行政院國家科學委員會專題研究計畫 成果報告

不受分配影響的模型設定之檢定方法
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 98-2410-H-004-056-
執 行 期 間 ：98年08月01日至99年07月31日
執 行 單 位 ：國立政治大學經濟學系

計畫主持人：林馨怡

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 99年08月19日

1 Introduction

Many economic and econometric models are represented by conditional moment restrictions, for example, the rational expectation model, the market disequilibrium model, the conditional probability model, the discrete choice model and the nonlinear simultaneous equations model. The validity of these types of model is determined by testing conditional moment restrictions. Examples of such tests include conditional moment tests or M-test developed by Newey (1985), Tauchen (1985), and White (1987). However, such conditional moment tests may not be consistent because only necessary conditions of conditional moment restrictions are checked. There is an abundance of literature on constructing consistent conditional moment tests. One technique is to employ a nonparametric test. See, for example, Delgado and González Manteiga (2001), Li, Hsiao, and Zinn (2003), Horowitz and Spokoiny (2001), Tripathi and Kitamura (2003), and Zheng (2000), among others. The nonparametric tests are usually subjective in choosing smoothing parameters and may be computationally costly. Another technique for constructing a consistent conditional moment test is based on infinitely many unconditional orthogonality restrictions with uncountably many weighted functions indexed by continuous nuisance parameters (Stichcombe and White, 1998). This technique is called the integrated function approach because it uses the integrated measures of dependence of orthogonal restrictions. For these types of tests, when determining the weighted functions, Bierens (1982, 1984, 1990), Bierens and Ploberger (1997), Bierens and Ginther (2001), and de Jong (1996), employ the exponential function, while Koul and Stute (1999), Stute (1997), Stute, Thies and Zhu (1998), and Stute and Zhu (2002) employ the indicator function.

It is noted, that generally, tests based on integrated function approach are not asymptotically pivotal. That is, their limiting distributions depend on model characteristics and critical values cannot be tabulated. For example, the limiting distribution for tests employing the exponential weight function depend on the data generating process (DGP) of the auxiliary nuisance parameters. Although Bierens and Ploberger (1997) have derived case-independent upper bounds of critical values to solve the limiting distribution problem, their test may be too conservative in practice. Meanwhile, the limiting distribution for tests employing the indicator weight function is not asymptotically pivotal because of estimation effects (Durbin, 1973) and being case dependent. Dominguez and Lobato (2006), Stute, González Manteiga and Presedo Quindïmil (1998), and Whang (2000, 2001, 2004) try to avoid the problem by using bootstrapping techniques to approximate the limiting distribution. Specifically, Khmaladze and Koul (2004), Stute, Thies and Zhu (1998), Koul and Stute (1999), Stute and Zhu (2002), and Song (2009) employ the martingale transformation technique of Khmaladze (1981) to obtain asymptotically distribution-free test statistics. However, these tests usually encounter the poor finite sample performance due to the curse of dimensionality. Recently, Excanciano (2006) and Lavergne and Patilea (2008) propose tests breaking the curse of dimensionality. The former test is based on the integrated function technique and uses projections, while the latter test is based on the smoothing nonparametric technique.

Accordingly, this paper proposes a consistent conditional moment test that is asymptotically pivotal. The proposed test is based on the integrated function approach and the test statistic is obtained through a subsampling marked empirical process, using sample size b instead of the whole sample size n such that $b < n$. Subsampling, as defined by Politis and Romano (1994) and Politis, Romano and Wolf (1999) is a method for estimating the distribution of an estimator or test statistic by drawing subsamples from the original data. Andrews and Guggenberger (2005), Chernozhukov and Fernández-Val (2005), Guggenberger and Wolf (2004), Hong and Scailet (2006), Linton, Massoumi and Whang (2005) and Whang (2004) have employed subsampling techniques for estimating the distribution of estimators. Instead of computing the sample average of the conditional moment function with the whole sample, the test statistic is obtained by the subsampling marked empirical process with subsample size b. The estimation effect disappears when the relative sample size of subsampling to that of the whole sample is zero asymptotically. Therefore, the proposed test does not suffer from the estimation effect problem and is asymptotically pivotal. Further, multiple regressors may be employed in the test. Thus, the proposed test can be viewed as the complement of Escanciano (2006) and Lavergne and Patilea (2008) for breaking the curse of dimensionality. Additionally, any $\sqrt{n}$-consistent estimator and different estimation methods may be employed to compute the test statistic. Bootstrapping, martingale transformation or nonparametric techniques are not required, thus, simplifying computation of test statistics. However, the proposed test is powerful against local alternatives at rates $b^{-1/2}$, but the proposed test is incapable of detecting local alternatives at rate $n^{-1/2}$. When performing Monte Carlo simulation, it was shown that good finite sample performances were obtained and the proposed test was robust with respect to different values of b.

Following arrangement of this paper is as follows. Section 2 presents the conditional moment restriction and the proposed test. Section 3 shows the consistency of the proposed test and the asymptotic behavior given different local alternatives. Section 4 shows the results of Monte Carlo simulation. Lastly, Section 5 is the conclusion. All proofs are presented in the Appendix.

2 A New Test

2.1 Conditional Moment Restrictions

Consider the general conditional moment restrictions

$$\mathbb{E}[m(Y, X, \theta_o)\big|X] = 0, \tag{1}$$

where $\mathbb{E}[\cdot|X]$ denotes the expectation conditional on the information set of X, the function $m(\cdot)$ is well-defined, $\{Y, X\}$ is a sequence of random variables with $X = (X_1, \cdots, X_k)'$ and parameters $\theta \in \Theta$ with $\Theta \in \boldsymbol{R}^k$. The conditional moment restrictions can be obtained from existing models such as the parametric nonlinear regression model where $m(Y, X, \theta_o)$ is the difference between Y and $g(X', \theta)$, with $g(\cdot)$ being a nonlinear function. To test the condition moment restrictions, the null and alternative hypotheses are as follows. The null hypothesis is the conditional moment function being equal to zero:

$$H_0 : P\Big\{\mathbb{E}(m(Y, X, \theta_o)\big|X) = 0\Big\} = 1, \text{ for some } \theta_o \in \Theta,$$

and the alternative hypothesis is, for all $\theta \in \Theta$, $\mathbb{E}(m(Y, X, \theta)\big|X) \neq 0$ with a positive probability:

$$H_1 : P\Big\{\mathbb{E}(m(Y, X, \theta)\big|X) = 0\Big\} < 1, \text{ for all } \theta \in \Theta,$$

with $\Theta \in \boldsymbol{R}^k$ a compact set.

As previously proposed by Stinchcombe and White (1998), the conditional moment condition (1) equals infinitely many unconditional moment functions

$$\mathbb{E}[m(Y, X, \theta_o)\omega(X, x)] = 0, \forall x \in \boldsymbol{R}^k, \tag{2}$$

where $\omega(\cdot)$ is an infinite set indexed by continuous parameters x and $\omega(\cdot)$ may be any analytic function that is not polynomial. A consistent conditional moment test can be constructed by testing (2). For example, Bierens (1982, 1984, 1990), de Jong (1996) and Bierens and Ploberger (1997) and Bierens and Ginther (2001) employ the exponential weighted function $\omega(X, x) = \exp(X'x)$ for their integrated conditional moment test. Meanwhile, Stute (1997), Stute, Thies and Zhu (1998), Koul and Stute (1999) and Stute and Zhu (2002) employ the indicator function

$$\omega(X, x) = \mathbf{1}_{\{X \leq x\}} := \mathbf{1}_{\{X_1 \leq x_1\}} \cdots \mathbf{1}_{\{X_k \leq x_k\}},$$

where $\mathbf{1}_A$ denotes the indicator function of even A. This paper proposes employing the indicator function and the conditional moment restrictions (1) can be rewritten by the infinitely many unconditional moment functions as follows:

$$\mathbb{E}[m(Y, X, \theta_o)\mathbf{1}_{\{X \leq x\}}] = 0, \forall x = (x_1, \cdots, x_k)' \in \boldsymbol{R}^k, \tag{3}$$

wherein multivariate regressors may be employed; see Khmaladze and Koul (2004), Escanciano (2006), and Song (2009).

2.2 Test Statistics

The specification test employed in this paper examines infinitely many unconditional moment functions (3) that are equivalent to the conditional moment restriction (1). Thus, the specification test is a consistent conditional moment test. To test whether the moment function $\mathbb{E}[m(Y,X,\theta_o)\mathbf{1}_{\{X\leq x\}}]$ equals to zero, the normalized sample average of the moment function:

$$M_n(x;\theta_o) := \frac{1}{\sqrt{n}}\sum_{t=1}^{n} m\big(Y_i, X_i, \theta_o\big)\mathbf{1}_{\{X_i\leq x\}},$$

with $\{Y_i, X_i\}_{i=1}^{n}$ a sequence of random variable, and $\mathbf{1}_{\{X_i\leq x\}} = 1_{\{X_{i1}\leq x_1\}}\cdots 1_{\{X_{ik}\leq x_k\}}$, is employed. The function $m\big(Y_i, X_i, \theta\big)\mathbf{1}_{\{X_i\leq x\}}$ is the marked empirical process with the marks given by the moment function m. The function M_n is the average of the marked empirical process with sample size n. If $M_n(x;\theta_o)$ is close to zero, then the null hypothesis is not rejected. Otherwise, the null hypothesis is rejected and the conditional moment restriction does not hold.

Since the true parameter θ_o is unknown, we replace θ_o by its consistent estimator, $\hat{\theta}_n$. Thus the sample average of the marked empirical process is:

$$M_n(x;\hat{\theta}_n) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} m(Y_i, X_i, \hat{\theta}_n)\mathbf{1}_{\{X_i\leq x\}}.$$

By rewriting the process M_n based on

$$M_n(x;\hat{\theta}_n) = M_n(x;\theta_o) + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(m(Y_i, X_i, \hat{\theta}_n) - m(Y_i, X_i, \theta_o)\big)\mathbf{1}_{\{X_i\leq x\}},$$

if $m(Y_i, X_i, \theta)$ is once differentiable with first derivative $\nabla_\theta m(Y_i, X_i, \theta_o)$, then

$$\begin{aligned} M_n(x;\hat{\theta}_n) &= M_n(x;\theta_o) + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nabla_\theta m(Y_i, X_i, \theta_o)(\hat{\theta}_n - \theta_o)\mathbf{1}_{\{X_i\leq x\}} + o_p(1) \\ &= M_n(x;\theta_o) + \sqrt{n}(\hat{\theta}_n - \theta_o)\frac{1}{n}\sum_{i=1}^{n}\nabla_\theta m(Y_i, X_i, \theta_o)\mathbf{1}_{\{X_i\leq x\}} + o_p(1). \end{aligned}$$

Thus, $M_n(x;\hat{\theta}_n)$ and $M_n(x;\theta_o)$ are not asymptotically equivalent due to the presence of the second term on the right hand side of the second equality. This term is the estimation effect presented in Durbin (1973), wherein the presence of the second term depends on a model characteristic that makes the test based on $M_n(x;\hat{\theta}_n)$ not asymptotically pivotal. To obtain an asymptotically distribution-free test, Stute, Thies and Zhu (1998), Koul and

Stute (1999) and Stute and Zhu (2002) employ the martingale transformation technique for univariate regressors and Khmaladze and Koul (2004) and Song (2009) employ the same technique for multivariate regressors. Note that because using a nonparametric estimation of the conditional moment function is required, it is complicated to compute a high dimensional nonparametric estimation and is subjective to user-chosen parameters employing martingale transformation technique. In addition, the finite sample performance is poor due to the curse of dimensionality. To solve the subjective choice of parameters problem and the curse of dimensionality, Escanciano (2006) proposes a consistent conditional moment test using the projections technique and his test presents excellent empirical powers in finite sample. However, the limiting distribution of Escanciano's test should be obtained by bootstrapping technique and is not asymptotically pivotal.

Thus, this paper employs a subsampling version of the M_n process to construct a consistent conditional moment test which is asymptotically pivotal. Instead of employing the whole sample size n to compute the marked empirical process, a subsample size b is employed to compute the sample average and construct the process, for $b < n$:

$$M_b(x; \hat{\theta}_n) := \frac{1}{\sqrt{b}} \sum_{i=1}^{b} m\big(Y_i, X_i, \hat{\theta}_n\big) \mathbf{1}_{\{X_i \leq x\}},$$

where $\hat{\theta}_n$ can be any $\sqrt{n}$-consistent estimator associated with the model of interest with sample size n. Thus, by employing M_b the following equation is provided:

$$\begin{aligned} M_b(x; \hat{\theta}_n) &= M_b(x; \theta_o) + \frac{1}{\sqrt{b}} \sum_{i=1}^{b} \nabla_\theta m(Y_i, X_i, \theta_o)(\hat{\theta}_n - \theta_o) \mathbf{1}_{\{X_i \leq x\}} + o_p(1) \\ &= M_b(x; \theta_o) + \sqrt{\frac{b}{n}} \sqrt{n}(\hat{\theta}_n - \theta_o) \left[\frac{1}{b} \sum_{i=1}^{b} \nabla_\theta m(Y_i, X_i, \theta_o) \mathbf{1}_{\{X_i \leq x\}} \right] + o_p(1). \end{aligned} \tag{4}$$

If $b \to \infty$, $n \to \infty$ and $b/n \to 0$, and there exist some regularity conditions, then the second term on the right-hand-side of the second equality of (4) converges to zero. Thus, $M_b(x; \hat{\theta}_n)$ and $M_b(x; \theta_o)$ are asymptotically equivalent. Subsampling the marked empirical process eliminates the estimation effect. Assume $D(\boldsymbol{R}^k)$ to be the space of the cadlag function on $\boldsymbol{R}^k$ endowed with the Skorohod topology. Here, M_b is in $D(\boldsymbol{R}^k)$. Assume also, that $\Rightarrow$ denotes the convergence in distribution, and $\xrightarrow{p}$ denotes the convergence in probability. The following assumptions are sufficient for the weak convergence of the subsampling marked empirical process.

[A1] $\{Y_i, X_i\}_{i=1}^n$ is independent and identically distributed (i.i.d.) where X_i has the bounded and continuous distribution function F and the density function is f.

[A2] (i) $\mathbb{E}[m(Y_i, X_i, \theta)^2 | X_i] < \infty$,

(ii) $\mathbb{E}[m(Y_i, X_i, \theta)^4] = \kappa < \infty$,

(iii) $\mathbb{E}[m(Y_i, X_i, \theta)^4 ||X_i||^{1+\eta}] < \infty$, for some $\eta > 0$.

[A3] $m(\cdot)$ is once continuously differentiable in a neighborhood θ_o and satisfies

$$\mathbb{E}\left[\sup_{\theta \in \Theta_o} |\nabla_\theta m(Y_i, X_i, \theta)|\right] < \infty,$$

where Θ_o denotes a neighborhood of θ_o.

[A4] $\hat{\theta}_n$ is a $\sqrt{n}$-consistent estimator; that is $\sqrt{n}\big(\hat{\theta}_n - \theta_o\big) = O_p(1)$.

The assumptions in [A2] restrict the dependence of the moment function. Given [A2] (i), the conditional variance function $\sigma^2(X_i)$ of $m(Y_i, X_i, \theta)$ is defined with

$$\sigma^2(u) := \text{var}\big[m(Y_i, X_i, \theta)\big|X_i = u\big].$$

For $x_i = (x_1, \cdots, x_k)'$ and $u = (u_1, \cdots, u_k)'$:

$$V(x) := \mathbb{E}\left[\sigma^2(X_i)\mathbf{1}_{\{X_i \leq x\}}\right] = \int_{-\infty}^{x} \sigma^2(u) F(du),$$

is defined with $\int_{-\infty}^{x} := \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k}$. Assumptions [A1] together with [A2] are required to obtain the uniform tightness in the space $D[-\infty, \infty]$. Assumption [A3] is a standard smoothness assumption. [A3] can be relaxed as a non-smooth moment function when considering the stochastic equicontiunity of m. Assumption [A4] is weak and may be applied to most existing estimation methods. Following, the weak convergence of M_b is obtained.

Theorem 2.1. *Under H_0 and given assumptions* [A1]-[A4], *if $b \to \infty$, $n \to \infty$ and $b/n \to 0$, then one has:*

$$M_b(x; \hat{\theta}_n) \Rightarrow B\big(V(x)\big),$$

where $B(\cdot)$ is a Gaussian process with mean zero and covariance function $V(x_1 \wedge x_2)$.

The limiting distribution of M_b is a centered Gaussian process which is a multi-parameter Brownian motion process on $[0, 1]^k$ with covariance function

$$V(x_1 \wedge x_2) = \int_{-\infty}^{x_1 \wedge x_2} \sigma^2(u) F(du),$$

where $\int_{-\infty}^{x_1 \wedge x_2} = \int_{-\infty}^{x_{11} \wedge x_{21}} \cdots \int_{-\infty}^{x_{1k} \wedge x_{2k}}$. In particular, when X_i is univariate, the process B is the standard Brownian motion process. The limit of $M_b(x; \theta_o)$ and that of $M_b(x; \hat{\theta}_n)$ are the same and the estimation effect problem of Durbin (1973) is eliminated because the convergence rate of b to infinity is slower that of n to infinity. Note that $V(x)$ plays an important role in the proposed test. Since $V(x)$ still depends on the distribution of X_i and σ^2, the process $M_b(x;\hat{\theta}_n)$ is not asymptotically distribution-free. For a general conditional heteroskedasticity case, the scaled invariant version of subsampling marked empirical process is considered as follows:

$$\tilde{M}_b(x;\hat{\theta}_n) := \frac{1}{\sqrt{b}}\sum_{i=1}^{b}\hat{\sigma}(X_i)^{-1}m\big(Y_i, X_i, \hat{\theta}_n\big)\mathbf{1}_{\{X_i\le x\}},$$

where $\hat{\sigma}(X_i)^2$ is a consistent estimator of $\sigma(X_i)^2$. The scaled version of the statistic is also considered in many research, such as Khmaladze and Koul (2004), Koul and Stute (1999), Stute (1997), Stute, Thies and Zhu (1998), and Song (2009). When $\sigma^2(X_i) = \sigma_0^2$ (the conditional homoskedasticity case), which is a constant, $\tilde{M}_b(x;\hat{\theta}_n)$ simplifies to

$$\frac{1}{\sqrt{b}}\hat{\sigma}_b^{-1}\sum_{i=1}^{b}m\big(Y_i, X_i, \hat{\theta}_n\big)\mathbf{1}_{\{X_i\le x\}},$$

with $\hat{\sigma}_b^2 = b^{-1}\sum_{i=1}^{b}m(Y_i, X_i, \hat{\theta}_n)^2$ a consistent estimator for σ_0^2.

Theorem 2.2. *Under H_0 and given assumptions* [A1]-[A4]*, if $b\to\infty$, $n\to\infty$, $b/n\to 0$ and $\hat{\sigma}(X_i)^2-\sigma(X_i)^2 = o_p(b^{-1/2})$, then*

$$\tilde{M}_b(x;\hat{\theta}_n) \Rightarrow B\big(F(x)\big),$$

with $B(\cdot)$ a Gaussian process with mean zero and covariance function $F(x_1\wedge x_2)$.

The computational counterpart of the scaled invariant version of $\tilde{M}_b(x;\hat{\theta})$ is as follows:

$$\tilde{M}_b(X_j;\hat{\theta}_n) := \frac{1}{\sqrt{b}}\sum_{i=1}^{b}\hat{\sigma}(X_i)^{-1}m\big(Y_i, X_i, \hat{\theta}_n\big)\mathbf{1}_{\{X_i\le X_j\}}, j=1,\cdots,n,$$

where each realization X_j is used as an x in the indicator function. Consider two goodness-of-fit statistics, the Kolmogorov-Smirnov and Cramer-von Mises test statistics:

$$KS_n = \sup_{X_j\in\boldsymbol{R}^k}\big|\tilde{M}_b(X_j;\hat{\theta}_n)\big|,$$

and

$$CM_n = \frac{1}{n}\sum_{j=1}^{n}\tilde{M}_b(X_j;\hat{\theta}_n)^2.$$

Employing Theorem 2.2 and the continuous mapping theorem, for a large n, with $\tau\in[0,1]^k$, then

$$KS_n \Rightarrow \sup_{x\in\boldsymbol{R}^k}\big|B\big(F(x)\big)\big| = \sup_{\tau\in[0,1]^k}\big|B(\tau)\big|,$$

and

$$CM_n = \int_{-\infty}^{\infty} \tilde{M}_b(X;\hat{\theta}_n)^2 F(dx) \Rightarrow \int_{-\infty}^{\infty} B\big(F(x)\big)^2 F(dx) = \int_{[0,1]^k} B(\tau)^2 d\tau.$$

The critical values of the test statistics KS_n and CM_n can be found in existing literature such as Shorack and Wellner (1986) and Khmaladze and Koul (2004).[1] Note that the proposed test of this paper is asymptotically pivotal and the limiting distribution of the proposed test does not depend on a DGP. Therefore, the following corollary is as follows.

Corollary 2.3. *Under all the assumptions in Theorem* 2.2*.*

$$KS_n \Rightarrow \sup_{\tau \in [0,1]^k} \big|B(\tau)\big|,$$

and

$$CM_n \Rightarrow \int_{[0,1]^k} B(\tau)^2 d\tau,$$

where $B(\cdot)$ *is a Gaussian process with mean zero and covariance* $(\tau_1 \wedge \tau_2)$*.*

3 Power of the Tests

To investigate the power performance of the proposed test, two types of alternatives are considered. One is the general alternative:

$$H_1 : \mathbb{E}[m(Y,X,\theta_o)|X] = \mu(X) \neq \mathbf{0},$$

and the other is the local alternatives:

$$H_1^L : \mathbb{E}\big[m(Y,X,\theta_o)\big|X\big] = \frac{\delta(X)}{\sqrt{b}},$$

with $\delta(X) \neq 0$. Under H_1, the limiting distribution of the proposed test statistic diverges, in which the power of the test is obtained.

Theorem 3.1. *Assume all the conditions of Theorem 2.2 hold and* $b \to \infty$*,* $n \to \infty$ *and* $b/n \to 0$*. Therefore:*

(i) *Under the fixed alternative* H_1*:*

$$\tilde{M}_b(x;\hat{\theta}_n) \to \infty.$$

[1]See also www.mcs.vuw.ac.nz/ ray/Brownian.

(ii) *Under the local alternatives* H_1^L*:*

$$\tilde{M}_b(x;\hat{\theta}_n) \Rightarrow B(F(x)) + \mathbb{E}[\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i \le x\}}].$$

Employing Theorem 3.1 and the continuous mapping theorem, then under the fixed alternative H_1: $KS_n \to \infty$, and $CM_n \to \infty$. Therefore, the consistency of the proposed test is obtained. Moreover, under the local alternatives H_1^L,

$$KS_n \Rightarrow \sup_{x \in \boldsymbol{R}^k} \left| B(F(x)) + \mathbb{E}[\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i \le x\}}] \right|,$$

and

$$CM_n \Rightarrow \int_{[0,1]^k} \left(B(F(x)) + \mathbb{E}[\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i \le x\}}] \right)^2 d\tau.$$

This shows that the proposed test has nontrivial powers against local alternatives H_1^L at rate $b^{-1/2}$. Note that there may exist local alternatives at rate $n^{-1/2}$ as follows.

$$H_2^L : \mathbb{E}\big[m(Y, X, \theta_o)\big|X\big] = \frac{\delta(X)}{\sqrt{n}}.$$

Under the local alternatives H_2^L, the limiting distribution of $\tilde{M}_b(x;\hat{\theta}_n)$ is the same as that under the null hypothesis (see Theorem 2.1). The proposed test is not powerful against local alternatives at rate $n^{-1/2}$.

Theorem 3.2. *Assume assumptions* [A1]-[A4] *hold and* $b \to \infty$*,* $n \to \infty$ *and* $b/n \to 0$*. Under the local alternatives* H_2^L*:*

$$\tilde{M}_b(x;\hat{\theta}_n) \Rightarrow B(F(x)).$$

4 Monte Carlo Simulations

The finite sample performance of the test statistic KS_n is examined. The following null DGPs is as follows.

(A) $y_i = x_{i1} + e_i$,

(B) $y_i = x_{i1} + 5 + e_i$,

(C) $y_i = x_{i1} + \exp(z_i) + e_i$,

(D) $y_i = x_{i2} + x_{i3} + e_i$,

(E) $y_i = x_{i2} + x_{i3} + 5 + e_i$,

(F) $y_i = x_{i2} + x_{i3} + \exp(z_i) + e_i$.

Here x_{i1}, x_{i2}, x_{i3} and z_i are i.i.d. $N(0,1)$ distribution and e_i is i.i.d. $N(0,\sigma_0^2)$ with $\sigma_0^2 = 1,2,3,4$. The test statistic KS_n for one regressor is:

$$KS_1 = \max_j \left| \frac{1}{\sqrt{b}} \hat{\sigma}_1^{-1} \sum_{i=1}^{b} (y_i - x_{i1}' \hat{\beta}_1) \mathbf{1}_{\{x_{i1} \le x_j\}} \right|,$$

for DGPs (A), (B) and (C) and the test statistic for two regressors is:

$$KS_2 = \max_j \left| \frac{1}{\sqrt{b}} \hat{\sigma}_2^{-1} \sum_{i=1}^{b} (y_i - x_{i2}' \hat{\beta}_2 - x_{i3}' \hat{\beta}_3) \mathbf{1}_{\{x_{i2} \le x_{j1}\}} \mathbf{1}_{\{x_{i3} \le x_{j2}\}} \right|,$$

for DGPs (D), (E) and (F) where $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ are least square estimates, $\hat{\sigma}_1^2 = b^{-1} \sum_{i=1}^{b} (y_i - x_{i1}' \hat{\beta}_1)^2$ and $\hat{\sigma}_2^2 = b^{-1} \sum_{i=1}^{b} (y_i - x_{i2}' \hat{\beta}_2 - x_{i3}' \hat{\beta}_3)^2$. In each simulation experiment, the number of replications is 2000 and the significance level is 0.05. Different values of b are employed in this simulation. The choice of b is considered for the formula $b = n^p$ with $p = 0.5, 0.55, \cdots, 0.95$.

Table 1, given $\sigma_0^2 = 1$, reports the rejection frequencies of the tests for different values of n and p. For DGPs (A) and (D), the rejection values are finite sample sizes of the test. In the column of DGP (A), all values are close to the significance level 0.05 except for the values of $p = 0.5$. However, in the column of DGP (D), the proposed test is under-sized for a large p. Thus, when the number of regressors of the regression increases, or if b increases, then the finite sample sizes of the test are lower. For DGPs (B), (C), (E) and (F), the rejection rates are the finite sample powers of the proposed test. In columns of DGPs (B) and (E) that have fixed alternatives, the finite sample powers are 1. Thus, the test has good power performances with different values of n and b(or p). In addition, the values in the columns of DGPs (C) and (F) determine that the test performs well when the alternatives are a random variable. For DGP (C), there are good power performances of the test for large values of p. When n increases, the powers of the test are closer to 1. For DGP (F), finite sample powers are lower for $n = 100$, and as n and b(or p) increase, the finite sample powers increase. Thus, the proposed test has correct finite sample sizes for one regressor and is slightly under-sized for two regressors in the regression model. When there are fixed alternatives, the power performances are very good. The finite sample powers of the test increase along with both n and b. Table 2 reports the rejection frequencies of the test for six DGPs with different σ_0^2 and p. The sample size is 500. The finite sample performances in Table 2 are similar to those in Table 1. Moreover, when the variety of error term increases, the finite sample powers of the test decrease.

Table 1: Rejection frequencies of the conditional moment tests

		KS_1			KS_2		
n	p	(A)	(B)	(C)	(D)	(E)	(F)
100	0.50	0.076	1.000	0.742	0.057	1.000	0.596
	0.55	0.067	1.000	0.817	0.043	1.000	0.661
	0.60	0.057	1.000	0.877	0.037	1.000	0.775
	0.65	0.048	1.000	0.947	0.038	1.000	0.847
	0.70	0.044	1.000	0.977	0.036	1.000	0.919
	0.75	0.047	1.000	0.992	0.031	1.000	0.968
	0.80	0.049	1.000	0.999	0.024	1.000	0.985
	0.85	0.042	1.000	1.000	0.021	1.000	0.995
	0.90	0.046	1.000	0.999	0.025	1.000	0.999
	0.95	0.041	1.000	1.000	0.018	1.000	0.999
200	0.50	0.063	1.000	0.863	0.046	1.000	0.783
	0.55	0.045	1.000	0.937	0.038	1.000	0.867
	0.60	0.051	1.000	0.978	0.040	1.000	0.950
	0.65	0.053	1.000	0.992	0.031	1.000	0.981
	0.70	0.039	1.000	0.997	0.035	1.000	0.990
	0.75	0.054	1.000	1.000	0.023	1.000	0.998
	0.80	0.048	1.000	1.000	0.028	1.000	1.000
	0.85	0.043	1.000	1.000	0.023	1.000	1.000
	0.90	0.042	1.000	1.000	0.025	1.000	1.000
	0.95	0.049	1.000	1.000	0.022	1.000	1.000
500	0.50	0.068	1.000	0.964	0.057	1.000	0.950
	0.55	0.042	1.000	0.989	0.030	1.000	0.982
	0.60	0.047	1.000	0.998	0.035	1.000	0.994
	0.65	0.044	1.000	0.999	0.036	1.000	0.998
	0.70	0.045	1.000	1.000	0.040	1.000	0.999
	0.75	0.040	1.000	1.000	0.030	1.000	1.000
	0.80	0.055	1.000	1.000	0.028	1.000	1.000
	0.85	0.040	1.000	1.000	0.029	1.000	1.000
	0.90	0.036	1.000	1.000	0.019	1.000	1.000
	0.95	0.035	1.000	1.000	0.028	1.000	1.000

Note: The significant level is 0.05. $b = n^p$. The values in the 3rd and 6th columns are the finite sample sizes and the values in the 4th, 5th, 7th and 8th columns are the finite sample powers of the proposed test.

Table 2: Rejection frequencies of the conditional moment tests

		KS_1			KS_2		
σ_0^2	p	(A)	(B)	(C)	(D)	(E)	(F)
2	0.50	0.058	1.000	0.906	0.037	1.000	0.862
	0.55	0.053	1.000	0.972	0.041	1.000	0.955
	0.60	0.043	1.000	0.995	0.036	1.000	0.984
	0.65	0.044	1.000	0.999	0.037	1.000	0.998
	0.70	0.049	1.000	1.000	0.039	1.000	0.999
	0.75	0.044	1.000	1.000	0.029	1.000	1.000
	0.80	0.052	1.000	1.000	0.030	1.000	1.000
	0.85	0.040	1.000	1.000	0.033	1.000	1.000
	0.90	0.040	1.000	1.000	0.022	1.000	1.000
	0.95	0.036	1.000	1.000	0.023	1.000	1.000
3	0.50	0.060	1.000	0.843	0.050	1.000	0.797
	0.55	0.050	1.000	0.942	0.039	1.000	0.915
	0.60	0.054	1.000	0.987	0.037	1.000	0.972
	0.65	0.051	1.000	0.999	0.039	1.000	0.994
	0.70	0.048	1.000	1.000	0.034	1.000	0.999
	0.75	0.037	1.000	1.000	0.033	1.000	1.000
	0.80	0.046	1.000	1.000	0.038	1.000	1.000
	0.85	0.052	1.000	1.000	0.030	1.000	1.000
	0.90	0.040	1.000	1.000	0.023	1.000	1.000
	0.95	0.041	1.000	1.000	0.027	1.000	1.000
4	0.50	0.048	1.000	0.776	0.042	1.000	0.703
	0.55	0.053	1.000	0.897	0.042	1.000	0.845
	0.60	0.044	1.000	0.969	0.037	1.000	0.949
	0.65	0.042	1.000	0.992	0.031	1.000	0.988
	0.70	0.048	1.000	0.999	0.033	1.000	0.999
	0.75	0.046	1.000	1.000	0.035	1.000	1.000
	0.80	0.053	1.000	1.000	0.038	1.000	1.000
	0.85	0.047	1.000	1.000	0.027	1.000	1.000
	0.90	0.043	1.000	1.000	0.024	1.000	1.000
	0.95	0.042	1.000	1.000	0.029	1.000	1.000

Note: The significant level is 0.05. $b = n^p$. The values in the 3rd and 6th columns are the finite sample sizes and the values in the 4th, 5th, 7th and 8th columns are the finite sample powers of the proposed test.

Table 3: Empirical powers of tests

n	100		200		500	
	KS_2	ES	KS_2	ES	KS_2	ES
(G)	0.970	0.949	0.973	0.953	0.965	0.947
(H)	0.980	0.944	0.979	0.944	0.947	0.949

Note: The significant level is 0.05. $b = n^{0.8}$. The values are the finite sample powers of the proposed test and Escanciano's (2006) test.

Then the finite sample powers of the proposed test and Escanciano's (2006) test are compared. In Escanciano's test, the wild bootstrapping technique is required and the number of the wild bootstrapping in the simulation is 500. In addition, to make the computation simpler, $A_{ijr}^{(0)} = \pi$ is employed. Two DGPs with two regressors considered are:

(G) $y_i = (x_{i1} + x_{i2}) + (x_{it} + x_{i2})exp(-0.1(x_{i1} + x_{i2})^2) + e_i,$

(H) $y_i = (x_{i1} + x_{i2}) + x_{it}x_{i2} + e_i,$

with x_{i1}, x_{i2}, e_i i.i.d. $N(0, 1)$. The finite sample powers of the proposed test and Escanciano's test are reported in Table 3 with different sample sizes $n = 100, 200$, and 500. The finite sample powers of the proposed test are higher than those of Escanciano's test in all scenarios, except when $n = 500$ for DGP (H). This result shows that the proposed test has good finite sample power.

5 Conclusions

This paper proposes a consistent conditional moment test based on infinitely many unconditional moment restrictions. The test statistic is a subsampling marked empirical process and an asymptotically pivotal test is obtained. The proposed test is consistent against a general type of alternatives and is powerful against local alternatives at rates $b^{-1/2}$. In addition, the test performs well in finite sample simulations and the power performances are good with most values of b. However, the proposed test still suffers from choosing b and a future work might consider an optimal choice for b.

Appendix

Proof of Theorem 2.1. Given assumption [A3], the subsampling marked empirical process M_b permits the Taylor expansion:

$$\begin{aligned}&\frac{1}{\sqrt{b}}\sum_{i=1}^{b} m\big(Y_i, X_i, \hat{\theta}_n\big)\,\mathbf{1}_{\{X_i\le x\}}\\ &= \frac{1}{\sqrt{b}}\sum_{i=1}^{b} m\big(Y_i, X_i, \theta_o\big)\,\mathbf{1}_{\{X_i\le x\}} + \frac{1}{\sqrt{b}}\sum_{i=1}^{b} \nabla_\theta m\big(Y_i, X_i, \theta_o\big)\big(\hat{\theta}_n-\theta_o\big)\,\mathbf{1}_{\{X_i\le x\}} + o_p(1).\end{aligned}$$

Because $b/n \to 0$ and given assumption [A4],

$$\sqrt{b}(\hat{\theta}_n-\theta_o) = \sqrt{\frac{b}{n}}\sqrt{n}(\hat{\theta}_n-\theta_o) \xrightarrow{p} 0.$$

In addition, given assumptions [A1] and [A4], and Hölder's inequality, the following law of large numbers of i.i.d. sequence is:

$$\frac{1}{b}\sum_{i=1}^{b} \nabla_\theta m\big(Y_i, X_i, \theta_o\big)\,\mathbf{1}_{\{X_i\le x\}} \xrightarrow{p} \mathbb{E}\left[\nabla_\theta m\big(Y_i, X_i, \theta_o\big)\,\mathbf{1}_{\{X_i\le x\}}\right],$$

Then we obtain

$$\frac{1}{\sqrt{b}}\sum_{i=1}^{b} \nabla_\theta m\big(Y_i, X_i, \theta_o\big)\big(\hat{\theta}_n-\theta_o\big)\,\mathbf{1}_{\{X_i\le x\}} = \left[\frac{1}{b}\sum_{i=1}^{b} \nabla_\theta m\big(Y_i, X_i, \theta_o\big)\,\mathbf{1}_{\{X_i\le x\}}\right]\sqrt{b}(\hat{\theta}_n-\theta_o) \xrightarrow{p} 0.$$

Therefore,

$$\frac{1}{\sqrt{b}}\sum_{i=1}^{b} m\big(Y_i, X_i, \hat{\theta}_n\big)\,\mathbf{1}_{\{X_i\le x\}} = \frac{1}{\sqrt{b}}\sum_{i=1}^{b} m\big(Y_i, X_i, \theta_o\big)\,\mathbf{1}_{\{X_i\le x\}} + o_p(1).$$

$M_b(x;\hat{\theta}_n)$ and $M_b(x;\theta_o)$ are asymptotically equivalent. Thus, the estimating parameter θ does not affect the limiting distribution of the statistic and the estimation effect problem does not appear.

The process M_b belongs to the Shorohod space $D(\boldsymbol{R}^k)$ and the weak convergence of $M_b(x;\theta_o)$ in the space $D(\boldsymbol{R}^k)$ to a continuous limit is determined by the tightness of M_b and the finite dimensional convergence of $M_b(x;\theta_o)$. In the following, Bickel and Wichura (1971), Koul and Stute (1999) and Domínguez and Lobato (2006) are employed to show the tightness of M_b and then the weak convergence of $M_b(x;\theta_o)$. $I_1 = (s^1, t^1] = \times_{j=1}^{k}(s_j^1, t_j^1]$, and $I_2 = (s^2, t^2] = \times_{j=1}^{k}(s_j^2, t_j^2]$ are defined as the two subsets in $\boldsymbol{R}^k$. Then I_1 and I_2 are neighbor subsets if and only if for some $j^* \in \{1, 2, \cdots, k\}$, $(s_{j^*}^1, t_{j^*}^1] \neq (s_{j^*}^2, t_{j^*}^2]$, $\times_{j\neq j^*}^{k}(s_j^1, t_j^1] = \times_{j\neq j^*}^{k}(s_j^2, t_j^2]$ and $t_{j^*}^1 = s_{j^*}^2$. That is, they are next to each other and share the j^*th face.

Thus, the process M_b indexed by a parameter in $\boldsymbol{R}^k$ has an associated process indexed by the intervals as follows, wherein $h = 1, 2$,

$$\begin{aligned}
M_b(I_h;\theta) &:= \frac{1}{\sqrt{b}}\sum_{i=1}^{b} m(Y_i, X_i;\theta)\mathbf{1}_{\{X_i\in I_h\}} \\
&= \sum_{e_1=0}^{1}\cdots\sum_{e_k=0}^{1}(-1)^{k-\sum_{j=1,\cdots,k}e_j}M_b(s_1^h + e_1(t_1^h - s_1^h), \cdots, s_k^h + e_k(t_k^h - s_k^h);\theta),
\end{aligned}$$

which is the increment of M_b around I_h. Denote $m(Y_i, X_i;\theta) = m_i$. Employing Bickel and Wichura (1971, Theorem 3 and example II), if

$$\mathbb{E}\left(M_b(I_1;\theta)^2, M_b(I_2;\theta)^2\right) = \frac{1}{b^2}\mathbb{E}\left(\left[\sum_{i=1}^{b} m_i\mathbf{1}_{\{X_i\in I_1\}}\right]^2\left[\sum_{i=1}^{b} m_i\mathbf{1}_{\{X_i\in I_2\}}\right]^2\right).$$

is bounded, then for any $\lambda > 0$ and $\gamma > 1$,

$$P(M_b \geq \lambda) \leq \lambda^{-4}\mu(I_1 \cup I_2)^{\gamma},$$

with some measure μ. Thus, as show the process M_b is tight.

Under H_0 and given assumption [A1], when a subindex appears once in the summation, the corresponding term is zero by the law of iterated expectation and the i.i.d. assumption. Moreover, since I_1 and I_2 are disjoint sets, when a subindex appears more than twice, the corresponding term is zero. Therefore,

$$\begin{aligned}
&\mathbb{E}\left(M_b(I_1;\theta)^2, M_b(I_2;\theta)^2\right) \\
&= \frac{1}{b^2}\mathbb{E}\left[\sum_{i=1}^{b} m_i^2\mathbf{1}_{\{X_i\in I_1\}}\left(\sum_{j=1}^{i-1} m_j\mathbf{1}_{\{X_j\in I_2\}}\right)^2\right] + \frac{1}{b^2}\mathbb{E}\left[\sum_{i=1}^{b} m_i^2\mathbf{1}_{\{X_i\in I_2\}}\left(\sum_{j=1}^{i-1} m_j\mathbf{1}_{\{X_j\in I_1\}}\right)^2\right].
\end{aligned}$$

The first and the second terms in the above equation are similar and the only difference is the indexing set I_h; we then focus on the first term. Under H_0 and given assumption [A2](i),

$$\begin{aligned}
&\frac{1}{b^2}\sum_{i=1}^{b}\mathbb{E}\left[m_i^2\mathbf{1}_{\{X_i\in I_1\}}\left(\sum_{j=1}^{i-1} m_j\mathbf{1}_{\{X_j\in I_2\}}\right)^2\right] \\
&= \frac{1}{b^2}\sum_{i=1}^{b}\mathbb{E}\left[\sigma^2(X_i)\mathbf{1}_{\{X_i\in I_1\}}\left(\sum_{j=1}^{i-1} m_j\mathbf{1}_{\{X_j\in I_2\}}\right)^2\right] \\
&= \frac{1}{b^2}\sum_{i=1}^{b}\mathbb{E}\left[\int_{I_1}\sigma^2(u)f(u)du\left(\sum_{j=1}^{i-1} m_j\mathbf{1}_{\{X_j\in I_2\}}\right)^2\right].
\end{aligned}$$

Given Fubini's Theorem, the above equation equals to:

$$\frac{1}{b^2}\sum_{i=1}^{b}\int_{I_1}\mathbb{E}\left[\sigma^2(u)f(u)\left(\sum_{j=1}^{i-1}m_j\mathbf{1}_{\{X_j\in I_2\}}\right)^2\right]du.$$

Given Cauchy-Schwarz's inequality, the following is

$$\begin{aligned}&\frac{1}{b^2}\sum_{i=1}^{b}\int_{I_1}\mathbb{E}\left[\sigma^2(u)f(u)\left(\sum_{j=1}^{i-1}m_j\mathbf{1}_{\{X_j\in I_2\}}\right)^2\right]du\\&\leq\frac{1}{b^2}\sum_{i=1}^{b}\int_{I_1}\left[\left\{\mathbb{E}\left[\sigma^2(u)f(u)\right]^2\right\}^{1/2}\left\{\mathbb{E}\left(\sum_{j=1}^{i-1}m_j\mathbf{1}_{\{X_j\in I_2\}}\right)^4\right\}^{1/2}\right]du.\end{aligned}$$

Given Burkholder's inequality and the moment inequality yield, with some constant C,

$$\mathbb{E}\left(\sum_{j=1}^{i-1}m_j\mathbf{1}_{\{X_i\in I_2\}}\right)^4\leq C\,\mathbb{E}\left(\sum_{j=1}^{i-1}m_j^2\mathbf{1}^2_{\{X_i\in I_2\}}\right)^2\leq C(i-1)^2\,\mathbb{E}(m_1^4\mathbf{1}_{\{X_1\in I_2\}}).$$

Thus,

$$\begin{aligned}&\frac{1}{b^2}\sum_{i=1}^{b}\mathbb{E}\left[m_i^2\mathbf{1}_{\{X_i\in I_1\}}\left(\sum_{j=1}^{i-1}m_j\mathbf{1}_{\{X_j\in I_2\}}\right)^2\right]\\&\leq\frac{1}{b^2}\sum_{i=1}^{b}\int_{I_1}\left[\left\{\mathbb{E}\left[\sigma^2(u)f(u)\right]^2\right\}^{1/2}\left\{C(i-1)^2\,\mathbb{E}(m_1^4\mathbf{1}_{\{X_1\in I_2\}})\right\}^{1/2}\right]du\\&=\frac{1}{b^2}\left[C\,\mathbb{E}(m_1^4\mathbf{1}_{\{X_1\in I_2\}})\right]^{1/2}\sum_{i=1}^{b}(i-1)\int_{I_1}\left\{\mathbb{E}\left[\sigma^2(u)f(u)\right]^2\right\}^{1/2}du.\end{aligned}$$

$\mathbb{E}(m_1^4\mathbf{1}_{\{X_1\in I_2\}})\leq\mathbb{E}(m_1^4)$ which is bounded by assumption [A2] (ii). In addition, from Koul and Stute (1999), $\int_{I_1}\{\mathbb{E}[\sigma^2(u)f(u)]^2\}^{1/2}du$ is bounded by assumptions [A1] and [A2](iii). Therefore, under H_0 and given assumptions [A1]–[A2], the process M_b is tight. Note that our assumption [A2] (ii) and (iii) are similar to the assumption (A)(a) in Koul and Stute (1999).

Given assumptions [A1] and [A2] (i), and by a central limit theorem for i.i.d. sequence, we have for any $x\in\boldsymbol{R}^k$,

$$M_b(x;\theta_o)\Rightarrow N(0,V(x)).$$

For $x_1, x_2 \in \boldsymbol{R}^k$,

$$\begin{aligned}
&\operatorname{Cov}\big(M_b(x_1;\theta_o), M_b(x_2;\theta_o)\big) \\
&\quad = \frac{1}{b}\sum_{i=1}^{b} \mathbb{E}\big[m(Y_i, X_i;\theta_o)^2 \mathbf{1}_{\{X_i \le x_1\}} \mathbf{1}_{\{X_i \le x_2\}}\big] \\
&\quad \xrightarrow{p} \int_{-\infty}^{x_1 \wedge x_2} \sigma^2(u)\, F(\mathrm{d}u) \\
&\quad = V(x_1 \wedge x_2),
\end{aligned}$$

where the first equality holds by the property of i.i.d. sequence. Since $V(x)$ is nondecreasing and nonnegative, M_b is an asymptotically distributed $B(V(x))$, where $B(\cdot)$ is a multi-parameter Brownian motion process. □

Proof of Theorem 2.2. Herein, it is shown that a consistent estimator $\hat{\sigma}(X_i)^2$ to replace $\sigma(X_i)^2$ does not affect the asymptotics of the scale invariant subsampling marked empirical process. Thus, the process $\tilde{M}_b(x;\hat{\theta}_n)$ may be rewritten as:

$$\begin{aligned}
&\frac{1}{\sqrt{b}}\sum_{i=1}^{b} \hat{\sigma}(X_i)^{-1} m\big(Y_i, X_i, \hat{\theta}_n\big)\, \mathbf{1}_{\{X_i \le x\}} \\
&= \frac{1}{\sqrt{b}}\sum_{i=1}^{b} \big(\hat{\sigma}(X_i)^{-1} - \sigma(X_i)^{-1}\big) m\big(Y_i, X_i, \hat{\theta}_n\big)\, \mathbf{1}_{\{X_i \le x\}} + \frac{1}{\sqrt{b}}\sum_{i=1}^{b} \sigma(X_i)^{-1} m\big(Y_i, X_i, \hat{\theta}_n\big)\, \mathbf{1}_{\{X_i \le x\}}.
\end{aligned}$$

The first term of the above equation

$$\begin{aligned}
&\left|\frac{1}{\sqrt{b}}\sum_{i=1}^{b} \big(\hat{\sigma}(X_i)^{-1} - \sigma(X_i)^{-1}\big) m\big(Y_i, X_i, \hat{\theta}_n\big)\, \mathbf{1}_{\{X_i \le x\}}\right| \\
&\le \sqrt{b} \sup_{X_i} |\hat{\sigma}(X_i)^{-1} - \sigma(X_i)^{-1}| \sup_{Y_i, X_i} |m\big(Y_i, X_i, \hat{\theta}_n\big)|.
\end{aligned}$$

Therefore, given $\hat{\sigma}(X_i) - \sigma(X_i) = o_p(b^{-1/2})$ and assumption [A2] (i),

$$\frac{1}{\sqrt{b}}\sum_{i=1}^{b} \hat{\sigma}(X_i)^{-1} m\big(Y_i, X_i, \hat{\theta}_n\big)\, \mathbf{1}_{\{X_i \le x\}} = \frac{1}{\sqrt{b}}\sum_{i=1}^{b} \sigma(X_i)^{-1} m\big(Y_i, X_i, \hat{\theta}_n\big)\, \mathbf{1}_{\{X_i \le x\}} + o_p(1).$$

Let $\tilde{M}_b^{\sigma}(x;\theta) := b^{-1/2}\sum_{i=1}^{b} \sigma(X_i)^{-1} m(Y_i, X_i, \theta)\, \mathbf{1}_{\{X_i \le x\}}$. Similar to the proof of Theorem 2.1, replacing θ_o by $\hat{\theta}_n$ in $\tilde{M}_b^{\sigma}(x;\hat{\theta}_n)$ does not affect its asymptotics. It suffices to focus on the limiting behavior of $\tilde{M}_b^{\sigma}(x;\theta_o)$. The tightness of the process can be obtained in Theorem 2.1 as $\sigma(X_i)^2$ is continuous. Using Lindeberg-Lévy central limit theorem for i.i.d. sequence, we obtain the limiting distribution, which is a Gaussian process with zero mean and for

$x_1, x_2 \in \boldsymbol{R}^k$,

$$\begin{aligned}
&\mathrm{Cov}\Big(\tilde{M}_b^\sigma(x_1,\theta_o), \tilde{M}_b^\sigma(x_2,\theta_o)\Big)\\
&\quad = \frac{1}{b}\sum_{i=1}^{b} \mathbb{E}\big[\sigma(X_i)^{-2} m(Y_i, X_i; \theta_o)^2 \mathbf{1}_{\{X_i \le x_1\}} \mathbf{1}_{\{X_i \le x_2\}}\big]\\
&\quad \xrightarrow{p} \int_{-\infty}^{x_1 \wedge x_2} F(\mathrm{d}u)\\
&\quad = F(x_1 \wedge x_2).
\end{aligned}$$

Hence, $\tilde{M}_b(x;\hat{\theta}_n) \Rightarrow B(F(x))$, with B a multi-parameter Brownian motion process. □

Proof of Theorem 3.1. Following Theorem 2.2, $\tilde{M}_b(x;\hat{\theta}_n)$ and $\tilde{M}_b^\sigma(x;\theta_o)$ are asymptotically equivalent. It suffices to discuss the limit of $\tilde{M}_b^\sigma(x;\theta_o)$ under two different types of alternatives.

For part (i), $\tilde{M}_b^\sigma(X;\theta_o)$ may be rewritten as:

$$\begin{aligned}
&\frac{1}{\sqrt{b}}\sum_{i=1}^{b} \sigma(X_i)^{-1} m\big(Y_i, X_i, \theta_o\big) \mathbf{1}_{\{X_i \le x\}}\\
&= \frac{1}{\sqrt{b}}\sum_{i=1}^{b} \sigma(X_i)^{-1} \left[m\big(Y_i, X_i, \theta_o\big) - \mu(X_i)\right] \mathbf{1}_{\{X_i \le x\}} + \frac{1}{\sqrt{b}}\sum_{i=1}^{b} \sigma(X_i)^{-1}\mu(X_i)\mathbf{1}_{\{X_i \le x\}}.
\end{aligned}$$

Under H_1 and given assumptions [A1]–[A4], by the previous proofs, the first part of the above equation converges to $B(F(x))$. In addition, if $\mathbb{E}|\sigma(X_i)^{-1}\mu(X_i)\mathbf{1}_{\{X_i \le x\}}| < \infty$ from the i.i.d. assumption, the probability limit of $b^{-1/2}\sum_{i=1}^{b}\sigma(X_i)^{-1}\mu(X_i)\mathbf{1}_{\{X_i \le x\}}$ will be

$$\sqrt{b}\,\mathbb{E}[\sigma(X_i)^{-1}\mu(X_i)\mathbf{1}_{\{X_i \le x\}}].$$

As $b \to \infty$, $\tilde{M}_b^\sigma(X;\theta_o) \to \infty$, thus

$$\tilde{M}_b(x;\hat{\theta}_n) \to \infty.$$

For part (ii), $\tilde{M}_b^\sigma(X;\theta_o)$ may be rewritten as:

$$\begin{aligned}
&\frac{1}{\sqrt{b}}\sum_{i=1}^{b} \sigma(X_i)^{-1} m\big(Y_i, X_i, \theta_o\big) \mathbf{1}_{\{X_i \le x\}}\\
&= \frac{1}{\sqrt{b}}\sum_{i=1}^{b} \sigma(X_i)^{-1} \left[m\big(Y_i, X_i, \theta_o\big) - \frac{\delta(X_i)}{\sqrt{b}}\right] \mathbf{1}_{\{X_i \le x\}} + \frac{1}{b}\sum_{i=1}^{b} \sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i \le x\}}.
\end{aligned}$$

Under H_1^L and given assumptions [A1]–[A4], if $\mathbb{E}|\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i \le x\}}| < \infty$, the probability limit of $b^{-1}\sum_{i=1}^{b}\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i \le x\}}$ will be $\mathbb{E}[\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i \le x\}}]$. Therefore, under H_1^L, $\tilde{M}_b(x;\hat{\theta}_n)$ converges to a multiparameter Brownian motion process plus a non-zero constant term $\mathbb{E}[\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i\le x\}}]$. $\square$

Proof of Theorem 3.2. Proof of Theorem 3.2 is similar to the proof of Theorem 3.1 wherein $\tilde{M}_b^{\sigma}(X;\theta_o)$ may be rewritten as:

$$\begin{aligned}&\frac{1}{\sqrt{b}}\sum_{i=1}^{b}\sigma(X_i)^{-1}m\big(Y_i,X_i,\theta_o\big)\mathbf{1}_{\{X_i\le x\}}\\&=\frac{1}{\sqrt{b}}\sum_{i=1}^{b}\sigma(X_i)^{-1}\left[m\big(Y_i,X_i,\theta_o\big)-\frac{\delta(X_i)}{\sqrt{n}}\right]\mathbf{1}_{\{X_i\le x\}}+\frac{1}{\sqrt{b}\sqrt{n}}\sum_{i=1}^{b}\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i\le x\}}.\end{aligned}$$

The probability limit of the second term on the right-hand-side of the above equation will be

$$\frac{1}{\sqrt{b}\sqrt{n}}\sum_{i=1}^{b}\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i\le x\}}=\frac{\sqrt{b}}{\sqrt{n}}\left[\frac{1}{b}\sum_{i=1}^{b}\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i\le x\}}\right]\xrightarrow{p}0,$$

with $b/n\to 0$ and $b^{-1}\sum_{i=1}^{b}\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i\le x\}}\xrightarrow{p}\mathbb{E}[\sigma(X_i)^{-1}\delta(X_i)\mathbf{1}_{\{X_i\le x\}}]$. Therefore, $\tilde{M}_b(x;\hat{\theta}_n)$ converges to a multi-parameter Brownian motion process under both H_0 and H_2^L. $\square$

References

Andrews, D. and P. Guggenberger (2005). Hybrid and size-corrected subsample methods, unpublished manuscript, Cowels Foundation, Yale University.

Bickel, P. and M. Wichura (1971). Convergence criteria for multiparameter stochastic processes and some applications, *Annals of Mathematical Statistics*, **42**, 1656–1670.

Bierens, H. (1982). Consistent model specification tests, *Journal of Econometrics*, **20**, 105–134.

——— (1984). Model specification testing of time series regressions, *Journal of Econometrics*, **26**, 323–353.

——— (1990). A consistent conditional moment test of functional form, *Econometrica*, **58**, 1443–1458.

Bierens, H. and D. Ginther (2001). Integrated conditional moment testing of quantile regression models, *Empirical Economics*, **26**, 307–324.

Bierens, H. and W. Ploberger (1997). Asymptotic theory of integrated conditional moment tests, *Econometrica*, **65**, 1129–1151.

Chernozhukov, V. and I. Fernández-Val (2005). Subsampling inference on quantile regression process, *Sankha*, **67**, 253–276.

de Jong, R. (1996). On the Bierens test under data dependence, *Journal of Econometrics*, **72**, 1–32

Delgado, M. and W. González-Manteiga (2001). Significance testing in nonparametric regression based on the bootstrap, *Annals of Statistics*, **29**, 1469–1507.

Domínguez, M and I. Lobato (2006). Consistent estimation of models defined by conditional moment restrictions, *Econometrica*, **72**, 1601–1615.

Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated, *Annals of Statistics*, **1**, 279–290.

Escancciano, C. (2006). A consistent diagnostic test for regression models using projections, *Econometrics Theory*, **22**, 1030–1051.

Guggenberger, P. and M. Wolf (2004). Subsampling tests of parameter hypotheses and overifentifying restrictions with possible failure of identification, working paper, Department of Economics, UCLA.

Hong, H. and O. Scaillet (2006). A fast subsampling method for nonlinear dynamics models, *Journal of Econometrics*, **133**, 557–578.

Horowitz, J. and V. Spokoiny (2001). An adaptive, rate-optimal test of a parametric mean=regression model against a nonparametric alternative, *Econometrica*, **69**, 599–631.

Khmaladze, E. (1981). Martingale approach in the theory of goodness-of-fit tests, *Theory of Probability and its applications*, **XXVI**, 240–257.

Khmaladze, E. and H. Koul (2004). Martingale transforms goodness-of-fit tests in regression models, *Annals of statistics*, **32**, 995–1034.

Koul, H. and E. Stute (1999). Nonparametric model checks for time series, *Annals of Statistics*, **27**, 204–236.

Lavergne, P. and V. Patilea (2008). Breaking the curse of dimensionality in nonparametric testing, *Journal of Econometrics*, **143**, 103–122.

Li, W, C. Hsiao, and J. Zinn (2003). Consistent specification tests for semiparametric/nonparametric models based on series estimation methods, *Journal of Econometrics*, **112**, 295–325.

Linton, O., Maasoumi, E. and Y.-J. Whang (2005). Consistent testing for stochastic dominance under general sampling schemes, *Review of Economic Studies*, *72*, 735–765.

Newey, W. (1985). Maximum likelihood specification testing and conditional moment tests, *Econometrica*, **53**, 1047–1070.

Politis, D. N. and J. P. Romano (1994). Large sample confidence regions based on subsamples under minimal assumptions, *Annals of Statistics*, **22**, 2031–2050.

Politis, D. N., J. P. Romano, and M. Wolf (1999). *Subsampling*, New York: Springer.

Shorack, G. and J. Wellner (1986). *Empirical Processes with Applications to Statistics*, New York: John Wiley & Sons.

Song. K. (2009). Testing semiparametric conditional moment restrictions using conditional martingale transforms, *Journal of Econometrics*, forthcoming.

Stichcombe, M. and H. White (1998). Consistent specification testing with nuisance parameters present only under the alternative, *Econometric theory*, **14**, 295–325.

Stute, W. (1997). Nonparametric model checks for regression, *Annals of Statistics*, **25**, 613–641.

Stute, W., W. González Manteiga and M. Presedo Quindïmil (1998). Bootstrap approximations in model check for regression, *Journal of the American Statistical Association*, **93**, 141–149.

Stute, W., S. Thies, and L. Zhu (1998). Model checks for regression: an innovation process approach, *Annals of Statistics*, **26**, 1916–1934.

Stute, W. and L. Zhu (2002). Model checks for generalized linear models, *Scandinavian Journal of Statistics*, **29**, 535–545.

Tauchen, G. (1985). Diagnostic testing and evaluation of maximum likelihood models, *Journal of Econometrics*, **30**, 415–443.

Tripathi, G. and Y. Kitamura (2003). Testing conditional moment restrictions, *Annals of Statistics*, **31**, 2059–2095.

Whang, Y. (2000). Consistent bootstrap tests of parametric regression functions, *Journal of Econometrics*, **98**, 27–46.

——— (2001). Consistent specification testing for conditional moment restrictions, *Economics Letters*, **71**, 299–306.

——— (2004). Consistent specification testing for quantile regression models, working paper.

White, H. (1987). Specification testing in dynamic models. In T. Bewley (ed.), *Advances in Econometrics–Fifth World Congress*, **1**, New York: Cambridge University Press.

Zheng, J. (2000). A consistent test of conditional parametric distributions, *Econometric Theory*, **16**, 667–691.

無研發成果推廣資料

98 年度專題研究計畫研究成果彙整表

計畫主持人：林馨怡			計畫編號：98-2410-H-004-056-				
計畫名稱：不受分配影響的模型設定之檢定方法							
成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	100	100	100%		本報告正在投稿中.
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：□已發表 □未發表之文稿 ■撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）