BACKTESTING PARAMETRIC VALUE-AT-RISK WITH ESTIMATION RISK

Jose Olmo

Department of Economics

City University, London

(joint work with Juan Carlos Escanciano, Indiana University)

Outline

- Motivation of the econometric problem and definitions

- The effect of Estimation Risk in unconditional backtesting

- The effect of Estimation Risk in conditional backtesting

- A Monte-Carlo simulation exercise for size and power of uncorrected and corrected tests.

- Empirical application to $S\&P500$.

- Conclusions

Motivation:

- In the aftermath of a series of bank failures during the seventies a group of ten countries (G-10) decided to create a committee to set up a regulatory framework intended to prevent financial institutions, in particular banks, from operating without effective supervision.

- The subsequent documents derived from this commitment focused on the imposition of capital requirements for internationally active banks intending to act as provisions for losses from adverse market fluctuations, concentration of risks or simply bad management of institutions.

- The risk measure agreed to determine the amount of capital on hold was the Value-at-Risk (VaR).

- Large financial institutions gained the possibility of computing their own risk measures and hence to err on the side of using models infra-estimating risk.

- To monitor these models the Basel Accord (1996a), and the Ammendment of Basel Accord (1996b) developed a backtesting procedure to assess the accuracy and quality of different risk measurement techniques.

- This process is used both by banks' internal units to measure the accuracy of their models and by external regulators that on the basis of backtesting performance set appropriate punishments reflected on additional capital requirements applicable in case of failure.

- The essence of backtesting is the **out-of-sample** comparison of actual trading results with model-generated risk measures. If the comparison uncovers sufficient differences between both figures either the risk model or the assumptions on the backtesting technique should be subject to revision by the corresponding regulatory body.

What is Value-at-Risk?

Definition.- The α-th conditional VaR of Y_t given W_{t-1} is the measurable function $q_\alpha(W_{t-1})$ satisfying the equation

$$P(Y_t \leq q_\alpha(W_{t-1}) \mid W_{t-1}) = \alpha, \text{ almost surely (a.s.)}, \alpha \in (0,1), \ \forall t \in \mathbb{Z}. \qquad (1)$$

Furthermore, in *parametric* VaR inference one assumes the existence of

$$\mathcal{M} = \{m_\alpha(\cdot, \theta) : \theta \in \Theta \subset \mathbb{R}^p\}$$

and proceeds to make VaR forecasts using the model $\mathcal{M}$.

Inferences within the model depend on the hypothesis that $q_\alpha \in \mathcal{M}$, i.e., if there exists some $\theta_0 \in \Theta$ such that

$$m_\alpha(W_{t-1}, \theta_0) = q_\alpha(W_{t-1}).$$

Different methodologies for calculating VaR

We use $m_\alpha(W_{t-1}, \theta_0)$ to denote the conditional version of $VaR_p(y_{t+1})$. Then,

- **Historical Simulation:**

$$m_\alpha(W_{t-1}, \theta_0) = \theta_0 \equiv F_Y^{-1}(\alpha),$$

where $F_Y^{-1}(\alpha)$ denotes the unconditional quantile function of Y_t evaluated at α.

- **Location-scale models:** The returns process can be expressed as follows

$$Y_t = \mu(W_{t-1}, \beta_0) + \sigma(W_{t-1}, \beta_0)\varepsilon_{t+1},$$

then,

$$m_\alpha(W_{t-1}, \theta_0) = \mu(W_{t-1}, \beta_0) + \sigma(W_{t-1}, \beta_0)F_\varepsilon^{-1}(\alpha)$$

with $\mu(W_{t-1}, \beta_0)$ the conditional mean process, $\sigma(W_{t-1}, \beta_0)$ the conditional volatility process and $F_\varepsilon^{-1}(\alpha)$ the unconditional quantile of the error distribution.

- **Regression quantile models:** Koenker and Basset (1978).

The returns process can be expressed as follows

$$Y_t = q_\alpha(W_{t-1}) + \varepsilon_t,$$

with ε_t iid, independent of W_{t-1}, and satisfying $q_\alpha(\varepsilon_t \mid W_{t-1}) = 0$.

A parametric VaR model will be well specified if

$$m_\alpha(W_{t-1}, \theta_0) = q_\alpha(W_{t-1})$$

- **Other methods** as Gourieroux and Jasiak (2006), CaViAR of Engle and Manganelli (2004), or methods derived from Extreme Value Theory, see McNeil and Frey (2001).

In our simulation and empirical application sections we will concentrate on the second class of models (location-scale) given their simplicity and wide applicability in financial econometrics and risk management.

What is Backtesting?

It is the comparison of actual trading results with model-generated risk measures when measuring the number of exceedances (failures) of the VaR risk measure.

In other words, test the condition

$$P(Y_t \leq m_\alpha(W_{t-1}, \theta_0) \mid W_{t-1}) = \alpha, \text{ almost surely (a.s.)}, \ \alpha \in (0,1), \ \forall t \in \mathbb{Z}$$

The corresponding hypothesis test is the following:

$$E[I_{t,\alpha}(\theta_0) \mid W_{t-1}(\theta_0)] = \alpha, \text{ a.s. for some } \theta_0 \in \Theta,$$

where

$$I_{t,\alpha}(\theta_0) = I(Y_t \leq m_\alpha(I_{t-1}, \theta_0)),$$

with $I(\cdot)$ the indicator function.

Existing backtesting techniques:

These first tests assume independence among observations. This implies that the condition to be tested is

$$H_{0u} : E[I_{t,\alpha}(\theta_0)] = \alpha.$$

- Basel Committee in Banking Supervision (Basel Accord (1996)) proposes a one-sided test assuming that $\sum_{t=R+1}^{n} I_{t,\alpha}(\theta_0)$ follows a binomial $bin(P, p)$. Then

$$H_0 : p \leq \alpha \text{ vs } H_1 : p > \alpha,$$

with R the in-sample size and $P = n - R$ out-of-sample size. This test is overconservative.

- Kupiec (1995) proposed the following test statistic:

$$\frac{1}{\sqrt{P}} \sum_{t=R+1}^{P} (I_{t,\alpha}(\theta_0) - \alpha)$$

that is asymptotically normal $N(0, \alpha(1-\alpha))$ under H_{0u}.

- Christoffersen (1998) proposed a Likelihood ratio (LR) test for the unconditional coverage:

$$LR_u = -2log\frac{L(\alpha;\{I_{t,\alpha}(\theta_0)\}_{t=R+1}^{n})}{L(\hat{\pi};\{I_{t,\alpha}(\theta_0)\}_{t=R+1}^{n})},$$

with

$$L(\hat{\pi};\{I_{t,\alpha}(\theta_0)\}_{t=R+1}^{n}) = (1-\hat{\pi})^{n_0}\hat{\pi}^{n_1},$$

with

$$\hat{\pi} = n_1/(n_0+n_1),$$

and n_1 denoting the number of VaR exceedances and $n_0 = n - R - n_1$.

* Other methods:

- Christoffersen, Hahn and Inoue (2001), that use nonnested VaR hypothesis testing and information criteria.

- Christoffersen and Gon çalves (2005), that propose bootstrap confidence intervals for VaR.

- Christoffersen and Pelletier (2006) based on duration models.

Motivation for this work

- All the former methods assume θ_0 known.

- In practice however this is not the case, and the parameter is replaced by $\sqrt{R}$-consistent estimators, θ_R, of it.

- Furthermore, the existing backtesting procedures are all based on testing

$$E[I_{t,\alpha}(\theta_0) \mid \widetilde{I}_{t-1}(\theta_0)] = \alpha, \text{ a.s. for some } \theta_0 \in \Theta, \tag{2}$$

  where $\widetilde{I}_{t-1}(\theta_0) = (I_{t-1,\alpha}(\theta_0), I_{t-2,\alpha}(\theta_0)...)'$.

- Note that this equation is a necessary but not sufficient condition to test the correct specification of the VaR model.

- This has important consequences in terms of the power performance of the backtesting procedures.

Backtesting procedures free of estimation risk

The unconditional composite hypothesis: Consider

$$H_{0u} : E[I_{t,\alpha}(\theta_0)] = \alpha \qquad \text{for some } \theta_0 \in \Theta.$$

A natural two-sided test for H_{0u} is based on rejecting for large values of $|S_P|$, where

$$S_P = \frac{1}{\sqrt{P}} \sum_{t=R+1}^{n} (I_{t,\alpha}(\widehat{\theta}_t) - \alpha),$$

with P the out-of-sample size, and $\widehat{\theta}_t$ a consistent estimator of θ_0.

We quantify in the next slides the effect of the estimation risk in S_P.

Assumptions:

Assumption A1: $\{Y_t, Z_t'\}_{t\in\mathbb{Z}}$ is strictly stationary and ergodic.

Assumption A2: The family of distributions functions $\{F_x, x \in \mathbb{R}^\infty\}$ has Lebesgue densities $\{f_x, x \in \mathbb{R}^\infty\}$ that are uniformly bounded

$$\sup_{x\in\mathbb{R}^\infty, y\in\mathbb{R}} |f_x(y)| \leq C$$

and equicontinuous: for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\sup_{x\in\mathbb{R}^\infty, |y-z|\leq\delta} |f_x(y) - f_x(z)| \leq \epsilon.$$

Assumption A3: The model $m_\alpha(W_{t-1}, \theta)$ is continuously differentiable in θ (a.s.) with derivative $g_\alpha\left(W_{t-1}, \theta\right)$ such that $E\left[\sup_{\theta\in\Theta_0} |g_\alpha(W_{t-1}, \theta)|^2\right] < C$, for a neighborhood Θ_0 of θ_0.

Assumption A4: The parameter space Θ is compact in $\mathbb{R}^p$.

The true parameter θ_0 belongs to the interior of Θ.

The estimator $\widehat{\theta}_t$ satisfies the asymptotic expansion

$\widehat{\theta}_t - \theta_0 = H(t) + o_P(1)$,

where

$H(t) = t^{-1} \sum_{s=1}^{t} l(Y_s, W_{s-1}, \theta_0)$, for the recursive scheme,

$R^{-1} \sum_{s=t-R+1}^{t} l(Y_s, W_{s-1}, \theta_0)$ for the rolling scheme, and

$R^{-1} \sum_{s=1}^{R} l(Y_s, W_{s-1}, \theta_0)$ for the fixed scheme.

We also assume standard regularity conditions on $l(Y_t, W_{t-1}, \theta_0)$.

Assumption A5: $R, P \to \infty$ as $n \to \infty$, and $\lim_{n\to\infty} P/R = \pi$, $0 \leq \pi < \infty$.

Theoretical foundation for the paper:

Define the process

$$K_n(c) := \frac{1}{\sqrt{P}} \sum_{t=R+1}^{n} \left[I_{t,\alpha}(\theta_0 + c(t-1)^{-1/2}) - F_{W_{t-1}}(\theta_0 + c(t-1)^{-1/2}) \right].$$

If $\widehat{c}$ is bounded in probability, $\widehat{c} = O_P(1)$, then $|K_n(\widehat{c}) - K_n(0)| = o_P(1)$.

Applying this argument to $\widehat{c} := \max_{R \leq t \leq n} \sqrt{t}(\widehat{\theta}_t - \theta_0)$, for our three forecasting schemes yields

$$\frac{1}{\sqrt{P}} \sum_{t=R+1}^{n} \left[I_{t,\alpha}(\theta_0 + c(t-1)^{-1/2}) - F_{W_{t-1}}(\theta_0 + c(t-1)^{-1/2}) \right] - \frac{1}{\sqrt{P}} \sum_{t=R+1}^{n} \left[I_{t,\alpha}(\theta_0) - F_{W_{t-1}}(\theta_0) \right] = o_p(1).$$

This result allows us to establish the first theorem of the paper:

THEOREM 1: *Under Assumptions A1-A5,*

$$S_P = \frac{1}{\sqrt{P}} \sum_{t=R+1}^{n} \left[I_{t,\alpha}(\theta_0) - F_{W_{t-1}}(m_\alpha(W_{t-1},\theta_0)) \right]$$

$$\underbrace{+ E\left[g'_\alpha(W_{t-1},\theta_0) f_{W_{t-1}}(m_\alpha(W_{t-1},\theta_0)) \right] \frac{1}{\sqrt{P}} \sum_{t=R+1}^{n} H(t-1)}_{\text{Estimation Risk}}$$

$$\underbrace{+ \frac{1}{\sqrt{P}} \sum_{t=R+1}^{n} \left[F_{W_{t-1}}(m_\alpha(W_{t-1},\theta_0)) - \alpha \right]}_{\text{Model Risk}} + o_P(1).$$

Theorem 1 quantifies both *estimation risk* and *model risk* in the unconditional coverage test introduced before.

COROLLARY 1: *Under Assumptions A1-A5 and* $E[I_{t,\alpha}(\theta_0) \mid W_{t-1}(\theta_0)] = \alpha$,

$$S_P \xrightarrow{d} N(0, \sigma_u^2),$$

where

$$\sigma_u^2 = \alpha(1-\alpha) + 2\lambda_{hl} A\rho + \lambda_{ll} AVA',$$

with $\rho = E[(I_{t,\alpha}(\theta_0) - \alpha)\, l(Y_t, W_{t-1}, \theta_0)]$, $A := E\left[g'_\alpha(W_{t-1}, \theta_0) f_{W_{t-1}}(m_\alpha(W_{t-1}, \theta_0))\right]$ *and where*

Scheme	λ_{hl}	λ_{ll}
Recursive	$1 - \pi^{-1}\ln(1+\pi)$	$2\left[1 - \pi^{-1}\ln(1+\pi)\right]$
Rolling, $\pi \leq 1$	$\pi/2$	$\pi - \pi^2/3$
Rolling, $1 < \pi < \infty$	$1 - (2\pi)^{-1}$	$1 - (3\pi)^{-1}$
Fixed	0	π

(3)

Estimating the asymptotic variance of S_P

The vector A can be consistently estimated by

$$\widehat{A}_\tau = -\frac{1}{P}\sum_{t=R+1}^{n}\frac{1}{\tau}\exp\left[\left(Y_t - m_\alpha(W_{t-1},\widehat{\theta}_{t-1})\right)/\tau\right] I_{t,\alpha}(\widehat{\theta}_{t-1}) g'_\alpha(W_{t-1},\widehat{\theta}_{t-1}), \qquad (4)$$

with $\tau \to 0$ as $n \to \infty$, see Giacomini and Komunjer (2005) for encompassing tests of different conditional quantile forecasts.

Natural estimators of $\lambda_{hl} = \lambda_{hl}(\pi)$ and $\lambda_{ll} = \lambda_{ll}(\pi)$ are $\widehat{\lambda}_{hl} = \lambda_{hl}(\widehat{\pi})$ and $\widehat{\lambda}_{ll} = \lambda_{ll}(\widehat{\pi})$, where the parameter π is approximated by $\widehat{\pi} = P/R$.

Hence, the asymptotic variance σ_u^2 can be consistently estimated by

$$\widehat{\sigma}_u^2 := \alpha(1-\alpha) + 2\widehat{\lambda}_{hl}\widehat{A}_\tau\widehat{\rho} + \widehat{\lambda}_{ll}\widehat{A}_\tau\widehat{V}\widehat{A}'_\tau,$$

where

$$\widehat{\rho} = \frac{1}{P} \sum_{t=R+1}^{n} \Big(I_{t,\alpha}(\widehat{\theta}_{t-1}) - \alpha \Big) \, l(Y_t, W_{t-1}, \widehat{\theta}_{t-1}),$$

and

$$\widehat{V} = \frac{1}{P} \sum_{t=R+1}^{n} l(Y_t, W_{t-1}, \widehat{\theta}_{t-1}) l'(Y_t, W_{t-1}, \widehat{\theta}_{t-1}),$$

are consistent estimators for ρ and V, respectively. Then, valid inference can be accomplished by the *corrected unconditional backtesting test* statistic

$$\widetilde{S}_P \equiv \widetilde{S}(P, R, \widehat{\theta}_{t-1}) = \frac{1}{\widehat{\sigma}_u \sqrt{P}} \sum_{t=R+1}^{n} (I_{t,\alpha}(\widehat{\theta}_{t-1}) - \alpha),$$

which converges to a standard normal $r.v$ as shown in the next corollary.

COROLLARY 2: *Under Assumptions A1-A5,* $E[I_{t,\alpha}(\theta_0) \mid W_{t-1}(\theta_0)] = \alpha$, *and that* $\tau \to 0$ *as* $n \to \infty$,

$$\widetilde{S}_P \xrightarrow{d} N(0,1).$$

The independence hypothesis

Aim: Testing the hypothesis

$$\{I_{t,\alpha}(\theta_0)\}_{t=R+1}^{n} \text{ are } iid. \tag{5}$$

Christoffersen (1998) introduces in his seminal paper a likelihood ratio (LR) test where the alternative hypothesis is embedded in a first-order Markov model.

More general tests for (5) have been based on

$$\xi_j = Cov(I_{t,\alpha}(\theta_0), I_{t-j,\alpha}(\theta_0)) \qquad j \geq 1, \tag{6}$$

at different lags j, which can be consistently estimated (under $E[I_{t,\alpha}(\theta_0)] = \alpha$) by

$$\xi_{P,j} = \frac{1}{P-j} \sum_{t=R+j+1}^{n} (I_{t,\alpha}(\theta_0) I_{t-j,\alpha}(\theta_0) - \alpha^2) \quad \text{for} \quad j \geq 1.$$

In practice, however, tests for (5) need to be based on estimates of the relevant parameters, such as

$$\widehat{\xi}_{P,j} = \frac{1}{P-j} \sum_{t=R+j+1}^{n} (I_{t,\alpha}(\widehat{\theta}_{t-1}) I_{t-j,\alpha}(\widehat{\theta}_{t-j-1}) - \alpha^2)$$

and

$$\widehat{\gamma}_{P,j} = \frac{1}{P-j} \sum_{t=R+j+1}^{n} (I_{t,\alpha}(\widehat{\theta}_{t-1}) - \alpha)(I_{t-j,\alpha}(\widehat{\theta}_{t-j-1}) - \alpha).$$

THEOREM 2: *Under Assumptions A1-A5 and* $E[I_{t,\alpha}(\theta_0) \mid W_{t-1}(\theta_0)] = \alpha$, *for any* $j \geq 1$,

(i) $\sqrt{P-j}(\widehat{\xi}_{P,j} - \xi_{P,j}) = B\frac{1}{\sqrt{P-j}}\sum_{t=R+j+1}^{n} H(t-j-1) + o_P(1)$.

(ii) $\sqrt{P-j}(\widehat{\gamma}_{P,j} - \gamma_{P,j}) = \{B - 2\alpha A\}\frac{1}{\sqrt{P-j}}\sum_{t=R+j+1}^{n} H(t-j-1) + o_P(1)$.

(iii) $\sqrt{P-j}\widehat{\xi}_{P,j} \xrightarrow{d} N(0, \sigma_c^2)$, *where*

$$\sigma_c^2 = \alpha^2(1-\alpha)^2 + 2\lambda_{hl}B\eta + \lambda_{ll}BVB',$$

with λ_{hl} *and* λ_{ll} *as in the previous table,* and

$B_j := E[g'_\alpha(W_{t-1}, \theta_0) f_{W_{t-1}}(m_\alpha(W_{t-1}, \theta_0))\{I_{t-j,\alpha}(\theta_0) + \alpha\}]$ and

$\eta \equiv \eta_j =: E[(I_{t,\alpha}(\theta_0) I_{t-j,\alpha}(\theta_0) - \alpha^2) l(Y_t, W_{t-1}, \theta_0)]$.

Estimating the asymptotic variance of $\widehat{\xi}_{P,j}$

The vector B can be consistently estimated by $\widehat{B}_\tau$, where

$$\widehat{B}_\tau = -\frac{1}{P-j}\sum_{t=R+j+1}^{n}\frac{1}{\tau}\exp\left[\left(Y_t - m_\alpha(W_{t-1},\widehat{\theta}_{t-1})\right)/\tau\right] I_{t,\alpha}(\widehat{\theta}_{t-1})\{I_{t-j,\alpha}(\widehat{\theta}_{t-j-1})+\alpha\}g'_\alpha(W$$

with $\tau \to 0$ as $n \to \infty$.

Then

$$\widehat{\sigma}_c^2 := \alpha^2(1-\alpha)^2 + 2\widehat{\lambda}_{hl}\widehat{B}_\tau\widehat{\eta} + \widehat{\lambda}_{ll}\widehat{B}_\tau\widehat{V}\widehat{B}'_\tau,$$

where

$$\widehat{\eta} = \frac{1}{P-j}\sum_{t=R+j+1}^{n}(I_{t,\alpha}(\widehat{\theta}_{t-1})I_{t-j,\alpha}(\widehat{\theta}_{t-j-1}) - \alpha^2)l(Y_t, W_{t-1}, \widehat{\theta}_{t-1}).$$

Then, valid inference can be accomplished by the *corrected independence backtesting test* statistic

$$\widetilde{\xi}_{P,j} \equiv \widetilde{\xi}(P, R, \widehat{\theta}_{t-1}) = \frac{1}{\widehat{\sigma}_c \sqrt{P-j}} \sum_{t=R+j+1}^{n} (I_{t,\alpha}(\widehat{\theta}_{t-1}) I_{t-j,\alpha}(\widehat{\theta}_{t-j-1}) - \alpha^2),$$

that satisfies

COROLLARY 3: *Under Assumptions A1-A5,* $E[I_{t,\alpha}(\theta_0) \mid W_{t-1}(\theta_0)] = \alpha$ *and that* $\tau \to 0$ *as* $n \to \infty$, *for each* $j \geq 1$

$$\widetilde{\xi}_{P,j} \xrightarrow{d} N(0,1).$$

Example: Location-Scale Models

These models are defined as

$$Y_t = \mu(W_{t-1}, \beta_0) + \sigma(W_{t-1}, \beta_0)\varepsilon_t, \tag{7}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are specifications for the conditional mean and standard deviation of Y_t given W_{t-1}, respectively, and ε_t are the standardized innovations which are usually assumed to be iid, and independent of W_{t-1}.

Under such assumptions the α-th conditional VaR is given by

$$m_\alpha(W_{t-1}, \theta_0) = \mu(W_{t-1}, \beta_0) + \sigma(W_{t-1}, \beta_0)F_\varepsilon^{-1}(\alpha), \tag{8}$$

where $F_\varepsilon^{-1}(\alpha)$ denotes a univariate quantile function of ε_t and the nuisance parameter is $\theta_0 = (\beta_0, F_\varepsilon^{-1}(\alpha))$.

Estimation risk for a location-scale model

If the parameters β_0 are estimated using a fixed forecasting scheme the estimation risk term for the unconditional test takes this form:

$$\sqrt{\pi}\sqrt{R}(F_{\varepsilon,R}^{-1}(\alpha) - F_{\varepsilon}^{-1}(\alpha))f_{\varepsilon}(F_{\varepsilon}^{-1}(\alpha)) + \sqrt{\pi}\sqrt{R}(\widehat{\beta}_R - \beta_0)' a(\alpha, \beta_0), \tag{9}$$

where $F_{\varepsilon,R}^{-1}(\alpha)$ is an α-quantile estimator of the innovation distribution, and

$$a(\alpha, \beta_0) := f_{\varepsilon}(F_{\varepsilon}^{-1}(\alpha))E\left[a_{1,t}(\beta_0)\right] + f_{\varepsilon}(F_{\varepsilon}^{-1}(\alpha))F_{\varepsilon}^{-1}(\alpha)E\left[a_{2,t}(\beta_0)\right], \tag{10}$$

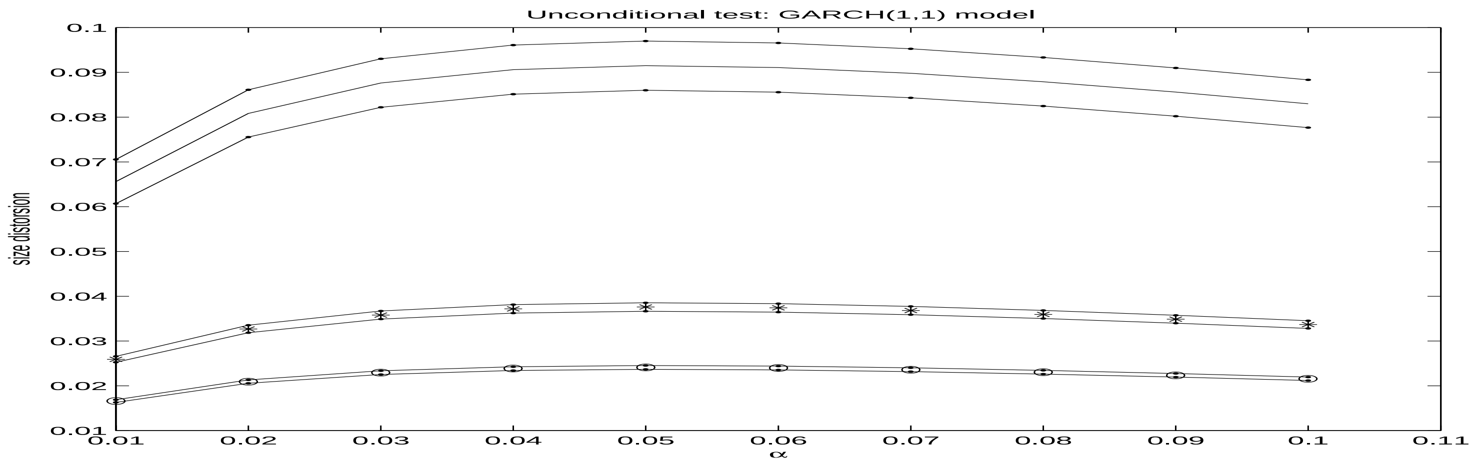
with

$$a_{1,t}(\beta) = \dot{\mu}_t(\beta)/\sigma(W_{t-1}, \beta), \qquad a_{2,t}(\beta) = \dot{\sigma}_t(\beta)/\sigma(W_{t-1}, \beta),$$

and where $\dot{\mu}_t(\beta) = \partial\mu(W_{t-1}, \beta)/\partial\beta$ and $\dot{\sigma}_t(\beta) = \partial\sigma(W_{t-1}, \beta)/\partial\beta$.

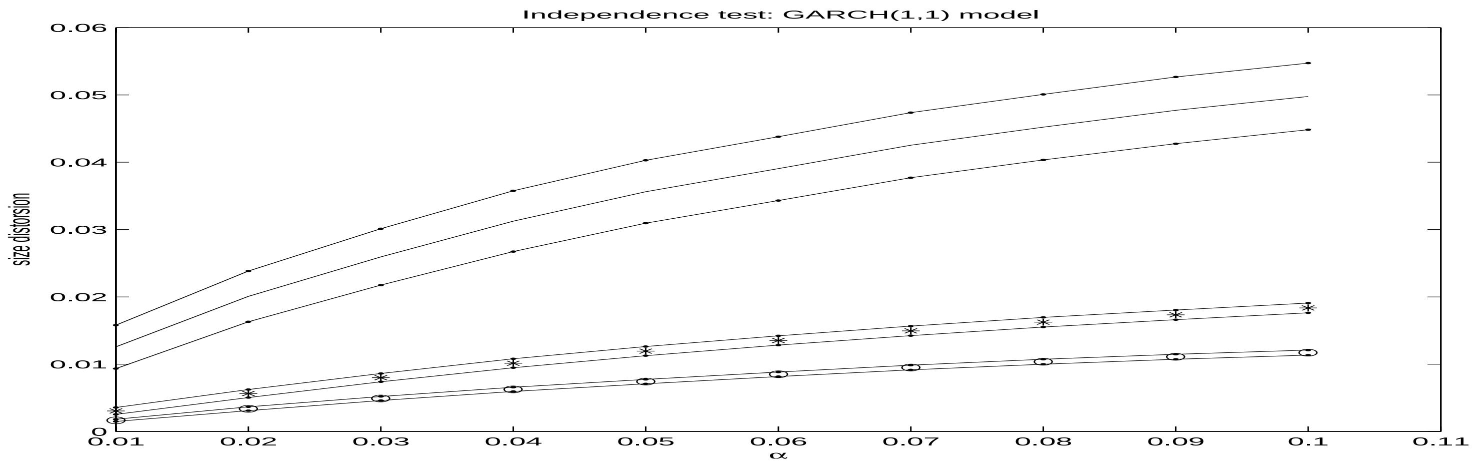
Note there are two sources of estimation risk in this model, one from estimating $F_{\varepsilon}^{-1}(\alpha)$ and other from estimating β_0.

Size distortion for Unconditional test: $d(\alpha) = 2\left(1 - \Phi\left(\frac{\alpha^{1/2}(1-\alpha)^{1/2}}{\widehat{\sigma}_u(\alpha)} z_{0.025}\right)\right) - 0.05$.



$$P_{H_0}\{|\frac{S_P}{\sqrt{\alpha(1-\alpha)}}| > N(0,1)\} - P_{H_0}\{|\widetilde{S}_P| > N(0,1)\}.$$

Size distortion for Dependence test: $d(\alpha) = 2\left(1 - \Phi\left(\frac{\alpha^{1/2}(1-\alpha)^{1/2}}{\widehat{\sigma}_C(\alpha)} z_{0.025}\right)\right) - 0.05.$



$P_{H_0}\{\left|\frac{\widehat{\xi}_{P,j}}{\alpha(1-\alpha)}\right| > N(0,1)\} - P_{H_0}\{\left|\widetilde{\xi}_{P,j}\right| > N(0,1)\}$, with j denoting the number of lags.

Simulation Exercise: GARCH(1,1) for financial returns:

$$Y_t = \sigma(W_{t-1}, \beta_0)\varepsilon_t, \qquad \sigma^2(W_{t-1}, \beta_0) = \eta_{00} + \eta_{10}Y_{t-1}^2 + \eta_{20}\sigma^2(W_{t-2}, \beta_0),$$

where $\{\varepsilon_t\}$ are iid t_ν standardized disturbances.

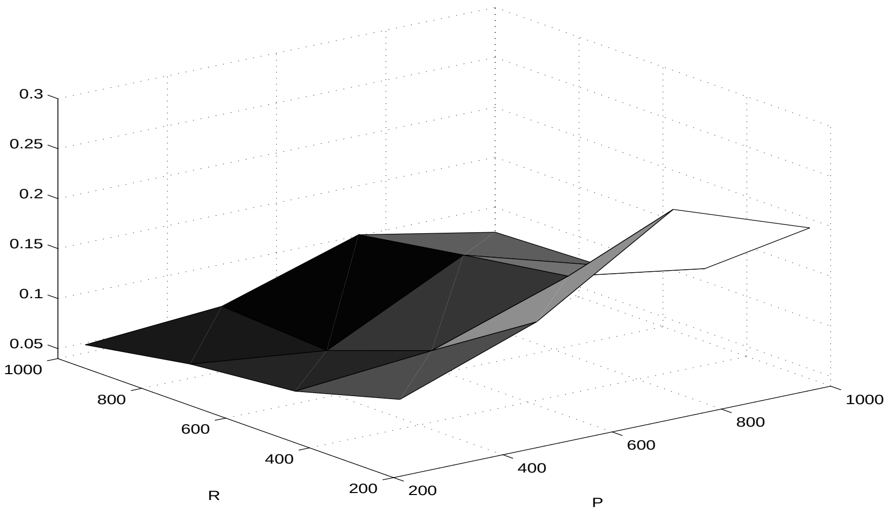
$\beta_0 \equiv (\eta_{00}, \eta_{10}, \eta_{20}) = (0.05, 0.1, 0.85)$, and $\nu = 30, 5$.

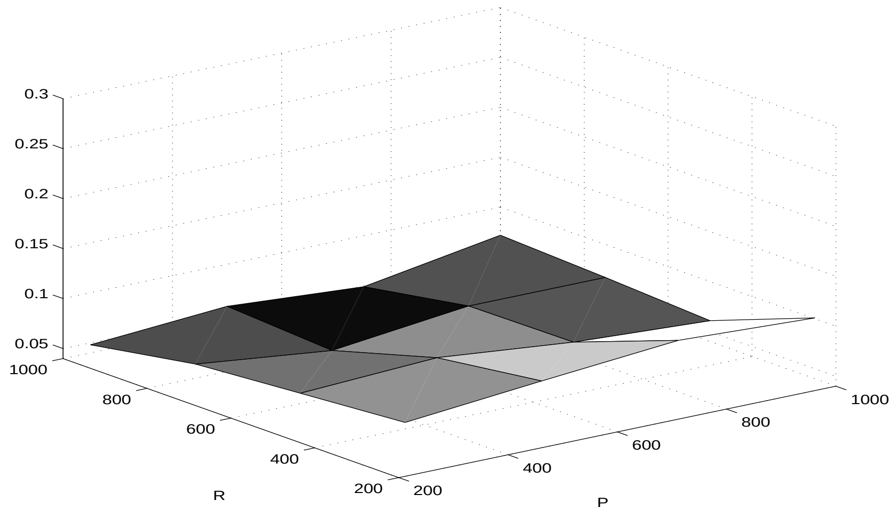
In-sample size using for fixed estimator: $R = [250, 500, 750, 1000]$.

Out-of-sample size for testing forecasting ability:

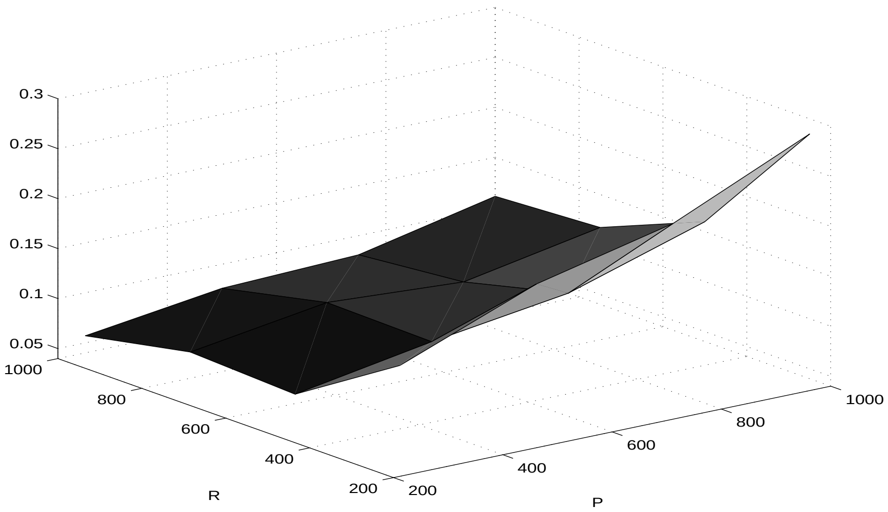
$P = [250, 500, 750, 1000]$.

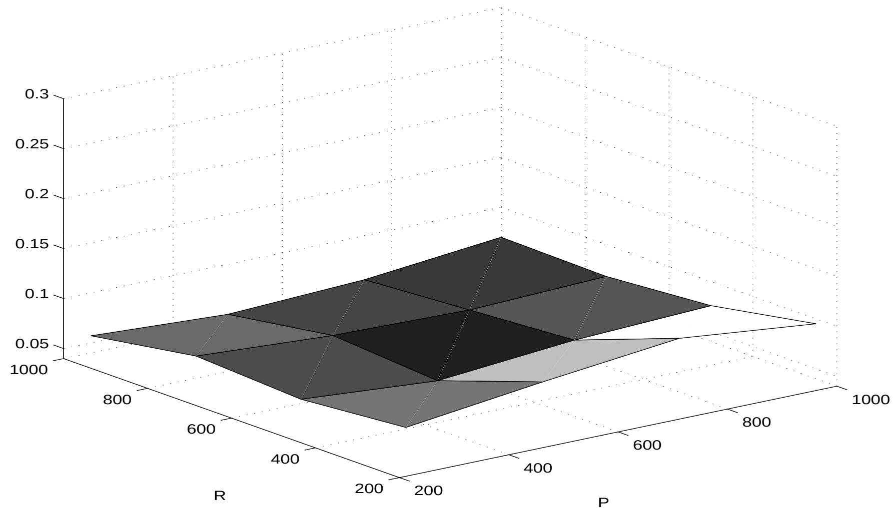
M=500 Monte-Carlo simulations.



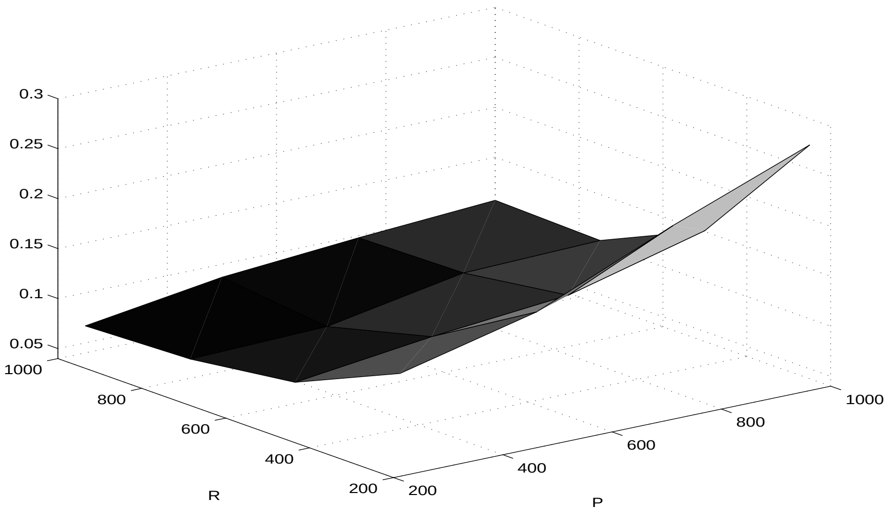


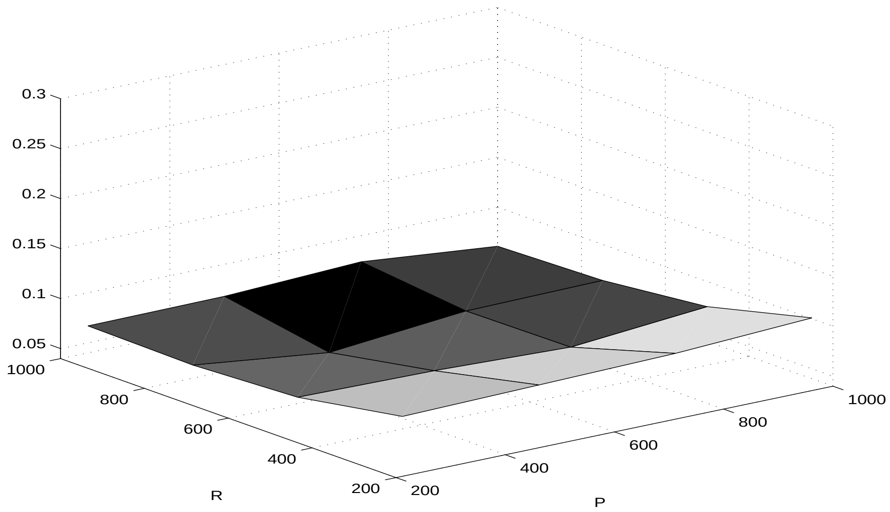
Simulated size at 0.05 for S_P and $\widetilde{S}_P$. $\alpha = 0.01$, $\nu = 30$



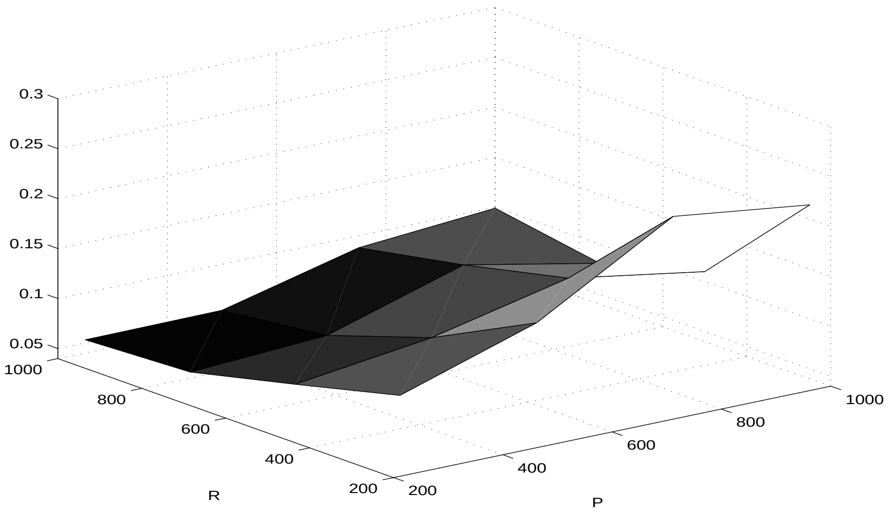


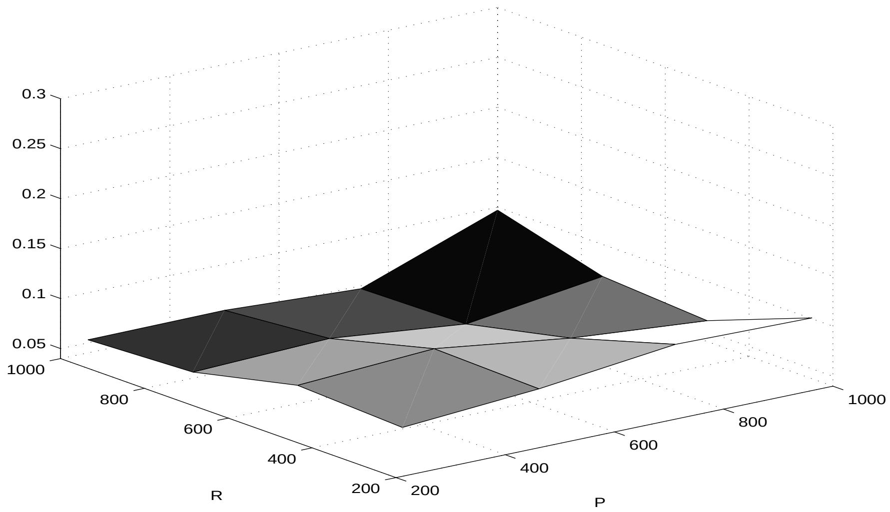
Simulated size at 0.05 for S_P and $\widetilde{S}_P$. $\alpha = 0.05$, $\nu = 30$



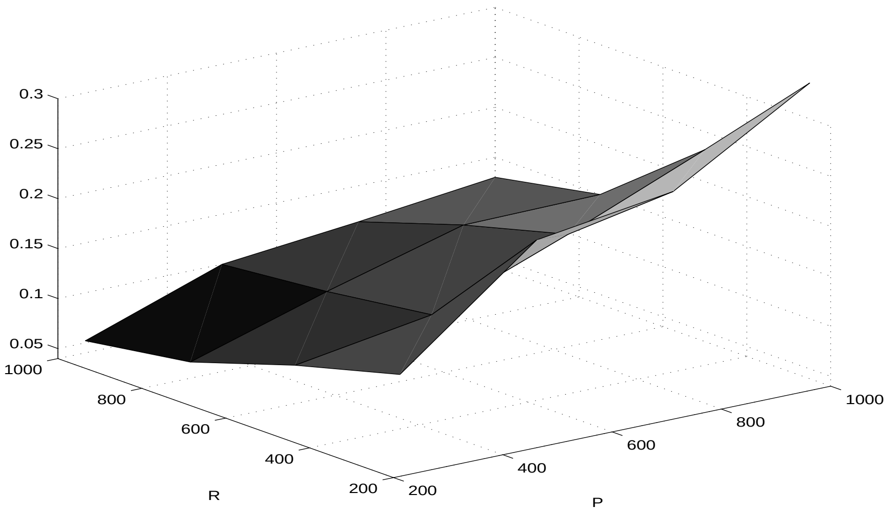


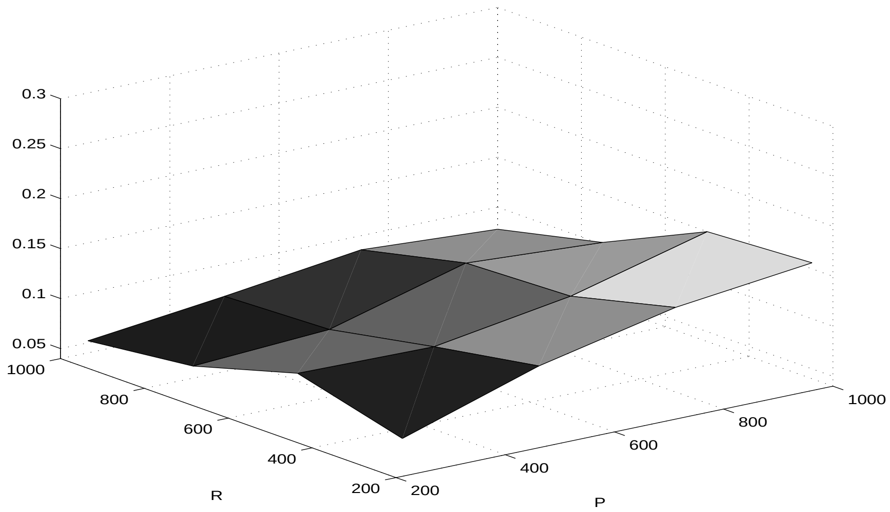
Simulated size at 0.05 for S_P and $\widetilde{S}_P$. $\alpha = 0.1$, $\nu = 30$



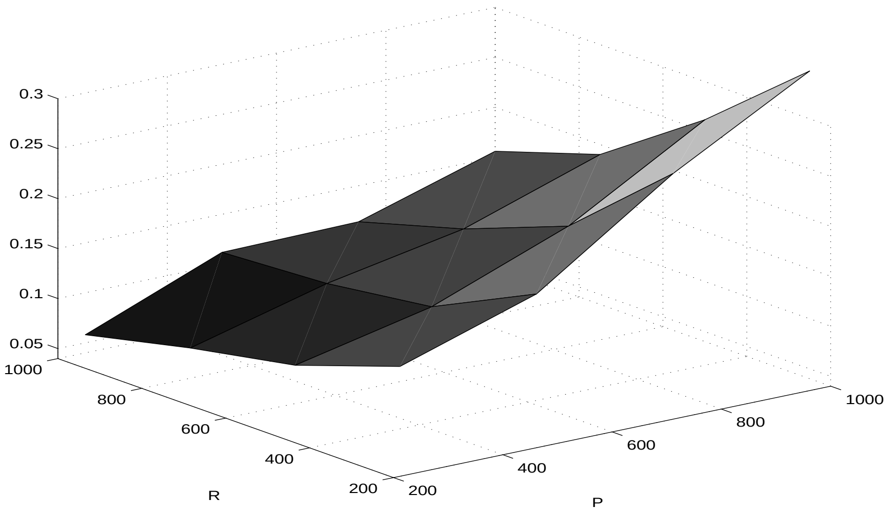


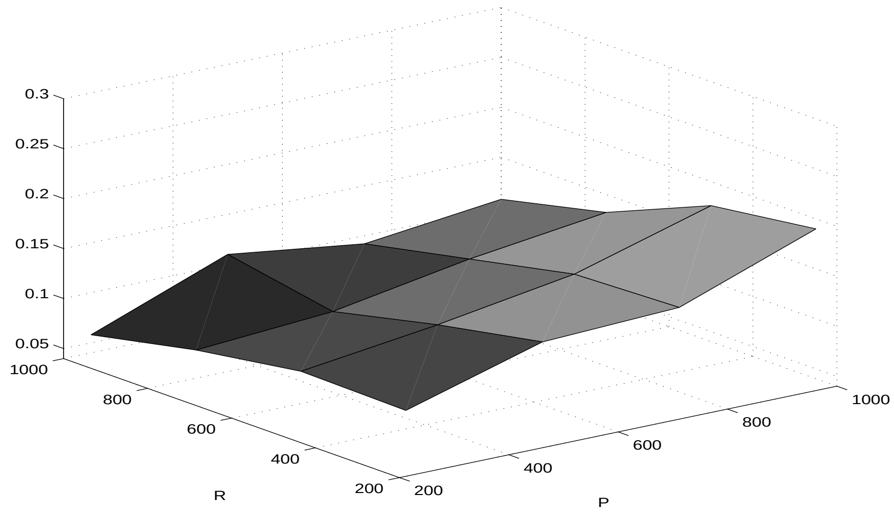
Simulated size at 0.05 for S_P and $\widetilde{S}_P$. $\alpha = 0.01$, $\nu = 5$





Simulated size at 0.05 for S_P and $\widetilde{S}_P$. $\alpha = 0.05$, $\nu = 5$

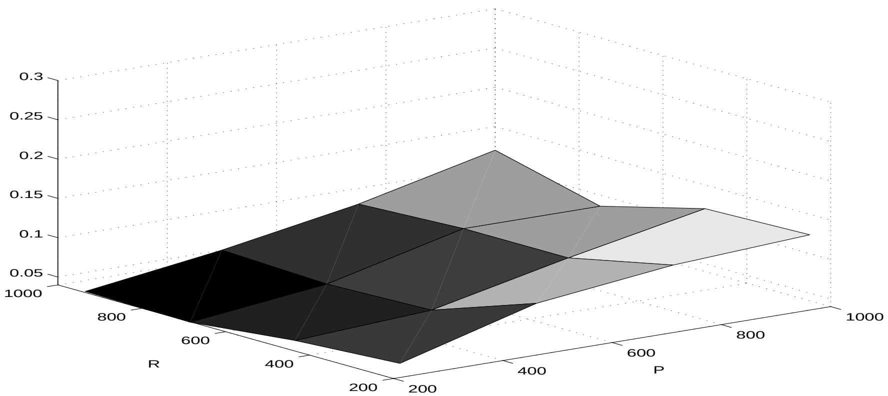


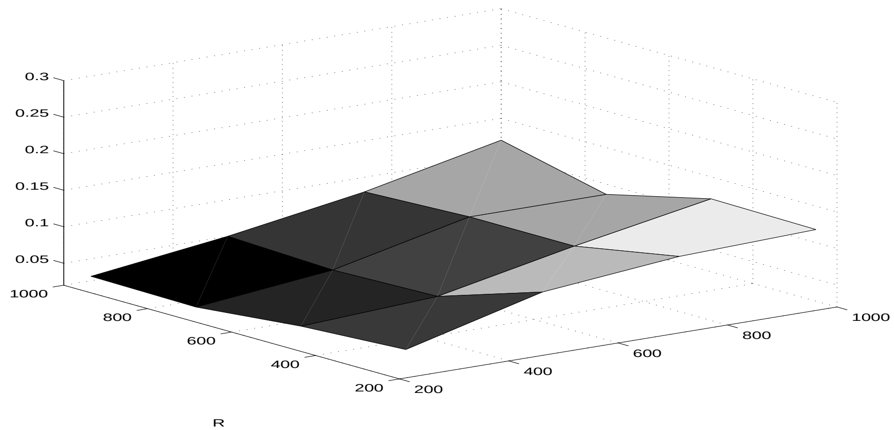


Simulated size at 0.05 for S_P and $\widetilde{S}_P$. $\alpha = 0.1$, $\nu = 5$

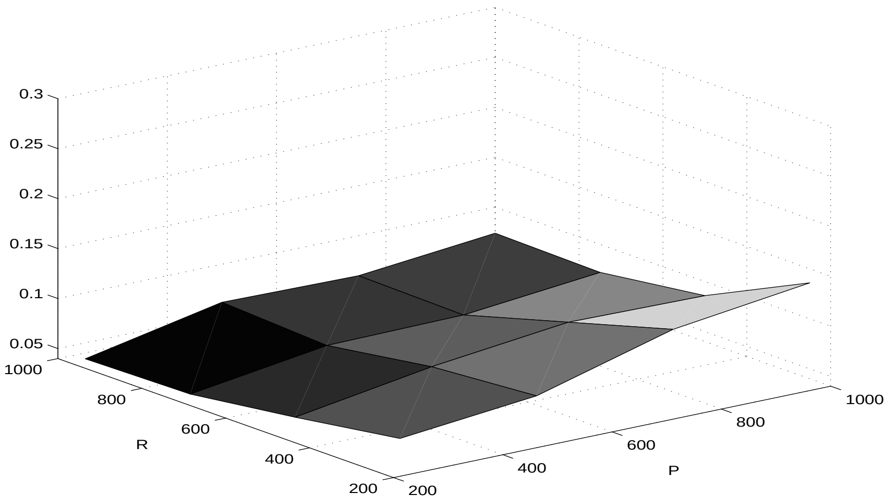
Simulated size for independence tests:

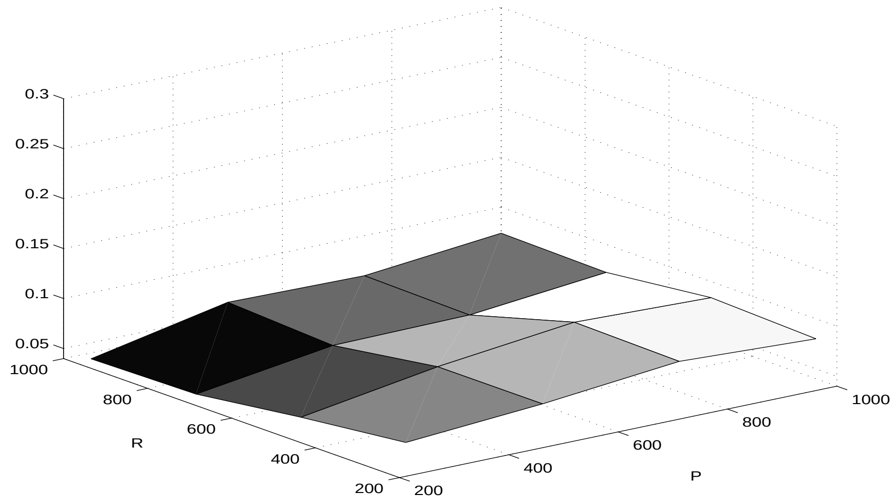
$$C_{P,j} = (P-j)\frac{\widehat{\xi}^2_{P,j}}{\alpha^2(1-\alpha^2)} \qquad \textbf{and} \qquad \widetilde{C}_{P,j} = \widetilde{\xi}^2_{P,j}$$



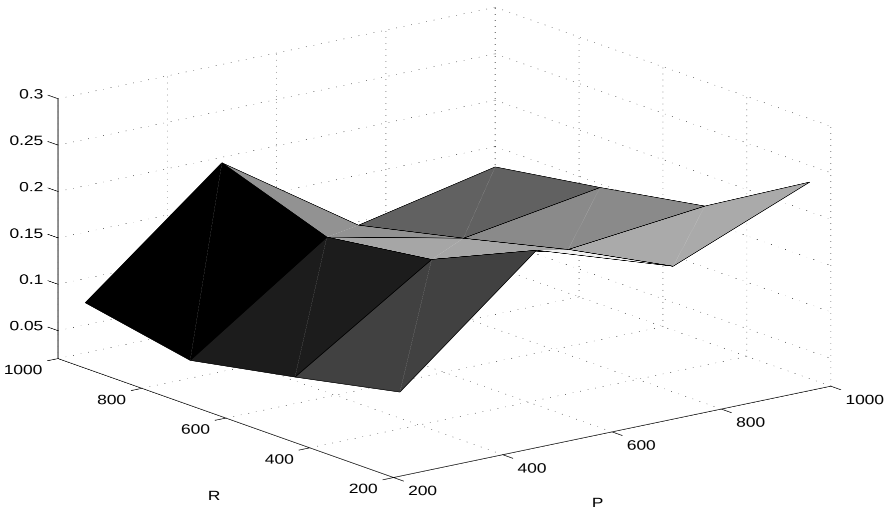


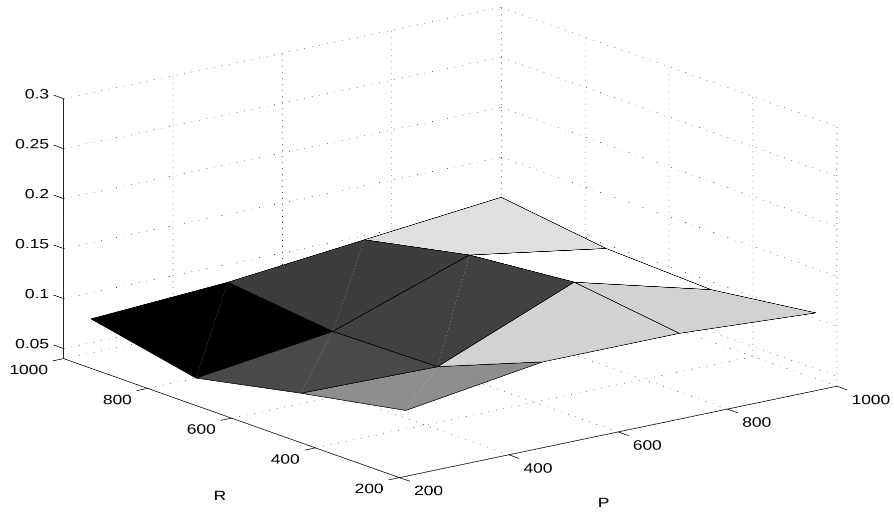
Simulated size at 0.05 for $C_{P,j}$ and $\widetilde{C}_{P,j}$. $\alpha = 0.01$, $\nu = 30$



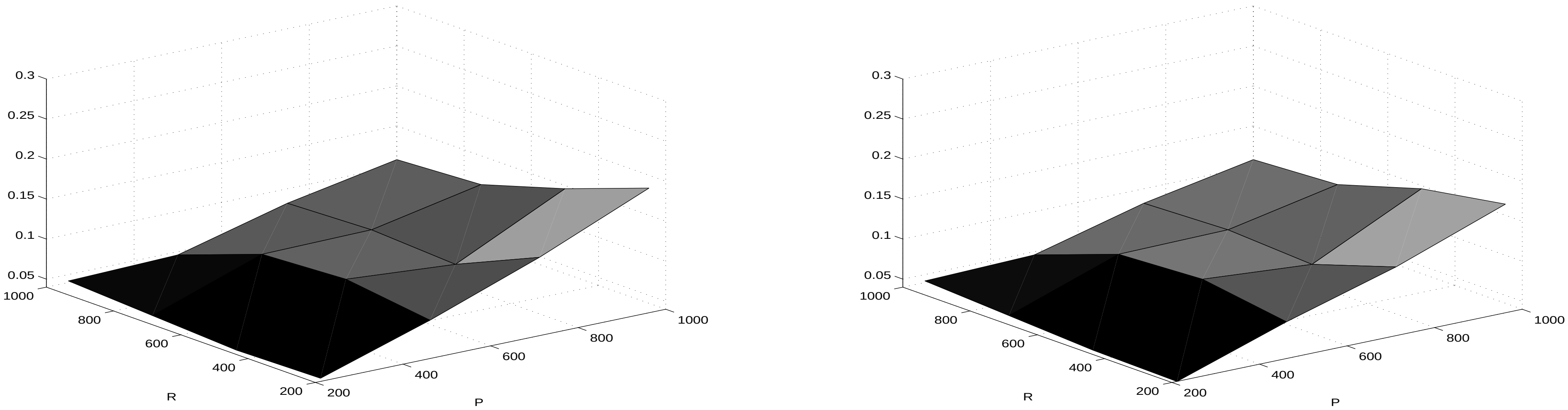


Simulated size at 0.05 for $C_{P,j}$ and $\widetilde{C}_{P,j}$. $\alpha = 0.05$, $\nu = 30$

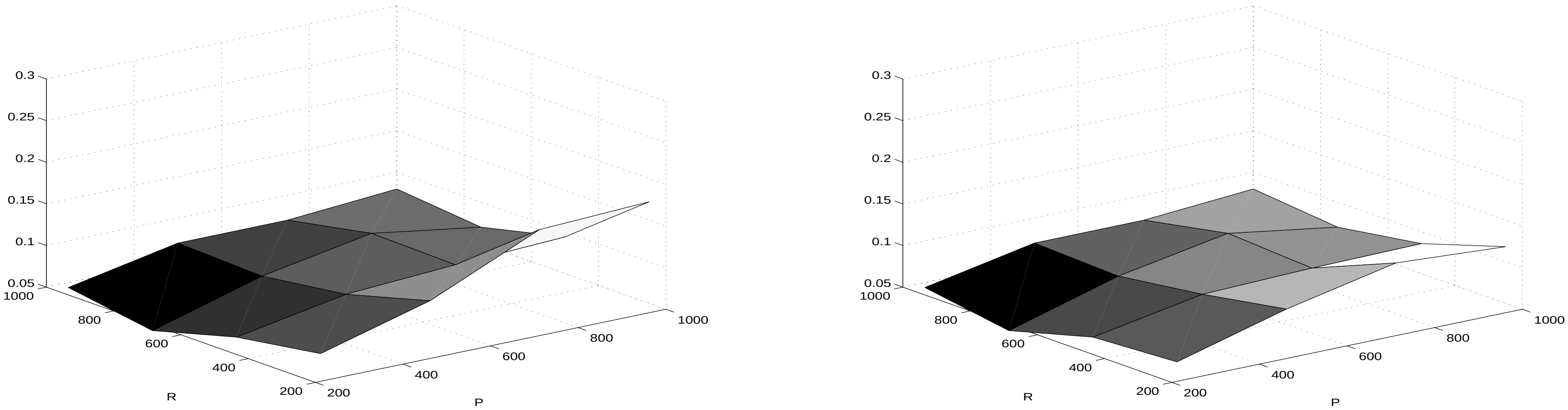




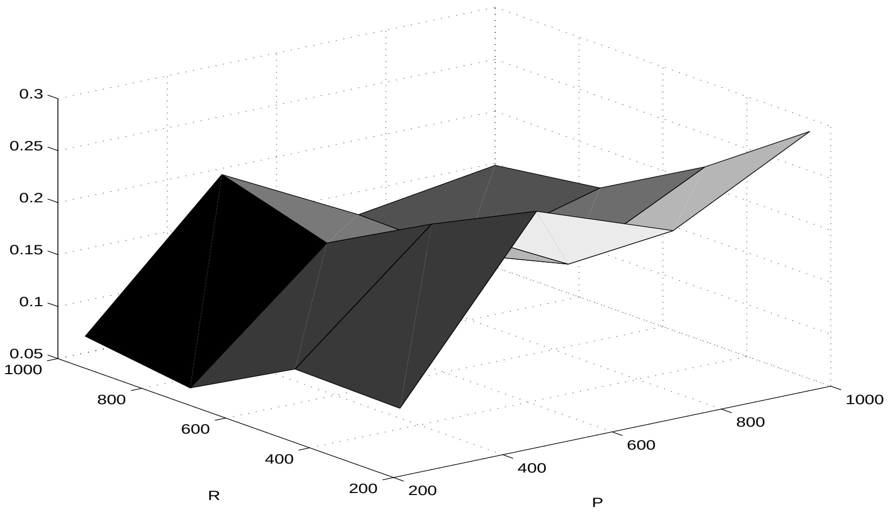
Simulated size at 0.05 for $C_{P,j}$ and $\widetilde{C}_{P,j}$. $\alpha = 0.1$, $\nu = 30$

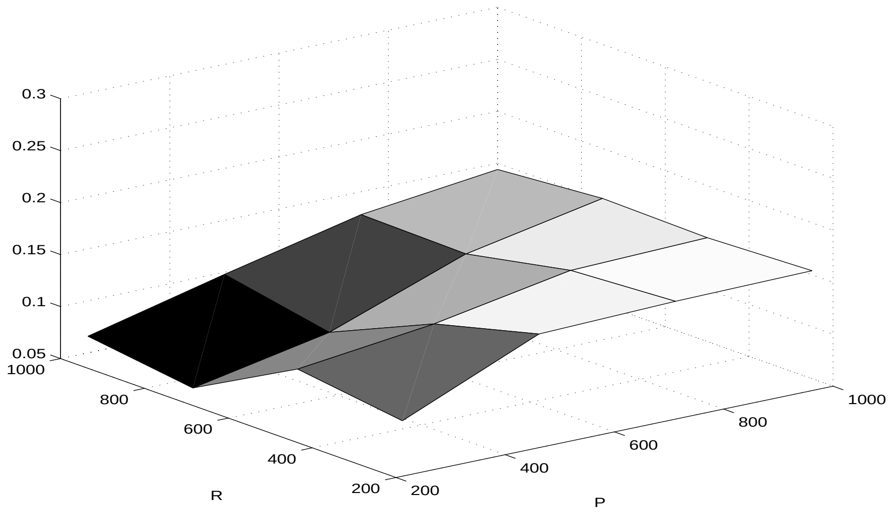


Simulated size at 0.05 for $C_{P,j}$ and $\widetilde{C}_{P,j}$. $\alpha = 0.01$, $\nu = 5$



Simulated size at 0.05 for $C_{P,j}$ and $\widetilde{C}_{P,j}$. $\alpha = 0.05$, $\nu = 5$





Simulated size at 0.05 for $C_{P,j}$ and $\widetilde{C}_{P,j}$. $\alpha = 0.1$, $\nu = 5$

$\alpha = 0.01$		GARCH-M			TAR			EGARCH		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
$P = 250$	S_P	0.000	0.000	0.000	0.176	0.176	0.026	0.622	0.622	0.280
	$\widetilde{S}_P$	0.208	0.000	0.000	0.342	0.176	0.056	0.700	0.620	0.396
	$C_{P,1}$	1.000	1.000	0.000	1.000	1.000	0.001	1.000	1.000	0.100
	$\widetilde{C}_{P,1}$	0.539	0.519	0.001	0.881	0.854	0.056	0.582	0.525	0.261
$P = 500$	S_P	0.193	0.193	0.000	0.234	0.234	0.050	0.806	0.806	0.462
	$\widetilde{S}_P$	0.384	0.131	0.000	0.324	0.230	0.080	0.842	0.772	0.558
	$C_{P,1}$	1.000	0.006	0.006	1.000	0.148	0.148	1.000	0.587	0.587
	$\widetilde{C}_{P,1}$	0.548	0.006	0.003	0.924	0.147	0.117	0.729	0.586	0.411

$\alpha = 0.05$		GARCH-M			TAR			EGARCH		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
$P = 250$	S_P	0.805	0.501	0.071	0.748	0.578	0.320	0.348	0.220	0.094
	$\widetilde{S}_P$	0.778	0.487	0.064	0.748	0.578	0.320	0.344	0.220	0.094
	$C_{P,1}$	0.007	0.007	0.000	0.831	0.831	0.499	0.660	0.660	0.234
	$\widetilde{C}_{P,1}$	0.004	0.003	0.000	0.831	0.822	0.498	0.612	0.497	0.214
$P = 500$	S_P	0.954	0.823	0.304	0.842	0.750	0.484	0.422	0.262	0.102
	$\widetilde{S}_P$	0.938	0.869	0.508	0.842	0.770	0.606	0.406	0.300	0.154
	$C_{P,1}$	0.006	0.002	0.000	0.926	0.837	0.619	0.786	0.658	0.365
	$\widetilde{C}_{P,1}$	0.005	0.001	0.000	0.926	0.837	0.730	0.756	0.658	0.490

$\alpha = 0.01$		SV			BIL			NLMA		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
$P = 250$	S_P	0.320	0.320	0.066	0.004	0.004	0.000	0.002	0.002	0.000
	$\widetilde{S}_P$	0.420	0.320	0.116	0.290	0.004	0.002	0.392	0.002	0.000
	$C_{P,1}$	1.000	1.000	0.002	1.000	1.000	0.000	1.000	1.000	0.000
	$\widetilde{C}_{P,1}$	0.272	0.105	0.026	0.957	0.947	0.000	0.960	0.939	0.000
$P = 500$	S_P	0.510	0.510	0.144	0.084	0.084	0.000	0.164	0.164	0.000
	$\widetilde{S}_P$	0.594	0.508	0.220	0.296	0.084	0.000	0.432	0.164	0.000
	$C_{P,1}$	1.000	0.125	0.125	1.000	0.000	0.000	1.000	0.000	0.000
	$\widetilde{C}_{P,1}$	0.268	0.124	0.065	0.978	0.000	0.000	0.980	0.000	0.000

$\alpha = 0.05$		SV			BIL			NLMA		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
$P = 250$	S_P	0.134	0.040	0.002	0.526	0.226	0.016	0.694	0.364	0.034
	$\widetilde{S}_P$	0.122	0.040	0.002	0.526	0.226	0.016	0.694	0.364	0.034
	$C_{P,1}$	0.067	0.067	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	$\widetilde{C}_{P,1}$	0.046	0.010	0.002	0.000	0.000	0.000	0.000	0.000	0.000
$P = 500$	S_P	0.164	0.050	0.002	0.764	0.490	0.056	0.926	0.730	0.176
	$\widetilde{S}_P$	0.144	0.074	0.014	0.764	0.610	0.206	0.926	0.812	0.392
	$C_{P,1}$	0.061	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$\widetilde{C}_{P,1}$	0.049	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Findings from the simulation experiment

- The corrected unconditional and independence tests have, for the models considered and uniformly in all VaR levels, better size performance than the respective uncorrected tests.
- The larger the parameter $\widehat{\pi} = P/R$ the higher the distortions. These distortions are of positive sign, $i.e.$ overrejection, which is consistent with our asymptotic theory.
- For the independence test, the greater the α the more important is the correction.
- The size improvement of $\widetilde{S}_P$ is without sacrificing power for the unconditional hypothesis.

- The larger the VaR level α the better is the approximation by the asymptotic theory of the finite sample distributions.
- The approximation for $\alpha = 0.01$ is poor and may lead to misleading conclusions for common in-sample and out-of-sample sizes.
- In view of these Monte Carlo experiments, our recommendation is then to consider VaR levels $\alpha \geq 0.05$ and corrected tests to develop out-of-sample inferences in VaR models.

Application to financial data

Methodology:

- We entertain a pure Gaussian GARCH(1,1) model for the log-returns Y_t, that yields the following VaR risk model, where $\Phi_{\varepsilon}^{-1}(\alpha)$ is the α-quantile of the Gaussian error distribution.

- The parameters are estimated by $QMLE$ using $R = 250$ observations.

- The out-of-sample period is also $P = 250$ observations, thus, with the data set available we have repeated the backtesting experiment for five different periods starting in February 2000.

$$m_\alpha(W_{t-1}, \theta_0) = \sigma_t \Phi_\varepsilon^{-1}(\alpha), \qquad \sigma_t^2 = \eta_{00} + \eta_{10} Y_{t-1}^2 + \eta_{20} \sigma_{t-1}^2,$$

GARCH(1,1)	P1	P2	P3	P4	P5
$\widehat{\eta}_{00}$	0.176	0.115	0.112	0.010	0.263
	(0.111)	(0.073)	(0.085)	(0.015)	(0.357)
$\widehat{\eta}_{10}$	0.222	0.095	0.098	0.057	0.061
	(0.053)	(0.037)	(0.044)	(0.026)	(0.063)
$\widehat{\eta}_{20}$	0.706	0.835	0.861	0.930	0.431
	(0.084)	(0.068)	(0.061)	(0.035)	(0.703)

GARCH(1,1)	vio	S_P	$\widetilde{S}_P$	$C_{P,1}$	$\widetilde{C}_{P,1}$
$\alpha = 0.01$					
P1	4	0.953	0.869	39.11^{**}	37.39^{**}
P2	3	0.317	0.278	0.025	0.025
P3	0	-1.589	-1.228	0.025	0.025
P4	2	-0.317	-0.261	0.025	0.025
P5	1	-0.953	-0.870	0.025	0.025
$\alpha = 0.05$					
P1	14	0.435	0.380	0.254	0.223
P2	16	1.015	0.838	3.391	2.923
P3	4	-2.466^{**}	-1.733	0.692	0.643
P4	14	0.435	0.331	3.391	2.857
P5	8	-1.305	-1.143	0.692	0.652

Conclusion

- The standard unconditional and independence backtesting used by banks and regulators to assess dynamic parametric VaR estimates may be very misleading in composite environments.
- Any conclusion regarding the validity of these risk models based on standard backtesting procedures may be spurious.
- We find in this paper the appropriate cut-off point by correcting the variance in the relevant test statistics corresponding to the recursive, rolling and fixed out-of-sample forecasting schemes.
- We find evidence in the simulation exercises of significant size distortions for the Kupiec uncorrected test:
  - For independence tests the distortions are only significant for moderate and large values of α such as $\alpha = 0.1$.

- These distortions are remarkably important for backtesting exercises using large out-of-sample sizes and small in-sample sizes for estimating the parameters.
- Our simulations indicate that the approximation by the asymptotic theory is not accurate for small values of α such as $\alpha = 0.01$.

- We find in the empirical application that the standard unconditional backtesting procedure with VaR calculated with the fixed forecasting scheme overstates risk exposure yielding in the third period under study to a spurious rejection of VaR at 5% for the GARCH(1,1) model.
- These findings somehow support the scepticism of American regulators about the implementation of Basel II risk measurement and risk monitoring techniques, and should help to restore their confidence on internal risk management systems validated by this new corrected backtesting procedure.