Exact optimal inference in regression models under heteroskedasticity and non-normality of unknown form

Jean-Marie Dufour [*] McGill University

Abderrahim Taamouti [†] Universidad Carlos III de Madrid

June 2009

[*] William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 8879; FAX: (1) 514 398 4938; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

[†] Economics Department, Universidad Carlos III de Madrid. Adresse and e-mail: Departamento de Economía Universidad Carlos III de Madrid Calle Madrid, 126 28903 Getafe (Madrid) España. TEL: +34-91 6249863; FAX: +34-91 6249329; e-mail: ataamout@eco.uc3m.es.

ABSTRACT

Simple point-optimal sign-based tests are developed for inference on linear and nonlinear regression models with non-Gaussian heteroskedastic errors. The tests are exact, distribution-free, robust to heteroskedasticity of unknown form, and may be inverted to build confidence regions for the parameters of the regression function. Since point-optimal sign tests depend on the alternative hypothesis considered, an adaptive approach based on a split-sample technique is proposed in order to choose an alternative that brings power close to the power envelope. The performance of the proposed *quasi-point-optimal* sign tests with respect to size and power is assessed in a Monte Carlo study. The power of quasi-point-optimal sign tests is typically close to the power envelope, when approximately 10% of the sample is used to estimate the alternative and the remaining sample to compute the test statistic. Further, the proposed procedures perform much better than common least-squares-based tests which are supposed to be robust against heteroskedasticity.

Keywords: sign test; point-optimal test; nonlinear model; heteroskedasticity; exact inference; distribution-free; power envelope; split-sample; adaptive method; projection.

Contents

1. Introduction **1**

2. General framework **2**

3. POS tests in linear and nonlinear regression models **5**
- 3.1. Testing the zero coefficient hypothesis in linear regressions 6
- 3.2. Testing general full coefficient hypotheses in nonlinear regressions 7

4. Choice of the optimal alternative hypothesis **9**
- 4.1. Power envelope of POS tests 9
- 4.2. An adaptive approach to choose an optimal alternative 11

5. POS confidence regions **15**

6. Monte Carlo study **19**
- 6.1. Simulated models 20
- 6.2. Simulation results 21

7. Conclusion **29**

A. Appendix: Proofs **31**

B. Appendix: Additional simulation results using nonlinear regression model **32**

C. Acknowledgments **32**

List of Tables

1 Power comparisons: different tests. Normal and Cauchy error distributions 22
2 Power comparisons: different tests. Mixture and normal distribution with break . 24
3 Power comparisons: different tests. GARCH error distributions 26
4 True weights versus normal weights 30

List of Figures

1 Power comparisons: different alternatives. Normal and Cauchy error distributions 12
2 Power comparaisons: different alternatives. Mixture and normal error distribution with break 13
3 Power comparisons: different alternatives. GARCH error distributions 14
4 Power comparisons: different sample splits . Normal and Cauchy error distribution 16
5 Power comparisons: different sample splits. Mixture and normal error distribution with break 17
6 Power comparisons: different sample splits. GARCH error distributions 18
7 Power comparisons: different tests. Normal and Cauchy error distributions 23
8 Power comparisons: different tests. Mixture and normal error distribution with break 25
9 Power comparisons: different tests. GARCH error distributions 27
10 Power comparisons: different tests. Student and normal error distribution with exponential variance 28
11 Power comparaison using different tests: Nonlinear regression model 33

1. Introduction

Regression errors in economic data frequently exhibit non-normal distributions and heteroskedasticity. In the presence of several types of heteroskedasticity, usual "robust" tests – such as tests based on White (1980)-type variance corrections - remain plagued by poor size control and/or low power. This is the case, in particular, when there is a break in the disturbance variance or with a GARCH structure with one or several outliers. Further, the available *exact parametric* tests typically assume Gaussian disturbances. The latter assumption is often unrealistic and, in the presence of heavy tails and asymmetric distributions, the associated tests may easily not perform well in terms of size control or power. Furthermore, statistical procedures for inference on parameters of *nonlinear* models are typically based on asymptotic approximations, which may easily not be reliable in finite samples [see Dufour (2003)].

The present paper proposes simple point-optimal sign-based tests in linear and nonlinear regression models, which are valid under non-normality and heteroskedascticity of unknown form, provided the errors have median zero conditional on the explanatory variables. The proposed tests are exact, distribution-free, robust against heteroskedasticity of unknown form, and may be inverted to build confidence regions for the vector of unknown parameters. The setup and the type of procedures we consider are motivated in at least two ways.

First, it is well known that hypotheses on means (or moments) are not testable in nonparametric setups even under the apparently restrictive assumption that observation are independent and identically distributed (*i.i.d.*): if a test has level α for testing the null hypothesis that the mean of *i.i.d.* observations has a given value, then its power cannot larger than the level α under any alternative of the mean; see Bahadur and Savage (1956). Similar results hold for coefficients of regression models; see Dufour, Jouneau and Torrès (2008). In other words, moments are not empirically meaningful in many common nonparametric models. This provides a strong reason for focusing on quantile parameters (such as medians) in nonparametric models – instead of moments – because quantiles are not affected by such problems of nontestability.

Second, in the presence of general heteroskedasticity, Lehmann and Stein (1949) and Pratt and Gibbons (1981) show that sign methods are the only possible way of producing valid inference in finite samples; see also Dufour and Hallin (1991) and Dufour (2003). If a test has level α for testing the null hypothesis that observations are independent each with a distribution symmetric about zero, then its level must be equal to α conditional on the absolute values of the observations: in other words, it must be a *sign test*. For a more detailed discussion of statistical inference impossibilities in nonparametric models, see Dufour (2003) and Dufour et al. (2008).

A number of sign-based test procedures have been developed in the literature. In the presence of only one explanatory variable, Campbell and Dufour (1991, 1995, 1997) propose nonparametric analogues of the *t-test*, based on sign and signed rank statistics, which are applicable when regressors involve feedback of the type considered by Mankiw and Shapiro (1986). These tests are exact even when the disturbances are asymmetric, non-normal, and heteroskedastic. Boldin, Simonova and Tyurin (1997) propose locally optimal sign-based inference and estimation for linear models. Coudin and Dufour (2008) extend the work by Boldin et al. (1997) to account for serial dependence and discrete distributions. Wright (2000) proposes variance-ratio tests based on the signs and ranks to test the null hypothesis that the series of interest is a martingale difference sequence. For other sign-based test procedures, the reader can consult Capanu, Jones and Randles (2006) and Gerard and Schucany (2007) among others.

The present paper focuses on the optimality of sign tests and derives point-optimal tests based on sign statistics. Point-optimal tests are useful in a number of ways and they are particularly attractive when testing an economic theory against another one. An important feature of these tests comes from the fact that they trace out the *power envelope*, *i.e.* the maximum achievable power for a given testing problem. The power envelope provides an obvious benchmark against which test procedures can be evaluated. An early review and discussion of point-optimal tests is available in King (1987-88). More recently, this technique has been exploited in several papers in order to improve power. Dufour and King (1991) use point-optimal tests to do inference on the autocorrelation coefficient of a linear regression model with first-order autoregressive normal disturbances. Elliott, Rothenberg and Stock (1996) derive the asymptotic power envelope for point-optimal tests of a unit root in the autoregressive representation of a Gaussian time series under various trend specifications. Jansson (2005) derives an asymptotic Gaussian power envelope for tests of cointegration and proposes a feasible point-optimal cointegration test whose local asymptotic power function is found to be close to the asymptotic Gaussian power envelope. Begum and King (2005) propose a new approach for testing a composite null against a composite alternative hypothesis based on the generalized Neyman-Pearson lemma and maximizes average power subject to controlling average size over different subsets of the null hypothesis parameter space. Liang, Huang and Yang (2008) suggests locally optimal tests for exponential distributions with type-I censoring.

Since point-optimal sign (hereafter POS) tests depend on the alternative hypothesis, we propose an adaptive approach based on a split-sample technique [Dufour and Torrès (1998), Dufour and Jasiak (2001)] to choose an alternative that makes the power curve of the POS test close to the power envelope. The idea consists in dividing the sample into two independent parts and use the first one to estimate the value of the alternative hypothesis and the second to compute the POS test statistic [Dufour and Taamouti (2003), Dufour and Iglesias (2008)]. The simulation results show that using approximately 10% of sample to estimate the alternative yields a power function which is typically very close to the power envelope. We present a Monte Carlo study assessing the performance of the proposed "quasi-POS" test by comparing their size and power to those of some common tests which are supposed to be robust against heteroskedasticity. The results show that our procedures work quite well.

The plan of the paper is as follows. In Section 2, we present a general framework for deriving POS tests. In Section 3, we propose POS tests in the context of linear and nonlinear regression models. In Section 4, we study the power properties of the POS tests and propose an adaptive approach to choose an optimal alternative. In Section 5, we discuss the construction of the POS confidence regions using projection techniques. In Section 6, we present a Monte Carlo study assessing the performance of POS tests by comparing their size and power to those of some popular tests. We conclude in Section 7. Proofs are presented in Appendix A.

2. General framework

In this section, we describe a framework for deriving POS tests in the context of general hypothesis testing problem. The point-optimal tests are useful in a number of ways and they are most attractive for problems in which the parameter space can be restricted by theoretical considerations. They would ensure optimal power at given point and, depending on the structure of the problem, they can have power over the entire parameter space.

We consider here a random sample $\{y_t\}_{t=1}^{n}$ such that

$$\begin{aligned} &y_1, \ldots , \ y_n \text{ are independent with} \\ &\mathsf{P}[y_t \geq 0] = p_t \,, \ t = 1, \ldots , n. \end{aligned} \tag{2.1}$$

We define the following vector of signs

$$U(n) = \big(s(y_1), \ldots , s(y_n)\big)'$$

where

$$s(y_t) = \begin{cases} 1, if \ y_t \geq 0 \\ 0, if \ y_t < 0 \end{cases}, \ t = 1, \ldots , n.$$

We assume also that the y_t have no mass at zero, *i.e.*

$$\mathsf{P}[y_t = 0] = 0 \,, \ t = 1, \ldots , n, \tag{2.2}$$

which holds automatically when each y_t has a continuous distribution.

We wish to test the null hypothesis

$$H_0 : \mathsf{P}[s(y_t) = 1] = p_{t0}, \ t = 1, \ldots , n, \tag{2.3}$$

where $0 < p_{t0} < 1$, $t = 1, \ldots , n$, against the alternative hypothesis

$$H_1 : \mathsf{P}[s(y_t) = 1] = p_{t1}, \ t = 1, \ldots , n, \tag{2.4}$$

where $0 < p_{t1} < 1$, $t = 1, \ldots , n$. We consider optimal tests (in the Neyman-Pearson sense) which maximize the power function under the constraint $\mathsf{P}\,[\text{reject } H_0 \mid H_0] \leq \alpha$; see Lehmann (1959, page 65). The latter allows one to work with the log-likelihood function and simplify the expression of POS test statistics. The following theorem gives a POS test to test the null hypothesis H_0 against the alternative hypothesis H_1.

Theorem 2.1 *Under the assumptions* (2.1)-(2.2), *let* H_0 *and* H_1 *be defined by* (2.3) - (2.4),

$$S_n[p_0(n), \, p_1(n)] = \sum_{t=1}^{n} \ln \left[\frac{p_{t1}(1 - p_{t0})}{p_{t0}(1 - p_{t1})}\right] \, s(y_t) \tag{2.5}$$

where $p_0(n) = \big(p_{10}, \ldots , p_{n0}\big)'$ *and* $p_1(n) = \big(p_{11}, \ldots , p_{n1}\big)'$, *and suppose the constant* c_1 *satis-*

fies $\mathsf{P}\left[S_n[p_0(n),\, p(n)] > c_1\right] = \alpha$ *under* H_0, *with* $0 < \alpha < 1$. *Then the test with critical region*

$$S_n[p_0(n),\, p_1(n)] > c_1 \tag{2.6}$$

is most powerful for testing H_0 *against* H_1 *among level-*α *tests based on the signs* $\big(s(y_1),\, \ldots\, ,\, s(y_n)\big)'$.

PROOF. The likelihood function of the random sample $\{y_t\}_{t=1}^{n}$ is

$$L\big(U(n),\, p(n)\big) = \prod_{t=1}^{n} \Big\{\mathsf{P}[y_t \geq 0]^{s(y_t)}(1 - \mathsf{P}[y_t \geq 0])^{1-s(y_t)}\Big\} \tag{2.7}$$

where $p(n) = \big(p_1,\, \ldots\, ,\, p_n\big)'$. Under H_0, $L\big(U(n),\, p(n)\big)$ takes the form

$$L\big(U(n),\, p_0(n)\big) = \prod_{t=1}^{n} p_{t0}^{s(y_t)}(1 - p_{t0})^{1-s(y_t)}, \tag{2.8}$$

while under H_1,

$$L\big(U(n),\, p_1(n)\big) = \prod_{t=1}^{n} p_{t1}^{s(y_t)}(1 - p_{t1})^{1-s(y_t)}. \tag{2.9}$$

The log-likelihood ratio is then

$$\ln\left\{\frac{L\big(U(n),\, p_1(n)\big)}{L\big(U(n),\, p_0(n)\big)}\right\} = \sum_{t=1}^{n} a_t(1|0)s(y_t) + b(n) \tag{2.10}$$

where

$$a_t(1|0) = \ln\left(\frac{p_{t1}}{p_{t0}}\right) - \ln\left(\frac{1-p_{t1}}{1-p_{t0}}\right)\,,\; b(n) = \sum_{t=1}^{n} \ln\left(\frac{1-p_{t1}}{1-p_{t0}}\right). \tag{2.11}$$

Using the Neyman-Pearson lemma [see Lehmann (1959, page 65)], the most powerful level-α test of H_0 against H_1 rejects H_0 when

$$\sum_{t=1}^{n} \ln\left[\frac{p_{t1}(1-p_{t0})}{p_{t0}(1-p_{t1})}\right] s(y_t) > c_1 \equiv c - b(n)\,.$$

$\square$

In the case where $p_{t1} = p_1$, $p_{t0} = p_0$, for all t, with $p_1 > p_0 > 0$, the critical region in (2.6) can be written as

$$\sum_{t=1}^{n} s(y_t) > c_1\,.$$

Similarly, for $p_{t1} = p_1, p_{t0} = p_0$ and $0 < p_1 < p_0$, the critical region (2.6) takes the form

$$\sum_{t=1}^{n} s(y_t) < \bar{c}_1$$

for some appropriate constant $\bar{c}_1$. In both cases, *i.e.* for $p_1 > p_0 > 0$ and $0 < p_1 < p_0$, the test statistic is

$$S_n = \sum_{t=1}^{n} s(y_t). \tag{2.12}$$

Under H_0, S_n follows a binomial distribution $Bi(n, p_0)$, *i.e.* $\mathsf{P}(S_n = k) = C_n^k p_0^k (1-p_0)^{n-k}$, where $C_n^k = n!/[k!(n-k)!]$. Since the test statistic (2.12) does not depend on the alternative hypothesis p_1, the above test corresponds to a *uniformly most powerful* test.

Example 2.1 BACKTESTING VALUE-AT-RISK Consider daily ex post portfolio returns, say R_t, and daily ex ante Value-at-Risk forecasts, say $VaR_t(p)$, with promised coverage rate p, such that $\mathsf{P}_{t-1}[R_t < VaR_t(p)] = p$. Define the hit sequence of $VaR_t(p)$ violations as

$$I_t = \begin{cases} 1, \text{ if } R_t < VaR_t(p) \\ 0, \text{ otherwise.} \end{cases}$$

Backtesting Value-at-Risk consists in testing whether the coverage rate of Value-at-Risk (VaR) is correct [see Christoffersen (1998)]. It is a key part of the internal model's approach to market risk management as laid out by the Basel Committee on Banking Supervision (1996). Testing the unconditional coverage of VaR is equivalent to testing the null hypothesis

$$H_0 : I_t \overset{iid}{\sim} B(p) \tag{2.13}$$

against the alternative hypothesis

$$H_1 : I_t \overset{iid}{\sim} B(\bar{p}) \tag{2.14}$$

where $B(p)$ represents a Bernoulli random variable such that $\mathsf{P}[B(p) = 1] = 1 - \mathsf{P}[B(p) = 0] = p$. Under H_0, the likelihood function of the random sequence $\{I_t\}_{t=1}^{T}$ is given by

$$L_0(I_1, \ldots , I_T, p) = \prod_{t=1}^{T} p^{I_t}(1-p)^{1-I_t} = p^{S_T}(1-p)^{n-S_T}$$

where $S_T = \sum_{t=1}^{T} I_t$. Under the alternative, the likelihood function is

$$L_1(I_1, \ldots , I_T, \bar{p}) = \bar{p}^{S_T}(1-\bar{p})^{n-S_T}.$$

Using the Neyman-Pearson lemma and the previous results, a test statistic for the null hypothesis (2.13) against the alternative hypothesis (2.14) is given by $S_T = \sum_{t=1}^{T} I_t$, where under $H_{0,}$ S_T follows a binomial distribution $Bi(T, p)$.

3. POS tests in linear and nonlinear regression models

This section proposes exact POS-based tests in the context of linear and nonlinear regression models where regressors can be taken as fixed. We consider in turn two problems. The first one consists in testing whether the conditional median of a vector of observation is zero against a linear regression alternative. The second one tests whether the coefficients of a possibly nonlinear median regression function have a given value against another nonlinear median regression. The first problem is a special case of the second one, but it will be useful from an expositional viewpoint to study the simpler problem first. Both problems can be viewed as special cases of the general setup in Section 2.

3.1. Testing the zero coefficient hypothesis in linear regressions

Suppose the variable y_t can be explained by a linear function of the vector x_t :

$$y_t = x_t^{'}\beta + \varepsilon_t\,,\ t = 1,\ \dots\ , n, \tag{3.1}$$

where x_t is a $k \times 1$ vector of explanatory variables, $\beta \in \mathbb{R}^k$ is an unknown parameter vector, and the errors $\varepsilon_1,\ \dots\ , \varepsilon_n$ are independent conditional on X with

$$\mathsf{P}[\varepsilon_t > 0 \mid X] = \mathsf{P}[\varepsilon_t < 0 \mid X] = \frac{1}{2},\ t = 1,\ \dots\ , n, \tag{3.2}$$

where $X = [x_1,\ \dots\ , x_n]^{'}$ is an $n \times k$ matrix. Note (3.2) entails that ε_t has no mass at zero, *i.e.* $\mathsf{P}[\varepsilon_t = 0 \mid X] = 0$ for all t.

We wish to test the null hypothesis

$$H_0 : \beta = 0 \tag{3.3}$$

against the alternative hypothesis

$$H_1 : \beta = \beta_1. \tag{3.4}$$

Under (3.1), the hypothesis testing problem given by (3.3)-(3.4) is a special case of the one defined by (2.3)-(2.4) where

$$p_t = \mathsf{P}[y_t \geq 0 \mid X] = 1 - \mathsf{P}[\varepsilon_t < -\beta^{'} x_t \mid X].$$

Under H_0,

$$p_{t0} = 1 - \mathsf{P}[\varepsilon_t < 0 \mid X] = \frac{1}{2} \tag{3.5}$$

while, under H_1,

$$p_{t1} = 1 - \mathsf{P}[\varepsilon_t < -\beta_1^{'} x_t \mid X]. \tag{3.6}$$

Thus, a POS test for the null hypothesis (3.3) against the alternative hypothesis (3.4) can be deduced from Theorem **2.1** using the equations (3.5)-(3.6). We then have the following result.

Proposition 3.1 *Under the assumptions* (3.1) *and* (3.2)*, let* H_0 *and* H_1 *be defined by* (3.3) - (3.4),

$$SL_n(\beta_1) = \sum_{t=1}^{n} a_t(\beta_1)\, s(y_t)$$

where

$$a_t(\beta_1) = \ln\left[\frac{1 - \mathsf{P}[\varepsilon_t \le -x_t'\beta_1 \mid X]}{\mathsf{P}[\varepsilon_t \le -x_t'\beta_1 \mid X]}\right], \tag{3.7}$$

and suppose the constant $c_1(\beta_1)$ *satisfies* $\mathsf{P}\left[\sum_{t=1}^{n} a_t(\beta_1)s(y_t) > c_1(\beta_1)\right] = \alpha$ *under* H_0*, with* $0 < \alpha < 1$. *Then the test that rejects* H_0 *when*

$$SL_n(\beta_1) > c_1(\beta_1) \tag{3.8}$$

is most powerful (*conditional on* X) *for testing* H_0 *against* H_1 *among level-*α *tests based on the signs* $\big(s(y_1), \ldots, s(y_n)\big)'$.

Under the null hypothesis, the signs $s(y_1), \ldots, s(y_n)$ are i.i.d. according to a Bernoulli $Bi(1, 0.5)$. So the distribution of the test statistic only depends on the weights $a_t(\beta_1)$ and thus does not involve any nuisance parameter under the null hypothesis. In view of the nonparametric nature of assumption (3.2), this means that tests based on $SL_n(\beta_1)$, such as the test given by (3.8), are distribution-free and robust against heteroskedasticity of unknown form. It is a nonparametric *pivotal function*. Under the alternative hypothesis, however, the power function of the test depends on the form of the distribution function of ε_t.

An interesting special case is the one where $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. according to a $N(0, 1)$ distribution. Then the optimal test statistic $SL_n(\beta_1)$ takes the form:

$$SL_n^*(\beta_1) = \sum_{t=1}^{n} \ln\left[\frac{\Phi(x_t'\beta_1)}{1 - \Phi(x_t'\beta_1)}\right] s(y_t) \tag{3.9}$$

where $\Phi(\cdot)$ is the standard normal distribution function.

In view of the above characterization of the distribution of $SL_n(\beta_1)$, its distribution can be simulated under the null hypothesis and the relevant critical values can be evaluated to any degree of precision with a sufficient number of replications. It is also possible to run exact Monte Carlo tests (corrected for the discrete nature of the test statistic) as described in Dufour (2006).

3.2. Testing general full coefficient hypotheses in nonlinear regressions

We consider now a nonlinear regression model:

$$y_t = f(x_t, \beta) + \varepsilon_t,\ t = 1, \ldots, n, \tag{3.10}$$

where x_t is an observable $k \times 1$ vector of fixed explanatory variables, $f(\cdot)$ is a scalar function, $\beta \in \mathbb{R}^k$ is an unknown vector of parameters, and the errors $\varepsilon_1, \ldots, \varepsilon_n$ are independent conditional on X with a distribution that satisfies (3.2). We do not require that the parameter vector β be identified.

We consider the problem of testing the null hypothesis

$$H(\beta_0) : \beta = \beta_0 \tag{3.11}$$

against the alternative hypothesis

$$H(\beta_1) : \beta = \beta_1. \tag{3.12}$$

A test for $H(\beta_0)$ against $H(\beta_1)$ can be constructed as in Section 3.1. First, we note that model (3.10) is equivalent to the transformed model

$$\tilde{y}_t = g(x_t,\, \beta,\, \beta_0) + \varepsilon_t,$$

where $\tilde{y}_t = y_t - f(x_t, \beta_0)$ and $g(x_t, \beta, \beta_0) = f(x_t, \beta) - f(x_t, \beta_0)$. Under assumption (2.1) and conditional on X, $\tilde{y}_1, \ldots , \tilde{y}_n$ are independent. Second, testing $H(\beta_0)$ against $H(\beta_1)$ is equivalent to testing

$$\bar{H}_0 : g(x_t,\, \beta, \beta_0) = 0,\; t = 1, \ldots , n,$$

against

$$\bar{H}_1 : g(x_t, \beta, \beta_0) = f(x_t, \beta_1) - f(x_t, \beta_0),\; t = 1, \ldots , n.$$

Finally, the likelihood function of new random sample $\{\tilde{y}_t\}_{t=1}^{n}$ is given by

$$L(\tilde{U}(n),\, \beta,\, X) = \prod_{t=1}^{n} \left\{ \mathsf{P}[\tilde{y}_t \geq 0 \mid X]^{s(\tilde{y}_t)} (1 - \mathsf{P}[\tilde{y}_t \geq 0 \mid X])^{1-s(\tilde{y}_t)} \right\}$$

where the elements of the sign vector $\tilde{U}(n) = \big(s(\tilde{y}_1), \ldots , s(\tilde{y}_n)\big)$ are

$$s(\tilde{y}_t) = \left\{ \begin{array}{l} 1, \text{ if } \tilde{y}_t \geq 0 \\ 0, \text{ if } \tilde{y}_t < 0 \end{array} \right. , \text{ for } t = 1, \ldots , n.$$

Thus, we can use the result of Proposition **3.1** to derive a sign-based test for the null hypothesis $H(\beta_0)$ against $H(\beta_1)$. This yields the following result.

Proposition 3.2 *Under the assumptions* (3.10) *and* (3.2), *let* $H(\beta_0)$ *and* $H(\beta_1)$ *be defined by* (3.11) - (3.12),

$$SN_n(\beta_0|\beta_1) = \sum_{t=1}^{n} \tilde{a}_t(\beta_0|\beta_1)\, s\left(y_t - f(x_t, \beta_0)\right) \tag{3.13}$$

where

$$\tilde{a}_t(\beta_0|\beta_1) = \ln \left[\frac{1 - p(x_t, \beta_0,\, \beta_1 \mid X)}{p(x_t, \beta_0,\, \beta_1 \mid X)} \right],$$

and suppose the constant $c_1(\beta_0,\, \beta_1)$ *satisfies* $\mathsf{P}\left[\sum_{t=1}^{n} a_t(\beta_1) s(y_t) > c_1(\beta_0,\, \beta_1)\right] = \alpha$ *under*

$H(\beta_0)$, with $0 < \alpha < 1$. Then the test that rejects $H(\beta_0)$ when

$$SN_n(\beta_0|\beta_1) > c_1(\beta_0,\, \beta_1)$$

is most powerful (conditional on X) for testing $H(\beta_0)$ against $H(\beta_1)$ among level-α tests based on the signs $\big(s(\tilde{y}_1),\, \ldots\, ,\, s(\tilde{y}_n)\big)^{'}$.

If we consider a linear function $f(x_t, \beta) = x_t^{'}\beta$ and assume that under the alternative hypothesis ε_t follows $N(0, 1)$, then the test statistic for the null hypothesis $H(\beta_0)$ against the alternative hypothesis $H(\beta_1)$ is given by:

$$SN_n^*(\beta_0|\beta_1) = \sum_{t=1}^{n} \ln\left[\frac{\Phi\big(x_t'(\beta_1 - \beta_0)\big)}{1 - \Phi\big(x_t'(\beta_1 - \beta_0)\big)}\right] s(y_t - x_t^{'}\beta_0) \tag{3.14}$$

where $\Phi(.)$ is the standard normal distribution function. The test statistic $SN_n^*(\beta_0|\beta_1)$ depends on a particular alternative hypothesis β_1. In practice, the latter is supposed to be unknown which makes the proposed POS test unfeasible. However, in the next section we propose a new approach which can be use to choose an optimal alternative β_1 at which the power of the test is maximized.

4. Choice of the optimal alternative hypothesis

In this section, we study the power properties of the proposed POS test. We derive its power envelope and analyze the impact of the alternative hypothesis β_1 on its power function. Since the latter depends on the alternative hypothesis, we propose an approach (hereafter adaptive approach) to choose the alternative β_1 at which the power of POS test is close to the power envelope.

4.1. Power envelope of POS tests

We derive an upper bound (hereafter power envelope) of the power function of POS test. It is well known, see for example King (1987-88), that point-optimal tests can be used to trace out the maximum attainable power envelope for a given testing problem. This power envelope provides a natural benchmark against which test procedures can be compared.

We know from Section 3 that the POS test statistic is a function of β_1

$$SN_n^*(\beta_0|\beta_1) = \sum_{t=1}^{n} \ln\left[\frac{1 - p(x_t, \beta_0,\, \beta_1 \mid X)}{p(x_t, \beta_0,\, \beta_1 \mid X)}\right] s\big(y_t - f(x_t, \beta_0)\big).$$

Its power function, say $\Pi(\beta,\, \beta_1)$, is also a function of β_1 :

$$\Pi(\beta,\, \beta_1) = \mathsf{P}\left[SN_n^*(\beta_0|\beta_1) > c_1(\beta_0,\, \beta_1)\right]$$

where $c_1(\beta_0,\, \beta_1)$ satisfies $\mathsf{P}[SN_n^*(\beta_0|\beta_1) > c_1(\beta_0,\, \beta_1) \mid H_0] \leq \alpha$. The following theorem provides a theoretical formula for power function of POS test.

Theorem 4.1 *Under assumptions* (2.1), (3.2) *and* (3.10)*, the power function of POS test at* β_1 *is given by*

$$\Pi(\beta, \beta_1) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\mathrm{Im}\left\{\exp\left(-iuc_1(\beta_0,\, \beta_1)\right) \phi_{SN_n^*}(u)\right\}}{u} du$$

where, for $u \in \mathbb{R}$,

$$\phi_{SN_n^*}(u) = \prod_{t=1}^{n} \left[1 + \left(\exp\left(iu\, \ln\left[\frac{1 - p(x_t, \beta_0,\, \beta_1 \mid X)}{p(x_t, \beta_0,\, \beta_1 \mid X)}\right]\right) - 1\right) \left(1 - p(x_t, \beta_0,\, \beta_1 \mid X)\right)\right]$$

$p(x_t, \beta_0,\, \beta_1 \mid X) = \mathsf{P}[\varepsilon_t \leq f(x_t, \beta_0) - f(x_t, \beta_1) \mid X]$, $i = \sqrt{-1}$, *and* $\mathrm{Im}\{z\}$ *denotes the imaginary part of a complex number* z*. The critical value* $c_1(\beta_0,\, \beta_1)$ *is chosen so that* $\mathsf{P}[SN_n^*(\beta_0|\beta_1) > c_1(\beta_0,\, \beta_1) \mid H_0] \leq \alpha$*, where* α *is an arbitrary significance level.*

The proof of this theorem is given in Appendix A. Since the test statistic $SN_n^*(\beta_0|\beta_1)$ is optimal against the alternative β_1, the envelope power function, say $\bar{\Pi}(\beta)$, is a function which associates the value $\Pi(\beta,\, \beta_1)$ to each element $\beta \in \mathbb{R}^k$:

$$\bar{\Pi}(\beta) = \Pi(\beta,\, \beta) = \mathsf{P}[SN_n^*(\beta) > c_1(\beta_0,\, \beta_1)]. \tag{4.1}$$

The objective now is to find a value of β_1 at which the power curve of POS test remains close to the relevant power envelope. For a given value Π of power function and level α of POS test, an alternative, say $\beta_1(\Pi, \alpha)$, can be determined by inverting the power envelope function $\bar{\Pi}(\beta)$. For any value $\Pi \in (\alpha, 1)$, the family of POS test statistics can be written as follows:

$$\left\{SN_n^*(\Pi) = \sum_{t=1}^{n} \ln\left[\frac{1 - p(x_t, \beta_0,\, \boldsymbol{\beta}_1(\mathbf{\Pi}, \boldsymbol{\alpha}) \mid X)}{p(x_t, \beta_0,\, \boldsymbol{\beta}_1(\mathbf{\Pi}, \boldsymbol{\alpha}) \mid X)}\right] s\left(y_t - f(x_t, \beta_0)\right), \text{ for } \Pi \in (\alpha, 1)\right\}.$$

Although every member of this family is admissible, it is possible that some values of Π may yield tests whose power functions lie close to the power envelope over a considerable range. Past research suggests that values of Π near one-half often have this property, see for example King (1987-88), Dufour and King (1991) and Elliott et al. (1996). Consequently, one can choose as an optimal alternative the one which corresponds to $\Pi = 0.5$. From Theorem **4.1** and equation (4.1), the value of β_1 which corresponds to $\Pi = 0.5$ is the solution of the following equation

$$\int_0^\infty \frac{\mathrm{Im}\left\{\exp\left(-iuc_1(\beta_0,\, \beta_1)\right) \phi_{SN_n^*}(u)\right\}}{u} du = 0 \tag{4.2}$$

where $c_1(\beta_0,\, \beta_1)$ and $\phi_{SN_n^*}(u)$ are defined in Theorem **4.1**. Using the properties of the cumulative density function (monotonically increasing, continuous, $\lim\limits_{c \to -\infty} \Pr(z < c) = 0$ and $\lim\limits_{c \to +\infty} P_t(z < c) = 1$) one can show that equation (4.2) has a unique solution. However, an exact solution for this equation is not feasible, since it is not easy to find an expression for $\mathrm{Im}\{\cdot\}$ and the integral $\int_0^\infty \mathrm{Im}\{\cdot\} du$ is difficult to evaluate. The latter can be approximated using results from Imhof

(1961), Davies (1973, 1980), among others, who propose a numerical approximation for the distribution function using the

characteristic function. The proposed approximation introduces two types of errors: discretization and truncation errors. Davies (1973), proposes a criterion to control for discretization error, and Davies (1980) proposes three different bounds to control for truncation error. Another alternative way to solve the power envelope function for β_1 is to use simulations [see Elliott et al. (1996)]. We can use simulations to approximate the power envelope function and calculate the optimal alternative which corresponds to the value of $\bar{\Pi}(\beta_1)$ near one-half.

Let us now examine the impact of the alternative hypothesis β_1 on the power function. Using simulations, we compare the power curves of POS test to the power envelope (PE) under different alternatives and data generating processes (hereafter DGPs). We consider a linear regression model with one regressor and an error term which follows one of the following distributions (DGPs): normal distribution, Cauchy distribution, mixture of normal and Cauchy distributions, and normal distribution with a break in variance. We also consider other DGPs [normal distribution with GARCH$(1, 1)$ plus jump variance and normal distribution with non stationary GARCH$(1, 1)$ variance] which do not satisfy they key assumption (2.1) and the results seem interesting. A more detailed description of these DGPs is given in Section 6. The simulations results [Figures 4.1-4.1] show that the alternative hypothesis β_1 affects the power function. Particularly, when β_1 is far from the null hypothesis, here $\beta = 0$, the power curve of POS test moves away from the power envelope curve.

Since the previous approach to finding the optimal alternative is somewhat arbitrary, in the next subsection we propose an adaptive approach based on split-sample technique to estimate the optimal alternative.

4.2. An adaptive approach to choose an optimal alternative

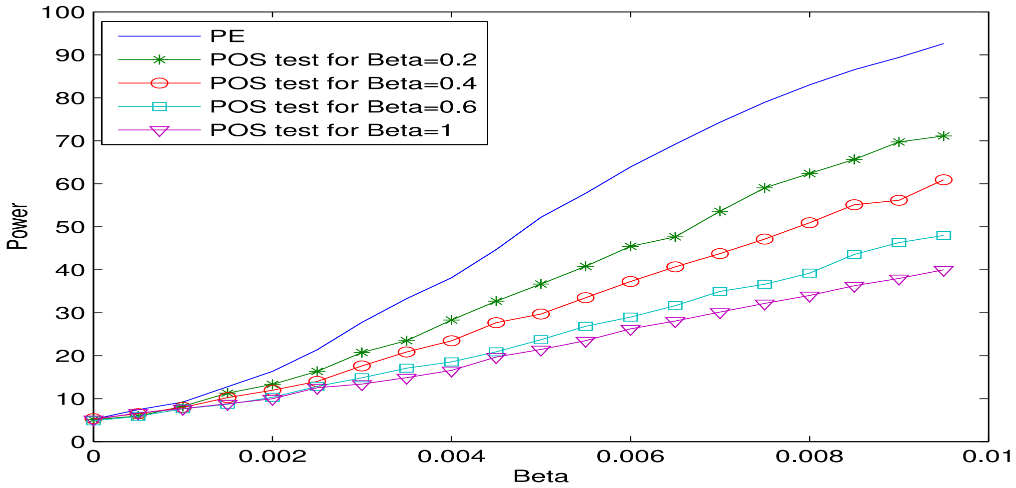
Existing adaptive statistical methods use the data to determine which statistical procedure is most appropriate for a specific testing problem. These methods usually involve two steps. In the first step a selection statistic is computed that estimates the shape of the error distribution. In the second step the selection statistic is used to determine an effective statistical procedure for the error distribution. For more details about the adaptive statistical methods, the reader can consult O'Gorman (2004).

The adaptive approach we consider here is an extension of the adaptive approach suggested in Dufour and Taamouti (2003) and Dufour and Iglesias (2008) for tests in parametric models involving nonstandard distributions. We propose a split-sample technique [Dufour and Jasiak (2001)] to choose β_1 such that the power of POS test is close to the power envelope. The alternative hypothesis β_1 is unknown and a practical problem consists in finding its independent estimate. To make size control easier, we estimate β_1 from a sample which is independent of the one used to compute the POS test statistic. This can be easily done by splitting the sample. The idea is to divide the sample into two independent parts and use the first one to estimate the value of the alternative and the second one to compute the POS test statistic.

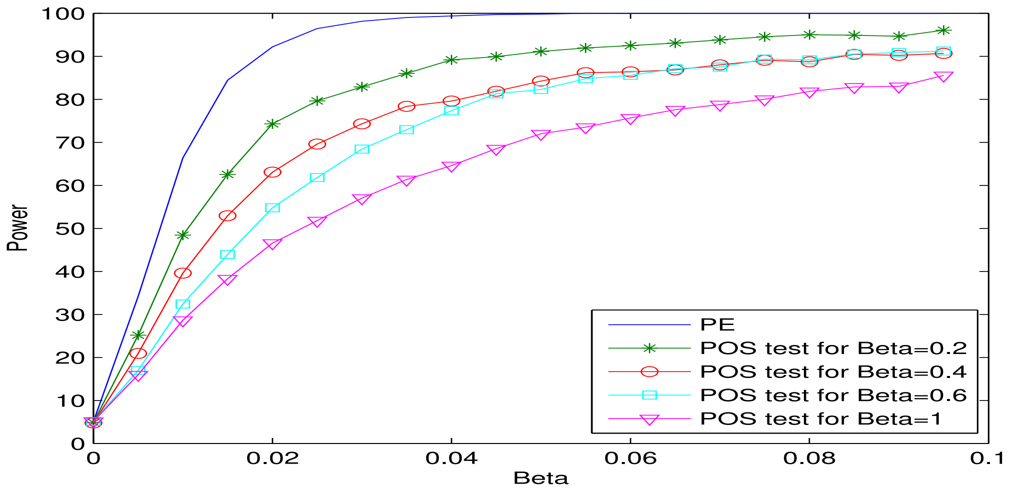
Let $n = n_1 + n_2$, $y = (y_{(1)}^{'}, y_{(2)}^{'})^{'}$, $X = (X_{(1)}^{'}, X_{(2)}^{'})^{'}$, and $\varepsilon = (\varepsilon_{(1)}^{'}, \varepsilon_{(2)}^{'})^{'}$, where the matrices $y_{(i)}$, $X_{(i)}$, and $\varepsilon_{(i)}$ have n_i, $i = 1, 2$, rows. When $f(x_t, \beta)$ is a linear function of β (linear regression

Figure 1. **Power comparisons: different alternatives**
Normal and Cauchy error distributions

A. Normal distribution



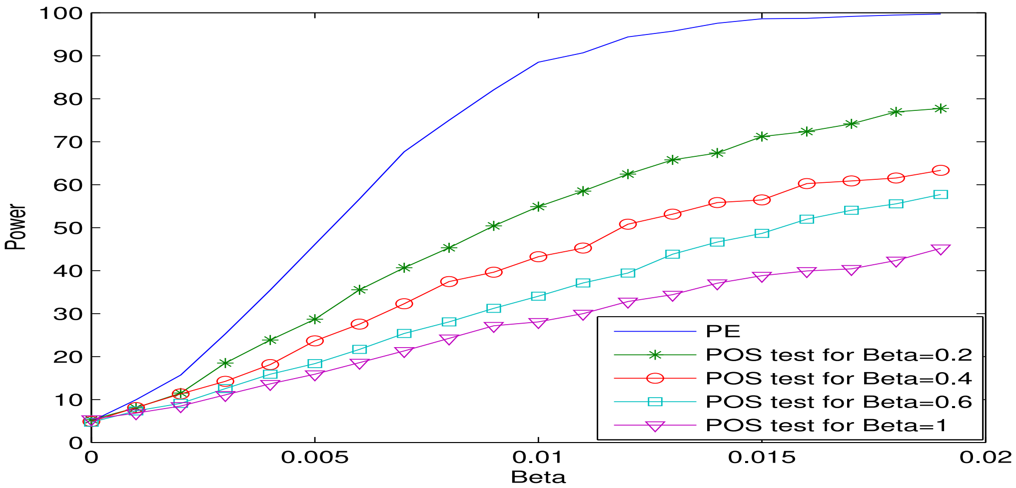
B. Cauchy distribution



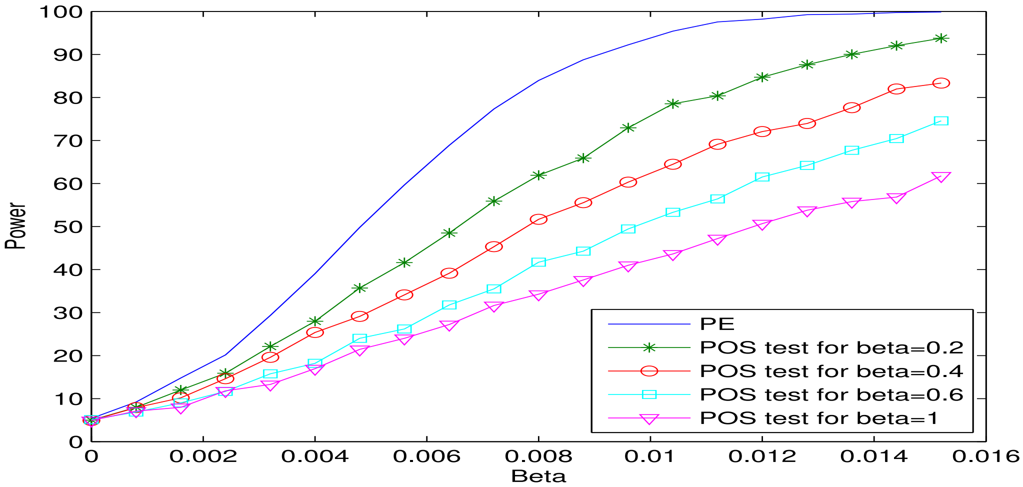
Note: These figures compare the power of POS test under different alternatives. Panel A corresponds to the case where the error term ε_t in the model (6.1) is homoskedastic and normally distributed. Panel B corresponds to the case where ε_t is homoskedastic and follows a Cauchy distribution. PE corresponds to the power envelop.

Figure 2. **Power comparisons: different alternatives**
Mixture and normal error distribution with break

A. Mixture distribution



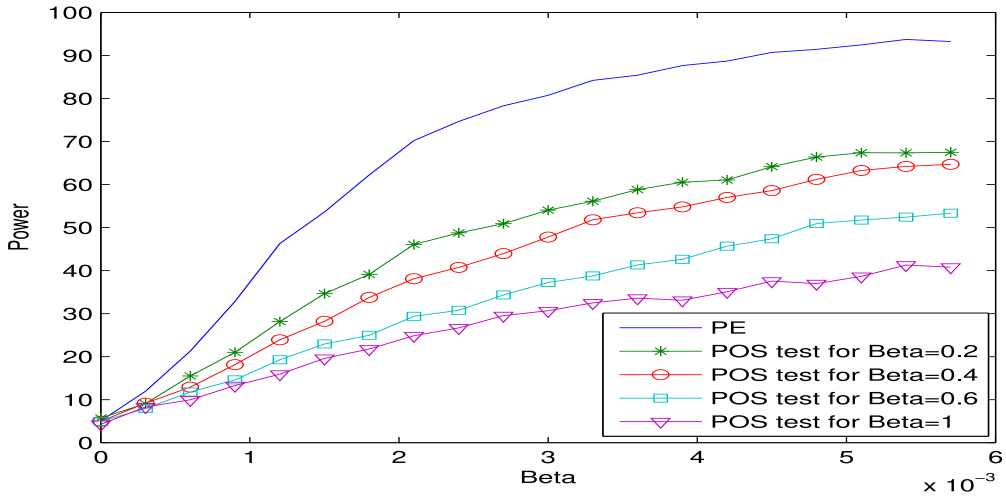
B. Normal distribution with break in variance



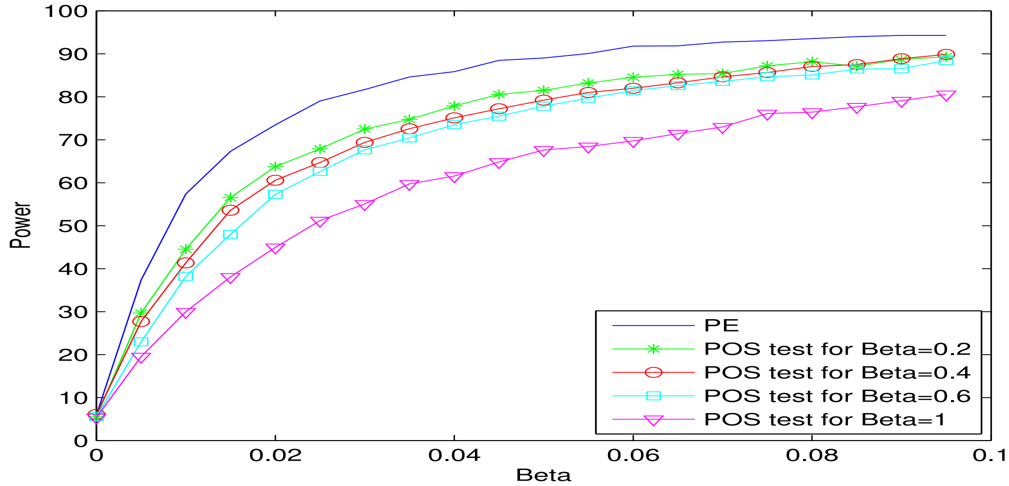
Note: These figures compare the power of POS test under different alternatives. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a mixture of normal and Cauchy distributions. Panel B corresponds to the case where ε_t follows a normal distribution with break in variance. PE corresponds to the power envelop.

Figure 3. **Power comparisons: different alternatives**
GARCH error distributions

A. Normal distribution with GARCH$(1,\ 1)$ plus jump variance



B. Normal distribution with non stationary GARCH$(1,\ 1)$ variance



Note: These figures compare the power of POS test under different alternatives. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a normal distribution with GARCH$(1,\ 1)$ plus jump variance and Panel B corresponds to the case where ε_t follows a normal distribution with non stationary GARCH$(1,\ 1)$ variance. PE corresponds to the power envelop.

model), we can use the first n_1 observations, $y_{(1)}$ and $X_{(1)}$, to estimate the alternative hypothesis β_1 using OLS

$$\hat{\beta}_{(1)} = (X'_{(1)} X_{(1)})^{-1} X'_{(1)} y_{(1)}.$$

Because $\hat{\beta}_1$ is independent of $X_{(2)}$, we can use the last n_2 observations, $y_{(2)}$ and $X_{(2)}$, to calculate the test statistic and get a valid POS test

$$SN_n^*(\beta_0|\hat{\beta}_1) = \sum_{t=n_1+1}^{n} \ln\left[\frac{1 - p(x_t, \beta_0,\, \hat{\beta}_{(1)} \mid X)}{p(x_t, \beta_0,\, \hat{\beta}_{(1)} \mid X)}\right] s(y_t - x_t'\beta_0)$$

where $p(x_t, \beta_0, \beta \mid X) = \mathsf{P}[\varepsilon_t \leq x_t'(\beta_0 - \beta) \mid X]$. However, the OLS estimator is known to be very sensitive to outliers and non-normal errors, consequently it is important to choose a more appropriate method to estimate β_1. In the presence of outliers many estimators are proposed to estimate the coefficients in regression model such that the least median of squares (LMS) estimator [see Rousseeuw and Leroy (1987)], the S-estimators [see Rousseeuw and Yohai (1984)], and the τ-estimators [see Yohai and Zamar (1988)].

Now, when $f(x_t, \beta)$ is a nonlinear function of β (nonlinear regression model), the above OLS method can not be used to estimate β_1. We will need to use for example nonlinear least squares or maximum likelihood method to estimate the alternative hypothesis β_1. This case will typically require an iterative procedure for solution. As for linear regression model, we can use the first n_1 observations, $y_{(1)}$ and $X_{(1)}$, to estimate the alternative hypothesis β_1 using nonlinear least squares method:

$$\hat{\beta}_1 = \arg\min_{\beta_1} \sum_{t=1}^{n_1} [y_t - f(x_t, \beta_1)]^2$$

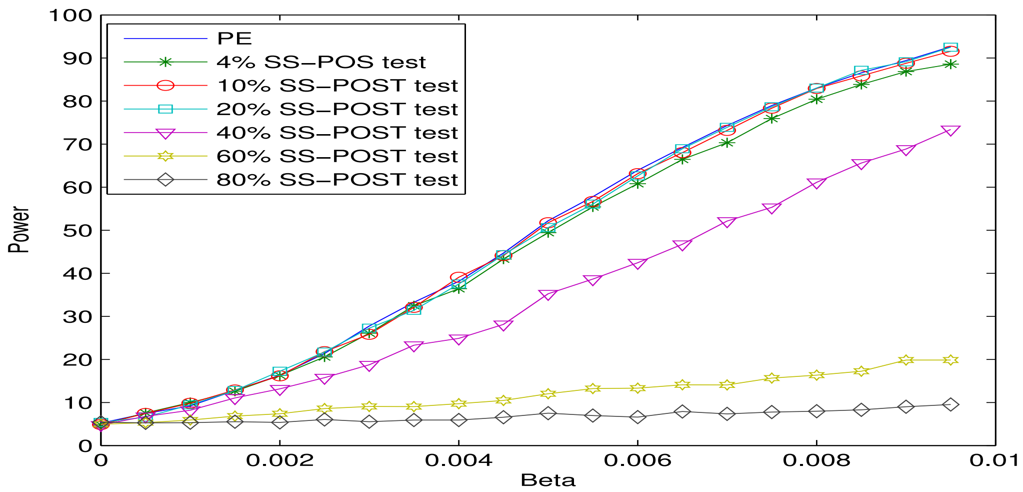
and the second last n_2 observations, $y_{(2)}$ and $X_{(2)}$, to calculate the test statistic:

$$SN_n^*(\beta_0|\hat{\beta}_1) = \sum_{t=n_1+1}^{n} \ln\left[\frac{1 - p(x_t, \beta_0,\, \hat{\beta}_{(1)} \mid X)}{p(x_t, \beta_0,\, \hat{\beta}_{(1)} \mid X) \mid X]}\right] s\left(y_t - f(x_t, \beta_0)\right)$$

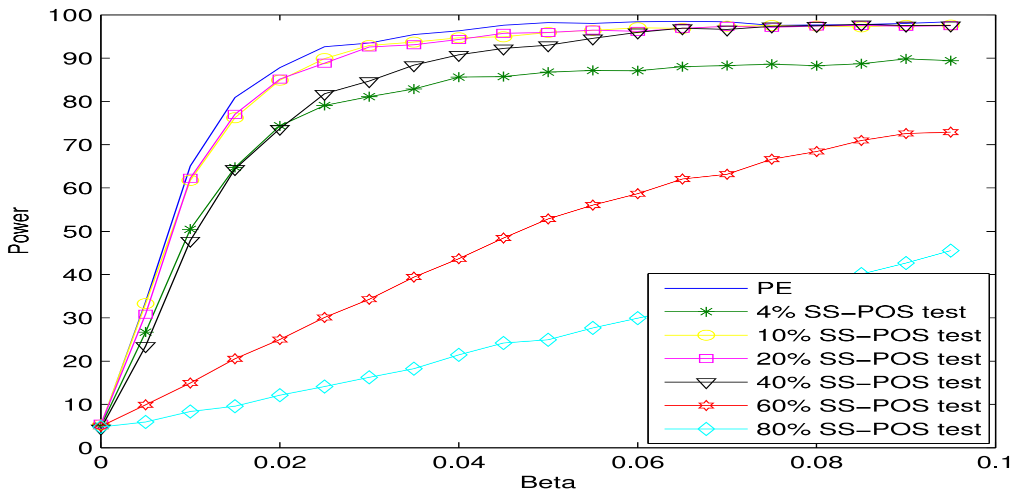
where $p(x_t, \beta_0, \beta \mid X) = \mathsf{P}[\varepsilon_t \leq f(x_t, \beta_0) - f(x_t, \beta) \mid X]$. Different choices for n_1 and n_2 are clearly possible. Alternatively, we could select randomly the observations assigned to the vectors $y_{(1)}$ and $y_{(2)}$. As we will show latter the number of observations retained for the first and the second subsamples have a direct impact on the power of the test. In particular, it seems that we could get more powerful test when we use a relatively small number of observations for computing the alternative hypothesis and keep more observations for the calculation of test statistic. This point is illustrated below in the context of a linear regression model. We use simulations to compare the power curves of split-sample-based POS test (hereafter SS-POS test) to the power envelope (hereafter PE) under different split-sample sizes and using different DGPs [see Section 6]. The results [Figures 4.2-4.2] show that using approximately 10% of sample to estimate the alternative yields a power which is typically very close to the power envelope. This is true for all DGPs considered in our simulation study.

Figure 4. **Power comparisons: different sample splits**
Normal and Cauchy error distributions

A. Normal distribution



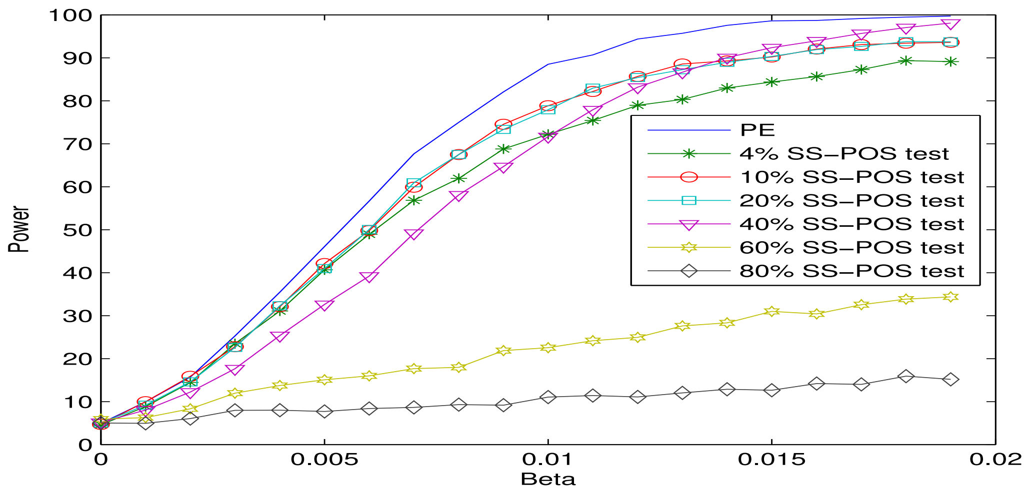
B. Cauchy distribution



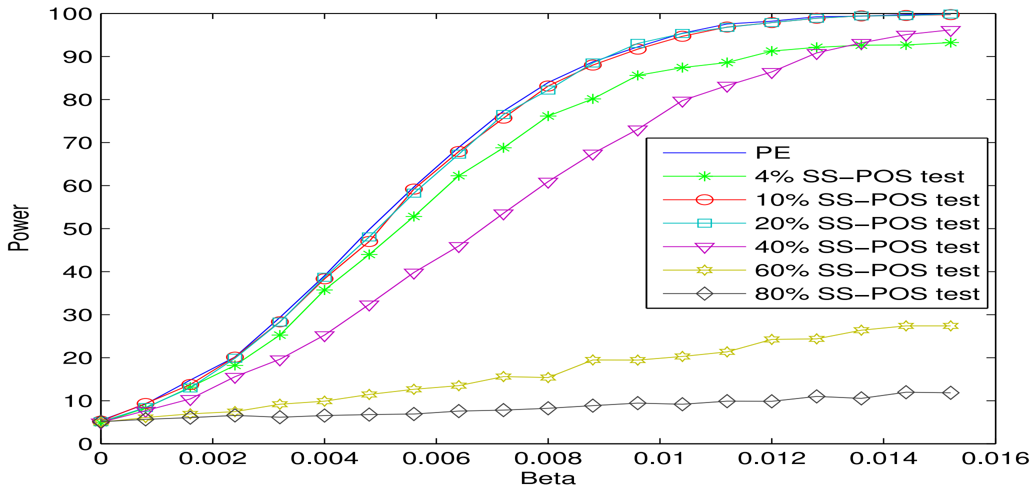
Note: These figures compare the power of POS test using different split-samples (SS-POS test); 4%, 10%, 20%, 40%, 60%, and 80%. Panel A corresponds to the case where the error term ε_t in the model (6.1) is homoskedastic and normally distributed. Panel B corresponds to the case where ε_t is homoskedastic and follows a Cauchy distribution. PE corresponds to the power envelop.

Figure 5. **Power comparisons: different sample splits**
Mixture and normal distribution with break

A. Mixture distribution



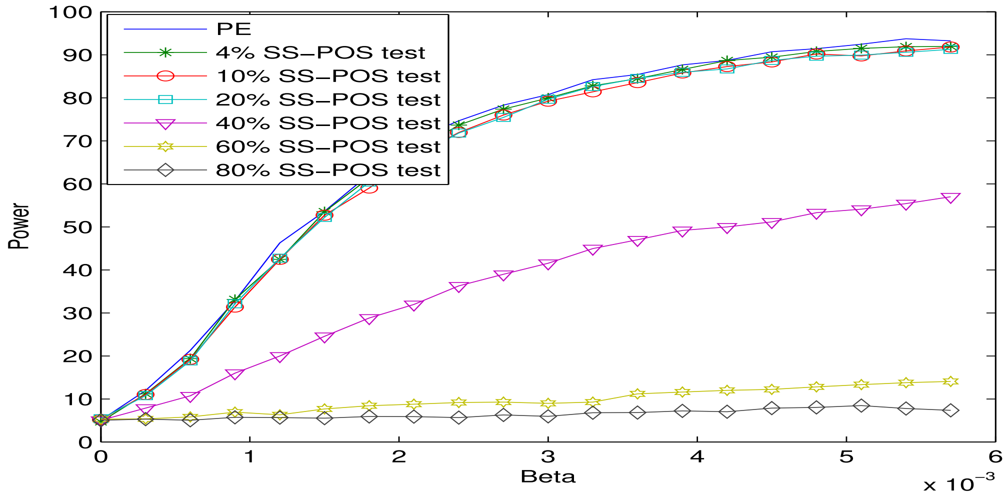
B. Normal distribution with Break in variance



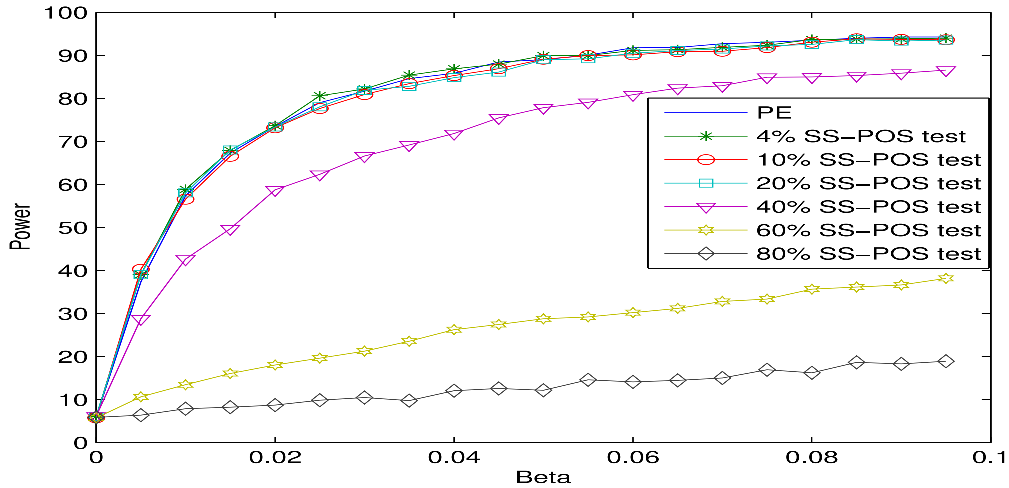
Note: These figures compare the power of POS test using different split-samples (SS-POS test); 4%, 10%, 20%, 40%, 60%, and 80%. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a mixture of normal and Cauchy distributions. Panel B corresponds to the case where ε_t follows a normal distribution with break in variance. PE corresponds to the power envelop.

Figure 6. **Power comparisons: different sample splits**
GARCH error distributions

A. Normal distribution with GARCH$(1,\ 1)$ plus jump variance



B. Normal distribution with non stationary GARCH$(1,\ 1)$ variance



Note: These figures compare the power of POS test using different split-samples (SS-POS test); 4%, 10%, 20%, 40%, 60%, and 80%. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a normal distribution with GARCH$(1,\ 1)$ plus jump variance. Panel B corresponds to the case where ε_t follows a normal distribution with non stationary GARCH$(1,\ 1)$ variance. PE corresponds to the power envelop.

5. POS confidence regions

In this section, we briefly describe how to build confidence regions with known significance level α, say $C_\beta(\alpha)$, for a vector of unknown parameters β using the proposed POS tests. Consider the regression model (3.10) and suppose we wish to test the null hypothesis (3.11) against the alternative hypothesis (3.12). The idea consists in finding all the values of $\beta_0 \in \mathbb{R}^k$ such that

$$SN_n^*(\beta_1)^{(0)} = \sum_{t=1}^{n} \left\{ \ln \left[\frac{1 - p(x_t, \beta_0,\, \beta_1 \mid X)}{p(x_t, \beta_0,\, \beta_1 \mid X)} \right] s\left(y_t - f(x_t, \beta_0)\right) \right\} < c_1(\beta_0,\, \beta_1)$$

where $SN_n^*(\beta_1)^{(0)}$ is the observed value of $SN_n^*(\beta_0|\beta_1)$ and the critical value $c_1(\beta_0,\, \beta_1)$ is given by the smallest constant $c_1(\beta_0,\, \beta_1)$ such that

$$\mathsf{P}[SN_n^*(\beta_0|\beta_1) > c_1(\beta_0,\, \beta_1) \mid \beta = \beta_0] \le \alpha.$$

The confidence region $C_\beta(\alpha)$ of the vector of parameters β can be defined as follows:

$$C_\beta(\alpha) = \left\{ \beta_0 : SN_n^*(\beta_1)^{(0)} < c_1(\beta_0,\, \beta_1) \mid \mathsf{P}\left[SN_n^*(\beta_0|\beta_1) > c_1(\beta_0,\, \beta_1) \mid \beta = \beta_0\right] \le \alpha \right\}.$$

Further, given the confidence region $C_\beta(\alpha)$, we can also derive confidence intervals for the components of vector β using the projection techniques. The latter can be used to find confidence sets, say $g(C_\beta(\alpha))$, for general transformations g of β in $\mathbb{R}^m$. Since, for any set $C_\beta(\alpha)$,

$$\beta \in C_\beta(\alpha) \Rightarrow g(\beta) \in g(C_\beta(\alpha)) \tag{5.1}$$

we have

$$\mathsf{P}[\beta \in C_\beta(\alpha)] \ge 1 - \alpha \Rightarrow \mathsf{P}[g(\beta) \in g(C_\beta(\alpha))] \ge 1 - \alpha, \tag{5.2}$$

where

$$g(C_\beta(\alpha)) = \{\delta \in \mathbb{R}^m : \exists\, \beta \in C_\beta(\alpha),\ g(\beta) = \delta\}.$$

From (5.1) and (5.2), the set $g(C_\beta(\alpha))$ is a conservative confidence set for $g(\beta)$ with level $1 - \alpha$. If $g(\beta)$ is a scalar, then we have:

$$\mathsf{P}\left[\inf\{g(\beta_0), \text{for } \beta_0 \in C_\beta(\alpha)\} \le g(\beta) \le \sup\{g(\beta_0),\ \text{for } \beta_0 \in C_\beta(\alpha)\}\right] > 1 - \alpha.$$

More details about the projection technique can be find in Dufour (1997), Abdelkhalek and Dufour (1998), Dufour and Kiviet (1998), Dufour and Jasiak (2001), and Dufour and Taamouti (2005).

6. Monte Carlo study

We present simulation results illustrating the performance of the statistical procedures defined in the previous sections. Since the number of tests and alternative models is so large, we have limited our results to two groups of data generating processes (DGPs) which correspond to different symmetric and asymmetric distributions and different forms of heteroskedasticity. Further, because for nonlinear regression models an iterative procedure is required for the estimation of β_1, which makes the convergence of our simulations slow, we restrict our simulations to the linear regression model where only an analytical formula is needed to estimate β_1 (OLS estimator).[1] However, other simulations results using an exponential regression model [$f(x_t, \beta) = \exp(\beta x_t)$], which show that the proposed tests perform quite well, can be found in Appendix **??**.

6.1. Simulated models

We assess the performance of the proposed POS test by comparing its size and power to those of some other tests, under various general DGPs. We choose our DGPs to illustrate performance in different contexts encountered in practice. We consider the following linear regression model

$$y_t = x_t \beta + \varepsilon_t, \ t = 1, \ \dots \ , n, \tag{6.1}$$

where β is an unknown parameter and the errors $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ are independent and follow different distributions (DGPs), so they are not necessarily identically distributed. The first group of DGPs that we examine represents different symmetric and asymmetric distributions of the error term ε_t:
1. normal distribution: $\varepsilon_t \sim N(0, 1)$;
2. Cauchy distribution: $\varepsilon_t \sim Cauchy$;
3. Student t distribution with two degrees of freedom: $\varepsilon_t \sim t(2)$;
4. Mixture of normal and Cauchy distributions: $\varepsilon_t \sim s_t \mid \varepsilon_t^C \mid -(1 - s_t) \mid \varepsilon_t^N \mid$, where ε_t^C follows Cauchy distribution, ε_t^N follows $N(0, 1)$ distribution, and

$$\mathsf{P}(s_t = 1) = \mathsf{P}(s_t = 0) = \frac{1}{2}.$$

The second group of DGPs represents different forms of heteroskedasticity:
5. break in variance:

$$\varepsilon_t \sim \begin{cases} N(0,1) & \text{for } t \neq 25 \\ \sqrt{1000}N(0,1) & \text{for } t = 25 \end{cases};$$

6. exponential variance: $\varepsilon_t \sim N(0, \sigma_\varepsilon^2(t))$ and $\sigma_\varepsilon(t) = \exp(0.5\, t)$;
7. GARCH$(1,\ 1)$ plus jump variance:

$$\sigma_\varepsilon^2(t) = 0.00037 + 0.0888\varepsilon_{t-1}^2 + 0.9024\sigma_\varepsilon^2(t-1)\,,$$

$$\varepsilon_t \sim \begin{cases} N(0, \sigma_\varepsilon^2(t)) & \text{for } t \neq 25 \\ 50N(0, \sigma_\varepsilon^2(t)) & \text{for } t = 25 \end{cases};$$

[1] We use GAUSS for the simulations. For nonlinear regression model, it takes around 5 days and 7 hours to calculate the empirical size and power, whereas for linear model it takes 2 days and 3 hours. Some characteristics of the computer hardware employed are:
(1) Memory (RAM): 3.00 GB;
(2) AMD Athlon(tm) 64X2 Dual Core Processor 4200+ 2.21 GHz.

8. nonstationary GARCH$(1,\,1)$ variance: $\varepsilon_t \sim N(0, \sigma^2_\varepsilon(t))$ and

$$\sigma^2_\varepsilon(t) = 0.75\varepsilon^2_{t-1} + 0.75\sigma^2_\varepsilon(t-1)\,.$$

We use POS test and other tests, which are supposed to be robust against heteroskedasticity and non-normality, to test the null hypothesis $H_0 : \beta = 0$. We run Monte Carlo simulations to compare the size and power of 10% split-sample POS tests (hereafter 10% SS-POS test) to those of T-test, T-test based on White's (1980) variance correction (hereafter WT-test), and sign-based test proposed by Campbell and Dufour (1995) (hereafter CD95 test). In what follows, the notations CT-test and CWT-test refer to the T-test and WT-test after size correction, respectively. For some DGPs, T-test and WT-test may not control size and we adjust the power functions such that CT-test and CWT-test control their size. In our simulations the explanatory variable x_t is generated from a mixture of normal and χ^2 distributions. We perform $M_1 = 10000$ simulations to evaluate the probability distribution of POS test statistic and $M_2 = 5000$ simulations to estimate the power functions of POS test and other tests. All simulated samples are of size $n = 50$. The sign-based test statistic of Campbell and Dufour (1995) has a discrete distribution and it is not possible (without randomization) to obtain test whose size is precisely 5%. In our simulations study, the size of this test is 5.95% for $n = 50$.

6.2. Simulation results

Monte Carlo simulation results are presented in Tables 6.1-6.1 and Figures 6.1-6.1. These results correspond to different DGPs described in Section 6.1. Tables 6.1-6.1 show the power envelope of POS test, the size and power of POS test under different alternative hypotheses and using different split-sample sizes, and size and power of T-test (CT-test), WT-test (CWT-test), and CD95 test. Figures 6.1-6.1 compare the power of 10% SS-POS test, T-test (CT-test), WT-test (CWT-test), and CD95 test to the power envelope. The results are detailed below.

First, Panel A of Table 6.1 and Panel A of Figure 6.1 correspond to the case where the error term ε_t in the model (6.1) is normally distributed. Panel A of Table 6.1 shows that the power of POS test depends on the alternative hypothesis β_1. When the latter is far from the null hypothesis, the POS test power's curve moves away from the power envelope [see also Panel A of Figure 4.1]. However, using approximately 10% of sample to estimate β_1 yields a power which is typically very close to the power envelope. Thus, split-sample approach represents a good way to select the appropriate alternative hypothesis at which the power of POS test is maximized.

The T-test based on White's (1980) variance correction, say WT-test, does not control size and its power after size correction is presented in the last column of Panel A of Table 6.1. Panel A of Figure 6.1 shows that T-test is more powerful than 10% SS-POS test, CWT-test, and CD95 test. We expect to get the latter result, since under normality T-test is the most powerful test. However, the power of 10% SS-POS test is very close to the power envelope and does better than CD95 test.

Second, Panel B of Table 6.1 and Panel B of Figure 6.1 and Panel A of Figure 6.1 correspond to the cases where the error term ε_t follows Cauchy distribution and Student's distribution with two degrees of freedom, respectively. We see again that the power of POS test depends on the alternative hypothesis β_1. Particularly, when the alternative hypothesis is far from the null hypothesis, the

Table 1. **Power comparisons: different tests**
Normal and Cauchy error distributions

A. Normal distribution

		POS test		SS-POS test				Other tests			
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD95 test	T-test	WT-test	CWT-test
0.0000	5.20	5.14	5.34	4.82	4.88	5.36	4.78	5.94	4.88	7.52	4.94
0.0005	7.44	5.96	6.50	7.58	7.44	6.62	6.78	6.96	7.42	10.70	7.30
0.0010	9.20	8.24	7.96	9.98	9.82	9.48	8.20	8.24	11.40	15.40	11.50
0.0015	12.78	11.28	10.24	12.60	12.90	12.76	11.04	10.06	16.24	20.08	16.50
0.0020	16.34	13.34	11.96	16.28	16.18	17.26	13.18	11.02	21.70	26.78	20.68
0.0025	21.38	16.36	14.02	20.56	21.80	21.70	15.76	14.12	29.42	34.42	27.74
0.0030	27.74	20.74	17.62	26.08	25.84	27.26	18.74	17.02	39.32	41.20	34.24
0.0035	33.26	23.48	20.86	32.44	32.08	31.42	23.28	19.22	45.22	49.16	43.48
0.0040	38.14	28.28	23.46	36.40	39.08	37.52	24.88	21.56	55.36	58.52	52.38
0.0045	44.68	32.68	27.68	43.28	44.10	44.30	28.14	23.46	62.38	66.96	57.44
0.0050	52.20	36.68	29.70	49.44	51.74	50.60	35.24	27.50	71.04	73.16	67.32
0.0055	57.76	40.78	33.50	55.42	56.68	56.06	38.64	29.80	79.16	79.92	74.70
0.0060	63.92	45.44	37.26	60.78	63.12	62.62	42.44	32.30	84.18	85.70	80.84
0.0065	69.22	47.66	40.68	66.44	68.00	68.90	46.74	34.78	89.58	89.74	85.06

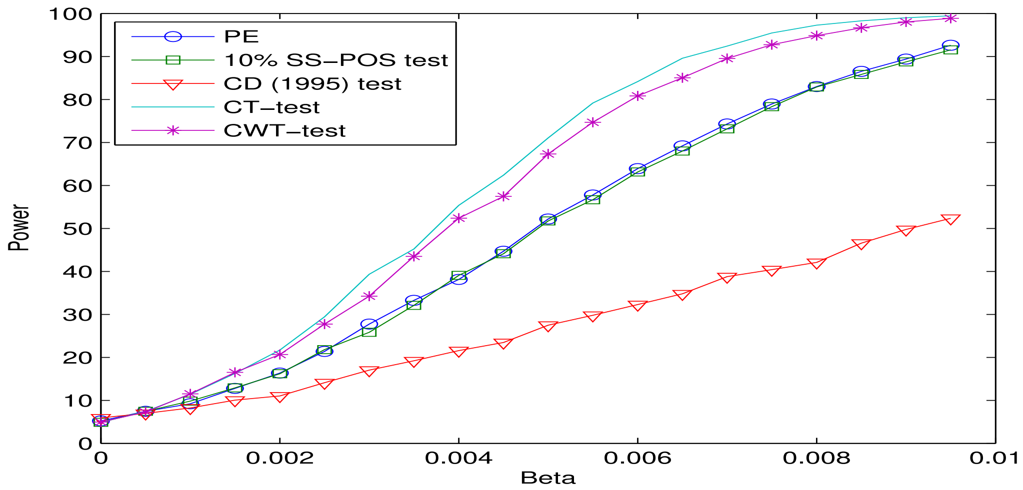
B. Cauchy distribution

		POS test		SS-POS test				Other tests		
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD95 test	T-test	WT-test
0.000	5.10	4.88	4.80	5.02	5.30	5.48	4.46	5.78	5.68	3.94
0.005	34.22	25.18	20.94	26.72	33.30	30.86	23.48	18.44	9.50	15.00
0.010	66.38	48.42	39.58	50.46	61.74	62.28	47.86	35.16	16.60	28.92
0.015	84.44	62.56	52.94	64.74	76.24	77.02	64.38	48.90	25.76	43.82
0.020	92.20	74.30	63.08	74.36	84.90	85.14	73.70	60.36	36.28	54.72
0.025	96.44	79.62	69.60	79.06	89.88	88.82	81.78	69.58	42.74	62.08
0.030	98.12	82.86	74.30	81.08	92.92	92.58	84.70	76.60	50.14	67.06
0.035	99.00	86.02	78.36	82.86	93.70	93.10	88.38	81.88	56.00	70.72
0.040	99.36	89.16	79.60	85.62	94.70	94.30	90.76	86.42	60.56	73.34
0.045	99.68	89.92	81.88	85.74	94.92	95.74	92.24	88.84	63.30	77.18
0.050	99.80	91.12	84.24	86.76	95.92	95.92	93.00	91.18	66.60	78.70
0.055	99.98	91.94	86.20	87.14	96.42	96.48	94.56	92.98	69.88	81.30
0.060	99.94	92.50	86.38	87.08	97.02	96.18	95.96	94.16	72.72	82.96
0.065	99.94	93.08	86.84	88.02	96.86	96.90	96.92	94.68	74.10	83.22

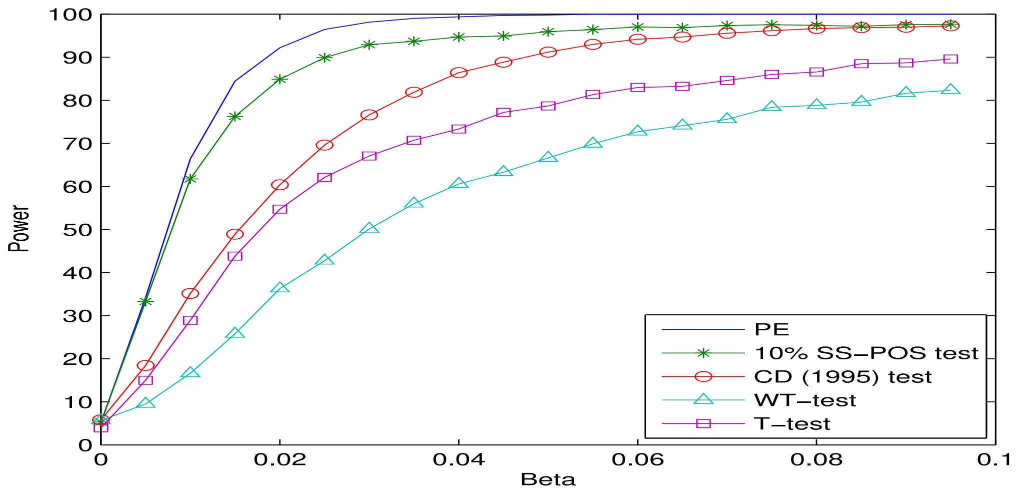
Note: These tables show the power envelope of POS test (PE) and the power of: **(1)** POS test under different alternative hypotheses (POS test); **(2)** POS test using different split-sample sizes (SS-POS test); **(3)** sign-based test of Campbell and Dufour (1995) [CD95 test]; **(4)** T-test; **(5)** T-test based on White's (1980) variance correction (WT-test); and **(6)** WT-test after size correction (CWT-test). Panel A corresponds to the case where the error term ε_t in the model (6.1) is homoskedastic and normally distributed and Panel B corresponds to the case where ε_t is homoskedastic and follows a Cauchy distribution.

Figure 7. **Power comparisons: different tests**
Normal and Cauchy error distributions

A. Normal distribution



B. Cauchy distribution



Note: These figures compare the power envelope (PE) to: **(1)** the power curves of 10% split-sample POS test [10% SS-POS test]; **(2)** T-test (or CT-test); **(3)** sign-based test proposed by Campbell and Dufour (1995) [CD95 test]; and **(4)** the T-test based on White's (1980) variance correction [WT-test or CWT-test]. Panel A corresponds to the case where the error term ε_t in the model (6.1) is homoskedastic and normally distributed and Panel B corresponds to the case where ε_t is homoskedastic and follows Cauchy distribution.

Table 2. **Power comparisons: different tests**
Mixture and normal distribution with break

A. Mixture distribution

		POS test		SS-POS test				Other tests				
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD95 test	T-test	WT-test	CT-test	CWT-test
0.000	4.96	5.30	4.90	4.58	4.70	5.02	5.18	5.98	9.92	10.74	5.08	5.04
0.001	9.96	8.08	8.14	8.86	9.98	9.16	8.02	8.94	11.28	13.12	5.90	7.92
0.002	15.70	11.52	11.30	14.46	15.90	14.60	12.24	11.76	13.98	18.88	7.50	12.94
0.003	25.26	18.48	14.24	22.00	24.76	24.60	19.64	15.72	16.90	25.76	10.10	18.74
0.004	35.46	23.84	18.12	29.60	34.08	34.28	27.36	21.00	20.68	31.76	11.82	25.68
0.005	46.08	28.70	23.66	39.16	44.14	42.96	34.60	26.24	24.32	40.04	14.64	31.74
0.006	56.68	35.52	27.56	47.44	51.78	52.06	41.22	29.72	28.24	47.06	18.16	37.82
0.007	67.64	40.66	32.30	55.34	61.90	61.84	51.16	34.06	33.00	51.22	21.92	44.76
0.008	75.00	45.32	37.46	60.44	69.48	69.50	60.10	38.96	36.62	56.70	24.56	49.14
0.009	82.06	50.40	39.64	67.28	76.52	75.32	66.68	44.22	40.16	60.50	30.18	54.60
0.010	88.48	54.90	43.24	70.70	80.84	79.90	73.68	49.58	45.86	63.74	33.64	58.80
0.011	90.68	58.48	45.24	73.92	84.16	84.94	79.92	52.40	48.60	66.90	38.06	61.70
0.012	94.38	62.44	50.78	77.44	87.66	87.42	85.18	58.54	51.16	69.26	39.72	65.62
0.013	95.70	65.76	53.12	78.82	90.54	89.22	88.64	60.10	55.26	72.16	43.66	67.42

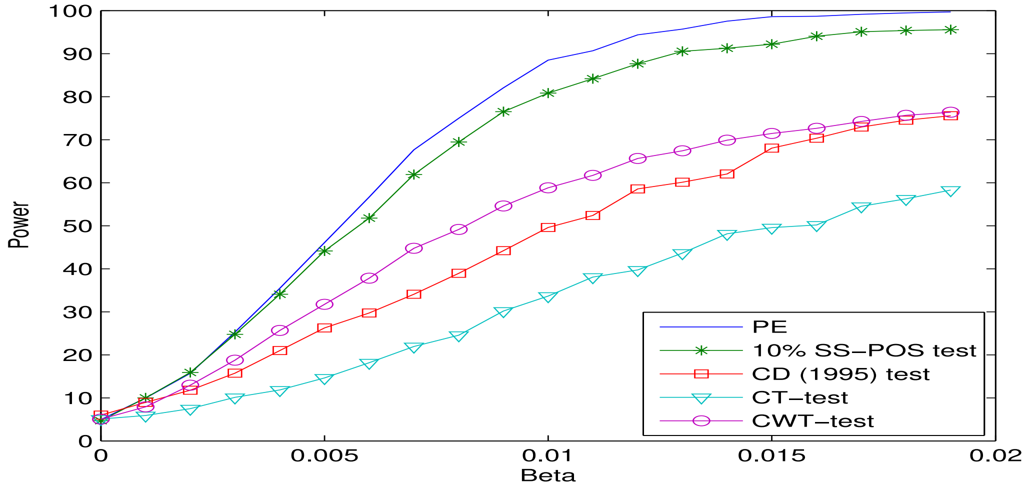
B. Normal distribution with break in variance

		POS test		SS-POS test				Other tests		
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD95 test	T-test	WT-test
0.0000	5.40	4.98	4.92	4.84	5.24	5.10	4.96	5.78	0.01	0.16
0.0008	9.22	7.96	7.90	8.28	9.32	8.38	7.68	8.24	0.04	0.42
0.0016	14.78	12.00	10.18	13.12	13.76	12.98	10.42	10.44	0.06	0.60
0.0024	20.16	15.88	14.62	18.20	20.12	19.86	15.58	12.98	0.12	1.08
0.0032	29.32	22.12	19.60	25.24	28.34	28.26	19.64	17.34	0.30	1.62
0.0040	39.04	27.96	25.38	35.72	38.32	38.68	25.24	21.40	0.22	1.86
0.0048	49.78	35.70	29.12	43.98	47.00	48.06	32.38	26.12	0.46	2.30
0.0056	59.66	41.62	34.12	52.82	59.16	58.24	39.78	30.42	0.84	3.60
0.0064	68.88	48.50	39.14	62.30	67.90	67.28	45.96	34.78	0.78	4.58
0.0072	77.32	55.90	45.30	68.78	75.66	76.50	53.54	38.38	0.94	4.88
0.0080	83.96	61.90	51.68	76.14	83.14	82.20	60.92	42.72	0.94	5.88
0.0088	88.76	65.90	55.52	80.14	88.00	88.50	67.46	47.04	1.22	6.54
0.0096	92.22	72.94	60.32	85.60	91.70	93.02	73.06	51.76	1.50	8.14
0.0104	95.42	78.52	64.48	87.42	94.68	95.34	79.76	55.02	1.42	7.88

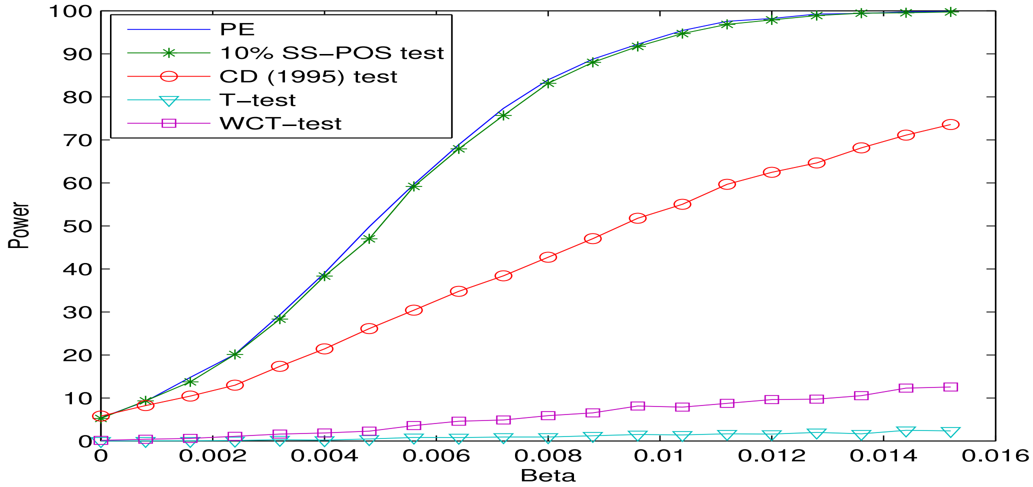
Note: These tables show the power envelope of POS test (PE) and the power of: **(1)** POS test under different alternative hypotheses (POS test); **(2)** POS test using different split-sample sizes (SS-POS test); **(3)** sign-based test of Campbell and Dufour (1995) [CD95 test]; **(4)** T-test; **(5)** T-test based on White's (1980) variance correction (WT-test); **(6)** T-test after size correction (CT-test); and **(7)** WT-test after size correction (CWT-test). Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a mixture of normal and Cauchy distributions and Panel B corresponds to the case where ε_t follows a normal distribution with Break in variance.

Figure 8. **Power comparisons: different tests**
Mixture and normal error distribution with break

A. Mixture distribution



B. Normal distribution with Break in variance



Note: These figures compare the power envelope (PE) to: **(1)** the power curves of 10% split-sample POS test [10% SS-POS test]; **(2)** T-test (or CT-test); **(3)** sign-based test proposed by Campbell and Dufour (1995) [CD95 test]; and **(4)** the T-test based on White's (1980) variance correction [WT-test or CWT-test]. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a mixture of normal and Cauchy distributions and Panel B corresponds to the case where ε_t follows a normal distribution with break in variance.

Table 3. **Power comparisons: different tests**
GARCH error distributions

A. Normal distribution with GARCH$(1,\ 1)$ plus jump variance

		POS test		SS-POS test				Other tests		
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD95 test	T-test	WT-test
0.0000	5.07	5.74	4.98	4.70	5.24	5.40	5.04	6.42	1.22	4.96
0.0003	11.98	9.06	9.16	11.18	11.02	10.76	7.86	8.06	2.36	8.92
0.0006	21.28	15.50	12.90	19.38	19.20	18.84	10.74	12.18	5.00	14.60
0.0009	32.80	21.00	18.14	33.12	31.34	32.12	15.98	17.24	8.90	21.20
0.0012	46.28	28.14	23.90	42.46	42.46	42.72	19.98	21.90	13.36	27.16
0.0015	53.62	34.62	28.20	53.52	52.70	52.20	24.56	25.86	16.76	30.86
0.0018	62.24	39.10	33.74	61.36	59.00	60.40	28.80	30.12	19.06	36.58
0.0021	70.22	46.06	38.10	67.52	66.44	66.14	31.96	34.44	24.20	42.58
0.0024	74.66	48.74	40.72	73.66	71.94	71.80	36.28	37.68	27.26	45.10
0.0027	78.28	50.88	43.94	77.36	75.98	75.44	38.98	40.12	29.22	48.82
0.0030	80.72	54.04	47.76	79.96	79.22	79.66	41.54	44.32	32.40	51.02
0.0033	84.22	56.12	51.80	82.76	81.38	82.62	44.96	46.72	36.10	55.08
0.0036	85.42	58.82	53.44	84.46	83.52	84.50	47.00	47.84	38.32	56.42
0.0039	87.66	60.52	54.78	86.58	85.76	85.94	49.18	51.04	41.22	60.18

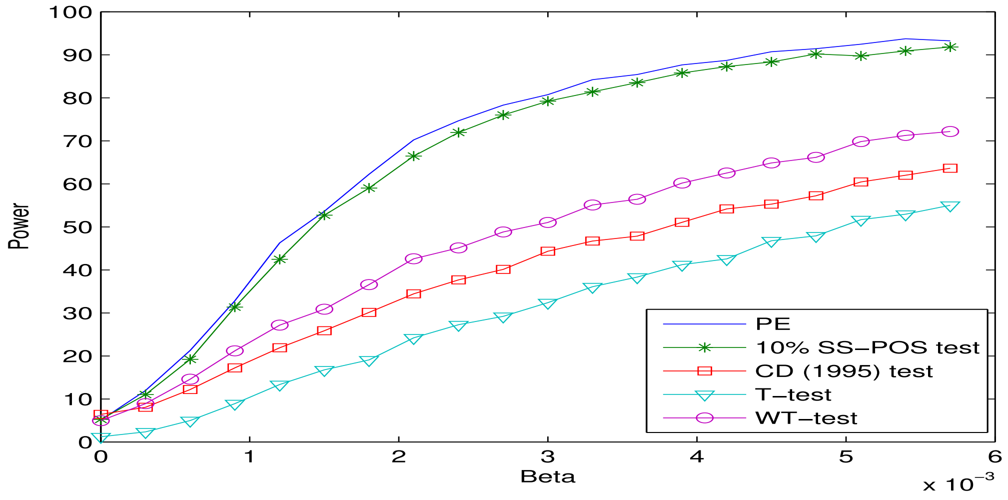
B. Normal distribution with non stationary GARCH$(1,\ 1)$ variance

		POS test		SS-POS test				Other tests		
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD95 test	T-test	WT-test
0.000	5.95	5.58	6.08	6.02	5.76	6.04	6.16	6.26	0.94	5.00
0.005	37.34	29.68	27.72	39.04	40.28	39.00	28.78	23.58	14.26	34.18
0.010	57.36	44.54	41.36	58.86	56.58	58.04	42.64	39.78	27.00	51.22
0.015	67.30	56.54	53.58	67.92	66.54	68.00	49.70	49.84	35.00	60.44
0.020	73.46	63.76	60.56	73.64	73.16	73.36	58.74	58.04	42.04	67.28
0.025	79.02	67.86	64.70	80.60	77.64	78.04	62.34	65.88	47.16	72.36
0.030	81.66	72.50	69.38	82.18	80.88	81.88	66.60	69.72	50.90	75.14
0.035	84.58	74.72	72.56	85.40	83.42	82.80	69.18	74.78	54.22	78.24
0.040	85.82	77.86	75.08	86.86	85.30	84.82	71.84	77.82	57.52	80.04
0.045	88.46	80.52	77.20	87.98	86.90	86.12	75.46	80.44	61.18	82.96
0.050	89.02	81.48	79.22	89.92	89.10	88.98	77.84	83.04	62.48	84.34
0.055	90.04	83.20	81.00	89.94	89.94	89.22	79.08	83.82	64.16	84.88
0.060	91.76	84.52	81.96	91.14	90.10	90.50	80.86	85.70	67.20	87.26
0.065	91.82	85.22	83.22	91.30	90.86	91.12	82.38	87.00	68.80	88.22

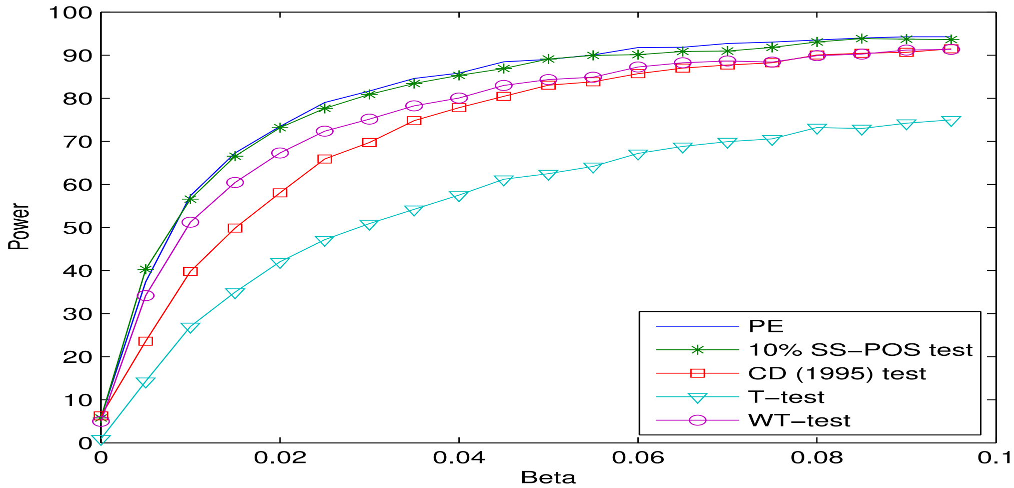
Note: These tables show the power envelope of POS test (PE) and the power of: **(1)** POS test under different alternative hypotheses (POS test); **(2)** POS test using different split-sample sizes (SS-POS test); **(3)** sign-based test of Campbell and Dufour (1995) [CD95 test]; **(4)** T-test; and **(5)** T-test based on White's (1980) variance correction (WT-test). Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a normal distribution with GARCH$(1,\ 1)$ plus jump variance and Panel B corresponds to the case where ε_t follows a normal distribution with non stationary GARCH$(1,\ 1)$ variance.

Figure 9. **Power comparisons: different tests**
GARCH error distributions

A. Normal distribution with GARCH$(1,\ 1)$ plus jump variance



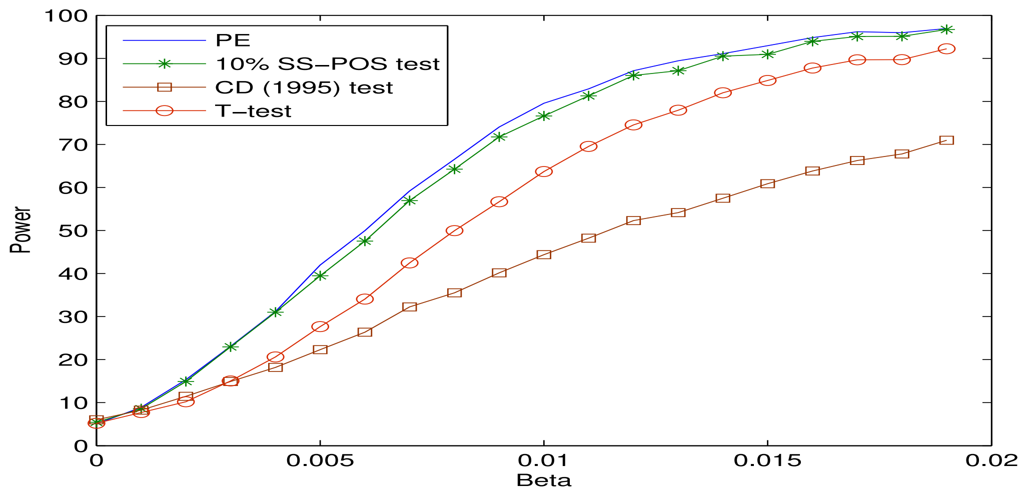
B. Normal distribution with non stationary GARCH$(1,\ 1)$ variance



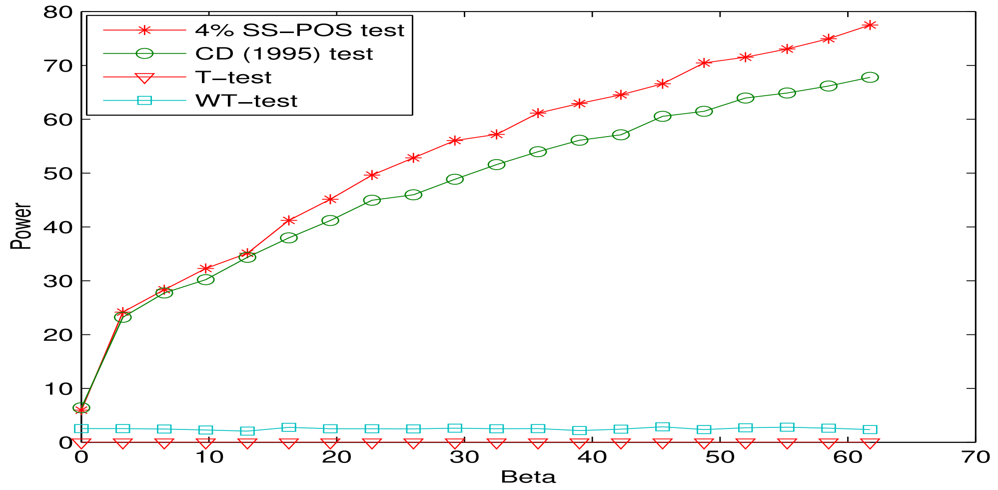
Note: These figures compare the power envelope (PE) to: **(1)** the power curves of 10% split-sample POS test [10% SS-POS test]; **(2)** T-test (or CT-test); **(3)** sign-based test proposed by Campbell and Dufour (1995) [CD95 test], and **(4)** the T-test based on White's (1980) variance correction [WT-test or CWT-test]. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows normal distribution with GARCH$(1,\ 1)$ plus jump variance and Panel B corresponds to the case where ε_t follows normal distribution with non stationary GARCH$(1,\ 1)$ variance.

Figure 10. **Power comparisons: different tests**
Student and normal error distribution with exponential variance

A. Student distribution



B. Normal distribution with exponential variance



Note: These figures compare the power envelope (PE) to: **(1)** the power curves of 10% split-sample POS test [10% SS-POS test]; **(2)** T-test (or CT-test); **(3)** sign-based test proposed by Campbell and Dufour (1995) [CD95 test]; and **(4)** the T-test based on White's (1980) variance correction [WT-test or CWT-test]. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a student distribution with degree of freedom 2 and Panel B corresponds to the case where ε_t follows a normal distribution with exponential variance.

power curve of POS test moves away from the power envelope [see Panel B of Table 6.1]. We also see that 10% represents the appropriate proportion of sample that we need to use for the estimation of β_1. Further, Panel B of Figure 6.1 and Panel A of Figure 6.1 shows that 10% SS-POS test is more powerful than T-test, WT-test, and CD95 test, and is close to the power envelope.

Third, Panel A of Table 6.1 and Figure 6.1, Panels A and B of Table 6.1, and Panels A and B of Figure 6.1 correspond to the cases where the error term ε_t follows a mixture of normal and Cauchy distributions, normal distribution with GARCH$(1,\ 1)$ plus jump variance, and normal distribution with non stationary GARCH$(1,\ 1)$ variance, respectively. The results, in terms of the impact of β_1 on the power function of POS test and the appropriate proportion of sample to use in estimating β_1, are similar to those of previous cases. Further, Panel A of Figure 6.1 and Panels A and B of Figure 6.1 show that 10% SS-POS test is again more powerful than T-test, WT-test, CD95 test, and is very close to the power envelope. When ε_t follows the mixture distribution, WT-test and T-test do not control size and we adjust their power functions such that CWT-test and CT-test control size. Interestingly, even if GARCH$(1,\ 1)$ and non stationary GARCH$(1,\ 1)$ models do not satisfy they key assumption (2.1), POS test still controls size and has very good power.

Finally, Panel B of Table 6.1 and Figure 6.1 and Panel B of Figure 6.1 correspond the cases where ε_t follows normal distribution with a break in variance and an exponential variance, respectively. In these cases, the powers of T-test and WT-test are very weak and flat, whereas the 10% SS-POS test does well and is more powerful than sign-based test proposed by Campbell and Dufour (1995).

From the previous results we draw the following conclusions. First, it is clear that the alternative hypothesis has an impact on the power function of POS test. Second, the adaptive approach based on split-sample technique allows to choose an optimal value of the alternative hypothesis at which the power of POS test is maximized. We should use a small part, approximately 10%, of sample to estimate the alternative hypothesis and the rest, 90%, to compute the test statistic of POS test. Third, when the error term ε_t follows normal and heteroskedastic distributions, the power of 10% SS-POS test is close to the power envelope. For non-normal errors this is not the case and the power of 10% SS-POS test is somewhat far from the power envelope. Finally, except for a normally and homoskedastic distributed error, 10% SS-POS test performs better than T-test (CT-test), WT-test (CWT-test), and CD95 test.

We also use simulations to compare the power of 10% SS-POS test calculated using the true weights with the power of 10% SS-POS test computed using normal weights. The weights $a_t(\beta_1)$ are computed using homoskedastic and normal distribution. The results are presented in Table 6.2. We see that using the true weights may improve the power of 10% SS-POS test. However, the power loss when we substitute the true weights by normal weights is very small.

7. Conclusion

We propose exact POS-based tests to test the parameters in the context of linear and nonlinear regression models with fixed regressors. These tests are distribution-free, robust against heteroskedasticity of an unknown form, and they may be inverted to obtain confidence sets for the vector of unknown parameters.

Table 4. **True weights versus normal weights**

A. True weights using Cauchy distribution

		SS-POS test using true weights		SS-POS test using normal weights	
β	PE	10%	20%	10%	20%
0.000	5.10	5.16	5.16	5.30	5.48
0.005	34.22	33.58	31.18	33.30	30.86
0.010	66.38	61.94	62.47	61.74	62.28
0.015	84.44	80.32	80.32	76.24	77.02
0.020	92.20	89.76	89.76	84.90	85.14
0.025	96.44	95.22	95.22	89.88	88.82
0.030	98.12	96.98	96.98	92.92	92.58
0.035	99.00	98.26	98.26	93.70	93.10
0.040	99.36	99.14	99.14	94.70	94.30
0.045	99.68	99.30	99.30	94.92	95.74
0.050	99.80	99.44	99.44	95.92	95.92
0.055	99.98	99.70	99.70	96.42	96.48
0.060	99.94	99.82	99.82	97.02	96.18
0.065	99.94	99.90	99.90	96.86	96.90

B. True weights using mixture distribution

		SS-POS test with true weights		SS-POS test with normal weights	
β	PE	10%	20%	10%	20%
0.000	4.96	4.74	5.26	4.70	5.02
0.001	9.96	8.96	9.08	9.98	9.16
0.002	15.70	14.34	16.70	15.90	14.60
0.003	25.26	24.84	24.67	24.76	24.60
0.004	35.46	34.52	34.46	34.08	34.28
0.005	46.08	44.26	44.06	44.14	42.96
0.006	56.68	53.24	54.96	51.78	52.06
0.007	67.64	62.92	62.88	61.90	61.84
0.008	75.00	71.66	70.14	69.48	69.50
0.009	82.06	79.24	79.54	76.52	75.32
0.010	88.48	85.52	84.34	80.84	79.90
0.011	90.68	88.80	89.22	84.16	84.94
0.012	94.38	92.06	91.50	87.66	87.42
0.013	95.70	94.32	94.62	90.54	89.22

Note: These tables summarize the results of the comparison between the power of 10% split-sample POS test (SS-POS test) calculated using the true weights $a_t(\beta_1)$ with the power of 10% split-sample POS test calculated using normal weights. In Panel A the true weights correspond to the case where the error term ε_t in the model (6.1) follows a Cauchy distribution and in Panel B the true weights correspond to the case where ε_t follows a mixture of normal and Cauchy distributions. SS-POS test corresponds to split-sample POS test. PE corresponds to the power envelop.

Since the proposed POS test maximizes the power at a given value of the alternative, we suggest an approach based on split-sample technique to choose an optimal alternative such that the power of POS test is close to the power envelope. The simulation results show that using approximately 10% of sample to estimate the alternative hypothesis and the rest (90%) to compute the test statistic of POS test, yields a power which is typically very close to the power envelope.

To assess the performance of POS test we run a Monte Carlo simulation study and compare its size and power to those of some other tests, under various general DGPs. We consider different DGPs to illustrate different contexts that one can encounter in practice. We use two groups of DGPs which correspond to different symmetric and asymmetric distributions and different heteroskedasticity forms. The results show that 10% split-sample POS test is more powerful than T-test, Campbell and Dufour's (1995) sign-based test, T-test with White's (1980) variance correction, and it is close to the power envelope.

The present paper could be generalized to the case where the explanatory variables are stochastic by relaxing the assumption (2.1). This issue is the topic of on-going research.

A. Appendix: Proofs

PROOF OF THEOREM **4.1**. Conditionally on X the characteristic function of $SN_n^*(\beta_0|\beta_1)$ is given by:

$$\phi_{SN_n^*}(u) = E_X\left[\exp(iu\, SN_n^*(\beta_0|\beta_1))\right] = E_X\left[\prod_{t=1}^{n}\exp\left(i\,u\ln\left[\frac{1-p(x_t,\beta_0,\,\beta_1\mid X)}{p(x_t,\beta_0,\,\beta_1\mid X)}\right]s\left(\tilde{y}_t\right)\right)\right],$$

where $p(x_t,\beta_0,\,\beta_1\mid X)=\mathsf{P}[\varepsilon_t\le f(x_t,\beta_0)-f(x_t,\beta_1)\mid X]$, $u\in\mathbb{R}$, $\tilde{y}_t=y_t-f(x_t,\beta_0)$ and the complex number $i=\sqrt{-1}$. Since conditional on X the random variables $\tilde{y}_t$, for $t=1,\,\dots\,,\,n$, are independent

$$\begin{aligned}
\phi_{SN_n^*}(u) &= \prod_{t=1}^{n}E_X\left[\exp\left(iu\;\ln\left[\frac{1-p(x_t,\beta_0,\,\beta_1\mid X)}{p(x_t,\beta_0,\,\beta_1\mid X)}\right]s\left(\tilde{y}_t\right)\right)\right]\\
&= \prod_{t=1}^{n}\left\{\sum_{j=0}^{1}\mathsf{P}\left(s\left(\tilde{y}_t\right)=j\mid X\right)\exp\left(iu\;\ln\left[\frac{1-p(x_t,\beta_0,\,\beta_1\mid X)}{p(x_t,\beta_0,\,\beta_1\mid X)}\right]j\right)\right\}\\
&= \prod_{t=1}^{n}\left[1+\left(\exp\left(iu\;\ln\left[\frac{1-p(x_t,\beta_0,\,\beta_1\mid X)}{p(x_t,\beta_0,\,\beta_1\mid X)}\right]\right)-1\right)(1-\mathsf{P}[\varepsilon_t\le f(x_t,\beta_0)-f(x_t,\beta)\mid X])\right]\\
&= \prod_{t=1}^{n}\left[1+\left(\exp\left(iu\;\ln\left[\frac{1-p(x_t,\beta_0,\,\beta_1\mid X)}{p(x_t,\beta_0,\,\beta_1\mid X)}\right]\right)-1\right)\left(1-p(x_t,\beta_0,\,\beta\mid X)\right)\right] \qquad \text{(A.1)}
\end{aligned}$$

Given the conditional characteristic function (A.1), a standard Fourier-inversion formula [see Gil-Pelaez (1951)] implies that the conditional distribution function of $SN_n^*(\beta_0|\beta_1)$ evaluated at

$c_1(\beta_0,\, \beta_1)$, for $c_1(\beta_0,\, \beta_1) \in \mathbb{R}$, is given by:

$$\mathsf{P}(SN_n^*(\beta_0|\beta_1) \le c_1(\beta_0,\, \beta_1) \mid X) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left\{\exp\left(-iuc_1(\beta_0,\, \beta_1)\right)\phi_{SN_n^*}(u)\right\}}{u}du, \tag{A.2}$$

where, $\forall\, u \in \mathbb{R}$,

$$\phi_{SN_n^*}(u) = \prod_{t=1}^{n}\left[1 + \left(\exp\left(iu\, \ln\left[\frac{1 - p(x_t, \beta_0,\, \beta_1 \mid X)}{p(x_t, \beta_0,\, \beta_1 \mid X)}\right]\right) - 1\right)(1 - p(x_t, \beta_0,\, \beta \mid X))\right],$$

and $\mathrm{Im}\{z\}$ denotes the imaginary part of a complex number z. Thus, the power function of POS test is given by the following probability function:

$$\mathit{\Pi}(\beta, \beta_1) = \mathsf{P}\left[SN_n^*(\beta_0|\beta_1) > c_1(\beta_0,\, \beta_1)\right] = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left\{\exp\left(-iuc_1(\beta_0,\, \beta_1)\right)\phi_{SN_n^*}(u)\right\}}{u}du.$$

□

B. Appendix: Additional simulation results using nonlinear regression model

In this appendix, we consider a nonlinear DGP to assess the performance (size and power) of the proposed POS test:

$$y_t = \exp(\beta x_t) + \varepsilon_t, \tag{B.3}$$

where we assume three different distributions for the error term ε_t :
1. Normal distribution: $\varepsilon_t \sim N(0, 1)$;
2. Mixture of normal and Cauchy distributions: $\varepsilon_t \sim s_t \mid \varepsilon_t^C \mid -(1 - s_t) \mid \varepsilon_t^N \mid$, where ε_t^C follows Cauchy distribution, ε_t^N follows $N(0, 1)$ distribution, and $\mathsf{P}\,(s_t = 1) = \mathsf{P}\,(s_t = 0) = \frac{1}{2}$;
3. GARCH$(1,\, 1)$ plus jump variance:

$$\varepsilon_t \sim \begin{cases} N(0, \sigma_\varepsilon^2(t)) & \text{for } t \neq 25 \\ 50\, N(0, \sigma_\varepsilon^2(t)) & \text{for } t = 25 \end{cases}$$

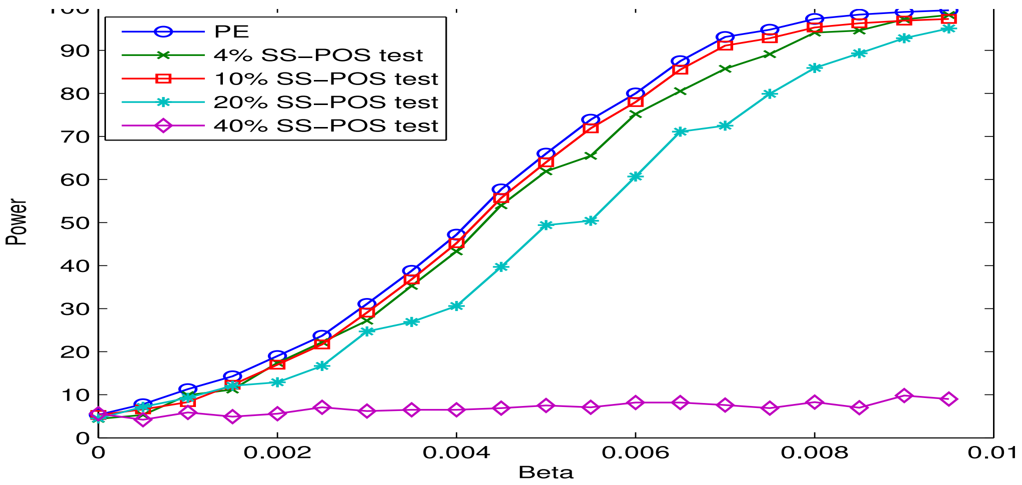
and

$$\sigma_\varepsilon^2(t) = 0.00037 + 0.0888\varepsilon_{t-1}^2 + 0.9024\sigma_\varepsilon^2(t-1).$$

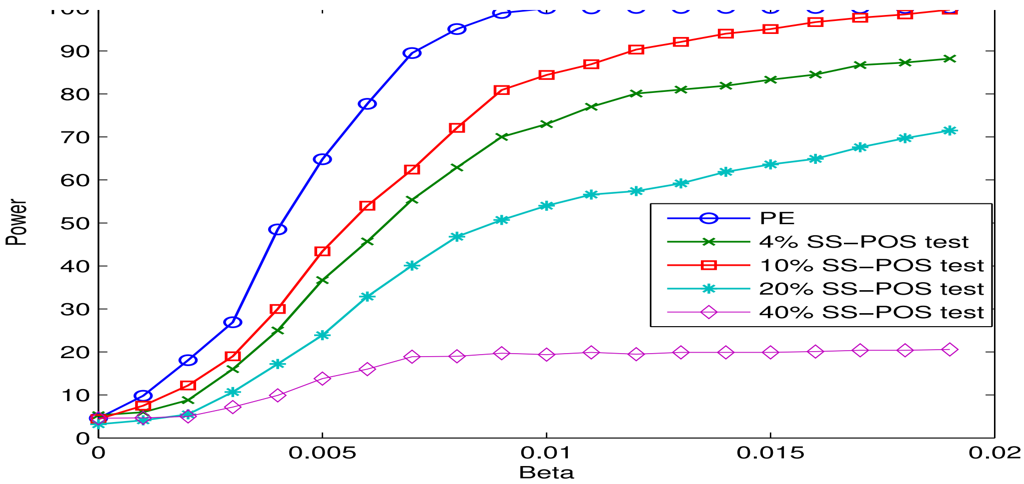
The results are presented in Figure 6.1. The latter show that the proposed tests perform quite well even in the context of a nonlinear model. We also see that 10% represents the appropriate proportion of sample that one needs to use for the estimation of the alternative hypothesis β_1.

Figure 11. **Power comparisons: different sample splits**
Normal, Mixture and GARCH with jump error distributions

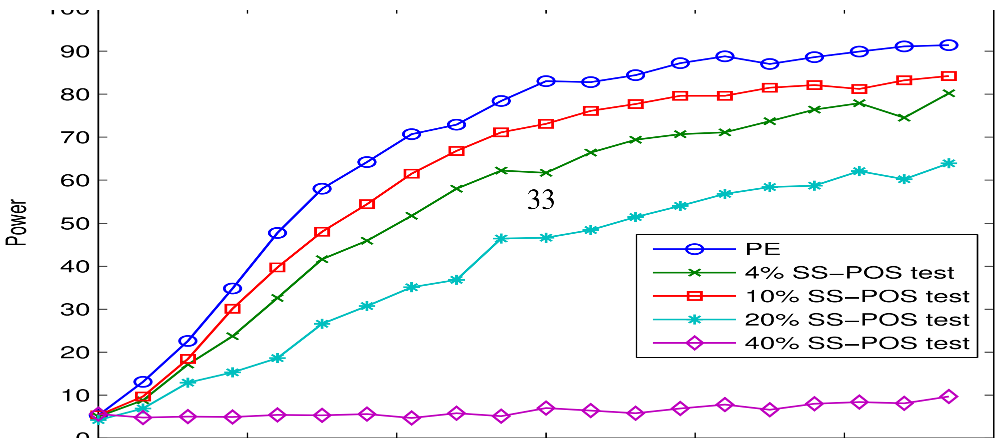
A. Normal distribution



B. Mixture distribution



C. Normal distribution with GARCH$(1,\ 1)$ plus jump variance



C. Acknowledgments

The authors thank Lynda Khalaf, Frédéric Jouneau-Sion, Benoit Perron, Thi Thuy Anh Vo, Carlos Velasco, an anonymous referee, an Associate Editor, and the Co-Editor Erricos John Kontoghiorghes for several useful comments. Earlier versions of this paper were presented at the 46th Annual Meeting of the Société canadienne de science économique at Montreal (2006), the 61th European Meeting of the Econometric Society at Vienna (2006), and the 2nd International Workshop on Computational and Financial Econometrics (CFE'08) at Neuchâtel (Switzerland) (2008). This work was supported by the William Dow Chair in Political Economy (McGill University), the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Bank of Canada (Research Fellowship), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Institut de finance mathématique de Montréal (IFM2), the Canadian Network of Centres of Excellence [program on Mathematics of Information Technology and Complex Systems (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de recherche sur la société et la culture (Québec). Financial support from the Spanish Ministry of Education through grants SEJ 2007-63098 is acknowledged.

References

Abdelkhalek, T. and Dufour, J.-M. (1998), 'Statistical inference for computable general equilibrium models, with application to a model of the Moroccan economy', *Review of Economics and Statistics* **80**, 520–534.

Bahadur, R. and Savage, L. J. (1956), 'The nonexistence of certain statistical procedures in nonparametric problems', *Annals of Mathematical Statistics* **27**, 1115–22.

Begum, N. and King, M. (2005), 'Most mean powerful test of a composite null against a composite alternative', *Journal Computational Statistics & Data Analysis* **49(4)**, 1079–1104.

Boldin, M. V., Simonova, G. and Tyurin, Y. N. (1997), 'Sign-based methods in linear statistical models', *Translations of Mathematical Monographs, American Mathematical Society, Vol. 162.* .

Campbell, B. and Dufour, J.-M. (1991), 'Over-rejections in rational expectations models: A nonparametric approach to the Mankiw-Shapiro problem', *Economics Letters* **35**, 285–290.

Campbell, B. and Dufour, J.-M. (1995), 'Exact nonparametric orthogonality and random walk tests', *Review of Economics and Statistics* **77**, 1–16.

Campbell, B. and Dufour, J.-M. (1997), 'Exact nonparametric tests of orthogonality and random walk in the presence of a drift parameter', *International Economic Review* **38**, 151–173.

Capanu, M., Jones, G. A. and Randles, R. H. (2006), 'Testing for preference using a sum of Wilcoxon signed rank statistics', *Journal Computational Statistics & Data Analysis* **51(2)**, 793–796.

Christoffersen, P. F. (1998), 'Evaluating interval forecasts', *International Economic Review* **39**, 841–862.

Coudin, E. and Dufour, J.-M. (2008), 'Finite-sample distribution-free inference in linear median regressions under heteroskedasticity and nonlinear dependence of unknown form', *Econometrics Journal, 10th anniversary special edition* **forthcoming**.

Davies, R. (1973), 'Numerical inversion of a characteristic function', *Biometrika* **60**, 415–417.

Davies, R. (1980), 'The distribution of a linear combination of chi-squared random variables', *Applied Statistics* **29**, 323–333.

Dufour, J. and Iglesias, E. M. (2008), Finite sample and optimal inference in possibly nonstationary general volatility models with gaussian and heavy-tailed errors, Technical report, Centre interuniversitaire de recherche en analyse des organisations (CIRANO) and Centre interuniversitaire de recherche en économie quantitative (CIREQ), Université de Montréal, Montréal, Québec.

Dufour, J.-M. (1997), 'Some impossibility theorems in econometrics with applications to structural and dynamic models', *Econometrica* **65**, 1365–1389.

Dufour, J.-M. (2003), 'Identification, weak instruments and statistical inference in econometrics', *Canadian Journal of Economics* **36(4)**, 767–808.

Dufour, J.-M. (2006), 'Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics in econometrics', *Journal of Econometrics* **133**(2), 443–477.

Dufour, J.-M. and Hallin, M. (1991), 'Nonuniform bounds for nonparametric t tests', *Econometric Theory* **7**, 253–263.

Dufour, J.-M. and Jasiak, J. (2001), 'Finite sample limited information inference methods for structural equations and models with generated regressors', *International Economic Review* **42**, 815–844.

Dufour, J.-M., Jouneau, F. and Torrès, O. (2008), On (non)-testability : Applications to linear and nonlinear semiparametric and nonparametric regression models, Technical report, Department of Economics and CIREQ, McGill University, and GREMARS, Université de Lillie 3.

Dufour, J.-M. and King, M. L. (1991), 'Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors', *Journal of Econometrics* **47**, 115–143.

Dufour, J.-M. and Kiviet, J. F. (1998), 'Exact inference methods for first-order autoregressive distributed lag models', *Econometrica* **66**, 79–104.

Dufour, J.-M. and Taamouti, M. (2003), Projection-based statistical inference in linear structural models with possibly weak instruments, Technical report, C.R.D.E., Université de Montréal.

Dufour, J.-M. and Taamouti, M. (2005), 'Projection-based statistical inference in linear structural models with possibly weak instruments', *Econometrica* **73**(4), 1351 – 1365.

Dufour, J.-M. and Torrès, O. (1998), Union-intersection and sample-split methods in econometrics with applications to SURE and MA models, *in* D. E. A. Giles and A. Ullah, eds, 'Handbook of Applied Economic Statistics', Marcel Dekker, New York, pp. 465–505.

Elliott, G., Rothenberg, T. J. and Stock, J. H. (1996), 'Efficient tests for an autoregressive unit root', *Econometrica* **64**(4), 813–836.

Gerard, P. and Schucany, W. (2007), 'An enhanced sign test for dependent binary data with small numbers of clusters', *Journal Computational Statistics & Data Analysis* **51(9)**, 4622–4632.

Gil-Pelaez, J. (1951), 'Note on the inversion theorem', *Biometrika* **38**, 481–482.

Imhof, J. P. (1961), 'Computing the distribution of quadratic forms in normal variables', *Biometrika* **48**, 419–426.

Jansson, M. (2005), 'Point optimal tests of the null hypothesis of cointegration', *Journal of Econometrics* **124**, 187–201.

King, M. L. (1987-88), 'Towards a theory of point optimal testing (with comments)', *Econometric Reviews* **6**, 169–255.

Lehmann, E. L. (1959), *Testing Statistical Hypotheses*, New York: John Wiley.

Lehmann, E. and Stein, C. (1949), 'On the theory of some non-parametric hypotheses', *Annals of Mathematical Statistics* **20**, 28–45.

Liang, T., Huang, W. and Yang, K. (2008), 'Locally optimal tests for exponential distributions with type-i censoring', *Journal Computational Statistics & Data Analysis* **52(7)**, 3603–3615.

Mankiw, N. G. and Shapiro, M. (1986), 'Do we reject too often? small sample properties of tests of rational expectations models', *Economic Letters* **20**, 139–145.

O'Gorman, T. (2004), *Applied Adaptive Statistical Methods: Tests of Significance and Confidence Intervals*, Society for Industrial and Applied Mathematics Philadelphia, PA, USA.

Pratt, J. and Gibbons, J. (1981), *Concepts of Nonparametric Theory*, New York: Springer-Verlag.

Rousseeuw, P. J. and Leroy, A. (1987), *Robust Regression and Outlier Detection*, Wiley Series in Probability and Mathematical Statistics. Wiley, New York.

Rousseeuw, P. J. and Yohai, V. (1984), *Robust Regression by Means of S-Estimators*, in Robust and Nonlinear Time Series Analysis, edited by J. Franke, W. Hardle, and D. Martin, Lecture Notes in Statistics No. 26, Springer Verlag, Berlin.

White, H. (1980), 'A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity', *Econometrica* **48**, 817–838.

Wright, J. H. (2000), 'Alternative variance-ratio tests using ranks and signs', *Journal of Business and Economic Statistics* **18(1)**, 1–9.

Yohai, V. J. and Zamar, R. (1988), 'High breakdown point estimates of regression by means of the minimization of an efficient scale', *Journal of the American Statistical Association* **83**, 406–413.