Consequences and detection of invalid exogeneity conditions

Niemczyk, J.

Chapter 4

Evaluating the performance of tests of overidentifying restrictions

4.1 Introduction

In linear regression models with endogenous explanatory variables a researcher utilizes additional variables, so-called instruments (which can also include some of the regressors), that have known correlation with the regression error term, and provide necessary information for the consistent estimation of the unknown parameter(s). So far, the most common estimator used to tackle the linear model is the Instrumental Variables estimator which can be viewed as a special case of the Generalized Method of Moments estimator (GMM) for possibly non linear models with non i.i.d. disturbances, introduced by Hansen (1982) in his seminal paper. In that case the validity of the exploited moment conditions (restrictions), provided by the instruments, is investigated via the Hansen (1982) J statistic.

More recently, based on the concept of Empirical Likelihood introduced by Owen (1991), an alternative method was proposed by Qin and Lawless (1994), Imbens (1997). Empirical Likelihood (EL) finds an estimator together with 'empirical probabilities' that maximize the 'empirical likelihood function' such that the moment conditions are exactly satisfied in the sample (which, in general, is not the case with the GMM estimator for which the implied 'empirical probabilities' are all equal). Here the maximized criterion function (empirical likelihood) provides the natural statistic for testing validity of overidentifying restrictions (via the 'Empirical Likelihood Ratio' test, ELR). Other estimators, such as Exponential Tilting and the Continuous Updating Estimator (CUE), were further proposed. They are all special cases of the so called Generalized Empirical Likelihood (GEL) estimator, see Smith (1997), Newey and Smith (2004) and citations therein. The consistency of the overidentifying restriction tests based on GEL was proven by Smith (1997). Optimality of EL for testing moment conditions was shown by Kitamura (2001). The study of Newey and Smith (2004) suggests that GEL can have better finite sample properties relative to GMM.

In this Chapter, for a simple linear model, we examine the finite sample properties of several procedures for testing overidentifying restrictions via Monte Carlo simulation. We compare several versions of known GMM statistics (that differ with respect to weighting matrices applied) with GEL type Likelihood Ratio tests. We analyze the behavior of the tests when the instruments are either weak or strong. We also examine the incremental version of those overidentifying restriction tests, that is the difference between test statistics of the validity of a set of instruments and of a subset of those instruments. By exploiting the validity of a subset of the instruments this incremental version should lead to a 'local power' improvement, see Hall (2005).

Finite sample properties of the GMM tests can probably be improved by applying bootstrap procedures. We examine an implementation suggested in Hall and Horowitz (1996) and Brown and Newey (2002) and some modifications of those. We also analyze a simple residual type bootstrap.

We find that the Hall and Horowitz (1996) implementations, in terms of size, are working well for large samples and rather strong instruments. However, the residual bootstrap performs much better here, for both small and large samples and under weak or strong instruments. Brown and Newey's (2002) implementation does not perform well in our examples, which is probably due to numerical problems.

In Section 2, we present the GMM and GEL type tests for overidentifying restrictions. In Section 3, we describe bootstrap procedures for correcting GMM type test statistics. In Section 4 we illustrate the size and power of those tests and the performance of bootstrap procedures for a simple linear model.

4.2 Test statistics for overidentifying restrictions

For some stationary data vector $\mathcal{X}_i$, $i = 1, \ldots, n$, from $\mathcal{X} = [\mathcal{X}_1, \ldots, \mathcal{X}_n]'$, where n is the sample size, we denote a particular $l \times 1$ vector function of the data by $g_i(\theta) \equiv g(\mathcal{X}_i, \theta)$ and the corresponding sample moment function by $\bar{g}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} g_i(\theta)$, where θ is a $p \times 1$ vector containing all the parameters. We assume that under the true but unknown data generating process $\mathrm{E}g_i(\theta) = g(\theta)$ and that $\bar{g}_n(\theta) \xrightarrow{p} g(\theta)$ for every $\theta \in \Theta \subseteq \mathbb{R}^p$. We aim to estimate a unique θ_0 for which $g(\theta_0) = 0$.

A popular estimation procedure for estimating θ_0 is the Hansen (1982) GMM method, which minimizes a particular quadratic form of the sample moment function. Although having attractive asymptotic properties, GMM can perform poorly in finite samples. Especially when the parameter is weakly identified (when identifying conditions are close to being violated), see Andrews and Stock (2007) and references therein.

Alternatives to Hansen's (1982) GMM estimator and the test statistics for overidentifying restrictions include: empirical likelihood (which finds an estimator that maximizes the likelihood function of the data subject to the moment restrictions being satisfied in the sample), exponential tilting and the continuous updating estimator, see Newey and Smith (2004). These are members of so called generalized empirical likelihood estimators (GEL).

Below we describe those procedures in some more detail.

4.2.1 GMM statistics

The GMM estimator is

$$\tilde{\theta} \equiv \underset{\theta \in \Theta}{\operatorname{argmin}}\ \bar{g}_n(\theta)' W(\mathcal{X}) \bar{g}_n(\theta), \tag{4.1}$$

where $W(\mathcal{X}) = O_p(1)$ is a $l \times l$ positive semi-definite weighting matrix. The efficient GMM estimator is obtained in one, two or several "steps". In the first step we obtain (4.1) using some initial $W(\mathcal{X})$ (for instance the identity matrix, but in particular cases an optimal weighting matrix can be derived analytically), in the second step we re-compute (4.1)

$$\hat{\theta} \equiv \underset{\theta \in \Theta}{\text{argmin}}\ \bar{g}_n(\theta)'\hat{\Omega}^{-1}(\tilde{\theta})\bar{g}_n(\theta), \tag{4.2}$$

with $W(\mathcal{X}) = \hat{\Omega}^{-1}(\tilde{\theta})$ the inverse of a consistent estimator of the asymptotic variance of $\sqrt{n}\bar{g}_n(\theta_0)$. The third stage estimator, $\hat{\theta}_3$, would be based on $\hat{\Omega}^{-1}(\hat{\theta})$. For the theoretical derivation of the consistency and asymptotic normality of the GMM estimator see Hansen (1982).

As recommended by Andrews (1999) and further justified by Hall (2000) (because it may lead to local power improvement), we shall examine the following 'adapted' form of the covariance estimator,

$$\hat{\Omega}_a(\tilde{\theta}) = \frac{1}{n}\sum_{i=1}^{n} g_i(\tilde{\theta})g_i(\tilde{\theta})' - \bar{g}_n(\tilde{\theta})\bar{g}_n(\tilde{\theta})', \tag{4.3}$$

next to the 'standard' estimator of the covariance matrix

$$\hat{\Omega}_s(\tilde{\theta}) = \frac{1}{n}\sum_{i=1}^{n} g_i(\tilde{\theta})g_i(\tilde{\theta})'. \tag{4.4}$$

Hansen's two-step test statistic for overidentifying restrictions is

$$J_n \equiv n\bar{g}_n(\hat{\theta})'\hat{\Omega}^{-1}(\hat{\theta})\bar{g}_n(\hat{\theta}). \tag{4.5}$$

Using (4.3) or (4.4), this specializes to

$$J_n^a = n\bar{g}_n(\hat{\theta}_a)'\hat{\Omega}_a^{-1}(\hat{\theta}_a)\bar{g}_n(\hat{\theta}_a) \tag{4.6}$$

or

$$J_n^s = n\bar{g}_n(\hat{\theta}_s)'\hat{\Omega}_s^{-1}(\hat{\theta}_s)\bar{g}_n(\hat{\theta}_s), \tag{4.7}$$

with

$$\hat{\theta}_a \equiv \underset{\theta \in \Theta}{\text{argmin}}\ \bar{g}_n(\theta)'\hat{\Omega}_a^{-1}(\tilde{\theta})\bar{g}_n(\theta), \tag{4.8}$$

$$\hat{\theta}_s \equiv \underset{\theta \in \Theta}{\operatorname{argmin}} \; \bar{g}_n(\theta)' \hat{\Omega}_s^{-1}(\tilde{\theta}) \bar{g}_n(\theta). \tag{4.9}$$

Sometimes, for brevity, we will simply write $\hat{\theta}$ for $\hat{\theta}_s$ or $\hat{\theta}_a$, although in general those estimators are not equal. Also note that we used $\hat{\Omega}^{-1}(\hat{\theta})$ in (4.5) and not $\hat{\Omega}^{-1}(\tilde{\theta})$, so in general (4.5) will be different from, but asymptotically equivalent to, the minimized criterion function, $n\bar{g}_n(\hat{\theta})'\hat{\Omega}^{-1}(\tilde{\theta})\bar{g}_n(\hat{\theta})$.

Also the difference between the expressions (4.3) and (4.4), $\bar{g}_n(\tilde{\theta})\bar{g}_n(\tilde{\theta})'$, tends to zero if the population moments are satisfied. Hence, it will not change the limiting null distribution of the test statistic. However, if some population moment conditions are invalid, this factor does not disappear in the limit, and thus can lead to power improvements of the test, see Hall (2000).

When all the moment conditions are valid, i.e. $g(\theta_0) = 0$ for a unique θ_0, then we can test whether $(l-k)$ overidentifying moment restrictions are satisfied. Under appropriate regularity conditions, see Hansen (1982), (4.6) and (4.7) are asymptotically distributed as $\chi^2(l-k)$. The procedures for testing overidentifying restrictions based on (4.5) are consistent, see Andrews (1999).

Linear model

In a linear model we have $\mathcal{X}_i = (y_i, x_i', z_i')$, where z_i is an $l \times 1$ vector of alleged instruments and x_i is a $k \times 1$ vector of regressors (if some of the regressors are 'exogenous' or 'predetermined' then x_i and z_i can share the same elements), $i = 1, \ldots, n$. Let $Z = [z_1, \ldots, z_n]'$ be the $(n \times l)$ matrix of 'instruments', $X = [x_1, \ldots, x_n]'$ the $(n \times k)$ matrix of regressors and $y = [y_1, \ldots, y_n]'$ the $(n \times 1)$ vector of dependent variables. The GMM estimator is based on the following population moment conditions

$$g(\theta_0) = \mathrm{E} z_i (y_i - x_i'\theta_0) = 0.$$

For an initial consistent estimator, $\tilde{\theta}$, let $u(\tilde{\theta}) = y - X\tilde{\theta}$, we then have

$$\bar{g}_n(\tilde{\theta}) = \frac{1}{n}\sum_{i=1}^{n} z_i (y_i - x_i'\tilde{\theta}) = \frac{1}{n} Z' u(\tilde{\theta}), \tag{4.10}$$

$$\hat{\Omega}_a(\tilde{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i'\tilde{\theta})^2 z_i z_i' - \frac{1}{n^2}Z'u(\tilde{\theta})u(\tilde{\theta})'Z, \tag{4.11}$$

and

$$\hat{\Omega}_s(\tilde{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i'\tilde{\theta})^2 z_i z_i'. \tag{4.12}$$

The second stage GMM estimator is

$$\hat{\theta} = (X'Z\hat{\Omega}^{-1}X'Z)^{-1}X'Z\hat{\Omega}^{-1}Z'y \tag{4.13}$$

with $\hat{\Omega}$ being either (4.11) or (4.12). For testing moment conditions we are using (4.6) or (4.7).

The instrumental variables (IV) estimator is

$$\tilde{\theta} = (X'P_Z X)^{-1}X'P_Z y, \tag{4.14}$$

where $P_Z = Z(Z'Z)^{-1}Z'$. It results from the minimization of (4.1) using $W(\mathcal{X}) = [\frac{1}{n}Z'Z]^{-1}$, where under conditional homoscedasticity, i.e. $\mathrm{E}((y_i - x_i'\theta_0)^2|z_i) = \sigma_0^2$, it is an optimal choice. The unconditional covariance matrix of the moment conditions is $\Omega = \sigma^2\Sigma_{Z'Z}$. Using, for the weighting matrix in the second stage, the structure of this matrix we can apply

$$\dot{\Omega}_s(\tilde{\theta}) = \frac{u(\tilde{\theta})'u(\tilde{\theta})}{n}\frac{1}{n}Z'Z \tag{4.15}$$

or

$$\dot{\Omega}_a(\tilde{\theta}) = \frac{u(\tilde{\theta})'u(\tilde{\theta})}{n}\frac{1}{n}Z'Z - \frac{1}{n^2}Z'u(\tilde{\theta})u(\tilde{\theta})'Z, \tag{4.16}$$

instead of (4.12) or (4.11).

However, from the form of (4.13) we can easily see that updating the estimator that uses $\dot{\Omega}_s \equiv \dot{\Omega}_s(\tilde{\theta})$ will not affect the second stage estimator. Hence, when using $\dot{\Omega}_s$, one obtains

$$\dot{\theta}_s = (X'Z\dot{\Omega}_s^{-1}Z'X)^{-1}X'Z\dot{\Omega}_s^{-1}Z'y = \tilde{\theta}.$$

We now show that it is also true when using $\dot{\Omega}_a \equiv \dot{\Omega}_a(\tilde{\theta})$.

Lemma 4.2 *For* $\hat{\Omega} = \dot{\Omega}_a(\tilde{\theta})$ *in* (4.13) *we have*

$$\dot{\theta}_a = \tilde{\theta}. \tag{4.17}$$

Proof. For simplicity, write

$$\dot{\Omega}_a = \dot{\Omega}_s - \tilde{g}\tilde{g}', \tag{4.18}$$

with $\tilde{g} = \frac{1}{n}Z'u(\tilde{\theta})$. The resulting estimator is

$$\dot{\theta}_a = (X'Z\dot{\Omega}_a^{-1}Z'X)^{-1}X'Z\dot{\Omega}_a^{-1}Z'y \tag{4.19}$$

Applying a known matrix result, viz.

$$\dot{\Omega}_a^{-1} = (\dot{\Omega}_s - \tilde{g}\tilde{g}')^{-1} = \dot{\Omega}_s^{-1} + (1 - \tilde{g}'\dot{\Omega}_s^{-1}\tilde{g})^{-1}\dot{\Omega}_s^{-1}\tilde{g}\tilde{g}'\dot{\Omega}_s^{-1},$$

we obtain

$$\begin{aligned}
\dot{\theta}_a &= \{X'Z[\dot{\Omega}_s^{-1} + (1 - \tilde{g}'\dot{\Omega}_s^{-1}\tilde{g})^{-1}\dot{\Omega}_s^{-1}\tilde{g}\tilde{g}'\dot{\Omega}_s^{-1}]Z'X\}^{-1} \\
&\quad \times X'Z[\dot{\Omega}_s^{-1} + (1 - \tilde{g}'\dot{\Omega}_s^{-1}\tilde{g})^{-1}\dot{\Omega}_s^{-1}\tilde{g}\tilde{g}'\dot{\Omega}_s^{-1}]Z'y \\
&= \{X'Z\dot{\Omega}_s^{-1}Z'X + (1 - \tilde{g}'\dot{\Omega}_s^{-1}\tilde{g})^{-1}X'Z\dot{\Omega}_s^{-1}\tilde{g}\tilde{g}'\dot{\Omega}_s^{-1}Z'X\}^{-1} \\
&\quad \times \{X'Z\dot{\Omega}_s^{-1}Z'y + (1 - \tilde{g}'\dot{\Omega}_s^{-1}\tilde{g})^{-1}X'Z\dot{\Omega}_s^{-1}\tilde{g}\tilde{g}'\dot{\Omega}_s^{-1}Z'y\} \\
&= \{X'Z\dot{\Omega}_s^{-1}Z'X\}^{-1}X'Z\dot{\Omega}_s^{-1}Z'y = \tilde{\theta}.
\end{aligned}$$

The third equality is due to

$$X'Z\dot{\Omega}_s^{-1}\tilde{g} = \frac{1}{n}X'Z\dot{\Omega}_s^{-1}(Z'y - Z'X\tilde{\theta}) = \left(\frac{u(\tilde{\theta})'u(\tilde{\theta})}{n}\right)^{-1}X'P_Z(y - X\tilde{\theta}) = 0.$$

■

The 'standard' Sargan test (Sargan (1958)) arises from the application of (4.15) in (4.7), giving

$$S_n^s \equiv n\tilde{g}'\dot{\Omega}_s^{-1}\tilde{g} = n\frac{u(\tilde{\theta})'P_Z u(\tilde{\theta})}{u(\tilde{\theta})'u(\tilde{\theta})}. \tag{4.20}$$

Even though the estimator $\dot{\theta}_a = \dot{\theta}_s = \tilde{\theta}$, the test statistic

$$S_n^a \equiv n\tilde{g}'\dot{\Omega}_a^{-1}\tilde{g} \tag{4.21}$$

is not equivalent to S_n^s. In fact we have,

$$\begin{aligned} S_n^a &= n\tilde{g}'[\dot{\Omega}_s^{-1} + (1 - \tilde{g}'\dot{\Omega}_s^{-1}\tilde{g})^{-1}\dot{\Omega}_s^{-1}\tilde{g}\tilde{g}'\dot{\Omega}_s^{-1}]\tilde{g} \\ &= n\tilde{g}'\dot{\Omega}_s^{-1}\tilde{g} + (1 - \tilde{g}'\dot{\Omega}_s^{-1}\tilde{g})^{-1}n\tilde{g}'\dot{\Omega}_s^{-1}\tilde{g}\tilde{g}'\dot{\Omega}_s^{-1}\tilde{g} \\ &= S_n^s + \frac{S_n^s}{n - S_n^s}S_n^s = S_n^s\frac{n}{n - S_n^s}. \end{aligned} \tag{4.22}$$

Since $0 \leq S_n^s \leq n$ (equality can happen in the extreme cases when $u(\tilde{\theta})$ is either orthogonal to or completely spanned by Z), we have $S_n^a \geq S_n^s$. Hence, for a given critical value, S_n^a will never reject less often than S_n^s.

4.2.2 GEL statistics

Here we will shortly describe GEL and the resulting test statistic for overidentifying restrictions. For the analytical development of the following results see Smith (1997). For some more refined results on GEL see also Newey and Smith (2004).

Like GMM, GEL estimation is based on moment conditions, $\mathrm{E}g_i(\theta_0) = 0$. It assigns multinomial weights $\{\pi_i\}_{i=1}^n$ to each of the observations $\{\mathcal{X}_i\}_{i=1}^n$. This allows the GEL estimator to estimate the empirical (implied) probabilities $\{\hat{\pi}_i\}_{i=1}^n$ such that

$$\sum_{i=1}^n \hat{\pi}_i g_i(\hat{\theta}_{GEL}) = 0.$$

Note that for GMM we have $\hat{\pi}_i \equiv \frac{1}{n}$ and in general $\frac{1}{n}\sum_{i=1}^n g_i(\hat{\theta}_{GMM}) \neq 0$. For a concave scalar function $\rho(v)$ of the scalar v in an open interval $\mathcal{V}$ containing zero

$$\hat{\theta}_{GEL} \equiv \operatorname*{argmin}_{\theta\in\Theta} \sup_{\lambda\in\Lambda(\theta)} \sum_{i=1}^n \rho(\lambda' g_i(\theta)), \tag{4.23}$$

where $\Lambda(\theta) = \{\lambda : \lambda' g_i(\theta) \in \mathcal{V},\ i = 1, \ldots, n\}$. For example, the Empirical Likelihood estimator applies $\rho(v) = \ln(1 - v)$ with $\mathcal{V} = (-\infty, 1)$, the Exponential Tilting estimator uses $\rho(v) = -\exp(v)$ with $\mathcal{V} = \mathbb{R}$ and the Continuous Updating estimator employs $\rho(v) = -(1 + v)^2/2$ with $\mathcal{V} = \mathbb{R}$.

Let $\hat{g}_i \equiv g_i(\hat{\theta}_{GEL})$. The empirical probabilities of the observations associated with the GEL are

$$\hat{\pi}_i \equiv \frac{\rho'(\hat{\lambda}'\hat{g}_i)}{\sum_{j=1}^{n} \rho'(\hat{\lambda}'\hat{g}_j)}, \tag{4.24}$$

where $\rho'(v)$ is the first order derivative of ρ, and $\hat{\lambda} = \underset{\lambda \in \Lambda(\hat{\theta}_{GEL})}{\operatorname{argmax}} \sum_{i=1}^{n} \rho(\lambda'\hat{g}_i)$. These probabilities are important for the bootstrap procedure introduced by Brown and Newey (2002) which we will describe later.

The GEL likelihood ratio test statistic is

$$GELR_n \equiv 2(\sum_{i=1}^{n} \rho(\hat{\lambda}'\hat{g}_i) - n\rho(0))$$

and has asymptotic distribution $\chi^2(l - k)$, when all the moment conditions are valid.

In the simulations we will analyze the size and power properties of this likelihood ratio test using Empirical Likelihood (we will call it ELR) and Exponential Tilting (ETR). We will also analyze a version that applies the GMM estimator (4.8) instead of (4.23). Then $\hat{\lambda} = \underset{\lambda \in \Lambda(\hat{\theta}_a)}{\operatorname{argmax}} \sum_{i=1}^{n} \rho(\lambda' g_i(\hat{\theta}_a))$, and for ρ corresponding to either Empirical Likelihood or Exponential Tilting we will call the modified LR test $ELR(\hat{\theta})$ or $ETR(\hat{\theta})$ respectively.

4.3 Bootstrap procedures

Below we describe bootstrap procedures of Hall and Horowitz (1996) and Brown and Newey (2002) for improving the finite sample properties of the GMM overidentifying restrictions test statistics. The Hall and Horowitz (1996) version of the bootstrapped test statistic uses a 're-centered' moment function such that it satisfies the moment restrictions in the bootstrap samples. Brown and Newey (2002) propose resampling the data according to the probabilities associated with the observations that arise from the computation of the empirical likelihood evaluated at the GMM estimator. That way the moments exploited by GMM are also satisfied in the bootstrap samples, and in theory this procedure should lead to improvements with respect to the Hall and Horowitz (1996) procedure.

4.3.1 HH type bootstrap

We adopt here the bootstrap procedure of Hall and Horowitz (1996) which is originally designed to handle dependent data. Assuming that the data is i.i.d., we have $\mathrm{E}(g(\mathcal{X}_i, \theta_0) g(\mathcal{X}_j, \theta_0)') = O$ for $i \neq j$. The version of the GMM statistic they consider is the one that uses the inverse of (4.4) for the weighting matrix.

A bootstrap sample, $\mathcal{X}^*$, ($\mathcal{X}_i^*$ $i = 1, \ldots, n$) is obtained by drawing independently with replacement from $\mathcal{X}$ ($\mathcal{X}_i$ $i = 1, \ldots, n$). Let $\hat{\theta}$ be the estimator for which we would like to bootstrap the test statistic. Define

$$g_i^{h*}(\theta) \equiv g(\mathcal{X}_i^*, \theta) - \mathrm{E}^* g(\mathcal{X}_i^*, \hat{\theta}), \tag{4.25}$$

where $\mathrm{E}^*(\cdot) = \mathrm{E}(\cdot|\mathcal{X})$. We have

$$\mathrm{E}^* g(\mathcal{X}_i^*, \hat{\theta}) = \int g(x, \hat{\theta}) d\hat{F}(x) = \bar{g}_n(\hat{\theta})$$

and

$$\mathrm{E}^* g(\mathcal{X}_i^*, \hat{\theta}) g(\mathcal{X}_i^*, \hat{\theta})' = \int g(x, \hat{\theta}) g(x, \hat{\theta})' d\hat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} g_i(\hat{\theta}) g_i(\hat{\theta})',$$

where $\hat{F}(x)$ is the EDF of $\mathcal{X}$. Writing $g_i^*(\theta) \equiv g(\mathcal{X}_i^*, \theta)$ we get

$$g_i^{h*}(\theta) = g_i^*(\theta) - \bar{g}_n(\hat{\theta}). \tag{4.26}$$

Clearly

$$\mathrm{E}^* g_i^{h*}(\hat{\theta}) = \bar{g}_n(\hat{\theta}) - \bar{g}_n(\hat{\theta}) = 0, \tag{4.27}$$

which clarifies the structure of the bootstrap moment function (4.25). For that choice the population moments exploited in the estimation, $\mathrm{E}g_i(\theta_0) = 0$, are satisfied exactly in the bootstrap samples at $\hat{\theta}$. Because of the independence of the bootstrap samples

($\mathrm{E}^* g_i^{h*}(\hat\theta) g_j^{h*}(\hat\theta)' = 0$ for $i \neq j$) the bootstrap variance of $\sqrt{n}\bar g_n^{h*}(\hat\theta)$ is

$$\begin{aligned} n\mathrm{E}^* \bar g_n^{h*}(\hat\theta)\bar g_n^{h*}(\hat\theta)' &= \frac{1}{n}\sum_{i,j=1}^n \mathrm{E}^* g_i^{h*}(\hat\theta) g_j^{h*}(\hat\theta)' = \mathrm{E}^* g_i^{h*}(\hat\theta) g_i^{h*}(\hat\theta)' \\ &= \mathrm{E}^*[g_i^*(\hat\theta) - \bar g_n(\hat\theta)][g_i^*(\hat\theta) - \bar g_n(\hat\theta)]' \\ &= \mathrm{E}^* g_i^*(\hat\theta) g_i^*(\hat\theta)' - \bar g_n(\hat\theta)\bar g_n(\hat\theta)' \\ &= \frac{1}{n}\sum_{i=1}^n g_i(\hat\theta) g_i(\hat\theta)' - \bar g_n(\hat\theta)\bar g_n(\hat\theta)' = \hat\Omega_a(\hat\theta). \end{aligned}$$

The bootstrap version of (4.4) is

$$\begin{aligned} \hat\Omega_s^{h*}(\tilde\theta^*) &= \frac{1}{n}\sum_{i=1}^n g_i^{h*}(\tilde\theta^*) g_i^{h*}(\tilde\theta^*)' \\ &= \frac{1}{n}\sum_{i=1}^n g_i^*(\tilde\theta^*) g_i^*(\tilde\theta^*)' - \bar g_n^*(\tilde\theta^*)\bar g_n(\hat\theta)' \\ &\quad -\bar g_n(\hat\theta)\bar g_n^*(\tilde\theta^*)' + \bar g_n(\hat\theta)\bar g_n(\hat\theta)' \end{aligned} \tag{4.28}$$

and the bootstrap version of (4.3) is

$$\begin{aligned} \hat\Omega_a^{h*}(\tilde\theta^*) &= \frac{1}{n}\sum_{i=1}^n g_i^{h*}(\tilde\theta^*) g_i^{h*}(\tilde\theta^*)' - \bar g_n^{h*}(\tilde\theta^*)\bar g_n^{h*}(\tilde\theta^*)' \\ &= \frac{1}{n}\sum_{i=1}^n g_i^*(\tilde\theta^*) g_i^*(\tilde\theta^*)' - \bar g_n^*(\tilde\theta^*)\bar g_n^*(\tilde\theta^*)', \end{aligned} \tag{4.29}$$

where $\tilde\theta^*$ is obtained from

$$\tilde\theta^* \equiv \operatorname*{argmin}_{\theta\in\Theta} \bar g_n^{h*}(\theta)' W^* \bar g_n^{h*}(\theta), \tag{4.30}$$

for some initial weighting matrix W^*. Since $\mathrm{E}^* g_i^*(\hat\theta) g_i^*(\hat\theta)' = \frac{1}{n}\sum_{i=1}^n g_i(\hat\theta) g_i(\hat\theta)'$ and $\mathrm{E}^* g_i^*(\hat\theta) = \bar g_n(\hat\theta)$, it is easy to see that

$$\mathrm{E}^*\hat\Omega_s^{h*}(\hat\theta) = \hat\Omega_a(\hat\theta). \tag{4.31}$$

Lemma 4.3 *For 4.29 evaluated at $\hat{\theta}$, we have*

$$\mathrm{E}^*\hat{\Omega}_a^{h*}(\hat{\theta}) = \frac{n-1}{n}\hat{\Omega}_a(\hat{\theta}). \tag{4.32}$$

Proof. Let $\hat{A} \equiv \frac{1}{n}\sum_{i=1}^n g_i(\hat{\theta})g_i(\hat{\theta})'$. Now consider

$$\mathrm{E}^*\hat{\Omega}_a^{h*}(\hat{\theta}) = \mathrm{E}^*\big\{\frac{1}{n}\sum_{i=1}^n g_i^*(\hat{\theta})g_i^*(\hat{\theta})' - \bar{g}_n^*(\hat{\theta})\bar{g}_n^*(\hat{\theta})'\big\}.$$

Because of independence of the terms in the summation and because $\mathrm{E}^*g_i^*(\hat{\theta})g_i^*(\hat{\theta})' = \hat{A}$, we have $\mathrm{E}^*\big\{\frac{1}{n}\sum_{i=1}^n g_i^*(\hat{\theta})g_i^*(\hat{\theta})'\big\} = \hat{A}$. Now

$$\begin{aligned}
\mathrm{E}^*\bar{g}_n^*(\hat{\theta})\bar{g}_n^*(\hat{\theta})' &= \frac{1}{n^2}\mathrm{E}^*\sum_{i,j=1}^n \big\{g_i^*(\hat{\theta})g_j^*(\hat{\theta})'\big\} \\
&= \frac{1}{n^2}\mathrm{E}^*\sum_{i=j}\big\{g_i^*(\hat{\theta})g_j^*(\hat{\theta})'\big\} + \frac{1}{n^2}\mathrm{E}^*\sum_{i\neq j}\big\{g_i^*(\hat{\theta})g_j^*(\hat{\theta})'\big\} \\
&= \frac{1}{n}\hat{A} + \frac{n^2-n}{n^2}\mathrm{E}^*g_i^*(\hat{\theta})g_j^*(\hat{\theta})' = \frac{1}{n}\hat{A} + \frac{n-1}{n}\bar{g}_n(\hat{\theta})\bar{g}_n(\hat{\theta})'.
\end{aligned}$$

The last equality comes from the independence of $g_i^*(\hat{\theta})$ and $g_j^*(\hat{\theta})$ for $i \neq j$, then $\mathrm{E}^*g_i^*(\hat{\theta})g_j^*(\hat{\theta})' = \mathrm{E}^*g_i^*(\hat{\theta})\mathrm{E}^*g_j^*(\hat{\theta})'$. Summing up, we get

$$\mathrm{E}^*\hat{\Omega}_a^{h*}(\hat{\theta}) = \hat{A} - \frac{1}{n}\hat{A} - \frac{n-1}{n}\bar{g}_n(\hat{\theta})\bar{g}_n(\hat{\theta})' = \frac{n-1}{n}\hat{\Omega}_a(\hat{\theta}).$$

■

Hence, for the W^* in (4.30), we shall use the inverse of $\hat{\Omega}_s^{h*}(\hat{\theta})$ (if we bootstrap the 'standard' test statistic) or $\hat{\Omega}_a^{h*}(\hat{\theta})$ (for the 'adaptive' one, where the multiplicative factor resulting from (4.32), $\frac{n}{n-1}$, does not affect the estimator (and thus can be omitted) but it will influence the bootstrap version of the test statistic). We will also examine how, using the 'updated' (4.28) and (4.29), changes the results with respect to this initial choice.

The bootstrap versions of (4.6) and (4.7) are

$$J_a^* = (n-1)\bar{g}_n^{h*}(\tilde{\theta}_a^*)'\hat{\Omega}_a^{h*}(\hat{\theta}_a)^{-1}\bar{g}_n(\tilde{\theta}_a^*) \tag{4.33}$$

and

$$J_s^* = n\bar{g}_n^{h*}(\tilde{\theta}_s^*)'\hat{\Omega}_s^{h*}(\hat{\theta}_s)^{-1}\bar{g}_n(\tilde{\theta}_s^*), \tag{4.34}$$

where the factor $n-1$ in (4.33) is justified by (4.32). If we 'update' $\hat{\theta}$ in the weighting matrix, the second stage versions are

$$J_a^{**} = (n-1)\bar{g}_n^{h*}(\hat{\theta}_a^*)'\hat{\Omega}_a^{h*}(\hat{\theta}_a^*)^{-1}\bar{g}_n(\hat{\theta}_a^*) \tag{4.35}$$

and

$$J_s^{**} = n\bar{g}_n^{h*}(\hat{\theta}_s^*)'\hat{\Omega}_s^{h*}(\hat{\theta}_s^*)^{-1}\bar{g}_n(\hat{\theta}_s^*). \tag{4.36}$$

Here, ignoring the subscript a or s, $\tilde{\theta}^*$ is obtained from (4.30) using for the weighting matrix $\hat{\Omega}^{h*}(\hat{\theta})^{-1}$ and $\hat{\theta}^*$ using $\hat{\Omega}^{h*}(\tilde{\theta}^*)^{-1}$. In the simulation, we will refer to the bootstrap critical values obtained from J^* or J^{**} as 1 step and 2 step critical values.

Having designed their procedure primarily for dependent data, Hall and Horowitz (1996) are concerned with the fact that the block bootstrap does not replicate the dependence of the true data generating process. To overcome this issue they apply a particular transformation to (4.28). Since we are not dealing here with dependent data we do not need to apply this transformation here.

Linear model

In the linear case we have

$$\bar{g}_n^{h*}(\theta) = \frac{1}{n}Z^{*'}u^*(\theta) - \frac{1}{n}Z'u(\hat{\theta}),$$

where $u^*(\theta) \equiv y^* - X^*\theta$. For a given weighting matrix W^*, say the inverse of $\hat{\Omega}_s^{h*}(\hat{\theta})$ or $\hat{\Omega}_a^{h*}(\hat{\theta})$, (4.30) becomes

$$\tilde{\theta}^* = (X^{*'}Z^*W^*Z^{*'}X^*)^{-1}X^{*'}Z^*W^*(Z^{*'}y^* - Z'(y - X\hat{\theta})). \tag{4.37}$$

The bootstrap version of the weighting matrix used for the Sargan test statistic based on (4.28) or (4.29) will use the inverse of

$$\begin{aligned}\dot{\Omega}_s^{h*}(\tilde{\theta}^*) &= \frac{u^*(\tilde{\theta}^*)'u^*(\tilde{\theta}^*)}{n}\frac{1}{n}Z^{*'}Z^* - \frac{1}{n^2}Z^{*'}u^*(\tilde{\theta}^*)u(\hat{\theta})'Z \\ &\quad -\frac{1}{n^2}Z'u(\hat{\theta})u^*(\tilde{\theta}^*)'Z^* + \frac{1}{n^2}Z'u(\hat{\theta})u(\hat{\theta})'Z,\end{aligned} \tag{4.38}$$

or

$$\dot{\Omega}_a^{h*}(\tilde{\theta}^*) = \frac{u^*(\tilde{\theta}^*)'u^*(\tilde{\theta}^*)}{n}\frac{1}{n}Z^{*'}Z^* - \frac{1}{n^2}Z^{*'}u^*(\tilde{\theta}^*)u^*(\tilde{\theta}^*)'Z. \tag{4.39}$$

Hence, the bootstrap versions of the standard or alternative Sargan tests (S_a^*, S_s^*, S_a^{**} and S_s^{**}) will have the same structure as (4.33), (4.34), (4.35), (4.36) but with the weighting matrix replaced with 'the dotted' version.

In the MC experiments we will examine how, for a given version of the test statistic, the bootstrap critical values obtained from the bootstrapped statistics (J_a^*, J_s^*, J_a^{**}, J_s^{**}, S_a^*, S_s^*, S_a^{**} or S_s^{**}) perform relative to the standard asymptotic critical values.

4.3.2 EL type bootstrap

An alternative approach to the one of Hall and Horowitz (1996) was proposed by Brown and Newey (2002). Instead of drawing (with replacement) bootstrap samples $\mathcal{X}_i^* = (y_i^*, x_i^{*'}, z_i^{*'})$ from $\mathcal{X}$, with each $\mathcal{X}_i$ having equal probability, they suggest using the probabilities obtained from the calculation of the empirical likelihood at the given GMM estimator, $\hat{\theta}$. These are

$$\hat{\pi}_i = \frac{1}{n(1-\hat{\lambda}'g_i(\hat{\theta}))}, \tag{4.40}$$

with

$$\hat{\lambda}' = \underset{\lambda' g_i(\hat{\theta})<1}{\arg\max} \sum \ln(1-\lambda' g_i(\hat{\theta})).$$

For that choice we then have

$$\sum_{i=1}^{n} \hat{\pi}_i g_i(\hat{\theta}) = 0.$$

The bootstrap procedure then goes as follows:

- For a given GMM estimator $\hat{\theta}$ and the 'resulting' statistic $J_n(\hat{\theta})$ obtain the probabilities (4.40)

- obtain bootstrap samples $\mathcal{X}^{b*}$ by drawing with replacement from $\mathcal{X}$, where $\mathcal{X}_i$ has the probability of being drawn $\hat{\pi}_i$, $i = 1, \ldots, n$ $b = 1, \ldots, B$.

- compute the statistic $J_n^{b*}(\hat{\theta})$ exactly the same way $J_n(\hat{\theta})$ was obtained but using the bootstrapped data $\mathcal{X}^{b*}$, instead of $\mathcal{X}$.

- the bootstrap α level critical value is the $100(1-\alpha)\%$ quantile of the bootstrap distribution: $cv^* \approx J_n^{[(1-\alpha)B]*}$

A modification to that procedure could use probabilities derived from another GEL member. For example, for the exponential tilting, where $\rho(v) = -\exp\{v\}$, we would use

$$\hat{\pi}_i = \frac{\exp(\hat{\lambda}' g_i(\hat{\theta}))}{\sum_{j=1}^{n} \exp(\hat{\lambda}' g_j(\hat{\theta}))}, \tag{4.41}$$

with

$$\hat{\lambda}' = \underset{\lambda \in \Lambda(\hat{\theta})}{\operatorname{argmax}} \Big\{ -\sum_{i=1}^{n} \exp(\lambda' g_i(\hat{\theta})) \Big\}.$$

4.3.3 Residual type bootstrap

For the linear model (4.44) that we are going to analyze, the residual type bootstrap consists of:

- obtain $\hat{\Pi}_{ols} = (Z'Z)^{-1}Z'X$ and the residuals $\hat{V} = M_Z X$ (scaled by $\sqrt{\frac{n}{n-l}}$)

- obtain $\hat{\theta}$ from (4.13) (we used (4.11) for the weighting matrix) and the residuals $\hat{u} = y - X\hat{\theta}$

- from $V = (\hat{u}, \hat{V})$, for $b = 1, \ldots, B$ obtain bootstrap versions of the disturbances

$V^{b*} = (\hat{u}^{b*}, \hat{V}^{b*})$ by re-drawing 'row-wise'; generate

$$X^{b*} = Z\hat{\Pi}_{ols} + \hat{V}^{b*} \tag{4.42}$$

$$y^{b*} = X^{b*}\hat{\theta} + \hat{u}^{b*} \tag{4.43}$$

- compute the bootstrap versions of the Sargan or Hansen tests (from y^{b*}, X^{b*} and Z)
- the bootstrap critical value is obtained from the $100(1-\alpha)\%$ quantile of the bootstrap versions of the statistic.

By 'fixing' Z in (4.42) we produce exogeneity of the instruments in the resampling scheme. By drawing 'row-wise' from the residual matrix V we preserve in (4.43) possible simultaneity of the regressor and heteroscedasticity of the disturbances.

4.4 Illustrations

Here we present illustrations. We will analyze two linear examples, one involving homoscedastic errors the other heteroscedastic ones. For a $(n \times l)$ matrix Z we denote by Z_i its i'th column and its t'th row we denote by z_t'.

4.4.1 Homoscedastic Example

Let us consider the simple linear model

$$\begin{aligned} y_t &= x_t\theta + u_t && (4.44)\\ x_t &= \bar{z}_t'\pi + v_t, && (4.45) \end{aligned}$$

where y_t, x_t are scalar endogenous variables, $t = 1, \ldots, n$, π is a $(l \times 1)$ vector of reduced form parameters. The instruments $\bar{Z} = [\bar{z}_1, \ldots, \bar{z}_T]'$ are exogenous, $\mathrm{E}(u|\bar{Z}) = 0$ with $\mathrm{Var}(\bar{z}_t) = I_l$. Let

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim \mathrm{IIN}\big(0, \Sigma\big), \quad \Sigma = \begin{pmatrix} \sigma_u^2 & \rho_{uv}\sigma_u\sigma_v \\ \rho_{uv}\sigma_u\sigma_v & \sigma_v^2 \end{pmatrix}. \tag{4.46}$$

We take $l = 3$ and create one invalid instrument (say the first one, Z_1) by generating Z_1 according to

$$Z_1 = \sqrt{1 - \rho_{Z_1u}^2}\bar{Z}_1 + \rho_{Z_1u}u. \tag{4.47}$$

We have $\mathrm{Var}(z_{t1}) = 1$. For the normalization we take $\sigma_u^2 = 1$, $\theta = 1$. We will choose values for π and σ_v^2 via population versions of the concentration parameter and the signal to noise ratio, which we define below.

The concentration parameter in the valid instruments case for (4.45) is

$$\mu^2 = \pi'\bar{Z}'\bar{Z}\pi/\sigma_v^2. \tag{4.48}$$

Denote the 'population' version of the concentration parameter by

$$\mu_p^2 = n\pi'\Sigma_{\bar{Z}'\bar{Z}}\pi/\sigma_v^2 = \frac{n}{\sigma_v^2}\sum_{i=1}^{l}\pi_i^2. \tag{4.49}$$

Denote $\mathrm{Var}(x_t)$ by σ_x^2. We then have

$$\sigma_x^2 = \pi'\Sigma_{\bar{Z}'\bar{Z}}\pi + \sigma_v^2 = \sum_{i=1}^{l}\pi_i^2 + \sigma_v^2, \tag{4.50}$$

and combining (4.49) with (4.50) we get

$$\sigma_v^2 = \frac{\sigma_x^2}{\mu_p^2/n + 1}. \tag{4.51}$$

We define the signal to noise ratio of (4.44) by $\eta^2 = \frac{\mathrm{Var}(x_t\theta)}{\sigma_u^2}$, then $\eta^2 = \sigma_x^2$.

Taking $\pi_i = \pi_0$ in (4.49) for $i = 1, \ldots, l$, we get $\mu_p^2\sigma_v^2 = nl\pi_0^2$. Hence, for a given sample size n, μ_p^2, and η^2 we can calculate σ_v^2 from (4.51) and π_0 from $\pi_0^2 = \mu_p^2\sigma_v^2/(nl)$.

4.4.2 Heteroscedastic Example

In the above example we have conditional homoscedasticity, because $\mathrm{E}(u_t^2|\bar{z}_t) = 1$. The Sargan test seems the most appropriate one to use. Here we will generalize the previous model by introducing conditional heteroscedasticity.

Let $w_t \equiv (\bar{z}_t'\bar{z}_t)/l$. Because $\bar{z}_t \sim N(0, I_l)$, we obtain $\mathrm{E}(w_t) = 1$ and $\mathrm{E}(w_t^2) = 5/3$ for $l = 3$ (as $w_t \sim \chi^2(3)/3$). Let $\bar{w}_t \equiv \frac{\alpha_1+\alpha_2 w_t}{\sqrt{\alpha_1^2+\frac{5}{3}\alpha_2^2+2\alpha_1\alpha_2}}$, for some scalars α_1, α_2. By construction $\mathrm{E}\bar{w}_t^2 = 1$ and

$$\mathrm{E}\bar{w}_t = \frac{\alpha_1 + \alpha_2}{\sqrt{\alpha_1^2 + \frac{5}{3}\alpha_2^2 + 2\alpha_1\alpha_2}} \equiv \gamma.$$

In the extended Monte Carlo design below the $\bar{w}_t$ will be used to introduce conditional heteroscedasticity given by $\bar{w}_t^2$, by generating the disturbances as $\tilde{u}_t = \bar{w}_t u_t$. Note that γ can be equal to one or minus one only if $\alpha_2 = 0$. Then $\bar{w}_t$ is in fact nonrandom and we again have conditional homoscedasticity. For $\alpha_2 \neq 0$ we can rewrite the above as

$$\bar{w}_t = \frac{\kappa + w_t}{\sqrt{\kappa^2 + 2\kappa + 5/3}} \tag{4.52}$$

and

$$\gamma = \mathrm{E}\bar{w}_t = \frac{\kappa + 1}{\sqrt{\kappa^2 + 2\kappa + 5/3}}, \tag{4.53}$$

where $\kappa = \alpha_1/\alpha_2$. One of the solutions to (4.53) is

$$\kappa = \frac{\gamma}{\sqrt{1-\gamma^2}}\sqrt{\frac{2}{3}} - 1, \tag{4.54}$$

which we will use to parameterize the conditional heteroscedasticity in our simulations.

Now, for $|\gamma| \neq 1$, we generate

$$y_t = x_t'\theta + \tilde{u}_t \tag{4.55}$$

$$x_t = \bar{z}_t'\pi + v_t, \tag{4.56}$$

where $\tilde{u}_t = \bar{w}_t u_t$ (hence yielding conditional heteroscedasticity, $E(\tilde{u}_t^2|\bar{z}_t) = \bar{w}_t^2$). When $|\gamma| = 1$ then we would take $\tilde{u}_t = u_t$ and we would be back in the previous example. We have

$$\mathrm{E}\tilde{u}_t = \mathrm{E}\mathrm{E}(\tilde{u}_t|\bar{z}_t) = \mathrm{E}\bar{w}_t\mathrm{E}(u_t|\bar{z}_t) = 0$$

$$\mathrm{E}\tilde{u}_t^2 = \mathrm{E}\mathrm{E}(\tilde{u}_t^2|\bar{z}_t) = \mathrm{E}\bar{w}_t{}^2\mathrm{E}(u_t^2|\bar{z}_t) = 1.$$

$$\mathrm{E}\tilde{u}_t^3 = \mathrm{E}\mathrm{E}(\tilde{u}_t^3|\bar{z}_t) = \mathrm{E}\bar{w}_t{}^3\mathrm{E}(u_t^3|\bar{z}_t) = 0.$$

Hence, the unconditional first three moments correspond to those from the previous example. The fourth unconditional moment, however, is

$$\mathrm{E}\tilde{u}_t^4 = \mathrm{E}\mathrm{E}(\tilde{u}_t^4|\bar{z}_t) = \mathrm{E}\bar{w}_t{}^4\mathrm{E}(u_t^4|\bar{z}_t) = 3\mathrm{E}\bar{w}_t{}^4 =$$

$$3\mathrm{E}\frac{\kappa^4 + 4\kappa^3 w_t + 6\kappa^2 w_t^2 + 4\kappa w_t^3 + w_t^4}{(\kappa^2 + 2\kappa + 5/3)^4} = 3\frac{\kappa^4 + 4\kappa^3 + 10\kappa^2 + 15\frac{5}{9}\kappa + 11\frac{2}{3}}{(\kappa^2 + 2\kappa + 5/3)^4}.$$

Via (4.54) $\mathrm{E}\tilde{u}_t^4$ is a function of γ. It can be shown that $\mathrm{E}\tilde{u}_t^4 > 3$ for $|\gamma| < 1$. Figure

4.1 shows the graph of $\mathrm{E}\tilde{u}_t^4$ for different γ's. Hence, for $|\gamma| < 1$ the distribution is more

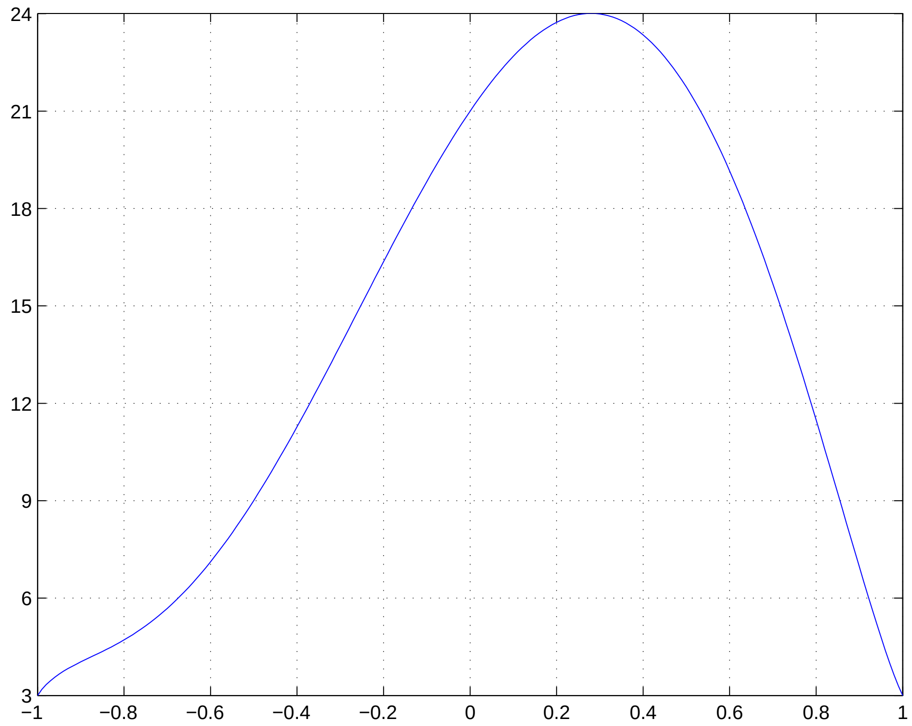

Figure 4.1: $\mathrm{E}\tilde{u}_t^4$ for different γ's.

peaked around the mean with fatter tails than the standard normal, as can be seen from Figure 4.2, which presents the shape of the density of $\tilde{u}_t$ in comparison to the standard normal. Figure 4.3 shows realizations of u_t and $\tilde{u}_t$.

From (4.56), we now have

$$\mathrm{E}x_t\tilde{u}_t = \mathrm{E}\mathrm{E}(\bar{z}_t'\pi\bar{w}_t u_t|\bar{z}_t) + \mathrm{E}\mathrm{E}(\bar{w}_t u_t v_{2t}|\bar{z}_t) =$$

$$\mathrm{E}(\bar{z}_t'\pi\bar{w}_t)\mathrm{E}(u_t|\bar{z}_t) + \mathrm{E}\bar{w}_t\mathrm{E}(u_t v_{2t}|\bar{z}_t) = \gamma\rho_{uv}\sigma_v.$$

Because $\mathrm{Var}x_t = \eta^2$ and $\mathrm{Var}\tilde{u}_t = 1$ then

$$\rho_{x\tilde{u}} = \frac{\gamma\rho_{uv}\sigma_v}{\eta}.$$

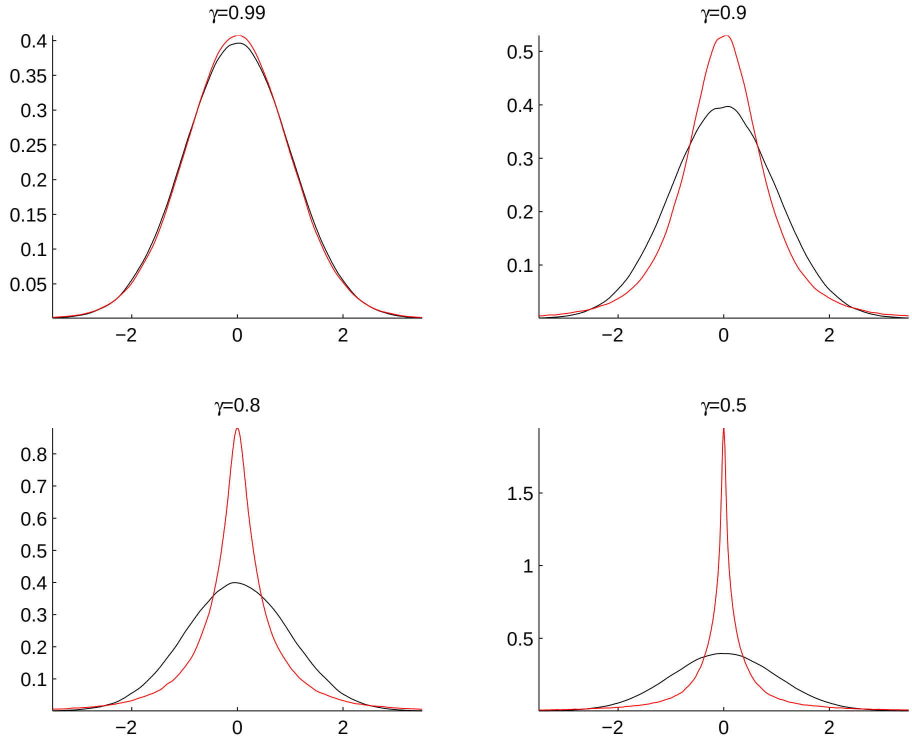

Figure 4.2: Densities of u_t (black) and $\tilde{u}_t$ (red) for different γ's.

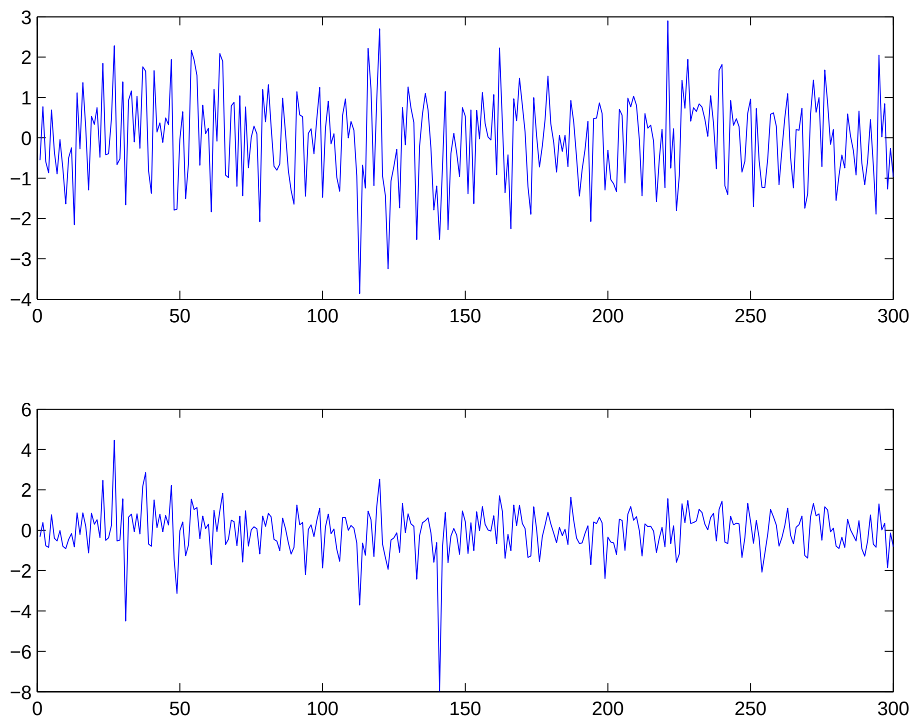

Figure 4.3: A realization of u_t (upper panel) and $\tilde{u}_t$ (lower panel) for $\gamma = 0.9$.

Similarly to (4.47) we generate the invalid instrument, Z_1, according to

$$Z_1 = \sqrt{1 - \rho^2_{Z_1\tilde{u}}}\bar{Z}_1 + \rho_{Z_1\tilde{u}}\tilde{u}. \tag{4.57}$$

Results

For all overidentification tests presented above and for different values of the concentration parameter μ_p^2, signal to noise ratio η^2, sample size n, simultaneity ρ_{xu} (or $\rho_{x\tilde{u}}$ when $\gamma \neq 1$), we will present our findings on size distortion and size corrected power by varying ρ_{Z_1u} (the degree of invalidity of the instrument when $\gamma = 1$) or $\rho_{Z_1\tilde{u}}$ (when $\gamma \neq 1$).

For the heteroscedastic case we use $\gamma = 0.9$. For that choice the fourth moment of $\tilde{u}_t$ is approximately 6.96 and the peak of the distribution of the disturbances is moderate. For given model parameters we generate $\bar{z}_t' \sim N(0, I_l)$ and the disturbances (4.46) (there we use $\rho_{uv}\sigma_v = \rho_{x\tilde{u}}\eta/\gamma$), we obtain (4.44), (4.45), (4.47) for the homoscedastic case and (4.55), (4.56), (4.57) for heteroscedastic disturbances.

We take the number of Monte Carlo replications $MC = 190000$ (we made that choice to achieve 'the accuracy' of 0.0005 for $\alpha = 0.05$, see Table 4.2) for the 'GMM' results (Sargan and Hansen tests) and $MC = 50000$ (to achieve 'the accuracy' of 0.001 for $\alpha = 0.05$) for the results involving 'GEL' (ELR, ETR, $ELR(\hat{\theta})$, $ETR(\hat{\theta})$), whereas the number of bootstrap replications $B = 100$.

Whenever computations involve Empirical Likelihood or Exponential Tilting, we used the Matlab optimization toolbox to obtain $\hat{\theta}$ (function `fminsearch.m` modifying the default accuracy criteria (`'TolFun'` and `'TolX'`) from 10^{-4} to 10^{-10} setting `'MaxIter'`=100) and adopted the `elm.m` program from A. Owen's webpage to calculate $\hat{\lambda}$ for Empirical Likelihood. Moreover, we used B. Hansen's lecture notes for writing the Newton algorithm for $\hat{\lambda}$ in Exponential Tilting cases. For the weak instrument case ($\mu_p^2/3 = 1$), $2.7\% - 4.5\%$ of the simulations where skipped (hence the number of MC accordingly is smaller) when Empirical Likelihood or Exponential Tilting were being computed, due to lack of convergence of the optimization procedure. For the strong instrument case ($\mu_p^2/3 = 20$), this percentage was almost zero (about 5 simulations out of 50000 were skipped).

For $n = 50,\ 200$ (small and larger sample), $\mu_p^2/3 = 20,\ 1$ (rather strong and weak instruments), and $\gamma = 1,\ 0.9$ (homoscedastic and heteroscedastic disturbances) Tables 4.3 and 4.4 contain the Monte Carlo estimate of $\frac{P(T_n > cv) - \alpha}{\alpha}$, i.e. the fraction of the size distortion for the given nominal significance level $\alpha = 0.05$ of the standard Sargan and Hansen tests, S_n^s and J_n^s, and the adapted versions, S_n^a and J_n^a. The critical values employed are either based on asymptotic theory (cv_∞), Hall and Horowitz (1996) (referred to as HH from below on) 1 step (cv^*) or 2 steps (cv^{**}) bootstrap, or the residual type bootstrap (cv_R). In the column indicating instruments used by a test (say T_n), [123] means that all the three instruments are taken, [23] only the second and the third, [123]|[23] means that the test used is the incremental one ($T_n([123]) - T_n([23])$). We do not show the results for [13] or [12] in the size distortion Tables 4.3, 4.4 and 4.5 since all three instruments are valid and i.i.d. here, they would yield the same (similar) results as $T_n([23])$. Likewise, as far as type I errors are concerned, $T_n([123]) - T_n([23])$ gives similar results as $T_n([123]) - T_n([12])$ or $T_n([123]) - T_n([13])$. In fact, for the size distortion results, we averaged the results for [12] and [23] (hence the result is based on 2MC replications). We did the same for [123]|[12] and [123]|[23].

Given the 'true' rejection probability p, the MC estimates of the size distortion results would have standard deviation equal to $\sqrt{\frac{p(1-p)}{MC}}/\alpha$. Table 4.1 shows this standard deviation for a given p and actual number of Monte Carlo simulations. For 'GEL' results this is (almost) 50000 for [123] and (almost) 100000 for [12] and [123]|[12]. For 'GMM' results it is 190000 or 380000 respectively.

Table 4.1: Standard Deviation of MC estimate of $\frac{p-0.05}{0.05}$: $\sqrt{\frac{p(1-p)}{MC}}/0.05$

	p				
MC	0.01	0.05	0.1	0.2	0.3
50000	0.0028	0.0195	0.0268	0.0358	0.0410
100000	0.0020	0.0138	0.0190	0.0253	0.0290
1900000	0.0015	0.0100	0.0138	0.0184	0.0210
3800000	0.0010	0.0071	0.0097	0.0130	0.0149

As we will see from the size distortion tables, the GMM tests will mainly under-reject (hence the *true* p is less than 0.05). That means that the standard errors of the MC results are less than or about 0.01 (i.e. 1%). When the tests over-reject, but less than

Table 4.2: Standard Deviation of MC estimate of p: $\sqrt{\frac{p(1-p)}{MC}}$

	p				
MC	0.01	0.05	0.2	0.3	0.5
50000	0.0004	0.0010	0.0018	0.0020	0.0022
1900000	0.0002	0.0005	0.0009	0.0011	0.0011

100%, the standard errors are about 0.02 (i.e. 2%). When tests over-reject by more than 100%, this occurs only on 3 occasions for LR tests and almost always for the Sargan tests in the heteroscedastic case, the standard errors do not exceed 0.04 (i.e. are less than 4% and this accuracy suffices).

Tables 4.6 to 4.9 show the size corrected power of the analyzed tests. Here, we can not merge different results, hence, the number of Monte Carlo replications is 50000 for all the GEL results and 190000 for the GMM results. There, if the true rejection probability is p, the 'power' results have standard deviation $\sqrt{\frac{p(1-p)}{MC}}$. Table 4.2 shows values of the standard deviation for given p and MC. We notice that the GMM results will be twice as accurate. The highest standard errors are obtained when $p = 0.5$. Hence, the standard errors of the results are never higher than the values in the last column of Table 4.2.

Tables interpretation

From Tables 4.3 ($\gamma = 1$) and 4.4 ($\gamma = 0.9$) we see that as far as size distortion is concerned, 2 stage HH bootstrap performs better than 1 step, except for the $\gamma = 1$, $n = 200$, $\mu_p^2/3 = 20$ case.

For the homoscedastic case (both for the Sargan and Hansen tests), we notice that the standard asymptotic critical values produce size distortions of more than 10% in absolute value in the cases when we have 'weak' instruments (for both sample sizes we consider) or 'strong' instruments (for $n = 50$). In those cases HH bootstrap gives improvements with respect to the asymptotic results, but we notice that the residual bootstrap performs even better. When we have the 'large' sample ($n = 200$) and strong instruments the standard asymptotics works very well and bootstrapping does not produce better results. Fortunately in this case, bootstrap size distortions never exceed 10%. Hence, for both

Sargan and Hansen tests the residual bootstrap would be preferable to the HH in the homoscedastic case, since it seems to perform better in 'weak' instruments cases and in small samples with 'strong' instruments.

For the heteroscedastic example, however, when applied to the Sargan tests, the residual bootstrap performs equally badly as its standard asymptotic implementation. For that case HH bootstrap (2 step) gives improvements, the nicest for the large samples ($< 10\%$). For the Hansen tests HH bootstrap does not give improvements in the cases examined, but the residual bootstrap does. Note that for the Sargan tests HH bootstrap size distortions are smaller than the corresponding asymptotic ones for the Hansen test, and clearly smaller than the corresponding HH ones. They are also smaller or about the same as the corresponding residual bootstrap size distortions for Hansen tests when $n = 200$. Hence it seems that, even though asymptotic χ^2 critical values are not appropriate for the Sargan test here, the HH bootstrap produces valid critical values (replicates well the distribution of the Sargan statistic).

Table 4.5 shows results for the procedures applying Empirical Likelihood or Exponential Tilting. The first columns present outcomes of applying the BN bootstrap for the alternative Sargan and Hansen tests, with either Empirical Likelihood or Exponential Tilting implied probabilities. We notice that this bootstrap procedure underrejects by at least 50%. (The results for the 'standard' Sargan and Hansen tests were almost identical) The right-hand columns show the LR type tests. We see that, for the strong instrument case, ELR is less distorted than $ELR(\hat{\theta})$, similarly for ETR and $ETR(\hat{\theta})$, and this size distortion decreases with the sample size. For the weak instrument case ELR and ETR are worse than the versions evaluated at the GMM estimators, which is probably due to the optimization procedure applied to find Empirical Likelihood or Exponential Tilting estimators. Apart from a few (regular) cases, the distortions are rather high.

Tables 4.6 to 4.9 show the size corrected power of the analyzed GMM tests and the LR type tests considered in Table 4.5. They all seem to perform about equally. We have highlighted the entries with the highest values (row-wise).

For the homoscedastic case, we notice that the Sargan tests are the best, overall. From [123]|[12] entries we conclude that the 'adapted' version is probably better to reject [123].

Unfortunately, we would probably conclude here that it is due to the 'third' instrument, instead of the 'first'. We notice that for all the tests, the power is substantial only for rather strong invalidity of the instrument ($\rho_{Z_1u} = 0.5$), where for the weak instruments we have only about 50% chance to be right in rejecting (similarly for $\rho_{Z_1u} = 0.2$, large sample and strong instrument).

The corresponding heteroscedastic results show the power loss with respect to the homoscedastic case. Here the power is high only for the large sample, strong instrument and serious invalidity. There is no uniformly best test in the heteroscedastic case.

For the values of n, γ and $\mu_p^2/3$ analyzed in the previous Tables, Figures 4.4 (for $\gamma = 1$) and 4.5 (for $\gamma = 0.9$) each show four panels of four plots (2 'Sargans' and 2 'Hansens') for 3 values of invalidity of the instrument (0.1, 0.2, 0.5) on the horizontal axes. They depict the worst case among the different 'instrument settings' ([123], [12], [123]||[12] and [123]||[23]) of $\frac{\hat{p}-\hat{p}_c}{\hat{p}_c}$ - percentage distortion of the 'power' rejection probability, $\hat{p}$, obtained using different critical values (asymptotic, HH and the residual bootstrap) with respect to the 'size corrected power', $\hat{p}_c$. First, we notice that using asymptotic critical values is almost never the best choice (in terms of being closest to the 'size corrected' power line). For the J tests it seems that the residual bootstrap is the best, except for the large sample and weak instrument case. The percentage distortion of HH bootstrap looks similar to the size distortion for the Hansen tests. Residual bootstrap also works well for the Sargan tests in the homoscedastic case. For the heteroscedastic case, the HH bootstrap is the only good option for the Sargan tests. Here also the 'power percentage distortion' looks similar to the 'size percentage distortion'.

4.5 Conclusions

In this Chapter we analyzed and compared several versions of overidentifying restriction tests. They are based on the GMM criterion function and GEL type Likelihood Ratio tests. We used different bootstrap schemes to improve on GMM type tests. For a simple linear homoscedastic model we saw that the Sargan tests perform very well in terms of power and when supported with the residual bootstrap their size distortion does not exceed 10%. Hansen tests are best corrected for size by the residual bootstrap.

For the heteroscedastic case the residual bootstrap does not work for the Sargan test, but the HH (2 step) bootstrap does work well. In the same circumstances, HH bootstrap performs better for Sargan than for Hansen tests. The residual bootstrap does best for the Hansen tests.

The BN bootstrap does not perform well in the examples we considered. Also the LR type tests have size problems, which can be fixed possibly by using bootstrap techniques. Size corrected versions do not perform better than Sargan tests in the homoscedastic case, but for the heteroscedastic case they show some potential (but with no obvious winner among the LR type tests). There is no clear winner for the heteroscedastic case anyway, and together with rather substantial size problems, the GMM (Sargan) tests seem to be the safest option for the model and circumstances we analyzed.

Kitamura (2001) analyzes the performance of the EL and Hansen tests in a non-linear example from Hall and Horowitz (1996). It shows that EL performs better than GMM tests. Here we demonstrate that in a linear model the Sargan test can be a better option.

instr.	S_n^s cv_∞	cv^*	cv^{**}	cv_R	S_n^a cv_∞	cv^*	cv^{**}	cv_R
			$\mu_p/3 = 20$		$n = 50$			
[123]	-0.02	-0.1	-0.089	0.084	0.38	-0.1	-0.098	0.083
[12]	-0.023	-0.054	-0.0082	0.074	0.16	-0.055	-0.013	0.074
[123]|[23]	0.052	-0.073	-0.071	0.045	0.35	-0.089	-0.092	0.04
			$\mu_p/3 = 1$		$n = 50$			
[123]	-0.3	-0.25	-0.16	-0.082	0.0088	-0.25	-0.17	-0.083
[12]	-0.34	-0.22	-0.12	-0.11	-0.21	-0.22	-0.12	-0.11
[123]|[23]	-0.059	-0.26	-0.16	-0.1	0.19	-0.27	-0.17	-0.11
			$\mu_p/3 = 20$		$n = 200$			
[123]	-0.02	0.04	0.071	0.085	0.064	0.04	0.071	0.085
[12]	-0.042	0.026	0.073	0.075	0.001	0.026	0.073	0.075
[123]|[23]	-0.0028	0.04	0.044	0.053	0.064	0.038	0.042	0.051
			$\mu_p/3 = 1$		$n = 200$			
[123]	-0.31	-0.17	-0.079	-0.086	-0.24	-0.17	-0.08	-0.086
[12]	-0.36	-0.2	-0.087	-0.095	-0.33	-0.2	-0.088	-0.095
[123]|[23]	-0.094	-0.18	-0.1	-0.1	-0.038	-0.18	-0.1	-0.11

instr.	J_n^s cv_∞	cv^*	cv^{**}	cv_R	J_n^a cv_∞	cv^*	cv^{**}	cv_R
			$\mu_p/3 = 20$		$n = 50$			
[123]	-0.18	-0.36	-0.3	0.07	0.2	-0.37	-0.31	0.068
[12]	-0.11	-0.19	-0.11	0.074	0.076	-0.2	-0.11	0.074
[123]|[23]	-0.024	-0.3	-0.25	0.037	0.27	-0.33	-0.28	0.037
			$\mu_p/3 = 1$		$n = 50$			
[123]	-0.42	-0.47	-0.35	-0.071	-0.14	-0.49	-0.36	-0.072
[12]	-0.4	-0.34	-0.19	-0.092	-0.27	-0.35	-0.2	-0.094
[123]|[23]	-0.12	-0.45	-0.32	-0.088	0.12	-0.47	-0.34	-0.093
			$\mu_p/3 = 20$		$n = 200$			
[123]	-0.06	0.012	0.048	0.085	0.028	0.011	0.047	0.085
[12]	-0.055	0.012	0.061	0.075	-0.015	0.012	0.06	0.075
[123]|[23]	-0.011	0.0062	0.021	0.043	0.057	0.0038	0.018	0.044
			$\mu_p/3 = 1$		$n = 200$			
[123]	-0.35	-0.2	-0.1	-0.086	-0.28	-0.2	-0.1	-0.087
[12]	-0.38	-0.22	-0.11	-0.087	-0.35	-0.22	-0.11	-0.088
[123]|[23]	-0.11	-0.22	-0.14	-0.099	-0.054	-0.22	-0.14	-0.099

Table 4.3: Size distortion: For different n and $\mu_p/3$, the table presents $\frac{P(T_n > cv) - \alpha}{\alpha}$ (relative deviation of the actual rejection probabilities from the nominal level $\alpha = 0.05$), where T_n is either S_n^s, S_n^a in the upper part of the table or J_n^s, J_n^a in the lower part. cv stands for the different critical values (at the given significance α level used): cv_∞ asymptotic critical value, cv^* based on 'one step' version of Hall and Horowitz (1996), cv^{**} based on 'two steps', cv_R is based on the residual type bootstrap alternative.

instr.	S_n^s cv_∞	cv^*	cv^{**}	cv_R	S_n^a cv_∞	cv^*	cv^{**}	cv_R
			$\mu_p/3 = 20$		$n = 50$			
[123]	2.6	-0.32	-0.34	2.7	3.4	-0.32	-0.36	2.7
[12]	1.7	-0.23	-0.2	1.8	2	-0.23	-0.21	1.8
[123]|[23]	1.9	-0.24	-0.25	1.8	2.5	-0.29	-0.31	1.9
			$\mu_p/3 = 1$		$n = 50$			
[123]	1.5	-0.4	-0.3	1.8	2.1	-0.41	-0.31	1.8
[12]	0.79	-0.32	-0.21	1	1	-0.32	-0.21	1
[123]|[23]	1.4	-0.4	-0.28	1.2	1.9	-0.43	-0.3	1.3
			$\mu_p/3 = 20$		$n = 200$			
[123]	2.8	-0.02	0.013	3	3	-0.02	0.011	3
[12]	1.9	-0.0011	0.065	2	1.9	-0.0014	0.064	2
[123]|[23]	2	-0.0019	0.005	2	2.1	-0.0098	0.0011	2
			$\mu_p/3 = 1$		$n = 200$			
[123]	1.7	-0.2	-0.084	2	1.8	-0.2	-0.086	2
[12]	0.85	-0.21	-0.085	1.2	0.91	-0.21	-0.086	1.2
[123]|[23]	1.6	-0.21	-0.11	1.4	1.7	-0.22	-0.11	1.4

instr.	J_n^s cv_∞	cv^*	cv^{**}	cv_R	J_n^a cv_∞	cv^*	cv^{**}	cv_R
			$\mu_p/3 = 20$		$n = 50$			
[123]	-0.36	-0.76	-0.73	-0.013	-0.0099	-0.77	-0.74	-0.022
[12]	-0.21	-0.52	-0.46	0.0087	-0.04	-0.52	-0.46	0.0048
[123]|[23]	-0.084	-0.67	-0.63	-0.017	0.2	-0.7	-0.67	-0.03
			$\mu_p/3 = 1$		$n = 50$			
[123]	-0.55	-0.77	-0.69	-0.19	-0.29	-0.78	-0.7	-0.2
[12]	-0.48	-0.56	-0.43	-0.19	-0.35	-0.56	-0.44	-0.19
[123]|[23]	-0.19	-0.72	-0.63	-0.18	0.053	-0.75	-0.66	-0.18
			$\mu_p/3 = 20$		$n = 200$			
[123]	-0.16	-0.3	-0.28	0.028	-0.063	-0.3	-0.28	0.027
[12]	-0.084	-0.15	-0.098	0.059	-0.039	-0.15	-0.099	0.058
[123]|[23]	-0.045	-0.25	-0.22	0.00037	0.023	-0.25	-0.23	-0.0038
			$\mu_p/3 = 1$		$n = 200$			
[123]	-0.41	-0.41	-0.33	-0.17	-0.35	-0.41	-0.33	-0.17
[12]	-0.41	-0.32	-0.21	-0.17	-0.38	-0.32	-0.21	-0.17
[123]|[23]	-0.15	-0.4	-0.31	-0.18	-0.087	-0.4	-0.31	-0.18

Table 4.4: Size distortion: Same as Table 4.3, but with conditional heteroscedasticity where $\gamma = 0.9$

	S_n^a		J_n^a					
instr.	cv_{EL}	cv_{ET}	cv_{EL}	cv_{ET}	$ELR(\hat{\theta})$	$ETR(\hat{\theta})$	ELR	ETR
$\gamma = 1$								
$\mu_p/3 = 20$					$n = 50$			
[123]	-0.51	-0.49	-0.78	-0.74	0.85	0.6	0.75	0.52
[12]	-0.65	-0.6	-0.79	-0.77	0.37	0.27	0.28	0.21
[123]|[23]	-0.71	-0.68	-0.84	-0.83	0.78	0.58	0.73	0.54
$\mu_p/3 = 1$					$n = 50$			
[123]	-0.67	-0.62	-0.86	-0.85	0.35	0.15	-0.45	-0.56
[12]	-0.77	-0.75	-0.88	-0.87	-0.054	-0.13	-0.61	-0.65
[123]|[23]	-0.74	-0.72	-0.87	-0.86	0.52	0.36	-0.14	-0.26
$\mu_p/3 = 20$					$n = 200$			
[123]	-0.53	-0.5	-0.57	-0.58	0.14	0.14	0.11	0.11
[12]	-0.68	-0.65	-0.71	-0.7	0.04	0.042	0.013	0.017
[123]|[23]	-0.72	-0.7	-0.74	-0.74	0.14	0.14	0.12	0.12
$\mu_p/3 = 1$					$n = 200$			
[123]	-0.69	-0.68	-0.71	-0.72	-0.19	-0.19	-0.73	-0.73
[12]	-0.79	-0.79	-0.81	-0.8	-0.31	-0.31	-0.74	-0.74
[123]|[23]	-0.76	-0.75	-0.77	-0.77	0.011	0.0096	-0.51	-0.51
$\gamma = 0.9$								
$\mu_p/3 = 20$					$n = 50$			
[123]	-0.36	-0.27	-0.91	-0.86	1.7	1	1.6	0.88
[12]	-0.54	-0.47	-0.89	-0.87	0.87	0.52	0.72	0.42
[123]|[23]	-0.6	-0.55	-0.93	-0.89	1.5	0.95	1.4	0.89
$\mu_p/3 = 1$					$n = 50$			
[123]	-0.54	-0.47	-0.95	-0.92	1	0.44	-0.21	-0.47
[12]	-0.7	-0.64	-0.93	-0.91	0.28	0.03	-0.5	-0.6
[123]|[23]	-0.65	-0.6	-0.94	-0.92	1.1	0.68	0.13	-0.12
$\mu_p/3 = 20$					$n = 200$			
[123]	-0.46	-0.41	-0.69	-0.68	0.55	0.39	0.48	0.34
[12]	-0.62	-0.58	-0.76	-0.75	0.28	0.2	0.22	0.16
[123]|[23]	-0.67	-0.64	-0.8	-0.8	0.51	0.37	0.46	0.35
$\mu_p/3 = 1$					$n = 200$			
[123]	-0.61	-0.58	-0.75	-0.74	0.088	-0.028	-0.65	-0.67
[12]	-0.74	-0.71	-0.81	-0.82	-0.15	-0.19	-0.68	-0.7
[123]|[23]	-0.72	-0.69	-0.78	-0.78	0.28	0.18	-0.39	-0.43

Table 4.5: Brown and Newey (2002) bootstrap: relative size distortion: $\frac{P(T_n > cv) - \alpha}{\alpha}$; for different T_n with the critical values cv_{EL}, cv_{ET} obtained applying Empirical Likelihood (4.40) or Exponential Tilting (4.41) implied probabilities. Last 4 columns show relative distortion of $ELR(\hat{\theta})$, $ETR(\hat{\theta})$, ELR and ELR type Likelihood Ratio tests.

instr.	S_n^s	S_n^a	J_n^s	J_n^a	$ELR(\hat{\theta})$	$ETR(\hat{\theta})$	ELR	ETR
				$\mu_p = 20$				
				$\rho_{Z_1u} = 0.1$				
[123]	**0.0727**	**0.0727**	0.0697	0.0697	0.0711	0.0708	0.0726	0.0716
[12]	0.0751	0.0751	0.0734	0.0734	0.0746	0.0741	**0.0763**	0.075
[123]|[12]	0.0588	**0.0598**	0.0569	0.0577	0.0596	0.0585	0.0596	0.0596
[123]|[23]	**0.0834**	0.0829	0.081	0.0809	0.0818	0.0815	0.0833	0.0831
				$\rho_{Z_1u} = 0.2$				
[123]	**0.151**	**0.151**	0.138	0.138	0.143	0.143	0.145	0.145
[12]	**0.155**	**0.155**	0.148	0.147	0.152	0.152	0.154	0.153
[123]|[12]	0.087	0.093	0.0779	0.0823	0.0921	0.0874	**0.094**	0.09
[123]|[23]	**0.192**	0.191	0.176	0.176	0.184	0.184	0.187	0.187
				$\rho_{Z_1u} = 0.5$				
[123]	**0.718**	**0.718**	0.646	0.645	0.675	0.675	0.674	0.673
[12]	**0.615**	**0.615**	0.579	0.579	0.593	0.591	0.583	0.582
[123]|[12]	0.339	**0.41**	0.23	0.274	0.376	0.34	0.395	0.36
[123]|[23]	0.797	**0.797**	0.733	0.735	0.761	0.76	0.758	0.759
				$\mu_p = 1$				
				$\rho_{Z_1u} = 0.1$				
[123]	**0.0648**	**0.0648**	0.0639	0.0639	0.0642	0.064	0.0647	0.064
[12]	0.0662	0.0662	0.0653	0.0655	0.0664	**0.0669**	0.0668	0.0666
[123]|[12]	0.0575	0.058	0.0558	0.0566	0.0583	**0.0587**	0.0583	0.0578
[123]|[23]	0.071	0.0707	0.0693	0.0691	0.0709	**0.0714**	0.0707	0.0697
				$\rho_{Z_1u} = 0.2$				
[123]	**0.122**	**0.122**	0.115	0.115	0.118	0.118	0.111	0.109
[12]	**0.122**	**0.122**	0.118	0.118	0.121	0.122	0.115	0.115
[123]|[12]	0.0853	0.0882	0.08	0.0824	**0.089**	0.0878	0.0842	0.0833
[123]|[23]	**0.147**	0.147	0.137	0.137	0.142	0.143	0.131	0.131
				$\rho_{Z_1u} = 0.5$				
[123]	**0.43**	**0.43**	0.397	0.396	0.412	0.411	0.303	0.302
[12]	**0.328**	**0.328**	0.317	0.316	0.323	0.323	0.26	0.259
[123]|[12]	0.241	**0.264**	0.204	0.22	0.256	0.242	0.193	0.187
[123]|[23]	**0.485**	0.484	0.45	0.45	0.467	0.468	0.354	0.352

Table 4.6: Size corrected power: $P(T_n > c_\alpha|\rho_{Z_1u})$; $\alpha = 5\%$, n=50, $\gamma = 1$.

instr.	S_n^s	S_n^a	J_n^s	J_n^a	$ELR(\hat{\theta})$	$ETR(\hat{\theta})$	ELR	ETR
				$\mu_p = 20$				
				$\rho_{Z_1u} = 0.1$				
[123]	**0.156**	**0.156**	0.153	0.153	0.153	0.152	0.155	0.154
[12]	**0.163**	**0.163**	0.161	0.161	0.159	0.16	0.161	0.161
[123]|[12]	0.0895	0.0908	0.0869	0.0883	0.0906	0.0904	**0.0914**	0.091
[123]|[23]	**0.205**	0.204	0.199	0.199	0.196	0.198	0.197	0.199
				$\rho_{Z_1u} = 0.2$				
[123]	**0.502**	**0.502**	0.489	0.489	0.491	0.491	0.493	0.493
[12]	**0.465**	**0.465**	0.458	0.458	0.457	0.458	0.458	0.459
[123]|[12]	0.226	**0.236**	0.207	0.216	0.229	0.227	0.234	0.231
[123]|[23]	**0.607**	0.607	0.593	0.593	0.589	0.593	0.59	0.594
				$\rho_{Z_1u} = 0.5$				
[123]	**0.996**	**0.996**	0.994	0.994	0.994	0.994	0.993	0.993
[12]	**0.949**	**0.949**	0.945	0.945	0.945	0.945	0.94	0.94
[123]|[12]	0.875	**0.902**	0.772	0.808	0.857	0.858	0.864	0.874
[123]|[23]	0.998	**0.998**	0.997	0.997	0.997	0.997	0.997	0.997
				$\mu_p = 1$				
				$\rho_{Z_1u} = 0.1$				
[123]	**0.124**	**0.124**	0.122	0.122	0.12	0.12	0.112	0.112
[12]	**0.122**	**0.122**	0.121	0.121	0.119	0.119	0.111	0.111
[123]|[12]	0.0866	**0.0873**	0.0851	0.0857	0.086	0.0858	0.0846	0.0836
[123]|[23]	0.151	**0.151**	0.147	0.147	0.145	0.145	0.131	0.132
				$\rho_{Z_1u} = 0.2$				
[123]	**0.322**	**0.322**	0.315	0.315	0.314	0.315	0.251	0.252
[12]	**0.266**	**0.266**	0.263	0.263	0.261	0.263	0.217	0.217
[123]|[12]	0.186	**0.189**	0.179	0.182	0.188	0.188	0.162	0.161
[123]|[23]	**0.381**	0.381	0.373	0.373	0.374	0.374	0.301	0.301
				$\rho_{Z_1u} = 0.5$				
[123]	**0.479**	**0.479**	0.474	0.474	0.474	0.476	0.365	0.366
[12]	**0.332**	**0.332**	0.329	0.329	0.327	0.328	0.287	0.287
[123]|[12]	0.342	**0.349**	0.324	0.331	0.343	0.342	0.234	0.232
[123]|[23]	**0.533**	0.533	0.527	0.527	0.528	0.529	0.422	0.422

Table 4.7: Size corrected power: $P(T_n > c_\alpha | \rho_{Z_1u})$; $\alpha = 5\%$, n=200, $\gamma = 1$.

instr.	S_n^s	S_n^a	J_n^s	J_n^a	$ELR(\hat{\theta})$	$ETR(\hat{\theta})$	ELR	ETR
				$\mu_p = 20$				
				$\rho_{Z_1\tilde{u}} = 0.1$				
[123]	**0.0622**	**0.0622**	0.0593	0.0591	0.0599	0.0594	0.0608	0.0604
[12]	0.0637	0.0637	0.0605	0.0606	0.0619	0.0625	0.0629	**0.0641**
[123]|[12]	0.0544	**0.0559**	0.0528	0.0531	0.0538	0.0538	0.054	0.0544
[123]|[23]	**0.0684**	0.0679	0.0636	0.0639	0.0652	0.0646	0.0663	0.0659
				$\rho_{Z_1\tilde{u}} = 0.2$				
[123]	**0.102**	**0.102**	0.0896	0.0894	0.0951	0.0928	0.0961	0.0945
[12]	**0.106**	**0.106**	0.0964	0.0964	0.102	0.102	0.103	0.104
[123]|[12]	0.0682	**0.0735**	0.0624	0.0644	0.0697	0.0683	0.0706	0.0685
[123]|[23]	**0.124**	0.124	0.107	0.107	0.116	0.115	0.117	0.117
				$\rho_{Z_1\tilde{u}} = 0.5$				
[123]	0.441	0.441	0.375	0.373	**0.446**	0.441	0.441	0.44
[12]	0.387	0.387	0.372	0.371	**0.412**	0.41	0.398	0.406
[123]|[12]	0.192	0.247	0.139	0.156	0.239	0.214	**0.253**	0.224
[123]|[23]	0.526	0.528	0.463	0.464	**0.534**	0.532	0.526	0.529
				$\mu_p = 1$				
				$\rho_{Z_1\tilde{u}} = 0.1$				
[123]	0.0561	0.0561	0.0557	0.0556	0.0548	0.0552	**0.0576**	0.0575
[12]	0.0566	0.0566	0.0569	0.0568	0.0564	0.0568	0.0588	**0.0592**
[123]|[12]	0.053	0.0536	0.0522	0.0524	0.0529	0.0526	**0.0548**	0.0537
[123]|[23]	0.0585	0.0588	0.0577	0.0578	0.0576	0.0579	0.0601	**0.0611**
				$\rho_{Z_1\tilde{u}} = 0.2$				
[123]	**0.0825**	**0.0825**	0.0781	0.0779	0.0791	0.0792	0.0812	0.0821
[12]	0.0828	0.0828	0.082	0.0819	0.0833	0.084	0.0836	**0.0841**
[123]|[12]	0.0655	0.0681	0.0628	0.0636	0.0664	0.0653	**0.069**	0.0684
[123]|[23]	0.0936	**0.0942**	0.089	0.0887	0.0906	0.0917	0.0916	0.0934
				$\rho_{Z_1\tilde{u}} = 0.5$				
[123]	0.263	0.263	0.26	0.258	**0.284**	0.284	0.223	0.233
[12]	0.221	0.221	0.24	0.239	0.244	**0.246**	0.204	0.208
[123]|[12]	0.15	0.167	0.139	0.146	**0.179**	0.169	0.149	0.149
[123]|[23]	0.308	0.309	0.306	0.306	0.33	**0.33**	0.261	0.271

Table 4.8: Size corrected power: $P(T_n > c_\alpha|\rho_{Z_1\tilde{u}})$; $\alpha = 5\%$, n=50, $\gamma = 0.9$.

instr.	S_n^s	S_n^a	J_n^s	J_n^a	$ELR(\hat{\theta})$	$ETR(\hat{\theta})$	ELR	ETR
				$\mu_p = 20$				
				$\rho_{Z_1\tilde{u}} = 0.1$				
[123]	**0.104**	**0.104**	0.1	0.1	0.102	0.102	0.103	0.103
[12]	**0.108**	**0.108**	0.106	0.106	0.105	0.105	0.107	0.106
[123]|[12]	0.0711	**0.0724**	0.0678	0.0684	0.0713	0.0697	0.0712	0.0703
[123]|[23]	**0.129**	0.129	0.124	0.124	0.125	0.126	0.126	0.127
				$\rho_{Z_1\tilde{u}} = 0.2$				
[123]	0.284	0.284	0.274	0.274	0.281	0.284	0.282	**0.285**
[12]	0.273	0.273	0.273	0.273	0.274	0.275	0.275	**0.277**
[123]|[12]	0.141	**0.149**	0.124	0.127	0.144	0.139	0.146	0.142
[123]|[23]	0.361	0.361	0.349	0.35	0.355	0.36	0.357	**0.363**
				$\rho_{Z_1\tilde{u}} = 0.5$				
[123]	0.93	0.93	0.946	0.946	0.954	**0.954**	0.947	0.951
[12]	0.799	0.799	0.854	0.854	**0.859**	0.858	0.845	0.848
[123]|[12]	0.654	**0.706**	0.491	0.52	0.689	0.669	0.69	0.683
[123]|[23]	0.96	0.961	0.967	0.967	0.974	**0.975**	0.97	0.972
				$\mu_p = 1$				
				$\rho_{Z_1\tilde{u}} = 0.1$				
[123]	0.0823	0.0823	**0.0835**	0.0834	0.0827	0.0832	0.0817	0.0813
[12]	0.0821	0.0821	0.0851	0.085	0.0849	**0.0852**	0.0842	0.085
[123]|[12]	0.0668	0.0675	0.0665	0.0669	0.0676	0.0676	**0.0697**	0.0689
[123]|[23]	0.0954	0.0954	**0.0976**	0.0974	0.0964	0.0965	0.0927	0.0938
				$\rho_{Z_1\tilde{u}} = 0.2$				
[123]	0.178	0.178	**0.195**	0.195	0.192	0.194	0.17	0.172
[12]	0.161	0.161	**0.18**	0.18	0.177	0.178	0.16	0.162
[123]|[12]	0.116	0.119	0.118	0.119	**0.126**	0.125	0.117	0.117
[123]|[23]	0.216	0.216	0.237	0.237	0.237	**0.238**	0.206	0.211
				$\rho_{Z_1\tilde{u}} = 0.5$				
[123]	0.317	0.317	0.429	**0.429**	0.423	0.425	0.322	0.329
[12]	0.224	0.224	0.307	**0.307**	0.301	0.303	0.264	0.269
[123]|[12]	0.221	0.229	0.265	0.271	**0.293**	0.289	0.206	0.21
[123]|[23]	0.365	0.365	0.481	**0.481**	0.477	0.48	0.379	0.388

Table 4.9: Size corrected power: $P(T_n > c_\alpha | \rho_{Z_1\tilde{u}})$; $\alpha = 5\%$, n=200, $\gamma = 0.9$.

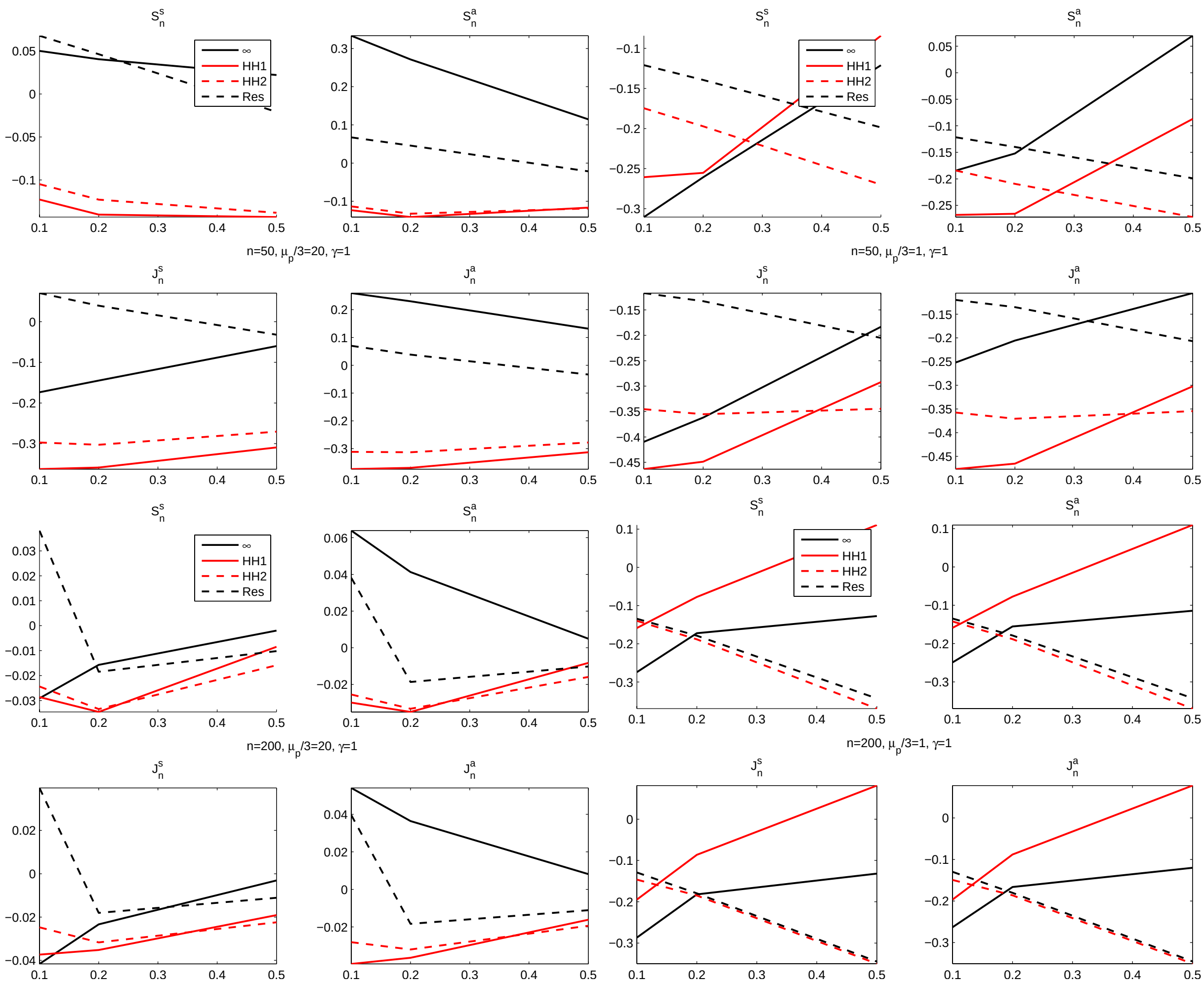

Figure 4.4: Percentage deviation of rejection frequencies from the ‘size corrected’ power when using different critical values: asymptotic (∞), 1,2 step HH, and based on the residual bootstrap; $\gamma = 1$.

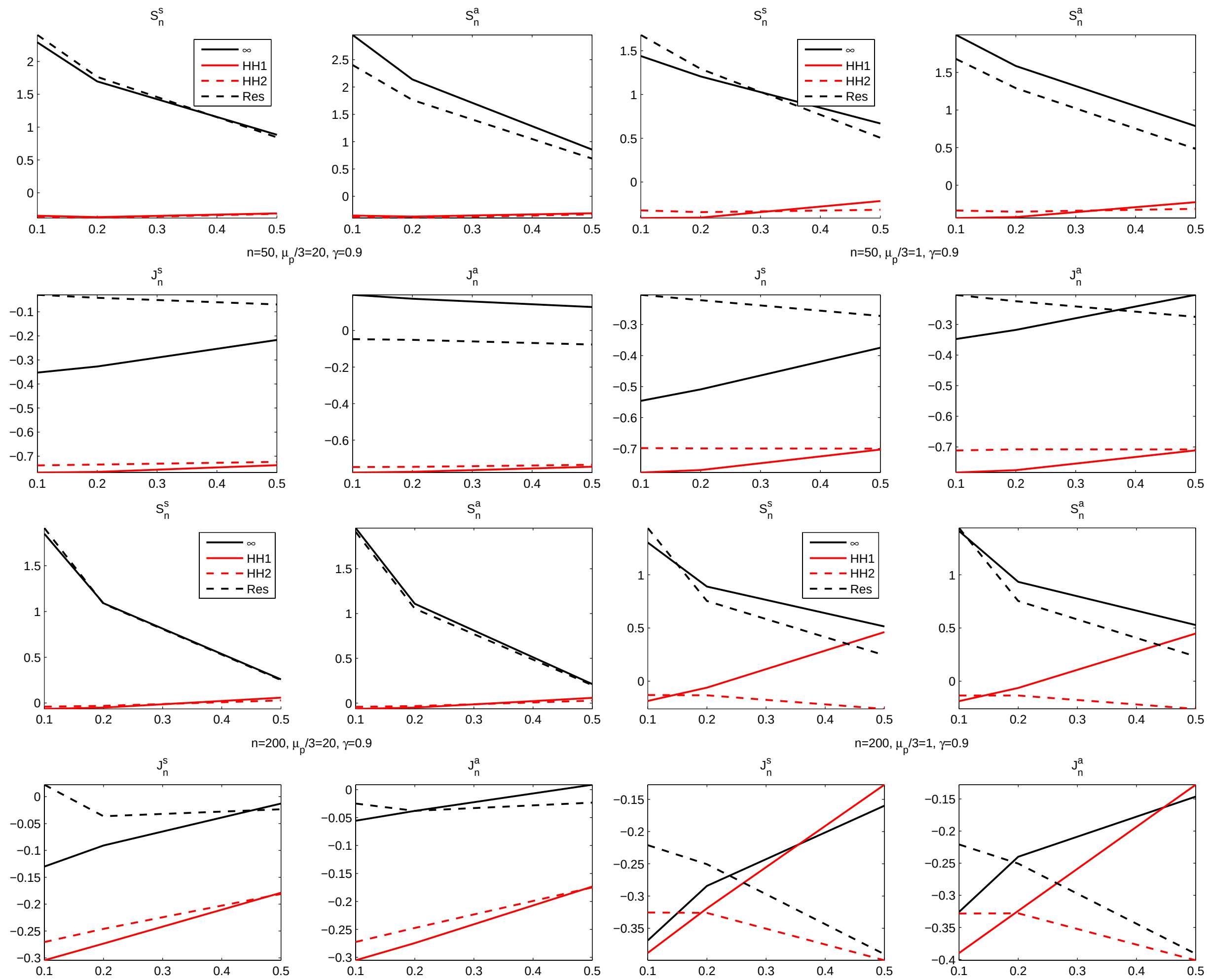

Figure 4.5: Percentage deviation of rejection frequencies from the ‘size corrected’ power when using different critical values: asymptotic (∞), 1,2 step HH, and based on the residual bootstrap; $\gamma = 0.9$.