*Article*

# Improved Inference on Cointegrating Vectors in the Presence of a near Unit Root Using Adjusted Quantiles

**Massimo Franchi [1] and Søren Johansen [2,3,\*]**

[1] Department of Statistical Sciences, Sapienza University of Rome, P.le A. Moro 5, 00198 Rome, Italy; massimo.franchi@uniroma1.it

[2] Department of Economics, University of Copenhagen, Øster Farimagsgade 5 Building 26, 1353 Copenhagen K, Denmark

[3] CREATES, Department of Economics and Business, Aarhus University, Building 1322, DK-8000 Aarhus C, Denmark

\* Correspondence: soren.johansen@econ.ku.dk; Tel.: +45-35-323-071

**Abstract:** It is well known that inference on the cointegrating relations in a vector autoregression (CVAR) is difficult in the presence of a near unit root. The test for a given cointegration vector can have rejection probabilities under the null, which vary from the nominal size to more than 90%. This paper formulates a CVAR model allowing for multiple near unit roots and analyses the asymptotic properties of the Gaussian maximum likelihood estimator. Then two critical value adjustments suggested by McCloskey (2017) for the test on the cointegrating relations are implemented for the model with a single near unit root, and it is found by simulation that they eliminate the serious size distortions, with a reasonable power for moderate values of the near unit root parameter. The findings are illustrated with an analysis of a number of different bivariate DGPs.

**Keywords:** long-run inference; test on cointegrating relations; likelihood inference; vector autoregressive model; near unit roots; Bonferroni type adjusted quantiles

**JEL Classification:** C32

## 1. Introduction

Elliott (1998) and Cavanagh et al. (1995) investigated the test on a coefficient of a cointegrating relation in the presence of a near unit root in a bivariate cointegrating regression. They show convincingly that when inference on the coefficient is performed as if the process has a unit root, then the size distortion is serious, see top panel of Figure A1 for a reproduction of their results. This paper analyses the $p$-dimensional cointegrated VAR model with $r$ cointegrating relations under local alternatives

$$\Delta y_t = (\alpha\beta' + T^{-1}\alpha_1 c\beta_1')y_{t-1} + \varepsilon_t, \qquad t = 1, \dots, T, \tag{1}$$

where $\alpha, \beta$ are $p \times r$ and $\varepsilon_t$ is i.i.d. $N_p(0, \Omega)$. It is assumed that $\alpha_1$ and $\beta_1$ are known $p \times (p - r)$ matrices of rank $p - r$, and $c$ is $(p - r) \times (p - r)$ and an unknown parameter, such that the model allows for a whole matrix, $c$, of near unit roots. We consider below the likelihood ratio test, $Q_\beta$, for a given value of $\beta$, calculated as if $c = 0$, that is, as if we have a CVAR with rank $r$. The properties of the test $Q_\beta$ can be very bad, when the actual data generating process (DGP) is a slight perturbation of the process generated by the model specified by $\alpha\beta'$. The matrix $\alpha\beta'$ describes a surface in the space of $p \times p$ matrices of dimension $p^2 - (p - r)^2$. Therefore a model is formulated that in some particular

"directions", given by the matrix $\alpha_1 c \beta_1'$, has a small perturbation of the order of $T^{-1}$ and $(p-r)^2$ extra parameters, $c$, that are used to describe the near unit roots.

A similar model could be suggested for near unit roots in the $I(2)$ model, see Di Iorio et al. (2016), but this will not be attempted here.

The model (1) contains as a special case the DGP used for the simulations in Elliott (1998), whe the errors are i.i.d. Gaussian and no deterministic components are present. The likelihood ratio test, $Q_\beta$, for $\beta$ equal to a given value, is derived assuming that $c = 0$ and analyzed when in fact near unit roots are present, $c \neq 0$. The parameters $\alpha, \beta,$ and $\Omega$ can be estimated consistently, but $c$ cannot, and this is what causes the bad behaviour of $Q_\beta$.

The matrix $\Pi(\alpha, \beta, c) = \alpha\beta' + T^{-1}\alpha_1 c \beta_1'$ is an invertible function of the $p^2$ parameters $(\alpha, \beta, c)$, see Lemma 1, so that the Gaussian maximum likelihood estimator in model (1) is least squares, and their limit distributions are found in Theorem 2. The main contribution of this paper, however, is a simulation study for the bivariate VAR with $p = 2, r = 1$. It is shown that two of the methods introduced by McCloskey (2017, Theorems Bonf and Bonf -Adj), for allowing the critical value for $Q_\beta$ to depend on the estimator of $c$, give a much better solution to inference on $\beta$, in the case of a near unit root. The results of McCloskey (2017) also allow for multivariate parameters and for more complex adjustments, but in the present paper we focus for the simulations on the case with $p = 2$ and $r = 1$, so there is only one parameter in $c$. In case $r = 1$, the matrix $I_p + \Pi$ is linear in $c \in \mathbb{R}$, and for $c = 0$, it has an extra unit root. Therefore there is a near unit root for $c \neq 0$, and we choose the vector $\alpha_1$ such $c \geq 0$ corresponds to the non-explosive near unit roots of interest.

The assumption that $\alpha_1$ and $\beta_1$ are known is satisfied under the null, in the DGP analyzed by Elliott, see (15) and (16). This is of course convenient, because $\alpha_1, \beta_1$ as free parameters, are not estimable.

Let $\theta$ denote the parameters $\alpha, \beta$ and $\Omega$ and let $\hat\theta$ and $\hat c$ denote the maximum likelihood estimators in model (1). For a given $\eta$ (here 5% or 10%), the quantile $c_{\theta,\eta}(c)$ is defined by $P_{c,\theta}\{\hat c \leq c_{\theta,\eta}(c)\} = \eta$. Simulations show that the quantile is increasing in $c$, and solving the inequality for $c$, a $1 - \eta$ confidence interval, $[0, c_{\theta,\eta}^{-1}(\hat c)]$, is defined for $c$. For given $\xi$ (here 90% or 95%) the quantile $q_{\theta,\xi}(c)$ is defined by $P_{c,\theta}\{Q_\beta \leq q_{\theta,\xi}(c)\} = \xi$, and McCloskey (2017) suggests replacing the critical value $q_{\theta,\xi}(c)$, by the stochastic critical value $q_{\theta,\xi}(c_{\theta,\eta}^{-1}(\hat c))$, or introducing the optimal $\xi$ by solving the equation

$$\max_{0 \leq c \leq \infty} P_{c,\theta}\left\{Q_\beta \leq q_{\theta,\xi}\left(c_{\theta,\eta}^{-1}(\hat c)\right)\right\} = v,$$

for a given nominal size $v$ (here 10%).

These methods are explained and implemented by a simulation study, and it is shown that they offer a solution to the problem of inference on $\beta$ in the presence of a near unit root.

## 2. The Vector Autoregressive Model with near Unit Roots

### 2.1. The Model

The model is given by (1) and the following standard $I(1)$ assumptions are made.

**Assumption 1.** *It is assumed that $r < p$, $c$ is $(p - r) \times (p - r)$, and that the equation*

$$\det\left(I_p(1 - z) - \alpha\beta'z\right) = 0$$

*has $p - r$ roots equal to one, and the remaining roots are outside the unit circle, such that $|eigen(I_r + \beta'\alpha)| < 1$. Moreover $\Pi = \alpha\beta' + T^{-1}\alpha_1 c \beta_1'$ has rank $p$ and*

$$\det\left(I_p(1 - z) - \alpha\beta'z - T^{-1}\alpha_1 c \beta_1'z\right) = 0, \tag{2}$$

*has all roots outside the unit circle for all $T \geq T_0$.*

For the asymptotic analysis we need condition (2) to hold for $T$ tending to $\infty$, and for the simulations, we need it to hold for $T = 100$. In model (1) with cointegrating rank $r$ and $\alpha_1$ and $\beta_1$ known, the number of free parameters in $\alpha$ and $\beta$ is $2pr - r^2 = p^2 - (p - r)^2$. The next result shows how the parameters $\alpha, \beta, c$ are calculated from $\Pi$. For any $p \times m$ matrix of rank $m < p$, we use the notation $a_\perp$ to indicate a $p \times (p - m)$ matrix of rank $p - m$, for which $a'_\perp a = 0$, and the notation $\bar{a} = a(a'a)^{-1}$.

**Lemma 1.** *Let* $\Pi = \alpha\beta' + T^{-1}\alpha_1 c\beta'_1$ *and let Assumption 1 be satisfied. Then, for $\beta$ normalized as $\beta'b = I_r$,*

$$\alpha = \Pi\beta_{1\perp}(\alpha'_{1\perp}\Pi\beta_{1\perp})^{-1}\alpha'_{1\perp}\Pi b, \tag{3}$$

$$\beta' = (\alpha'_{1\perp}\Pi b)^{-1}\alpha'_{1\perp}\Pi, \tag{4}$$

$$c = T(\beta'_1\Pi^{-1}\alpha_1)^{-1}. \tag{5}$$

To discuss the estimation we introduce the product moments of $\Delta y_t$ and $y_{t-1}$

$$S_{00} = T^{-1}\sum_{t=1}^{T}\Delta y_t \Delta y'_t, \qquad S_{11} = T^{-1}\sum_{t=1}^{T}y_{t-1}y'_{t-1}, \qquad S_{10} = S'_{01} = T^{-1}\sum_{t=1}^{T}y_{t-1}\Delta y'_t.$$

**Theorem 2.** *In model (1) with $\alpha_1$ and $\beta_1$ known, the Gaussian maximum likelihood estimator of $\Pi = \alpha\beta' + T^{-1}\alpha_1 c\beta'_1$ is the coefficient in a least squares regression of $\Delta y_t$ on $y_{t-1}$. For $\beta$ normalized on some $p \times r$ matrix $b$, $\beta'b = I_r$, the maximum likelihood estimators $(\alpha, \beta, c)$ are given in (3)–(5) by inserting $\hat{\Pi}$.*

*For $c = 0$, such that the rank of $\Pi$ is $r$, the likelihood ratio test for a given value of $\beta$ is*

$$Q_\beta = T\log\frac{\det\left(S_{00} - S_{01}\beta(\beta'S_{11}^{-1}\beta)\beta'S_{10}\right)}{\det\left(S_{00} - S_{01}\breve{\beta}(\breve{\beta}'S_{11}^{-1}\breve{\beta})\breve{\beta}'S_{10}\right)}, \tag{6}$$

*where the maximum likelihood estimator $\breve{\beta}$ is determined by reduced rank regression assuming the rank is $r$.*

### 2.2. Asymptotic Distributions

The basic asymptotic result for the analysis of the estimators and the test statistic is that $\alpha'_\perp y_t$ converges to an Ornstein-Uhlenbeck process. This technique was developed by Phillips (1988), and Johansen (1996, chp. 14) is used as a reference for details related to the CVAR. The results for the test statistic can be found in Elliott (1998).

Under Assumption 1, the process given by (1) satisfies

$$T^{-1/2}\alpha'_\perp y_{[Tu]} \xrightarrow{D} K(u),$$

where $K$ is the Ornstein-Uhlenbeck process

$$K(u) = \alpha'_\perp \int_0^u \exp\left\{\alpha_1 c\beta'_1 C(u - s)\right\}dW_\varepsilon(s),$$

$C = \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp$ and $W_\varepsilon$ is Brownian motion generated by the cumulated $\varepsilon_t$.

**Theorem 3.** *The test $Q_\beta$ for a given value of $\beta$, derived assuming $c = 0$, see (6), satisfies*

$$Q_\beta \xrightarrow{D} \chi^2_{(p-r)r} + B, \tag{7}$$

*where the stochastic noncentrality parameter*

$$B = tr\left\{ \beta_1 c' \zeta c \beta_1' \beta_\perp (\alpha_\perp' \beta_\perp)^{-1} \left( \int_0^1 KK' du \right) (\beta_\perp' \alpha_\perp)^{-1} \beta_\perp' \right\},\tag{8}$$

is independent of the $\chi^2$ distribution and has expectation

$$E(B) = tr\left\{ \beta_1 c' \zeta c \beta_1' C \left( \int_0^1 (1-v) \exp(v\tau C) \Omega \exp(vC'\tau') dv \right) C' \right\}.\tag{9}$$

Here $\zeta = \alpha_1' \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \alpha_1$ and $\tau = \alpha_1 c \beta_1'$, so it follows that $E(B) = 0$ if and only if $\alpha_1' \Omega^{-1} \alpha = 0$, in which case $Q_\beta \xrightarrow{D} \chi^2_{(p-r)r}$.

Let $\beta$ be normalized as $\beta' \beta_{1\perp} = I_r$. The asymptotic distribution of the estimators, $\hat{\alpha}$, $\hat{\beta}$, $\hat{c}$, see (3)–(5), are given as

$$T^{1/2}(\hat{\alpha} - \alpha) \xrightarrow{D} N_{p \times r}(0, \Sigma_{\beta\beta}^{-1} \otimes \Omega),\tag{10}$$

$$T(\hat{\beta} - \beta)' \beta_\perp \xrightarrow{D} (\alpha_{1\perp}' \alpha)^{-1} \alpha_{1\perp}' \int_0^1 (dW_\varepsilon) K' \left( \int_0^1 KK' du \right)^{-1} \alpha_\perp' \beta_\perp,\tag{11}$$

$$\hat{c} - c \xrightarrow{D} (\alpha_\perp' \alpha_1)^{-1} \alpha_\perp' \int_0^1 (dW_\varepsilon) K' \left( \int_0^1 KK' du \right)^{-1} \alpha_\perp' \beta_\perp (\beta_1' \beta_\perp)^{-1}.\tag{12}$$

Note that the asymptotic distributions of $\hat{\beta}$ and $\hat{c}$ given in (11) and (12) are not mixed Gaussian, because $\alpha_{1\perp}' W_\varepsilon(u)$ and $\alpha_\perp' W_\varepsilon(u)$ are not independent of $K(u)$, which is generated by $\alpha_\perp' \varepsilon_t$.

**Corollary 1.** *In the special case where $r = p - 1$, we choose $\alpha_1$ so that $c \geq 0$, and find*

$$E(B) = \frac{e^{2\delta c} - 1 - 2\delta c}{(2\delta)^2} \kappa \zeta,\tag{13}$$

*where*

$$\delta = \beta_1' C \alpha_1, \quad \kappa = \beta_1' C \Omega C' \beta_1, \quad \zeta = \alpha_1' \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \alpha_1.$$

## 3. Critical Value Adjustment for Test on $\beta$ in the CVAR with near Unit Roots

### 3.1. Bonferroni Bounds

In this section the method of McCloskey (2017, Theorem Bonf) is illustrated by a number of simulation experiments. The simulations are performed with data generated by a bivariate model (1), where $p = 2$ and $r = 1$. The direction $\alpha_1$ is chosen such that $c \geq 0$. The test $Q_\beta$ for a given value of $\beta$, is calculated assuming $c = 0$, see (6). The simulations of Elliott (1998), see Section 3.3, show that there may be serious size distortions of the test, depending on the value of $c$ and $\rho$, if the test is based on the quantiles from the asymptotic $\chi^2(1)$ distribution.

The methods of McCloskey (2017) consists in this case of replacing the $\chi^2(1)$ critical value with a stochastic critical value depending on $\hat{c}$, in order to control the rejection probability under the null hypothesis.

Let $\theta = (\alpha, \beta, \Omega)$ and let $P_{c,\theta}$ denote the probability measure corresponding to the parameters $c, \theta$. The method consists of finding the $\eta$ quantile of $\hat{c}$, see (5) with $\Pi$ replaced by $\hat{\Pi}$, as defined by

$$P_{c,\theta}\left( \hat{c} \leq c_{\theta,\eta}(c) \right) = \eta,$$

for $\eta = 5\%$ or $10\%$, say, and the $\xi$ quantile $q_{\theta,\xi}(c)$ of $Q_\beta$ as defined by

$$P_{c,\theta}\left(Q_\beta \leq q_{\theta,\xi}(c)\right) = \xi,$$

for $\xi = 90\%$ or $95\%$, say.

By simulation for given $\theta$ and a grid of given of values $c \in (c_1, \ldots, c_n)$, the quantiles $c_{\theta,\eta}(c_i)$ and $q_{\theta,\xi}(c_i)$ are determined. It turns out, that both $c_{\theta,\eta}(c)$ and $q_{\theta,\xi}(c)$ are increasing in $c$, see Figure A2. Therefore, a solution $c_{\theta,\eta}^{-1}(\hat{c})$ can be found such that

$$P_{c,\theta}\left\{\hat{c} > c_{\theta,\eta}(c)\right\} = P_{c,\theta}\left\{c \leq c_{\theta,\eta}^{-1}(\hat{c})\right\} = 1 - \eta. \tag{14}$$

This gives a $1 - \eta$ confidence interval $[0, c_{\theta,\eta}^{-1}(\hat{c})]$ for $c$, based on the estimator $\hat{c}$. Note that for $c \leq c_{\theta,\eta}^{-1}(\hat{c})$ it holds by monotonicity of $q_{\theta,\xi}(\cdot)$ that $q_{\theta,\xi}(c) \leq q_{\theta,\xi}(c_{\theta,\eta}^{-1}(\hat{c}))$, such that

$$P_{c,\theta}\left[Q_\beta > q_{\theta,\xi}\left\{c_{\theta,\eta}^{-1}(\hat{c})\right\} \text{ and } c \leq c_{\theta,\eta}^{-1}(\hat{c})\right] \leq P_{c,\theta}\left\{Q_\beta > q_{\theta,\xi}(c)\right\} \leq 1 - \xi,$$

but we also have

$$P_{c,\theta}\left[Q_\beta > q_{\theta,\xi}\left\{c_{\theta,\eta}^{-1}(\hat{c})\right\} \text{ and } c > c_{\theta,\eta}^{-1}(\hat{c})\right] \leq P_{c,\theta}\left[c > c_{\theta,\eta}^{-1}(\hat{c})\right] = \eta,$$

such that

$$P_{c,\theta}\left[Q_\beta > q_{\theta,\xi}\left\{c_{\theta,\eta}^{-1}(\hat{c})\right\}\right] \leq 1 - \xi + \eta.$$

In the paper from McCloskey (2017) it is proved under suitable conditions that we have the much stronger result

$$1 - \xi \leq \limsup_{T\to\infty} \sup_{0\leq c<\infty} P_{c,\theta}\left[Q_\beta > q_{\hat{\theta},\xi}\left\{c_{\hat{\theta},\eta}^{-1}(\hat{c})\right\}\right] \leq 1 - \xi + \eta.$$

Thus, the limiting rejection probability, for given $\theta$, of the test on $\beta$, calculated as if $c = 0$, but replacing the $\chi_{\xi}^2(1)$ quantile by the estimated stochastic quantile $q_{\hat{\theta},\xi}(c_{\hat{\theta},\eta}^{-1}(\hat{c}))$, lies between $1 - \xi$ and $1 - \xi + \eta$. In the simulations we set $\eta = 0.05$ and $\xi = 0.95$, so that the limiting rejection probability is bounded by 10%.

Note that $\theta$ is replaced by the consistent estimator $\hat{\theta}$. It obviously simplifies matters that in all the examples we simulate, it turns out that $c_{\theta,\eta}(c)$ is approximately linear and increasing in $c$, and $q_{\theta,\xi}(c)$ is approximately quadratic and increasing in $c$ for the relevant values of $c$, see Figure A2.

### 3.2. Adjusted Bonferroni Bounds

McCloskey (2017, Theorem Bonf-Adj) suggests determining by simulation on a grid of values of $c$ and $\xi$, the quantitity

$$\bar{P}_{\theta,\eta}(\xi) = \max_{0\leq c\leq\infty} P_{c,\theta}\left(Q_\beta > q_{\theta,\xi}(c_{\theta,\eta}^{-1}(\hat{c}))\right).$$

It turns out that $\bar{P}_{\theta,\eta}(\xi)$ is monotone in $\xi$, and we can determine for a given nominal size $\upsilon$ (here 10%)

$$\xi_{opt} = \bar{P}_{\theta,\eta}^{-1}(\upsilon).$$

The Adjusted Bonferroni quantile is then

$$q_{\theta,\xi_{opt}}\left(c_{\theta,\eta}^{-1}(\hat{c})\right),$$

and we find

$$P_{c,\theta}\left(Q_\beta > q_{\theta,\xi_{opt}}(c_{\theta,\eta}^{-1}(\hat{c}))\right) \leq v.$$

The result of McCloskey (2017, Theorem Bonf-Adj) is that under suitable assumptions

$$\limsup_{T\to\infty} \sup_{0\leq c<\infty} P_{c,\theta}\left[Q_\beta > q_{\hat{\theta},\xi_{opt}}\left\{c_{\hat{\theta},\eta}^{-1}(\hat{c})\right\}\right] = v,$$

where we illustrate the upper bound.

*3.3. The Simulation Study of Elliott (1998)*

The DGP is defined by the equations,

$$y_{1t} = \left(1 - \frac{c}{T}\right)y_{1t-1} + u_{1t}, \tag{15}$$

$$y_{2t} = \gamma y_{1t} + u_{2t}. \tag{16}$$

It is assumed that $u_t = (u_{1t}, u_{2t})'$ are i.i.d. $N_2(0, \Omega_u)$ with

$$\Omega_u = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

and the initial values are $y_{10} = y_{20} = 0$. The data $y_1, \ldots, y_T$ are generated from (15) and (16), and the test statistic $Q_\beta$ for the hypothesis $\gamma = 0$, is calculated using (6).

The DGP defined by (15) and (16) is contained in model (1) for $p = 2$. Note that $y_{2t} = \gamma(1 - c/T)y_{1t-1} + \gamma u_{1t} + u_{2t}$ such that

$$\alpha = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \beta = \begin{pmatrix} \gamma \\ -1 \end{pmatrix}, \alpha_1 = \begin{pmatrix} -1 \\ -\gamma \end{pmatrix}, \beta_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{17}$$

where the sign on $\alpha_1$ has been chosen such that $c \geq 0$. Finally $\varepsilon_{1t} = u_{1t}$ and $\varepsilon_{2t} = u_{2t} + \gamma u_{1t}$, and therefore

$$\Omega = \begin{pmatrix} 1 & \rho + \gamma \\ \rho + \gamma & 1 + \gamma^2 + 2\gamma\rho \end{pmatrix}.$$

For $c = 0$, the process $y_t = (y_{1t}, y_{2t})'$ is $I(1)$ and $\gamma y_{1t} - y_{2t}$ is stationary, and if $c/T$ is close to zero, $y_t$ has a near unit root.

Applying Corollary 1 to the DGP (15) and (16), the expectation of the test statistic $Q_\beta$ is found to be

$$E(Q_\beta) = p - 1 + \frac{e^{-2c} - 1 + 2c}{4}\frac{\rho^2}{1-\rho^2}, \tag{18}$$

which increases approximately linearly in $c$.

Based on $N = 1000$ simulations of errors $u_1, \ldots, u_T$, $T = 100$, the data $y_1, \ldots, y_T$, are constructed from the DGP for each combination of the parameters

$$(\gamma, c, \rho) \in [-0.5 : (0.01) : 0.5] \times [1 : (1) : 20] \times [-0.9 : (0.1) : 0.9],$$

where $[a : (b) : c]$ indicates the interval from $a$ to $c$ with step $b$. Based on each simulation, $\hat{c}$ and the test $Q_\beta$ for $\gamma = 0$ are calculated.

Top panel of Figure A1 shows the rejection probabilities of the test $Q_\beta$ as a function of $(c, \rho)$, using the asymptotic critical value, $\chi^2_{0.90}(1) = 2.71$, for a nominal rejection probability of 10%. The rejection probability increases with $|\rho|$ and with $c$. When $c = 10$ (corresponding to an autoregressive coefficient of $c/T = 0.9$) and $|\rho| = 0.7$, the size of the test $Q_\beta$ is around 50%, as found in Elliott (1998). The results are analogous across models with an unrestricted constant term, or with a constant restricted to the cointegrating space. In the paper by Elliott (1998) a number of tests are analyzed, and it was found that they were quite similar in their performance and similar to the above likelihood ratio test $Q_\beta$ from the CVAR with rank equal to 1.

### 3.4. Results with Bonferroni Quantiles and Adjusted Bonferroni Quantiles for $Q_\beta$

Data are simulated as above and first the rank test statistic, $Q_r$, see Johansen (1996, chp. 11) for rank equal to 1, is calculated. The rejection probabilities for a 5% test using $Q_r$ are given in the bottom panel of Figure A1 and they show that for $c = 20$, the hypothesis that the rank is 1, is practically certain to be rejected. If $c = 8$, the probability of rejecting that the rank is 1 is around 50%, so that plotting the rejection probabilities for $0 \le c \le 10$, covers the relevant values, see Figure A3.

For $\eta = 5\%$ and 10%, the quantiles $c_\eta(c)$ of $\hat{c}$ are reported in Figure A2 as a function of $c$. The quantiles $c_\eta(c)$ are nearly linear in $c$, and they are approximated by

$$\tilde{c}_\eta(c) = a_\eta + b_\eta c,$$

where the coefficients $(a_\eta, b_\eta)$ depend on $\eta$, which is used to construct the upper confidence limit in (14) as

$$\tilde{c}_\eta^{-1}(\hat{c}) = (\hat{c} - a_\eta)b_\eta^{-1}.$$

For $\xi = 90\%$ and 95%, the quantiles $q_{\rho,\xi}(c)$ of $Q_\beta$ are reported in Figure A2 as function of $c$ for four values of $\rho$. It is seen that for given $\rho$, the quantiles $q_{\rho,\xi}(c)$ are monotone and quadratic in $c$, for relevant values of $c$, and hence they can be approximated by

$$\tilde{q}_{\rho,\xi}(c) = f_{\rho,\xi} + g_{\rho,\xi}c + h_{\rho,\xi}c^2, \tag{19}$$

where the coefficients $(f_{\rho,\xi}, g_{\rho,\xi}, h_{\rho,\xi})$ depend on $\rho$ and $\xi$. The modified critical value is then constructed replacing $(c, \rho)$ by $(\tilde{c}_\eta^{-1}(\hat{c}), \hat{\rho})$ in (19), and thus one finds the adjusted critical value

$$\tilde{q}_{\hat{\rho},\xi,\eta}(\hat{c}) = f_{\hat{\rho},\xi} + g_{\hat{\rho},\xi}(\hat{c} - a_\eta)b_\eta^{-1} + h_{\rho,\xi}\left((\hat{c} - a_\eta)b_\eta^{-1}\right)^2 \tag{20}$$

which depends on estimated values, $\hat{c}$ and $\hat{\rho}$, and on discretionary values, $\xi$ and $\eta$.

The adjusted Bonferroni quantile is explained in Section 3.2. Simulations show that $\bar{P}_{\theta,\eta}(\xi)$ is linear in $\xi$ and the solution of the equation

$$\bar{P}_{\theta,\eta}(\xi) = v,$$

where $v = 0.10$ is the nominal size of the test, determines $\xi_{opt}$; the adjusted Bonferroni $q$-quantile is then found like (20) as

$$\tilde{q}_{\hat{\rho},\xi_{opt},\eta}(\hat{c}) = f_{\hat{\rho},\xi_{opt}} + g_{\hat{\rho},\xi_{opt}}(\hat{c} - a_\eta)b_\eta^{-1} + h_{\rho,\xi_{opt}}\left((\hat{c} - a_\eta)b_\eta^{-1}\right)^2, \tag{21}$$

where $\eta = 0.05$.

The rejection frequency of $Q_\beta$, the test for $\gamma = 0$, calculated using the $\chi^2(1)_{0.90}$ quantile, the Bonferroni quantile in (20) for $\xi = 95\%$ and $\eta = 5\%$ and the adjusted Bonferroni quantile in (21) for $\eta = 5\%$ is reported as a function of $c$ for four values of $\rho$ in Figure A3. For both corrections the rejection frequency is below the nominal size of 10%; hence both procedures are able to eliminate the serious size-distortions of the $\chi^2$ test. While the Bonferroni adjustment leads to rather conservative test with rejection frequency well below the nominal size, the adjusted Bonferroni procedure is closer to the nominal value. The power of the two procedures is shown in Figures A4 and A5 for values of $|\gamma| \leq 1/2$. It is seen that the better rejection probabilities in Figure A3 are achieved together with a reasonable power for $c \leq 5$, where the probability of rejecting the hypothesis of $r = 1$ is around 30%, see bottom panel of Figure A1. Notice that both tests become slightly biased, that is, the power functions are not flat around the null $\gamma = 0$.

In conclusion, the simulations indicate that the adjusted Bonferroni procedure works better than the simple Bonferroni, the reason being that the former relies on the joint distribution of $Q_\beta$ and $\hat{c}$.

### 3.5. A Few Examples of Other DGPs

Four other data generating processes are defined in Table 1, to investigate the role of different choices of $\alpha_1$ and $\beta_1$ for the results on improving the rejection probabilities for test on $\beta$ under the null and alternative. The DGPs all have $\alpha = -\beta = (-1, 1)'/2$. The vectors $\alpha_1$ and $\beta_1$ are chosen to investigate different positions of the near unit root in the DGP.

The choice of DGP turns out to be important also for the test, $Q_r$, for $r = 1$. In fact the probability of rejecting $r = 1$ is around 50% for DGP 1 if $c = 4$, for DGP 2 if $c = 20$, whereas for DGP 3 and 4 the 50% value value is 8.

The rejection probabilities in Figure A6 are plotted for $0 \leq c \leq 10$, to cover the most relevant values.

The results are summarized in Figures A6–A8. It is seen that the conclusions from the study of the DGP analyzed by Elliott seem to be valid also for other DGPs. For moderate values of $c$, using the Bonferroni quantiles gives a rather conservative test while the adjusted Bonferroni procedure is closer to the nominal size and the power curves look reasonable for $c \leq 5$, although the tests are slightly biased, except for DGP 1. For this DGP, $\alpha_1 = \beta_1 = (1, 1)'$, $\Omega = I_2$, such that $\alpha_1' \Omega^{-1} \alpha = 0$, which means that the asymptotic distribution of $Q_\beta$ is $\chi^2(1)$, see Theorem 3, despite the near unit root. It is seen from Figure A6, there is only moderate distortion of the rejection probability in this case and in Figures A7 and A8, the power curves are symmetric around $\gamma = 0$, so the tests are approximately unbiased.

**Table 1.** The matrix $\Pi$ for four different DGPs given by $\alpha = -\beta = (-1, 1)'/2$ which are the basis for the simulations of rejection probabilities for the adjusted test for $\beta = (1, -1)'/2$. The positions of $c/T$ give the different $\alpha_1$ and $\beta_1$.

| Four DGPs Allowing for near Unit Roots, $\Omega = I_2$ |
| :---: |

1: $\begin{pmatrix} -\frac{1}{4} - c/T & \frac{1}{4} - c/T \\ \frac{1}{4} - c/T & -\frac{1}{4} - c/T \end{pmatrix}$　　2: $\begin{pmatrix} -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} - c/T \end{pmatrix}$

3: $\begin{pmatrix} -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} - c/T & -\frac{1}{4} - c/T \end{pmatrix}$　　4: $\begin{pmatrix} -\frac{1}{4} - c/T & \frac{1}{4} - c/T \\ \frac{1}{4} & -\frac{1}{4} \end{pmatrix}$

## 4. Conclusions

It has been demonstrated that for the DGP analyzed by Elliott (1998), it is possible to apply the methods of McCloskey (2017) to adjust the critical value in such a way that the rejection probabilities of the test for $\beta$ are very close to the nominal values. By simulating the power of the test for $\beta$, it is seen that for $c \leq 5$, the test has a reasonable power. Some other DGPs have been investigated and similar results have been found.

**Author Contributions:** Both authors contributed equally to the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Proofs

**Proof of Lemma 1.** Multiplying $\Pi = \alpha\beta' + T^{-1}\alpha_1 c\beta_1'$ by $\beta_{1\perp}$, we find

$$\Pi\beta_{1\perp} = \alpha\beta'\beta_{1\perp}. \tag{A1}$$

Multiplying $\Pi$ by $\alpha_{1\perp}'$ we find

$$\alpha_{1\perp}'\Pi = \alpha_{1\perp}'\alpha\beta' = \alpha_{1\perp}'\Pi\beta_{1\perp}(\beta'\beta_{1\perp})^{-1}\beta'. \tag{A2}$$

Multiplying by $b$ we find

$$(\beta'\beta_{1\perp})^{-1} = (\alpha_{1\perp}'\Pi\beta_{1\perp})^{-1}\alpha_{1\perp}'\Pi b.$$

It follows that from (A2)

$$\beta' = \beta'\beta_{1\perp}(\alpha_{1\perp}'\Pi\beta_{1\perp})^{-1}\alpha_{1\perp}'\Pi = (\alpha_{1\perp}'\Pi b)^{-1}\alpha_{1\perp}'\Pi$$

and from (A1)

$$\alpha = \Pi\beta_{1\perp}(\beta'\beta_{1\perp})^{-1} = \Pi\beta_{1\perp}(\alpha_{1\perp}'\Pi\beta_{1\perp})^{-1}\alpha_{1\perp}'\Pi b,$$

which proves (3) and (4).

Inserting these results in the expression for $\Pi$, we find using $\alpha_{1\perp}'\Pi b\beta' = \alpha_{1\perp}'\alpha\beta'b\beta' = \alpha_{1\perp}'\Pi$

$$\Pi = \alpha\beta' + T^{-1}\alpha_1 c\beta_1' = \Pi\beta_{1\perp}(\alpha_{1\perp}'\Pi\beta_{1\perp})^{-1}\alpha_{1\perp}'\Pi + T^{-1}\alpha_1 c\beta_1'. \tag{A3}$$

Next $\Pi$ is decomposed using

$$\Pi = \Pi\beta_{1\perp}(\alpha_{1\perp}'\Pi\beta_{1\perp})^{-1}\alpha_{1\perp}'\Pi + \alpha_1(\beta_1'\Pi^{-1}\alpha_1)^{-1}\beta_1', \tag{A4}$$

which is proved by premultiplying (A4) by $\alpha_{1\perp}'$ and $\beta_1'\Pi^{-1}$. Subtracting (A3) and (A4) and multiplying by $\bar{\alpha}_1'$ and $\bar{\beta}_1$, it is seen that

$$(\beta_1'\Pi^{-1}\alpha_1)^{-1} = c/T.$$

$\square$

**Proof of Theorem 2.** The unrestricted maximum likelihood estimator of $\Pi$ is $\hat{\Pi} = S_{01}S_{11}^{-1}$, and $\hat{\Omega} = S_{00} - S_{01}S_{11}^{-1}S_{10}$, and the results for $\hat{\alpha}, \hat{\beta}, \hat{c}$ follow from Lemma 1. If $c = 0$, the maximum likelihood estimator $\breve{\beta}$ can be determined by reduced rank regression, see (Johansen (1996, chp. 6)). $\square$

**Proof of Theorem 3.** Proof of (7) and (8): The limit results for the product moments are given first, using the normalization matrix $C_T = (\beta, T^{-1/2}\alpha_\perp)$ and the notation $S_{1\varepsilon} = T^{-1}\sum_{t=1}^T y_{t-1}\varepsilon_t'$,

$$C_T'S_{11}C_T = \begin{pmatrix} \beta'S_{11}\beta & T^{-1/2}\beta'S_{11}\alpha_\perp \\ T^{-1/2}\alpha_\perp'S_{11}\beta & T^{-1}\alpha_\perp'S_{11}\alpha_\perp \end{pmatrix} \xrightarrow{D} \begin{pmatrix} \Sigma_{\beta\beta} & 0 \\ 0 & \int_0^1 KK' du \end{pmatrix}, \tag{A5}$$

$$T^{1/2}C_T'S_{1\varepsilon} = \begin{pmatrix} T^{1/2}\beta'S_{1\varepsilon} \\ T^{-1}\alpha_\perp'S_{1\varepsilon} \end{pmatrix} \xrightarrow{D} \begin{pmatrix} N_{r\times p}(0, \Omega\otimes\Sigma_{\beta\beta}) \\ \int_0^1 K(dW_\varepsilon)' \end{pmatrix}. \tag{A6}$$

The test for a known value of $\beta$ is given in (6). It is convenient for the derivation of the limit distribution of $Q_\beta$, to normalize $\breve{\beta}$ on the matrix $\alpha(\beta'\alpha)^{-1}$, such that $\breve{\beta}'\alpha(\beta'\alpha)^{-1} = I_r$, and define $\breve{\theta} = (\beta'_\perp \alpha_\perp)^{-1}\beta'_\perp (\breve{\beta} - \beta)$. This gives the representation

$$\breve{\beta} - \beta = \alpha_\perp (\beta'_\perp \alpha_\perp)^{-1}\beta'_\perp (\breve{\beta} - \beta) + \beta(\alpha'\beta)^{-1}\alpha'(\breve{\beta} - \beta) = \alpha_\perp \breve{\theta}.$$

The proof under much weaker conditions can be found in Elliott (1998), and is just sketched here. The estimator for $\theta$ for known $\alpha, \Omega$ and $c = 0$, is given by the equation

$$T\breve{\theta} = (\alpha'_\perp T^{-1}S_{11}\alpha_\perp)^{-1}(\alpha'_\perp S_{1\varepsilon} + \alpha'_\perp T^{-1}S_{11}\beta_1 c\alpha'_1)\alpha_\Omega,$$

where $\alpha_\Omega = \Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1}$. The limit distribution of $T\breve{\theta}$ follows from (A5) and (A6) as follows. Because $T^{-1}\alpha'_\perp S_{11}\beta \xrightarrow{P} 0$ it follows that

$$\alpha'_\perp T^{-1}S_{11}\beta_1 c\alpha'_1 = \alpha'_\perp T^{-1}S_{11}\left(\alpha_\perp(\beta'_\perp \alpha_\perp)^{-1}\beta'_\perp + \beta(\alpha'\beta)^{-1}\alpha'\right)\beta_1 c\alpha'_1$$
$$\xrightarrow{D} \left(\int_0^1 KK'du\right)(\beta'_\perp \alpha_\perp)^{-1}\beta'_\perp \beta_1 c\alpha'_1,$$

and from $\alpha'_\perp S_{1\varepsilon} \xrightarrow{D} \int_0^1 K(dW_\varepsilon)$, it is seen that

$$T\breve{\theta} \xrightarrow{D} \left(\int_0^1 KK'du\right)^{-1}\left(\int_0^1 K(dW_\varepsilon) + \left(\int_0^1 KK'du\right)(\beta'_\perp \alpha_\perp)^{-1}\beta'_\perp \beta_1 c\alpha'_1\right)\alpha_\Omega = U,$$

say. Conditional on $K$, the distribution of $U$ is Gaussian with variance $(\alpha'\Omega^{-1}\alpha)^{-1} \otimes (\int_0^1 KK'du)^{-1}$ and mean $(\beta'_\perp \alpha_\perp)^{-1}\beta'_\perp \beta_1 c\alpha'_1\alpha_\Omega$. The information about $\theta$ satisfies

$$T^{-2}I_{\theta\theta} = tr\left\{\Omega^{-1}\alpha(d\theta)'\alpha'_\perp S_{11}\alpha_\perp(d\theta)\alpha'\right\} \xrightarrow{D} tr\left\{\alpha'\Omega^{-1}\alpha(d\theta)'\int_0^1 KK'du(d\theta)\right\},$$

and inserting $U$ for $(d\theta)$ determines the asymptotic distribution of $Q_\beta$. Conditional on $K$, this has a noncentral $\chi^2((p-r)r)$ distribution with noncentrality parameter

$$B = tr\left\{(\beta'_\perp \alpha_\perp)^{-1}\beta'_\perp \beta_1 c'\zeta c\beta'_1\beta_\perp(\alpha'_\perp \beta_\perp)^{-1}\int_0^1 KK'du\right\},$$

where $\zeta = \alpha'_1\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}\alpha_1$, which proves (8). The marginal distribution is therefore a noncentral $\chi^2$ distribution with a stochastic noncentrality parameter, which is independent of the $\chi^2$ distribution, as shown by Elliott (1998).

*Proof of (9):* For $\tau = \alpha_1 c\beta'_1$ it is seen that

$$Etr\left\{(\beta'_\perp \alpha_\perp)^{-1}\beta'_\perp \beta_1 c'\zeta c\beta'_1\beta_\perp(\alpha'_\perp \beta_\perp)^{-1}\int_0^1 KK'du\right\}$$
$$= Etr\left\{\beta_1 c'\zeta c\beta'_1 C\int_0^1 \left(\int_0^u \exp\left(\tau C(u-s)\right)dW(s)\right)\left(\int_0^u dW(t)'\exp\left(C'\tau'(u-t)\right)\right)duC'\right\}$$
$$= tr\left\{\beta_1 c'\zeta c\beta'_1 C\int_0^1 \left(\int_0^u \exp\left(\tau C(u-s)\right)\Omega\exp(C'\tau'(u-s))ds\right)duC'\right\}$$
$$= tr\left\{\beta_1 c'\zeta c\beta'_1 C\left(\int_0^1 (1-v)\exp(v\tau C)\Omega\exp(vC'\tau')dv\right)C'\right\},$$

which proves (9). Note that this expression is zero if and only if $\zeta = 0$, or $\alpha_1' \Omega^{-1} \alpha = 0$, in which case the asymptotic distribution of $Q_\beta$ is $\chi^2$.

*Proof of (10) and (11):*

It follows that $\hat{\Pi} = S_{01} S_{11}^{-1}$ can be expressed as

$$
\begin{aligned}
\hat{\Pi} &= \alpha \beta' + T^{-1} \alpha_1 c \beta_1' + S_{\varepsilon 1} S_{11}^{-1} \\
&= \alpha \beta' + T^{-1} \alpha_1 c \beta_1' + T^{-1/2} (T^{1/2} S_{\varepsilon 1} C_T)(C_T' S_{11} C_T)^{-1} \left( \beta, T^{-1/2} \alpha_\perp \right)' \\
&= \alpha \beta' + T^{-1} \alpha_1 c \beta_1' + T^{-1/2} M_{1T} \beta' + T^{-1} M_{2T} \alpha_\perp',
\end{aligned}
\tag{A7}
$$

where, using (A5) and (A6),

$$
M_{1T} \overset{D}{\to} M_1 = N_{p \times r}(0, \Sigma_{\beta\beta}^{-1} \otimes \Omega), \tag{A8}
$$

$$
M_{2T} \overset{D}{\to} M_2 = \int_0^1 dW_\varepsilon K' \left( \int_0^1 KK' du \right)^{-1}. \tag{A9}
$$

From $\hat{\beta}' = (\alpha_{1\perp}' \hat{\Pi} b)^{-1} \alpha_{1\perp}' \hat{\Pi}$, it follows that

$$
\begin{aligned}
T(\hat{\beta} - \beta)' \beta_\perp &= T(\alpha_{1\perp}' \hat{\Pi} b)^{-1} \alpha_{1\perp}' (\hat{\Pi} - \alpha \beta') \beta_\perp \\
&= (\alpha_{1\perp}' \hat{\Pi} b)^{-1} \alpha_{1\perp}' (T^{1/2} M_{1T} \beta' + M_{2T} \alpha_\perp') \beta_\perp \\
&\overset{D}{\to} (\alpha_{1\perp}' \alpha \beta' b)^{-1} \alpha_{1\perp}' M_2 \alpha_\perp' \beta_\perp = (\alpha_{1\perp}' \alpha)^{-1} \alpha_{1\perp}' M_2 \alpha_\perp' \beta_\perp,
\end{aligned}
$$

where $T^{1/2} M_{1T} \beta' \beta_\perp = 0$, $\alpha_{1\perp}' \alpha_1 c \beta_1' = 0$ and $\beta' b = I_r$. This proves (11).

From the normalization $\hat{\beta}' b = I_r$ we find, replacing $\hat{\beta}$ by $\beta$

$$
\begin{aligned}
T^{1/2}(\hat{\alpha} - \alpha) &= T^{1/2}(\hat{\Pi} \beta_{1\perp} (\hat{\beta}' \beta_{1\perp})^{-1} - \Pi \beta_{1\perp} (\beta' \beta_{1\perp})^{-1}) \\
&= T^{1/2}(T^{-1/2} M_{1T} + T^{-1} M_{2T} \alpha_\perp' \beta_{1\perp} (\beta' \beta_{1\perp})^{-1}) + o_P(1) \\
&= M_{1T} + T^{-1/2} M_{2T} \alpha_\perp' \beta_{1\perp} (\beta' \beta_{1\perp})^{-1} + o_P(1) \overset{D}{\to} M_1,
\end{aligned}
$$

which proves (10).

*Proof of (12):* To analyse the limit distribution of $\hat{c}$, define

$$
A_T = (T^{-1/2} \bar{\alpha}, \alpha_\perp) \text{ and } B_T = (T^{-1/2} \bar{\beta}, \beta_\perp),
$$

and write

$$
\hat{c} = T(\beta_1' \hat{\Pi}^{-1} \alpha_1)^{-1} = (\beta_1' B_T (A_T' T \hat{\Pi} B_T)^{-1} A_T' \alpha_1)^{-1}.
$$

The expansion (A7), and the limits (A8) and (A9) are then applied to give the limit results

$$
\begin{aligned}
T^{-1/2} \bar{\alpha}' (T \hat{\Pi}) T^{-1/2} \bar{\beta} &= I_r + O(T^{-1}) + O_P(T^{-1}), \\
T^{-1/2} \bar{\alpha}' (T \hat{\Pi}) \beta_\perp &= 0 + O(T^{-1/2}) + O_P(T^{-1/2}), \\
\alpha_\perp' (T \hat{\Pi}) \bar{\beta} T^{-1/2} &= 0 + O(T^{-1/2}) + \alpha_\perp' M_{1T}, \\
\alpha_\perp' (T \hat{\Pi}) \beta_\perp &= 0 + \alpha_\perp' \alpha_1 c \beta_1' \beta_\perp + \alpha_\perp' M_{2T} \alpha_\perp' \beta_\perp.
\end{aligned}
$$

Thus

$$A_T'(T\hat{\Pi})B_T \xrightarrow{D} \begin{pmatrix} I_r & 0 \\ \alpha_\perp' M_1 & \alpha_\perp' \alpha_1 c\beta_1'\beta_\perp + \alpha_\perp' M_2\alpha_\perp'\beta_\perp \end{pmatrix},$$

$$B_T(A_T'(T\hat{\Pi})B_T)^{-1}A_T' \xrightarrow{D} (0,\beta_\perp) \begin{pmatrix} I_r & 0 \\ \alpha_\perp' M_1 & \alpha_\perp' \alpha_1 c\beta_1'\beta_\perp + \alpha_\perp' M_2\alpha_\perp'\beta_\perp \end{pmatrix}^{-1} (0,\alpha_\perp)'$$

$$= \beta_\perp(\alpha_\perp'\alpha_1 c\beta_1'\beta_\perp + \alpha_\perp' M_2\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'.$$

Multiplying by $\beta_1'$ and $\alpha_1$ and inverting, it is seen that because $\beta_1'\beta_\perp$ and $\alpha_1'\alpha_\perp$ are $(p-r) \times (p-r)$ of full rank,

$$\hat{c} = (\beta_1' B_T(A_T' T\hat{\Pi}B_T)^{-1}A_T'\alpha_1)^{-1} \xrightarrow{D} [\beta_1'\beta_\perp(\alpha_\perp'\alpha_1 c\beta_1'\beta_\perp + \alpha_\perp' M_2\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'\alpha_1]^{-1}$$

$$= (\alpha_\perp'\alpha_1)^{-1}(\alpha_\perp'\alpha_1 c\beta_1'\beta_\perp + \alpha_\perp' M_2\alpha_\perp'\beta_\perp)(\beta_1'\beta_\perp)^{-1}$$

$$= c + (\alpha_\perp'\alpha_1)^{-1}\alpha_\perp' M_2\alpha_\perp'\beta_\perp(\beta_1'\beta_\perp)^{-1},$$

which proves (12). □

**Proof of Corollary 1.** *Proof of (13):* If $r = p - 1$, the expression (9) can be reduced as follows. For $\tau = \alpha_1 c\beta_1'$

$$(\tau C)^2 = c\alpha_1 c(\beta_1' C\alpha_1)\beta_1' C = c(\beta_1' C\alpha_1)\alpha_1 c\beta_1' C = c\delta\tau C,$$

for $\delta = \beta_1' C\alpha_1$, and in general for $n \geq 0$, it is seen that

$$(\tau C)^{n+1} = (c\delta)^n \tau C.$$

Therefore, using $\beta_1 c'\zeta c\beta_1' = \beta_1 c'\tau\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}\tau'$,

$$E(B) = tr\{\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}\tau C \left( \int_0^1 (1-v)\exp(v\tau C)\Omega\exp(vC'\tau')dv \right) C'\tau'\}.$$

The integral can be calculated by the expansion

$$\tau C\exp(v\tau C)\Omega\exp(vC'\tau')C'\tau' = \sum_{n,m=0}^\infty \frac{v^n}{n!}(\tau C)^{n+1}\Omega(C'\tau')^{m+1}\frac{v^m}{m!}$$

$$= \sum_{n,m=0}^\infty \frac{(vc\delta)^{n+m}}{n!m!}\tau C\Omega C'\tau' = \exp(2vc\delta)c^2\kappa\alpha_1\alpha_1',$$

where $\kappa = \beta_1' C\Omega C'\beta_1$. This allows the integral to be calculated

$$\tau C \left( \int_0^1 (1-v)\exp(v\tau C)\Omega\exp(vC'\tau')dv \right) C'\tau'$$

$$= \left( \int_0^1 (1-v)\exp(2vc\delta)dv \right) c^2\kappa\alpha_1\alpha_1' = \frac{e^{2\delta c} - 1 - 2c\delta}{(2\delta c)^2}c^2\kappa\alpha_1\alpha_1'.$$

Therefore

$$E(B) = \frac{e^{2c\delta} - 1 - 2c\delta}{(2\delta)^2}\kappa\alpha_1'\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}\alpha_1 = (e^{2c\delta} - 1 - 2c\delta)\frac{\kappa\zeta}{(2\delta)^2},$$

where $\zeta = \alpha_1'\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}\alpha_1$. □

## Appendix B. Figures



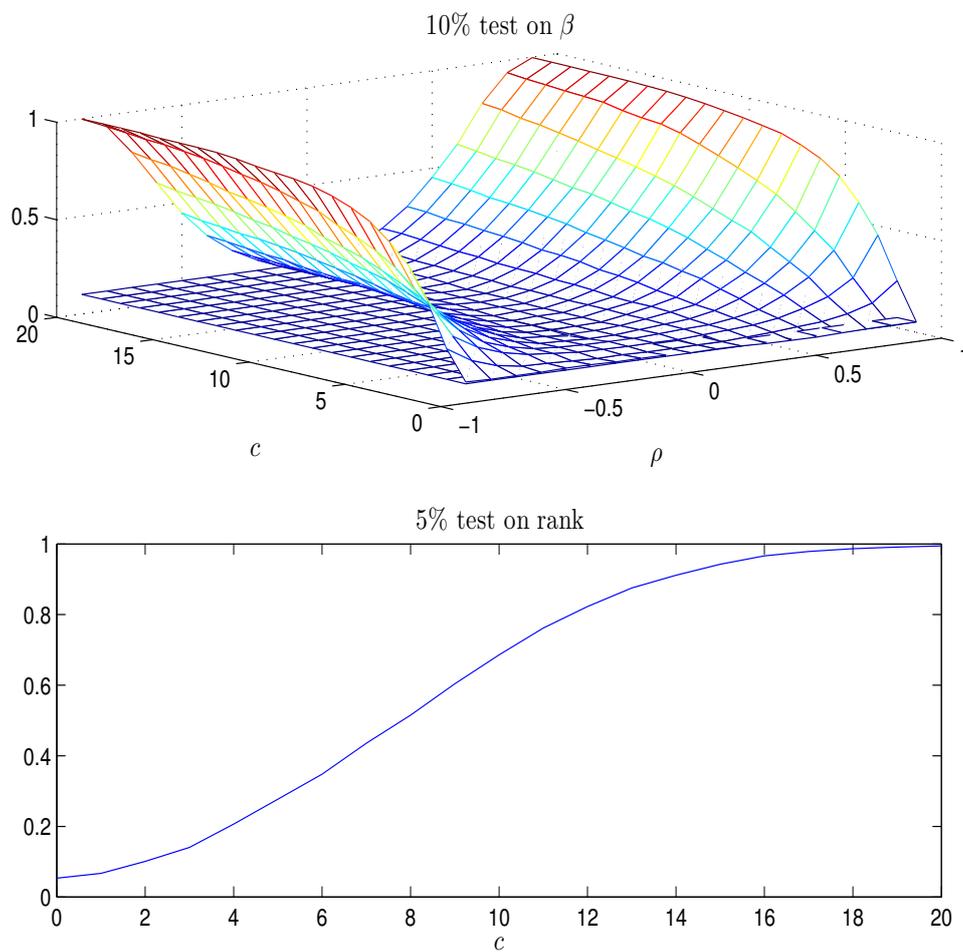**Figure A1.** Top panel: Rejection frequency of the test $Q_\beta$ for $\gamma = 0$ using the $\chi^2(1)_{0.90}$ quantile as a function of $c$ and $\rho$. Bottom panel: Rejection frequency of the 5% test $Q_r$ for $r = 1$ using Table 15.1 in Johansen (1996) as a function of $c$. $N = 1000$ simulations of $T = 100$ observations from the DGP (15) and (16).
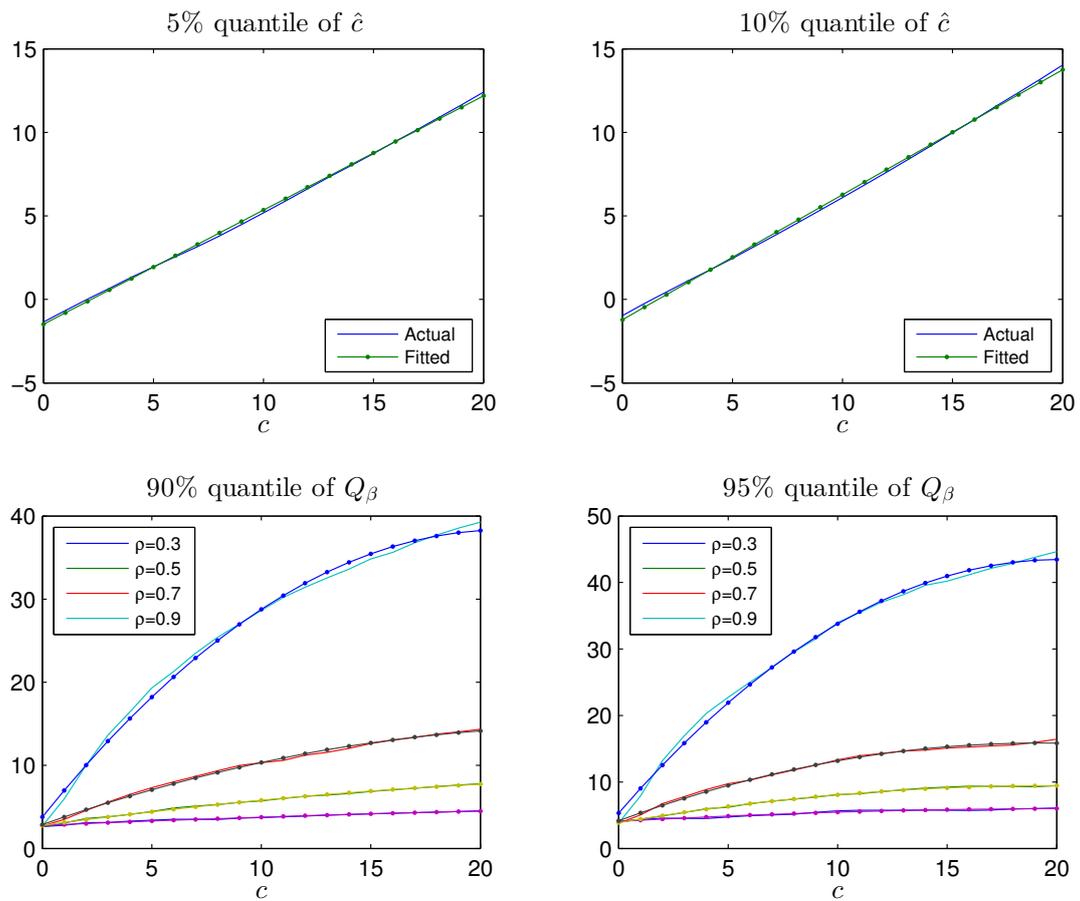
**Figure A2.** Quantiles and fitted values in the distributions of $\hat{c}$ and $Q_\beta$ as a function of $c$ for different values of $\rho$; $N = 1000$ simulations of $T = 100$ observations from the DGP (15) and (16).
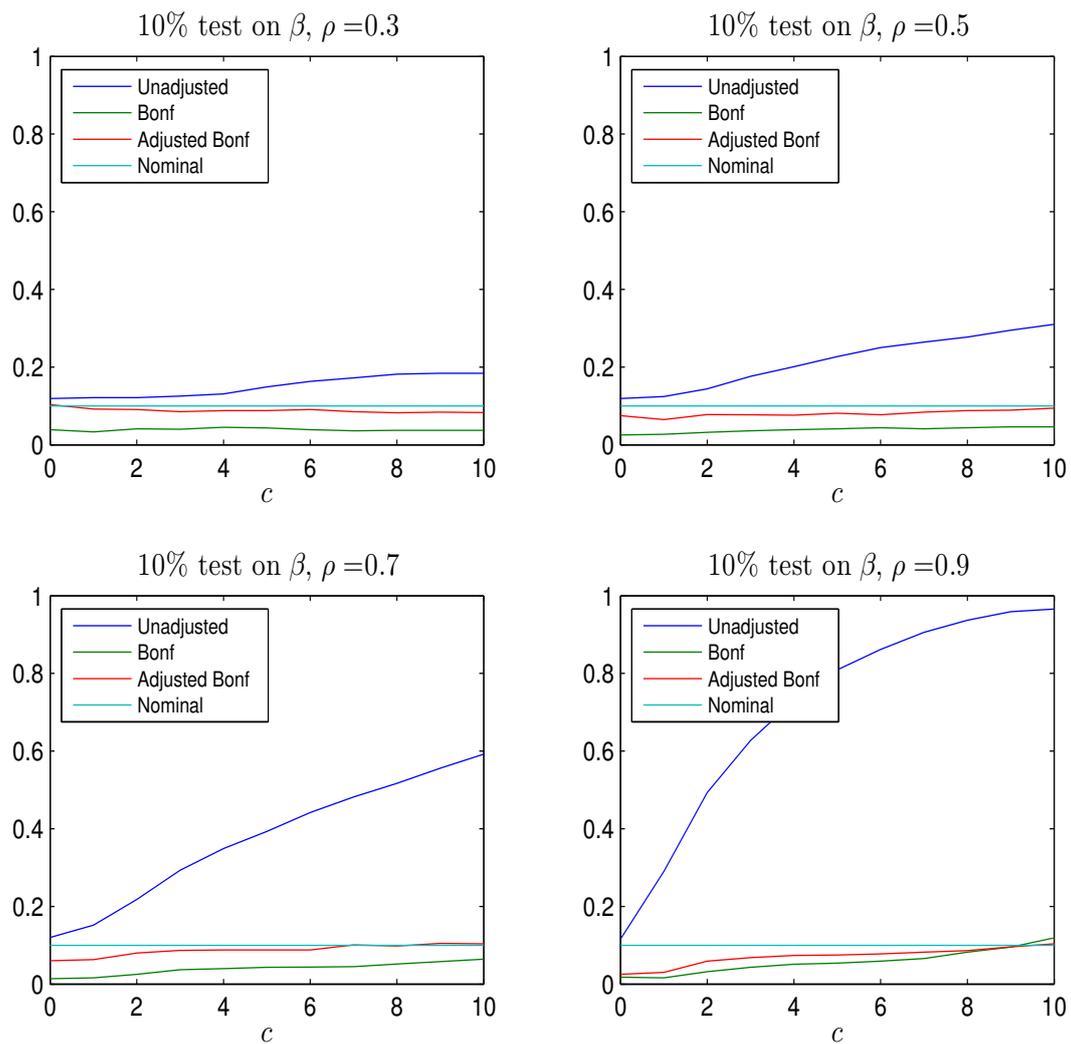
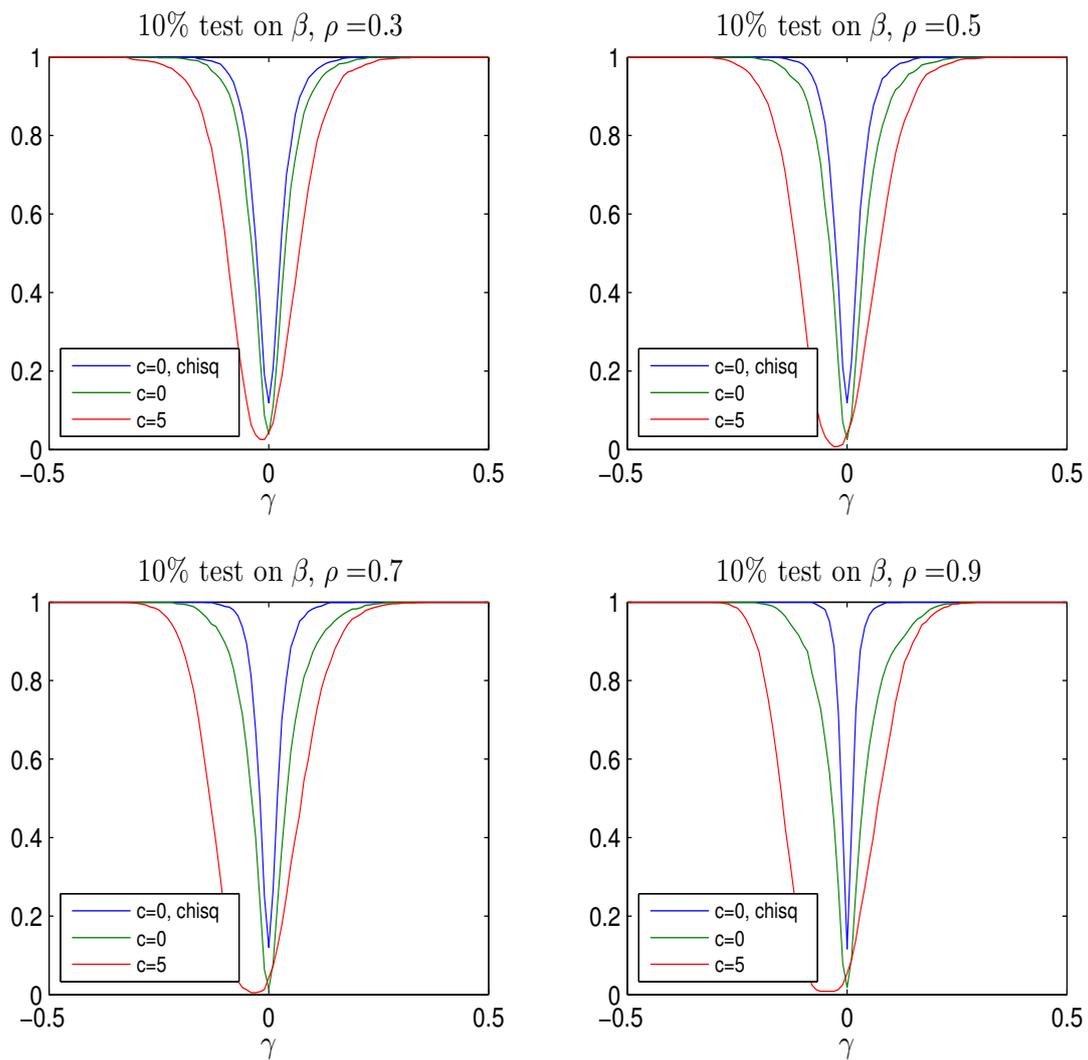**Figure A3.** Rejection frequency of the test $Q_\beta$ for $\gamma = 0$ using the $\chi^2(1)_{0.90}$ quantile (Unadjusted), the Bonferroni quantile in (20) for $\xi = 95\%$ and $\eta = 5\%$ (Bonf) and the adjusted Bonferroni quantile in (21) for $\eta = 5\%$ (Adjusted Bonf) as a function of $c$ for different values of $\rho$; $N = 1000$ simulations of $T = 100$ observations from the DGP (15) and (16).

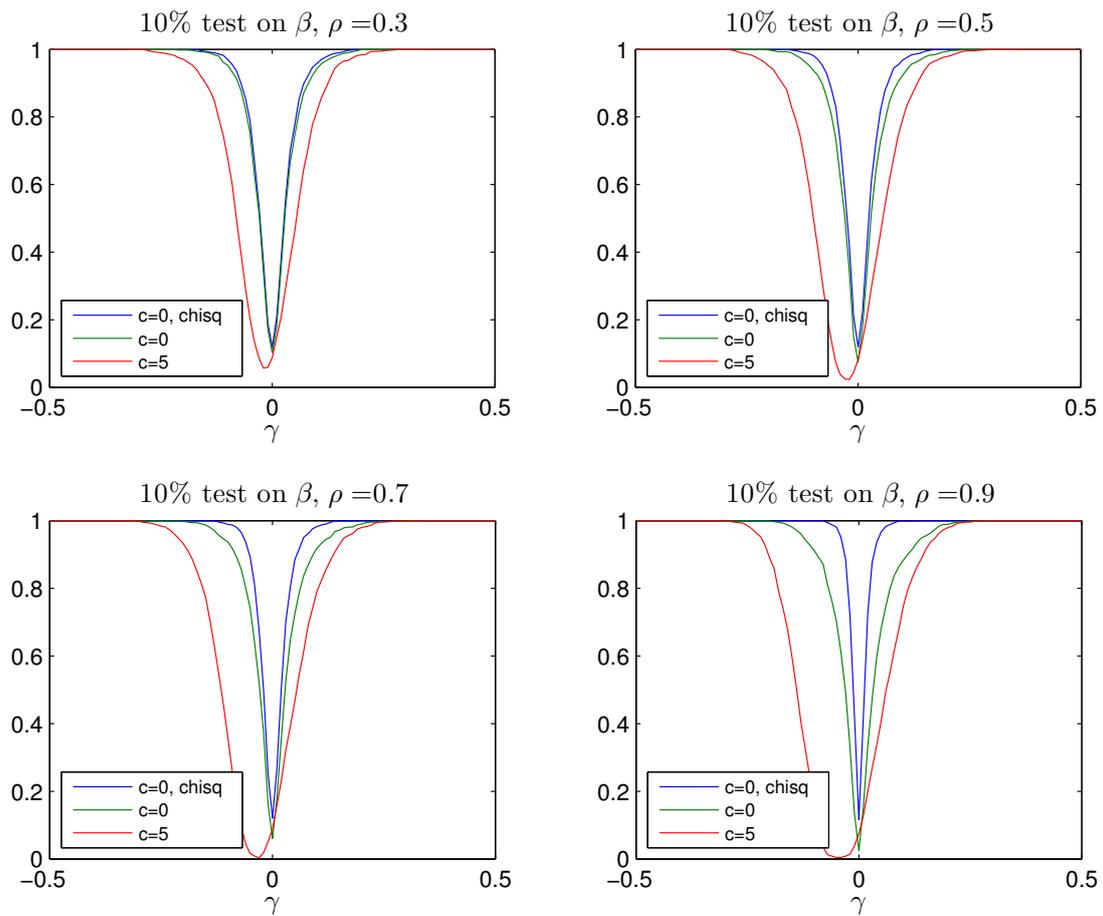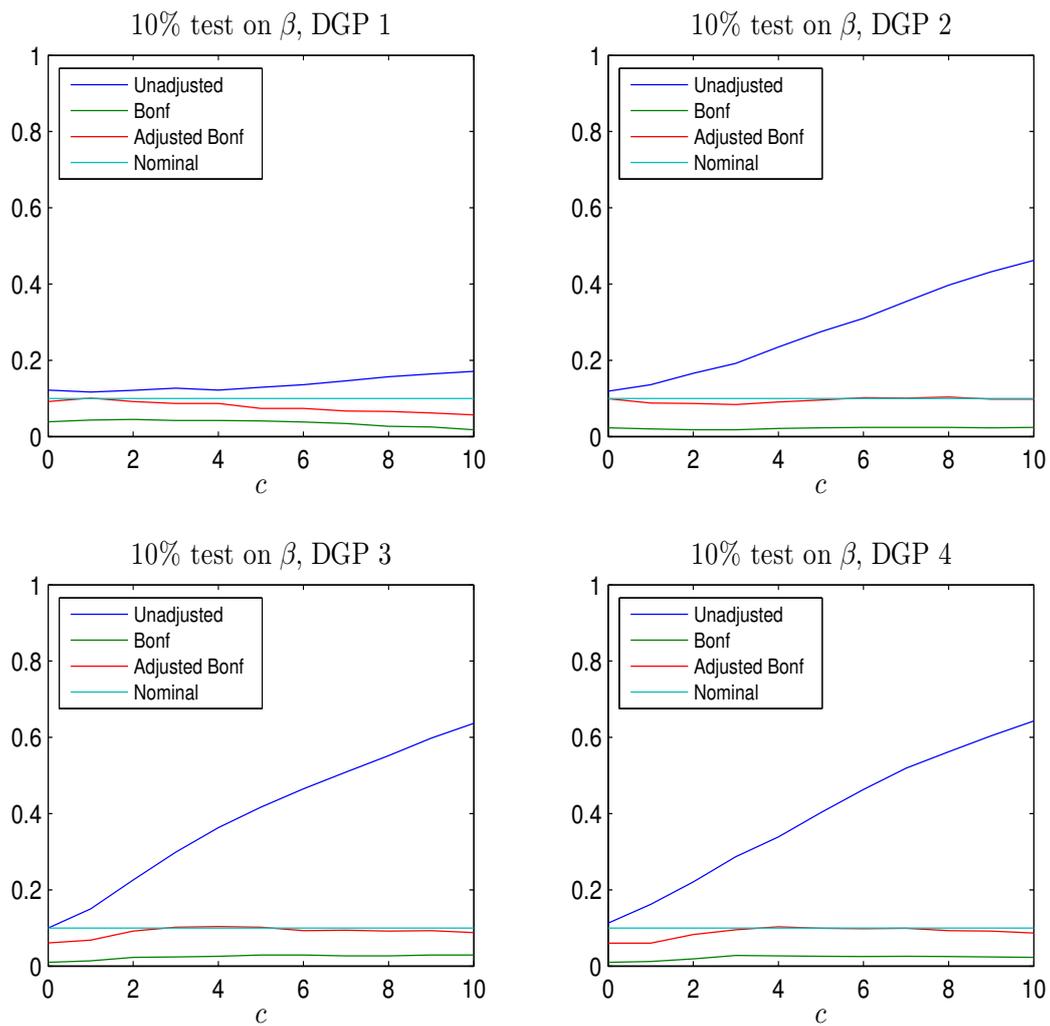**Figure A4.** Rejection frequency of the test $Q_\beta$ for $\gamma = 0$ using the $\chi^2(1)_{0.90}$ quantile (c = 0, chisq) and the Bonferroni quantile in (20) for $\xi = 95\%$ and $\eta = 5\%$, as a function of $\gamma$ for different values of $c$ and $\rho$; $N = 1000$ simulations of $T = 100$ observations from the DGP (15) and (16).

**Figure A5.** Rejection frequency of the test $Q_\beta$ for $\gamma = 0$ using the $\chi^2(1)_{0.90}$ quantile (c=0, chisq) and the adjusted Bonferroni quantile in (21) for $\eta = 5\%$ as a function of $\gamma$ for different values of $c$ and $\rho$; $N = 1000$ simulations of $T = 100$ observations from the DGP (15) and (16).

**Figure A6.** Rejection frequency of the test $Q_\beta$ for $\gamma = 0$ using the $\chi^2(1)_{0.90}$ quantile (Unadjusted), the Bonferroni quantile in (20) for $\xi = 95\%$ and $\eta = 5\%$ (Bonf) and the adjusted Bonferroni quantile in (21) for $\eta = 5\%$ (Adjusted Bonf) as a function of $c$; $N = 1000$ simulations of $T = 100$ observations from the DGPs in Table 1.
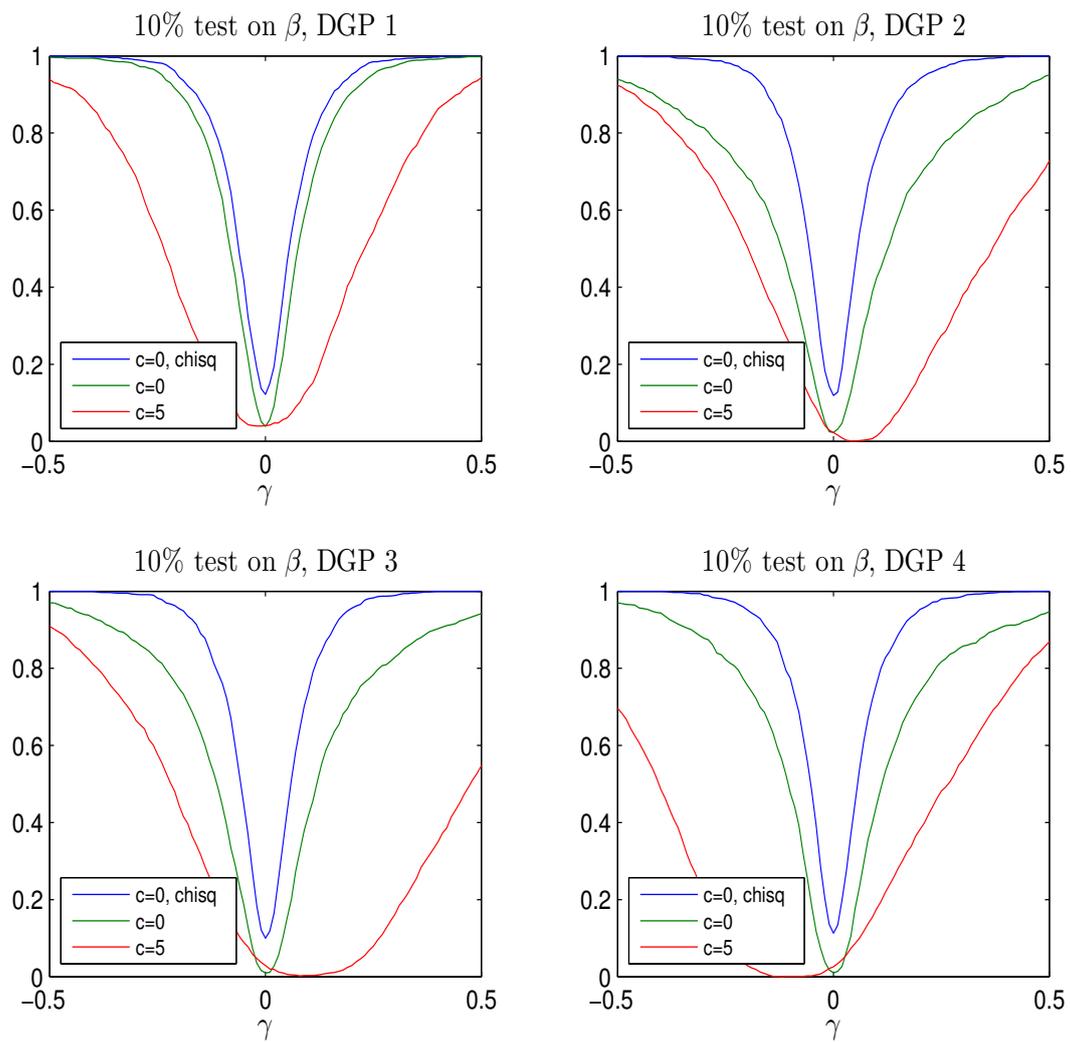
**Figure A7.** Rejection frequency of the test $Q_\beta$ for $\gamma = 0$ using the $\chi^2(1)_{0.90}$ quantile (c = 0, chisq) and the Bonferroni quantile in (20) for $\xi = 95\%$ and $\eta = 5\%$, as a function of $\gamma$ for different values of $c$; $N = 1000$ simulations of $T = 100$ observations from the DGPs in Table 1.
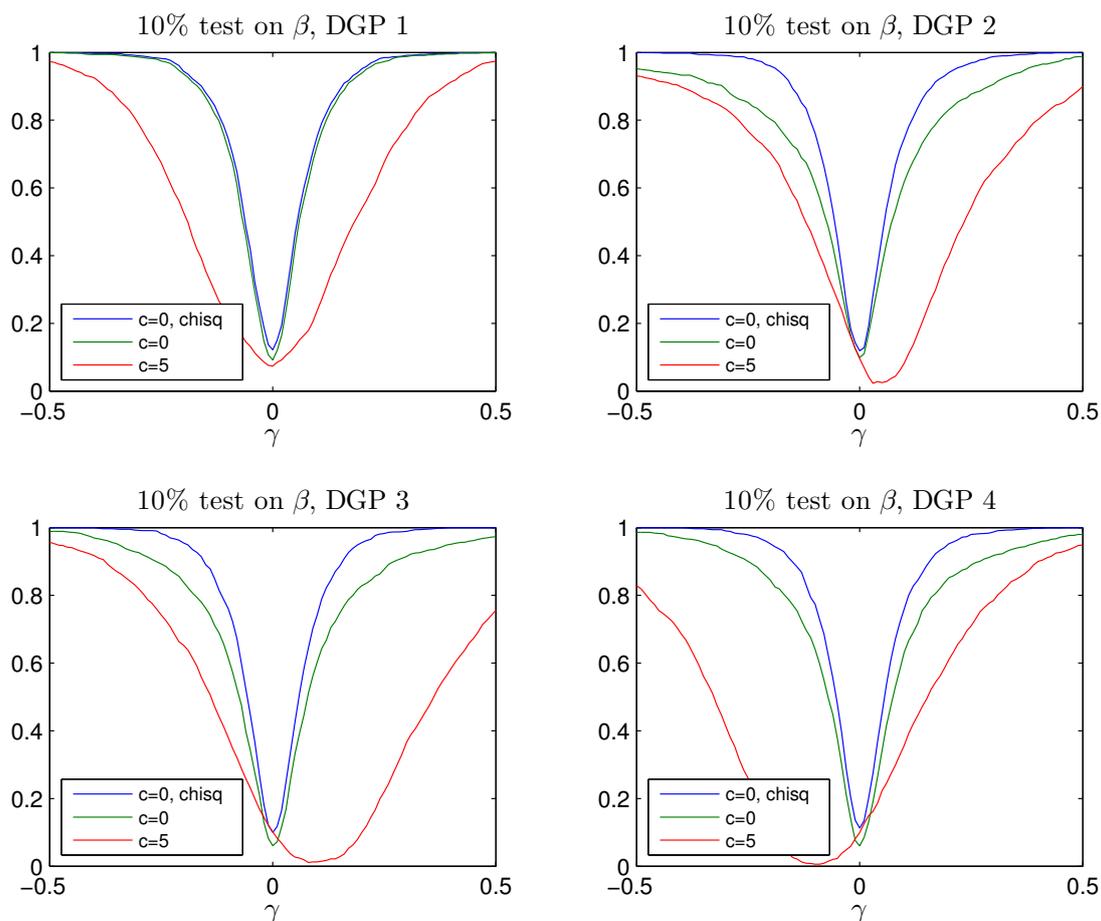
**Figure A8.** Rejection frequency of the test $Q_\beta$ for $\gamma = 0$ using the $\chi^2(1)_{0.90}$ quantile (c = 0, chisq) and the adjusted Bonferroni quantile in (21) for $\eta = 5\%$ as a function of $\gamma$ for different values of $c$; $N = 1000$ simulations of $T = 100$ observations from the DGPs in Table 1.

## References

Cavanagh, Christopher L., Graham Elliott, and James H. Stock. 1995. Inference in Models with Nearly Integrated Regressors. *Econometric Theory* 11: 1131–47.

Di Iorio, Francesca, Stefano Fachin, and Riccardo Lucchetti. 2016. Can you do the wrong thing and still be right? Hypothesis testing in I(2) and near-I(2) cointegrated VARs. *Applied Economics* 48: 3665–78. doi:10.1080/00036846.2016.1142660.

Elliott, Graham. 1998. On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica* 66: 149–58.

Johansen, Søren. 1996. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

McCloskey, Adam. 2017. Bonferroni-based size-correction for nonstandard testing problems. *Journal of Econometrics* in press. Available online: http://www.sciencedirect.com/science/article/pii/S0304407617300556 (accessed on 12 June 2017).

Phillips, Peter C. B. 1988. Regression theory for near integrated time series. *Econometrica* 56: 1021–44.