GREDI

Groupe de Recherche en Économie
et Développement International

Cahier de Recherche / Working Paper
14-06

Bootstrap tests in linear models with many regressors

Patrick Richard

UNIVERSITÉ DE
SHERBROOKE

Bootstrap tests in linear models with many regressors

Patrick Richard
Département de sciences économiques
Faculté d'administration, Université de Sherbrooke
2500 Boulevard de l'Université
Sherbrooke, Québec, Canada
J1K 2R1
patrick.richard2@usherbrooke.ca
P:(819) 821-80000 ext. 63233
F:(819) 821-7934

Abstract

This paper is concerned with bootstrap hypothesis testing in high dimensional linear regression models. Using a theoretical framework recently introduced by Anatolyev (2012), we show that bootstrap F, LR and LM tests are asymptotically valid even when the numbers of estimated parameters and tested restrictions are not asymptotically negligible fractions of the sample size. These results are derived for models with iid error terms, but Monte Carlo evidence suggests that they extend to the wild bootstrap in the presence of heteroskedasticity and to bootstrap methods for heavy tailed data.

Keywords: bootstrap, linear regressions, high dimension.

JEL codes: C12; C14; C15.

August 2014

I am grateful to Victoria Zinde-Walsh, Russell Davidson and participants at the CEA 2014 Vancouver conference for helpful discussions and comments.

1 Introduction

Hypothesis testing in linear regression models is a fundamental part of statistical inference and is conducted in virtually every applied econometrics study. The most common approach consists of using critical values obtained from (or calculating a P value using) an asymptotic approximation to the test statistics' null distribution. Most conditions for the validity of this approach, such as basic properties of the error terms and regressors, are well known and understood by applied researchers. For instance, any serious economist knows that a strongly skewed error distribution or one with very fat tails compromises the finite sample accuracy of first order asymptotic approximations.

However, not all necessary conditions are as well known. For instance, while most applied researchers are wary of using a large number of explanatory variables in a regression model for fear of increasing estimation error, very few are aware of the size distortion problems related with conventional asymptotic testing in this situation. Indeed, concrete steps to mitigate potential size problems related to this issue rarely appear in econometrics textbooks.

Yet simple intuition warns against blindly trusting conventional asymptotic approximations when carrying-out hypothesis testing in such circumstances. Indeed, standard theory assumes that both the numbers of parameters to be estimated and restrictions to be tested are asymptotically negligible quantities compared to the sample size. What we might call the "high dimensionality" of the testing problem, that is, the fact that these quantities are not negligible in finite samples, is therefore abstracted from asymptotically.

Consequences of this are well documented in the theoretical econometrics literature. For instance Evans and Savin (1982) find substantial size distortions for the Wald (W), Likelihood Ratio (LR) and Lagrange Multiplier (LM) tests in linear regression models when there are many parameters and restrictions and propose some simple corrections to improve finite sample testing accuracy. Rothenberg (1984) also notices that Wald tests based on chi-square critical values are likely to be quite inaccurate in small samples and proposes second order Edgeworth corrected versions of the W, LR and LM statistics. Important work on the consistency and asymptotic distribution of M and likelihood estimators in high dimension models has been done by, among others, Portnoy (1984, 1985, 1988), Mammen (1989) and Koenker and Machado (1999). With regards to testing, the main conclusion of this literature is that standard asymptotic approximations remain valid, provided the number of parameters and/or restrictions grows no faster than a given rate of the sample size. Yet, that assumption, while it allows for parameters and restrictions to be infinitely numerous asymptotically, still assumes that the number of degrees of freedom available to test each restriction tends to infinity. That still misrepresents the finite sample reality of large dimensional regression models.

Using an alternative framework, to be fully described in the next section, within which the asymptotic ratios of the number of regressors and the number of restrictions to the sample size are allowed to be non-zero constants, Anatolyev (2012) shows that the W, LR and LM tests are wrongly sized asymptotically in the presence of many restrictions, while the F test using critical values from Fisher's distribution have correct size. He derives the actual asymptotic size distortions of the W, LR and LM tests, and proposes alternative versions that have correct asymptotic size.

Anatolyev's (2012) analysis is purely asymptotic, meaning that he does not provide much insight into the finite sample accuracy of his corrected tests. The first part of the present paper seeks to fill this gap by providing extensive Monte Carlo evidence. This is done in section 2, and it is found that Anatolyev's tests have good finite sample properties, except in very extreme cases (ie, with a very large number of restrictions and / or very small samples).

The second contribution of this paper is to derive the asymptotic properties of the bootstrap F, LR and LM tests under the many regressors and many restrictions framework. We consider both the parametric and semiparametric bootstraps.[1]

The statistics literature contains some interesting theoretical work on the bootstrap in linear regression models with many regressors. Early key papers are Bickel and Freedman (1983) and Mammen (1993) while a more recent contribution is Chatterjee and Bose (2002). These papers are mostly concerned with estimating the distribution of a vector of contrast of the OLS estimator when the number of regressors increases with the sample size, though Mammen (1993) also considers the distribution of the F test. A common feature of these papers is that they allow the number of estimated parameters to tend to infinity, as long as the ratio of some power of the number of parameters to the sample size tends to 0. Under these assumptions, they establish the asymptotic validity of different types of bootstraps (iid resampling bootstrap, pairs bootstrap and wild bootstrap).

[1]By parametric bootstrap, we mean that the bootstrap DGP uses draws from a fully parametric distributions (usually the Normal), with parameters estimated from the sample. The semiparametric bootstrap is the more usual resampling bootstrap, wherein bootstrap error terms are drawn with replacement from the empirical distribution function of the residuals.

Our theoretical analysis, which is contained in section 3, employs Anatolyev's (2012) framework, which permits the ratio of the number of parameters to the sample size, and that of the number of tested restrictions to the sample size, to be non-zero asymptotically. Assuming exogenous regressors and iid errors, we show that the bootstrap versions of the F, LR and LM tests are valid, in the sense that their respective asymptotic distributions are the same as those of the sample statistics derived by Anatolyev (2012). Monte Carlo simulations confirm this and show that the bootstrap often provides more accurate tests than the asymptotic corrections of Anatolyev (2012).

In the fourth section of the paper, further Monte Carlo simulations are used to explore the robustness of the bootstrap and asymptotic tests to departures from key assumptions of Anatolyev's (2012) framework. Special attention is given to heavy tailed and highly skewed distributions and to heteroskedasticity. Results indicate that using the appropriate bootstrap method yields valid tests.

2 Monte Carlo for asymptotic tests

This section provides some Monte Carlo results to illustrate the small sample accuracy of the corrected tests proposed by Anatolyev (2012) and compare them with other versions of the F (or Wald), LR and LM tests. Subsection 2.1 describes the relevant asymptotic framework. Subsection 2.2 introduces the test statistics and considers the situation of many regressors but few restrictions, while subsection 2.3 considers cases where there are many restrictions being tested.

2.1 The asymptotic framework

Consider the linear regression model

$$y = X\beta + u, \tag{1}$$

where y and u are $n \times 1$ vectors and X is a $n \times m$ matrix of exogenous regressors. It is assumed that the error terms are iid with variance σ^2. The test being considered is that of the r restrictions

$$H_0 : R\beta = q,$$

against

$$H_1 : R\beta \neq q,$$

where R is a $r \times m$ matrix, q is a $r \times 1$ vector and $r \leq m$. The ratio of the number of regressors to the sample size and that of the number of restrictions to the sample size will be of considerable interest for the following analysis. Anatolyev (2012) suggests an asymptotic framework in which, contrarily to what is found in the literature cited in the introduction, these two quantities do not converge to 0 as n increases. His key assumptions in this respect are:

Assumption 1. a) As $n \to \infty$, $m/n = \mu + o(r^{-1/2})$, where μ is a constant such that $0 < \mu < 1$.

Assumption 1. b) As $n \to \infty$, either r is fixed (so that $r/n \to 0$) or $r/n = \rho + o(r^{-1/2})$, where ρ is a constant such that $0 < \rho \leq \mu$.

In what follows, we refer to the case where $\mu > 0$ and $\rho = 0$ as the "many regressors" case. Likewise, "many restrictions" refers to the case where $\rho > 0$. Another very important quantity is the asymptotic ratio of the number of restrictions to the number of degrees of freedom, denoted by λ:

$$\lambda = \frac{\rho}{1-\mu}.$$

Its empirical counterpart is

$$\hat{\lambda} = \frac{r}{n-m}.$$

Further statistical assumptions about the elements of model (1) are

Assumption 2. $E\left(u_i^4\right) = \kappa < \infty$, where u_i is element i of the vector u.

Assumption 3. Let $\Xi_I = (X^\top X)^{-1}$ and $\Xi_R = (X^\top X)^{-1} R^\top \left(R(X^\top X)^{-1} R^\top\right)^{-1} R(X^\top X)^{-1}$. Then,

$$\max_{1 \leq i \leq n} |X_i \Xi_I X_i^\top - \mu| \to 0,$$

and

$$\max_{1 \leq i \leq n} |X_i \Xi_R X_i^\top - \rho| \to 0,$$

where X_i is row i of the matrix X.

Assumption 2 is quite common in econometrics and needs no special explanation. Assumption 3 is a bit less intuitive, but it is crucial for the derivation of Anatolyev's (2012) main results. One way to look at it is through the leverage observation i has on the restricted and unrestricted OLS estimators of β. Indeed, notice that the term $X_i \Xi_I X_i^\top$ in the first part of the assumption is the i^{th} element of the diagonal of $P_X = X(X^\top X)^{-1}X^\top$. It is a well-known fact that the sum of these diagonal elements is m and that their average is m/n (see Davidson and MacKinnon, 2004, chapter 2). Thus, the first part of assumption 3 simply states that no single observation has excessive leverage asymptotically on the unrestricted OLS estimator. A similar argument can be made about the second part of the assumption. The appendix in Anatolyev (2012) provides a complete discussion of this assumption and some evidence to the effect that it is not overly restrictive.

2.2 Tests with many regressors and few restrictions.

With the asymptotic framework described by assumptions 1, 2 and 3, Anatolyev (2012) derives several results for the F, LR and LM tests. The Wald test is not studied because it is obtained by multiplying the rF statistic by $n/(n-m)$, so any result for the F test also applies to it. In all that follows, the expression "exact F test", which will sometimes be denoted by EF, refers to a F test carried-out using critical values from the $F(r, n-m)$ distribution. Assuming first that $\rho = 0$, Anatolyev (2012) establishes the following theorem.

Theorem 1 (Anatolyev, 2012). *Under assumptions 1, 2 and 3 and assuming that $\rho = 0$, if H_0 is true, then*

$$rF \overset{a}{\sim} \chi^2(r),$$
$$\left(1 - \frac{m}{n}\right) LR \overset{a}{\sim} \chi^2(r),$$
$$\left(1 - \frac{m}{n}\right) LM \overset{a}{\sim} \chi^2(r),$$

also, the exact F test is asymptotically valid.

Thus, according to theorem 1, simple degree of freedom adjustments to the LM and LR tests suffice to obtain the usual asymptotic chi-square distribution while no adjustment at all is necessary for the rF and EF tests. This makes intuitive sense since the F statistic uses a degree of freedom corrected estimator of the errors' variance, while the LR and LM statistics do not. To illustrate this, consider the LM statistic

$$LM = (R\hat{\beta} - q)^\top \left[\frac{\tilde{u}^\top \tilde{u}}{n} R(X^\top X)^{-1} R^\top\right]^{-1} (R\hat{\beta} - q), \tag{2}$$

where $\hat{\beta}$ is the unrestricted OLS estimator and $\tilde{u}$ is the vector of residuals from OLS estimation of the restricted model. It is well known that $\tilde{u}^\top \tilde{u}/n$ is a biased estimator of the error terms variance and that, under H_0, an unbiased estimator is given by $\tilde{u}^\top \tilde{u}/(n-m+r)$. Usual asymptotic theory assumes that both m and r are fixed so that the difference between n and $n-m+r$ vanishes. Under the asymptotic framework of theorem 1, r is asymptotically negligible, but m is not. The corrected statistic obtained in theorem 1 is

$$\frac{n-m}{n} LM = (R\hat{\beta}-q)^\top \left[\frac{\tilde{u}^\top \tilde{u}}{n-m} R(X^\top X)^{-1} R^\top\right]^{-1} (R\hat{\beta}-q).$$

Thus, it simply corrects the variance estimator given the asymptotic importance of m.

Modifications such as those proposed by theorem 1 are not entirely new in the econometric literature. For instance, Evans and Savin (1982) discuss the following modified LR and LM statistics and derive their asymptotic behavior:

$$\left(1-\frac{m-r/2+1}{n}\right) LR \overset{a}{\sim} \chi^2(r), \tag{3}$$

$$\left(1-\frac{m-r}{n}\right) LM \overset{a}{\sim} \chi^2(r). \tag{4}$$

The correction to the LM statistic aims to eliminate the bias of the error variance estimator while the correction to the LR statistic is derived from an Edgeworth expansion, see Anderson (1958, section 8.6) for a detailed derivation. Evans and Savin (1982, p. 744) report excellent results for the modified LR test even in quite small samples with rather large $\hat{\lambda}$ ($n = 33$ and $m = r = 8$, so $\hat{\lambda} = 0.32$).[2] On the other hand, they find that the modified LM test underrejects in small samples and it is shown that it might actually have zero power against all alternatives when $\hat{\lambda}$ is too large (Evans and Savin, 1982, page 743). They also analyse the small sample properties of the rF test and find that it overrejects when $\hat{\lambda} = 0.32$.

Given these problems, Evans and Savin (1982) suggest using the rF and modified LM statistics (4) along with Edgeworth corrected critical values. Let $q_\alpha^{\chi^2(r)}$ denote the $1-\alpha$ quantile of the $\chi^2(r)$ distribution (that is, the α-level critical value). Then, the Edgeworth corrected critical value for the rF test is

$$q_\alpha^{\chi^2(r)}\left(1+\frac{q_\alpha^{\chi^2(r)}-r+2}{2(n-m)}\right), \tag{5}$$

[2] Evans and Savin (1982) use Laguerre series and numerical integration to study the size and power of their tests. Their results are similar to the Monte Carlo results we present later.

while that of the modified LM test is

$$q_{\alpha}^{\chi^2(r)}\left(1-\frac{q_{\alpha}^{\chi^2(r)}-r-2}{2(n-m)}\right). \qquad (6)$$

These new versions of the rF and LM tests have excellent finite sample properties according to the calculations of Evans and Savin (1982). The results they report are, however, limited to only a few cases of sample size, number of regressors and restrictions. In particular, they do not consider high values of m/n or $r/(n-m)$, the largest being 8/25. We now provide some Monte Carlo results for a wider variety of cases of values of m/n. Unless otherwise specified, every Monte Carlo experiment in this paper is based on the linear model (1) with $u \sim N(0,1)$ and all the regressors, except for a constant, being generated as independent draws from a standard normal distribution. Further, the null hypothesis considered is $H_0 : \beta_2 = 0$ against the null $H_1 : \beta_2 \neq 0$, where β_2 is a $r \times 1$ subvector of β.

Figure 1 shows the null rejection frequencies (RF) of the exact F test (EF), using the $F(r, n-m)$ critical values, the rF test, using the $\chi^2(r)$ critical values and the rF test using the Edgeworth-corrected critical values (5), denoted by rFe. There were 1 000 000 Monte Carlo samples of size $n = 20$, r is either 1 or 3 and μ varies along the horizontal axis. It can be seen that the EF test is indeed exact while the rF test tends to overreject, especially for very large values of μ. The Edgeworth corrected version has much smaller size distortion than the simple rF version but it also overrejects quite severely when μ is large.

Figures 2 and 3 show the same information for the LR and LM tests respectively, in their modified versions given in theorem 1 (MLR and MLM) and their Edgeworth-corrected versions (LMe and LRe) proposed by Evans and Savin (1982), that is, statistics (3) and (4) using critical values from the $\chi^2(r)$ distribution or given in (6), respectively. The sample size is again 20 observations. The LR and LM tests overreject very badly and more so when the number of regressors increases. This is, of course, exactly what should have been expected from theorem 1, which implies that the actual critical values of the LR and LM tests are larger than those of the $\chi^2(r)$ distribution. These results are also very similar to those reported in table 1 of Evans and Savin (1982). Of course, overrejection increases with r.

The Edgeworth-corrected statistics and Anatolyev (2012)'s modified statistics both provide adequate corrections for the LR test when μ is small but loose some accuracy as μ tends to 1. Using the Edgeworth correction, which takes into account r, seems to increase small sample accuracy of the LR test, but the improvement over Anatolyev's (2012) correction is somewhat marginal. As for the LM test, figure 3 shows that the Edgeworth correction is very accurate when μ is small, but severely over or underrejects when μ is large. In comparison, Anatolyev's (2012) correction is not very accurate, even when μ is rather small.

Figures 4, 5 and 6 contain the same informations for samples of 100 observations. As Anatolyev's (2012) theorem 1 states, the rF test improves as n increases, but the LR and LM tests don't. The size distortions of the modified and Edgeworth-corrected LR and LM tests for large μ are much smaller with $n = 100$, which reflects the asymptotic validity of these tests. Notice that the LR tests are systematically more accurate than their LM counterparts.

The results presented in figures 1 to 6 therefore confirm the numerical analysis of Evans and Savin (1982) and illustrate the implications of Anatolyev's (2012) theorem 1. They also show that the corrected, asymptotically valid, tests may not work well in finite samples when the number of regressors is very large relative to the sample size.

2.3 Tests with many regressors and many restrictions

This subsection first provides simulation results for the EF, rF, LR and LM tests when $\rho > 0$, that is, when the number of restrictions is allowed to increase at a fixed rate of the sample size. It then investigates the accuracy of some corrected versions. Table 1 shows the null rejection frequency of the 4 tests for different sample sizes and values of μ and ρ, obtained from 1 000 000 Monte Carlo samples. As expected from theorems 2 and 3 of Anatolyev (2012), the EF test has correct size regardless of μ and ρ while the other tests have severe size distortion. Notice that the size distortion of the LR and LM tests for a given λ actually gets worst for larger sample sizes. This is in accordance with Anatolyev (2012)'s corollary 2, which establishes that the rF, LR and LM tests have incorrect asymptotic sizes. The last few lines of table 1 give the theoretical limiting sizes computed from Anatolyev's (2012) corollary. It can be seen that the simulated rejection frequencies tend to these values.

Anatolyev (2012) proposes testing procedures that have correct asymptotic size. The corrected rF test (crF) rejects whenever $rF > q_{\alpha^F}^{\chi^2(r)}$, where $q_{\alpha^F}^{\chi^2(r)}$ is the $1-\alpha^F$ quantile of the $\chi^2(r)$ distribution and

$$\alpha^F = \Phi\left(\Phi^{-1}(\alpha)\sqrt{1+\hat{\lambda}}\right),$$

where Φ is the standard normal cumulative distribution function. The corrected LR test (cLR) rejects when

$$LR > \frac{n}{r}\ln(1+\hat{\lambda})q^{\chi^2(r)}_{\alpha^{LR}},$$

where

$$\alpha^{LR} = \Phi\left(\frac{\Phi^{-1}(\alpha)}{\sqrt{1+\hat{\lambda}}}\frac{\hat{\lambda}}{\ln(1+\hat{\lambda})}\right).$$

Finally, the corrected LM test (cLM) rejects when

$$LM > \frac{n}{n-m+r}q^{\chi^2(r)}_{\alpha^{LM}},$$

where

$$\alpha^{LM} = \Phi\left(\frac{\Phi^{-1}(\alpha)}{\sqrt{1+\hat{\lambda}}}\right).$$

Corollary 3 of Anatolyev (2012) shows that these corrected tests have asymptotic size α. Monte Carlo results based on 1 000 000 simulated samples are presented in table 2 and confirm this. However, it can be seen that they might have fairly large size distortions in small and moderate samples, especially when λ is high.

It is interesting to compare these corrected tests with those using the Edgeworth corrections discussed in the previous subsection. Under the asymptotic framework described by assumptions 1, 2 and 3 with $\rho > 0$, corollary 4 in Anatolyev (2012) establishes that the Edgeworth corrected tests are not asymptotically valid. Specifically, the asymptotic sizes of the rFe, LRe and LMe tests are shown to be, respectively,

$$\Phi\left((1+\lambda/2)/\sqrt{1+\lambda}\Phi^{-1}(\alpha)\right) < \alpha, \tag{7}$$

$$\Phi\left(\sqrt{1+r}(1-\lambda/2)\Phi^{-1}(\alpha)\right) > \alpha, \tag{8}$$

and

$$\Phi\left(\sqrt{\frac{1+\lambda}{2}}\left(\frac{\ln(1+\lambda)}{\lambda} - \frac{1}{1+\lambda/2}\right)\sqrt{r} + o(\sqrt{r})\right). \tag{9}$$

The actual values of these limits for the λs considered here are given in the last 6 lines of table 2. It can be seen that the Monte Carlo simulations illustrate these results since the finite sample rejection frequencies appear to converge to the theoretical limits.

Thus, the simulations presented in this section are consistent with the theoretical results of Anatolyev (2012). Interestingly, they show that the asymptotically corrected tests may still have noticable size distortions in small and medium samples with large numbers of regressors or restrictions. Of course, because the regressors are exogenous and the errors are iid normal, the EF test performs extremely well. Things might be different with non iid normal data. Also, we did not so far consider the power of the tests. These issues will be taken up in section 4. In the mean time, the next section establishes some results for bootstrap versions of the tests.

3 Bootstrap tests

This section considers the use of the bootstrap under the asymptotic framework described in section 2.1.[3] It is shown in subsection 3.1 that the standard parametric and semiparametric bootstraps remain asymptotically valid when there are many regressors and many restrictions. Subsection 3.2 provides Monte Carlo results that confirm these theoretical results.

3.1 Asymptotic validity of the bootstrap

This subsection shows that the bootstrap remains valid under the many regressors and many restrictions asymptotic framework. This is not a trivial result because when $\mu > 0$ and $\rho < \mu$, the restricted OLS estimator of β with which the bootstrap DGP is built is not consistent. Validity of the bootstrap nevertheless obtains from showing that the theorems of Anatolyev (2012) apply to the bootstrap statistics, which implies that their limiting distributions are the same as those of the sample rF, LR and LM statistics.

Consider again regression model (1). The first step toward running a bootstrap test is to estimate (1) under the null hypothesis. It is well known that the restricted OLS estimator of β is

$$\tilde{\beta} = \hat{\beta} - (X^\top X)^{-1} R^\top \left[R(X^\top X)^{-1} R^\top \right]^{-1} \left(R\hat{\beta} - q \right), \tag{10}$$

[3] Notice that Mammen (1993) provides a proof of the asymptotic validity of the semiparametric bootstrap F test with nonrandom regressors and iid error terms under a different set of assumptions, including that $m/n \to 0$.

where

$$\hat{\beta} = (X^\top X)^{-1} X^\top y.$$

Let $\tilde{\sigma}^2$ denote the degree of freedom corrected variance estimator under H_0,

$$\tilde{\sigma}^2 = \frac{1}{n-m+r} \sum_{i=1}^{n} \tilde{u}_i^2,$$

where $\tilde{u}_i$ is element i of the restricted residuals vector $\tilde{u}$,

$$\tilde{u} = y - X\tilde{\beta}.$$

The bootstrap DGP is then

$$y_i^\star = X_i \tilde{\beta} + u_i^\star, \tag{11}$$

where $u_i^\star$ is either drawn from a normal distribution with mean 0 and variance $\tilde{\sigma}^2$ or drawn with replacement from the empirical density function (EDF) of

$$\left\{ \sqrt{\frac{n}{n-m+r}} \left(\tilde{u}_i - \bar{u} \right) \right\}_{i=1}^{n}, \tag{12}$$

and $\bar{u}$ is the sample average of the $\tilde{u}_i$. It is customary to call the first type of procedure a parametric bootstrap and the second one a semiparametric bootstrap. Notice that recentering and rescaling the residuals in (12) insures that $Var(u_i^\star) = E(u_i^{\star 2}) = \tilde{\sigma}^2$, see Davidson and MacKinnon (2004, section 4.6). Once a bootstrap sample has been drawn from (11), a bootstrap test statistic is calculated. The bootstrap versions of the F, LR and LM statistics are

$$F^\star = \frac{\left(R\hat{\beta}^\star - q \right)^\top \left(\hat{\sigma}^{2\star} R (X^\top X)^{-1} R^\top \right)^{-1} \left(R\hat{\beta}^\star - q \right)}{r}, \tag{13}$$

$$LR^\star = n \ln \left(\frac{\tilde{u}^{\star\top} \tilde{u}^\star}{\hat{u}^{\star\top} \hat{u}^\star} \right), \tag{14}$$

$$LM^\star = \left(R\hat{\beta}^\star - q \right)^\top \left(\tilde{\sigma}^{2\star} R (X^\top X)^{-1} R^\top \right)^{-1} \left(R\hat{\beta}^\star - q \right), \tag{15}$$

where

$$\hat{\beta}^\star = \left(X^\top X \right)^{-1} X^\top y^\star,$$

$$\hat{u}^\star = y^\star - X\hat{\beta}^\star,$$

$$\tilde{u}^\star = y^\star - X\tilde{\beta}^\star,$$

$$\hat{\sigma}^{2\star} = \frac{\hat{u}^{\star\top} \hat{u}^\star}{n-m},$$

$$\tilde{\sigma}^{2\star} = \frac{\tilde{u}^{\star\top} \tilde{u}^\star}{n},$$

and $\tilde{\beta}^\star$ is the restricted estimator (10) obtained from the bootstrap data. In order to prove the consistency of the bootstrap tests based of these three statistics, one must show that their asymptotic distributions are those given in theorem 1 of Anatolyev (2012) for the cases of many regressors and few restrictions ($\mu > 0$ and $\rho = 0$) and theorem 2 for many regressors and many restrictions ($\mu > 0$ and $0 < \rho \leq \mu$). We now state formally these results:

Theorem 2. Under assumptions 1 to 3, the bootstrap rF, LR and LM tests are asymptotically valid when $\mu > 0$ and $\rho = 0$, that is

$$rF^\star \overset{a}{\sim} \chi^2(r),$$
$$\left(1 - \frac{m}{n}\right) LR^\star \overset{a}{\sim} \chi^2(r),$$
$$\left(1 - \frac{m}{n}\right) LM^\star \overset{a}{\sim} \chi^2(r).$$

Proof: see appendix.

Theorem 3. Under assumptions 1 to 3 with $\mu > 0$ and $0 < \rho \leq \mu$, the bootstrap rF, LR and LM tests are asymptotically valid, that is,

$$\sqrt{r}\,(F^\star - 1) \overset{a}{\sim} N(0, 2(1+\lambda)),$$
$$\sqrt{r}\left(\frac{LR^\star}{n} - \ln(1+\lambda)\right) \overset{a}{\sim} N\left(0, \frac{2\lambda^2}{1+\lambda}\right),$$
$$\sqrt{r}\left(\frac{LM^\star}{n} - \frac{\lambda}{1+\lambda})\right) \overset{a}{\sim} N\left(0, \frac{2\lambda^2}{(1+\lambda)^3}\right).$$

Proof: see appendix.

As it turns out, the proofs only require minor adaptations of the key results and proofs of Anatolyev (2012). These proofs rely on assumptions 1, 2 and 3 as well as on exogeneity of the regressors and the assumption that the error terms are iid. Evidently, because of the nature of the bootstrap DGP, the bootstrap errors are iid. Furthermore, if the regressors are exogenous, then they remain so in the bootstrap DGP. Assumption 1 obviously also holds for the bootstrap DGP if it holds in the original sample. Assumption 3 is concerned only with some properties of the regressors, so it also necessarily is true in the bootstrap DGP if it is so in the original sample.

Assumption 2, finiteness of the fourth moment, that is $E(u_i^{\star 4}) = \kappa^\star < \infty$, requires a little bit more thought. For the parametric bootstrap, $\kappa^\star = 3\tilde{\sigma}^4$, so $\kappa^\star$ is finite if $\tilde{\sigma}$ is. Under assumption 2, σ is finite, so one way to show that $\tilde{\sigma}$ is finite is to show that it is a consistent estimator of σ. This is quite easy to do. Indeed, by the properties of orthogonal projection matrices, and assuming that H_0 is true,

$$\begin{aligned}\tilde{\sigma}^2 &= \frac{u^\top \left[I - X\left(\Xi_{Im} - \Xi_R\right)X^\top\right]u}{n-m+r} \\ &= \frac{n}{n-m+r}\left[\frac{u^\top u}{n} - \frac{m-r}{n}\frac{u^\top X\left(\Xi_{Im} - \Xi_R\right)X^\top u}{m-r}\right] \\ &\rightarrow \frac{1}{1-\mu+\rho}\left(\sigma^2 - (\mu-\rho)\sigma^2\right) = \sigma^2,\end{aligned}$$

where the last line follows from lemma 1 in Anatolyev (2012). As for the semi-parametric bootstrap, let $\{\tilde{u}_i'\}_{i=1}^n$ denote the recentered and rescaled residuals (12), which are all finite numbers. Then, $\kappa^\star = \frac{1}{n}\sum_{i=1}^n \tilde{u}_i'^4 < \infty$.

Thus, if the sample data's DGP respects all the necessary conditions for Anatolyev's (2012) results to hold, then the bootstrap DGP also respects these conditions, so theorems 1 and 2 follow.[4]

The simulation results presented in the rest of the paper are based on the semi-parametric bootstrap only, which is much more commonly used than the gaussian parametric bootstrap. Consequently, the term bootstrap should henceforth be understood as referring to the semi-parametric bootstrap. Because that procedure involves drawing bootstrap errors from the EDF of the restricted models' residuals, its finite sample accuracy is affected by how well the EDF estimates the true error distribution. Mammen (1996) shows that the EDF of the maximum likelihood residuals of a m dimensional linear regression with $m^2/n \rightarrow \infty$ is asymptotically biased.[5] This bias is shown to be a function of several variables and a bias corrected estimator is proposed. Interestingly, it turns out that the proper bias correction in the case of a well specified gaussian model happens to be exactly the residuals rescaling (12). Thus, whenever the null model H_0 is well specified and the errors are close to gaussians, the EDF of the rescaled residuals (12) should provide a reasonably accurate estimator of the actual errors' distribution.

[4]Detailed proofs are given in the appendix.

[5]It is easy to see that if $m/n \rightarrow \mu$, $0 < \mu < 1$, then $m^2/n \rightarrow \infty$. However, some of Mammen's results also require that $m^2/n = o(n^{1/5})$, which is not true if $\mu > 0$. Nevertheless, the results from that paper are still of considerable interest in finite samples because they allow for quite a large number or regressors.

3.2 Monte Carlo results

The last column of table 2 shows Monte Carlo simulation results about the size of the bootstrap test.[6] Each bootstrap test was conducted using 499 bootstrap samples. In order to keep the experiment's running time within reasonable limits, the number of Monte Carlo samples varied with the sample size, so that 500 000 samples were used when $n = 20$, 100 000 when $n = 100$ and 10 000 when $n = 250$. The computational time of similar experiments with samples of 1000 observations was somewhat prohibitive. We therefore made use of the "warp speed" method developed by Giacomini, Politis and White (2013). This consists of drawing a single bootstrap sample (instead of B samples) for each Monte Carlo repetition and use their empirical distribution to compute a critical value which is then used in each Monte Carlo sample. As Davidson and MacKinnon (2002, 2007) argue and Giacomini et al (2013) rigorously show, such a procedure is valid under fairly weak conditions, including independence between the bootstrap and the original Monte Carlo statistics. To judge whether this works in the present case, we first reran every experiment with smaller samples using this fast Monte Carlo method. Since we obtained results almost identical to those from standard Monte Carlo, we expect that the results reported for $n = 1000$ can be relied upon. Every such fast Monte Carlo experiment was carried-out using 10 000 simulated samples.

It can be seen in table 2 that the bootstrap works extremely well even in small samples. Contrarily to the corrected versions of the rF, LR and LM tests, the bootstrap test does not have severe size distortion when the number of restrictions is large relative to the number of degrees of freedom. This is most likely due to the fact that the bootstrap test, just like the EF test, uses critical values that take into account, through the bootstrap DGP, the number of parameters, the sample size and the number of restrictions rather than just the number of restrictions.

4 Power and other issues

This section uses yet some more Monte Carlo experiments to address three important aspects of testing with many regressors and restrictions. The first is the power of the various available tests, for which Anatolyev (2012) gives some theoretical

[6] All four bootstrap tests have identical rejection frequencies because the rF, LR and LM statistics are simple monotone transformations of the F statistic. Thus, the location of any given original sample statistic in the bootstrap distribution is the same as that of the other three. The bootstrap P values are therefore identical.

results. Then the tests are analyzed with DGPs with skewed and heavy tailed regressors and errors, violating assumptions 2 and 3. Finally, heteroskedastic errors are considered.

4.1 Power

Theorem 4 of Anatolyev (2012) establishes that the three corrected tests and the EF test all have equivalent and non-trivial power against a sequence of local alternatives. Precisely, letting δ denote a $m \times 1$ vector of constants such that

$$\Delta = \lim_{n\to\infty} \frac{\delta^\top R^\top \left(R\left(X^\top X\right)^{-1} R^\top\right)^{-1} R\delta}{r^2}$$

is finite, a sequence of drifting DGPs may be defined by the vector of parameters $\check{\beta} = \beta + \delta/r^{3/4}$. Under this setting, with $\rho > 0$ and when assumptions 1 to 3 hold, the asymptotic rejection probability of the EF and corrected rF, LR and LM tests conducted at nominal level α is given by

$$\Phi\left(\frac{\Delta}{\sigma^2\sqrt{2(1+\lambda)}} - \Phi^{-1}(1-\alpha)\right), \tag{16}$$

see Anatolyev (2012, p. 374). Notice that the null hypothesis corresponds to $\Delta = 0$, in which case (16) indeed yields a rejection probability of $\alpha\%$. Expression (16) implies quite common features, such as the fact that larger departures from the null hypothesis, that is, larger values of Δ, increase the power of the tests and that higher error variance diminishes it. It also explicitly shows that the number of restrictions per degree of freedom is negatively related to power.

Table 3 provides Monte Carlo results for a variety of alternative DGPs. It can be seen that power is indeed negatively related to λ, which illustrates the theoretical result (16). The corrected rF, LR and LM tests have better power than the EF and bootstrap tests when λ is large, but it is only so because they over-reject under H_0, recall table 2. The only other remarkable feature here is that the bootstrap tests have power quite similar to that of the EF test.

4.2 Non Gaussian errors and regressors

We now check whether deviations from assumptions 2 and 3 greatly affect the finite sample performances of the bootstrap and corrected asymptotic tests. In order to do this, we use the same data generating process as before with $\mu = 0.5$ and $\rho = 0.3$, which is a case in which the corrected tests accuracy increases quickly with the sample size, see table 2.

The results shown in table 2 were obtained by drawing the regressors and the error terms from independent standard normal distributions. In table 4, we provide results for the EF test, corrected rF, LR and LM tests and the bootstrap when X and u are drawn from other distributions that violate assumptions 2 and 3. On the one hand, to isolate the effect of high kurtosis, we draw from a Cauchy (that is, Student with one degree of freedom) distribution, which has infinite fourth moment, in clear violation of assumption 2. Results are also presented for drawings from Student distributions with 2, 3 and 4 degrees of freedom in order to investigate the effect of an infinite fourth moment with somewhat thinner tails than the Cauchy distribution's. Then, to investigate the impact of skewness, and departures from assumption 3, we draw numbers from the lognormal distribution. This yields draws from a thin tailed distribution strongly skewed to the right with high probability of generating higly leverage observations.

Panels 1 to 4 of table 4 show that violation of assumption 2 may result in important accuracy losses in small samples and asymptotic invalidity of the testing procedures since all the tests' rejection frequencies diverge away from 0.05 as n increases. This is most clearly seen with thicker tailed distributions. It is worth noting that even tough it also is asymptotically invalid, the bootstrap test is considerably more accurate in finite samples than any other test, including the EF test, when the tails are very thick.

There exists a variety of resampling methods that remain valid with thick tail data. Among these is the semiparametric bootstrap proposed by Davidson and Flachaire (2007), which consists of combining usual iid drawings from the EDF in the middle part of the distribution with drawings from a parametric estimate of the tails of the distribution. As an illustration, we have run some Monte Carlo simulations using this method with data from the Cauchy distribution. We arbitrarily defined the tail areas as the $\sqrt{n}$ smallest and largest residuals and used these observations to obtain parametric estimates of the left and right tails of the distribution in precisely the same way as Davidson and Flachaire (2007), to whom we refer the interested reader for details. In those experiments, the null rejection frequency of this bootstrap test is 0.0669 in samples of 20 observations, 0.0537 with 100 observations and 0.0442 with 250 observations. These figures are markedly closer to 5% than those of the ordinary bootstrap presented in panel 1 of table 4, which suggests that the semiparametric bootstrap of Davidson and Flachaire (2007) is valid. Limited simulation evidence which we do not report here suggests that subsampling methods also remain valid.

Panel 5 shows that skewed data adversely affects the tests' finite sample accuracy since the rejection frequencies reported there are somewhat larger than those displayed in table 2.

4.3 Heteroskedasticity

We now consider the situation in which the error terms in model (1) are heteroskedastic, which violates the iid assumption made in section 2.1. The regression model is then

$$y_i = X_i\beta + u_i, \quad E(u_i) = 0, \quad E(u_i^2) = \sigma_i^2. \tag{17}$$

It has been known since White (1980) that, in the presence of conditional heteroskedasticity, that is, heteroskedasticity depending on the elements of the matrix of regressors X, one can obtain an asymptotically valid and heteroskedasticity-robust Wald test defined as

$$W_{HR} = \left(R\hat{\beta} - q\right)^\top \left[R\hat{V}R^\top\right]^{-1} \left(R\hat{\beta} - q\right), \tag{18}$$

where $\hat{V}$ is a heteroskedasticity consistent covariance matrix estimator (HCCME). Generally, this is definned as

$$\hat{V} = (X^\top X)^{-1}(X^\top \hat{\Omega} X)(X^\top X)^{-1} \tag{19}$$

where $\hat{\Omega}$ is a diagonal matrix the elements of which are $f(\hat{u}_i^2)$, where $\hat{u}_i$ is the ith residual and f is a given transformation. Different such transformations have been studied by MacKinnon and White (1985) in the context of heteroskedasticity robust covariance matrix estimation. The simplest one is no transformation at all, $f(\hat{u}_i^2) = \hat{u}_i^2$, which is often referred to as HC0. Other transformations include

$$\text{HC1}: \quad f(\hat{u}_i^2) = \frac{n}{n-m}\hat{u}_i^2,$$

$$\text{HC2}: \quad f(\hat{u}_i^2) = \frac{1}{1-h_i}\hat{u}_i^2,$$

$$\text{HC3}: \quad f(\hat{u}_i^2) = \frac{1}{(1-h_i)^2}\hat{u}_i^2,$$

where h_i is the ith diagonal element of the orthogonal projection matrix $X(X^\top X)^{-1}X^\top$. Transformation HC1 aims to compensate for the fact that OLS residuals are on average smaller than the unobserved errors, while transformations HC2 and HC3 take into account the leverage effect of each individual observation, see MacKinnon (2013) for a detailed treatment and a recent survey of heteroskedasticity robust methods.

In the present section, we are mainly interested in learning something about the finite sample accuracy of the bootstrap under these circumstances. The naive bootstrap considered so far is of course inappropriate in the presence of heteroskedasticity, so we consider the wild bootstrap instead. The wild bootstrap DGP is

$$y_i^\star = X_i\tilde{\beta} + g(\tilde{u}_i)\eta_i^\star,$$

where g is the square root of one of the a transformations HC0, HC1, HC2 or HC3 described above, $\eta_i^\star$ is a random variables with expectation 0 and variance 1 and the tilde indicate estimates under the null. Following Davidson and Flachaire (2008), the random variables $\eta_i^\star$, $i = 1, ..., n$ are independent draws from the two-point Rademacher distribution and constrained residuals are employed in forming the wild bootstrap samples and to compute the test statistics. Using unconstrained residuals to build the test statistics had only a small incidence on our results. MacKinnon's (2013) results indicate that the most accurate wild bootstrap in finite samples is the one where the bootstrap samples are built using transformation HC3 with restricted residuals while the test statistics are built using transformation HC1. Thus, all the figures reported here were obtained in this way. However, many other specifications yielded quite similar results.

Since Anatolyev (2012) assumes homoskedasticity, we are also interested in the finite sample accuracy of his corrected tests with heteroskedastic errors. We also look at the null rejection frequency of a corrected naive heteroskedasticity-robust F test based on the crF test statistic. This test is carried-out by comparing the W_{HC} statistic to the critical value $q_{\alpha^F}^{\chi^2(r)}$ defined before, multiplied by $n/(n-m)$. We make no claim as to the asymptotic validity of this procedure nor its robustness to heteroskedasticity, but we hope that our simulations may provide some interesting guidelines for practical applications.

Running Monte Carlo experiments for heteroskedastic models is not as straightforward as it might seem. Indeed, as argued by MacKinnon (2013), the choice of DGP can considerably affect the outcome of the experiment. The key feature is the largest h_i, which measures the leverage of observation i on the regression line, and what happens to it as $n \to \infty$. Indeed, it has been shown that a large value or h_i along with a high σ_i can severely distort HCCME based inference, see Chesher and Jewitt (1987). Several authors, including MacKinnon and White (1985) and Davidson and Flachaire (2008), use an experimental design in which a matrix of regressors X is first created for a small sample, often with a single extreme observation, and the variances of the error terms is tied to the contemporaneous values of these regressors. Then, for larger samples, this matrix is simply repeated as many times as necessary. It follows from that construction that the largest value of h_i must decrease as a function of $1/n$. This implies that the severeness of the heteroskedasticity problem decreases as a function of the sample size and the accuracy of heteroskedasticity-robust methods should be expected to increase rapidly with n.

As noted by MacKinnon (2013), this design may yield overly optimistic results. For this reason, his simulations are based on a DGP allowing much more severe heteroskedasticity in large samples. In that DGP, y_i is generated from a linear model such as (17) with

$$\sigma_i = z(\gamma) \left(\sum_{k=1}^{m} X_{i,k} \right)^{\gamma}, \tag{20}$$

where all but the first column of X, which is a constant, contain independent draws from a lognormal distribution, γ is a parameter controlling the degree of heteroskedasticity and $z(\gamma)$ is a scaling factor chosen so that the average variance is 1. This DGP has the advantage of allowing one to control the strength of heteroskedasticity by varying the coefficient γ. The fact that the regressors are drawn from a lognormal distribution creates extreme observations which are likely to become evermore extreme as n increases, so that there are high leverage observations with strong heteroskedasticity even in large samples.

Table 5 reports the results of some simulations carried-out using a linear regression model such as (17) with lognormal regressors, heteroskedastic normal errors and skedastic function (20). The values of γ range from weak ($\gamma = 0.25$) to moderate ($\gamma = 1$) and severe ($\gamma = 3$) heteroskedasticity. Again, we consider the case $\mu = 0.5$ and $\rho = 0.3$. As may be seen, the presence of heteroskedasticity has a strong impact on the EF test and the three corrected tests, who all over-reject moderately when the heteroskedasticity is weak but quite a lot when the errors are strongly heteroskedastic. The wild bootstrap, however, retains a roughly correct size, which indicates it is still valid in the presence of many regressors and restrictions. Unsurprisingly perhaps, the (naively) corrected heteroskedasticity robust F tests (labeled CF0, CF1, CF2 and CF3) do not seem to work well at all.

5 Conclusion

This paper explores the properties of bootstrap tests in linear regressions with large numbers of parameters and restrictions. Using the asymptotic framework of Anatolyev (2012), we find that bootstrap F, LR and LM tests provide asymptotically valid inferences without the necessity of any correction to account for the high dimensionality of the model. Monte Carlo simulations indicate that these tests achieve the same degree of finite sample accuracy than the F test based on critical values from Fisher's distribution when that test is exact. Simulations further show that bootstrap methods designed to be robust to heavy tailed data or heteroskedastic errors remain correctly sized while the F test no longer is exact.

6 Appendix: Detailed proofs.

All proofs provided in this section are almost identical to those provided by Anatolyev (2012). As stated in the main body of the text, they rely on the fact that the bootstrap error terms are iid draws with variance $\tilde{\sigma}^2$. Notice that all the following statements and proofs are stated conditional on the bootstrap DGP (11) and on the regressor matrix X, thus making constant use of the exogeneity assumption.

Lemma 1. Under assumptions 1, 2 and 3,

a) $\frac{u^{\star\top} X \Xi_p X^\top u^\star}{p\tilde{\sigma}^2} \xrightarrow{p} 1$

and

b) $\frac{u^{\star\top} X \Xi_p X^\top u^\star}{p\tilde{\sigma}^2} - 1 = O_p\left(p^{-1/2}\right).$

Proof of a)

$$E\left[\frac{u^{\star\top} X \Xi_p X^\top u^\star}{p\tilde{\sigma}^2}\right] = \frac{1}{p\tilde{\sigma}^2} E\left[tr\left(u^{\star\top} X \Xi_p X^\top u^\star\right)\right]$$

$$= \frac{1}{p\tilde{\sigma}^2} tr\left(E(u^\star u^{\star\top}) X \Xi_p X^\top\right)$$

$$= \frac{1}{p} tr\left(X \Xi_p X^\top\right)$$

$$= \frac{1}{p} tr\left(\Xi_p X^\top X\right)$$

$$= \frac{1}{p} tr\left((X^\top X)^{-1} P^\top \left[P(X^\top X)^{-1} P^\top\right]^{-1} P(X^\top X)^{-1} X^\top X\right)$$

$$= \frac{1}{p} tr\left((X^\top X)^{-1} P^\top \left[P(X^\top X)^{-1} P^\top\right]^{-1} P\right)$$

$$= \frac{1}{p} tr\left(P(X^\top X)^{-1} P^\top \left[P(X^\top X)^{-1} P^\top\right]^{-1}\right)$$

$$= \frac{1}{p} tr\left(I_p\right)$$

$$= 1. \quad \diamondsuit$$

Proof of b) It can easily be seen that

$$\frac{u^{\star\top} X \Xi_p X^\top u^\star}{\tilde{\sigma}^2 p} - 1 = \frac{1}{p} \sum_{i=1}^{n} \sum_{j=1}^{n} X_i^\top \Xi_p X_j \frac{u_i^\star u_j^\star}{\tilde{\sigma}^2} - 1.$$

$$= A_1 + A_2$$

where

$$A_1 = \frac{1}{p} \sum_{i=1}^{n} X_i^\top \Xi_p X_i \left(\frac{u_i^{\star 2}}{\tilde{\sigma}^2} - 1\right)$$

and

$$A_2 = \frac{1}{p} \sum_{i \neq j} X_i^\top \Xi_p X_j \frac{u_i^\star u_j^\star}{\tilde{\sigma}^2}.$$

We will show that the variance of $A_1 + A_2$ is $O(1/p)$. Because the $u_t^\star$ are drawn from iid resampling, Cov(A_1,A_2)=0. The variance of $A_1 + A_2$ is therefore the sum of their variances. We start with A_1.

$$Var(A_1) = E\left[\frac{1}{p} \sum_{i=1}^{n} X_i^\top \Xi_p X_i \left(\frac{u_i^{\star 2}}{\tilde{\sigma}^2} - 1\right)\right]^2$$

$$= \frac{1}{p^2} \sum_{i=1}^{n} \left(X_i^\top \Xi_p X_i\right)^2 E\left(\frac{u_i^{\star 4}}{\tilde{\sigma}^4} - 1\right)$$

$$= \left(\frac{\kappa^\star}{\tilde{\sigma}^4} - 1\right) \frac{1}{p^2} \sum_{i=1}^{n} \left(X_i^\top \Xi_p X_i\right)^2,$$

where $\kappa^\star$ is the fourth moment of $u^\star$,

$$= \left(\frac{\kappa^\star}{\tilde{\sigma}^4} - 1\right) \frac{1}{p} \sum_{i=1}^{n} \left(X_i^\top \Xi_p X_i\right)^2 \frac{1}{p/n} \frac{1}{n}$$

$$= \left(\frac{\kappa^\star}{\tilde{\sigma}^4} - 1\right) \frac{1}{pn} \left[\sum_{i=1}^{n} \left(X_i^\top \Xi_p X_i\right)^2\right] \frac{1}{\pi + o(r^{-1/2})}.$$

Now, for any given i,

$$\begin{aligned}(X_i^\top \Xi_p X_i)^2 &= (\pi + X_i^\top \Xi_p X_i - \pi)^2 \\ &\leq (\pi + |X_i^\top \Xi_p X_i - \pi|)^2 \\ &= \pi^2 + o(1),\end{aligned}$$

where we used the triangle inequality and assumption 3. Thus,

$$\begin{aligned}Var(A_1) &= \left(\frac{\kappa^\star}{\tilde{\sigma}^4} - 1\right) \frac{1}{p} \frac{n(\pi^2 + o(1))}{n} \frac{1}{\pi + o(r^{-1/2})} \\ &= \left(\frac{\kappa^\star}{\tilde{\sigma}^4} - 1\right) \frac{1}{p} \frac{\pi^2 + o(1)}{\pi + o(r^{-1/2})} \\ &= O(1/p).\end{aligned}$$

For A_2, we have

$$\begin{aligned}Var(A_2) &= \frac{1}{p^2} E\left[\left(\sum_{i \neq j} X_i^\top \Xi_p X_j \frac{u_i^\star u_j^\star}{\tilde{\sigma}^2}\right)^2\right] \\ &= \frac{1}{p^2} E\left[\sum_{i \neq j} \sum_{k \neq l} X_i^\top \Xi_p X_j X_k^\top \Xi_p X_l \frac{u_i^\star u_j^\star}{\tilde{\sigma}^2} \frac{u_k^\star u_l^\star}{\tilde{\sigma}^2}\right] \\ &= \frac{1}{p^2} \sum_{i \neq j} \sum_{k \neq l} X_i^\top \Xi_p X_j X_k^\top \Xi_p X_l E\left[\frac{u_i^\star u_j^\star u_k^\star u_l^\star}{\tilde{\sigma}^2 \tilde{\sigma}^2}\right].\end{aligned}$$

In this last expression, except when $i = k$ and $j = l$ or $i = l$ and $j = k$, the expectation is 0 because the $u_t^\star$ are obtained by iid resampling. When $i = k$ and $j = l$ or $i = l$ and $j = k$, we have

$$E\left[\frac{u_i^\star u_j^\star u_k^\star u_l^\star}{\tilde{\sigma}^2 \tilde{\sigma}^2}\right] = E\left[\frac{u_i^{\star 2} u_j^{\star 2}}{\tilde{\sigma}^2 \tilde{\sigma}^2}\right] = 1.$$

It therefore follows that

$$Var(A_2) = \frac{2}{p^2}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\left(X_i^\top \Xi_p X_j\right)^2$$

$$= \frac{2}{p^2}\sum_{i=1}^{n} X_i^\top \Xi_p \left(\sum_{j=1,j\neq i}^{n} X_j X_j^\top\right)\Xi_p X_i$$

$$= \frac{2}{p^2}\sum_{i=1}^{n}\left(X_i^\top \Xi_p X_i - \left(X_i^\top \Xi_p X_i\right)^2\right)$$

$$= \frac{2}{p^2}\sum_{i=1}^{n} X_i^\top \Xi_p X_i - \frac{2}{p^2}\sum_{i=1}^{n}\left(X_i^\top \Xi_p X_i\right)^2$$

$$= \frac{2}{p} - \frac{2}{p^2}\sum_{i=1}^{n}\left(X_i^\top \Xi_p X_i\right)^2$$

$$< \frac{2}{p} = O(1/p),$$

where the next to last line uses the fact that

$$\sum_{i=1}^{n} X_i^\top \Xi_p X_i = p$$

and the inequality comes from the fact that $Var(A_2) > 0$ and

$$\frac{2}{p^2}\sum_{i=1}^{n}\left(X_i^\top \Xi_p X_i\right)^2 > 0.$$

Putting this result together with that on $Var(A_1)$ yields $Var(A_1 + A_2) = O(1/p)$, which completes the proof. $\diamondsuit$

Lemma 2 Under assumptions 1, 2 and 3,

a) $\hat{\sigma}^{\star 2} \xrightarrow{p} \tilde{\sigma}^2$

and

b) $\frac{\hat{\sigma}^{\star 2}}{\tilde{\sigma}^2} - 1 = O_p(n^{-1/2})$.

Proof of a)

$$\hat{\sigma}^{\star 2} = \frac{u^{\star\top}\left[I - X(X^\top X)^{-1}X^\top\right]u^\star}{n-m}$$

$$= \frac{n}{n-m}\left[\frac{u^{\star\top}u^\star}{n} - \frac{m}{n}\frac{u^{\star\top}X\Xi_I X^\top u^\star}{m}\right].$$

Applying lemma 1 to the second term in the square bracket and a weak law of large numbers to the first term, we get

$$\hat{\sigma}^{\star 2} \xrightarrow{p} \frac{n/n}{n/n - m/n}\left[\tilde{\sigma}^2 - \frac{m}{n}\tilde{\sigma}^2\right]$$
$$\xrightarrow{p} \frac{1}{1-\mu}\left[\tilde{\sigma}^2(1-\mu)\right] = \tilde{\sigma}^2. \quad \diamondsuit$$

Proof of b)

$$\hat{\sigma}^{\star 2} - \tilde{\sigma}^2 = \frac{u^{\star\top}\left[I - X(X^\top X)^{-1}X^\top\right]u^\star}{n-m} - \tilde{\sigma}^2$$
$$= \frac{u^{\star\top}u^\star}{n-m} - \frac{u^{\star\top}X\Xi_I X^\top u^\star}{n-m} + \frac{m-n}{n-m}\tilde{\sigma}^2$$
$$= \frac{n}{n-m}\left[\frac{u^{\star\top}u^\star}{n} - \tilde{\sigma}^2 - \frac{m}{n}\left(\frac{u^{\star\top}X\Xi_I X^\top u^\star}{m} - \tilde{\sigma}^2\right)\right]$$
$$= \frac{1}{1-\mu+o(r^{-1/2})}\left[O_p(n^{-1/2}) - \frac{m}{n}\left(\frac{u^{\star\top}X\Xi_I X^\top u^\star}{m} - \tilde{\sigma}^2\right)\right],$$

where we used the fact that $u^{\star\top}u^\star/n$ is root-n consistent. Then, using the second part of lemma 1,

$$\hat{\sigma}^{\star 2} - \tilde{\sigma}^2 = \frac{1}{1-\mu+o(r^{-1/2})}\left[O_p(n^{-1/2}) - \frac{m}{n}O_p(m^{-1/2})\right]$$
$$= \frac{1}{1-\mu+o(r^{-1/2})}\left[O_p(n^{-1/2}) - (\mu+o(r^{-1/2}))O_p(m^{-1/2})\right]$$
$$= O_p(n^{-1/2}). \quad \diamondsuit$$

Theorem 2. Under assumptions 1 to 3, the bootstrap rF, LR and LM tests are valid when $\mu > 0$ and $\rho = 0$.

Proof.

Define H_n and G_n such that $H_n^\top H_n = \left(X^\top X\right)^{-1}$ and $G_n G_n^\top = \Upsilon_R$, where

$$\Upsilon_R = H_n R^\top \left[R\left(X^\top X\right)^{-1}R^\top\right]^{-1} R H_n^\top.$$

Note that, as long as X is exogenous, neither G_n nor H_n depend on the bootstrap DGP. Consequently, the results derived by Anatolyev for the triangular array $\Pi_n = XH_n^\top G_n$ apply here without modifications. In particular, under assumption 3 and $\rho = 0$,

$$\max_{1\le i\le n} | \, [\Pi_n]_{ij} \, | \le \sqrt{\max_{1\le i\le n} X_i^\top \Xi_R X_i} \to 0. \tag{21}$$

It is easy to check that

$$rF^\star = \left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^{\star 2}}\right) \frac{u^{\star\top}}{\tilde{\sigma}} \Pi_n \Pi_n^\top \frac{u^\star}{\tilde{\sigma}}.$$

By lemma 2,

$$rF^\star \overset{a}{=} \frac{u^{\star\top}}{\tilde{\sigma}} \Pi_n \Pi_n^\top \frac{u^\star}{\tilde{\sigma}}.$$

Because of (21) and because the $u_t^\star$ are iid with variance $\tilde{\sigma}^2$, a CLT (Pötsher and Prucha, 2001, theorem 30) implies

$$rF^\star \xrightarrow{d} \chi^2(r).$$

Thus, the bootstrap rF test is asymptotically valid when $\mu > 0$ and $\rho = 0$ because the $rF^\star$ statistic has the same asymptotic null distribution than the sample-based rF statistic. The fact that the LR and LM test statistics are simple functions of the rF statistic implies that their bootstrap versions also yield valid tests. For instance, for the bootstrap LR test,

$$LR^\star = n \ln\left(1 + \frac{1}{n-m} rF^\star\right) \to \frac{n}{n-m} rF^\star$$

because of the properties of the natural logarithm. Then, we may write

$$LR^\star \overset{a}{=} \frac{1}{1 - m/n} rF^\star,$$

so that

$$\left(1 - \frac{m}{n}\right) LR^\star \overset{a}{=} rF^\star,$$

which implies

$$\left(1 - \frac{m}{n}\right) LR^\star \xrightarrow{d} \chi^2(r),$$

which accords with Anatoliev's theorem 1 and shows that the bootstrap LR test is asymptotically valid. A similar result obtains for the LM test by using

$$LM^\star = \frac{n}{(n-m)(1 + rF^\star/(n-m))} rF^\star. \quad \diamondsuit$$

Theorem 3. Under assumptions 1 to 3 with $\mu > 0$ and $0 < \rho \leq \mu$, the bootstrap rF, LR and LM tests are asymptotically valid.

Proof. We will show that the normalized and recentered versions of the bootstrap test statistics have the same limiting distributions as their sample counter-parts. Precisely, we will show that

$$\sqrt{r}(F^\star - 1) \xrightarrow{d} N(0, 2(1+\lambda),$$

$$\sqrt{r}\left(\frac{LR^\star}{n} - \ln(1+\lambda)\right) \xrightarrow{d} N\left(0, \frac{2\lambda^2}{1+\lambda}\right),$$

and

$$\sqrt{r}\left(\frac{LM^\star}{n} - \frac{\lambda}{1+\lambda}\right) \xrightarrow{d} N\left(0, \frac{2\lambda^2}{(1+\lambda)^3}\right),$$

which are the limits derived by Anatolyev. We first consider the $F^\star$ statistic. We can write

$$F^\star = \frac{\tilde{\sigma}^2}{\hat{\sigma}^{\star 2}} \frac{u^{\star\top} X \Xi_R X^\top u^\star}{r\tilde{\sigma}^2}.$$

By lemmas 1 and 2, $F^\star \xrightarrow{p} 1$. Then, using lemma 1, lemma 2, a central limit theorem and a Taylor expansion, we can obtain

$$\sqrt{r}(F^\star - 1) = A^\star + o_p(r^{-1/2}),$$

where

$$A^\star = \sqrt{r}\left[\left(\frac{u^{\star\top} X \Xi_R X^\top u^\star}{r\tilde{\sigma}^2} - 1\right) - \frac{1}{1-\mu}\left(\frac{u^{\star\top} u^\star}{n\tilde{\sigma}^2} - 1\right) + \frac{\mu}{1-\mu}\left(\frac{u^{\star\top} X \Xi_{Im} X^\top u^\star}{m\tilde{\sigma}^2} - 1\right)\right].$$

Details of these derivations may be found in Anatolyev (2012), pages 377-378. They are not reproduced here because they obtain from simply replacing population quantities by the corresponding bootstrap ones. The variable $A^\star$ can be decomposed as follows:

$$A^\star = A_1^\star + A_2^\star + o_p(1),$$

where

$$A_1^\star = \sum_{i=1}^{n} \frac{1}{\sqrt{r}} \left(X_i^\top \Xi_R X_i + \lambda\left(X_i^\top \Xi_{Im} X_i - 1\right)\right)\left(\frac{u_i^{\star 2}}{\tilde{\sigma}^2} - 1\right)$$

and

$$A_2^\star = \sum_{i \neq j} \frac{1}{\sqrt{r}} X_i^\top \left(\Xi_R + \lambda \Xi_{Im}\right) X_j \frac{u_i^\star u_j^\star}{\tilde{\sigma}^2}.$$

It is easy to see that $E(A_1^\star) = 0$. Also,

$$Var(A_1^\star) = \left(\frac{\kappa^\star}{\tilde{\sigma}^4} - 1\right) \frac{1}{r} \sum_{i=1}^{n} \left(X_i^\top \Xi_R X_i + \lambda \left(X_i^\top \Xi_{Im} X_i - 1\right)\right)^2.$$

Then, by assumption 3 and the fact that $\kappa^\star < \infty$, we have

$$Var(A_1^\star) = \left(\frac{\kappa^\star}{\tilde{\sigma}^4} - 1\right) \frac{o(1)}{\rho + o(r^{-1/2})} = o(1),$$

see Anatolyev (2012), page 378 for details. Hence, $A_1^\star = o_p(1)$ and is therefore asymptotically negligible. Next, because the $u_t^\star$ are iid, we have $E(A_2^\star) = 0$. Also, we have,

$$\begin{aligned} Var(A_2^\star) &= \left(\frac{n}{r}\right) \frac{1}{n} E\left[\left(\sum_{i \neq j} X_i^\top \left(\Xi_R + \lambda \Xi_{Im}\right) X_j \frac{u_i^\star u_j^\star}{\tilde{\sigma}^2}\right)^2\right] \\ &= \left(\frac{n}{r}\right) \frac{2}{n} \sum_{i=1}^{n} \sum_{j=1, j \neq i} \left[X_i^\top \left(\Xi_R + \lambda \Xi_{Im}\right) X_j\right]^2, \end{aligned}$$

because the $u_i^\star$ are iid with variance $\tilde{\sigma}^2$. It therefore follows from assumption 3 that $Var(A_2^\star) \to 2(1+\lambda)$, see Anatolyev (2012) p. 378 for detailed derivations. Hence, the validity of the bootstrap F test will be established if we can show that a central limit theorem applies to $A_2^\star$. Anatolyev shows that the CLT of Kelejian and Prucha (2001) applies to the sample counterpart of $A_2^\star$, where $u_k^\star$ is replaced by u_k and σ^2 replaces $\tilde{\sigma}^2$. The first condition is that

$$\sup_{1 \leq j \leq n, n \geq 1} \sum_{i=1}^{n} \left| \frac{1}{\sqrt{r}} X_i^\top \left(\Xi_R + \lambda \Xi_{Im}\right) X_j \right| < \infty$$

and is shown to hold by Anatolyev (2012), p. 378. It is also necessary that $X_i^\top \left(\Xi_R + \lambda \Xi_{Im}\right) X_j$ be symmetric, which it obviously is. Next, the standardized bootstrap errors $\varepsilon_{i,n}^\star = u_i^\star / \tilde{\sigma}$ are iid, which satisfies another condition of the CLT. The last condition is

$$\sup_{1 \leq i \leq n, n \geq 1} E\left[|\varepsilon_{i,n}^\star|^{2+\eta}\right] < \infty$$

for some $\eta > 0$, which holds because $\kappa^\star < \infty$. Thus, the bootstrap F test is asymptotically valid when $\mu > 0$ and $0 < \rho \leq \mu$. Validity of the bootstrap LR and LM tests can be established by using the algebraic expressions linking them to the F statistic, see Anatolyev (2012), p. 378-379. $\diamondsuit$

Figures and tables

Figure 1. Null rejection frequencies of F tests, $n = 20$.

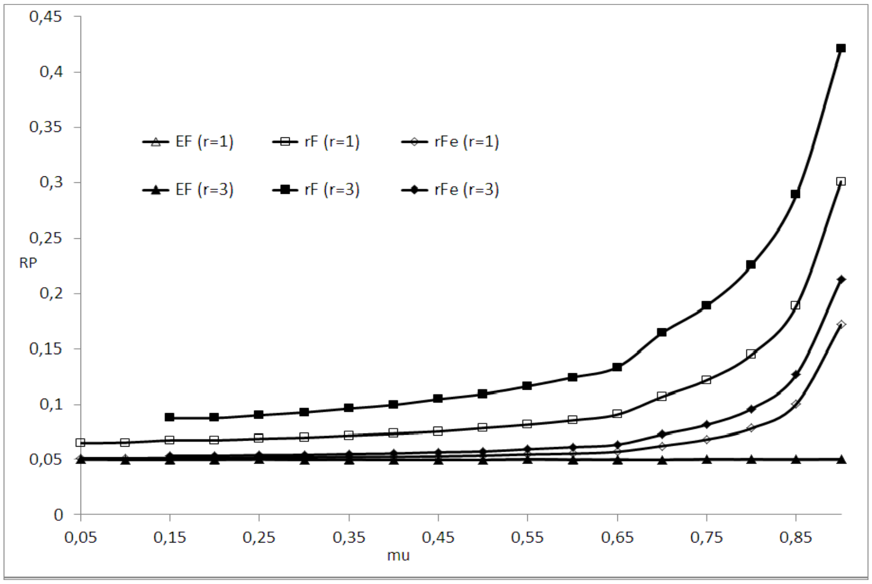

Figure 2. Null rejection frequencies of LR tests, $n = 20$.

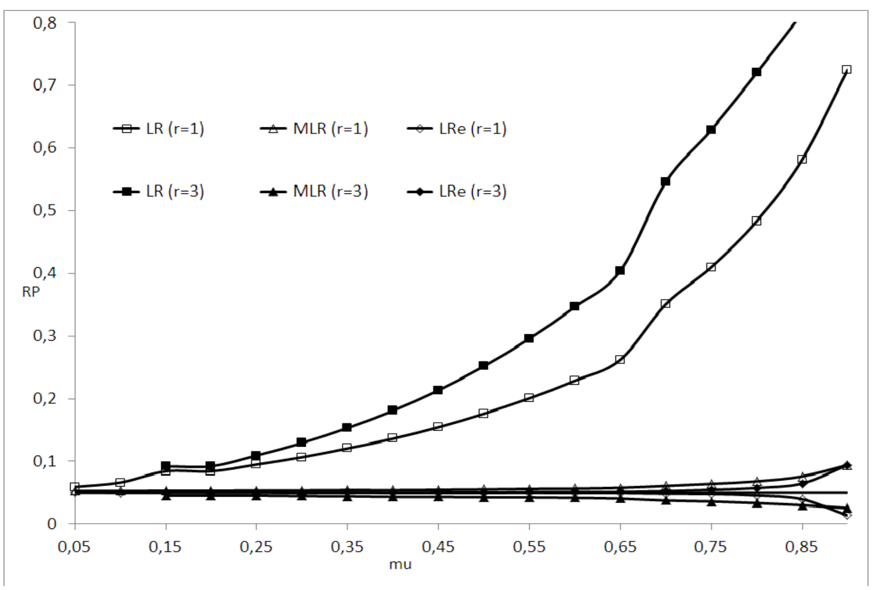

Figure 3. Null rejection frequencies of LM tests, $n = 20$.

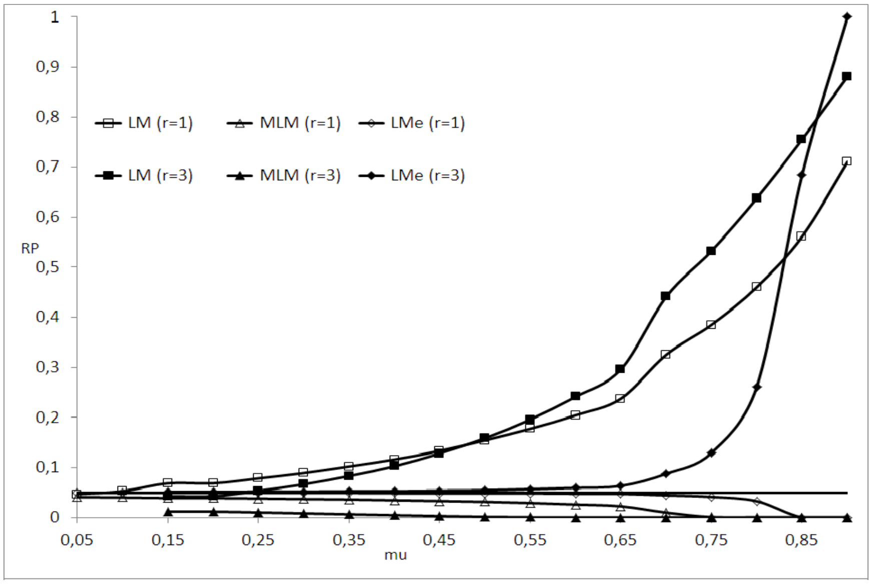

Figure 4. Null rejection frequencies of F tests, $n = 100$.

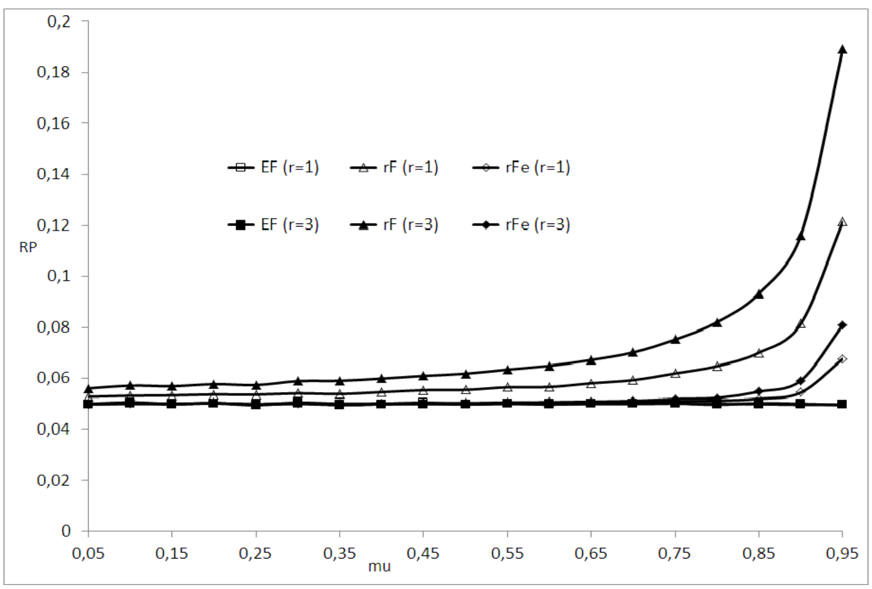

Figure 5. Null rejection frequencies of LR tests, $n = 100$.

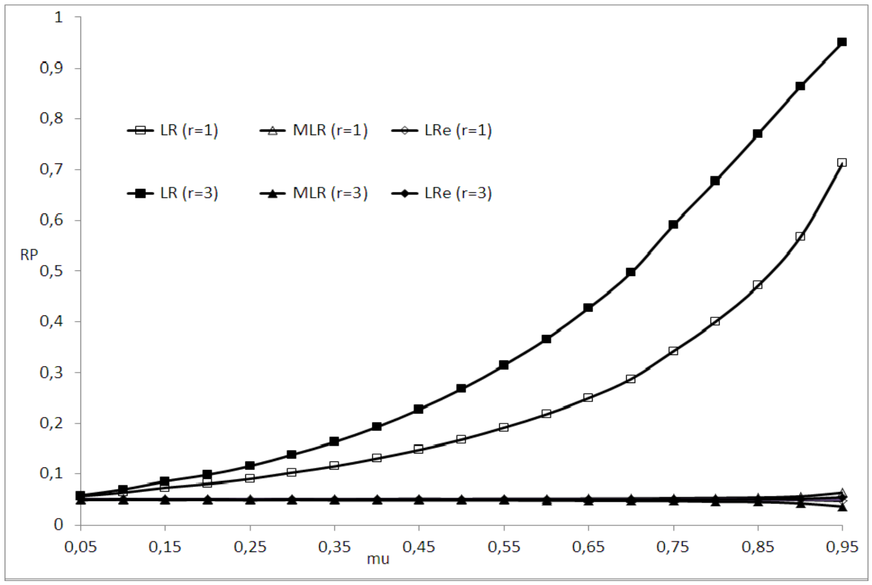

Figure 6. Null rejection frequencies of LM tests, $n = 100$.

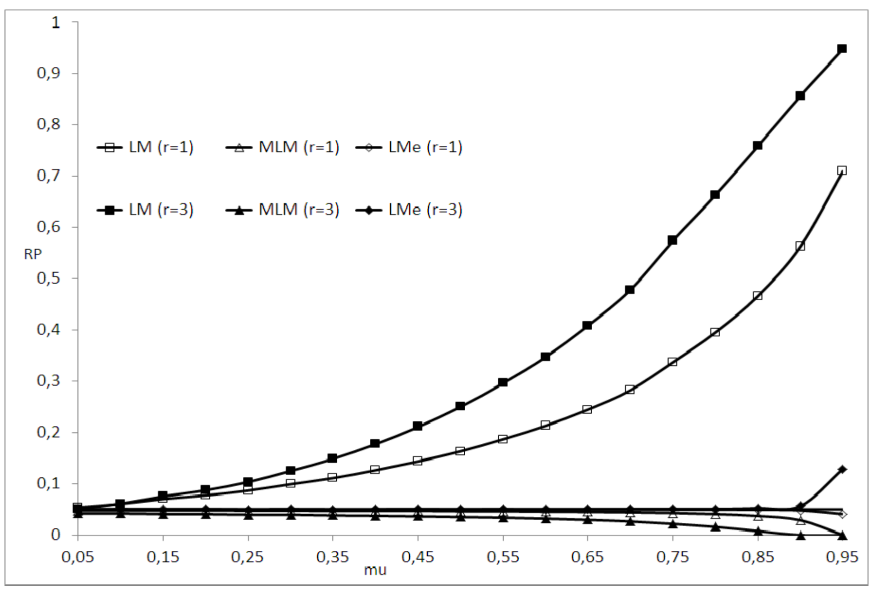

Table 1. Rejection frequencies of standard asymptotic tests.

n	μ	ρ	λ	EF	rF	LR	LM
20	0.1	0.1	0.1111	0.04987	0.07518	0.06719	0.04053
20	0.5	0.1	0.2	0.04971	0.09533	0.22367	0.16831
20	0.5	0.3	0.6	0.04981	0.15287	0.28432	0.03982
20	0.8	0.1	0.5	0.05010	0.16039	0.54946	0.49047
20	0.8	0.3	1.5	0.05007	0.26056	0.75013	0.40482
20	0.8	0.7	3.5	0.05027	0.33179	0.75024	0.00000
100	0.1	0.1	0.1111	0.04989	0.06618	0.07019	0.04054
100	0.5	0.1	0.2	0.05037	0.07928	0.45276	0.36585
100	0.5	0.3	0.6	0.05008	0.11821	0.59519	0.18829
100	0.8	0.1	0.5	0.04989	0.12008	0.92994	0.90381
100	0.8	0.3	1.5	0.05006	0.19354	0.99403	0.94649
100	0.8	0.7	3.5	0.04992	0.26492	0.99572	0.00701
250	0.1	0.1	0.1111	0.04968	0.06348	0.07799	0.04053
250	0.5	0.1	0.2	0.05029	0.07447	0.71964	0.62251
250	0.5	0.3	0.6	0.05018	0.11006	0.87622	0.40222
250	0.8	0.1	0.5	0.04998	0.10839	0.99830	0.99675
250	0.8	0.3	1.5	0.05016	0.17642	0.99999	0.99969
250	0.8	0.7	3.5	0.05011	0.24817	1.00000	0.07991
1000	0.1	0.1	0.1111	0.04927	0.06103	0.1072	0.04004
1000	0.5	0.1	0.2	0.05044	0.07116	0.99415	0.98219
1000	0.5	0.3	0.6	0.05075	0.10487	0.99985	0.91716
1000	0.8	0.1	0.5	0.05098	0.09878	1.00000	1.00000
1000	0.8	0.3	1.5	0.04924	0.16208	1.00000	1.00000
1000	0.8	0.7	3.5	0.04987	0.23403	1.00000	0.78232
∞	0.1	0.1	0.1111	0.05000	0.05930	1.00000	1.00000
∞	0.5	0.1	0.2	0.05000	0.06660	1.00000	1.00000
∞	0.5	0.3	0.6	0.05000	0.09670	1.00000	1.00000
∞	0.8	0.1	0.5	0.05000	0.08960	1.00000	1.00000
∞	0.8	0.3	1.5	0.05000	0.14910	1.00000	1.00000
∞	0.8	0.7	3.5	0.05000	0.21900	1.00000	1.00000

Table 2. Rejection frequencies of corrected and bootstrap tests.

n	μ	ρ	λ	crF	cLR	cLM	rFe	LRe	LMe	bootstrap
20	0.1	0.1	0.1111	0.06533	0.05804	0.05032	0.05206	0.04988	0.04995	0.04733
20	0.5	0.1	0.2	0.07743	0.06449	0.04934	0.05586	0.04971	0.05022	0.04985
20	0.5	0.3	0.6	0.09500	0.07108	0.04140	0.06067	0.05779	0.11112	0.04950
20	0.8	0.1	0.5	0.11865	0.08689	0.03840	0.07654	0.05010	0.06261	0.05036
20	0.8	0.3	1.5	0.14823	0.09669	0.01360	0.08918	0.08531	0.79869	0.04988
20	0.8	0.7	3.5	0.16059	0.10139	0.00528	0.08906	0.18249	1.00000	0.04990
100	0.1	0.1	0.1111	0.05649	0.05298	0.04928	0.05018	0.05023	0.05083	0.04968
100	0.5	0.1	0.2	0.06195	0.05579	0.04906	0.05120	0.05138	0.05336	0.05078
100	0.5	0.3	0.6	0.06761	0.05773	0.04706	0.04954	0.06148	0.08457	0.04927
100	0.8	0.1	0.5	0.07719	0.06242	0.04569	0.05379	0.05508	0.07367	0.05082
100	0.8	0.3	1.5	0.08745	0.06637	0.04176	0.05033	0.10101	0.47116	0.05100
100	0.8	0.7	3.5	0.09231	0.06851	0.03955	0.03969	0.30121	0.99959	0.05053
250	0.1	0.1	0.1111	0.05380	0.05154	0.04917	0.04966	0.05015	0.05057	0.04980
250	0.5	0.1	0.2	0.05737	0.05330	0.04932	0.05028	0.05167	0.05315	0.05270
250	0.5	0.3	0.6	0.06090	0.05471	0.04826	0.04773	0.06649	0.07978	0.04740
250	0.8	0.1	0.5	0.06639	0.05723	0.04734	0.05007	0.05731	0.07134	0.05060
250	0.8	0.3	1.5	0.07230	0.05964	0.04562	0.04311	0.12613	0.37968	0.04840
250	0.8	0.7	3.5	0.07536	0.06096	0.04454	0.02929	0.49546	0.99913	0.04990
1000	0.1	0.1	0.1111	0.05197	0.05085	0.04969	0.04987	0.05006	0.05009	0.04960
1000	0.5	0.1	0.2	0.05454	0.05239	0.05020	0.05003	0.05331	0.05356	0.04740
1000	0.5	0.3	0.6	0.05604	0.05278	0.04974	0.04713	0.08222	0.07670	0.04700
1000	0.8	0.1	0.5	0.05898	0.05432	0.04958	0.04879	0.06422	0.06952	0.04760
1000	0.8	0.3	1.5	0.05979	0.05368	0.04712	0.03756	0.22184	0.31395	0.04680
1000	0.8	0.7	3.5	0.06167	0.05502	0.04750	0.02176	0.91367	0.99800	0.05430
∞	0.1	0.1	0.1111	0.05000	0.05000	0.05000	0.04976	1.00000	0.05076	0.05000
∞	0.5	0.1	0.2	0.05000	0.05000	0.05000	0.04929	1.00000	0.05243	0.05000
∞	0.5	0.3	0.6	0.05000	0.05000	0.05000	0.04546	1.00000	0.07263	0.05000
∞	0.8	0.1	0.5	0.05000	0.05000	0.05000	0.04659	1.00000	0.0654	0.05000
∞	0.8	0.3	1.5	0.05000	0.05000	0.05000	0.03433	1.00000	0.25778	0.05000
∞	0.8	0.7	3.5	0.05000	0.05000	0.05000	0.01649	1.00000	0.99556	0.05000

Table 3. Power of corrected and bootstrap tests.

n	μ	ρ	λ	δ	EF	CrF	CLR	CLM	Bootstrap
20	0.5	0.1	0.2	0.25	0.08115	0.11979	0.10201	0.08063	0.08042
20	0.8	0.1	0.5	0.25	0.06037	0.13892	0.10289	0.04670	0.06066
20	0.8	0.7	3.5	0.25	0.05201	0.16527	0.10464	0.00547	0.05157
20	0.5	0.1	0.2	1	0.57475	0.66023	0.62520	0.57333	0.57042
20	0.8	0.1	0.5	1	0.22977	0.40987	0.33588	0.18886	0.22816
20	0.8	0.7	3.5	1	0.08071	0.23954	0.15656	0.00892	0.08085
20	0.5	0.1	0.2	2	0.96792	0.98049	0.97581	0.96768	0.96747
20	0.8	0.1	0.5	2	0.62109	0.80271	0.74143	0.56045	0.61647
20	0.8	0.7	3.5	2	0.19741	0.47201	0.34168	0.02836	0.19687
100	0.5	0.1	0.2	0.25	0.08239	0.09919	0.09013	0.08049	0.08395
100	0.8	0.1	0.5	0.25	0.06208	0.09430	0.07692	0.05709	0.06356
100	0.8	0.7	3.5	0.25	0.05242	0.09621	0.07150	0.04153	0.05033
100	0.5	0.1	0.2	1	0.77649	0.80486	0.79046	0.77302	0.76740
100	0.8	0.1	0.5	1	0.34532	0.43156	0.38779	0.32983	0.34230
100	0.8	0.7	3.5	1	0.09810	0.16688	0.12887	0.07966	0.09693
100	0.5	0.1	0.2	2	0.99989	0.99992	0.99990	0.99988	0.99993
100	0.8	0.1	0.5	2	0.92717	0.95388	0.94178	0.92077	0.92320
100	0.8	0.7	3.5	2	0.33546	0.47113	0.40101	0.29189	0.32527
250	0.5	0.1	0.2	0.25	0.08223	0.09239	0.08673	0.08087	0.08150
250	0.8	0.1	0.5	0.25	0.06234	0.08177	0.0709	0.05915	0.05750
250	0.8	0.7	3.5	0.25	0.05269	0.07887	0.06406	0.04691	0.04450
250	0.5	0.1	0.2	1	0.85742	0.87071	0.86350	0.85539	0.85750
250	0.8	0.1	0.5	1	0.39833	0.45400	0.42399	0.38806	0.38150
250	0.8	0.7	3.5	1	0.10319	0.14655	0.12227	0.09311	0.10050
250	0.5	0.1	0.2	2	1.00000	1.00000	1.00000	1.00000	1.00000
250	0.8	0.1	0.5	2	0.98264	0.98773	0.98523	0.98146	0.97860
250	0.8	0.7	3.5	2	0.39515	0.48219	0.43563	0.37211	0.38560
1000	0.5	0.1	0.2	0.25	0.08060	0.08560	0.08284	0.07948	0.08370
1000	0.8	0.1	0.5	0.25	0.06252	0.07176	0.06688	0.06100	0.06180
1000	0.8	0.7	3.5	0.25	0.05116	0.06328	0.05616	0.04904	0.05040
1000	0.5	0.1	0.2	1	0.92680	0.93112	0.92892	0.92588	0.92340
1000	0.8	0.1	0.5	1	0.45752	0.48644	0.47060	0.45204	0.45970
1000	0.8	0.7	3.5	1	0.10736	0.12880	0.11732	0.10292	0.10460
1000	0.5	0.1	0.2	2	1.00000	1.00000	1.00000	1.00000	1.00000
1000	0.8	0.1	0.5	2	0.99908	0.99928	0.99916	0.99908	0.99920
1000	0.8	0.7	3.5	2	0.45760	0.50040	0.47684	0.44788	0.45420

Table 4. Rejection frequencies with high skewness and kurtosis ($\mu = 0.5$, $\rho = 0.3$).

X and u are $t(1)$					
n	EF	CrF	CLR	CLM	Bootstrap
20	0.14781	0.19033	0.16938	0.13908	0.08640
100	0.24651	0.26035	0.25283	0.24392	0.09964
250	0.29684	0.30348	0.29974	0.29558	0.10560
1000	0.35159	0.35338	0.35232	0.35123	0.11300
X and u are $t(2)$					
n	EF	CrF	CLR	CLM	Bootstrap
20	0.07268	0.11794	0.09473	0.06421	0.05900
100	0.09947	0.11823	0.10803	0.09612	0.06753
250	0.11874	0.13007	0.12362	0.11665	0.07546
1000	0.15400	0.15800	0.15500	0.15300	0.08400
X and u are $t(3)$					
n	EF	CrF	CLR	CLM	Bootstrap
20	0.05723	0.10146	0.07838	0.04939	0.05746
100	0.06110	0.07943	0.06919	0.05798	0.05693
250	0.06206	0.07292	0.06671	0.06008	0.05626
1000	0.07000	0.07500	0.07200	0.06900	0.05300
X and u are $t(4)$					
n	EF	CrF	CLR	CLM	Bootstrap
20	0.05300	0.09658	0.07382	0.04540	0.05080
100	0.05346	0.07122	0.06124	0.05044	0.04880
250	0.05264	0.06344	0.05734	0.05073	0.05193
1000	0.04600	0.04800	0.04800	0.04500	0.04500
X and u are lognormal					
n	EF	CrF	CLR	CLM	Bootstrap
20	0.08331	0.12797	0.10546	0.07499	0.06610
100	0.09465	0.11402	0.10323	0.09111	0.06970
250	0.09298	0.10494	0.09822	0.09076	0.06480
1000	0.07850	0.08440	0.08090	0.07770	0.06100

Table 5. Rejection frequencies, heteroskedastic errors ($\rho = 0.5$, $\mu = 0.3$).

$\gamma = 0.25.$	EF	CrF	CLR	CLM	Bootstrap	CF0	CF1	CF2	CF3
$n = 20$	0.06808	0.12166	0.09414	0.05842	0.05344	0.51779	0.26274	0.15531	0.02085
$n = 100$	0.06292	0.08392	0.07203	0.05936	0.04520	0.76447	0.17980	0.10629	0.00083
$n = 250$	0.05964	0.07242	0.06519	0.05749	0.04400	0.92423	0.10815	0.06493	0.00000
$n = 1000$	0.05620	0.06260	0.05920	0.05520	0.04730	0.99980	0.02020	0.01120	0.00000
$\gamma = 1$	EF	CrF	CLR	CLM	Bootstrap	CF0	CF1	CF2	CF3
$n = 20$	0.17555	0.26148	0.21900	0.15852	0.05586	0.62372	0.37922	0.22684	0.03544
$n = 100$	0.12862	0.16119	0.14295	0.12270	0.04680	0.82403	0.26084	0.14783	0.00175
$n = 250$	0.10157	0.11942	0.10929	0.09872	0.04660	0.94425	0.15122	0.08562	0.00005
$n = 1000$	0.07800	0.08480	0.08080	0.07680	0.04630	1.00000	0.02840	0.01400	0.00000
$\gamma = 3$	EF	CrF	CLR	CLM	Bootstrap	CF0	CF1	CF2	CF3
$n = 20$	0.47458	0.56548	0.5222	0.4536	0.05140	0.71741	0.51395	0.29296	0.04267
$n = 100$	0.45018	0.50004	0.47287	0.44075	0.04450	0.88486	0.40726	0.20645	0.00572
$n = 250$	0.32388	0.35599	0.33817	0.31784	0.04380	0.96512	0.24815	0.11962	0.00025
$n = 1000$	0.16810	0.18030	0.17200	0.16640	0.04620	0.99999	0.04180	0.02030	0.00000

References

Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*, Wiley.

Anatolyev, S. (2012). "Inference in regression models with many regressors," *Journal of Econometrics*, **170**, 368-382.

Bickel, P.J. and D.A. Freedman (1983). "Bootstrapping regression models with many parameters," in P. Bickel, K. Doksum and J.L. Hodges (Eds), A Festschrift for Erich Lehmann, Wadsworth, Belmont, CA.

Chatterjee, S. and A. Bose (2002). "Dimension asymptotics for generalised bootstrap in linear regression," *Annals of the Institute of Statistical Mathematics,* **54**, 367-381.

Chescher, A. and I. Jewitt (1987). "The bias of a heteroskedasticity consistent covariance matrix estimator," *Econometrica*, **55**, 1217-1222.

Davidson, R. and E. Flachaire (2007). "Asymptotic and bootstrap inference for inequality and poverty measures," *Journal of Econometrics*, **141**, 141-166.

Davidson, R. and E. Flachaire (2008). "The wild bootstrap, tamed at last," *Journal of Econometrics*, **146**, 162-169.

Davidson, R. and J.G. MacKinnon (2002). "Fast double bootstrap tests of nonnested linear regression models," *Econometric Reviews*, **21**, 417-427.

Davidson, R. and J.G. MacKinnon (2004). *Econometric Theory and Methods*, Oxford University Press.

Davidson, R. and J.G. MacKinnon (2007). "Improving the reliability of bootstrap tests with the fast double bootstrap," *Computational Statistics and Data Analysis*, **51**, 3259-3281.

Evans, G.B.A. and N.E. Savin (1982). "Conflict among the criteria revisited: the W, LR and LM tests," *Econometrica*, **50**, 737-748.

Giacomini, R., D.N. Politis and H. White (2013). "A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators," *Econometric Theory*, **29**, 567-589.

Kelejian, H.H. and I.R. Prucha (2001). "On the asymptotic distribution of the Moran I test statistic with applications," *Journal of Econometrics*, **104**, 219-257.

Koenker, R. and J.A.F. Machado (1999). "GMM infernece when the number of moment conditions is large," *Journal of Econometrics*, **93**, 327-344.

Mammen, E. (1989). "Asymptotics with increasing dimension for robust regression with applications to the bootstrap," *Annals of Statistics*, **17**, 382-400.

Mammen, E. (1993). "Bootstrap and wild bootstrap for high-dimensional linear regression models," *Annals of Statistics*, **21**, 255-285.

Mammen, E. (1996). "Empirical process of residuals for high-dimensional linear models," *Annals of Statistics*, **24**, 307-335.

MacKinnon, J.G. and H. White (1985). "Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties," *Journal of Econometrics*, **29**, 305-325.

MacKinnon, J.G. (2013). "Thirty years of heteroskedasticity-robust inference," in Chen, X. and N.R. Swanson (Eds), Recent Advances and Future Directions in Causality, Prediction and Specification Analysis, Springer, New-York.

Portnoy, S. (1984). "Asymptotic behavior of M estimators of p regression parameters when p^2/n is large: I. Consistency," *Annals of Statistics*, **12**, 1298-1309.

Portnoy, S. (1985). "Asymptotic behavior of M estimators of p regression parameters when p^2/n is large: II. Normal approximation," *Annals of Statistics*, **13**, 1403-1417.

Portnoy, S. (1988). "Asymptotic behavior of likelihood methods for exponential families when the number of parameters tends to infinity," *Annals of Statistics*, **16**, 356-366.

Rothenberg, T. G. (1984). "Approximating the distributions of econometric estimators and test statistics," in Griliches, Z. and M.D. Intriligator (Eds.), Handbook of Econometrics, vol. 2. Norht-Holland, New-York.

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