MANCHESTER
1824

The University
of Manchester

Economics
Discussion Paper

EDP-0606

On the sensitivity of Kernel-based Conditional Moment Tests to Unconsidered Local Alternatives

by E. Fe-Rodríguez and C. Orme

March 2006

Correspondence email: e.ferodriguez@manchester.ac.uk

School of Social Sciences,
The University of Manchester
Oxford Road
Manchester M13 9PL
United Kingdom

Combining the strengths of UMIST and
The Victoria University of Manchester

On the sensitivity of Kernel-based Conditional Moment Tests to Unconsidered Local Alternatives

Eduardo Fe-Rodríguez*and Chris D. Orme
Economic Studies
The University of Manchester
Manchester, M13 9PL
United Kingdom

March 16, 2006

Abstract

This article investigates the sensitivity of kernel-based conditional moment tests to unconsidered misspecifications in the main parametric model under evaluation. We establish a relationship between the asymptotic insensitivity of a test to a particular source of misspecification, and its asymptotic independence with the corresponding kernel-based score test, of the restrictions imposed by the uconsidered misspecification.

*Corresponding author. This paper is part of Eduardo Fe's Ph.D. dissertation. He thanks Professor Javier Hidalgo and Simon Peters for helpful comments. Email: e.ferodriguez@manchester.ac.uk

1 Introduction

This article studies the sensitivity of kernel-based test to the so called *unconsidered local alternatives*. The term was introduced by Godfrey and Orme (1996) to refer to any alternative hypothesis locally distinct from that *implicit alternative* against which a particular test statistic has highest power (see Davidson and MacKinnon, 1987). These departures are local in a Pitman sense (Pitman, 1949), so that they will vanish asymptotically to a suitable rate, to prevent the power of the test converging to one as the sample size approaches infinity. In the framework of this article, the rate at which these alternatives converge to zero will be slower than the square root of the sample size.

The behavior of kernel-based conditional moment tests under unconsidered alternatives is of great interest. When the true process that generated the data is known, tests can be constructed so as to be robust against certain types of misspecification (see Bera and Yoon, 1993). In practice full knowledge of the generating process is hardly ever available, and therefore a test thought to be robust to a type of misspecification might not be so. Generally speaking, Conditional Moment tests can be constructed to work either as a test for general misspecification, or they can be devised to isolate and detect a particular type of misspecification (such as asymmetry or heteroskedasticity). If the test is intended to be a very general test for misspecification, then it is expected that the test is able to detect any arbitrary type of discrepancy between model and generating process, including unconsidered local alternatives. However, if the test is designed to isolate a particular type of misspecification, then it is desirable that the specification test be insensitive to local misspecification for which it was not designed. For example, one would want that a consistent conditional moment test for conditional heteroskedasticity be not affected by local misspecifications in the conditional mean of the distribution. Therefore, it is important to have a characterization of the conditions under which an arbitrary test is affected by unconsidered local alternatives.

In this article we are going to focus on the class of kernel-based, Consistent Conditional Moment Tests. Members of these class are Zheng (1996) and Fan and Li's (1996) test for regression analysis, Zheng's (2000) nonparametric tests for conditional symmetry in a density function, or Hsiao and Li (2001) test for conditional heteroskedasticity, to mention but a few. These are all tests of a single moment condition. However Delgado, Dominguez and Lavergne (2001) show that the method can be extended to evaluate several restrictions simultaneously, and we adopt their approach in this article. In general all these tests can be understood as nonparametric versions of Newey's (1985) class of asymptotic chi-square Conditional Moment tests. Unlike these methods, their nonparametric counterparts have an asymptotic normal distribution and they are consistent against any type of misspecification. However, a stylized fact about these tests, is that they are undersized, and their performance can be disappointing for moderate samples, specially if the number of conditioning variables is large. On the other hand, bootstrap methods have been successfully employed to correct these deficiencies (see, for example, Li and Wang, 1998), so that consistent methods remain a very attractive instrument for model discrimination.

This article characterizes under what circumstances a kernel-based Conditional Moment test will be affected by a type of unconsidered alternative. Our analysis can be understood as an analysis of asymptotic relative efficiency for the type of tests under consideration, by constructing a generating process dependent on one or several sequences of Pitman alternatives. However, we split the parameter space to allow the consideration to two important situations. Firstly, we could examine the behavior of a test when there are different types of departures from the null model. Thus, if we were testing for heteroskedasticity, our framework would allow to study the properties of the test when there are departures such as non-normality, omission of variables in the conditional mean, autocorrelation and so on. Secondly, the settings in this article allow to study the impact of underparameterisation. Finally, we provide the reader with a very simple check to detect when a particular kernel-based conditional moment test will be sensitive to a certain type of unconsidered alternative. As will be explained, this check relates the asymptotic insensitivity of a test with its asymptotic independence with a kernel-based *score* test of the restrictions implied by the unconsidered local alternatives. A number of examples are provided on how to use our check in practice.

The structure of the paper is as follows. Section 2 sets the framework for our analysis and introduces the main result. Conditions are maintained very general, so that our analysis applies to cross-section and time-series data. Section 3 establishes a relationship between the sensitivity of kernel-based tests and a nonparametric score test of the significance of the local misspecification. Section 4 provides some examples on how to implement our results in practice. Section 5 is the closing section and contains some final remarks. Throughout the paper, we use the following notation: $\sum_i = \sum_{i=1}^N$, $\sum_{i_1,\cdots,i_m} = \sum_{i_1}\cdots\sum_{i_m} = \sum_{i_1=1}^N\cdots\sum_{i_m=1}^N$, $\sum_{(i,j)} = \sum_{i=1}^N\sum_{j\neq i,j=1}^N$, etc. For any real number a, $|a| = \sqrt[+]{a^2}$ will denote its absolute value, whilst for and arbitrary vector, $\mathbf{a}$, or matrix, $\mathbf{A}$, $\|\mathbf{a}\| = \sqrt[+]{\mathbf{a}'\mathbf{a}}$, and $\|\mathbf{A}\| = \sqrt[+]{tr\left(\mathbf{A}'\mathbf{A}\right)}$.

2 Assumptions and Limit Distribution of the Test Statistic.

Suppose that an econometrician is interested in studying the validity of an particular model, and a random sample from a population $\{W_i\}_{i=1}^N = \{Y_i, X_i', Z_i'\}_{i=1}^N$ is available, where $y \in \mathcal{Y} \subseteq \mathbb{R}$, $x \in \mathcal{X} \subseteq \mathbb{R}^{d_x}$, $z \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$ and, therefore, $w \in \mathcal{W} = \mathcal{Y} \times \mathcal{X} \times \mathcal{Z} \subseteq \mathbb{R} \times \mathbb{R}^{d_x} \times \mathbb{R}^{d_z}$. For simplicity of exposition it is assumed that Y is the dependent variable of interest, which may be modelled conditionally on X and Z, and X is a continuous random variable with probability density function $f(x)$, although this could be weakened (see for example, Raccine and Li (2004)). Following Godfrey and Orme (1996, Section 3), it will be assumed that the Data Generation Process (DGP) is characterized by a finite dimensional parameter vector, $\theta \in \Theta$, with true value denoted by θ^*. The variables X and Z (or some subset of them) may inform estimation of the unknown parameter vector, whereas only X will be employed in the construction of the Kernel-based conditional moment tests. In general, however, the functional form for the conditional distribution of Y given $X = x$, $Z = z$, is unknown and estimators are obtained by maximizing some estimation criterion denoted

$$Q_N(\theta) \equiv Q_N(\theta; W_1, ..., W_N).$$

Clearly, $Q_N(\theta)$ accommodates maximum likelihood estimation (if the conditional distribution of Y were to be specified up to a knowledge of the unknown parameter vector) Non-Linear Least Squares and Generalized Method of Moments (GMM) (in which case $-Q_N(\theta)$ would be the appropriate quadratic form to optimize). It is not the purpose of this paper to dwell on the regularity conditions on $Q_N(\theta)$ which underpin consistent estimation of θ^*; rather, it is assumed that the implied estimator has the usual *root-n* consistency properties.

In order to undertake the sensitivity analysis of the test statistic to unconsidered local alternatives, the null model being estimated and tested needs to be defined and to do so it will be convenient to partition the unknown parameter vector as $\theta' = (\theta_0, \theta_1, ..., \theta_L)$, with $\theta_p \in \Theta_p$, $p = 0, 1, \cdots, L$, being $(d_p \times 1)$, and $d = \sum_{p=0}^{L} d_p$. The null model is then obtained by imposing the $(d - d_0)$ restrictions of $\mathcal{H}$: $\theta_1 = \theta_2 = ... = \theta_L = 0$, which provides the estimator $\hat{\theta}' = \left(\hat{\theta}_0, 0', ..., 0'\right)$ as a solution to

$$\max_{\theta_0} Q_N(\theta_0, 0, ..., 0).$$

It is important to recognize that $\mathcal{H}$ only imposes restrictions on θ and does not, in itself, imply anything about the conditional distribution of Y given X, under these restrictions, such as conditional moments.

The adequacy of this null model is to be tested using the framework of, for example, Zheng (1996), Delgado,Dominguez and Lavergne (2001) or Hsiao and Li (2001), in which the *null hypothesis* specifies a set of R conditional moment restrictions, in addition to the parametric restrictions of $\mathcal{H}$. Specifically, a set of R random functions, denoted $m(W;\theta)' = \left(m^{(1)}(W;\theta), ..., m^{(R)}(W;\theta)\right)'$, are defined such that the null and alternative hypotheses are, respectively,

$$H_0 \quad : \quad \Pr\left[E_{\mathcal{H}}\left[m\left(W,\theta^*\right)|X=x\right]=0\right]=1, \text{ for some } \theta_0^* \in \Theta_0 \tag{1}$$

$$H_1 \quad : \quad \Pr\left[E_{\mathcal{H}}\left[m\left(W,\theta\right)|X=x\right]=0\right]<1, \text{ for all } \theta_0 \in \Theta_0. \tag{2}$$

Note that $E_{\mathcal{H}}[.]$ implies that expectations are taken under the *parametric* restrictions of the *null model*, $\mathcal{H}$, so that the last $(d-d_0)$ elements in θ^* (and θ, respectively) are set to zero in (1) (and (2), respectively). The test procedure is a consistent test of H_0 against H_1, in that it will reject H_0 with probability approaching 1 as $N \to \infty$, if H_1 is true (where H_1 simply states that H_0 is false) within the restrictions imposed by $\mathcal{H}$. The test statistic employed is Kernel-based and defined by

$$\hat{T}_N = Nh^{d/2}\frac{\frac{1}{N(N-1)}\sum_r\sum_{(i,j)}t_{ij}^{(r)}(\hat{\theta})}{\sqrt{\Sigma(\hat{\theta}_0)}}$$

where: $t_{ij}^{(r)}(\hat{\theta}) = K_{h,ij}m^{(r)}(W_i,\hat{\theta})m^{(r)}(W_j,\hat{\theta})$, $K_{h,ij} = \frac{1}{h^d}K\left(\frac{x_i-x_j}{h}\right)$, $K(\zeta)$ is a Kernel function, h a bandwidth parameter and $\Sigma(\hat{\theta}_0)$ is any consistent estimator of

$$\Sigma(\theta_0^*) = 2\sum_{r,s}E\left[\left\{\mathcal{M}^{(rs)}\left(X\right)\right\}^2 f\left(X\right)\right]\int K^2\left(\zeta\right)d\zeta,$$

where $\mathcal{M}^{(rs)}\left(x\right) = E_{\mathcal{H}}\left[m^{(r)}(W;\theta^*)m^{(s)}(W;\theta^*)|X=x\right]$, $f(x)$ is the density of X, and notice again that expectations respect the parametric restrictions of $\mathcal{H}$. One such consistent estimator is $\Sigma(\hat{\theta}_0) = \frac{2h^d}{N(N-1)}\sum_{r,s}\sum_{(i,j)}t_{ij}^{(r)}(\hat{\theta})t_{ij}^{(s)}(\hat{\theta})$; see Delgado, Dominguez and Lavergne (2001). Under H_0, and various primitive regularity conditions depending on the nature of the data available (see, Zheng, 1996, Delgado, Dominguez and Lavergne, 2001, or Hsiao and Li, 2001, Racine and Li, 2004), $\hat{T}_N \xrightarrow{d} N(0,1)$, whilst under H_1, $\Pr\left(\hat{T}_N > \delta_N\right) \to 1$ for any positive stochastic sequence $\delta_N = o\left(Nh^{d/2}\right)$, which implies that the corresponding one-sided test procedure is consistent. An asymptotically equivalent version of the test statistic can be expressed, rather simply, as

$$\hat{T}_N^{\#} = \frac{\sum_r \sum_{(i,j)} t_{ij}^{(r)}(\hat{\theta})}{\sqrt{2\sum_{r,s}\sum_{(i,j)} t_{ij}^{(r)}(\hat{\theta}) t_{ij}^{(s)}(\hat{\theta})}}. \tag{3}$$

Following Godfrey and Orme (1996), by allowing $\theta_p \neq 0$, $p = 1, ..., L$, (in a manner described below), we will be able to characterize the limiting distribution of the test statistic associated with testing H_0 under sources of misspecification indexed by θ_p. Two scenarios can be investigated. Firstly, the behavior of the test statistic based on $V_N(\hat{\theta}) = \frac{1}{N(N-1)} \sum_r \sum_{(i,j)} t_{ij}^{(r)}(\hat{\theta})$ in the presence one or more departures from the null model defined by the restrictions of $\mathcal{H}$. For example, consider the null model which specifies $Y_i = X_{i1}'\theta_{01} + U_i$, $E_{\mathcal{H}}[U|X] = 0$ and $E_{\mathcal{H}}[U^2|X] = \theta_{02}$, where X_{i1} is a sub-vector of X_i and θ_{0j} is a sub-vector of θ_0, $j = 1, 2$. A consistent test for heteroskedasticty (see, for example, Delgado, Dominguez and Lavergne, 2001) can be analyzed in situations where $Y_i = X_{i1}'\theta_{01} + l(X_i)'\theta_1 + U_i$, where $l(X)$ is some vector function of X. Secondly, the impact of underparameterisation of the relevant type of departure can be analyzed. Thus, in the previous example, suppose that the alternative is $Y_i = X_{i1}'\theta_{01} + g(Z_i)'\theta_1 + U_i$, but that the test procedure is designed to check $E_{\mathcal{H}}[U|X] = 0$ (so that the wrong set of conditioning variables are employed in the Kernel-based test for misspecified functional form).

In order undertake this analysis, some additional structure is required. Firstly, it is assumed that

$$\Pr\left[E\left[m(W,\theta^*)|X=x\right]=0\right]=1,\ \text{for some } \theta^*\in\Theta \tag{4}$$

so that the set of R random functions do indeed have zero conditional mean when all possible sources of misspecification are accounted for. Thus, in the first example of the preceding paragraph (with $R=1$), $m^{(1)}(W;\theta)=\left(Y_i-X_{i1}'\theta_{01}-l(X_i)'\theta_1\right)^2-\theta_{02}$, whilst under the restrictions of $\mathcal{H}$, $m^{(1)}(W;\theta)=\left(Y_i-X_{i1}'\theta_{01}\right)^2-\theta_{02}$. Secondly, for $p=1,2,...,L$, $\theta_p^*=\gamma_p/\sqrt{Nh^{d/2}}$, where γ_p, $(d_p\times 1)$, satisfies that $0\leq\|\gamma_p\|<\infty$ and $\gamma_p\neq 0$ for at least one p. This defines a sequence of local unconsidered alternatives. In general, standard first order asymptotic theory reveals that $\hat{\theta}_0-\theta_0^*=O_p(N^{-1/2})$, under such a sequence.

In this paper we study the behavior of the test indicator

$$V_N(\hat{\theta})=\frac{1}{N\left(N-1\right)}\sum_r\sum_{(i,j)}t_{ij}^{(r)}(\hat{\theta})$$

under the framework described above. In order to undertake our analysis we employ some high level assumptions which justify the limiting distribution results and various mean value expansions employed.

Firstly, we assume that the Kernel satisfies standard conditions such that $K:\mathbb{R}^d\rightarrow\mathbb{R}$ is bounded and symmetric, with $\int K\left(\zeta\right)d\zeta=1$, $\int K\left(\zeta\right)\|\zeta\|\,d\zeta<\infty$. We require a central limit theorem for the second order degenerate *U-statistic*, $V_N(\theta^*)=\frac{1}{N(N-1)}\sum_r\sum_{(i,j)}t_{ij}^{(r)}(\theta^*)$. This is given by Hall(1984) and de Jong (1987) for independent data and generalized by Fan and Li (1999) for the weakly dependent data case. In particular, it is often assumed that $h \to 0$, $Nh^d \to \infty$ as $N \to \infty$, in the case of independent data, but Fan and Li (1999) require slightly stronger restrictions on h for the weakly dependent data case. These results imply that under the sequence of local alternatives

$$Nh^{d/2}V_N(\theta^*) \xrightarrow{d} N\left(0, \Sigma(\theta_0^*)\right).$$

In order to justify the second order mean value expansion employed in deriving the limiting distribution result, the following will suffice and are standard (see for example proofs in Zheng, 1996, Delgado, Dominguez and Lavergne, 2001, and Hsiao and Li, 2001), although they demand that the $m^{(r)}(W;\theta)$ are continuously differentiable in θ as many times as required,

$$\begin{aligned} & S_{1N}^{(r)}(\theta^*) \\ = \;& \frac{1}{N(N-1)} \sum_{(i,j)} K_{h,ij} \left\{ m_i^{(r)}(\theta^*) \left[\nabla m_j^{(r)}(\theta^*) \right] + m_j^{(r)}(\theta^*) \left[\nabla m_i^{(r)}(\theta^*) \right] \right\} \\ = \;& O_p\left(N^{-1/2}\right) \end{aligned}$$

$$\begin{aligned} & S_{2N}^{(r)}(\bar{\theta}) \\ = \;& \frac{1}{N(N-1)} \sum_{(i,j)} K_{h,ij} \left\{ \left[\nabla m_i^{(r)}(\bar{\theta}) \right]' \nabla m_j^{(r)}(\bar{\theta}) + m_i^{(r)}(\bar{\theta}) \left[\nabla^2 m_j^{(r)}(\bar{\theta}) \right] \right\} \\ = \;& O_p(1) \end{aligned}$$

where $\bar{\theta}$ lies on a line segment joining $\hat{\theta}$ and θ^* and $m_i^{(r)}(\theta) = m^{(r)}(W_i;\theta)$, $\nabla m_i^{(r)}(\theta) = \dfrac{\partial m^{(r)}(W_i;\theta)}{\partial \theta'}$, $\nabla^2 m_i^{(r)}(\theta) = \dfrac{\partial^2 m^{(r)}(W_i;\theta)}{\partial \theta \partial \theta'}$. These conditions actually derive from more primitive assumptions that would naturally be employed when constructing a Zheng type test statistic of (4) based the consistent estimator $\tilde{\theta} = \arg\max_\theta Q_N(\theta)$, which removes the parametric restrictions of $\mathcal{H}$. For example, $S_{1N}^{(r)}(\theta^*) = O_p(N^{-1/2})$ is just an extension of Zheng (1996, Lemma 3.3b) which itself derives from Theorem 3.1 in Powell, Stock and Stoker (1989). Furthermore, it is easy to see that $S_{2N}^{(r)}(\bar{\theta}) = S_{2N}^{(r)}(\theta^*) + O_p(N^{-1/2})$ and a result similar to the one which establishes that $S_{1N}^{(r)}(\theta^*) = O_p(1)$ reveals that

$$S_{2N}^{(r)}(\theta^*) = \frac{1}{N(N-1)} \sum_{(i,j)} K_{h,ij} \left[\nabla m_i^{(r)}(\theta^*)\right]' \nabla m_j^{(r)}(\theta^*) + o_p(1) = O_p(1)$$

Under these assumptions, we can now establish the main result of the article.

Theorem 2.1 *Consider the statistic $V_N(\hat{\theta})$ defined above and suppose that the various high level assumptions are satisfied. Then, under the null hypothesis,*

$$\begin{aligned} N h^{d/2} V_N(\hat{\theta}) &\xrightarrow{d} N\left(b(\theta_0^*), \Sigma(\theta_0^*)\right), \qquad (5) \\ b(\theta_0^*) &= \sum_r E\left[\left\{\Lambda^{(r)}(X)\right\}^2 f(X)\right] \end{aligned}$$

where $\Lambda^{(r)}(x) = E_{\mathcal{H}}\left[\sum_p \dfrac{\partial m^{(r)}(W_j;\theta^*)}{\partial \theta_p'} \gamma_p \,|X = x\right]$.

Proof 2.1 *A second order mean value expansion of $Nh^{d/2}V_N(\hat{\theta})$ about $\hat{\theta} = \theta^*$ yields*

$$\begin{aligned} Nh^{d/2}V_N(\hat{\theta}) &= Nh^{d/2}V_N(\theta^*) \\ &\quad + Nh^{d/2}\left\{\sum_r S_{1N}^{(r)}(\theta^*)\right\}(\hat{\theta} - \theta^*) \\ &\quad + Nh^{d/2}(\hat{\theta} - \theta^*)'\left\{\sum_r S_{2N}^{(r)}(\bar{\theta})\right\}(\hat{\theta} - \theta_*) \end{aligned}$$

where $(\hat{\theta} - \theta^*)' = \left((\hat{\theta}_0 - \theta_0^*)', \frac{-1}{\sqrt{Nh^{d/2}}}\gamma_1', ..., \frac{-1}{\sqrt{Nh^{d/2}}}\gamma_L'\right)$ *and* $\hat{\theta}_0 - \theta_0^* = O_p(N^{-1/2})$. *This implies that* $Nh^{d/2}\left\{\sum_r S_{1N}^{(r)}(\theta^*)\right\}(\hat{\theta} - \theta^*) = O_p(h^{d/2}) = o_p(1)$. *Furthermore, let* $\nabla_p m_i^{(r)}(\theta) = \dfrac{\partial m^{(r)}(W_i;\theta)}{\partial \theta_p'}$

$$\begin{aligned} & Nh^{d/2}(\hat{\theta} - \theta^*)'\left\{S_{2N}^{(r)}(\bar{\theta})\right\}(\hat{\theta} - \theta_*) \\ = \;& Nh^{d/2}(\hat{\theta} - \theta^*)'\left\{\frac{1}{N(N-1)}\sum_{(i,j)} K_{h,ij}\nabla' m_i^{(r)}(\theta^*)\nabla m_j^{(r)}(\theta^*)\right\}(\hat{\theta} - \theta_*) + o_p(1) \\ = \;& h^{d/2}\sqrt{N}(\hat{\theta}_0 - \theta_0^*)'\left\{\frac{1}{N(N-1)}\sum_{(i,j)} K_{h,ij}\nabla_0' m_i^{(r)}(\theta^*)\nabla_0 m_j^{(r)}(\theta^*)\right\}\sqrt{N}(\hat{\theta}_0 - \theta_0^*) \\ & + 2\sqrt{h^{d/2}}\left\{\frac{1}{N(N-1)}\sum_{(i,j)} K_{h,ij}\left\{\sum_{p=1}^{L}\gamma_p'\nabla_p' m_i^{(r)}(\theta^*)\right\}\nabla_0 m_j^{(r)}(\theta^*)\right\}\sqrt{N}(\hat{\theta}_0 - \theta_0^*) \\ & + \frac{1}{N(N-1)}\sum_{(i,j)} K_{h,ij}\left\{\sum_{p=1}^{L}\gamma_p'\nabla_p' m_i^{(r)}(\theta^*)\right\}\left\{\sum_{p=1}^{L}\nabla_p m_j^{(r)}(\theta^*)\gamma_p\right\} + o_p(1) \\ = \;& \frac{1}{N(N-1)}\sum_{(i,j)} K_{h,ij}\left\{\sum_{p=1}^{L}\gamma_p'\nabla_p' m_i^{(r)}(\theta^*)\right\}\left\{\sum_{p=1}^{L}\nabla_p m_j^{(r)}(\theta^*)\gamma_p\right\} + o_p(1). \end{aligned}$$

Finally, standard results on U-statistics reveal that, under the sequence of local al-

ternatives,

$$\frac{1}{N\left(N-1\right)}\sum_{(i,j)}K_{h,ij}\left\{\sum_{p=1}^{L}\gamma_p'\nabla_p' m_i^{(r)}(\theta^*)\right\}\left\{\sum_{p=1}^{L}\nabla_p m_j^{(r)}(\theta^*)\gamma_p\right\}$$
$$\xrightarrow{p} E\left\{\left\{E_{\mathcal{H}}\left[\left(\sum_{p=1}^{L}\nabla_p m^{(r)}(\theta^*)\gamma_p\right)|X\right]\right\}^2 f(X)\right\}$$

Thus,

$$Nh^{d/2}V_N(\hat{\theta}) = Nh^{d/2}V_N(\theta^*) + b(\theta_0^*).$$

Remark 2.1 *If generated regressors, rather than* X, *are to be used in the construction of the test statistic then the Kernel is required to be continuously differentiable and the corresponding Mean Value Expansion and definitions of* S_{1N} *and* S_{2N} *will need to be modified accordingly. The method of proof then follows the arguments of Hsiao and Li (Theorem 5.1, 2001).*

3 Insensitivity to Unconsidered Local Alternatives

The previous section has shown that unconsidered local alternatives shift the centre of the asymptotic distribution of kernel-based conditional moment tests by a quantity

$$b\left(\theta_0^*\right) = \sum_r E\left[\left\{\Lambda^{(r)}\left(X\right)\right\}^2 f\left(x\right)\right] \tag{6}$$

which is a function of $\Lambda^{(r)}\left(X\right) = E_{\mathcal{H}}\left[\left(\sum_{p=1}^{L}\nabla_p m^{(r)}(\theta^*)\gamma_p\right)|X\right]$. The magnitude of the shift is thus related to the individual marginal contribution of each departure to the r^{th} moment condition evaluated by the statistic $T_N^{\#}$. These marginal contributions are logically measured by the gradients $\nabla_p m^{(r)}(.)$.

Ultimately, the question of sensitivity to unconsidered local alternatives reduces to whether $b(\theta_0^*) = 0$ or not. If we were using the statistic $\hat{T}_N^{\#}$ above as a general procedure to evaluate the overall adequacy of the econometric model, then we would expect $b(\theta_0^*)$ to be different from zero whenever unconsidered local departures are present, so that the null model be rejected at some pre-established significance level α. However, if $\hat{T}_N^{\#}$ is used to isolate and identify particular errors of specification, then it is desirable that the test be insensitive to local misspecification for which it was not designed. In accordance to the previous discussion, $\hat{T}_N^{\#}$ will be insensitive to the whole collection of departures represented by $\theta_1, \cdots, \theta_P$ whenever $b(\theta_0^*) = 0$, and this condition is attainable if, for all p and r,

$$E_{\mathcal{H}}\left[\left(\sum_{p=1}^{L} \nabla_p m^{(r)}(\theta^*)\gamma_p\right)|X\right] = 0 \tag{7}$$

Note that a sufficient condition for this to happen is that all the gradients $\nabla_p m^{(r)}(\theta^*)$ are zero.

The above expression implies a result relating the asymptotic insensitivity of $\hat{T}_n^{\#}$ to the collection $\{\theta_p\}_{p=1}^{L}$ of local alternatives, and the asymptotic independence of $\hat{T}_n^{\#}$ with an equivalent *kernel-based score test* of the null hypothesis $\mathcal{H}\colon \theta_1 = \theta_2 = ... = \theta_L = 0$. To develop this result, let us assume that the sequence of local alternatives $\theta_1, \cdots, \theta_L$ had been considered in the model, so that the density function associated to the econometric model had been specified as $f(W;\theta)$, for $\theta' = (\theta_0', \theta_1', \cdots \theta_L')$. We can make $\theta_2 = \cdots = \theta_L = \mathbf{0}$ without any loss of generality, so that only θ_1 is present in the generating process and thus, the vector of parameters in the model can be partitioned as $\theta' = (\theta_0', \theta_1')$. The test $\hat{T}_n^{\#}$ can be easily adapted to evaluate the $d - d_1$ restrictions $\mathcal{H}\colon \theta_1 = 0$. In particular, a kernel-based score test could be constructed as

$$\hat{T}_N^S = \frac{\sum_{(i,j)} t_{ij}^S(\hat{\theta})}{\sqrt{2\sum_{(i,j)} \left[t_{ij}^S(\hat{\theta})\right]^2}}$$

where: $t_{ij}^S(\hat{\theta}) = K_{h,ij} \frac{\partial f\left(W_i;\hat{\theta}\right)}{\partial \theta_1'} \frac{\partial f\left(W_j;\hat{\theta}\right)}{\partial \theta_1}$. Under the alternative hypothesis,

$$E_{\mathcal{H}}\left[\frac{\partial f\left(W_i;\hat{\theta}\right)}{\partial \theta_1'} | X = x\right] \neq 0,$$

so that large values of $\hat{T}_N^S$ would be observed.

Consider now the arbitrary test $\hat{T}_N^{\#}$ in (3). From the Information Matrix equality, we know that

$$E_{\mathcal{H}}\left[\frac{\partial m^{(r)}\left(W,\theta\right)}{\partial \theta_1} | X = x\right] = E_{\mathcal{H}}\left[m^{(r)}\left(W,\theta\right)\frac{\partial f\left(W,\theta\right)}{\partial \theta_1} | X = x\right] \tag{8}$$

The above equality represents the expected marginal variation of the function $m^{(r)}\left(.\right)$ with respect to θ_1 as a function of the statistical correlation between the moment functions in the tests $\hat{T}_N^{\#}$ and $\hat{T}_N^S$. In particular, if $\hat{T}_N^{\#}$ is insensitive to the local departure θ_1, so that $E_{\mathcal{H}}\left[\frac{\partial m^{(r)}(W,\theta)}{\partial \theta_1} | X = x\right] = 0$ then $E_{\mathcal{H}}\left[m^{(r)}\left(W,\theta\right)\frac{\partial f(W,\theta)}{\partial \theta_1} | X = x\right] = 0$, ensuring that the functions $m^{(r)}$ and $\partial f / \partial \theta_1$are uncorrelated conditional on X. A suitable Law of Large Numbers will ensure that the asymptotic correlation between $\hat{T}_N^{\#}$ and $\hat{T}_N^S$ will tend to zero as $N \to \infty$, allowing us to establish the following, which is the main result of this article.

Lemma 3.1 A Kernel-Based Consistent Conditional Moment Test, $\hat{T}_N^{\#}$, of M moment restrictions will be insensitive to the p^{th} unconsidered local alternative, whenever $\hat{T}_N^{\#}$ is asymptotically uncorrelated with the corresponding nonparametric score test $\hat{T}_N^S$ of $\mathcal{H}\colon \theta_p = 0$.

In practice, the problem of studying the sensitivity of a kernel-based consistent conditional moment test to a collection of unconsidered local alternatives reduces to evaluating condition (7). Under the assumptions enforced to justify the asymptotic results, the functions $m^{(r)}(.)$ are differentiable, condition (7) can be studied. The only difficulty associated with the problem is to calculate the gradient of the vector $m^{(r)}(W, \theta)$.

The result established in Lemma (3.1) resembles the result obtained by Godfrey and Orme (1996) for parametric conditional moment tests. In particular, their result related the asymptotic insensitivity of a parametric conditional moment test to its asymptotic independence with the appropriate Lagrange Multiplier test of the null hypothesis $\mathcal{H}\colon \theta_j = 0$, where θ_j is a vector of unconsidered local alternatives. However, the applicability of their result is subject to the estimation method employed to approximate the values of the parameters in the model. Thus, while some of the results they present are satisfied in a Maximum Likelihood framework, the same result need not hold true when GMM is employed. Unlike in the mentioned article, result (3.1) does not depend on the loss function employed to estimate $\hat{\theta}$ and therefore, our results will be true whether $\hat{\theta}$ was calculated by Maximum Likelihood, GMM or any other method.

4 Examples

We finish this article by illustrating how to use our check in practice. The examples here are of limited scope, however our condition can be applied to much more complex testing environments.

Example 4.1 *The first example to illustrate the results introduced in the previous sections uses the following regression model,*

$$Y = X_1'\beta_1 + X_2'\beta_2 + u_i,\ u_i \sim IID\left(0, \sigma_u^2\right) \tag{9}$$

where $\beta_1 = O(1)$, $\beta_2 = O\left(Nh^{d/2}\right)^{-1}$, *and* X_1, X_2 *are vectors of regressors. While* X_1 *is incorporated to the researcher's econometric model,* X_2 *is not observed. Therefore the estimated model is* $Y = X_1'\beta_1 + u_i$, *and the term* $X_2'\beta_2$ *configures a sequence of unconsidered local alternatives.*

After the model is estimated, one might be interested in testing if the error terms are normally distributed. Following the article by Jarque and Bera (1980), the null hypothesis that u *has a Gaussian distribution function can be tested by studying the two following conditions simultaneously:*

$$H_0 \ : \ E\left[u^3|X\right] = 0 \ \textit{and} \ E\left[u^4|X\right] - 3\sigma_u^4 = 0 \tag{10}$$

$$H_a \ : \ \textit{Either} \ E\left[u^3|X\right] \neq 0 \ \textit{or} \ E\left[u^4|X\right] - 3\sigma_u^4 \neq 0 \tag{11}$$

A consistent kernel-based version of the Jarque-Bera statistic can be constructed by modifying in the appropriate way the statistic $T_N^{\#}$ *is defined in* (3). *Letting* $m^{(1)}(W,\theta) =$

u^3, $m^{(2)}(W,\theta) = u^4 - 3\sigma_u^2$ and substitution of u and σ_u^2 by, say, least squares estimators will lead to a feasible test for normality; namely,

$$T_N^{JB} = \frac{\sum_{(i,j)} t_{ij}^{JB}(\hat{\theta})}{\sqrt{2\sum_{(i,j)} \left[t_{ij}^{JB}(\hat{\theta})\right]^2}}$$

where $t_{ij}^{JB} = K_{h,ji}\left\{\hat{u}_i^3\hat{u}_j^3 + \left(\hat{u}_i^4 - 3\hat{\sigma}_u^4\right)\left(\hat{u}_j^4 - 3\hat{\sigma}_u^4\right)\right\}$. Under the null hypothesis the test has an asymptotic standard normal distribution, provided that the model for the conditional mean had been correctly specified. The question arising is, in what way the above test might be affected by potential local Pitman misspecification in the conditional mean. Whether the test is sensitive to such form of misspecification can be checked by examination of the condition in equation (2.1)*. It will be sufficient just to consider $E_{\mathcal{H}}\left[\left(\sum_{p=1}^{L}\nabla_p m_j^{(r)}(\theta^*)\right)|X\right]$, since γ_p is a nonzero vector of constants. Note that,*

$$\begin{aligned} E_{\mathcal{H}}\left[\left(\nabla_p m^{(1)}(\theta^*)\right)|X_1\right] &= E_{\mathcal{H}}\left[\left(\frac{\partial m^{(1)}(\beta_1,\beta_2)}{\partial \beta_2'}\right)|X_1\right] = E_{\mathcal{H}}\left[3u^2X_2|X_1\right] \neq 0 \\ E_{\mathcal{H}}\left[\left(\nabla_p m^{(2)}(\theta^*)\right)|X_1\right] &= E_{\mathcal{H}}\left[\left(\frac{\partial m^{(2)}(\beta_1,\beta_2)}{\partial \beta_2'}\right)|X_1\right] = E_{\mathcal{H}}\left[4u^3X_2|X_1\right] \neq 0 \end{aligned}$$

The overall consistent Jarque-Bera test is not insensitive to $\sqrt{Nh^{d/2}}$ local misspecification in the conditional mean, so that one won't be able to distinguish those situations when rejection of the null is due to the non-normality of the test from those situations when rejection of the null is due to some sort of local alternative present in the conditional mean of the regression model.

Example 4.2 *The second example consides a duration model. For instance, as-*

sume that T represents duration of unemployment and imagine, for simplicity, that a proportional hazard Weibull model has been proposed to undertake the study. The evolution of the variable across time can be explained by the influence of a set of regressors entering to the regression model below:

$$Y = X_1'\beta_1 + X_2'\beta_{2,N} + u \tag{12}$$

In the above equation, X_1,X_2 are vectors of regressors, such that $E\left[u|X_1, X_2\right] = 0$. The dependent variable in the above regression function is $Y = -\alpha \ln(T)$, where α is a parameter. The parameter vectors in the right hand side are such that $\beta_1 = O(1)$, and $\beta_{2,N} = O\left(Nh^{d/2}\right)^{-1}$. In particular, we assume that X_2 is unobserved, so that the term $X_2'\beta_{2,N}$ constitutes a sequence of unconsidered Pitman local alternatives. Therefore, the researcher only estimates $Y = X_1'\beta_1 + u$. The error term is assumed to follow a Type I extreme value distribution (see Kiefer, 1988), and then, the above model can be estimated by Maximum Likelihood.

A conditional moment test for the overall validity of the Weibull specification can be constructed by using the generalized residuals associated to the model, which are defined as $\varepsilon = \varepsilon(T|X_1, X_2; \beta_1, \beta_{2,N}, \alpha) = t^{\alpha} \exp(X_1'\beta_1 + X_2'\beta_{2,N})$, where $\varepsilon(.)$ denotes the integrated hazard function. It is well known that the above generalized residuals follow a unit exponential distribution (Kierfer, 1988), so that $E[\varepsilon^r] = r!$, and $E[\log(\varepsilon)] = -0.5772$, and, hence, the above functions should have zero expected value under the null hypothesis. Following Jaggia (1991) and Lancaster (1985) the *validity of the model can be tested by using the functions*

$$
\begin{aligned}
m^{(1)}(.) &= \varepsilon^2 - 2 && (13)\\
m^{(2)}(.) &= \log(\varepsilon) - \psi(1) && (14)\\
&= \alpha \ln(t) + \beta_1' X_1 + X_2' \beta_{2,N} - \psi(1)
\end{aligned}
$$

(for $\psi(r) = \frac{\partial \log \Gamma(r)}{\partial r}$ *and* $\Gamma(r)$ *denotes the gamma function) which have zero expected value under the null hypothesis that the Weibul model is correctly specified. Under the null hypothesis,* $E\left[m^{(1)}|X_1\right] = E\left[m^{(2)}|X_1\right] = 0$ *so that the appropriate test for the Weibul model can be obtained by letting*

$$
t_{ij} = K_{h,ji}\left\{\left(\hat{\varepsilon}_i^2 - 2\right)\left(\hat{\varepsilon}_j^2 - 2\right) + \left(\log(\hat{\varepsilon}_i) - \psi(1)\right)\left(\log(\hat{\varepsilon}_j) - \psi(1)\right)\right\} \qquad (15)
$$

in the test (18)*, where* $\hat{\varepsilon} = t^{\hat{\alpha}} \exp\left(\hat{\beta}_1' X_1\right)$ *and* $\hat{a}$*,* $\hat{\beta}_1$ *are the ML estimators of* α *and* β_1 *respectively.*

In a strictly nonparametric framework, Godfrey and Orme (1996) showed that the corresponding Conditional Moment test was insensitive to local departures vanishing at rate $\sqrt{N}$*. When the nonparametric equivalent is considered, we observe that*

$$
\begin{aligned}
E_{\mathcal{H}}\left(\frac{\partial m^{(1)}}{\partial \beta_{2,N}}|X_1\right) &= 2\xi(X_2)\, u^2 \neq 0 && (16)\\
E_{\mathcal{H}}\left(\frac{\partial m^{(2)}}{\partial \beta_{2,N}}|X_1\right) &= \xi(X_2) \neq 0 && (17)
\end{aligned}
$$

so that the test based on condition (15) *will be affected by the presence of the above unconsidered local alternatives. Therefore the above kernel-based test would be sensi-*

tive to local departures present in the regression function, so that the null hypothesis could be rejected, not because of an incorrect choice of the model, but because the presence of unconsidered local alternatives.

Example 4.3 *The third example considers tests for conditional heteroskedasticity. In the parametric literature there are plenty of competing tests for such hypothesis (see, for instance pioneering work by Breusch and Pagan (1979), Godfrey (1978), or White (1980). Many of these tests are subject to two problems of opposite sign. On the one hand, they might lack power when errors follow certain types of heteroskedastic pattern. On the other hand, these tests are known to be sensitive to local misspecifications in the conditional mean of order* $\sqrt{N}$*, what will result in rejection of the null hypothesis more often than expected. To illustrate these situations, consider the following simple example. A regression model such as* $Y_i = \beta_1 X_i + u_i$*, for* $u_i = h\left(\sigma_0^2 + \theta Z_i\right) e_i$*,* $e_i \sim n.i.d.\left(0, \sigma^2\right)$ *and* X_i*,* Z_i *are exogenous, and* $E\left[X_i u_i\right] = 0$*. A score test of homoskedasticity is based on* $N^{-1} \sum_{i=1}^n \hat{u}_i^2 \left(Z_i - \bar{Z}\right)$*, for* $\bar{Z} = N^{-1} \sum_{i=1}^n Z_i$*. Under* $H_o : \theta = 0$*,* $E\left[u_i^2 \left(Z_i - \bar{Z}\right)\right] = 0$*. Therefore,*

$$\frac{\sum_{i=1}^n \hat{u}_i^2 \left(Z_i - \bar{Z}\right)}{\sqrt{n\tau}} \overset{A}{\sim} N\left(0, 1\right)$$

for $\tau = var\left(u_i^2 \left(Z_i - \bar{Z}\right)\right)$*, so that, given a consistent estimator,* $\hat{\tau}$ *of* τ*, we have*

$$\left(\frac{\sum_{i=1}^n \hat{u}_i^2 \left(Z_i - \bar{Z}\right)}{\sqrt{n\hat{\tau}}}\right)^2 \overset{A}{\sim} \chi_1^2$$

Now, if $Z_i = X_i$*, then* $E\left(u_i^2 \left(X_i - \bar{X}\right)\right) = 0$ *under the alternative hypothesis, so* *that the numerator of the test will be close to zero. The null hypothesis will thus be accepted, even though disturbances are heteroskedastic. On the other hand, suppose that* $Y_i = \beta_i X_i + \delta_n n^{-1/2} + u_i$, *for* $\delta_n = O(1)$, *and* $u_i \sim n.i.d\,(0, \sigma_0^2)$. *It follows that, under the null hypothesis, for any exogenous* Z_i, $E\left[u_i^2\left(Z_i - \bar{Z}\right)\right] = \delta_n^2$, *so that*

$$\frac{\sum_{i=1}^{n} \hat{u}_i^2\left(Z_i - \bar{Z}\right)}{\sqrt{n\tau}} \overset{A}{\sim} N\left(\delta_n^2, 1\right)$$

so that the test will have a non-central chi-square distribution, leading to rejection of the null when one should actually accept it. Thus, this test is sensitive to local misspecification in the conditional mean.

Consistent kernel-based tests for heteroskedasticity can be constructed by appropriately modifying the statistic in (3) *-see Delgado et al (2001) and Hsiao and Li (2001). Unlike those parametric tests mentioned above, these tests exhibit nontrivial power under any type of heteroskedasticity. However, there is a second advantage associated to such a procedure. As we show following this test will be insensitive to local misspecifications in the mean of order* $O\left(nh^{d/2}\right)^{-1/2}$. *The result will follow by application of the result introduced in the previous section.*

Suppose, thus, that we have a regression model, such as $Y_t = r\left(X_t, \theta^*\right) + u_i$, *for some* $(k_1 \times 1)$ *vector of regressors* X_t. *The null hypothesis is* $H_o : E\left(u_t^2 | Z_t\right) = \sigma_u^2$, *where* Z_t *is a vector of random variables. Under the alternative,* $E\left[u_t^2 | Z_t\right] = q\left(Z_t\right) \neq \sigma_u^2$. *The variables in* Z_t *could be a set of exogenous, independent regressors (as in Delgado et al, 2001), but it could also include weakly dependent data and endogenous variables (as in Hsiao and Li (2001)). In particular,* Z_t *could include unobservable generated regressors which need to be estimated. One such situation would arise, for* *instance, if one suspects that the error process could follow an ARCH-type structure. Without loss of generality, we could write* $Z_t = V_t + \varepsilon_i$*, where* $\varepsilon_t \sim iid\,(0, \mathbf{\Omega})$ *for some diagonal variance matrix* $\mathbf{\Omega}$*. Now,* $T_N^{\#}$ *will be based on the moment function* $m\,(W_t, \theta^*) = u_t^2 - \sigma_u^2 = \left[Y_t - r\,(X_t, \theta^*)\right]^2 - \sigma_u^2$*, which under the null hypothesis will satisfy*

$$E\left[\left(u_t^2 - \sigma^2\right) E\left[u_t^2 - \sigma_u^2 | Z_t\right] f\,(Z_t)\right] = 0, \text{ a.s.}$$

Since u_t*,* σ_u^2 *are not observable, replace them by the squared residuals,* $\hat{u}_t^2$*, and any consistent estimator of* σ_u^2*, an obvious choice being* $\hat{\sigma}_u^2 = N^{-1} \sum \hat{u}_t^2$*. So, on letting* $\hat{m} = \hat{u}^2 - \hat{\sigma}_u^2$*, the appropriate Consistent Conditional Moment test is defined as follows:*

$$T_n^H = \frac{1}{N\,(N-1)} \sum_{(t,s)}^{N} \hat{K}_{h,ts} \left(\hat{u}_t^2 - \hat{\sigma}_u^2\right) \left(\hat{u}_s^2 - \hat{\sigma}_u^2\right) \tag{18}$$

where,

$$\hat{K}_{h,ts} = \frac{1}{h^d} K \left(\frac{\hat{Z}_t - \hat{Z}_s}{h}\right)$$

and $\hat{Z}$ *is an estimation of* Z*. Hsiao and Li show that*

$$\frac{\sqrt{N h^{d/2}} T_N^H}{\sqrt{\hat{\Sigma}_n}} \xrightarrow{d} N\,(0,1)$$

where

$$\hat{\Sigma}^H = \frac{2h^d}{N\,(N-1)} \sum_{(t,s)}^{N} \hat{K}_{h,ts}^2 \left(\hat{u}_t^2 - \hat{\sigma}_u^2\right)^2 \left(\hat{u}_s^2 - \hat{\sigma}_u^2\right)^2 \tag{19}$$

Obtaining the distribution of the statistic under so general conditions requires further structure on the behavior of the bandwidth parameter h *and the smoothness properties* *of the kernel density -see Hsiao and Li (2001) for the details. In particular* $h = O(n^{-\alpha})$, *for* $0 < \alpha < 7d/8$, *when dependent data appears in the sample, whilst if generated regressors are used, then* $nh^{4-d/2} \to \infty$ *is required to ensure the asymptotic negligibility of higher order terms in the asymptotic expansion of the test. In the case of independent regressors, only* $nh^{d/2} \to \infty$ *is required.*

Suppose now, that the conditional mean of the model is a sequence of Pitman DGP's with an structure similar to the following

$$Y_t = X_{1t}'\theta^* + X_{2t}'\delta_N + u, \text{ for } u \sim i.i.d\left(0, \sigma_u^2\right) \tag{20}$$

where δ_N *is a sequence of order* $O\left(Nh^{d/2}\right)^{-1/2}$*.We might suspect that the conditional variance of* u_t *might follow a process depending on some regressors* Z*, which could be, for example* u_{t-1}^2*. Under the local misspecification, we have*

$$m_t(\theta) = u_t^2 - \sigma_u^2 = \left[Y - X_{1t}'\theta^* - X_{2t}'\delta_N\right]^2 - \sigma_u^2 \tag{21}$$

In order to study if the test T_n^H *will be affected by misspecification, we only need to check condition* (7)*, obtaining that*

$$E_{\mathcal{H}}\left[\left(\nabla_p m(\theta^*)\delta_N\right)|Z\right] = \delta_N E_{\mathcal{H}}\left[\left(\nabla_p m(\theta^*)\right)|Z_t\right] = \delta_N E_{\mathcal{H}}\left[-2uX_2|Z\right] = 0$$

Therefore, we conclude that the test for heteroskedasticity will not be affected by local misspecifications in the mean of order $O\left(Nh^{d/2}\right)^{-1/2}$*.*

5 Conclusion

This paper studied the sensitivity of a Consistent Conditional Moment test to unconsidered local alternatives vanishing at rate $O\left(Nh^{d/2}\right)^{-1/2}$. In general, these tests will detect this type of local departures, so that the null hypothesis will be rightly rejected. When applied workers need to isolate a particular type of error in their model, kernel-based conditional moment tests must be designed so that they are insensitive to departures for which it they were not designed. In these situations, one can make use of the condition presented in this article in order to characterize the type of unconsidered departures that might distort the behavior of the test of interest. This allows to tailor tests so that their performance is not affected by certain types of unobserved local departures. We conjecture that the results presented here might also apply more generally to arbitrary consistent tests, not only those based on kernels. However a formal proof of this claims is required and we leave that for future research.

University of Manchester, United Kingdom

REFERENCES

Bera, A. K. and M. J. Yoon, "Specification testing with Locally Misspecified Alternatives", Econometric Theory 9 (1993), 649-658.

Breusch, T and A. Pagan, "A simple test for heteroskedasticity and random coefficient variation", Econometrica 47 (1978) , 1287-1249

Davidson, R. and J. G. MacKinnon . "Implicit Alternatives and the Local Power of Test Statistics", Econometrica 55 (1987), 1305-1329.

Delgado, M., M. Dominguez and P. Lavergne, " Consistent tests of conditional moment restrictions" Annales d'Economie et Statistique (2001).

De Jong, "A central limit theorem for generalized quadratic forms", Probability Theory and Related Fields 75 (1987), 261-277.

Fan, Y. and Q. Li, "Central Limit Theorem for Degenerate U-Statistics of Absolutely Regular Processes with Applications to Model Specification Testing", Nonparametric Statistics 10 (1999), 245-271.

Fan, Y. and Q. Li, "Consistent Model Specification Test: Omitted Variables and Semiparametric Functional Forms". Econometrica 64 (1996), 865-890

Godfrey, L. "Testing for Multiplicative Heteroskedasticity", Journal of Econometrics (1978), 8, 227-236.

Godfrey, L. and C. Orme, "On the behavior of Conditional Moment Tests in the Presence of Unconsidered Local Alternatives", International Economic Reviews 37 (1996), 263-281

Hall, P. "Central Limit Theorem for Integrated Square Error of Multivariate Nonparametric Density Estimators", Journal of Multivariate Analysis 14 (1984) , 1-16.

Hsiao, C. and Q. Li, "A consistent test for conditional heteroskedasticity in Time-Series regression models", Econometric Theory 17 (2001) , 188-221.

Jaggia, S. "Specification Tests Based on the Heterogeneous Generalised Gamma Model of Duration: with an Application to Kenna's Strike Data", Journal of Applied Econometrics 6 (1991) , 169-180

Jarque, C.M. and A. K. Bera, "Efficient Tests for Normality, Homoskedasticity and Serial Independence" Economics Letters 6 (1980), 255-259.

Kiefer, N. "Economic Duration Data and Hazard Functions", Journal of Economic Literature (1988), 26, 646-679.

Lancaster, T. "Generalised Residuals and Heterogeneous Duration Models: the Weibull Case", Journal of Econometrics 28, 155-169

Li, Q, and S. Wang. "A simple consistent bootstrap test for a parametric regression function.", Journal of Econometrics (1998), 87, 145-165

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

Pitman, E. J. G. "*Lecture Notes on Nonparametric Statistical Inference*", (Columbia University, 1949)

Powell, J., J.H. Stock, and T.M. Stoker, "Semiparametric Estimation of Index Coefficients", Econometrica 57 (1989), 1403-1430

Racine, J. and Q. Li, "Nonparametric estimation of regression functions with both categorical and continuous data", Journal of Econometrics 119 (2004) , 99-130.

White, H. "A heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", Econometrica 48 (1980), 817-838

Zheng, J.X. "A Consistent Test of Functional Form Via Non-Parametric Estimation Techniques", Journal of Econometrics 76 (1996), 263-289

Zheng, J. X. "A consistent Test of conditional parametric distributions", Econometric Theory 16 (2000), 667-691