Brazilian Journal of Probability and Statistics
2020, Vol. 34, No. 1, 18–34
https://doi.org/10.1214/18-BJPS412

Bootstrap-based testing inference in beta regressions

Fábio P. Lima and Francisco Cribari-Neto

Universidade Federal de Pernambuco

Abstract. We address the issue of performing testing inference in small samples in the class of beta regression models. We consider the likelihood ratio test and its standard bootstrap version. We also consider two alternative resampling-based tests. One of them uses the bootstrap test statistic replicates to numerically estimate a Bartlett correction factor that can be applied to the likelihood ratio test statistic. By doing so, we avoid estimation of quantities located in the tail of the likelihood ratio test statistic null distribution. The second alternative resampling-based test uses a fast double bootstrap scheme in which a single second level bootstrapping resample is performed for each first level bootstrap replication. It delivers accurate testing inferences at a computational cost that is considerably smaller than that of a standard double bootstrapping scheme. The Monte Carlo results we provide show that the standard likelihood ratio test tends to be quite liberal in small samples. They also show that the bootstrap tests deliver accurate testing inferences even when the sample size is quite small. An empirical application is also presented and discussed.

1 Introduction

The class of beta regression models introduced by Ferrari and Cribari-Neto (2004) is commonly used when the response variable is restricted to the interval $(0, 1)$, such as rates and proportions. In such a class of models, the response (y) is beta-distributed, that is, $y \sim \mathcal{B}(\mu, \phi)$. Its density function is given by

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \qquad 0 < y < 1,$$

where $\mu \in (0, 1)$ and $\phi > 0$ are the mean and the precision parameters, respectively. Here, $\mathbb{E}(y) = \mu$ and $\mathrm{Var}(y) = \mu(1-\mu)/(1+\phi)$. The mean response is related to a set of independent variables through a link function. In the fixed precision variant of the model, the precision parameter is taken to be constant for all observations. In the varying dispersion beta regression model, in contrast, such a parameter is not taken to be fixed and is modeled much in the same fashion as the mean parameter. Thus, the more general formulation of the beta regression model comprises of two submodels, one for the mean and another for the precision. In each submodel, the parameter is related to a linear predictor via a (strictly increasing and twice differentiable) link function. Such a formulation was considered by Smithson and Verkuilen (2006) and formally introduced by Simas, Barreto-Souza and Rocha (2010). Different residuals for the model were proposed by Espinheira, Ferrari and Cribari-Neto (2008) and diagnostic tools were developed by Ferrari, Espinheira and Cribari-Neto (2011). Beta regression model selection was considered by Bayer and Cribari-Neto (2015). Improved point and interval estimation in beta regressions was developed by Ospina, Cribari-Neto and Vasconcellos (2006) and non-nested testing inference in the same class of models was developed by Cribari-Neto and Lucena (2015). An extension of the model to accommodate the presence

Key words and phrases. Bartlett correction, beta regression, bootstrap, double bootstrap, fast double bootstrap, likelihood ratio test.

Received June 2017; accepted August 2018.

of zeros and/or ones in the data was proposed by Ospina and Ferrari (2012): the class of inflated beta regressions.

Testing inferences in beta regressions are usually carried out using the likelihood ratio test. Since the test statistic null distribution is typically unknown, the inference is based on an asymptotic approximation. As is well known, the test statistic is asymptotically distributed as χ^2_r under the null hypothesis, where r is the number of restrictions under evaluation. Such an approximation can be poor in small samples, thus rendering the test to be size distorted. Analytical finite samples have been derived and are available in the literature. For instance, Ferrari and Pinheiro (2011) derived Skovgaard's correction (Skovgaard, 2001) to the likelihood ratio test statistic and Bayer and Cribari-Neto (2013) obtained a Bartlett-corrected likelihood ratio test statistic (Bartlett, 1937). The latter, however, is only available for the fixed precision variant of the beta regression model. In this paper we consider an alternative approach, namely: the use of bootstrap resampling (Efron, 1979) to improve testing inferences accuracy. The underlying idea is to use pseudo-samples to estimate the test statistic exact null distribution, thus avoiding the use of an asymptotic approximation. The usual practice is to base the test decision on a bootstrap p-value which is then compared to the selected significance level. We go further and consider two alternative uses of bootstrap resampling, namely: (i) the use of data resampling to numerically estimate the Bartlett correction factor and (ii) the use of a second level of data resampling, in nested fashion, but in a less computer intensive variant known as "the fast double bootstrap". In short, we explore three different bootstrap-based testing strategies in the class of beta regression models. Their finite sample performances are evaluated via Monte Carlo simulations.

We note that it is important for practitioners to have at their disposal several reliable testing strategies that can deliver accurate inferences even when the sample size is small. For instance, in Section 7 we present an empirical application that involves modeling a natural gas usage simultaneity factor. There are only 42 observations available. The interest lies in determining whether or not precision is fixed. The likelihood ratio test cannot be trusted since it tends to be considerably liberal in small samples, as our simulation results indicate. The standard bootstrap p-value is equals 0.098, and as consequence the test is not conclusive at the 10% significance level. We then resort to the bootstrap Bartlett corrected test and to the fast double bootstrap test. Both tests indicate rejection of the null hypothesis (fixed precision) at the 10% significance level. For more details, see the empirical analysis in Section 7.

Our focus is on the likelihood ratio test because it is most commonly used test by practitioners. We note that alternative asymptotically chi-squared testing criteria are available in the literature, for example, the gradient, score and Wald tests. For a comparison of likelihood ratio, score and Wald testing inferences in beta regressions, see Cribari-Neto and Queiroz (2014). It is noteworthy that the bootstrap procedures we consider can be easily adapted for use with the aforementioned alternative tests.

The paper unfolds as follows. Section 2 introduces the class of beta regression models. The likelihood ratio test and its standard bootstrap variant are briefly reviewed in Section 3.

Section 4 presents the test based on the bootstrap Bartlett corrected likelihood ratio statistic. The fast double bootstrap test is presented in Section 5. In Section 6, we present and discuss Monte Carlo simulations results. Section 7 contains an empirical application. Finally, concluding remarks are offered in Section 8.

2 The beta regression model

Let $y_1, \ldots, y_n$ be a sample of independent random variables such that $y_t \sim \mathcal{B}(\mu_t, \phi_t)$, for $t = 1, \ldots, n$. The varying precision beta regression model is given by

$$g(\mu_t) = \sum_{i=1}^{k} x_{ti}\beta_i = x_t^\top \beta = \eta_t,$$
$$h(\phi_t) = \sum_{j=1}^{m} z_{tj}\gamma_j = z_t^\top \gamma = \delta_t,$$

where β and γ are unknown parameter vectors ($\beta \in \mathbb{R}^k$ and $\gamma \in \mathbb{R}^m$), $x_{t1}, \ldots, x_{tk}$ and $z_{t1}, \ldots, z_{tm}$ are fixed covariates ($k + m < n$) and $g(\cdot)$ and $h(\cdot)$ are link functions, which are strictly increasing and twice-differentiable.

Parameter estimation is typically carried out by maximum likelihood. The log-likelihood function based on an n-dimensional sample of independent beta responses is

$$\ell(\beta, \gamma) = \sum_{t=1}^{n} \ell_t(\mu_t, \phi_t),$$

where

$$\begin{aligned}\ell_t(\mu_t, \phi_t) = {} & \log \Gamma(\phi_t) - \log \Gamma(\mu_t\phi_t) - \log \Gamma\big((1-\mu_t)\phi_t\big) + (\mu_t\phi_t - 1)\log y_t \\ & + \big\{(1-\mu_t)\phi_t - 1\big\}\log(1 - y_t).\end{aligned}$$

The score functions are given by

$$U_\beta(\beta, \gamma) = X^\top V T (y^* - \mu^*),$$
$$U_\gamma(\beta, \gamma) = Z^\top H a,$$

where X is an $n \times k$ matrix and Z is an $n \times m$ matrix whose tth rows are $x_t^\top$ and $z_t^\top$, respectively, and V, T and H are diagonal matrices given by $V = \mathrm{diag}\{\phi_1, \ldots, \phi_n\}$, $T = \mathrm{diag}\{1/g'(\mu_1), \ldots, 1/g'(\mu_n)\}$ and $H = \mathrm{diag}\{1/h'(\phi_1), \ldots, 1/h'(\phi_n)\}$. Additionally, $y^* = \{y_1^*, \ldots, y_n^*\}^\top$ and $\mu^* = \{\mu_1^*, \ldots, \mu_n^*\}^\top$ with $y_t^* = \log(y_t/(1-y_t))$ and $\mu_t^* = \psi(\mu_t\phi_t) - \psi((1-\mu_t)\phi_t)$ where $\psi(\cdot)$ denotes the digamma function, that is, $\psi(w) = d\log\Gamma(w)/dw$, for $w > 0$, and $a = (a_1, \ldots, a_n)^\top$ with $a_t = \partial\ell_t(\mu_t, \phi_t)/\partial\phi_t$.

The maximum likelihood estimators, $\widehat{\beta}$ and $\widehat{\gamma}$, solve the system of equations given by $U_\beta(\beta, \gamma) = U_\gamma(\beta, \gamma) = 0$. Such a system of equations has no solution in closed-form and maximum likelihood estimates are typically obtained by numerically maximizing the log-likelihood function using a nonlinear optimization algorithm, such as the BFGS quasi-Newton algorithm; see Press et al. (1992) and Nocedal and Wright (2006) for details on numerical optimization.

Under regularity conditions (Serfling, 1980),

$$\begin{pmatrix}\widehat{\beta}\\ \widehat{\gamma}\end{pmatrix} \sim \mathcal{N}_{k+m}\left(\begin{pmatrix}\beta\\ \gamma\end{pmatrix}, K^{-1}\right),$$

where K^{-1} is the inverse information given by

$$K^{-1} = K^{-1}(\beta, \gamma) = \begin{pmatrix} K^{\beta\beta} & K^{\beta\gamma} \\ K^{\gamma\beta} & K^{\gamma\gamma}\end{pmatrix},$$

with

$$K^{\beta\beta} = (X^\top VWX - X^\top CTHZ(Z^\top DZ)^{-1}Z^\top HTC^\top X)^{-1},$$
$$K^{\beta\gamma} = (K^{\gamma\beta})^\top = -K^{\beta\beta}X^\top CTHZ(Z^\top DZ)^{-1},$$
$$K^{\gamma\gamma} = (Z^\top DZ)^{-1}\{I_m + (Z^\top HTC^\top X)K^{\beta\beta}X^\top CTHZ(Z^\top DZ)^{-1}\}.$$

Here, $W = \mathrm{diag}\{w_1, \ldots, w_n\}$, with

$$w_t = \phi_t\{\psi'(\mu_t\phi_t) + \psi'\big((1-\mu_t)\phi_t\big)\}\frac{1}{[g'(\mu_t)]^2},$$

$C = \mathrm{diag}\{c_1, \ldots, c_n\}$, with

$$c_t = \phi_t\big[\psi'(\mu_t\phi_t)\mu_t - \psi'\big((1-\mu_t)\phi_t\big)(1-\mu_t)\big],$$

and $D = \mathrm{diag}\{d_1, \ldots, d_n\}$, with

$$d_t = \psi'(\mu_t\phi_t)\mu_t^2 + \psi'\big((1-\mu_t)\phi_t\big)(1-\mu_t)^2 - \psi'(\phi_t),$$

where $\psi'(\cdot)$ is the trigamma function and I_m is the $m \times m$ identity matrix.

3 The likelihood ratio test and its bootstrap variant

Let $y = (y_1, \ldots, y_n)^\top$, where $y_t \sim \mathcal{B}(\mu_t, \phi_t)$, $t = 1, \ldots, n$, and let $\theta = (\beta^\top, \gamma^\top)^\top$ be the parameter vector that indexes the model. We partition θ as $\theta = (\omega^\top, \psi^\top)^\top$, where $\omega = (\omega_1, \ldots, \omega_r)^\top$ denotes the parameter of interest and $\psi = (\psi_1, \ldots, \psi_s)^\top$ denotes the nuisance parameter, with $r + s = m + k$. Suppose the interest lies in the test of $\mathcal{H}_0 : \omega = \omega^{(0)}$ against a two-sided alternative hypothesis, where $\omega^{(0)}$ is a given r-vector. The likelihood ratio (*LR*) test statistic is given by

$$LR = 2\big\{\ell(\widehat{\omega}, \widehat{\psi}) - \ell\big(\omega^{(0)}, \tilde{\psi}\big)\big\},$$

where hats (tildes) indicate evaluation at the unrestricted (restricted) maximum-likelihood estimator. Under regularity conditions and when the null hypothesis is true, the distribution of the likelihood ratio statistic converges to χ^2_r (Serfling, 1980). The test is then performed using approximate (asymptotic) χ^2 critical values, which can cause considerable size distortions in small samples.

In order to overcome this problem, Cribari-Neto and Queiroz (2014) consider a bootstrap version of the *LR* test for beta regression models, thus avoiding the use of asymptotic critical values. In this paper we shall go further and consider two additional bootstrap-based testing strategies, namely: the use of bootstrap Bartlett correction to the likelihood ratio test statistic and the fast double bootstrap. The relative merits of each approach in the class of beta regression model will be numerically evaluated via Monte Carlo simulations.

The decision rule of the bootstrap test can be expressed based on the bootstrap p-value. Let $y = (y_1, \ldots, y_n)^\top$ be a random sample from the random variable Y, τ be a test statistic, and $\widehat{\tau}$ be the realized value of τ for y. If $\mathbb{F}$ is the null cumulative distribution function of τ, the p-value of the test based on $\widehat{\tau}$ is

$$p(\widehat{\tau}) = 1 - \mathbb{F}(\widehat{\tau}).$$

However, in most cases, $\mathbb{F}$ is unknown. Bootstrap resampling can then be used to estimate $\mathbb{F}$, the resulting estimate being denoted by $\widehat{\mathbb{F}}^*_B$. The estimation procedure can be outlined as follows. At the outset and imposing the null hypothesis, we obtain B of pseudo-samples (or bootstrap samples) from the estimated model: $y^* = (y^*_1, \ldots, y^*_n)^\top$. Here, B is a large positive integer and by "estimated model" we mean the model with its parameters replaced by restricted estimates. We then compute the quantity of interest for each bootstrap sample, $\widehat{\tau}^* = \tau(y^*)$, and use the empirical distribution of $\widehat{\tau}^*$ as an estimate of the null distribution of τ.

The bootstrap p-value is given by

$$p^* = 1 - \widehat{\mathbb{F}}_B^*(\widehat{\tau}) = \frac{1}{B}\sum_{i=1}^{B} I\big(\widehat{\tau}_i^* > \widehat{\tau}\big), \tag{3.1}$$

where $I(\cdot)$ is an indicator function that equals 1 if its argument is true and 0 otherwise. We reject the null hypothesis if p^* is smaller than the test significance level.

The standard bootstrap version for the likelihood ratio test can be performed as follows:

(1) Compute the test statistic LR using the original sample;
(2) Generate a bootstrap sample y^* under the null hypothesis, with $y_t^* \sim \mathcal{B}(\tilde{\mu}_t, \tilde{\phi}_t)$, where $\tilde{\mu}_t = g^{-1}(x_t^\top \tilde{\beta})$ and $\tilde{\phi}_t = h^{-1}(z_t^\top \tilde{\gamma})$, $\tilde{\beta}$ and $\tilde{\gamma}$ being the restricted maximum likelihood estimators of β and γ, respectively;
(3) Fit the model using y^* and compute the test statistic LR^*;
(4) Execute steps 2 and 3 B times, where B is a large positive integer;
(5) Compute the bootstrap p-value:

$$p^*(LR) = \frac{1}{B}\sum_{i=1}^{B} I\big(LR_i^* > LR\big);$$

(6) Reject the null hypothesis if p^* is smaller than the test significance level.

4 Bootstrap Bartlett correction

Bartlett (1937) proposed a correction to the likelihood ratio statistic, which was later generalized by Lawley (1956). It makes it possible to reduces the order of the chi-square approximation error from $O(n^{-1})$ to $O(n^{-2})$. The modified LR statistic is given by

$$LR_{Bc} = \frac{LR}{c},$$

where $c = \mathbb{E}(LR)/r$ is known as the Bartlett correction factor. Oftentimes, however, the derivation of the Bartlett correction factor may be quite cumbersome. For instance, the Bartlett correction factor for beta regressions has only been obtained under fixed precision; see Bayer and Cribari-Neto (2013). The derivation of the Bartlett correction factor for varying precision beta regressions is quite cumbersome, and for that reason no such result is available so far. An alternative lies in the use of bootstrap resampling to estimate the correction factor; see Rocke (1989) and Bayer and Cribari-Neto (2013). The bootstrap Bartlett corrected LR statistic is given by

$$LR_{Bc}^* = \frac{rLR}{\overline{LR^*}},$$

where $\overline{LR^*} = \frac{1}{B}\sum_{i=1}^{B} LR_i^*$. For further details on Bartlett corrections, we refer readers to Cordeiro and Cribari-Neto (2014), Cribari-Neto and Cordeiro (1996) and the references therein.

The bootstrap Bartlett corrected likelihood ratio test can be performed as follows:

(1) Compute the test statistic LR using the original sample;

(2) Generate a bootstrap sample y^* under the null hypothesis, with $y_t^* \sim \mathcal{B}(\tilde{\mu}_t, \tilde{\phi}_t)$, where $\tilde{\mu}_t = g^{-1}(x_t^\top \tilde{\beta})$ and $\tilde{\phi}_t = h^{-1}(z_t^\top \tilde{\gamma})$. Here, $\tilde{\beta}$ and $\tilde{\gamma}$ are the restricted maximum likelihood estimators of β and γ, respectively;
(3) Fit the model using y^* and compute the test statistic LR^*;
(4) Execute steps 2 and 3 B times, where B is a large positive integer;
(5) Compute the bootstrap Bartlett corrected statistic:

$$LR_{Bc}^* = \frac{rLR}{\overline{LR^*}};$$

(6) Reject the null hypothesis at the α significance level if LR_{Bc}^* is larger than upper χ_r^2 $1-\alpha$ quantile, where r is the number of restrictions under test.

We note that an advantage of the above resampling strategy over the more traditional approach of using the bootstrap method to estimate the test critical value or its p-value is that here the bootstrap test statistic replicates are used to estimate a distribution mean: the mean of the null distribution of LR. A distribution mean can be more accurately estimated from a set of bootstrap test statistics replicates than a quantity that lies in the distribution tail.

5 The fast double bootstrap test

Davidson and MacKinnon (2000) proposed the fast double bootstrap (FDB) as an alternative to the standard double bootstrap, which is very costly computationally, thus allowing practitioners to achieve greater accuracy with reduced computational cost. This is possible because, whereas in the standard double bootstrap C second-level pseudo-samples are used for each first-level pseudo-sample, in the fast double bootstrap only one second-level bootstrap sample is used, which considerably reduces the computational burden of the resampling scheme.

We generate, under the null hypothesis, B first-level bootstrap samples y^* and compute $\widehat{\tau}_i^* = \tau(y_i^*)$, $i = 1, \dots, B$. Then, imposing the null hypothesis, we generate one second-level bootstrap sample y^{**} for each first-level pseudo-sample and compute $\widehat{\tau}_i^{**} = \tau(y_i^{**})$. The fast double bootstrap p-value is

$$p^{**}(\tau) = \frac{1}{B}\sum_{i=1}^{B} I\big(\widehat{\tau}_i^* > \widehat{Q}_B^{**}(1 - p^*(\tau))\big),$$

where $p^*(\tau)$ is as in Equation (3.1), and $\widehat{Q}^{**}(1 - p^*(\tau))$ is the $1 - p^*(\tau)$ quantile of all B second-level bootstrap test statistics $\widehat{\tau}_1^{**}, \dots, \widehat{\tau}_B^{**}$. It is noteworthy that the FDB based on B bootstrap replications is equivalent to the double bootstrap provided that the distribution of $\widehat{\tau}_{jl}^{**}$ does not depend on $\widehat{\tau}_j^*$, where $\widehat{\tau}_j^*$ and $\widehat{\tau}_{jl}^{**}$ are, respectively, the test statistics computed in the first and second levels of the double bootstrap mechanism and j and l index the first and second level bootstrap replications, respectively. Let B_1 and B_2 be the number of first and second level bootstrap replications in the double bootstrap scheme and let, as before, B denote the number of FDB replications. By equivalence we mean that the two p-values coincide when $B, B_1, B_2 \to \infty$. For further details, see MacKinnon (2006).

In the last few years, the FDB has been considered by several authors. Davidson and MacKinnon (2002) proposed a FDB version of the J test for use with non-nested linear regression models, Ouysse (2011) used the FDB for performing bias correction and Davidson and Trokic (2011) extended the FDB to higher orders of iteration. Other applications of the method can be found in Omtzigt and Fachin (2002), Davidson (2006) and MacKinnon (2006).

The fast double bootstrap version of the likelihood ratio test can be performed as follows:

(1) Compute the test statistic LR using the original sample;

(2) Generate a bootstrap sample y^* under the null hypothesis, with $y_t^* \sim \mathcal{B}(\tilde{\mu}_t, \tilde{\phi}_t)$, where $\tilde{\mu}_t = g^{-1}(x_t^\top \tilde{\beta})$ and $\tilde{\phi}_t = h^{-1}(z_t^\top \tilde{\gamma})$, $\tilde{\beta}$ and $\tilde{\gamma}$ being the restricted maximum likelihood estimators of β and γ, respectively;
(3) Fit the model using y^* and compute the test statistic LR^*;
(4) Generate a second-level bootstrap sample y^{**} under the null hypothesis, with $y_t^{**} \sim \mathcal{B}(\tilde{\mu}_t^*, \tilde{\phi}_t^*)$, where $\tilde{\mu}_t^* = g^{-1}(x_t^\top \tilde{\beta}^*)$ and $\tilde{\phi}_t^* = h^{-1}(z_t^\top \tilde{\gamma}^*)$. Here, $\tilde{\beta}^*$ and $\tilde{\gamma}^*$ are the restricted maximum likelihood estimators of β and γ, respectively, obtained using y^* as response;
(5) Fit the model using y^{**} and compute the test statistic LR^{**};
(6) Execute steps 2 through 5 B times, where B is a large positive integer;
(7) Compute the bootstrap p-value:

$$p^*(LR) = \frac{1}{B}\sum_{i=1}^{B} I\big(LR_i^* > LR\big);$$

(8) Compute $\widehat{Q}^{**}(1 - p^*(LR))$, the $1 - p^*(LR)$ quantile of all B second-level bootstrap test statistics $LR_1^{**}, \ldots, LR_B^{**}$;
(9) Compute the fast double bootstrap p-value:

$$p^{**}(LR) = \frac{1}{B}\sum_{i=1}^{B} I\big(LR_i^* > \widehat{Q}_B^{**}(1 - p^*(LR))\big);$$

(10) Reject the null hypothesis if p^{**} is smaller than the test significance level.

It is noteworthy that only one second level bootstrap sample is used at each first level bootstrap replication.

6 Numerical evidence

This section presents Monte Carlo simulation results on the finite sample performances of the likelihood ratio test (LR) and its bootstrap-based versions: standard bootstrap (LR^*), fast double bootstrap (LR^{**}) and bootstrap Bartlett corrected (LR^*_{Bc}) tests. The number of Monte Carlo replications is 5000, the tests nominal levels are $\alpha = 10\%$, 5% and 1%, and the number of bootstrap replications is 500.

At the outset, we consider the fixed precision beta regression model with logit link function given by

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}.$$

The covariates values were obtained as random standard uniform draws. We use three samples sizes, $n = 10, 20, 40$, three values for the precision parameter, $\phi = 20, 100, 500$, and three scenarios for μ: $\mu \in (0.020, 0.088)$, $\mu \in (0.20, 0.84)$ and $\mu \in (0.94, 0.98)$. The null hypothesis is $\mathcal{H}_0 : \beta_3 = 0$ which is tested against a two-sided alternative hypothesis. Since the precision is fixed, we also report results on the analytically Bartlett-corrected likelihood ratio test (LR_{Bc}); recall that such a correction is only available for fixed precision beta regressions (Bayer and Cribari-Neto, 2013).

Tables 1, 2 and 3 contain the tests null rejection rates. The results reveal that the bootstrap-based versions of the LR test outperform the test based on asymptotic critical values, which displays liberal finite sample behavior, that is, it tends to over-reject the null hypothesis. Moreover, although the LR_{Bc} test outperforms the usual LR test, it still displays some size distortion, especially when $n = 10$ and 20. For instance, when $\phi = 20$, $\mu \in (0.020, 0.088)$,

Table 1 *Null rejection rates (%) when testing $\mathcal{H}_0 : \beta_3 = 0$; $\phi = 20$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu \in (0.020, 0.088)$									
LR	18.80	14.20	11.52	12.04	8.40	6.18	4.34	2.32	1.36
LR_{Bc}	13.14	11.40	9.84	7.44	6.08	5.02	1.82	1.42	0.98
LR^*_{Bc}	11.04	10.66	9.96	5.60	5.58	5.02	1.26	1.28	1.00
LR^*	10.70	10.66	9.92	5.60	5.70	4.86	1.32	1.32	0.92
LR^{**}	10.86	10.36	10.12	5.86	5.42	5.08	1.36	1.22	0.96
$\mu \in (0.20, 0.84)$									
LR	18.52	12.86	11.84	11.18	6.90	5.94	3.80	1.86	1.46
LR_{Bc}	12.14	9.62	10.20	6.36	4.92	5.36	1.50	1.18	1.04
LR^*_{Bc}	10.02	9.20	10.26	4.86	4.66	5.14	1.04	1.10	1.00
LR^*	9.62	9.14	10.08	5.02	4.58	5.00	0.94	1.14	1.00
LR^{**}	10.16	9.34	10.14	5.02	4.84	4.98	1.00	1.16	1.00
$\mu \in (0.94, 0.98)$									
LR	16.84	13.18	11.92	10.08	6.82	6.22	3.34	2.04	1.50
LR_{Bc}	11.06	9.88	10.28	6.04	4.90	5.14	1.28	1.18	1.12
LR^*_{Bc}	9.22	9.46	10.14	4.44	4.72	5.12	1.04	1.06	1.08
LR^*	9.20	9.10	10.10	4.12	4.82	5.00	0.82	1.10	1.12
LR^{**}	9.52	9.40	10.26	4.64	4.76	5.20	0.98	1.28	1.06

Table 2 *Null rejection rates (%) when testing $\mathcal{H}_0 : \beta_3 = 0$; $\phi = 100$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu \in (0.020, 0.088)$									
LR	19.14	13.76	11.86	12.16	7.36	6.54	3.90	1.76	1.58
LR_{Bc}	13.10	10.48	10.30	6.98	4.98	5.22	1.62	0.96	1.18
LR^*_{Bc}	10.34	10.18	10.28	5.26	4.86	5.08	0.94	0.88	1.14
LR^*	10.20	9.96	10.24	5.18	4.70	5.16	1.04	0.98	1.12
LR^{**}	10.52	10.00	10.28	5.48	4.68	5.24	1.02	0.92	1.08
$\mu \in (0.20, 0.84)$									
LR	18.88	13.26	12.16	11.48	7.56	6.80	3.80	1.72	1.44
LR_{Bc}	12.20	10.32	10.64	6.66	5.38	5.84	1.72	1.02	1.06
LR^*_{Bc}	10.32	10.06	10.40	5.10	4.94	5.58	1.20	0.90	1.16
LR^*	10.04	9.90	10.32	4.96	4.96	5.50	1.06	0.90	1.12
LR^{**}	10.12	9.96	10.56	5.18	4.98	5.50	1.06	0.94	1.32
$\mu \in (0.94, 0.98)$									
LR	17.94	13.82	11.46	10.78	7.72	6.00	3.48	1.78	1.48
LR_{Bc}	11.88	10.82	9.94	6.52	5.38	5.02	1.34	1.10	1.04
LR^*_{Bc}	9.20	10.42	10.02	4.94	4.90	4.90	0.84	0.94	1.04
LR^*	9.16	10.20	9.84	4.70	4.84	4.90	0.80	0.92	1.12
LR^{**}	9.38	10.06	9.88	4.60	4.82	4.90	0.86	0.96	1.30

$\alpha = 10\%$ and $n = 20$, the null rejection rate of the LR test is 14.20% whereas those of LR_{Bc}, LR^*_{Bc}, LR^* and LR^{**} are 11.40%, 10.66%, 10.66% and 10.36%, respectively. It is noteworthy that, in most cases, the FDB and LR^*_{Bc} tests outperform the standard bootstrap test.

Table 3 *Null rejection rates (%) when testing $\mathcal{H}_0: \beta_3 = 0$; $\phi = 500$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu \in (0.020, 0.088)$									
LR	18.32	13.16	11.88	11.04	7.16	6.06	3.76	2.04	1.38
LR_{Bc}	11.86	10.20	10.36	6.54	5.32	4.88	1.56	1.26	1.06
LR^*_{Bc}	10.04	9.82	10.24	5.06	4.86	4.98	1.04	1.10	1.10
LR^*	10.04	9.78	10.16	5.10	4.66	4.94	1.22	1.10	1.10
LR^{**}	10.04	9.70	10.00	5.14	4.74	4.94	1.32	1.06	1.14
$\mu \in (0.20, 0.84)$									
LR	18.18	12.84	11.94	11.10	7.14	5.78	3.54	1.66	1.44
LR_{Bc}	12.12	10.26	10.38	6.66	4.90	4.94	1.64	0.94	1.08
LR^*_{Bc}	9.94	9.78	10.40	4.86	4.64	4.94	0.96	0.82	0.92
LR^*	10.08	9.82	10.20	4.82	4.54	4.80	0.88	0.78	0.88
LR^{**}	9.66	9.60	10.12	4.96	4.58	4.92	0.98	0.92	1.14
$\mu \in (0.94, 0.98)$									
LR	17.78	13.36	10.92	11.14	7.86	5.72	3.42	1.84	1.08
LR_{Bc}	11.88	10.56	9.56	5.80	5.74	4.82	1.70	1.18	0.80
LR^*_{Bc}	9.92	10.18	9.58	4.66	5.34	4.76	1.22	1.06	0.76
LR^*	9.56	10.16	9.66	4.44	5.14	4.72	1.06	1.08	0.64
LR^{**}	9.44	10.24	9.56	4.46	5.10	4.66	1.14	1.04	0.82

We have also carried out Monte Carlo simulations using a varying precision beta regression model. The following model was used:

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3},$$

$$\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2}.$$

The samples sizes are $n = 10, 20, 40$. We generated 10 values from the standard uniform distribution for $x_{ti} = z_{tj}$ and replicated them to get covariate values for the three samples sizes in order to keep the degree of heterogeneity (measured by $\lambda = \phi_{\max}/\phi_{\min}$) constant. We consider different scenarios for μ and three values for $\lambda = 20, 50, 100$. At the outset, we test $\mathcal{H}_0: \beta_3 = 0$ against $\mathcal{H}_1: \beta_3 \neq 0$.

Tables 4, 5 and 6 contain the tests null rejection rates. The asymptotic *LR* test is clearly liberal. For example, when $\lambda = 50$, $\mu \in (0.20, 0.80)$, $\alpha = 10\%$ and $n = 10$, its null rejection rate exceeds 22%. On the other hand, the bootstrap-based tests show minor size distortions relative to the asymptotic test, especially when $n = 10$ and 20. Moreover, in most cases, the LR^{**} test outperforms the LR^* and LR^*_{Bc} tests. For instance, when $\lambda = 100$, $\mu \in (0.95, 0.98)$, $n = 40$ e $\alpha = 10\%$, the null rejection rate of the LR^{**} test is 9.92% whereas those of LR^* and LR^*_{Bc} are 10.16% and 10.10%, respectively.

We shall now consider the test of $\mathcal{H}_0: \beta_2 = \beta_3 = 0$ (two restrictions). The results in the Tables 7, 8 and 9 show that the bootstrap-based tests outperform the asymptotic test. Moreover, the LR^* test is the best performer in two scenarios, namely: when $\lambda = 20$ and $\mu = 0.092$, and when $\lambda = 100$ and $\mu = 0.85$. In all other cases, the LR^{**} and LR^*_{Bc} tests outperform the standard bootstrap test.

We shall now perform testing inferences on the parameter vector that indexes the precision submodel. We consider the null hypothesis $\mathcal{H}_0: \gamma_2 = 0$ which is tested against $\mathcal{H}_1: \gamma_2 \neq 0$. The tests null rejection rates are presented in Table 10. The results show that in most cases the LR^{**} and LR^*_{Bc} tests outperform the LR^* test. For instance, when $\mu \in (0.92, 0.98)$, $\alpha = 5\%$ and $n = 10$, the null reject rate of the standard bootstrap test is 4.54% whereas those of

Table 4 *Null rejection rates (%) when testing $\mathcal{H}_0 : \beta_3 = 0$; $\lambda = 20$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu \in (0.020, 0.080)$									
LR	23.60	14.48	12.44	16.08	7.84	6.48	6.70	2.08	1.68
LR^*	10.02	9.86	10.50	4.86	4.78	5.30	1.10	1.16	1.20
LR^{**}	10.38	10.00	10.56	5.34	4.84	5.26	1.06	0.98	1.12
LR^*_{Bc}	9.92	9.70	10.40	4.76	4.92	5.30	0.98	1.10	0.90
$\mu \in (0.20, 0.80)$									
LR	23.22	15.12	11.94	15.30	8.22	6.40	5.94	2.22	1.50
LR^*	9.26	10.30	10.02	4.54	5.54	5.18	1.08	1.16	1.04
LR^{**}	9.60	10.14	10.10	4.84	5.14	5.06	0.96	0.96	1.04
LR^*_{Bc}	9.16	10.26	9.82	4.42	5.24	5.20	0.94	0.94	0.94
$\mu \in (0.95, 0.98)$									
LR	22.52	14.22	11.72	14.58	7.98	6.44	5.70	2.48	1.30
LR^*	8.94	9.96	10.16	4.58	5.12	4.90	1.10	1.50	0.90
LR^{**}	9.16	9.72	10.06	4.82	5.24	4.70	1.04	1.20	0.78
LR^*_{Bc}	8.90	9.86	10.30	4.54	5.02	4.78	0.86	1.32	0.76

Table 5 *Null rejection rates (%) when testing $\mathcal{H}_0 : \beta_3 = 0$; $\lambda = 50$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu \in (0.020, 0.080)$									
LR	22.82	14.82	12.52	15.22	8.88	6.82	5.98	2.58	1.72
LR^*	9.24	10.48	10.58	4.64	5.60	5.62	0.84	1.42	1.36
LR^{**}	9.44	10.32	10.56	4.70	5.48	5.68	0.86	1.10	1.16
LR^*_{Bc}	9.44	10.44	10.34	4.44	5.38	5.40	0.64	1.20	1.24
$\mu \in (0.20, 0.80)$									
LR	22.82	15.84	11.84	15.20	8.84	6.80	5.66	2.40	1.76
LR^*	9.88	10.56	10.06	4.78	5.16	5.42	1.28	1.30	1.26
LR^{**}	9.66	10.56	10.04	4.70	5.48	5.36	0.98	1.34	1.22
LR^*_{Bc}	9.86	10.36	10.06	4.48	5.10	5.40	1.04	1.02	1.14
$\mu \in (0.95, 0.98)$									
LR	23.36	14.92	11.76	15.28	9.40	6.46	5.90	2.52	1.58
LR^*	9.24	10.90	10.42	4.66	5.64	5.32	1.04	1.22	1.24
LR^{**}	9.36	10.88	10.30	4.50	5.34	5.46	0.92	1.28	1.04
LR^*_{Bc}	9.28	10.98	10.18	4.70	5.28	5.06	0.74	1.04	1.10

LR^{**} and LR^*_{Bc} are 4.98% and 4.74%, respectively. The asymptotic test is again considerably liberal.

Finally, we analyze the impact of the number of bootstrap replications on the bootstrap-based tests accuracy. The results in the Tables 11, 12 and 13 show that for $B = 250$, in most cases, the LR^*_{Bc} test outperforms the standard bootstrap and fast double bootstrap tests. However, with $B = 500$ the LR^{**} test outperforms the other two bootstrap-based tests. When $B = 1000$, the results are inconclusive. It is noteworthy that, in most cases, the FDB test based on only 500 replications outperforms the standard bootstrap test with $B = 1000$.

Table 6 *Null rejection rates (%) when testing* $\mathcal{H}_0 : \beta_3 = 0$; $\lambda = 100$

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu \in (0.020, 0.080)$									
LR	22.48	13.96	11.98	14.42	7.84	6.50	5.60	2.36	1.54
LR^*	8.96	9.28	10.14	4.56	4.74	5.22	1.10	1.14	1.20
LR^{**}	8.82	9.26	10.06	4.60	5.00	5.32	1.00	0.78	1.04
LR^*_{Bc}	9.00	9.36	9.84	4.46	4.72	5.10	0.88	0.86	1.04
$\mu \in (0.20, 0.80)$									
LR	23.88	15.02	12.30	15.92	8.42	6.96	6.02	2.48	1.74
LR^*	10.12	10.14	10.42	5.24	4.92	5.48	1.26	1.14	1.50
LR^{**}	9.98	10.06	10.44	5.22	4.96	5.56	1.12	1.06	1.16
LR^*_{Bc}	10.26	10.24	10.18	4.96	5.10	5.34	0.98	0.90	1.20
$\mu \in (0.95, 0.98)$									
LR	21.94	14.92	11.78	14.86	8.50	6.62	5.78	2.12	1.36
LR^*	9.62	10.10	10.16	4.88	5.12	5.36	0.96	1.06	1.02
LR^{**}	9.56	9.98	9.92	4.68	5.00	5.24	0.88	1.04	0.98
LR^*_{Bc}	9.64	9.96	10.10	4.88	4.90	5.16	0.82	0.94	0.80

Table 7 *Null rejection rates (%) when testing* $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$; $\lambda = 20$

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu = 0.092$									
LR	27.20	15.60	12.40	18.26	9.24	6.56	7.28	2.50	1.60
LR^*	9.84	9.86	9.68	4.98	4.86	5.18	1.26	1.28	1.14
LR^{**}	10.30	9.44	9.52	4.78	4.70	5.04	0.98	1.08	1.04
LR^*_{Bc}	9.56	9.50	9.54	4.76	4.66	4.90	0.96	1.10	0.88
$\mu = 0.85$									
LR	27.78	16.98	13.18	18.80	10.00	7.34	7.56	2.90	1.60
LR^*	10.24	10.74	10.62	5.24	5.20	5.72	1.02	1.32	1.26
LR^{**}	9.84	10.82	10.52	5.16	5.30	5.48	0.96	1.08	1.18
LR^*_{Bc}	10.06	10.48	10.48	5.06	5.12	5.38	0.86	1.20	1.04
$\mu = 0.95$									
LR	26.52	16.32	12.76	17.76	9.66	6.34	6.70	2.70	1.48
LR^*	9.48	10.40	9.94	4.38	4.94	4.98	0.98	1.12	1.06
LR^{**}	9.62	10.06	9.66	4.60	4.88	4.88	0.94	1.00	1.02
LR^*_{Bc}	9.48	10.26	10.04	4.16	4.86	5.00	0.86	0.92	0.90

7 Empirical application

In what follows, we shall present an empirical application of the tests considered in the previous sections. We use the data analyzed by Espinheira, Ferrari and Cribari-Neto (2014), which contain 42 observations on the distribution of natural gas for home usage in São Paulo, Brazil. Gas distribution is based on a simultaneity factor that assumes values in $(0, 1)$. It relates to the nominal power and to the number of natural gas-powered home appliances. Given these factors, the gas supplier seeks to forecast the probability of simultaneous appliances usage in order to decide how much gas to supply to a given residential unit. The response (y) is the simultaneity factor, which is used to obtains an indicator of gas release in a given tubulation section according to the following formula: $Q = F \times Q_{\max}$, where Q is the release, F is

Table 8 *Null rejection rates (%) when testing* $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$; $\lambda = 50$

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu = 0.092$									
LR	27.06	16.50	11.96	18.28	9.18	6.82	7.30	2.76	1.46
LR^*	9.90	9.80	10.12	5.10	5.02	5.54	1.32	1.26	1.16
LR^{**}	9.96	9.72	10.06	4.98	4.86	5.20	1.10	1.18	1.16
LR^*_{Bc}	9.76	9.82	10.06	4.88	5.02	5.30	1.02	0.94	0.96
$\mu = 0.85$									
LR	26.36	16.44	12.72	18.32	9.88	7.14	7.06	2.86	1.60
LR^*	9.86	10.58	10.14	4.94	5.52	5.30	1.04	1.22	1.24
LR^{**}	9.84	10.56	9.88	4.80	4.96	5.16	0.92	1.18	1.10
LR^*_{Bc}	10.02	10.50	9.84	4.84	5.64	5.10	0.64	1.10	0.98
$\mu = 0.95$									
LR	26.70	16.68	12.38	18.40	9.44	7.00	7.24	2.88	1.52
LR^*	10.38	10.18	10.02	5.20	5.36	5.40	1.18	1.26	1.24
LR^{**}	10.50	10.18	10.02	5.02	5.42	5.02	0.92	1.10	1.10
LR^*_{Bc}	10.36	10.00	9.96	5.12	5.32	5.08	0.96	1.00	1.02

Table 9 *Null rejection rates (%) when testing* $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$; $\lambda = 100$

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu = 0.092$									
LR	28.36	16.32	13.30	19.12	9.66	7.06	7.90	2.76	1.60
LR^*	10.90	10.34	10.54	5.46	5.22	4.92	1.24	1.18	1.20
LR^{**}	10.82	10.20	10.44	5.30	5.08	5.16	1.18	1.02	1.10
LR^*_{Bc}	10.80	10.02	10.28	5.24	4.84	4.84	0.92	0.98	1.08
$\mu = 0.85$									
LR	28.32	16.44	12.24	19.10	9.38	6.66	7.66	2.62	1.54
LR^*	10.40	10.10	9.64	5.50	5.02	5.06	1.04	1.04	1.20
LR^{**}	10.12	10.08	9.56	5.20	5.10	4.90	1.08	0.86	1.26
LR^*_{Bc}	10.30	10.08	9.50	5.18	4.80	4.88	0.90	1.00	1.08
$\mu = 0.95$									
LR	27.36	16.34	12.82	18.44	9.50	6.80	6.78	2.40	1.42
LR^*	9.96	10.18	10.16	4.90	4.80	5.08	1.12	1.08	1.14
LR^{**}	9.84	10.10	10.10	4.62	4.88	4.98	0.90	1.12	1.18
LR^*_{Bc}	9.72	10.22	10.22	4.80	4.54	4.96	0.80	0.92	0.98

a simultaneity factor and $Q_{\max}$ is the maximum possible release in the tubulation section. Espinheira, Ferrari and Cribari-Neto (2014) use as covariate the log of the release and model the data with the following fixed precision beta regression model:

$$M1 : \log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 \log(x_{t2}).$$

We shall consider the model given by

$$M2 : \log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 \log(x_{t2}),$$

$$\log(\phi) = \gamma_1 + \gamma_2 \log(x_{t2}).$$

Table 10 *Null rejection rates (%) when testing $\mathcal{H}_0 : \gamma_2 = 0$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$\mu \in (0.025, 0.13)$									
LR	26.36	14.88	11.96	18.76	8.98	6.82	8.80	2.70	1.46
LR^*	9.26	9.98	10.12	4.72	4.92	5.54	0.94	1.30	1.16
LR^{**}	9.42	10.18	10.06	4.90	4.86	5.20	0.92	1.16	1.16
LR^*_{Bc}	9.74	9.78	10.06	4.96	5.00	5.30	0.78	1.10	0.96
$\mu \in (0.28, 0.87)$									
LR	24.24	14.86	12.58	16.78	8.54	7.14	7.28	2.30	1.88
LR^*	9.32	9.96	10.86	4.62	4.82	5.62	1.12	1.06	1.50
LR^{**}	9.14	9.70	10.72	4.40	4.62	5.32	0.88	1.14	1.32
LR^*_{Bc}	9.54	9.92	10.62	5.02	4.72	5.44	1.18	1.00	1.26
$\mu \in (0.92, 0.98)$									
LR	24.84	14.34	12.50	17.54	8.32	6.66	8.04	2.22	1.56
LR^*	9.46	9.72	10.38	4.54	5.14	5.16	1.04	1.06	1.12
LR^{**}	10.24	9.66	10.32	4.98	5.12	5.20	0.90	0.96	1.04
LR^*_{Bc}	9.56	9.64	10.20	4.74	5.06	5.10	0.94	0.84	0.92

Table 11 *Null rejection rates (%) when $\mu \in (0.020, 0.080)$; $\lambda = 100$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$B = 250$									
LR^*	9.96	10.16	9.94	5.20	5.10	5.02	1.04	1.42	1.26
LR^{**}	10.24	9.94	9.70	5.42	5.14	4.98	1.38	1.28	1.18
LR^*_{Bc}	9.96	9.78	9.58	4.82	4.94	4.58	0.72	1.14	0.96
$B = 500$									
LR^*	8.96	9.28	10.14	4.56	4.74	5.22	1.10	1.14	1.20
LR^{**}	8.82	9.26	10.06	4.60	5.00	5.32	1.00	0.78	1.04
LR^*_{Bc}	9.00	9.36	9.84	4.46	4.72	5.10	0.88	0.86	1.04
$B = 1000$									
LR^*	9.62	9.92	10.06	4.82	4.72	5.20	0.98	0.90	1.14
LR^{**}	9.84	9.82	9.74	4.92	4.70	5.02	0.96	0.88	1.10
LR^*_{Bc}	9.76	9.80	9.96	4.74	4.64	5.14	0.80	0.74	1.08

Our interest lies in testing the null hypothesis of constant precision (i.e. $\mathcal{H}_0 : \gamma_2 = 0$). The null hypothesis is rejected at 10% significance level by the asymptotic version of the LR test, whose p-value equals 0.071. The bootstrap p-value equals 0.098; since it is very close to 0.10 the test is inconclusive at the 10% significance level. We shall then resort to the fast double bootstrap version of the LR test and also to the test based on the bootstrap Bartlett corrected LR statistic. The null hypothesis of constant precision is rejected by the LR^*_{Bc} test at de 10% significance level; its p-value equals 0.091. The conclusion is enhanced by the FDB p-value, 0.068, which allows one to reject the null hypothesis of fixed precision with more confidence. Table 14 presents parameter estimates and standard errors for Model $M2$.

Next, we evaluate the goodness-of-fit of the estimated varying precision regression model. At the outset, we carry out the RESET misspecification test (Ramsey, 1969, Pereira and Cribari-Neto, 2014). The test is performed by including $\widehat{\eta}_t^2$ as an additional regressor in the mean submodel and testing its exclusion. The test p-value is 0.929. Hence, there is no evidence of model misspecification. We also computed, for Models $M1$ and $M2$, the pseudo-R^2_{LR}

Table 12 *Null rejection rates (%) when $\mu \in (0.20, 0.80)$; $\lambda = 100$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$B = 250$									
LR^*	9.84	10.44	9.30	5.10	5.26	4.80	1.08	1.20	1.18
LR^{**}	9.70	10.02	9.08	4.74	5.42	4.78	0.98	1.16	1.16
LR^*_{Bc}	9.98	10.10	9.14	4.92	5.08	4.44	0.90	0.78	0.90
$B = 500$									
LR^*	10.12	10.14	10.42	5.24	4.92	5.48	1.26	1.14	1.50
LR^{**}	9.98	10.06	10.44	5.22	4.96	5.56	1.12	1.06	1.16
LR^*_{Bc}	10.26	10.24	10.18	4.96	5.10	5.34	0.98	0.90	1.20
$B = 1000$									
LR^*	10.40	10.70	10.24	5.54	5.62	5.20	1.08	1.36	1.26
LR^{**}	10.38	10.54	10.26	5.50	5.72	5.08	0.98	1.34	1.20
LR^*_{Bc}	10.48	10.60	10.28	5.40	5.56	5.12	1.06	1.20	1.14

Table 13 *Null rejection rates (%) when $\mu \in (0.95, 0.98)$; $\lambda = 100$*

	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
n	10	20	40	10	20	40	10	20	40
$B = 250$									
LR^*	10.14	10.06	10.64	5.12	5.06	5.82	1.22	0.96	1.28
LR^{**}	10.34	9.96	10.72	5.10	4.86	5.66	1.28	0.98	1.14
LR^*_{Bc}	9.94	9.78	10.28	4.92	5.04	5.30	0.84	0.72	0.92
$B = 500$									
LR^*	9.62	10.10	10.16	4.88	5.12	5.36	0.96	1.06	1.02
LR^{**}	9.56	9.98	9.92	4.68	5.00	5.24	0.88	1.04	0.98
LR^*_{Bc}	9.64	9.96	10.10	4.88	4.90	5.16	0.82	0.94	0.80
$B = 1000$									
LR^*	9.94	10.96	9.72	4.82	5.42	4.58	1.10	1.16	1.32
LR^{**}	10.10	10.64	9.50	4.60	5.44	4.56	0.82	1.08	1.10
LR^*_{Bc}	9.96	11.02	9.68	4.62	5.38	4.58	0.90	1.14	1.26

Table 14 *Parameter estimates and standard errors*

	β_1	β_2	γ_1	γ_2
Estimate	−1.717	−0.797	4.001	0.542
Standard error	0.091	0.085	0.326	0.296

(Nagelkerke, 1991), its version for varying precision beta regression models $\overline{R^2}_{LR}$ (Bayer and Cribari-Neto, 2017), and also the AIC (Akaike, 1973), the AICc (Hurvich and Tsai, 1989) and the BIC (Akaike, 1978, Schwarz, 1978). Table 15 contains the R^2_{LR}, $\overline{R^2}_{LR}$, AIC, AICc and BIC values for both models. The results favor Model $M2$. Finally, Figure 1 contains simulated envelope plots based on the standardized weighted residual 2 of Ferrari, Espinheira and Cribari-Neto (2011) for both models. Such plots indicate that both models fit the data well.

Table 15 *Pseudo-R^2 and model selection criteria*

Model	R^2_{LR}	$\overline{R^2}_{LR}$	AIC	AICc	BIC
$M1$	0.722	0.705	−170.779	−170.148	−165.566
$M2$	0.743	0.722	−172.031	−170.950	−165.081

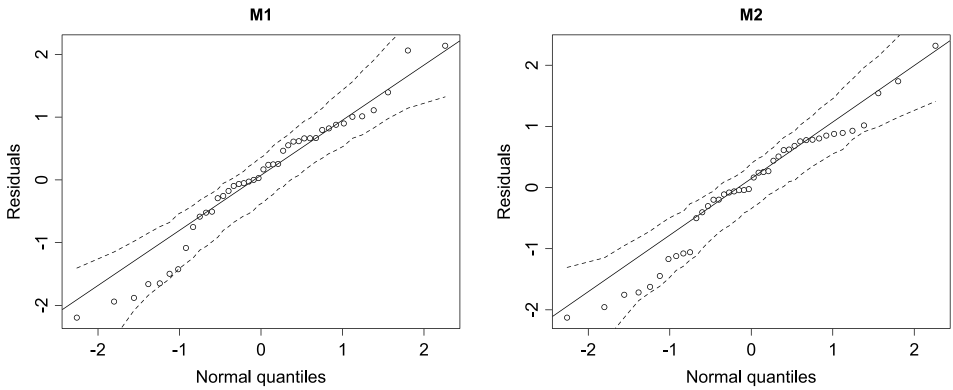

Figure 1 *Simulated envelope plots for Models $M1$ and $M2$.*

8 Concluding remarks

This paper addressed the issue of performing likelihood ratio testing inferences in beta regressions. In its standard formulation, the likelihood ratio test relies on an asymptotic approximation that may render the test inaccurate in small samples. In particular, the test may be considerably size distorted. Bootstrap resampling can be used to achieve more accurate inferences. The standard approach involves using the bootstrap test statistic replicates to compute a bootstrap p-value which can then be compared to the selected significance level. The null hypothesis is rejected if the former is smaller than the latter. In this paper, we consider two additional resampling strategies. The first involves using the bootstrap test statistic replicates to estimate the Bartlett correction factor which is then used to transform the test statistic. The second approach we pursue uses a nested resampling scheme in which one additional (second level) bootstrap sample is created based on each (first level) bootstrap sample. The inner bootstrap information is then used to obtain a p-value that is more accurate than the standard bootstrap p-value. The finite sample performances of the different bootstrap-based tests were evaluated using Monte Carlo simulations. An empirical application was presented and discussed.

Acknowledgments

The authors gratefully acknowledge partial financial support from CAPES and CNPq. We also thank an anonymous referee for useful comments and suggestions.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, 267–281. MR0483125

Akaike, H. (1978). A Bayesian analysis of the minimum AIC procedure. *Annals of the Institute of Statistical Mathematics* **30**, 9–14. MR0507075 https://doi.org/10.1007/BF02480194

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London Series A, Mathematical and Physical Sciences* **160**, 268–282.
Bayer, F. M. and Cribari-Neto, F. (2013). Bartlett corrections in beta regression models. *Journal of Statistical Planning and Inference* **143**, 531–547. MR2995113 https://doi.org/10.1016/j.jspi.2012.08.018
Bayer, F. M. and Cribari-Neto, F. (2015). Bootstrap-based model selection criteria for beta regressions. *Test* **24**, 776–795. MR3414517 https://doi.org/10.1007/s11749-015-0434-6
Bayer, F. M. and Cribari-Neto, F. (2017). Model selection criteria in beta regression with varying dispersion. *Communications in Statistics Simulation and Computation* **46**, 729–746. MR3563524 https://doi.org/10.1080/03610918.2014.977918
Cordeiro, G. M. and Cribari-Neto, F. (2014). *An Introduction to Bartlett Correction and Bias Reduction*, 1st ed. New York: Springer. MR3289997 https://doi.org/10.1007/978-3-642-55255-7
Cribari-Neto, F. and Cordeiro, G. M. (1996). On Bartlett and Bartlett-type corrections. *Econometric Reviews* **15**, 339–367. MR1423902 https://doi.org/10.1080/07474939608800361
Cribari-Neto, F. and Lucena, S. E. F. (2015). Nonnested hypothesis testing in the class of varying dispersion beta regressions. *Journal of Applied Statistics* **42**, 967–985. MR3315740 https://doi.org/10.1080/02664763.2014.993368
Cribari-Neto, F. and Queiroz, M. P. F. (2014). On testing inference in beta regressions. *Journal of Statistical Computation and Simulation* **84**, 186–203. MR3169320 https://doi.org/10.1080/00949655.2012.700456
Davidson, J. (2006). Alternative bootstrap procedures for testing cointegration in fractionally integrated processes. *Journal of Econometrics* **133**, 741–777. MR2252915 https://doi.org/10.1016/j.jeconom.2005.06.012
Davidson, R. and MacKinnon, J. G. (2000). Improving the reliability of bootstrap tests. Working Papers No. 995, Dept. Economics, Queen's Univ.
Davidson, R. and MacKinnon, J. G. (2002). Fast double bootstrap tests of nonnested linear regression models. *Econometric Reviews* **21**, 419–429. MR1951651 https://doi.org/10.1081/ETC-120015384
Davidson, R. and Trokic, M. (2011). The iterated bootstrap. Working Paper, McGill University.
Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* **7**, 1–26. MR0515681
Espinheira, P. L., Ferrari, S. L. P. and Cribari-Neto, F. (2008). On beta regression residuals. *Journal of Applied Statistics* **35**, 407–419. MR2420486 https://doi.org/10.1080/02664760701834931
Espinheira, P. L., Ferrari, S. L. P. and Cribari-Neto, F. (2014). Bootstrap prediction intervals in beta regressions. *Computational Statistics* **29**, 1263–1277. MR3266058 https://doi.org/10.1007/s00180-014-0490-5
Ferrari, S. L. P. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics* **31**, 799–815. MR2095753 https://doi.org/10.1080/0266476042000214501
Ferrari, S. L. P., Espinheira, P. L. and Cribari-Neto, F. (2011). Diagnostic tools in beta regression with varying dispersion. *Statistica Neerlandica* **65**, 337–351. MR2857878 https://doi.org/10.1111/j.1467-9574.2011.00488.x
Ferrari, S. L. P. and Pinheiro, E. C. (2011). Improved likelihood inference in beta regression. *Journal of Statistical Computation and Simulation* **81**, 431–443. MR2782138 https://doi.org/10.1080/00949650903389993
Hurvich, C. M. and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika* **76**, 297–307. MR1016020 https://doi.org/10.1093/biomet/76.2.297
Lawley, D. N. (1956). A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika* **43**, 295–303. MR0082237 https://doi.org/10.1093/biomet/43.3-4.295
MacKinnon, J. G. (2006). Applications of the fast double bootstrap. Working Papers No. 1023, Dept. Economics, Queen's Univ.
Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika* **78**, 691–692. MR1130937 https://doi.org/10.1093/biomet/78.3.691
Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*, 2nd ed. New York: Springer. MR2244940
Omtzigt, P. H. and Fachin, S. (2002). Bootstrapping and Bartlett corrections in the cointegrated VAR model. Econometrics Discussion Paper 15.
Ospina, R., Cribari-Neto, F. and Vasconcellos, K. L. P. (2006). Improved point and interval estimation for a beta regression model. *Computational Statistics & Data Analysis* **51**, 960–981. Errata: 55, 2011, 2445. MR2297500 https://doi.org/10.1016/j.csda.2005.10.002
Ospina, R. and Ferrari, S. L. P. (2012). A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis* **56**, 1609–1623. MR2892364 https://doi.org/10.1016/j.csda.2011.10.005
Ouysse, R. (2011). Computationally efficient approximation for the double bootstrap mean bias correction. *Economics Bulletin* **31**, 2388–2403.
Pereira, T. L. and Cribari-Neto, F. (2014). Detecting model misspecification in inflated beta regressions. *Communications in Statistics Simulation and Computation* **43**, 631–656. MR3200996 https://doi.org/10.1080/03610918.2012.712183
Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. New York: Cambridge University Press. MR1414682

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society B* **31**, 350–371. MR0290502

Rocke, D. M. (1989). Bootstrap Bartlett adjustment in seemingly unrelated regression. *Journal of the American Statistical Association* **84**, 598–601. MR1010351

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics* **6**, 461–464. MR0468014

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons. MR0595165

Simas, A. B., Barreto-Souza, W. and Rocha, A. V. (2010). Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis* **54**, 348–366. MR2756431 https://doi.org/10.1016/j.csda.2009.08.017

Skovgaard, I. M. (2001). Likelihood asymptotics. *Scandinavian Journal of Statistics* **28**, 3–32. MR1844348 https://doi.org/10.1111/1467-9469.00223

Smithson, M. and Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods* **11**, 54.

Departamento de Estatística
Universidade Federal de Pernambuco
50740-540, Recife/PE
Brazil
E-mail: fabiolima.fpl@gmail.com
cribari@gmail.com