UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 08-60
Economic Series (27)
November 2008

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

Exact optimal and adaptive inference in regression models under heteroskedasticity and non-normality of unknown forms*

Jean-Marie Dufour†
McGill University

Abderrahim Taamouti‡
Universidad Carlos III de Madrid

Abstract

In this paper, we derive simple point-optimal sign-based tests in the context of linear and nonlinear regression models with fixed regressors. These tests are exact, distribution-free, robust against heteroskedasticity of unknown form, and they may be inverted to obtain confidence regions for the vector of unknown parameters. Since the point-optimal sign tests depend on the alternative hypothesis, we propose an adaptive approach based on split-sample techniques in order to choose an alternative such that the power of point-optimal sign tests is close to the power envelope. The simulation results show that when using approximately 10% of sample to estimate the alternative and the rest to calculate the test statistic, the power of point-optimal sign test is typically close to the power envelope. We present a Monte Carlo study to assess the performance of the proposed "quasi"-point-optimal sign test by comparing its size and power to those of some common tests which are supposed to be robust against heteroskedasticity. The results show that our procedures are superior.

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

Journal of Economic Literature classification: C1; C12; C14; C15; C51.

* The authors thank Lynda Khalaf, Frédéric Jouneau-Sion, Benoit Perron, Thi Thuy Anh Vo, the participants of 46th Annual Meeting of the Société canadienne de science économique at Montreal (2006), 61th European Meeting of the Econometric Society at Vienna (2006), and 2nd International Workshop on Computational and Financial Econometrics (CFE'08) at Neuchâtel (Switzerland) (2008) for their helpful comments. 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.

† William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse desorganisations (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.

Contents

List of Tables

List of Figures

1. Introduction

Most economic data are heteroskedastic and non-normally distributed. In the presence of several types of heteroskedasticity, parametric tests proposed to improve inference may not control size and have very low power. This is the case, in particular, for common tests based on White (1980) variance correction – which is supposed to be robust against heteroskedasticity – when there is a break in the disturbance variance or with a GARCH structure with one or several outliers.[1] At the same time, many *exact* parametric tests developed in the literature typically assume normal disturbances. The latter assumption is unrealistic and in the presence of heavy tails and asymmetric distributions these tests may not perform very well in terms of power. Furthermore, the statistical procedures developed for the inference on parameters of *nonlinear* models are typically based on asymptotic approximations and there are only a few exact inference methods outside linear models framework. However, these approximations may be invalid in small samples and even in large samples [see Dufour (2003)]. The present paper aims to propose exact optimal tests which work under more realistic assumptions. We derive simple point-optimal sign-based tests which are valid under weak distributional assumptions such as heteroskedascticity of unknown form and non-normality.

Several authors have provided theoretical arguments for why the existing parametric tests about the mean of *i.i.d.* observations fail under weak distributional assumptions such as non-normality and heteroskedasticity of unknown form. Bahadur and Savage (1956) show that under weak distributional assumptions on the error terms, it is not possible to obtain a valid test for the mean of *i.i.d.* observations even for large samples. Many other hypotheses about various moments of *i.i.d.* observations lead to similar difficulties. This can be explained by the fact that the moments are not empirically meaningful in nonparametric models or models with weak assumptions. More discussion about the statistical inference problems in nonparametric models can be find in Dufour (2003). Further, Lehmann and Stein (1949) and Pratt and Gibbons (1981) show that sign methods were the only possible way of producing valid inference for finite sample procedures under conditions of heteroskedasticity of unknown form and non-normality.

This paper introduces new sign-based tests in the context of linear and nonlinear regression models with fixed regressors. The proposed tests are exact, distribution-free, robust against heteroskedasticity of unknown form, and they may be inverted to obtain confidence regions for the vector of unknown parameters. These tests are derived under assumptions that the disturbances in regression models are independent, but not necessarily identically distributed, with zero median conditional on the explanatory variables. A number of sign-based test procedures have been developed in the literature. In the presence of only one explanatory variable, Campbell and Dufour (1995) and Campbell and Dufour (1997) propose nonparametric analogues of *t-test*, based on sign and signed rank statistics, that are applicable to a specific class of feedback models including both Mankiw and Shapiro (1986) and random walk models. These tests are exact even when the disturbances are asymmetric, non-normal, and heteroskedastic. Boldin, Simonova and Tyurin (1997)

[1] See the simulation results in Section 6 of this paper. Financial markets are characterized by the presence of episodic occasional of crashes and rallies. The latter can be viewed as introducing outliers in volatility. Moreover, it may occur that financial returns series contain other atypical observations such as additive or innovation outliers. The reader can consult Hotta and Tsay (1998) for a recent classification of outliers in $GARCH$ volatility models and Friedman and Laibson (1989) for the economic arguments for the possible presence of atypical observations.

propose locally optimal sign-based inference and estimation for linear models. Coudin and Dufour (2008) extend the work by Boldin et al. (1997) to some forms of statistical dependence in the data. 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.

The present paper address the issue of optimality and seeks to derive point-optimal tests based on sign statistics. Point-optimal tests are useful in a number of ways and they are most attractive for problems in which the size of the parameter space can be restricted by theoretical considerations. Because of their power properties, these tests are particularly attractive when testing one economic theory against another, for example a new theory against an existing theory. They would ensure optimal power at given point and, depending on the structure of the problem, could give power over the entire parameter space. Another interesting feature of these tests is they can be used to trace out the maximum attainable *power envelope* for a given testing problem. This power envelope provides an obvious benchmark against which test procedures can be evaluated. More discussion about the usefulness of point-optimal tests can be found in King (1987-88). Many papers have proposed point-optimal tests to improve inference. 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 the null hypothesis 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.

Since point-optimal sign (hereafter POS) tests depend on the alternative hypothesis, we propose an adaptive approach based on split-sample technique to choose an alternative that makes the power curve of the POS test close to the power envelope.[2] 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 point-optimal sign test statistic. The simulation results show that using approximately 10% of sample to estimate the alternative yields a power which is typically very close to the power envelope. We present a Monte Carlo study to assess the performance of the proposed "quasi"-point-optimal sign test by comparing its size and power to those of some common tests which are supposed to be robust against heteroskedasticity. The results show that our procedures are superior.

The plan of the paper is as follows. In Section 2, we present a general framework for deriving POS tests. In Section 3, we derive 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 the optimal alternative. In Section 5, we discuss the construction of the point-optimal sign confidence regions using projection technique. In Section 6, we present a Monte Carlo simulation to assess the performance of POS tests by comparing their size and power to those of some popular tests. We conclude in Section 7. Technical proofs and simulation results are presented in appendices A and B, respectively.

[2]For more details about split-sample technique, the reader can consult Dufour and Torrès (1998) and Dufour and Jasiak (2001).

2. General framework

This section aims to introduce a framework for deriving point-optimal sign-based 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 size of 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 could give power over the entire parameter space. In what follows, we consider a random sample $\{y_t\}_{t=1}^{n}$ such that

$$y_1, ..., \ y_n \text{ are independent} \tag{2.1}$$

and we define the following vector of signs

$$U(n) = (s(y_1), ..., s(y_n))'$$

where

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

we assume that there is no probability mass at zero, which is true if y_t is a continuos variable. In the next two subsections, we use Neyman-Pearson lemma [see e.g. Lehmann (1959), page 65] to derive point-optimal sign-based tests to test simple constant and nonconstant hypotheses.

2.1. Point-optimal sign test for constant hypotheses

Let $(y_1, ..., y_n)'$ be an observable $n \times 1$ vector of independent random variables such that $\mathsf{P}[y_t \geq 0] = p$. Suppose we wish to test the null hypothesis

$$H_0 : p = p_0, \tag{2.2}$$

against the alternative hypothesis

$$H_1 : p = p_1 \tag{2.3}$$

where p_0 and p_1 are fixed and known. In what follows, we consider optimal tests, in the Neyman-Pearson sense, which maximize the power function (minimize the Type II error) under the constraint

$$\mathsf{P}\,[\text{reject } H_0 \mid H_0] \leq \alpha.$$

If we denote the likelihood function of $y = (y_1, ..., y_n)'$ under the null by $f(y \mid H_0)$ and its likelihood function under the alternative by $f(y \mid H_1)$, then the Neyman-Pearson lemma [see e.g. Lehmann (1959), page 65] implies that rejecting H_0 for large values of likelihood ratio

$$s = \frac{f(y \mid H_1)}{f(y \mid H_0)} \tag{2.4}$$

corresponds to the most powerful test. A critical value, say c, of test statistic (2.4) is given by the smallest constant c such that

$$\mathsf{P}[s > c \mid H_0] \leq \alpha,$$

where α is the desired level of significance or a Type I error. The choice of a significance level α is usually somewhat arbitrary, since in most situations there is no precise limit to the probability of a Type I error that can be tolerated. Standard values, such as 0.01 or 0.05, were originally chosen to effect a reduction in the tables needed for carrying out various test. However, the choice of significance level should take into account the power that the test will achieve against the alternative of interest. Rules for choosing α in relation to the attainable power are discussed in Lehmann (1958), Arrow (1960), Sanathanan (1974), and Lehmann and Romano (2005). For our testing problem, the likelihood function is given by:

$$L\left(U(n), p\right) = \mathop{\mathit{\Pi}}_{t=1}^{n} \mathsf{P}[y_t \geq 0]^{s(y_t)} \left(1 - \mathsf{P}[y_t \geq 0]\right)^{1-s(y_t)} \tag{2.5}$$

and the likelihood ratio is defined as follows:

$$\frac{L_1(U(n), p_1)}{L_0(U(n), p_0)} = \prod_{t=1}^{n} \left\{ \left(\frac{p_1}{p_0}\right)^{s(y_t)} \left(\frac{1-p_1}{1-p_0}\right)^{1-s(y_t)} \right\} = \left(\frac{p_1}{p_0}\right)^{S_n} \left(\frac{1-p_1}{1-p_0}\right)^{n-S_n}, \tag{2.6}$$

where

$$S_n = \sum_{t=1}^{n} s(y_t),$$

$L_0(.)$ and $L_1(.)$ represent the values of likelihood function (2.5) under the null and alternative hypotheses, respectively. For simplicity of exposition we assume that $p_0,\ p_1 \neq 0, 1$. The latter allows us to work with the log-likelihood function and simplify the expression of test statistics. However, for $p_0 = 0,\ 1$ we could work directly with likelihood function rather than log-likelihood function. We deduce the following log-likelihood ratio:

$$\ln\left\{\frac{L_1(U(n), p_1)}{L_0(U(n), p_0)}\right\} = S_n \left\{ \ln\left(\frac{p_1}{p_0}\right) - \ln\left(\frac{1-p_1}{1-p_0}\right) \right\} + n \ln\left(\frac{1-p_1}{1-p_0}\right).$$

According to Neyman-Pearson lemma, the best test of H_0 against H_1, based on the vector of signs $(s(y_1), ..., s(y_n))'$, rejects H_0 when

$$\ln\left\{\frac{L_1(U(n), p_1)}{L_0(U(n), p_0)}\right\} > c. \tag{2.7}$$

For $p_1 > p_0 > 0$, the test (2.7) is equivalent to rejecting H_0 when

$$S_n > \frac{c - n\ \ln\ (\frac{1-p_1}{1-p_0}\ )}{\ln\ (\frac{p_1}{p_0}\ ) - \ln\ (\frac{1-p_1}{1-p_0}\ )} \equiv c_1, \tag{2.8}$$

where c_1 satisfies

$$\mathsf{P}[S_n > c_1 \mid H_0] \leq \alpha.$$

The test (2.8) is the same for all $p_1 > p_0$. Further, under assumption (2.1) and for $p_1 > p_0 > 0$ the test with critical region

$$C = \left\{(y_1, ..., y_n)' : S_n > c_1\right\}$$

is the best POS test for the null hypothesis (2.2) against the alternative hypothesis (2.3). Similarly, for $0 < p_1 < p_0$, the test (2.7) is equivalent to rejecting H_0 when

$$S_n < \frac{c - n\ln\left(\frac{1-p_1}{1-p_0}\right)}{\ln\left(\frac{p_1}{p_0}\right) - \ln\left(\frac{1-p_1}{1-p_0}\right)} \equiv \bar{c}_1, \tag{2.9}$$

where $\bar{c}_1$ satisfies

$$\mathsf{P}[S_n < \bar{c}_1 \mid H_0] \leq \alpha.$$

The critical region which corresponds to the test (2.9) is defined as follows:

$$\bar{C} = \left\{(y_1, ..., y_n)' : S_n < \bar{c}_1\right\},$$

where the critical value $\bar{c}_1$ is chosen so that

$$\mathsf{P}\left[(y_1, ..., y_n)' \in \bar{C} \mid H_0\right] \leq \alpha.$$

In both cases, i.e. for $p_1 > p_0 > 0$ and $0 < p_1 < p_0$, the test statistic is given by:

$$S_n = \sum_{t=1}^{n} s(y_t). \tag{2.10}$$

Under H_0, S_n follows a binomial distribution $Bi(n, p_0)$, i.e.

$$\mathsf{P}(S_n = i) = C_n^i\, p_0^i\, (1 - p_0)^{n-i},$$

where $C_n^i = \frac{n!}{[i!(n-i)!]}$. Since the tests statistic (2.10) does not depend on the alternative hypothesis p_1, the above test correspond to *uniformly most powerful* tests.

Example 2.1 (Backtesting Value-at-Risk) Consider time series of 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 if on average 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

$$H_0 : I_t \sim i.i.d : Bernoulli(p) \tag{2.11}$$

against the alternative hypothesis

$$H_1 : I_t \sim i.i.d : Bernoulli(\bar{p}). \tag{2.12}$$

Under H_0, the likelihood function of random sequence $\{I_t\}_{t=1}^{T}$ is given by:

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

where

$$S_T = \sum_{t=1}^{T} I_t,$$

and under the alternative it is given by:

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

Using Neyman-Pearson lemma and the previous results, a test statistic for testing the null hypothesis (2.11) against the alternative hypothesis (2.12) is defined as follows:

$$S_T = \sum_{t=1}^{T} I_t,$$

where, under H_0, S_T follows a binomial distribution $Bi(T, p)$.

2.2. Point-optimal sign test for nonconstant hypotheses

Let $(y_1, ..., y_n)'$ be an observable $n \times 1$ vector of independent random variables such that $\mathsf{P}[y_t \geq 0] = p_t$. Suppose we wish now to test the null hypothesis

$$H_0 : \mathsf{P}[s(y_t) = 1] = p_{t,0},\ t = 1, ..., n, \tag{2.13}$$

against the alternative hypothesis

$$H_1 : \mathsf{P}[s(y_t) = 1] = p_{t,1},\ t = 1, ..., n. \tag{2.14}$$

For simplicity of exposition we assume again that $p_{t,0}$, $p_{t,1} \neq 0, 1$. Following the same steps as in Section 2.1, a point-optimal sign test to test the null hypothesis (2.13) against the alternative hypothesis (2.14) is given by the following theorem.

Theorem 2.2 *Under assumption (*2.1*) the test with critical region*

$$C = \left\{ (y_1, ..., y_n)' : \sum_{t=1}^{n} \ln \left[\frac{p_{t,1}(1 - p_{t,0})}{p_{t,0}(1 - p_{t,1})} \right] \; s(y_t) > c_1 \right\}$$

*is the best point-optimal sign test for the null hypothesis (*2.13*) against the alternative hypothesis (*2.14*). The critical value c_1 is chosen such that*

$$\mathsf{P}\left[(y_1, ..., y_n)' \in C \mid H_0 \right] \leq \alpha,$$

where α is an arbitrary significance level.

See proof of Theorem **2.2** in Appendix A. The corresponding test statistic is given by:

$$S_n^* = \sum_{t=1}^{n} \left[\frac{p_{t,1}(1 - p_{t,0})}{p_{t,0}(1 - p_{t,1})} \right] \; s(y_t). \tag{2.15}$$

Next section proposes exact optimal sign-based tests in the context of linear and nonlinear regression models with fixed regressors. The hypothesis testing problem in this section corresponds to a special case of the one defined before, since the null hypothesis $p_{t,0}$ and the alternative hypothesis $p_{t,1}$ take particular forms.

3. Point-optimal sign tests in linear and nonlinear regression models

In the presence of some types of heteroskedasticity, the parametric tests proposed to improve inference may not control size and have very low power. For example, when there is a break in the disturbances' variance the usual test statistic based on White (1980) variance correction, which is supposed to be robust against heteroskedasticity, has very poor power [see Section 6]. Other forms of heteroskedasticity for which the usual tests are less powerful are exponential variance and GARCH with one or several outliers [see Section 6]. At the same time, many *exact* parametric tests developed in the literature typically assume normal disturbance. The latter assumption is unrealistic and in the presence of heavy tails and asymmetric distributions these tests may not perform very well in terms of power. Furthermore, the statistical procedures developed for the inference on parameters of *nonlinear* models are typically based on asymptotic approximations and there are only a few exact inference methods outside linear models framework. However, these approximations may be invalid in small samples and even in large samples [see Dufour (1997)].

In this section, we propose exact optimal sign-based tests in the context of linear and nonlinear regression models with fixed regressors. The proposed tests are valid under weak distributional assumptions such as heteroskedascticity of unknown form and non-normality. In the next first subsection we derive POS tests to test zero coefficient hypothesis in the context of linear regression model, and in the second and last subsection, we propose POS tests to test nonzero coefficient hypothesis in the context of nonlinear regression model.

3.1. Testing zero coefficient hypothesis in linear regression model

We consider that the variable y_t can be linearly explained by another variable vector x_t

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

where y_t, for $t = 1, ..., n$, are independent random variables but not necessarily identically distributed, x_t is an observable $k \times 1$ vector of fixed explanatory variables, $\beta \in \mathbb{R}^k$ is an unknown vector of parameters, and ε_t is an error term such that

$$\varepsilon_t \mid X \sim F_t(. \mid X) \tag{3.2}$$

and

$$\mathsf{P}[\varepsilon_t \geq 0 \mid X] = \mathsf{P}[\varepsilon_t < 0 \mid X] = \frac{1}{2}, \tag{3.3}$$

where $F_t(.)$ is a distribution function and $X = [x_1, ..., x_n]'$ is an $n \times k$ matrix. Suppose we wish to test the null hypothesis

$$H_0 : \beta = 0 \tag{3.4}$$

against the alternative hypothesis

$$H_1 : \beta = \beta_1. \tag{3.5}$$

Under assumption (3.1), the hypothesis testing problem given by (3.4)-(3.5) is a special case of the one defined by (2.13)-(2.14) [see Section 2.2], where now

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

Under H_0

$$p_{t,0} = \mathsf{P}[y_t \geq 0 \mid X] = 1 - \mathsf{P}[\varepsilon_t < 0 \mid X] = \frac{1}{2} \tag{3.6}$$

and under H_1

$$p_{t,1} = \mathsf{P}[y_t \geq 0 \mid X] = 1 - \mathsf{P}[\varepsilon_t < -\beta_1' x_t \mid X]. \tag{3.7}$$

Consequently, a sign-based test for the null hypothesis (3.4) against the alternative hypothesis (3.5) can be deduced from Theorem **2.2** using equations (3.6) and (3.7). We have the following result.

Proposition 3.1 *Under assumptions (*2.1*), (*3.3*), and (*3.1*) the best point-optimal sign test for the null hypothesis (*3.4*) against the alternative hypothesis (*3.5*) rejects (*3.4*) when*

$$\sum_{t=1}^{n} a_t(0/1) s(y_t) > c_1(\beta_1),$$

where

$$a_t(0/1) = \ln \left[\frac{1}{\frac{1}{1 - \mathsf{P}[\varepsilon_t \leq -\beta_1' x_t \mid X]} - 1} \right]. \tag{3.8}$$

The critical value $c_1(\beta_1)$ is chosen such that

$$\mathsf{P}\left[\sum_{t=1}^{n} a_t(0/1)s(y_t) > c_1(\beta_1) \mid H_0\right] \leq \alpha,$$

where α is an arbitrary significance level.

The POS test of Proposition **3.1** controls size for any distribution of ε_t which satisfies the condition (3.3). Under the null hypothesis, the test is distribution-free and robust against heteroskedasticity of unknown form. However, under the alternative hypothesis the power function of the test will depend on the form of the distribution function of ε_t.

We assume that under the alternative hypothesis the error terms follow an homoskedastic normal distribution. In other words, we substitute the optimal weights (3.8) by the weights derived from an homoskedastic normal distribution. The latter assumption may affect the power of POS test, however our simulation results [see tables 7 and 8] show that there is almost no power loss when we misspecify the distribution function of ε_t. If we assume that under the alternative

$$\varepsilon_t \sim N(0,1),$$

then the test statistic is given by:

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

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

We use quantiles of random variable (3.9) to implement the POS test. Since the test statistic (3.9) is a continuous variable, its quantiles are easy to compute. To simulate (3.9) we first generate a sequence $\{s(y_t)\}_{i=1}^{n}$ under the null hypothesis. In particular, we generate a sequence $\{s(\varepsilon_i)\}_{i=1}^{n}$ which satisfies the condition (3.3). The variable $s(\varepsilon_t)$ takes only two values 0 and 1, thus the computation of test statistic (3.9) reduces to generating a sequence of Bernoulli random variables of given length with subsequent summation and the corresponding weights. The algorithm for implementing the POS test can be described as follows:

1. Compute the test statistic $S_n^*(\beta_1)$ based on the observed data, say $S_n^*(\beta_1)^0$;
2. Generate a sequence of Bernoulli random variables $\{s(\varepsilon_i)\}_{i=1}^{n}$ satisfying (3.3);
3. Compute $S_n^*(\beta_1)^j$ using the sequence $\{s(\varepsilon_i)\}_{i=1}^{n}$ and the corresponding weights $\{a_i(0/1)\}_{i=1}^{n}$;
4. Choose B such that $\alpha(B+1)$ is an integer and repeat steps $1-3$ B times;
5. Compute the $(1-\alpha)\%$ quantile, say $c(\beta_1)$, of the sequence $\left\{S_n^*(\beta_1)^j\right\}_{j=1}^{B}$;
6. Reject the null hypothesis at level α if $S_n^*(\beta_1)^0 \geq c(\beta_1)$.

3.2. Testing Nonzero coefficient hypothesis in nonlinear regression model

We consider now the following nonlinear regression model

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

where y_t, for $t = 1, ..., n$, are independent random variables but not necessarily identically distributed, x_t is an observable $k \times 1$ vector of fixed explanatory variables, $f(.)$ is a scaler function, $\beta \in \mathbb{R}^k$ is an unknown vector of parameters, and ε_t is an error term which satisfies the condition (3.3). Suppose now we wish to test the null hypothesis

$$H_0 : \beta = \beta_0 \tag{3.11}$$

against the alternative hypothesis

$$H_1 : \beta = \beta_1. \tag{3.12}$$

A test for H_0 against H_1 can be constructed in the same way as in the previous Section 3.1. First, notice that the model (3.10) is equivalent to the following transformed model

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

where

$$\tilde{y}_t = y_t - f(x_t, \beta_0) \text{ and } g(x_t, \beta, \beta_0) = f(x_t, \beta) - f(x_t, \beta_0).$$

Under assumption (2.1) and conditional on X we have

$$\tilde{y}_1, ..., \tilde{y}_n \text{ are independent.}$$

Second, testing (3.11) against (3.12) is equivalent to test

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

against

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

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

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

where the new vector of signs $\tilde{U}(n)$ is defined as follows:

$$\tilde{U}(n) = (s(\tilde{y}_1), ..., s(\tilde{y}_n))',$$

for

$$s(\tilde{y}_t) = \begin{cases} 1, if\ \tilde{y}_t \geq 0 \\ 0, if\ \tilde{y}_t < 0 \end{cases}.$$

Thus, a sign-based test for the null hypothesis (3.11) against the alternative hypothesis (3.12) can be derived using the above Proposition **3.1**. We have the following result.

Proposition 3.2 *Under assumptions (*2.1*), (*3.3*), and (*3.10*) the best point-optimal sign test for the null hypothesis (*3.11*) against the alternative hypothesis (*3.12*) reject (*3.12*) when*

$$\sum_{i=1}^{n} \ln \left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t \le f(x_t,\beta_0)-f(x_t,\beta_1)|X]} - 1} \right] s\left(y_t - f(x_t,\beta_0)\right) > c_1(\beta_1).$$

The critical value $c_1(\beta_1)$ *is chosen so that*

$$\mathsf{P}\left[\sum_{t=1}^{n} \tilde{a}_t(0/1) s\left(y_t - f(x_t,\beta_0)\right) > c_1(\beta_1) \mid H_0 \right] \le \alpha$$

and α *is an arbitrary significance level.*

If we consider a linear function $f(x_t,\beta) = \beta' x_t$ and assume that under the alternative hypothesis the error term ε_t follows $N(0,1)$ distribution, then the test statistic is given by:

$$S_n^*(\beta_1) = \sum_{t=1}^{n} \ln \left[\frac{1}{\frac{1}{\Phi((\beta_1-\beta_0)' x_t)} - 1} \right] s\left(y_t - \beta_0' x_t\right), \tag{3.13}$$

where $\Phi(.)$ is the standard normal distribution function. The test statistic (3.13) depends on a particular alternative hypothesis β_1. In practice, the latter is supposed to be unknown which makes the proposed POS test unfeasible. In the next section, we propose some additional techniques that we can use in order to choose an optimal alternative β_1 at which the power of 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 point-optimal sign test

We derive an upper bound (hereafter power envelope) of the power function of point-optimal sign test. It is well known 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.

According to Section 3, the test statistic of POS test is a function of β_1

$$S_n^*(\beta_1) = \sum_{t=1}^{n} \ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t \leq -\beta_1' x_t|X]} - 1}\right] s(y_t).$$

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

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

where c_1 satisfies

$$\mathsf{P}[S_n^*(\beta_1) > c_1 \mid H_0] \leq \alpha.$$

The following theorem provides a theoretical formula for power function of POS test.

Theorem 4.1 *Under assumptions (*2.1*) and (*3.3*), the power function of POS test at β_1 is given by*

$$\Pi(\beta, \beta_1) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \frac{I(u)}{u} du,$$

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

$$I(u) = \left(\frac{1}{2}\right)^n \operatorname{Im}\left\{\prod_{t=1}^{n}\left[\exp\left(-iu\frac{c_1}{n}\right) + \exp\left(iu\left(\ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t \leq -\beta_1' x_t|X]} - 1}\right] - \frac{c_1}{n}\right)\right)\right]\right\},$$

$i = \sqrt{-1}$ and $\operatorname{Im}\{z\}$ denotes the imaginary part of a complex number z. The critical value c_1 is chosen so that

$$\mathsf{P}[S_n^*(\beta_1) > c_1 \mid H_0] \leq \alpha,$$

where α is an arbitrary significance level.

See proof of Theorem **4.1** in Appendix A. Since the test statistic $S_n^*(\beta_1)$ is optimal against an 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 $\beta \in \mathbb{R}^k$,

$$\bar{\Pi}(\beta) = \Pi(\beta, \beta) = \mathsf{P}\left[S_n^*(\beta) > c_1\,\right]. \tag{4.1}$$

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

$$\left\{S_n^*(\Pi) = \sum_{t=1}^{n} \ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t \leq -\boldsymbol{\beta}_1(\boldsymbol{\Pi},\boldsymbol{\alpha})' x_t|X]} - 1}\right] s(y_t),\ \text{for } \Pi \in (\alpha, 1)\right\}.$$

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 $\Pi = 0.5$. Based on 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[3]

$$\int_0^\infty \mathrm{Im}\left\{\frac{\prod_{t=1}^{n}\left[\exp\left(-iu\frac{c_1}{n}\right)+\exp\left(iu\left(\ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t \le -\beta_1' x_t|X]}-1}\right]-\frac{c_1}{n}\right)\right)\right]}{u}\right\}du = 0. \quad (4.2)$$

An exact solution for equation (4.2) is not feasible, since it is not easy to find an expression for $\mathrm{Im}\{.\}$ and the integral $\int_0^\infty \mathrm{Im}\{.\}du$ is difficult to evaluate. The latter can be approximated using results from Imhof (1961), Bohman (1972), Davies (1973), and Davies (1980), 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 to solve the power envelope function for β_1 is to use simulations [see Elliott et al. (1996)]. We could 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's now examine the impact of the alternative hypothesis β_1 on the power function. We use simulations and plot the power curves of POS test under different alternatives and compare them to the power envelope. Our results are presented in figures 1-3.

Insert Figures 1-3.

The above figures compare the power curves of POS test to the power envelope under different alternatives and using different 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 are interesting. A description of these DGPs is given in Section 6. Based on simulation results, we find that the alternative hypothesis affects the power function. Particularly, when the alternative 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 way, in the

[3]Using the properties of the cumulative density function (monotonically increasing, continuous $\lim_{c\to-\infty} \Pr(z < c) = 0$, and $\lim_{c\to+\infty} P_t(z < c) = 1$) we can show that equation (4.2) has a unique solution.

next subsection we propose an adaptive approach based on split-sample technique to estimate the optimal alternative.

4.2. An adaptive approach to choose the optimal alternative

Existing adaptive statistical methods use the data to determine which statistical procedure is most appropriate for a specific testing problem. These methods are usually performed in 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 that we consider is somewhat different from the existing adaptive statistical approaches. We propose split-sample technique to choose an alternative hypothesis β_1 such that the power of POS test is close to the power envelope.[4] 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 from the one that we use 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.

Consider again the model given by (3.1) and 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. We use the first n_1 observations, $y_{(1)}$ and $X_{(1)}$, to estimate the alternative hypothesis β_1 using OLS

$$\hat{\beta}_1 = (X^{'}_{(1)} X_{(1)})^{-1} X^{'}_{(1)} y_{(1)}$$

and 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

$$S^*_n(\hat{\beta}_1) = \sum_{t=n_1+1}^{n} \ln \left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t \le -((X^{'}_{(1)} X_{(1)})^{-1} X^{'}_{(1)} y_{(1)})^{'} x_t | X]} - 1} \right] s(y_t).$$

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)].

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

[4]For more details about split-sample technique, the reader can consult Dufour and Torrès (1998) and Dufour and Jasiak (2001).

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 using simulation experiments. 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 are presented in figures 4-6.

Insert Figures 4-6.

>From the above figures we see 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 that we consider in the simulation study.

5. Point-optimal sign 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 previous linear regression model (3.1) and suppose we wish to test (3.11) against (3.12), then the idea consists in finding all the values of $\beta_0 \in \mathbb{R}^k$ such that

$$S_n^{*(0)}(\beta_1) = \sum_{t=1}^{n} \left\{ \ln \left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t \leq -(\beta_1-\beta_0)' x_t | X]} - 1} \right] s(y_t - \beta_0' x_t) \right\} < c(\beta_1),$$

where $S_n^{*(0)}(\beta_1)$ is the observed value of $S_n^*(\beta_1)$ and the critical value $c(\beta_1)$ is given by the smallest constant $c(\beta_1)$ such that

$$\mathsf{P}[S_n^*(\beta_1) > c(\beta_1) \mid \beta = \beta_0] \leq \alpha.$$

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

$$C_\beta(\alpha) = \left\{ \beta_0 : S_n^{*(0)}(\beta_1) < c(\beta_1) \mid \mathsf{P}\left[S_n^*(\beta_1) > c(\beta_1) \mid \beta = \beta_0 \right] \leq \alpha \right\}.$$

Further, given the confidence region $C_\beta(\alpha)$, we can also derive confidence intervals for the components of vector β using the projection techniques.[5] 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

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

[5] 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).

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

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

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.

6.1. Size and Power

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 that one can encounter in practice. We consider the following regression model

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

where β is an unknown parameter and the error terms ε_t, for $t = 1, ..., n$, 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

$$\varepsilon_t \sim \text{Cauchy};$$

3. Student's distribution with two degrees of freedom

$$\varepsilon_t \sim \text{Student } (2);$$

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

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

and

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

8. Non stationary GARCH$(1,\ 1)$ variance

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2(t))$$

and

$$\sigma_\varepsilon^2(t) = 0.75\varepsilon_{t-1}^2 + 0.75\sigma_\varepsilon^2(t-1).$$

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 test (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 CD (1995) test). In what follows, the notations CT-test and CWT-test correspond 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 sizes. 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. Results

Monte Carlo simulation results are presented in tables 1-6 and figures 7-10 of Appendix B. These results correspond to different DGPs described in Section 6.1. To summarize, tables 1-6 of Appendix B 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 CD (1995) test. Figures 7-10 of Appendix B compare the power of 10% SS-POS test, T-test (CT-test), WT-test (CWT-test), and CD (1995) test to the power envelope. The results are detailed below.

First, Table 1 and Panel A of Figure 7 correspond to the case where the error term ε_t in the model (6.1) is normally distributed. Table 1 shows that the power of POS test depends on the alternative hypothesis β_1. When the latter is far from the null hypothesis, here $\beta = 0$, the POS test power's curve moves away from the power envelope [see also Panel A of Figure 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 Table 1. Panel A of Figure 7 shows that T-test is more powerful than 10% SS-POS test, CWT-test, and CD (1995) 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 CD (1995) test.

Second, Table 2 and Panel B of Figure 7 and Panel A of Figure 10 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 power curve of POS test moves away from the power envelope [see Table 2]. 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 7 and Panel A of Figure 10 shows that 10% SS-POS test is more powerful than T-test, WT-test, and CD (1995) test, and is close to the power envelope.

Third, Table 3 and Panel A of Figure 8, Table 5 and Panel A of Figure 9, and Table 6 and Panel B of Figure 9 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 8 and Panels A and B of Figure 9 show that 10% SS-POS test is again more powerful than T-test, WT-test, CD (1995) 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, Table 4 and Panel B of Figure 8 and Panel B of Figure 10 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 CD (1995) 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.[6] The results are presented in tables 7 and 8 of Appendix B. 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 point-optimal sign-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.

Since the proposed point-optimal sign 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 point-optimal sign 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 point-optimal sign test, yields a power which is typically very close to the power envelope.

To assess the performance of point-optimal sign 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 point-optimal sign 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.

[6] Weights $a_t(0/1)$ computed using homoskedastic and normal distribution.

A. Appendix: Proofs

PROOF OF THEOREM **2.2**. The likelihood function of random sample $\{y_t\}_{t=1}^{n}$ is defined as follows:

$$L(U(n), p_t) = \mathop{\mathit{\Pi}}_{t=1}^{n} \mathsf{P}[y_t \geq 0]^{s(y_t)}(1 - \mathsf{P}[y_t \geq 0])^{1-s(y_t)}. \tag{A.1}$$

Under the null hypothesis H_0 the likelihood function (A.1) is given by:

$$L_0(U(n), p_{t,0}) = \mathop{\mathit{\Pi}}_{t=1}^{n} \left\{ p_{t,0}^{s(y_t)}(1 - p_{t,0})^{1-s(y_t)} \right\} \tag{A.2}$$

and under the alternative hypothesis H_1 it is given by:

$$L_1(U(n), p_{t,1}) = \prod_{t=1}^{n} \left\{ p_{t,1}^{s(y_t)}(1 - p_{t,1})^{1-s(y_t)} \right\}. \tag{A.3}$$

For simplicity of exposition we assume that $p_{t,0}, p_{t,1} \neq 0, 1$. However, for $p_{t,0}, p_{t,1} \neq 0, 1$ we could work directly with likelihood function rather than log-likelihood function. From equations (A.2) and (A.3), the log-likelihood ratio is given by:

$$\ln \left\{ \frac{L_1(U(n), p_{t,1})}{L_0(U(n), p_{t,0})} \right\} = \sum_{t=1}^{n} [q_t(1) - q_t(0)]\, s(y_t) + \sum_{t=1}^{n} q_t(0), \tag{A.4}$$

where

$$q_t(1) = \ln \left(\frac{p_{t,1}}{p_{t,0}} \right) \text{ and } q_t(0) = \ln \left(\frac{1 - p_{t,1}}{1 - p_{t,0}} \right).$$

The log-likelihood ratio (A.4) can also be written as follows:

$$\ln \left\{ \frac{L_1(U(n), p_1)}{L_0(U(n), p_0)} \right\} = \sum_{t=1}^{n} a_t(0/1) s(y_t) + b(n),$$

where

$$a_t(0/1) = q_t(1) - q_t(0) \text{ and } b(n) = \sum_{t=1}^{n} q_t(0).$$

Using Neyman-Pearson lemma [see e.g. Lehmann (1959), page 65], the best test of H_0 against H_1 rejects H_0 when

$$\sum_{t=1}^{n} \ln \left[\frac{p_{t,1}(1 - p_{t,0})}{p_{t,0}(1 - p_{t,1})} \right] s(y_t) + b(n) > c.$$

or equivalently when

$$\sum_{t=1}^{n} \ln \left[\frac{p_{t,1}(1 - p_{t,0})}{p_{t,0}(1 - p_{t,1})} \right] s(y_t) > c_1 \equiv c - b(n).$$

i.e. Theorem **2.2**. □

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

$$\phi_{S_n^*}(u) = E_X\left[\exp(iu\ S_n^*(\beta_1))\right] = E_X\left[\prod_{t=1}^{n}\exp\left(iu\ \ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t\le-\beta_1' x_t|X]}-1}\right]s(y_t)\right)\right],\ \forall\, u\in\mathbb{R}$$

where a complex number $i=\sqrt{-1}$. Since y_t, for $t=1,...,n$, are independent

$$\begin{aligned}
\phi_{S_n^*}(u) &= \prod_{t=1}^{n}E_X\left[\exp\left(iu\ \ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t\le-\beta_1' x_t|X]}-1}\right]s(y_t)\right)\right]\\
&= \prod_{t=1}^{n}\sum_{j=0}^{1}\mathsf{P}\left(s(y_t)=j\mid X\right)\exp\left(iu\ \ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t\le-\beta_1' x_t|X]}-1}\right]j\right)\\
&= \left(\frac{1}{2}\right)^n\prod_{t=1}^{n}\left[1+\exp\left(iu\ \ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t\le-\beta_1' x_t|X]}-1}\right]\right)\right].
\end{aligned}$$

According to Gil-Pelaez (1951), the conditional distribution function of $S_n^*(\beta_1)$ evaluated at c_1, for $c_1\in\mathbb{R}$, is given by:

$$\mathsf{P}(S_n^*(\beta_1)\le c_1\mid X) = \frac{1}{2}-\frac{1}{\pi}\int_0^{\infty}\frac{I(u)}{u}du, \tag{A.5}$$

where

$$I(u) = \left(\frac{1}{2}\right)^n \mathrm{Im}\left\{\prod_{t=1}^{n}\left[\exp\left(-iu\frac{c_1}{n}\right)+\exp\left(iu(\ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t\le-\beta_1' x_t|X]}-1}\right]-\frac{c_1}{n}\right)\right]\right\}.$$

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

$$\Pi(\beta,\beta_1) = \mathsf{P}\left[S_n^*(\beta_1)>c_1(\beta_1)\right] = 1-\mathsf{P}\left[S_n^*(\beta_1)\le c_1(\beta_1)\right] = \frac{1}{2}+\frac{1}{\pi}\int_0^{\infty}\frac{I(u)}{u}du,$$

where

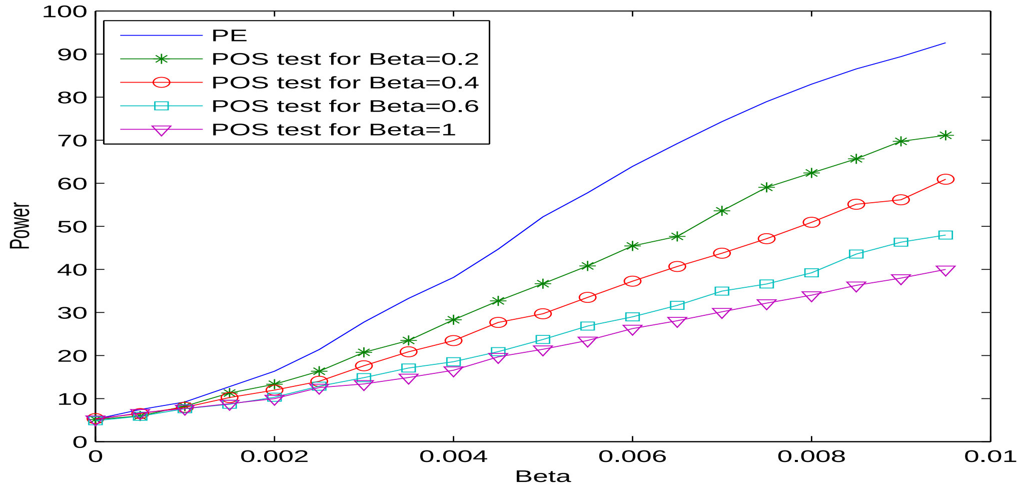
$$I(u) = \left(\frac{1}{2}\right)^n \mathrm{Im}\left\{\prod_{t=1}^{n}\left[\exp\left(-iu\frac{c_1}{n}\right)+\exp\left(iu(\ln\left[\frac{1}{\frac{1}{1-\mathsf{P}[\varepsilon_t\le-\beta_1' x_t|X]}-1}\right]-\frac{c_1}{n}\right)\right]\right\}.$$

i.e. Theorem **4.1**. □

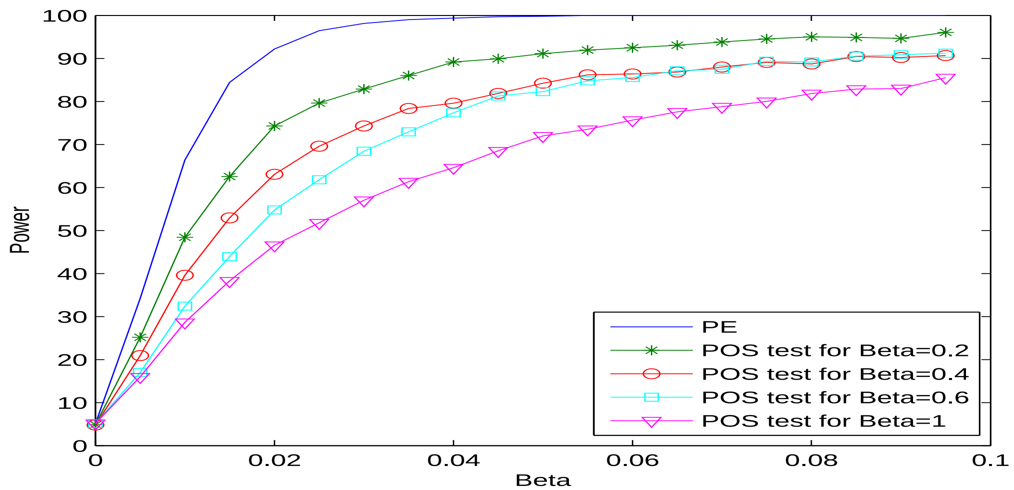
B. Appendix: Simulation results

Figure 1. **Power comparison under different alternatives.**

Panel A. Power comparison under different alternatives (Normal case)



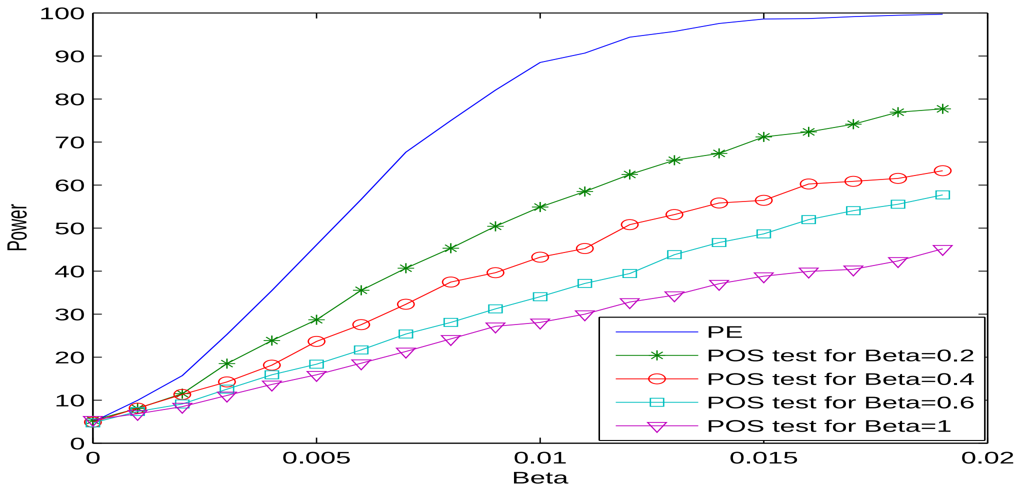
Panel B. Power comparison under different alternatives (Cauchy case)



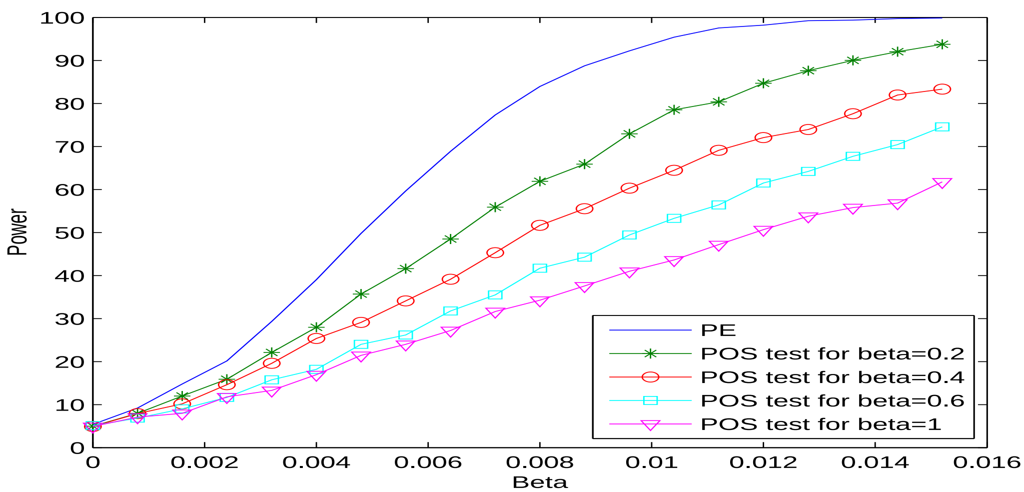
Note: This figure compares the power of POS test under different alternatives. Panel A corresponds to the case where the error term ε_t in the model (6.1) is homoskedastic and normally distributed and Panel B corresponds to the case where this error follows Cauchy distribution.

Figure 2. **(Continued). Power comparison under different alternatives**.

Panel A. Power comparison under different alternatives (Mixture case)



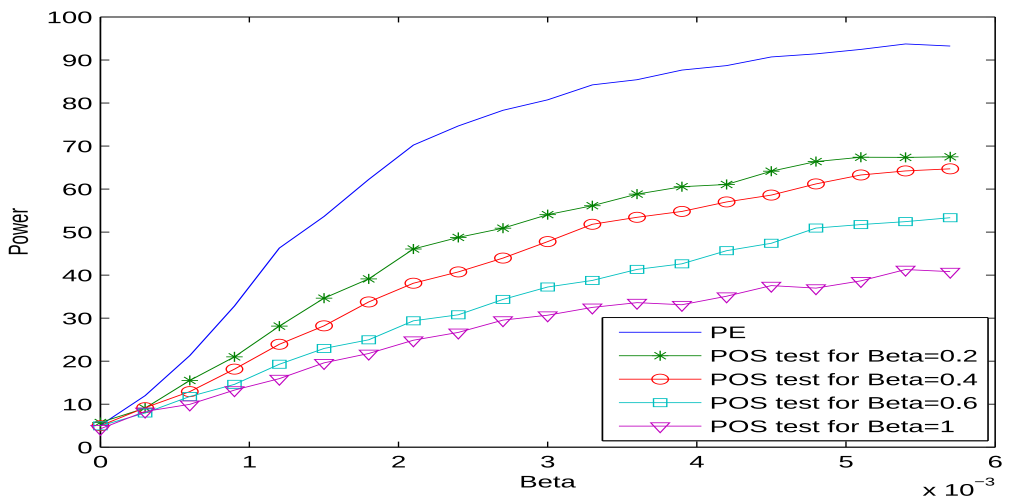
Panel B. Power comparison under different alternatives (Break in variance case)



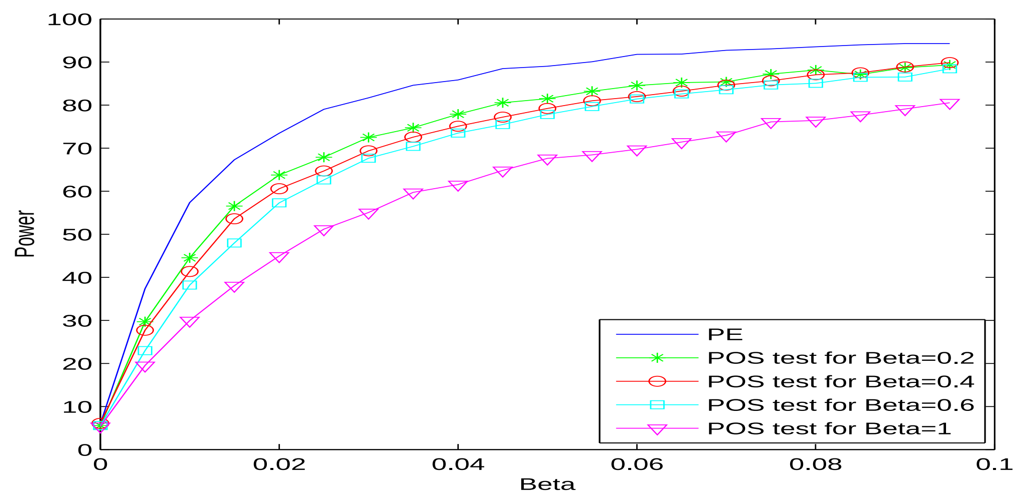
Note: This figure compares the power of POS test under different alternatives. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a mixture of normal and Cauchy distributions and Panel B corresponds to the case where this error follows normal distribution with break in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

Figure 3. **(Continued). Power comparison under different alternatives.**

Panel A. Power comparison under different alternatives (GARCH$(1,\ 1)$ plus jump case)



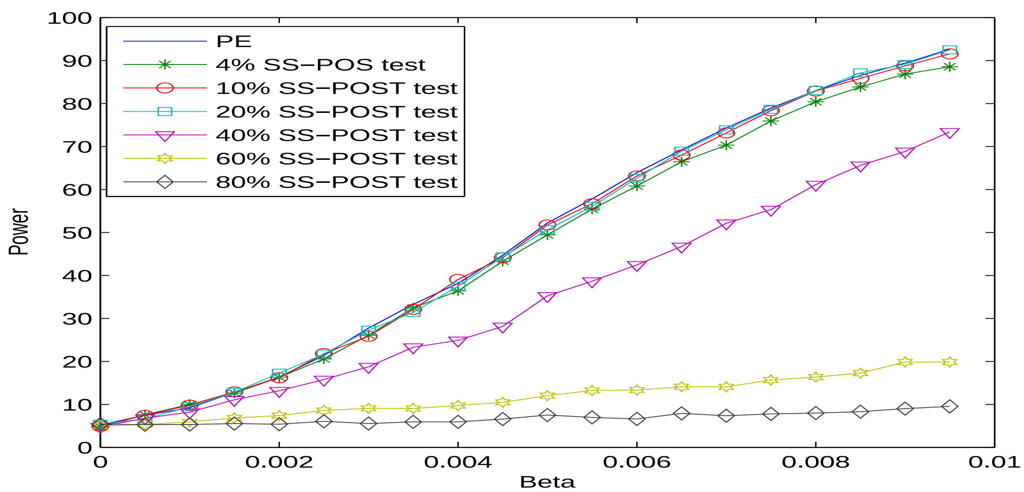
Panel B. Power comparison under different alternatives (Non stationary GARCH$(1,\ 1)$ case)



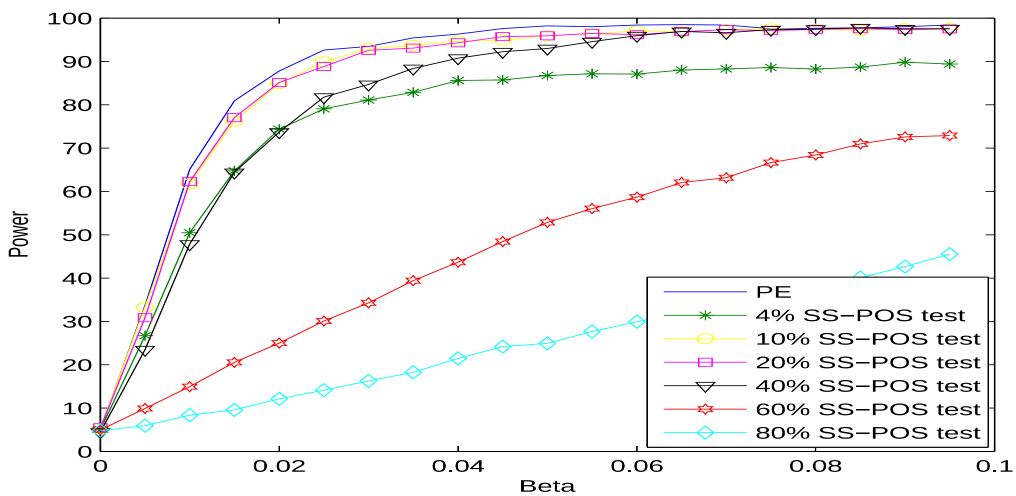
Note: This figure compares the power of POS test under different alternatives. Panel A corresponds to the case where the error term ε_t in the model (6.1) follows normal distribution with GARCH$(1,\ 1)$ plus jump variance and Panel B corresponds to the case where this error follows normal distribution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

Figure 4. **Power comparison using different split-samples.**

Panel A. Power comparison using different split-samples (Normal case)



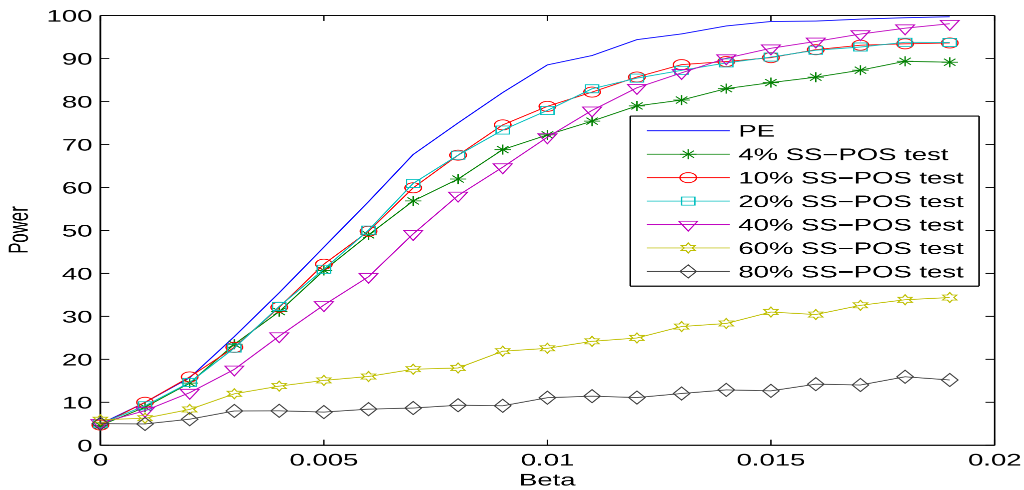
Panel B. Power comparison using different split-samples (Cauchy case)



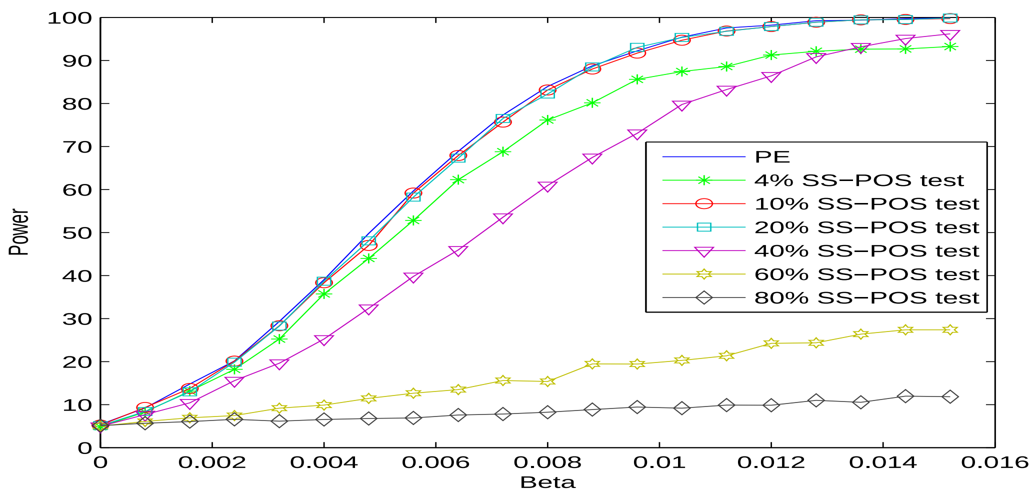
Note: This figure compares the power of POS test using different split-samples (4%, 10%, 20%, 40%, 60%, and 80%). Panel A corresponds to the case where the error term ε_t in the model (6.1) is homoskedastic and normally distributed and Panel B corresponds to the case where this error follows Cauchy distribution.

Figure 5. **(Continued). Power comparison using different split-samples.**

Panel A. Power comparison using different split-samples (Mixture case)



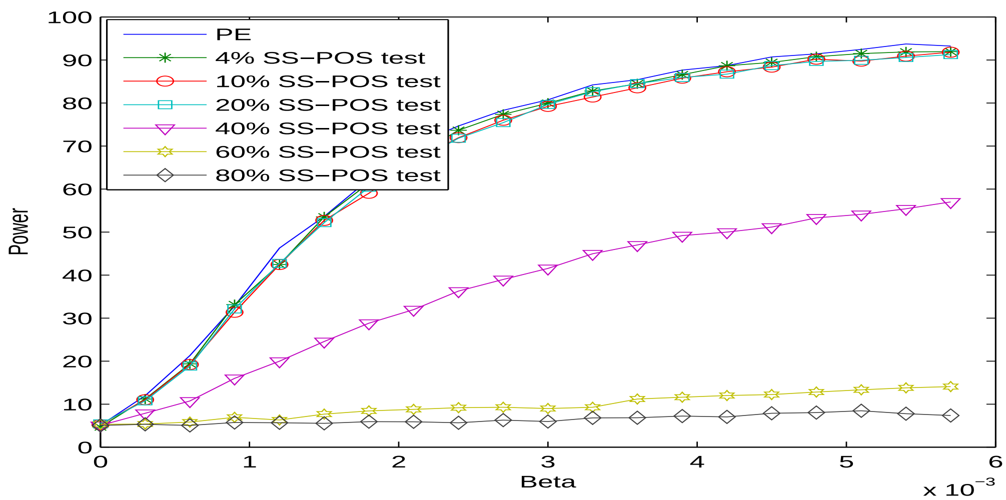
Panel B. Power comparison using different split-samples (Break in variance case)



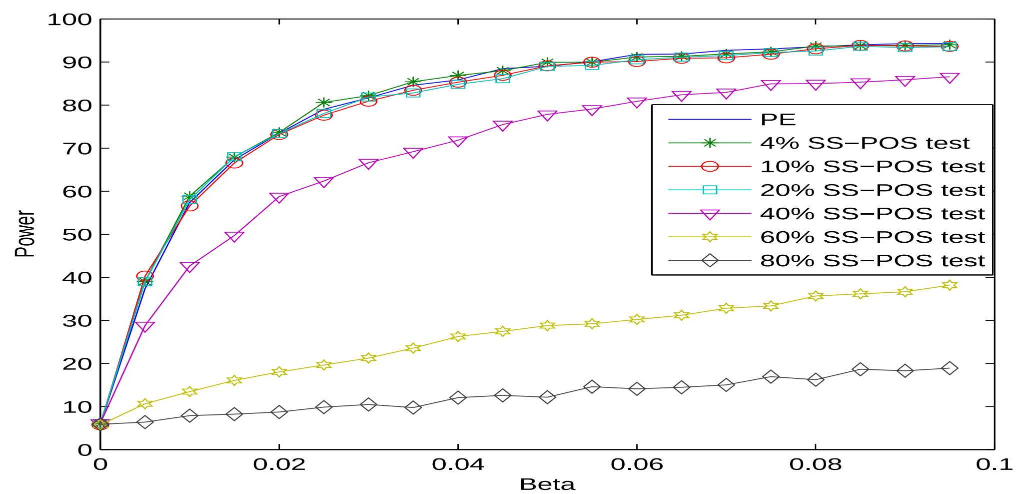
Note: This figure compares the power of POS test using different split-samples (4%, 10%, 20%, 40%, 60%, and 80%). Panel A corresponds to the case where the error term ε_t in the model (6.1) follows a mixture of normal and Cauchy distributions and Panel B corresponds to the case where this error follows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

Figure 6. **(Continued). Power comparison using different split-samples.**

Panel A. Power comparison using different split-samples (GARCH$(1,\ 1)$ plus jump case)



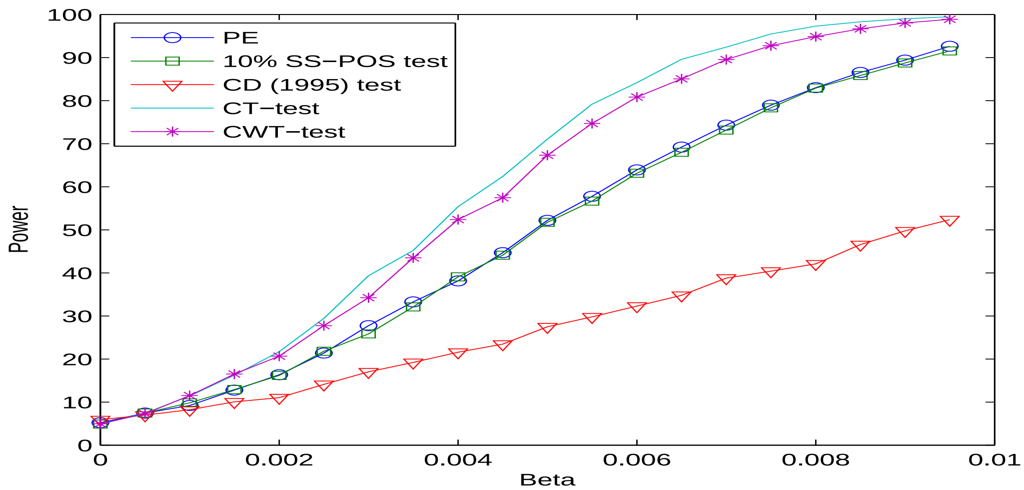
Panel B. Power comparison using different split-samples (Non stationary GARCH$(1,\ 1)$ case)



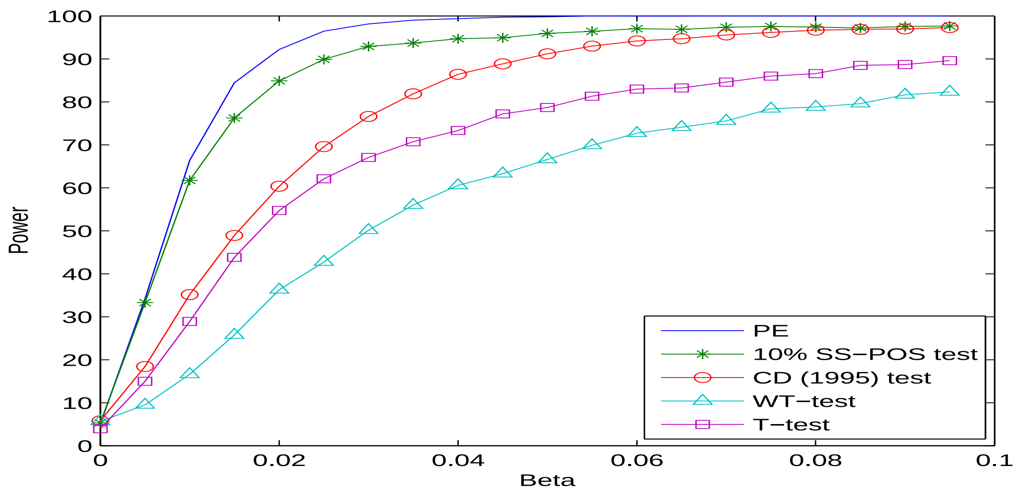
Note: This figure compares the power of POS test using different split-samples (4%, 10%, 20%, 40%, 60%, and 80%). Panel A corresponds to the case where the error term ε_t in the model (6.1) follows normal distribution with GARCH$(1,\ 1)$ plus jump variance and Panel B corresponds to the case where this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

Figure 7. **Power comparison using different tests.**

Panel A. Power comparison using different tests (Normal case)



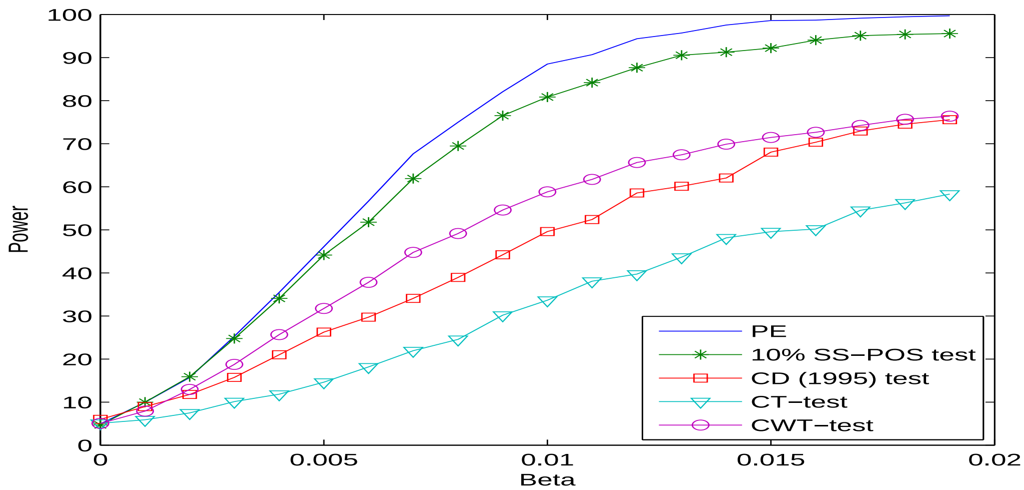
Panel B. Power comparison using different tests (Cauchy case)



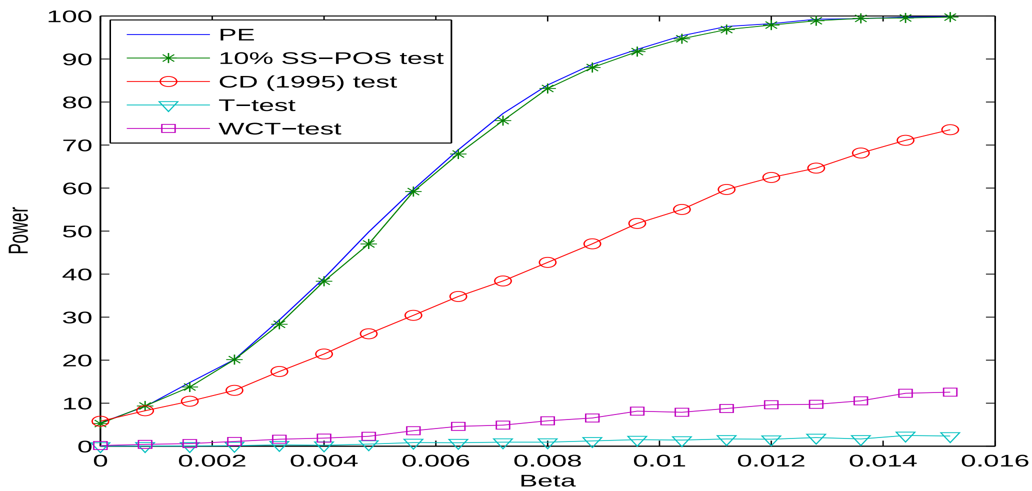
Note: This figure compares the power envelope to the power curves of 10% split-sample POS test [10% SS-POS test], T-test (or CT-test), sign-based test proposed by Campbell and Dufour (1995) [CD (1995) test], and the T-test based on White's (1980) variance correction [WT-test or CWT-test]. Panel A corresponds 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 this error follows Cauchy distribution.

Figure 8. **(Continued). Power comparison using different tests.**

Panel A. Power comparison using different tests (Mixture case)



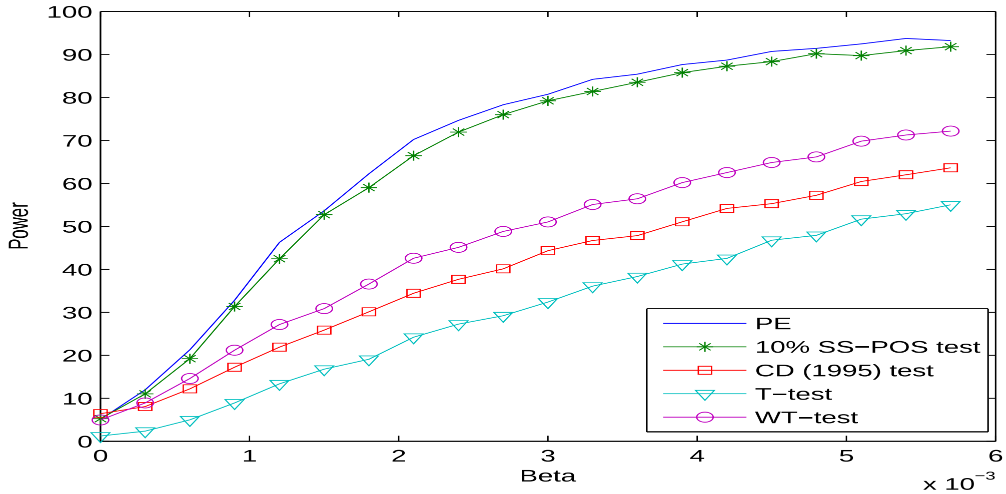
Panel B. Power comparison using different tests (Break in variance case)



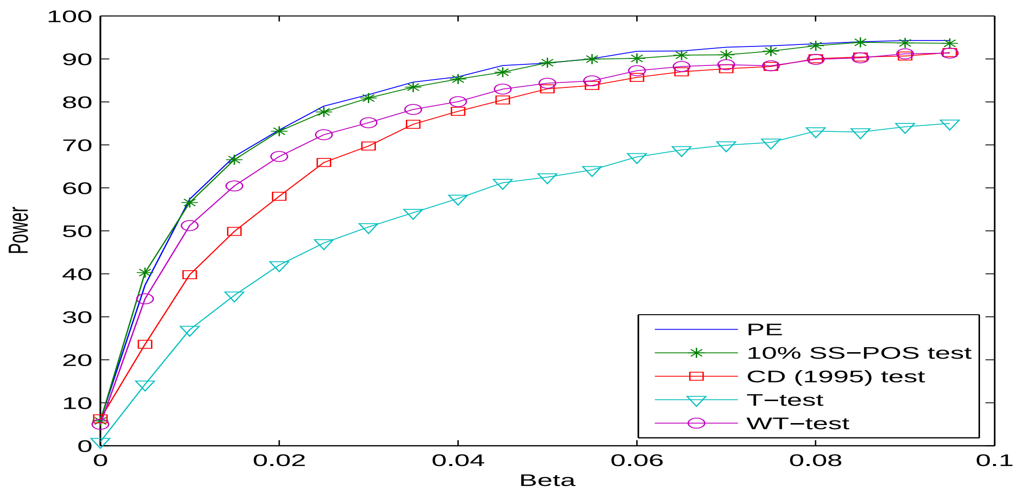
Note: This figure compares the power envelope to the power curves of 10% split-sample POS test [10% SS-POS test], T-test (or CT-test), sign-based test proposed by Campbell and Dufour (1995) [CD (1995) test], and the T-test based on White's (1980) variance correction [WT-test or CWT-test]. Panel A corresponds 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 this error follows normal distribution with break in variance.

Figure 9. **(Continued). Power comparison using different tests.**

Panel A. Power comparison using different tests (GARCH$(1,\ 1)$ plus jump case)



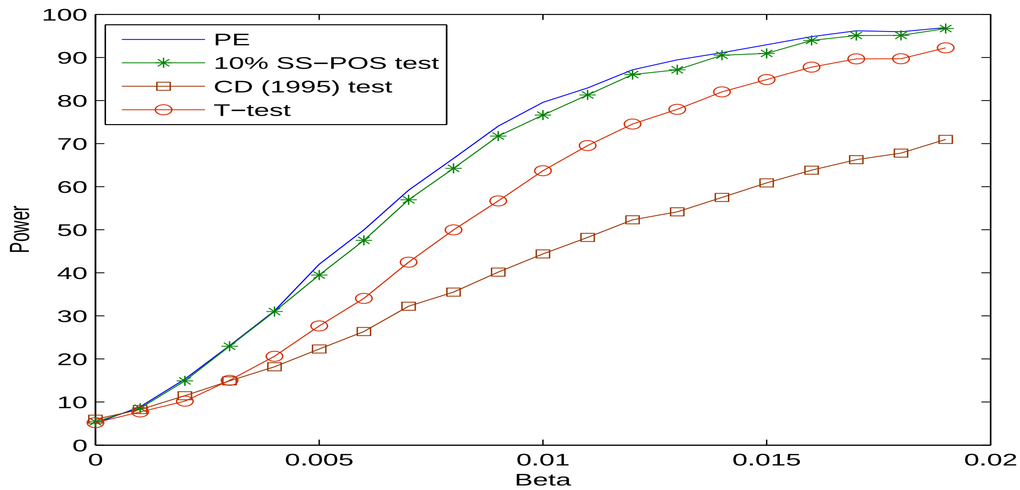
Panel B. Power comparison using different tests (Non stationary GARCH$(1,\ 1)$ case)



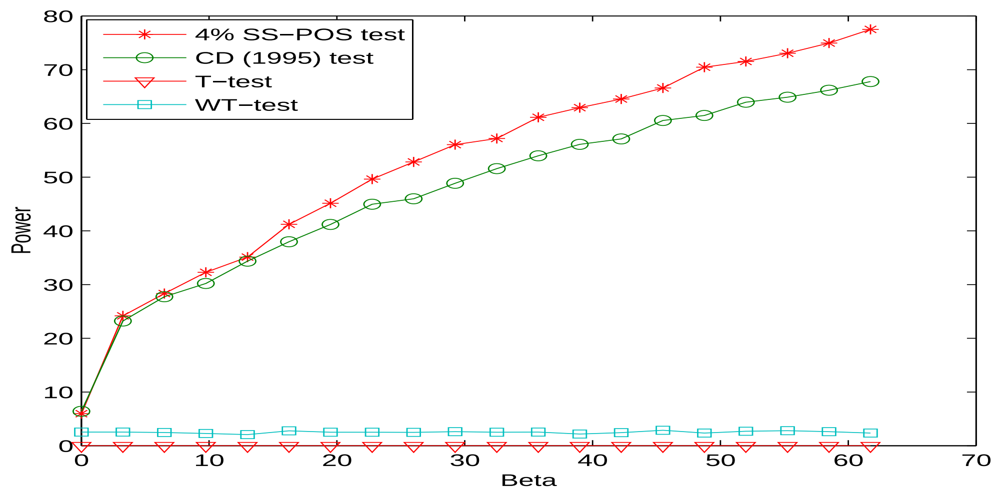
Note: This figure compares the power envelope to the power curves of 10% split-sample POS test [10% SS-POS test], T-test (or CT-test), sign-based test proposed by Campbell and Dufour (1995) [CD (1995) test], and the T-test based on White's (1980) variance correction [WT-test or CWT-test]. Panel A corresponds 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 this error follows normal distribution with non stationary GARCH$(1,\ 1)$ variance.

Figure 10. **(Continued). Power comparison using different tests.**

Panel A. Power comparison using different tests (Student case)



Panel B. Power comparison using different tests (Exponential variance case)



Note: This figure compares the power envelope to the power curves of 10% split-sample POS test [10% SS-POS test], T-test (or CT-test), sign-based test proposed by Campbell and Dufour (1995) [CD (1995) test], and 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 student distribution with degree of freedom 2 and Panel B corresponds to the case where this error follows normal distribution with exponential variance

Table 1. Power comparison: Normal distribution

		POS test		SS-POS test				Other tests			
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD (1995) test	T-test	WT-test	CWT-test
0	5.2	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	6.5	7.58	7.44	6.62	6.78	6.96	7.42	10.7	7.3
0.001	9.2	8.24	7.96	9.98	9.82	9.48	8.2	8.24	11.4	15.4	11.5
0.0015	12.78	11.28	10.24	12.6	12.9	12.76	11.04	10.06	16.24	20.08	16.5
0.002	16.34	13.34	11.96	16.28	16.18	17.26	13.18	11.02	21.7	26.78	20.68
0.0025	21.38	16.36	14.02	20.56	21.8	21.7	15.76	14.12	29.42	34.42	27.74
0.003	27.74	20.74	17.62	26.08	25.84	27.26	18.74	17.02	39.32	41.2	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
0.004	38.14	28.28	23.46	36.4	39.08	37.52	24.88	21.56	55.36	58.52	52.38
0.0045	44.68	32.68	27.68	43.28	44.1	44.3	28.14	23.46	62.38	66.96	57.44
0.005	52.2	36.68	29.7	49.44	51.74	50.6	35.24	27.5	71.04	73.16	67.32
0.0055	57.76	40.78	33.5	55.42	56.68	56.06	38.64	29.8	79.16	79.92	74.7
0.006	63.92	45.44	37.26	60.78	63.12	62.62	42.44	32.3	84.18	85.7	80.84
0.0065	69.22	47.66	40.68	66.44	68	68.9	46.74	34.78	89.58	89.74	85.06

Note: This table shows the power envelope of POS test (PE) and size and power of point-optimal sign test under different alternative hypotheses (POS test), point-optimal sign test using different split-sample sizes (SS-POS test), sign-based test of Campbell and Dufour (1995) [CD (1995) test], T-test, T-test based on White's (1980) variance correction (WT-test), and WT-test after size correction (CWT-test). These results correspond to the case where the error term ε_t in the model (6.1) is homoskedastic and normally distributed.

Table 2. Power comparison: Cauchy distribution

		POS test		SS-POS test				Other tests		
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD (1995) test	T-test	WT-test
0	5.1	4.88	4.8	5.02	5.3	5.48	4.46	5.78	5.68	3.94
0.005	34.22	25.18	20.94	26.72	33.3	30.86	23.48	18.44	9.5	15
0.01	66.38	48.42	39.58	50.46	61.74	62.28	47.86	35.16	16.6	28.92
0.015	84.44	62.56	52.94	64.74	76.24	77.02	64.38	48.9	25.76	43.82
0.02	92.2	74.3	63.08	74.36	84.9	85.14	73.7	60.36	36.28	54.72
0.025	96.44	79.62	69.6	79.06	89.88	88.82	81.78	69.58	42.74	62.08
0.03	98.12	82.86	74.3	81.08	92.92	92.58	84.7	76.6	50.14	67.06
0.035	99	86.02	78.36	82.86	93.7	93.1	88.38	81.88	56	70.72
0.04	99.36	89.16	79.6	85.62	94.7	94.3	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	63.3	77.18
0.05	99.8	91.12	84.24	86.76	95.92	95.92	93	91.18	66.6	78.7
0.055	99.98	91.94	86.2	87.14	96.42	96.48	94.56	92.98	69.88	81.3
0.06	99.94	92.5	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	96.9	96.92	94.68	74.1	83.22

Note: This table shows the power envelope of POS test (PE) and size and power of point-optimal sign test under different alternative hypotheses (POS test), point-optimal sign test using different split-sample sizes (SS-POS test), sign-based test of Campbell and Dufour (1995) [CD (1995) test], T-test, and T-test based on White's (1980) variance correction (WT-test). These results correspond to the case where the error term ε_t in the model (6.1) follows Cauchy distribution.

Table 3. Power comparison: Mixture distribution

		POS test		SS-POS test				Other tests				
β	PE	$\beta_1 = 0.2$	$\beta_1 = 0.4$	4%	10%	20%	40%	CD(1995) test	T-test	WT-test	CT-test	CWT-test
0	4.96	5.3	4.9	4.58	4.7	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	5.9	7.92
0.002	15.7	11.52	11.3	14.46	15.9	14.6	12.24	11.76	13.98	18.88	7.5	12.94
0.003	25.26	18.48	14.24	22	24.76	24.6	19.64	15.72	16.9	25.76	10.1	18.74
0.004	35.46	23.84	18.12	29.6	34.08	34.28	27.36	21	20.68	31.76	11.82	25.68
0.005	46.08	28.7	23.66	39.16	44.14	42.96	34.6	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	32.3	55.34	61.9	61.84	51.16	34.06	33	51.22	21.92	44.76
0.008	75	45.32	37.46	60.44	69.48	69.5	60.1	38.96	36.62	56.7	24.56	49.14
0.009	82.06	50.4	39.64	67.28	76.52	75.32	66.68	44.22	40.16	60.5	30.18	54.6
0.01	88.48	54.9	43.24	70.7	80.84	79.9	73.68	49.58	45.86	63.74	33.64	58.8
0.011	90.68	58.48	45.24	73.92	84.16	84.94	79.92	52.4	48.6	66.9	38.06	61.7
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	95.7	65.76	53.12	78.82	90.54	89.22	88.64	60.1	55.26	72.16	43.66	67.42

Note: This table shows the power envelope of POS test (PE) and size and power of point-optimal sign test under different alternative hypotheses (POS test), point-optimal sign test using different split-sample sizes (SS-POS test), sign-based test of Campbell and Dufour (1995) [CD (1995) test], T-test, T-test based on White's (1980) variance correction (WT-test), T-test after size correction (CT-test), and WT-test after size correction (CWT-test). These results correspond to the case where the error term ε_t in the model (6.1) follows a mixture of normal and Cauchy distributions.

Table 4. Power comparison: 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%	CD (1995) test	T-test	WT-test
0	5.4	4.98	4.92	4.84	5.24	5.1	4.96	5.78	0.01	0.16
0.0008	9.22	7.96	7.9	8.28	9.32	8.38	7.68	8.24	0.04	0.42
0.0016	14.78	12	10.18	13.12	13.76	12.98	10.42	10.44	0.06	0.6
0.0024	20.16	15.88	14.62	18.2	20.12	19.86	15.58	12.98	0.12	1.08
0.0032	29.32	22.12	19.6	25.24	28.34	28.26	19.64	17.34	0.3	1.62
0.004	39.04	27.96	25.38	35.72	38.32	38.68	25.24	21.4	0.22	1.86
0.0048	49.78	35.7	29.12	43.98	47	48.06	32.38	26.12	0.46	2.3
0.0056	59.66	41.62	34.12	52.82	59.16	58.24	39.78	30.42	0.84	3.6
0.0064	68.88	48.5	39.14	62.3	67.9	67.28	45.96	34.78	0.78	4.58
0.0072	77.32	55.9	45.3	68.78	75.66	76.5	53.54	38.38	0.94	4.88
0.008	83.96	61.9	51.68	76.14	83.14	82.2	60.92	42.72	0.94	5.88
0.0088	88.76	65.9	55.52	80.14	88	88.5	67.46	47.04	1.22	6.54
0.0096	92.22	72.94	60.32	85.6	91.7	93.02	73.06	51.76	1.5	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: This table shows the power envelope of POS test (PE) and size and power of point-optimal sign test under different alternative hypotheses (POS test), point-optimal sign test using different split-sample sizes (SS-POS test), sign-based test of Campbell and Dufour (1995) [CD (1995) test], T-test, and T-test based on White's (1980) variance correction (WT-test). These results correspond to the case where the error term ε_t in the model (6.1) follows a normal distribution with break in variance.

Table 5. Power comparison: 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%	CD (1995) test	T-test	WT-test
0	5.07	5.74	4.98	4.7	5.24	5.4	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	15.5	12.9	19.38	19.2	18.84	10.74	12.18	5	14.6
0.0009	32.8	21	18.14	33.12	31.34	32.12	15.98	17.24	8.9	21.2
0.0012	46.28	28.14	23.9	42.46	42.46	42.72	19.98	21.9	13.36	27.16
0.0015	53.62	34.62	28.2	53.52	52.7	52.2	24.56	25.86	16.76	30.86
0.0018	62.24	39.1	33.74	61.36	59	60.4	28.8	30.12	19.06	36.58
0.0021	70.22	46.06	38.1	67.52	66.44	66.14	31.96	34.44	24.2	42.58
0.0024	74.66	48.74	40.72	73.66	71.94	71.8	36.28	37.68	27.26	45.1
0.0027	78.28	50.88	43.94	77.36	75.98	75.44	38.98	40.12	29.22	48.82
0.003	80.72	54.04	47.76	79.96	79.22	79.66	41.54	44.32	32.4	51.02
0.0033	84.22	56.12	51.8	82.76	81.38	82.62	44.96	46.72	36.1	55.08
0.0036	85.42	58.82	53.44	84.46	83.52	84.5	47	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

Note: This table shows the power envelope of POS test (PE) and size and power of point-optimal sign test under different alternative hypotheses (POS test), point-optimal sign test using different split-sample sizes (SS-POS test), sign-based test of Campbell and Dufour (1995) [CD (1995) test], T-test, and T-test based on White's (1980) variance correction (WT-test). These results correspond 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.

Table 6. Power comparison: 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%	CD (1995) test	T-test	WT-test
0	5.95	5.58	6.08	6.02	5.76	6.04	6.16	6.26	0.94	5
0.005	37.34	29.68	27.72	39.04	40.28	39	28.78	23.58	14.26	34.18
0.01	57.36	44.54	41.36	58.86	56.58	58.04	42.64	39.78	27	51.22
0.015	67.3	56.54	53.58	67.92	66.54	68	49.7	49.84	35	60.44
0.02	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	64.7	80.6	77.64	78.04	62.34	65.88	47.16	72.36
0.03	81.66	72.5	69.38	82.18	80.88	81.88	66.6	69.72	50.9	75.14
0.035	84.58	74.72	72.56	85.4	83.42	82.8	69.18	74.78	54.22	78.24
0.04	85.82	77.86	75.08	86.86	85.3	84.82	71.84	77.82	57.52	80.04
0.045	88.46	80.52	77.2	87.98	86.9	86.12	75.46	80.44	61.18	82.96
0.05	89.02	81.48	79.22	89.92	89.1	88.98	77.84	83.04	62.48	84.34
0.055	90.04	83.2	81	89.94	89.94	89.22	79.08	83.82	64.16	84.88
0.06	91.76	84.52	81.96	91.14	90.1	90.5	80.86	85.7	67.2	87.26
0.065	91.82	85.22	83.22	91.3	90.86	91.12	82.38	87	68.8	88.22

Note: This table shows the power envelope of POS test (PE) and size and power of point-optimal sign test under different alternative hypotheses (POS test), point-optimal sign test using different split-sample sizes (SS-POS test), sign-based test of Campbell and Dufour (1995) [CD (1995) test], T-test, and T-test based on White's (1980) variance correction (WT-test). These results correspond to the case where the error term ε_t in the model (6.1) follows a normal distribution with non stationary GARCH$(1,\ 1)$ variance.

Table 7. True weights versus normal weights: Cauchy distribution

		$SS-POS\ test$ using true weights		$SS-POS\ test$ using normal weights	
β	PE	10%	20%	10%	20%
0	5.1	5.16	5.16	5.3	5.48
0.005	34.22	33.58	31.18	33.3	30.86
0.01	66.38	61.94	62.47	61.74	62.28
0.015	84.44	80.32	80.32	76.24	77.02
0.02	92.2	89.76	89.76	84.9	85.14
0.025	96.44	95.22	95.22	89.88	88.82
0.03	98.12	96.98	96.98	92.92	92.58
0.035	99	98.26	98.26	93.7	93.1
0.04	99.36	99.14	99.14	94.7	94.3
0.045	99.68	99.3	99.3	94.92	95.74
0.05	99.8	99.44	99.44	95.92	95.92
0.055	99.98	99.7	99.7	96.42	96.48
0.06	99.94	99.82	99.82	97.02	96.18
0.065	99.94	99.9	99.9	96.86	96.9

Note: This table summarizes the results of the comparison between the power of 10% split-sample POS test calculated using the true weights $a_t(0/1)$ with the power of 10% split-sample POS test calculated using normal weights. The true weights correspond to the case where the error term ε_t follows Cauchy distribution. The term SS-POS test corresponds to split-sample point-optimal sign test.

Table 8. True weights versus normal weights: Mixture distribution

		$SS-POS$ test with true weights		$SS-POS$ test with normal weights	
β	PE	10%	20%	10%	20%
0	4.96	4.74	5.26	4.7	5.02
0.001	9.96	8.96	9.08	9.98	9.16
0.002	15.7	14.34	16.7	15.9	14.6
0.003	25.26	24.84	24.67	24.76	24.6
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	61.9	61.84
0.008	75	71.66	70.14	69.48	69.5
0.009	82.06	79.24	79.54	76.52	75.32
0.01	88.48	85.52	84.34	80.84	79.9
0.011	90.68	88.8	89.22	84.16	84.94
0.012	94.38	92.06	91.5	87.66	87.42
0.013	95.7	94.32	94.62	90.54	89.22

Note: This table summarizes the results of the comparison between the power of 10% split-sample POS test calculated using the true weights $a_t(0/1)$ with the power of 10% split-sample POS test calculated using normal weights. The true weights correspond to the case where the error term ε_t follows a mixture of normal and Cauchy distributions. The term SS-POS test corresponds to split-sample point-optimal sign test.

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.

Arrow, K. (1960), *Decision Theory and the Choice of a Level of Significance for the T-Test*, in: Olkin, I. et al. (eds.), Contributions to Probability and Statistics: Essays in the Honor of Harold HotellingIn, Stanford: Stanford University Press., pp. 70–78.

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

Bohman, H. (1972), 'From characteristic function to distribution function via fourier analysis', *Nordisk Tidskr. Informationsbehandling (BIT)* **12**, 279–283.

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. (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.

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, Technical report, CREST and Université de Montréal.

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.-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. 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. 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. (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.

Friedman, B. M. and Laibson, D. I. (1989), 'Economic implications of extraordinary movements in stock prices', *Brookings Papers on Economic Activity* **2**, 137–189.

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

Hotta, L. K. and Tsay, R. S. (1998), 'Outliers in GARCH processes', *unpublished manuscript Graduate School of Business University of Chicago.* .

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. (1958), 'Significance level and power', *Annals of Mathematical Statistics* **29**(4), 1167–1176.

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

Lehmann, E. L. and Romano, J. P. (2005), *Testing Statistical Hypothesis*, third edn, Springer Texts in Statistics. Springer-Verlag, New York.

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

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.

Sanathanan, L. (1974), 'Critical power function and decision making', *Journal of the American Statistical Association* **69**, 398–402.

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.