appropriate changes all arguments can be extended to the discrete case). Assume the existence of finite moments of all order, μ_n, defined by $\mu_n = \mathbb{E}[y^n] = \int_{-\infty}^{\infty} y^n \cdot f(y)\, dy$, n=0,1,2.... In general f(y) may be a marginal or a conditional density, but for the purposes of this paper f(y) will be a conditional density, usually denoted by $f(y, X, \theta \mid X)$ where θ is an unknown parameter and X is a vector of observed explanatory variables. We use f(y) for generality and more compact notation.

Definition (Orthogonality): A system of orthogonal polynomials, henceforth abbreviated to OPS, $P_n(y)$ (or $P_n(y, X, \theta \mid X)$), degree $[P_n(y)] = n$, is called orthogonal with respect to f(y) (or $f(y, X, \theta \mid X)$) on the interval $a \le y \le b$ if

$$(2.1) \qquad \int_{-\infty}^{\infty} P_n(y) \cdot P_m(y) \cdot f(y)\, dy = \begin{cases} k_n & \text{if } m=n \\ 0 & \text{if } m \ne n \end{cases}$$

That is, $P_n(y)$ is a polynomial in y of degree n, a positive integer, satisfying the orthogonality condition

$$(2.2) \qquad \mathbb{E}[P_n(y) P_m(y)] = \delta_{mn} k_n, \quad k_n \ne 0 ,$$

where δ_{mn} is the Kronecker delta, $\delta_{mn} = 0$ if $m \neq n$, $\delta_{mn} = 1$ if $m = n$. In the special case of an *orthonormal polynomial* sequence, $k_n = 1$. An orthonormal polynomial is derived from an orthogonal polynomial by dividing by its standard deviation.

Let Δ be the matrix whose ij-element is μ_{i+j-2}, $i \geq 1$, $j \geq 1$. The necessary and sufficient condition for an arbitrary sequence $\{\mu_n\}$ to give rise to a sequence of orthogonal polynomials, unique up to an arbitrary constant, is that Δ is positive definite; see Cramer (1946, chapter 12.6). An orthonormal polynomial sequence $P_n(y)$ is complete if for every function $Q(y)$ with finite variance, $\mathrm{var}[Q(y)] = \sum_{n=1}^{\infty} a_n^2$, where $a_n = \mathbb{E}[Q(y)P_n(y)]$.

Theorem 1 *(Univariate Case)*: Let $\{P_0(y),\dots\}$ be the corresponding set of complete orthonormal polynomials for the density $f(y)$. Let $g(y)$ be another density which is ϕ^2-bounded in the sense that

$$(2.3) \qquad \phi^2 + 1 = \int_{-\infty}^{\infty} \{g(y)/f(y)\}^2 f(y)dy \ < \ \infty .$$

Then the following series expansion for $g(y)$ is formally valid:

$$(2.4) \qquad g(y) = f(y)\cdot\left[1 \ + \sum_{n=1}^{\infty} a_n P_n(y) \right] .$$

where $a_n = \int P_n(y)g(y)dy$, $\phi^2 = \sum_{n=1}^{\infty} a_n^2$. That is, the coefficients $\{a_n\}$ in the formal expansion are linear combinations of the moments of $g(y)$. ∎

For interpretation of ϕ^2 see Lancaster (1969, p.87). In (2.3), ϕ^2 is the variance of the probability ratio. See Ord (1972) for a proof.

An analogous result holds more generally, including discrete distributions. In the context of an expansion of type (2.4), we shall refer to $f(y)$ as the baseline density and to the sequence $\{P_n(y)\}$ as the associated orthonormal polynomial sequence (OPS).

Theorem 2 (Bivariate Case): Let $g(y_1, y_2)$ be a bivariate p.d.f. of continuous random variables y_1 and y_2 with respective marginal distributions $f_1(y_1)$ and $f_2(y_2)$ whose corresponding complete orthonormal polynomial sequences are, respectively, $Q_n(y_1)$ and $R_n(y_2)$, $n=0,1,\ldots\ldots$ If $g(y_1, y_2)$ is ϕ^2-bounded in the sense that

$$(2.5) \quad \phi^2 + 1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [g(y_1, y_2)/f_1(y_1)\cdot f_2(y_2)]^2 \cdot f_1(y_1)\cdot f_2(y_2)\cdot dy_1 \cdot dy_2 < \infty,$$

then the following expansion is formally valid:

$$(2.6) \qquad g(y_1, y_2) = f_1(y_1)\cdot f_2(y_2)\cdot \left[1 + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \rho_{nm} Q_n(y_1) R_m(y_2) \right],$$

where

$$(2.7) \qquad \rho_{nm} = E[Q_n(y_1)R_m(y_2)] = \int \int Q_n(y_1)\cdot R_m(y_2)\cdot g(y_1, y_2)\cdot dy_1 \cdot dy_2,$$

$$(2.8) \qquad \phi^2 = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \rho_{nm}^2 . \qquad \blacksquare$$

See Lancaster (1969, p.97, Theorem 4.1) for a proof.

For notational compactness we have suppressed X, but dependence of orthonormal polynomials on X variables is permissible, and is explicitly introduced in section 3.

3. TESTS OF INDEPENDENCE BASED ON ORTHOGONAL POLYNOMIALS

The results in (2.4) and (2.6) are the important ones for this paper. According to Theorem 2 a bivariate distribution is completely characterized almost everywhere by its marginal distributions and the matrix of correlations of all pairs of complete sets of orthonormal functions on the marginal distributions. Thus, in general, a test of independence in the bivariate case requires us to test H_0: $\rho_{nm} = 0$ (all n, m). This onerous task may be simplified in one of two ways. The null may be tested against an alternative in which dependence is restricted to be a function of a small number of parameters, usually just one. Or we may approximate the bivariate distribution by a series expansion with a smaller number of terms and then derive a score (LM) test of the null hypothesis H_0: $\rho_{nm} = 0$ (some n, m). For independence we require $\rho_{nm} = 0$ for all n and m. By testing only a subset of the restrictions, the hypothesis of approximate independence is tested.

3.1 Score tests based on orthogonal polynomial expansions for the joint p.d.f.

Consider whether a finite number of terms, say p, in the series expansion (2.6) provides an adequate approximation to $g(y_1, y_2)$, the unknown true joint p.d.f. If p=2, this is equivalent to the null hypothesis

(3.1) $$H_0: \rho_{11} = \rho_{22} = \rho_{12} = \rho_{21} = 0.$$

For general p we have the approximation

(3.2) $$\log g(y_1, y_2) = \log f_1(y_1) + \log f_2(y_2) + \log[1 + \sum_{n=1}^{p} \sum_{m=1}^{p} \rho_{nm} Q_n(y_1) R_m(y_2)]$$

(3.3) $$\nabla_{\rho_{nm}} \log g(y_1, y_2)\Big|_{\rho_{nm}=0} = Q_n(y_1) \cdot R_m(y_2), \quad n, m=1, 2, \ldots p,$$

where $\nabla_\rho \equiv \partial/\partial\rho$. To test H_0 we follow the score test approach. The score test will be based on $\mathbb{E}_0[\nabla_{\rho_{nm}} \log g(y_1,y_2)] = 0$, which implies that

$$\mathbb{E}_0[Q_n(y_1)\cdot R_m(y_2)] = 0, \qquad n,m=1,2\ldots,p, \tag{3.4}$$

where $\mathbb{E}_0$ denotes expectation under the null hypothesis of independence of y_1 and y_2.

Thus, given the orthonormal functions $Q_n(y_1)$, $R_m(y_2)$ corresponding to the marginals, a score test of the null hypothesis may be based on the expectation of the products of pairs of orthonormal functions. Since (3.4) must also hold when multiplied by a scalar (not a function of y_1 or y_2), we more generally consider tests based on the expectation of the products of pairs of orthogonal (not necessarily orthonormal) functions. These tests may also be interpreted as conditional moment (CM) tests in the sense of Newey (1985).

In some cases (see Section 3.4 for examples) the expansion (2.6) will simplify to the following:

$$g(y_1,y_2) = f_1(y_1)\cdot f_2(y_2)\cdot\left[\,1 + \sum_{n=1}^{\infty} \rho_n Q_n(y_1)\cdot R_n(y_2)\right], \tag{2.6a}$$

where $1 \ge \rho_1 \ge \rho_2 \ge \ldots\ldots \ge 0$. Then the test of independence will be based on reduced number of moment conditions $\mathbb{E}_0[Q_n(y_1)\cdot R_n(y_2)] = 0$, $n\ge1$. Furthermore, the rejection of H_0: $\rho_k = 0$ implies also the rejection of H_0: $\rho_n = 0$, $n > k$, because of the inequality relationship between the elements of the sequence $\{\rho_n\}$.

It is understandably more common to test for zero correlation than for independence in applied work, although the former does not imply the latter in non-Gaussian multivariate models. This corresponds to a test of $\rho_{11} = 0$. If the series expansion is truncated at p=1, then this is simply a test that the term for p=1 is negligible. If the series expansion is truncated at p=2, then for independence we test that $\rho_{11} = \rho_{12} = \rho_{21} = \rho_{22} = 0$. This test may be carried out by sequentially testing each separate hypothesis. Clearly $\rho_{11} = 0$ is in general a necessary, but not sufficient, condition for independence. Testing a very large number of terms will not be practicable, but tests up to second order dependence (p=2) may be useful in many applied situations. The analysis given above shows that the tests should be based on correlations between orthogonal (or orthonormal) polynomials corresponding to given marginal densities. This subsumes the common case of the multivariate normal regression model in which tests are based on regression residuals, which are simply the first order orthogonal polynomials. The appropriate polynomials to use in the case of tests of higher order dependence are less obvious; they are orthogonal polynomials given in the next section.

3.2 Implementation of the tests of independence

The test given above is implemented as follows:

(1) the marginals must be given;

(2) orthogonal polynomials corresponding to the marginals must be derived;

(3) the test statistic must be constructed from the appropriate sample quantities.

Step (1): Obvious. Section 3.5 discusses the impact of misspecification.

Step (2): For given marginals, orthogonal polynomial sequences may be constructed using knowledge of the moments of the marginals together with the definition of orthogonality given in (2.1) and (2.2). A general procedure and the expressions for orthogonal polynomials derived in this way are given in Cameron and Trivedi (1990). The polynomials up to order 2 are easily stated in terms of central moments, μ_i', of the baseline density, f(y):

$$(3.5) \qquad P_n(y) = \begin{cases} 1, & n=0 \\ y - \mu_1, & n=1 \\ (y-\mu_1)^2 - (\mu_3'/\mu_2')(y-\mu_1) - \mu_2', & n=2 \end{cases}$$

Note that in regression models $\mu_1 = \mathbb{E}[y \mid X,\theta]$, and $(y - \mu_1)$ will be usually interpreted as a regression residual; then the orthogonal polynomials are polynomials in these regression residuals. The orthonormal (unit variance) polynomials corresponding to $P_1(y)$ and $P_2(y)$, respectively, are $P_1(y)/\sqrt{\mu_2'}$ and $P_2(y)/\{\mu_4' - (\mu_3')^2/\mu_2' - (\mu_2')^2\}^{1/2}$. These orthogonal polynomials are not necessarily the obvious choice of polynomial in the regression residual. In particular, in going beyond tests based on the covariance of residuals (correlation) it is natural to consider tests based on the covariance of squared residuals, i.e. using $(y - \mu_1)^2 - \mu_2'$. By instead choosing the second order polynomial given in (3.5), we obtain a test that is statistically independent of the test based on the covariance of residuals, because $\mathbb{E}_0[P_1(y)P_2(y)] = 0$ by construction. Then the individual $\chi^2(1)$ tests given below will be jointly χ^2 with degrees of freedom given by the number of tests.

If the variable y is truncated or censored, then the polynomials $P_n(y)$ must be defined in terms of "generalized residuals" in the sense of Cox and Snell (1968), so that the polynomials have zero expectation. This adjustment is the same as used by Pagan and Vella (1989) and Cameron and Trivedi (1990); the former do not use orthogonal polynomials, the latter do.

An alternative equivalent procedure uses known generating functions for orthogonal polynomials for a given baseline density. These generating functions are known for the classical cases and for the Meixner class of distributions which includes the normal, gamma, Poisson, negative binomial and the binomial densities. The OPS corresponding to these cases are, respectively, Hermite, generalized Laguerre, Poisson-Charlier, Meixner and Krawtchouk polynomials. A summary table of their generating functions based on Meixner (1934), Eagleson (1964) and Griffiths (1985) is reproduced in Appendix A. To derive an orthogonal polynomial of order n, simply calculate $\nabla_z^n \Gamma(y;z)|_{z=0}$.

Step (3): The key moment restriction (3.4), written in full, is :

$$\mathbb{E}_0[Q_n(y_1, X_1, \theta_1 \mid X_1)\cdot R_m(y_2, X_2, \theta_2 \mid X_2)] = 0 , \tag{3.6}$$

where X_1 and X_2 are (not necessarily disjoint) subsets of regressor variables X, and θ_1 and θ_2 are (not necessarily disjoint) subsets of the parameter vector θ. Define

$$s_{nm}(y, X, \theta) = Q_n(y_1, X_1, \theta_1)\cdot R_m(y_2, X_2, \theta_2) . \tag{3.7}$$

By independence of R_m and Q_n, and conditional on X,

$$\begin{aligned} \mathrm{var}_0(s_{nm}(.)) &= (\mathbb{E}_0[R_m])^2\cdot \mathrm{var}_0(Q_n) + (\mathbb{E}_0[Q_n])^2\cdot \mathrm{var}_0(R_m) \\ &\quad + \mathrm{var}_0(Q_n)\cdot \mathrm{var}_0(R_m) , \\ &= \mathrm{var}_0(Q_n)\cdot \mathrm{var}_0(R_m) , \end{aligned} \tag{3.8}$$

when $[\mathbb{E}_0(R_m)] = [\mathbb{E}_0(Q_n)] = 0$, a property of orthogonal polynomials.

We begin by assuming that the parameters of the marginal distribution are known. By application of a central limit theorem, for orthogonal polynomials $Q_{n,t}$ and $R_{m,t}$ which have zero mean by construction, we obtain:

Proposition 1: For orthogonal polynomials, a test of the null hypothesis of

H_0: $\rho_{nm} = 0$ may be based on:

$$(3.9) \qquad \tau^2_{nm} = \Big(\sum_{t=1}^{T} Q_{n,t}R_{m,t}\Big)\cdot\Big(\sum_{t=1}^{T}(Q_{n,t}R_{m,t})^2\Big)^{-1}\cdot\Big(\sum_{t=1}^{T} Q_{n,t}R_{m,t}\Big).$$

Under H_0, τ^2_{nm} converges to the $\chi^2(1)$ distribution. ■

We note that τ^2_{nm} can be computed as T times the uncentered R^2 (equals the explained sum of squares) from the auxiliary regression of 1 on $Q_{n,t}R_{m,t}$.

When we use orthonormal polynomials, distinguised by an asterisk, $\mathbb{E}_0[(\sum_{t=1}^{T} Q^{*2}_{n,t})\cdot(\sum_{t=1}^{T} R^{*2}_{m,t})] = T^{-1}\mathbb{E}_0[\sum_{t=1}^{T}(Q^*_{n,t}R^*_{m,t})^2]$, by virtue of the homoscedasticity (and independence) of $Q^*_{n,t}$ and $R^*_{m,t}$. We obtain:

Proposition 2: For orthonormal polynomials, a test of the null hypothesis of H_0: $\rho_{nm} = 0$ may be based on

$$(3.10) \qquad r^2_{nm} = \Big(\sum_{t=1}^{T} Q^*_{n,t}R^*_{m,t}\Big)\cdot\Big(\Big(\sum_{t=1}^{T} Q^{*2}_{nt}\Big)\Big(\sum_{t=1}^{T} R^{*2}_{mt}\Big)\Big)^{-1}\cdot\Big(\sum_{t=1}^{T} Q^*_{n,t}R^*_{m,t}\Big).$$

Under H_0, T times r^2_{nm} converges to the $\chi^2(1)$ distribution. ■

Use of proposition 1 requires a zero mean for $Q_{n,t}$ and $R_{m,t}$ under H_0, i.e. correct specification of the first n moments of y_{1t} and m moments of y_{2t}. Use of proposition 2 additionally requires a constant variance for $Q^*_{n,t}$ and $R^*_{m,t}$ under H_0, which would be guaranteed by correct specification of the first 2n moments of y_{1t} and 2m moments of y_{2t}. In Section 6 these tests are compared in a Monte Carlo environment.

To operationalize these test statistics we need to consider the effect of substituting the estimated parameters $\hat{\theta}_1$ and $\hat{\theta}_2$ in place of the true population values θ_1 and θ_2. The following definitions will be used:

$$s_{nm,t}(\hat{\theta}) = s_{nm,t}(y, X, \hat{\theta} \mid X) , \tag{3.11}$$

$$s_{nm,T}(\hat{\theta}) = T^{-1} \sum_{t=1}^{T} s_{nm,t}(\hat{\theta}) . \tag{3.12}$$

We wish to test independence by testing the closeness of $s_{nm,T}(\hat{\theta})$ to zero, an example of a conditional moment test of Newey (1985) and Tauchen (1985). When $\hat{\theta}$ is root-T consistent and

$$\mathbb{E}_0[\nabla_\theta s_{nm}(y, X, \theta) \mid X_1] = 0 , \tag{3.13}$$

$T^{1/2}s_{nm,T}(\hat{\theta}) = T^{1/2}s_{nm,T}(\theta) + o_p(1)$, a result established by a first-order Taylor series expansion of $T^{1/2}s_{nm,T}(\hat{\theta})$. Thus we can immediately apply proposition 1, with $s_{nm,t}(\hat{\theta})$ in place of $s_{nm,t}(\theta)$. Letting $\hat{Q}_{nt} = Q_n(y_{1t}, X_{1t}, \hat{\theta}_1)$ and $\hat{R}_{mt} = R_m(y_{2t}, X_{2t}, \hat{\theta}_2)$, we can implement proposition 1 by computing T times the uncentered R^2 from regression of 1 on $\hat{Q}_{nt}$ and $\hat{R}_{mt}$. For orthonormal polynomials, we can alternatively use proposition 2 and compute T times the correlation coefficient between $\hat{Q}^*_{nt}$ and $\hat{R}^*_{mt}$.

When θ_1 and θ_2 in (3.7) are distinct, condition (3.13) will be satisfied under the null hypothesis that y_1 and y_2 are independent. To see this, observe that $\mathbb{E}_0[\nabla_{\theta_1} s_{nm}(y, X, \theta) \mid X] = \mathbb{E}_0[\nabla_{\theta_1} Q_n(y_1, X_1, \theta_1) \mid X] \cdot \mathbb{E}_0[R_m(y_2, X_2, \theta_2) \mid X] = 0$, since the orthogonal polynomial $R_m(y_2, X_2, \theta_2)$ has expectation zero by construction. Similarly $\mathbb{E}_0[\nabla_{\theta_2} s_{nm}(y, X, \theta) \mid X] = 0$. Thus the preceding theory can be directly applied.

For most of the tests considered in this paper, (3.13) holds and this simple form is adequate. The one exception is for tests of serial dependence in time series models. Then $(y_{1t}, y_{2t}) = (\varepsilon_t, \varepsilon_{t-1})$, where ε_t is the underlying error in the time series model and the conditioning variables are the current exogenous and lagged exogenous and endogenous variables. In this case $\theta_1 = \theta_2 = \theta$. Conditional moment tests in dynamic models are considered by White (1987), who shows that if $\hat{\theta}$ is the MLE then a $\chi^2(1)$ test statistic is computed by T times the uncentered R^2 from the regression of 1 on $s_{nm,t}(\hat{\theta})$ and the t-th score for the conditional density (given lagged endogenous variables). Implementation for the time series example is considered further in section 5.

3.3 Multivariate extensions

Since there are close parallels with the bivariate case, here we shall only sketch the argument. Suppose there are r random variables $(y_1, y_2, \ldots, y_r)$ with joint p.d.f. $g(y_1, y_2, \ldots, y_r)$ and r marginals $f_1(y_1), f_2(y_2), \ldots, f_r(y_r)$ whose respective OPS are denoted by $Q_n^i(y_i)$, where $n = 0, 1, \ldots \infty$; $i = 1, 2, \ldots r$. According to Lancaster (1969, Theorem 5.9, p.101), under the condition of ϕ^2-boundedness the joint p.d.f. admits a series expansion of the same type as (2.6), viz.,

$$(3.14) \qquad g(y_1, y_2, \ldots, y_r) = f_1(y_1)\cdot f_2(y_2)\cdot \ldots \cdot f_r(y_r)$$

$$\times \Big[1 + \sum_{\substack{i \\ i<j}} \sum_{j} \sum_{m} \sum_{n} \rho_{mn}^{ij}\cdot Q_m^i(y_i)\cdot Q_n^j(y_j) + \sum_{\substack{i \\ i<j<k}} \sum_{j} \sum_{k} \sum_{m} \sum_{n} \sum_{o} \rho_{mno}^{ijk}\cdot Q_m^i(y_i)\cdot Q_n^j(y_j)\cdot Q_o^k(y_k) + \ldots\Big]$$

where ρ_{mn}^{ij} denotes the correlation coefficient between orthonormal polynomials $Q_m^i(y_i)$ and $Q_n^j(y_j)$. Independence between the n random variables implies necessary conditions such as $\sum_{\substack{i \\ i<j}} \sum_{j} \sum_{m} \sum_{n} \rho_{mn}^{ij} = 0$ and $\sum_{\substack{i \\ i<j<k}} \sum_{j} \sum_{k} \sum_{m} \sum_{n} \sum_{o} \rho_{mno}^{ijk} = 0$.

Since a multivariate distribution is characterized by its marginal distributions and all possible pairs of complete sets of orthonormal functions on the marginal distributions, the computational burden of testing more than just a select few of these restrictions is likely to be formidable. Hence in most applications where r is large in the relevant sense, it would seem sensible to exploit prior knowledge and the structure of the problem in setting up a test based on a subset of the restrictions.

For example, for tests of independence based on the first-order orthogonal polynomials, i.e. covariances between y_i and y_j, there are potentially $r(r-1)/2$ pairs to test, each a $\chi^2(1)$ test. The joint test will be a $\chi^2(r(r-1)/2)$ test. A simpler problem is that of testing for zero correlation between two subsets of random variables, denoted by $\mathbf{Y}_1$ and $\mathbf{Y}_2$ with the covariance matrix $\Sigma = [\Sigma_{ij}]$, $i,j=1,2$, where $\text{rank}(\Sigma_{11}) = r_1$ and $\text{rank}(\Sigma_{22}) = r_2$. As an extension of the scalar case, following Hooper (1959) define the squared canonical correlation coefficient

$$(3.15) \qquad \rho_c^2 = (\text{vec}\ \Sigma_{21})'(\Sigma_{11}^{-1} \otimes \Sigma_{22}^{-1})\ (\text{vec}\ \Sigma_{21})$$

$$= \text{tr}\ (\Sigma_{11}^{-1}\ \Sigma_{12}\ \Sigma_{22}^{-1}\ \Sigma_{21}),$$

which is zero under the null hypothesis of independence of $\mathbf{Y}_1$ and $\mathbf{Y}_2$. Let r_c^2 denote the sample squared canonical correlation coefficient. Then, analogously to Proposition 2 we have

Proposition 3: Under the null hypothesis of independence, $T \cdot r_c^2 \sim \chi^2(r_1 r_2)$. ∎

This result has appeared in the literature in a number of places, including Jupp and Mardia (1980; Theorem 1) and Shiba and Tsurumi (1988; Proposition 3.)

3.4 Correlation and independence

Since tests of independence are onerous to apply, it seems natural to ask when a test of zero correlation is adequate as a test of independence. We are unaware of general and practically useful results along these lines, so the following discussion is restricted to a few interesting special cases.

Consider the generalized exponential family (Jupp and Mardia (1980))

$$(3.16)\quad g(y_1,y_2;\ \theta_1,\theta_2,\rho) = \exp\{\theta_1'v(y_1) + v(y_1)'\rho w(y_2) + \theta_2'w(y_2) - c(\theta_1,\theta_2,\rho) + d_1(y_1) + d_2(y_2)\}$$

where $v(.)$, $w(.)$ and $d_1(.)$ and $d_2(.)$ are some functions, and $c(\theta_1,\theta_2,0) = c_1(\theta_1)\cdot c_2(\theta_2)$. Under independence ρ=0 in which case the right hand side is a product of two exponential families. In such cases the null of independence can be tested through the single restriction $\rho = 0$, and hence the tests of zero correlation and independence are equivalent; see section 6 for an illustration.

Specific bivariate parametric families were considered by Eagleson (1964) who examined the Meixner class of distributions (see Appendix A) in which the correlated random variables were generated by considering sums of independent random variables with common components in the sums. He showed that in this class correlation is an adequate measure of dependence. In the bivariate Meixner class, with variables y_1 and y_2 defined as $y_1 = u + v$, $y_2 = v + w$, and u, v and w are independently distributed $\rho^2 = \mathrm{var}(v)/[\mathrm{var}(u+v)\cdot\mathrm{var}(v+w)]$.

When zero correlation implies independence we expect a classical statistical test, Wald, likelihood ratio or score, of correlation to be the same as a test of independence. However, the starting point in this paper is the univariate marginal distributions; no restrictions are placed on the joint distribution. Consequently we test the null against a wide range of joint distributions.

3.5 Test procedures when the marginals are misspecified

The preceding discussion leading to the derivation of the test statistics (3.9) and (3.10) was based on the assumption of correct specification of the marginals. We now consider the impact of misspecified marginals.

Under independence the true joint pdf $g(y_1,y_2)$ can be written as follows:

$$g(y_1,y_2) = f_1(y_1)\cdot f_2(y_2)$$

$$= f_1^*(y_1)\cdot f_2^*(y_2)\cdot\left[1 + \sum_{n=1}^{\infty} \alpha_n P_n(y_1)\right]\cdot\left[1 + \sum_{m=1}^{\infty} \beta_m Q_m(y_2)\right]$$

$$(3.17)\qquad = f_1^*(y_1)\cdot f_2^*(y_2)\left[1 + \sum_{n=1}^{\infty} \alpha_n P_n(y_1) + \sum_{m=1}^{\infty} \beta_m Q_m(y_2) + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_n\beta_m P(y_1)Q_m(y_2)\right]$$

where we have used (2.4) to write each *true* marginal as a series expansion around a baseline pdf, denoted by an asterisk. If the baseline marginals are correctly specified, then $\mathbb{E}[P_n(y_1)]$ and $\mathbb{E}[Q_m(y_2)]$ are both zero for all m and n, and $\alpha_n=0$, $\beta_m=0$, for all n and m; then (2.6) and (3.17) will coincide. The null hypothesis of independence implies $\mathbb{E}[P_n(y_1)\cdot Q_m(y_2)] = \mathbb{E}[P_n^*(y_1)\cdot Q_m^*(y_2)] = 0$, which is tested by (3.9) and (3.10) derived from the series expansion in (2.6). If both the marginals are mispecified, then $\mathbb{E}[P_n(y_1)]$ and $\mathbb{E}[Q_m(y_2)]$ will not be zero for all m and n, and $\alpha_n \neq 0$, $\beta_m \neq 0$, and $\mathbb{E}[P_n(y_1)\cdot Q_m(y_2)] \neq 0$ and $\mathbb{E}[P_n^*(y_1)\cdot Q_m^*(y_2)] \neq 0$. The tests may be significant because the baseline marginals are misspecified, not necessarily because the variables are dependent. In this case (3.9) and (3.10) will test whether the product of the misspecified marginals equals the joint pdf. (This is in the spirit of Hall's (1990) interpretation of his tests for the normal linear time series models. See also Robinson (1991).) Finally, note, rather remarkably, that if one marginal is misspecified, but the other is correctly specified the tests retain their validity as tests of independence. This result is not useful in dynamic time series models where misspecification of one marginal implies misspecification of the other.

If the main focus is on tests of *independence* rather than misspecification of the baseline marginals, then specification tests of the marginals, based on the relevant orthogonal polynomials (see Cameron and Trivedi (1990)) will be an appropriate preliminary to tests of independence. In Section 6 we examine the power of the independence tests under misspecification.

4. TESTS FOR SEEMINGLY UNRELATED REGRESSIONS

Consider the seemingly unrelated regression (SUR) multivariate normal model

$$(4.1) \qquad \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} \sim N(\mu,\ \Omega), \text{ where } \mu = (\mu_i),\ \Omega = [\omega_{ij}],\ i,j = 1,2,\ldots,r.$$

Suppose we wish to test the null hypothesis H_0: $\omega_{ij} = 0$, $i \neq j$, against the alternative that $\omega_{ij} \neq 0$. In the bivariate case the usual test for this is based on the sample covariance between the residuals in the two regressions; see Engle, Hendry and Richard (1983). In the multivariate case Breusch and Pagan (1980) consider the score test, and Shiba and Tsurumi (1988) develop Wald, likelihood ratio and score tests for block-diagonality of Ω. The score test is essentially the same as that discussed in section 3 with p=1.

The theory of section 3 permits extension of these tests in two dimensions. First, the usual score tests are tests of zero correlation rather than independence. This is a consequence of specifying the multivariate model to be joint normal, for which zero correlation and independence are equivalent, rather than some other distribution where the univariate marginal distributions are normal, but for which zero correlation does not necessarily imply independence. The tests of independence proposed here will not only use residuals $y_1 - \mu_1$ but also higher order orthogonal polynomials. For the normal the second order orthogonal polynomial is $(y_i - \mu_i) - \omega_{ii}$.

The second extension is to relax the assumption that the marginal distributions are normal. For example, sets of equations for different count data variables may be considered. Then the marginals may be Poisson, and the independence test will be based on cross products of orthogonal polynomials, the first two of which are $(y_i - \mu_i)$ and $(y_i - \mu_i)^2 - y_i$, where μ_i is the conditional mean of y_i, usually $\exp(x_i'\beta)$. This illustration is considered further in sections 5 and 7. Note that the independence tests do not require estimation or specification of multivariate models, such as the multivariate Poisson. Another non-normal example is sets of binary choice equations. Then we wish to know whether to simply model equations separately, with binomial (one trial) marginals, say the probit model, or to use a multivariate model such as the multivariate probit (Ashford and Sowden (1970)). Application of tests based on section 3 is again straight-forward as the condition (3.13) is satisfied.

5. TESTS OF INDEPENDENCE IN TIME SERIES MODELS

Consider time series observations on a scalar process $\{y_t, t=1,2,\dots,T\}$, where y_t has conditional mean $\mu_t = \mathbb{E}[y_t|F_{t-1}]$, and F_{t-1} denotes the information set comprising $y_{t-1}, y_{t-2}, \dots$ and current and lagged exogenous variables.

To test first order dependence we consider bivariate random variables $(y_{1t}, y_{2t}) = (y_t - \mu_t, y_{t-1} - \mu_{t-1}) = (\varepsilon_t, \varepsilon_{t-1})$. Applying the results of section 3.2, the test based on first order orthogonal polynomials (i.e. p=1) will be a test of whether the covariance of y_{1t} and y_{2t} equals zero, i.e. whether the errors ε_t and ε_{t-1} are correlated. The test is simply the usual test based on the first order serial correlation coefficient r_1: $T \cdot r_1^{\,2}$ is $\chi^2(1)$ under H_0.

Tests of independence often use L>1 successively lagged values of the variable y_t. Then we have an (L+1)-variate random variable $\{y_{1t}, y_{2t}, \ldots, y_{(L+1)t}\} = \{y_t, y_{t-1}, \ldots, y_{t-L}\}$. The independence tests of section 3.3 (the multivariate case) based on the first-order orthogonal polynomials will be tests of L(L+1)/2 correlations, but with the assumption of second order stationarity we need only consider the L correlations between y_{1t} and $y_{(i+1)t}$, i=1,...,L, or equivalently the correlation between ε_t and ε_{t-i}, i=1,...,L. Let $r_i = (\sum_t \varepsilon_t \varepsilon_{t-i})/(\sum_t \varepsilon_t^{\,2})$ denote the ith sample correlation coefficient. Then from section 3, $T \cdot r_i^{\,2}$ is $\chi^2(1)$. By the statistical independence of each test, the overall test based on the quantity

$$(5.1) \qquad \tau^2 = T \cdot \sum_{i=1}^{L} r_i^2,$$

is asymptotically $\chi^2(L)$. L is recommended to be chosen as some pre-determined fraction of the sample size. Ljung and Box (1979) have suggested an improved variant of this asymptotic test, which uses $\tau_a^2 = T(T+2) \cdot \sum_i (T-i)^{-1} r_i^2$.

The above standard tests use only the first order orthogonal polynomials. As a result they will not necessarily detect nonlinear forms of time dependence. For example, these tests do not test for an ARCH model, since in the ARCH model the errors are uncorrelated. A number of nonlinear models and tests for nonlinear dependence exist. For example, see Hsieh (1989), who applies a range of tests to daily exchange rate data. Here we propose testing for nonlinear dependence using higher order orthogonal polynomials. We do this for the simplest case where orthogonal polynomials up to second order are used and where only first order serial dependence is considered.

Specifically, tests of zero serial dependence are based on the restriction $\mathbb{E}_0[P_m(y_1)P_n(y_2)] = 0$, where, choosing p=2, the orthogonal polynomials $P_m(y_1)$ and $P_n(y_2)$, m,n = 1,2 are given in equation (3.5). This yields the following four conditional second-moment restrictions:

$$\mathbb{E}_0[\varepsilon_t\varepsilon_{t-1}] = 0 \qquad m = 1,\ n = 1 \tag{5.2}$$

$$\mathbb{E}_0[\varepsilon_t(\varepsilon^2_{t-1} - (\mu_3/\mu_2)\varepsilon_{t-1} - \mu_2)] = 0 \qquad m = 1,\ n = 2 \tag{5.3}$$

$$\mathbb{E}_0[\varepsilon_{t-1}(\varepsilon^2_t - (\mu_3/\mu_2)\varepsilon_t - \mu_2)] = 0 \qquad m = 2,\ n = 1 \tag{5.4}$$

$$\mathbb{E}_0[\varepsilon^2_t\varepsilon^2_{t-1} - (\mu_3/\mu_2)(\varepsilon_t\varepsilon^2_{t-1} + \varepsilon^2_t\varepsilon_{t-1}) - \mu_2(\varepsilon^2_t + \varepsilon^2_{t-1}) + \tag{5.5}$$

$$(\mu_3/\mu_2)^2(\varepsilon_t\varepsilon_{t-1}) + \mu_3(\varepsilon_t + \varepsilon_{t-1}) + \mu^2_2] = 0. \qquad m = 2,\ n = 2.$$

These moment restrictions essentially imply an absence of generalized serial dependence. Restriction (5.2) yields the test of zero first order autocovariance already discussed at the beginning of this section. Restriction (5.3) states that the covariance between the residual and lagged squared residual is zero. This restriction will be violated in a Gaussian model if, for example, the model displays an ARCH-M (ARCH-in-the-mean) effect (Engle, Lillien and Robins (1987), Hsieh (1989), Bollerslev, Chou and Kroner

(1992)). Thus, a test of (5.2) may be viewed as a test of the conditional mean specification of the maintained model, while a test of (5.3) may be viewed as a test of the conditional variance specification. Restriction (5.4) states that the covariance between current squared residual and lagged residual is non-zero. This restriction will be violated if the data display bilinear dependence or non-symmetric ARCH dependence. Finally, restriction (5.5) states that a generalized ARCH effect is absent.

The restrictions are stated in a general form and may be specialized to the Gaussian case by setting μ_3 to zero. They are suggestive of the type of nonlinear dependence one might test for in non-Gaussian nonlinear models of the type which have been discussed in the empirical finance literature; see Bollerslev *et. al.* (1990, sections 2.3-2.4 and 3.3-3.4).

Specializing to Gaussianity, the components of the dynamic information matrix (DIM) test of White (1987) yield restrictions similar to (5.2)~(5.5) with $\mu_3 = 0$, see White (1987). The case considered above corresponds to choosing lag of order 1 whereas the full DIM test is based on the matrix formed by the outer product of the vector of conditional scores and lagged conditional scores. Both White and Perez-Amaral (1989) discuss in detail the implementation of the full test and its components using regressions. Weiss (1986) and Perez-Amaral (1989) also discuss procedures for making these tests robust against departures from maintained distributional assumptions.

Tests based on (5.2) are widely used, while tests based on (5.3) and (5.4) are not. An important special case of (5.5) is the first order ARCH model, denoted ARCH1 (Engle (1982)). Assume that the distribution of y_t is symmetric, in which case $\mu_3 = 0$. Then the conditional moment restriction (5.5) collapses to

$$(5.5a) \qquad \mathbb{E}_0[\varepsilon_t^2\varepsilon_{t-1}^2 - \mu_2(\varepsilon_t^2 + \varepsilon_{t-1}^2) + \mu_2^2] = \mathbb{E}_0[(\varepsilon_t^2 - \mu_2)(\varepsilon_{t-1}^2 - \mu_2)] = 0.$$

The test of the restriction (5.5a) corresponds exactly to the test of independence against the alternative of serial dependence of first order ARCH type. Under Gaussianity, this test is conventionally implemented by running a regression of $\hat{\varepsilon}_t^2$ on 1 and $\hat{\varepsilon}_{t-1}^2$ and calculating TR^2 from this regression. But the version (5.5) is more general since it does not assume symmetry of the null distribution; see Engle (1982, sections 4-5) and Nelson (1991) on the role of symmetry in testing for ARCH. Incorrectly assuming symmetry when testing for ARCH will affect both the size and the power of the ARCH test. The present version also ensures orthogonality between the ARCH test and other tests of serial independence. It is easily implemented using the theory developed in section 3.2, in conjunction with tests of restriction (5.3) and (5.4) which test for a lower order of serial dependence than does (5.5). Extension to higher order processes is straight-forward, though cumbersome.

Observe that at least in "diagonal" models there is a close connection between an absence of symmetry and bilinearity, in that in general the latter implies the former (Granger and Anderson (1978, p.53)). Conventional ARCH tests assume symmetry and do not usually test for asymmetry, but the above discussion suggests that tests of skewness and bilinearity should precede tests of ARCH, and that the variant of ARCH test of this paper should be used if there is evidence of either.

6. A MONTE CARLO INVESTIGATION

This section reports the results of a Monte Carlo investigation of the properties of the regular and the orthonormalized variants of the independence tests in (3.9) and (3.10). To complement available extensive time series work and the empirical illustration of Section 7, the chosen context is that of discrete distributions in which the data are generated by bivariate Poisson and mixed Poisson regression models and tests are based on Poisson marginals. The maximum order of the polynomials is restricted to 2, so there are four tests in each category. The first two orthogonal polynomials are $P_1(y) = y-\mu$ and $P_2(y) = (y-\mu)^2-y$ and their orthonormalized counterparts are $P_1(y)/\sqrt{\mu}$ and $P_2(y)/(\sqrt{2}\mu)$, respectively. Substitution of these expressions in (3.9) and (3.10) yields the eight tests which will be considered.

6.1 Design of the Monte Carlo experiments

Eight models and two sample sizes, viz., 50 and 200, were used for size and power comparisons. Each simulation experiment was based on 500 paired replications. The regression component included a constant term and one explanatory variable, x, taken as a random draw from uniform [0,1] distribution and held fixed in all replications. Estimation is by maximum likelihood.

A comparison of the nominal and empirical size of the tests is carried out with Models 1 and 2 in which the variables y_1 and y_2 are generated as independent Poisson variates with parameters μ_1 and μ_2; in Model 1, $\mu_1=\mu_2=\mu$ where $\mu = \exp(\beta_0 + \beta_1 x)$; in Model 2, $\mu_2>\mu_1$. In Model 3 the variables y_1 and y_2 are again independent, but the marginals are not Poisson. The variables are dependent in the remaining cases. In the case of Models 4 and 5, data are generated via a bivariate Poisson distribution, with Poisson marginals. The tests are applied under correct specification of the marginals. In Models 6, 7 and 8, however, y_1 and y_2 are, respectively, Poisson$(\mu_1|\zeta_1)$ and Poisson$(\mu_2|\zeta_2)$ where

(6.1) $$\mu_1|\zeta_1 = \exp(\beta_0 + \zeta_1 + \beta_1 x)$$

(6.2) $$\mu_2|\zeta_2 = \exp(\beta_0 + \zeta_2 + \beta_1 x)$$

where ζ_1 and ζ_2 denote random terms corresponding to unobserved heterogeneity. When these random terms have a non-zero mean, α_0 will be unidentified, but this will not affect inference. The presence of the ζ_1 and ζ_2 induces overdispersion in the marginal distributions of y_1 and y_2. Dependence between y_1 and y_2 is induced by using correlated ζ_1 and ζ_2, so that in contrast to Models 4 and 5 in which the marginals are correctly specified, the assumption of Poisson marginals is a misspecification when the intercept term is random. In all cases except Model 2, $\beta_0 = \beta_1 = 1$; in Model 2, β_1 is 1 and 1.5, respectively, in each pair of replications. Additional details about data generation is given below along with the summary of the results.

6.2 The Results

Tables 1 and 2 present the rejection rates for the null hypothesis of independence, for N=50 and N=200, respectively, at nominal significance levels of 1, 5, and 10 per cent for two groups of tests, viz., $(T_{11}, T_{12}, T_{21}, T_{22})$ which correspond to the expression in (3.9), and $(T^n_{11}, T^n_{12}, T^n_{21}, T^n_{22})$ which correspond to the expression in (3.10) for the orthonormalized version. The results for Models 1-2 help to evaluate the match between nominal and empirical test size, and the rest throw light on the power properties.

Models 1 and 2: Since the data are conditionally independent the rejection frequency for all tests should equal the nominal significance level. The results for N=50 and N=200 suggest that for all tests the match between the nominal and empirical significance level is good at 1, 5 and 10 per cent levels in the sense that the observed divergences between the two are within the sampling variations expected from 500 replications.

Model 3: The variables y_1 and y_2 are independent but each is marginally overdispersed. This is achieved by adding to the intercept an exponentially distributed random variable with parameter $1/\lambda$, the λ being fixed at 2 when N=50, and 3 when N=200. This experiment enables us to observe the possible size distortions in tests of independence.

The results show that, for both sample sizes and for all orthonormal tests and for low order non-orthonormal ones also, the rejection frequency is considerably greater than the significance level of the test. For the tests based on low order polynomials the rejection frequency is close to 50 per cent even when N=50. In a sense this indicates size distortion, but it also indicates power against the alternative of misspecified marginals (as discussed earlier in Section 3.5).

Models 4 and 5: Here y_1 and y_2 are generated by a bivariate Poisson model where $y_1 = u + v$, $y_2 = v + w$, and u, v and w are independent Poisson distributed with parameters λ_1, λ_2 and λ_3, respectively. Then $y_1 \sim$ Poisson$(\lambda_1 + \lambda_2)$, $y_2 \sim$ Poisson$(\lambda_2 + \lambda_3)$, $\text{cov}(y_1, y_2) = \text{cov}(u+v, v+w) = \text{var}(v) = \lambda_2 > 0$, and $\rho^2 = \lambda_2^2/(\lambda_1+\lambda_2)(\lambda_2+\lambda_3)$; see Johnson and Kotz (1969), Gourieroux et. al. (1984). *In this case zero correlation implies independence.* Hence the tests based on low order polynomials should have high power. The degree of correlation may be controlled by varying λ_2, given other parameters. For N=50, the correlation is roughly 0.24 for Model 4, and 0.29 for Model 5; for N=200, the corresponding correlations are 0.19 and 0.21.

T_{11} and T_{11}^n have significantly higher rejection rates than the tests based on second order polynomials. Except when the sample sizes are large and and/or the correlations are high, tests based on higher order polynomials will have low power. The tests based on the first order orthonormal polynomials are preferred to the higher order tests.

Model 6: The random heterogeneity component in Models 6 through 8 was specified as follows: $\zeta_1 = \kappa(\eta^2-1)$ where $\eta \sim N(0,1)$; $\zeta_2 = \zeta_1$; $\kappa = 0.25$ (N=50) or 0.15 (N=200). The value of the scaling constant κ controls overdispersion of, and dependence between, y_1 and y_2. Because of the joint presence of overdispersion and dependence we expect the rejection frequencies to be high, and they are. The orthonormal version of the test has higher rejection frequencies and these are almost as large for tests based on second order polynomials as the first order ones. For the non-orthonormal tests the rejection frequecies for higher order tests are always lower but they rise sharply as N goes from 50 to 200.

The differences between the two versions of the test may be understood as follows. Overdispersion implies that $E[P_2(y)]$ will not be zero, as under the null. The tests based on orthonormal polynomials depend upon the higher moments of the assumed marginal distributions and consequently should be more affected by such a misspecification. The true variances of the orthonormal polynomials are understated under the null, and the tests based on them derive their power against this misspecification from this source. Viewed as tests against the joint null of independence and correct specification of the marginals, the orthonormal versions of the tests perform better.

Model 7: Here (6.1) and (6.2) are specialized as follows: $\zeta_1 = \kappa \cdot u$ where u~uniform(0,1); $\zeta_2 = \zeta_1$; $\kappa = 1.00$ (N=50) or 0.70 (N=200). In this case the random terms add to both the conditional mean and the variance of the model. The larger is κ, the greater the correlation.

The rejection frequencies are again consistently higher for the orthonormal version. Further, of the tests based on second order polynomials, only T^n_{22} has reasonable power compared with T_{11} and T^n_{11}. Once again this suggests that higher order tests may have limited utility if the sample is small.

Model 8: Here we construct a case in which the tests based on second order polynomials might have higher power than those based on the first order ones. This is done by specializing (6.1) and (6.2) such that the heterogeneity terms have close to zero correlation by construction but are not independent. Specifically, let: $\zeta_1 = \kappa(u_1 + u_2 - 1)$; $\zeta_2 = -\kappa(u_1 - u_2)$ where u_1~uniform(0,1) and u_2~uniform(0,1); $\kappa = 3.00$ (N=50) or $\kappa = 1.75$ (N=200). Tedious calculations show that the correlation coefficient is between -.0173 and -0.0179 when $\kappa=3$, and between -.0119 and -.0135 when $\kappa=1.75$. This construction is apparently successful in that it yields a case in which T_{22} is a considerably more powerful test than T_{11}. But this is also a case in which the orthonormal variant turns out to have close to zero power irrespective of whether it is based on the first or second order polynomials!

To summarize: Important differences between (3.9) and (3.10) arise when both marginals are misspecified; otherwise the differences are smaller and the latter is more powerful. Broadly speaking, the test based on (3.10) performs better than (3.9), but the margin of superiority declines as N increases. For the cases examined here the tests based on second order polynomials generally had lower power than the first order polynomials, but there are exceptions. The tests, especially (3.10), have power against the alternative of misspecified marginal. In applied work it will be helpful to compute both versions since large divergences between the two may be indicative of misspecification.

7. EMPIRICAL APPLICATIONS

Two empirical applications, one cross sectional and the other time series will be used to illustrate the use of independence tests.

7.1 Application: Tests of independence in a multivariate Poisson model.

In Cameron et. al. (1988) a microeconometric model was estimated in which the focus was the relationship between several measures of health care utilization and the insurance status of the household, controlling for a variety of socioeconomic and health status variables. The data were a sample of 5190 observations for single-person households from the *Australian Health Survey* 1977-78. The dependent variables were counts of events including the number of hospital admissions (HOSPADM) and the number of days spent in hospitals (HOSPDAYS) in a preceding 12 month period; and the number of prescribed and nonprescribed medicines taken (PRESC and NONPRESC, respectively) in the past 2 days. The explanatory variables included gender, age, age-squared, income, three categories of insurance status (FREEPOOR, FREEOTHER, LEVYPLUS) with LEVYPLUS denoting the highest level of insurance, and a set of five health status variables (ACTDAYS, ILLNESS, HSCORE, CHCOND1, CHCOND2).

As is standard for count data regression the conditional mean of the dependent variable is log-linear in the explanatory variables. The details of the data and the specification are given in Cameron et. al. (1988). The 1988 paper reported and compared Poisson and negative binomial regressions for seven variables, including the four mentioned above, but treated each utilization measure as conditionally independent of other measures. However, we can expect some utilization measures to be jointly dependent. For instance, hospital admission (HOSPADM) and the number of prescribed medicines (PRESC) are likely to be correlated with each other and with other measures of health care utilization. For this reason, it was decided to reexamine the data and to test for dependence.

Three possible univariate count data models were considered: the Poisson and two alternative specifications of the negative binomial, respectively named NEGBIN 1 and NEGBIN 2 in Cameron and Trivedi (1986). For the Poisson, $\mathrm{var}(y\,|X) = \mathbf{E}[y\,|X]$; for NEGBIN 1, $\mathrm{var}(y\,|X) = \alpha\cdot\mathbf{E}[y\,|X]$, where α is the overdispersion parameter; and for NEGBIN 2, $\mathrm{var}(y\,|X) = (1 + \delta\cdot\mathbf{E}(y\,|X))\cdot\mathbf{E}[y\,|X]$, where δ is the overdispersion parameter. Tests of independence up to order two may be based on the following orthogonal polynomials:

	Poisson	NEGBIN 1	NEGBIN 2
$P_1(y \mid X)$	$y - \mu$	$y - \mu$	$y - \mu$
$P_2(y \mid X)$	$(y - \mu)^2 - y$	$(y - \mu)^2 - (2\alpha-1)(y-\mu) - \alpha\mu$	$(y - \mu)^2 - (1+2\delta\mu)(y - \mu) - (1+\delta\mu)\mu$

where μ denotes $\mathbf{E}[y \mid X] = \exp(X'\beta)$.

In implementing the tests we can apply the simpler asymptotic theory with (3.13) satisfied. Thus for each of the four count data regressions, β, δ and α may be replaced by consistent estimates and subsequently treated as given. The Poisson quasi-MLE $\hat{\beta}$ is consistent for β in all three models. Letting $\hat{\mu}_t = \exp(X_t'\hat{\beta})$, a consistent two step estimator of α is $\tilde{\alpha} = (\sum_t \hat{\mu}_t^2)^{-1} \cdot (\sum_t \hat{\mu}_t(y_t - \hat{\mu}_t)^2)$, and a consistent two step estimator of δ is $\tilde{\delta} = (\sum_t \hat{\mu}_t^4)^{-1} \cdot (\sum_t \hat{\mu}_t^2((y_t - \hat{\mu}_t)^2 - y_t))$ (Gourieroux et. al. (1984)).

Although all calculations were carried out, only those based on the NEGBIN 1 specification are reported below. These are preferred to the Poisson as the data are overdispersed. The Poisson quasi-MLE of β for each of the four regressions, and the estimate of the overdispersion parameter α, are given in Table 3 below. (The reported t-values for $\hat{\beta}$ are those from a standard Poisson ML package. To obtain correct standard errors, assuming overdispersion of the NEGBIN 1 form, divide the reported t-ratios by $\sqrt{\tilde{\alpha}}$.).

To calculate the test of independence for each of the six possible bivariate pairs of variables we use the statistic τ^2_{nm} in (3.9). In Table 4 below we restrict the calculations to all combinations of the first and second order NEGBIN 1 orthogonal polynomials since these appear to be sufficient to reject the null of independence. For approximate independence it is required that all four test statistics in any row of Table 4 be small. A sufficient condition for dependence is that the test statistic in column (1) be large.

There seems clear evidence that (HOSPADM,HOSPDAYS) and (PRESC,NONPRESC) are two dependent pairs of count variables, and also strong evidence that (HOSPADM,PRESC) is a dependent pair.

A feature of the tests based on orthogonal polynomials is that by focusing on higher order dependencies between variables they may highlight possible patterns of dependence between dispersion or volatility of variables even if they are uncorrelated. For this data, tests using the second order polynomials suggest that additionally (HOSPDAYS, PRESC) and (HOSPADM, NONPRESC) may be dependent pairs, a dependence not detected by tests using first order polynomials alone.

For comparison we also computed the tests based on orthonormal polynomials, using r^2_{nm} in (3.10), corresponding to column 1 of Table 4. In some cases these turned out to be somewhat different. For example, the six values corresponding to column (1) were 1660.8, 18.18, 0.25, 18.87, 0.003, and 126.7. These differences, taken in conjunction with the results from the Monte Carlo, suggest that the specification of the marginals should be scrutinized further, especially for HOSPDAYS variable.

The findings of dependence are very plausible for the following reason. The overdispersion evident in each equation may be due to unobserved heterogeneity. Consequently, the negative binomial regression was preferred to the Poisson. Since the explanatory variables in all equations are the same, it seems very plausible that the neglected (unobserved) heterogeneity in different equations is correlated, as in the bivariate Poisson model. This will impart stochastic dependence between variables.

Overall, there is strong evidence that we are dealing with dependent counts for all four variables and that joint estimation of these equations is desirable. Unfortunately, multivariate models for dependent count variables do not permit flexible patterns of correlation between the variables, though semiparametric estimation of a multivariate model may be feasible.

7.2 Application: Tests of serial dependence in the IBM stock price series.

Following Weiss (1986), we shall analyze the time series of IBM daily common stock closing price for 150 trading days beginning May 17, 1961, which is part of a longer series of 369 observations given in Box and Jenkins (1976, p. 526). The efficient market hypothesis implies that the time series of percentage stock price changes, $y_t = (1-B)\log P_t$, where B denotes the backward shift operator, should be serially independent. In practice, the ACF of the series may appear to be that of white noise and the errors uncorrelated, yet they may not be serially independent. Therefore it is useful to go beyond tests of zero correlation.

Row 1 in Table 5 gives the ACF of y_t. This suggests an MA1 model which is evidence against unqualified efficient market hypothesis. Estimation of a MA1 model yields the following:

$$(7.1) \qquad \underset{(1.56)}{(y_t - 0.0015)} = (1 + \underset{(3.23)}{0.2598}B)\varepsilon_t \quad ,$$

where t-statistics are given in parantheses. The ACF of $\hat{\varepsilon}_t$ from this MA1 model, given in row 2 of Table 5, suggests the process is serially uncorrelated. We tested for symmetry, and since the estimated third moment was extremely small, we tested for ARCH-M, bilinearity and ARCH1 effects after imposing symmetry and using the tests of this paper and the tests given by Weiss (1986). Table 6 gives the results.

The results obtained, under either a white noise or MA1 specification of the null model, reject the null against the alternative of ARCH-M, or against ARCH1 alternative. This shows that ignoring serial correlation and using OLS rather than MA1 residuals to test for ARCH-M and ARCH had no impact. This nicely illustrates the orthogonality between the tests. Ignoring ARCH, assuming symmetry and testing the restrictions (5.3)-(5.4), leads to rejection of (5.3), but not (5.4), irrespective of whether the error is assumed white noise or MA1. These results are the same whether the version of the test used is Weiss's or this paper's. Since this suggests that the conditional mean has been misspecified, the interpretation of the outcome of an ARCH1 test is ambiguous. Ignoring the ARCH-M effect nevertheless, and testing against ARCH1 also leads to the rejection of the null. We may conclude that ARCH-M and/or ARCH1 effects are present in the data. The fact that Weiss's bilinearity test gives results similar to ARCH-M also suggests that it is difficult to distinguish between the two effects using that test.

8. SUMMARY AND CONCLUDING REMARKS

The orthogonal polynomial approach to testing for stochastic dependence is attractive for several reasons. For models popular in applied econometrics, parametric specification of a multivariate system may be difficult or not known, in which case Wald, likelihood ratio and score tests cannot be used. And for other models a multivariate system may exist but be very restrictive in the form of dependence admitted. The tests proposed here are often simple to implement, and have an orthogonality property that is likely to be useful in applied work. In Monte Carlo simulation with discrete dependent variables the two main tests of this paper showed significant differences in power against empirically interesting alternatives, suggesting that both might be useful in applied work. In all cases considered at least one of the two types of tests had high power. In application the approach suggested in this paper tests the independence hypothesis only approximately by testing only a small subset of restrictions under the null hypothesis of independence. But despite this limitation the present procedure goes beyond the more usual test of zero correlation.

TABLE 1 : PERCENTAGE REJECTIONS OF THE NULL HYPOTHESIS OF INDEPENDENCE

N = 50

Model	Significance level %	T_{11}	T_{12}	T_{21}	T_{11}	T^n_{11}	T^n_{12}	T^n_{21}	T^n_{22}
1	1	1.2	0.2	0.2	0.4	1.0	0.8	0.4	5.9
	5	4.8	4.0	2.6	2.2	4.6	4.8	4.2	4.2
	10	10.8	10.2	6.6	7.0	10.8	10.2	9.4	9.2
2	1	0.4	0.4	0.6	0.2	1.8	1.2	0.8	1.0
	5	5.6	4.3	4.2	2.6	5.8	6.6	5.0	3.4
	10	12.0	9.8	10.6	6.0	12.0	10.6	11.6	8.0
3	1	0.4	0.0	0.8	0.4	39.2	49.0	48.2	50.2
	5	3.6	0.6	2.6	1.0	49.9	50.4	51.4	51.6
	10	15.6	7.8	8.8	2.8	54.0	52.0	56.2	52.6
4	1	10.0	0.4	0.4	0.2	18.6	1.0	2.4	4.8
	5	33.2	3.8	2.4	3.4	41.0	6.4	8.4	11.2
	10	45.4	9.6	8.6	8.8	54.2	13.4	14.2	21.6
5	1	18.8	0.4	0.2	0.6	34.4	1.4	1.6	8.8
	5	51.8	3.0	3.6	4.6	60.8	6.4	8.6	14.4
	10	67.0	9.2	7.0	9.0	72.6	13.4	14.2	21.6
6	1	5.2	0.0	0.0	0.2	80.6	70.8	69.8	76.4
	5	24.4	0.8	0.4	0.8	90.2	78.4	77.6	82.2
	10	44.8	4.8	4.2	3.8	92.8	84.2	81.8	84.0
7	1	47.4	0.0	0.0	0.0	65.0	2.0	2.4	14.0
	5	81.8	1.8	1.4	4.4	84.8	7.8	8.6	27.6
	10	91.4	4.0	4.4	15.0	92.0	14.0	14.2	33.6
8	1	0.2	2.0	0.6	17.4	0.0	0.0	0.0	0.0
	5	6.2	10.2	11.2	63.0	0.2	0.0	0.0	0.0
	10	19.4	28.0	32.0	80.8	2.0	0.0	0.0	0.4

TABLE 2: PERCENTAGE REJECTIONS OF THE NULL HYPOTHESIS OF INDEPENDENCE

N = 200

Model	Significance level %	T_{11}	T_{12}	T_{21}	T_{11}	T^n_{11}	T^n_{12}	T^n_{21}	T^n_{22}
					Test statistics				
1	1	0.6	1.4	0.4	1.6	1.2	2.2	0.8	1.0
	5	4.8	6.8	5.6	6.8	4.6	5.8	4.8	4.8
	10	9.2	13.4	8.6	11.4	8.0	11.6	10.2	9.6
2	1	0.6	1.2	0.8	1.0	1.0	1.0	1.2	0.6
	5	5.4	4.6	5.4	5.6	4.4	4.4	6.8	3.2
	10	9.8	9.8	11.4	10.4	8.2	9.0	12.4	8.2
3	1	35.6	0.6	1.0	0.8	44.0	24.4	13.0	23.6
	5	48.2	10.8	4.6	5.4	49.2	40.4	23.6	29.8
	10	52.4	32.4	15.6	11.4	51.2	48.0	29.2	33.0
4	1	52.8	0.4	1.0	0.4	65.8	1.0	2.4	5.4
	5	76.6	4.2	4.2	4.0	85.6	4.8	7.2	11.6
	10	87.2	8.6	9.0	9.8	91.2	9.2	12.4	17.6
5	1	61 8	0.4	0.4	0.8	70.6	1.0	1.2	8.8
	5	83.6	3.0	2.4	4.0	88.0	6.0	6.6	16.0
	10	89.2	8.4	7.2	10.6	99.4	12.8	13.4	20.6
6	1	26.8	0.6	0.8	0.0	84.4	74.4	75.6	74.6
	5	59.8	8.2	9.4	2.4	91.8	81.4	83.8	80.6
	10	77.4	25.8	24.6	10.4	95.4	93.8	87.6	83.6
7	1	74.2	0.4	0.6	0.4	73.8	0.2	1.0	5.8
	5	89.2	4.2	2.8	4.8	88.6	3.8	5.6	10.0
	10	94.6	8.8	7.8	11.4	93.6	8.2	10.8	16.4
8	1	2.2	8.6	9.2	32.2	0.2	0.0	0.0	0.0
	5	13.0	29.8	32.4	65.6	2.6	1.6	1.6	0.2
	10	19.4	46.0	47.2	80.0	9.2	5.8	8.6	1.6

TABLE 3: NEGBIN 1 REGRESSIONS FOR HEALTH CARE VARIABLES

5190 observations

	HOSPADM		HOSPDAYS	
Explanatory variable	**Coeff.**	**t-ratio**	**Coeff.**	**t-ratio**
ONE	-1.09	[-2.43]	0.28	[0.44]
SEX	-5.05e-02	[-0.60]	-0.33	[-2.30]
AGE	-7.20	[-4.07]	-8.46	[-3.18]
AGESQ	7.76	[3.89]	9.77	[3.38]
INCOME	-0.30	[-2.30]	-0.47	[-2.20]
FREEPOOR	0.22	[0.60]	0.59	[1.07]
FREEOTH	1.49e-02	[0.07]	0.41	[1.36]
LEVYPLUS	5.52e-02	[0.14]	0.51	[0.97]
ACTDAYS	7.99e-02	[7.28]	0.10	[5.81]
ILLNESS	0.10	[2.98]	0.11	[2.01]
HSCORE	5.18e-02	[2.74]	1.24e-02	[0.31]
CHCOND1	0.33	[3.35]	0.79	[5.21]
CHCOND2	0.81	[6.77]	1.40	[6.92]
$\tilde{\alpha}$	1.33	[21.26]	31.69	[16.56]

	PRESC		NONPRESC	
Explanatory variable	**Coeff.**	**t-ratio**	**Coeff.**	**t-ratio**
ONE	-3.09	[-9.94]	-2.66	[-7.50]
SEX	0.66	[12.59]	0.40	[7.28]
AGE	1.98	[2.63]	6.23	[5.71]
AGESQ	-0.74	[-0.90]	-7.56	[-5.83]
INCOME	6.60e-02	[0.92]	5.25e-02	[0.62]
FREEPOOR	0.80	[2.74]	2.38e-02	[8.32]
FREEOTH	0.40	[2.63]	-6.03e-02	[-0.36]
LEVYPLUS	0.53	[1.84]	2.78e-02	[9.81]
ACTDAYS	1.91e-02	[2.88]	2.42e-02	[2.05]
ILLNESS	0.19	[13.92]	0.18	[9.34]
HSCORE	6.75e-03	[0.68]	3.08e-02	[2.43]
CHCOND1	0.78	[14.85]	0.33	[5.64]
CHCOND2	1.06	[16.02]	0.30	[3.50]
$\tilde{\alpha}$	1.44	[23.44]	1.42	[39.15]

<u>Notes</u>: The four dependent variables are number of hospital admissions (HOSPADM), the number of hospital days (HOSPDAYS), number of prescribed medicines taken (PRESC) and number of nonprescribed medicines taken (NONPRESC). The explanatory variables are gender (sex), age, age-squared (AGESQ), income, insurance type (FREEPOOR, FREEOTHER, LEVYPLUS) and health status variables: activity days lost due to illness (ACTDAYS), whether ill (ILLNESS), score on a general health questionnaire (HSCORE), presence of limiting or nonlimiting chronic conditions (CHCOND1, CHCOND2). See Cameron et. al. for a detailed description of the data and the econometric model.

TABLE 4: PAIRWISE TESTS OF INDEPENDENCE OF HEALTH CARE VARIABLES

Pair	(1)	(2)	(3)	(4)
HOSPADM, HOSPDAYS	189.6	72.43	275.2	94.70
HOSPADM, PRESC	20.28	22.88	4.82	0.18
HOSPADM, NONPRESC	0.20	16.26	0.82	0.05
HOSPDAYS, PRESC	0.18	10.06	1.09	9.23
HOSPDAYS, NONPRESC	0.01	1.91	0.55	0.16
PRESC, NONPRESC	9.20	9.85	4.35	4.07

Notes : This table gives the test statistic τ^2_{mn}; m,n = 1 in column (1), m,n = 2 in column (2), m = 1, n = 2, in column (3), and m = 2, n = 1, in column (4).

TABLE 5 : AUTOCORRELATIONS - IBM DATA

	Lag					
Estimator	1	2	3	4	5	6
OLS	.207	-.121	-.090	.124	.076	-.037
MA1	-.013	-.093	-.100	.135	.050	-.045

TABLE 6: TESTS OF INDEPENDENCE FOR IBM DATA

Null model	Alternative model[1]	Variant	Test statistic[2]
(0,1,1,0)	(0,1,1,1)	Weiss	12.99
(0,1,1,0)	(0,1,1,1)	This paper	12.07, 0.10
(0,1,1,0)	(0,1,1,0) + ARCH1	Weiss This paper	12.08
(0,1,0,0)	(0,1,0,1)	Weiss	14.53
(0,1,0,0)	(0,1,0,1)	This paper	11.21, 1.11
(0,1,0,0)	(0,1,0,0) + ARCH1	Weiss This paper	10.20

Notes: (1): Using standard nomenclature we refer to an ARMA (p,d,q) model with the hypothesized first order diagonal bilinear process a (p,d,q,1) model.

(2): Where two numbers are given, the first is a test of ARCH-M dependence and the second a test of non-symmetric ARCH.

Appendix A

(i) Orthogonal polynomials for distributions in the Meixner class have a generating function of the form

(A.1) $$\Gamma(y;z) = \sum_{n=0}^{\infty} P_n(y)z^n/n! = t(z)\exp(yu(z))$$

where t(z) and u(z) are functions with power series expansion in z (Meixner (1934)). Table A.1 presents the polynomial generating functions for the Meixner class and identifies the polynomial family by name.

Table A.1: Orthogonal polynomial functions for selected members of LEF-QVF

Distribution	*Generating Function:* $\Gamma(y;z)$
$(2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}\cdot(y-\mu)^2)$, $-\infty < y < \infty$ *(Normal)*	$\exp\{(y-\mu)z - \frac{1}{2}\sigma^2 z^2\}$ *(Hermite)*
$\frac{y^{\alpha-1}}{\Gamma(\alpha)}\exp(-y)$, $y>0$, $\alpha>0$ *(Gamma)*	$(1+z)^{-\alpha}\exp[yz/(1+z)]$ *(Generalized Laguerre)*
$\lambda^y e^{-\lambda}/y!$, $y=0,1,\ldots$ *(Poisson)*	$(1+z)^y e^{-\lambda z}$ *(Poisson-Charlier)*
$(1+\theta)^{-\alpha-y}\,\theta^y \binom{y+\alpha-1}{y}$ *(Negative binomial)*	$(1+z\theta)^{-y-\alpha}(1+z(1+\theta))^y$ *(Meixner)*
$\binom{n}{y} p^y(1-p)^{n-y}$ *(Binomial)*	$(1+(1-p)z)^y(1-pz)^{n-y}$ *(Krawtchouk)*

REFERENCES

Ashford, J.R., and R.R. Sowden (1970), "Multivariate Probit Analysis", *Biometrics*, 26, 535-546.

Brock, W. A., D. A. Hsieh, and B. LeBaron (1991), *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, MIT Press: Cambridge.

Brock, W. A., and S. Potter (1991), "Nonlinear Time Series and Macroeconometrics", unpublished paper.

Bollerslev, T., R. Y. Chou and K. F. Kroner (1990), "ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence", *Journal of Econometrics*, 52, 5-59.

Box, G. E. P., and G. M. Jenkins (1976), *Time Series Analysis: Forecasting and Control* (2nd Edition). San Fransisco: Holden Day.

Breusch, T., and A.R. Pagan (1980). "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics", *Review of Economic Studies*, 47, 239-253.

Cameron, A.C., and P.K., Trivedi (1986), "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests", *Journal of Applied Econometrics*, 1, 29-53.

__________________ (1990), "Conditional Moment Tests and Orthogonal Polynomials", Indiana University Working Papers in Economics, #90-051.

Cameron, A.C., P. K. Trivedi, F. Milne and J. Piggott (1988). "A Microeconometric Model of the Demand for Health Care and Health Insurance Australia", *Review of Economic Studies*, 55, 85-106.

Cox, D. R., and E. J. Snell (1968), "A General Definition of Residuals", *Journal of the Royal Statistical Society*, B, 30, 248-275.

Cramer, H. (1946), *Mathematical Methods of Statistics*, Princeton: Princeton

University Press.

Eagleson, G. K. (1964), "Polynomial Expansions of Bivariate Distributions", *Annals of Mathematical Statistics*, 35, 1208-1215.

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, 987-1007.

Engle, R. F., D. F. Hendry and J-F. Richard (1983), "Exogeneity", *Econometrica*, 51, 277-304.

Engle, R.F., D. Lilien and R. Robins (1987), "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model", *Econometrica*, 55, 391-407.

Griffiths, R.C. (1985), "Orthogonal Expansions", in Kotz, S. and N. L. Johnson, editors, *Encyclopaedia of Statistical Sciences*, 6, New York: John Wiley, 530-536.

Gourieroux, C., A. Monfort and A. Trognon (1984), "Pseudo Maximum Likelihood Methods: Applications to Poisson Models", *Econometrica*, 52, 701-720.

Granger, C. J. W., and A. P. Andersen (1978), *An Introduction to Bilinear Time Series Models*, Göttingen: Vandenhoeck & Ruprecht.

Hall, A., (1990), "Lagrange Multiplier Tests for Normality Against Seminonparametric Alternatives", *Journal of Business and Economic Statistics*, 8(4), 417-426.

Hooper, J. W. (1959), "Simultaneous Equations and Canonical Correlation Theory", *Econometrica*, 27, 245-56.

Hsieh, D. A. (1989), "Testing for Nonlinear Dependence in Daily Foreign Exchange Rates", *Journal of Business*, 62, 339-368.

Johnson, N. L. and S. Kotz (1969), *Discrete Distributions*, Boston: Houghton Mifflin Co.

Jupp, P. E. and K. V. Mardia (1980), "A General Correlation Coefficient for

Directional Data and Related Regression Problems", *Biometrika*, 67, 163-173.

Lancaster, H. O. (1958), "The Structure of Bivariate Distributions", *Annals of Mathematical Statistics*, 29, 719-736.

Lancaster, H. O. (1963), "Correlations and Canonical Forms of Bivariate Distributions", *Annals of Mathematical Statistics*, 34, 532-538.

Lancaster, H. O. (1969), *The Chi-squared Distribution*, New York: John Wiley.

Meixner, J. (1934), "Orthogonale Polynomsysteme mit einer besonderen Gestalt der erzeugenden Funktion", *Journal of London Mathematical Society*, 9, 6-13.

Nelson, D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach", *Econometrica*, 59, 347-370.

Newey, W.K. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests", *Econometrica*, 53, 1047-1070.

Ord, J. K. (1972) *Families of Frequency Distributions*, London: Charles Griffin & Co.

Pagan, A.R. and F. Vella (1989), "Diagnostic Tests for Models Based on Individual Data: A Survey", *Journal of Applied Econometrics*, Supplement, S4, S29-S59.

Perez-Amaral, T. (1989), "M-tests and Dynamic Information Matrix Tests with an Application to Linear Regression", Discussion paper 75, Universidad Complutense, Madrid, Spain.

Robinson, P. (1991), "Consistent Nonparametric Entropy-based Testing", *Review of Economic Studies*, 58, 437-453.

Shiba, T., and H. Tsurumi (1988), "Bayesian and Non-Bayesian Tests of Independence in Seemingly Unrelated Regressions", *International Economic Review*, 29, 377-395.

Tauchen, G. (1985), "Diagnostic Testing and Evaluation of Maximum Likelihood

Models", *Journal of Econometrics*, 30, 415,443.

Weiss, A. A. (1986), "ARCH and Bilinear Time Series Models: Comparison and Combination", *Journal of Business & Economic Statistics*, 4, 1, 59-70.

White, H. (1987), "Specification Testing in Dynamic Models", in Bewley, T. F., editor, *Advances in Econometrics, Fifth World Congress Volume 1*, 1-58.

Acknowledgment

The authors thank two anonymous referees and the Econometrics workshop participants at the University of Chicago and at Rice University for helpful suggestions for improvements of earlier drafts.

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
9116	N. Rankin	Exchange Rate Risk and Imperfect Capital Mobility in an Optimising Model
9117	E. Bomhoff	Currency Convertibility: When and How? A Contribution to the Bulgarian Debate!
9118	E. Bomhoff	Stability of Velocity in the G-7 Countries: A Kalman Filter Approach
9119	J. Osiewalski and M. Steel	Bayesian Marginal Equivalence of Elliptical Regression Models
9120	S. Bhattacharya, J. Glazer and D. Sappington	Licensing and the Sharing of Knowledge in Joint Ventures
9121	J.W. Friedman and L. Samuelson	An Extension of the "Folk Theorem" with Continuous Reaction Functions
9122	S. Chib, J. Osiewalski and M. Steel	A Bayesian Note on Competing Correlation Structures in the Dynamic Linear Regression Model
9123	Th. van de Klundert and L. Meijdam	Endogenous Growth and Income Distribution
9124	S. Bhattacharya	Banking Theory: The Main Ideas
9125	J. Thomas	Non-Computable Rational Expectations Equilibria
9126	J. Thomas and T. Worrall	Foreign Direct Investment and the Risk of Expropriation
9127	T. Gao, A.J.J. Talman and Z. Wang	Modification of the Kojima-Nishino-Arima Algorithm and its Computational Complexity
9128	S. Altug and R.A. Miller	Human Capital, Aggregate Shocks and Panel Data Estimation
9129	H. Keuzenkamp and A.P. Barten	Rejection without Falsification - On the History of Testing the Homogeneity Condition in the Theory of Consumer Demand
9130	G. Mailath, L. Samuelson and J. Swinkels	Extensive Form Reasoning in Normal Form Games
9131	K. Binmore and L. Samuelson	Evolutionary Stability in Repeated Games Played by Finite Automata
9132	L. Samuelson and J. Zhang	Evolutionary Stability in Asymmetric Games

No.	Author(s)	Title
9133	J. Greenberg and S. Weber	Stable Coalition Structures with Uni-dimensional Set of Alternatives
9134	F. de Jong and F. van der Ploeg	Seigniorage, Taxes, Government Debt and the EMS
9135	E. Bomhoff	Between Price Reform and Privatization - Eastern Europe in Transition
9136	H. Bester and E. Petrakis	The Incentives for Cost Reduction in a Differentiated Industry
9137	L. Mirman, L. Samuelson and E. Schlee	Strategic Information Manipulation in Duopolies
9138	C. Dang	The D'_2-Triangulation for Continuous Deformation Algorithms to Compute Solutions of Nonlinear Equations
9139	A. de Zeeuw	Comment on "Nash and Stackelberg Solutions in a Differential Game Model of Capitalism"
9140	B. Lockwood	Border Controls and Tax Competition in a Customs Union
9141	C. Fershtman and A. de Zeeuw	Capital Accumulation and Entry Deterrence: A Clarifying Note
9142	J.D. Angrist and G.W. Imbens	Sources of Identifying Information in Evaluation Models
9143	A.K. Bera and A. Ullah	Rao's Score Test in Econometrics
9144	B. Melenberg and A. van Soest	Parametric and Semi-Parametric Modelling of Vacation Expenditures
9145	G. Imbens and T. Lancaster	Efficient Estimation and Stratified Sampling
9146	Th. van de Klundert and S. Smulders	Reconstructing Growth Theory: A Survey
9147	J. Greenberg	On the Sensitivity of Von Neuman and Morgenstern Abstract Stable Sets: The Stable and the Individual Stable Bargaining Set
9148	S. van Wijnbergen	Trade Reform, Policy Uncertainty and the Current Account: A Non-Expected Utility Approach
9149	S. van Wijnbergen	Intertemporal Speculation, Shortages and the Political Economy of Price Reform
9150	G. Koop and M.F.J. Steel	A Decision Theoretic Analysis of the Unit Root Hypothesis Using Mixtures of Elliptical Models

No.	Author(s)	Title
9151	A.P. Barten	Consumer Allocation Models: Choice of Functional Form
9152	R.T. Baillie, T. Bollerslev and M.R. Redfearn	Bear Squeezes, Volatility Spillovers and Speculative Attacks in the Hyperinflation 1920s Foreign Exchange
9153	M.F.J. Steel	Bayesian Inference in Time Series
9154	A.K. Bera and S. Lee	Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis
9155	F. de Jong	A Univariate Analysis of EMS Exchange Rates Using a Target
9156	B. le Blanc	Economies in Transition
9157	A.J.J. Talman	Intersection Theorems on the Unit Simplex and the Simplotope
9158	H. Bester	A Model of Price Advertising and Sales
9159	A. Özcam, G. Judge, A. Bera and T. Yancey	The Risk Properties of a Pre-Test Estimator for Zellner's Seemingly Unrelated Regression Model
9160	R.M.W.J. Beetsma	Bands and Statistical Properties of EMS Exchange Rates: A Monte Carlo Investigation of Three Target Zone Models Zone Model
9161	A.M. Lejour and H.A.A. Verbon	Centralized and Decentralized Decision Making on Social Insurance in an Integrated Market Multilateral Institutions
9162	S. Bhattacharya	Sovereign Debt, Creditor-Country Governments, and
9163	H. Bester, A. de Palma, W. Leininger, E.-L. von Thadden and J. Thomas	The Missing Equilibria in Hotelling's Location Game
9164	J. Greenberg	The Stable Value
9165	Q.H. Vuong and W. Wang	Selecting Estimated Models Using Chi-Square Statistics
9166	D.O. Stahl II	Evolution of Smart$_n$ Players
9167	D.O. Stahl II	Strategic Advertising and Pricing with Sequential Buyer Search
9168	T.E. Nijman and F.C. Palm	Recent Developments in Modeling Volatility in Financial Data
9169	G. Asheim	Individual and Collective Time Consistency
9170	H. Carlsson and E. van Damme	Equilibrium Selection in Stag Hunt Games
9201	M. Verbeek and Th. Nijman	Minimum MSE Estimation of a Regression Model with Fixed Effects from a Series of Cross Sections

No.	Author(s)	Title
9202	E. Bomhoff	Monetary Policy and Inflation
9203	J. Quiggin and P. Wakker	The Axiomatic Basis of Anticipated Utility; A Clarification
9204	Th. van de Klundert and S. Smulders	Strategies for Growth in a Macroeconomic Setting
9205	E. Siandra	Money and Specialization in Production
9206	W. Härdle	Applied Nonparametric Models
9207	M. Verbeek and Th. Nijman	Incomplete Panels and Selection Bias: A Survey
9208	W. Härdle and A.B. Tsybakov	How Sensitive Are Average Derivatives?
9209	S. Albæk and P.B. Overgaard	Upstream Pricing and Advertising Signal Downstream Demand
9210	M. Cripps and J. Thomas	Reputation and Commitment in Two-Person Repeated Games
9211	S. Albæk	Endogenous Timing in a Game with Incomplete Information
9212	T.J.A. Storcken and P.H.M. Ruys	Extensions of Choice Behaviour
9213	R.M.W.J. Beetsma and F. van der Ploeg	Exchange Rate Bands and Optimal Monetary Accommodation under a Dirty Float
9214	A. van Soest	Discrete Choice Models of Family Labour Supply
9215	W. Güth and K. Ritzberger	On Durable Goods Monopolies and the (Anti-) Coase-Conjecture
9216	A. Simonovits	Indexation of Pensions in Hungary: A Simple Cohort Model
9217	J.-L. Ferreira, I. Gilboa and M. Maschler	Credible Equilibria in Games with Utilities Changing during the Play
9218	P. Borm, H. Keiding, R. Mclean, S. Oortwijn and S. Tijs	The Compromise Value for NTU-Games
9219	J.L. Horowitz and W. Härdle	Testing a Parametric Model against a Semiparametric Alternative
9220	A.L. Bovenberg	Investment-Promoting Policies in Open Economies: The Importance of Intergenerational and International Distributional Effects
9221	S. Smulders and Th. van de Klundert	Monopolistic Competition, Product Variety and Growth: Chamberlin vs. Schumpeter

No.	Author(s)	Title
9222	H. Bester and E. Petrakis	Price Competition and Advertising in Oligopoly
9223	A. van den Nouweland, M. Maschler and S. Tijs	Monotonic Games are Spanning Network Games
9224	H. Suehiro	A "Mistaken Theories" Refinement
9225	H. Suehiro	Robust Selection of Equilibria
9226	D. Friedman	Economically Applicable Evolutionary Games
9227	E. Bomhoff	Four Econometric Fashions and the Kalman Filter Alternative - A Simulation Study
9228	P. Borm, G.-J. Otten and Hans Peters	Core Implementation in Modified Strong and Coalition Proof Nash Equilibria
9229	H.G. Bloemen and A. Kapteyn	The Joint Estimation of a Non-Linear Labour Supply Function and a Wage Equation Using Simulated Response Probabilities
9230	R. Beetsma and F. van der Ploeg	Does Inequality Cause Inflation? - The Political Economy of Inflation, Taxation and Government Debt
9231	G. Almekinders and S. Eijffinger	Daily Bundesbank and Federal Reserve Interventions - Do they Affect the Level and Unexpected Volatility of the DM/$-Rate?
9232	F. Vella and M. Verbeek	Estimating the Impact of Endogenous Union Choice on Wages Using Panel Data
9233	P. de Bijl and S. Goyal	Technological Change in Markets with Network Externalities
9234	J. Angrist and G. Imbens	Average Causal Response with Variable Treatment Intensity
9235	L. Meijdam, M. van de Ven and H. Verbon	Strategic Decision Making and the Dynamics of Government Debt
9236	H. Houba and A. de Zeeuw	Strategic Bargaining for the Control of a Dynamic System in State-Space Form
9237	A. Cameron and P. Trivedi	Tests of Independence in Parametric Models: With Applications and Illustrations

P.O. BOX 90153. 5000 LE TILBURG. THE NETHERLANDS