Herwartz, Helmut

Working Paper

Exact inference in diagnosing value-at-risk estimates: A Monte Carlo device

Economics working paper / Christian-Albrechts-Universität Kiel, Department of Economics, No. 2008,16

Provided in cooperation with:
Christian-Albrechts-Universität Kiel (CAU)

Exact inference in diagnosing value-at-risk estimates – A Monte Carlo device

by Helmut Herwartz

C|A|U
Christian-Albrechts-Universität Kiel
Department of Economics

Economics Working Paper
No 2008-16

Exact inference in diagnosing value-at-risk estimates - A Monte Carlo device

Helmut Herwartz*

September 2, 2008

Abstract

In this note a Monte Carlo approach is suggested to determine critical values for diagnostic tests of Value-at-Risk models that rely on binary random variables. Monte Carlo testing offers exact significance levels in finite samples. Conditional on exact critical values the dynamic quantile test suggested by Engle and Manganelli (2004) turns out more powerful than a recently proposed Portmanteau type test (Hurlin and Tokpavi 2006).
Keywords: Value-at-Risk, Monte Carlo test.
JEL Classification: C22, C52, G28.

1 Introduction

Value-at-risk (VaR) is a widely used measure of portfolio risk. Formally, suppose that $\{y_t\}_{t=1}^T$ is a process of speculative returns. Then, conditional on the information set available in time $t-1$, Ω_{t-1}, the value-at-risk with coverage α, denoted $\text{VaR}_t(\alpha)$, is the quantile such that

$$\text{Prob}[y_t < -\text{VaR}_t(\alpha)|\Omega_{t-1}] = \alpha. \quad (1)$$

The computation of $\text{VaR}_t(\alpha)$ is challenging since the 'true' conditional distribution of returns is unknown. In the light of a plentitude of alternative approaches to VaR determination, model diagnosis is of essential importance in applied finance.

The remainder of this short note is organized as follows. Two particular tools for VaR diagnosis (Engle and Manganelli 2004, Hurlin and Tokpavi 2006) are sketched in the next Section. In light of their finite sample size biases a Monte Carlo approach to testing VaR accuracy is adopted in Section 3. Concluding remarks are provided in Section 4.

*Institut für Statistik und Ökonometrie, Christian-Albrechts-Universität zu Kiel, Germany, Herwartz@stat-econ.uni-kiel.de

2 VaR diagnosis

Starting from the definition of $\mathrm{VaR}_t(\alpha)$ in (1) so-called *hit* processes formalize the event of a conditional return shortfall. The centered *hit* process is

$$\delta_t^\alpha = \mathrm{Hit}_t(\alpha) \equiv I(y_t < -\mathrm{VaR}_t(\alpha)|\Omega_{t-1}) - \alpha, \tag{2}$$

where $I()$ is an indicator function. For diagnosing (condtional) $\mathrm{VaR}_t(\alpha)$ estimates it is important to verify that no piece of information contained in Ω_{t-1} carries explanatory content for the mean zero process δ_t^α. In particular, $\delta_{t-1}^\alpha, \delta_{t-2}^\alpha \ldots$ and lagged *hits* observed at other VaR nominal coverages $a \neq \alpha$ are part of Ω_{t-1}. For considering 'cross coverage' dynamics a *hit* vector is defined as $\delta_t = (\ \delta_t^{\alpha_1}, \delta_t^{\alpha_2}, \ldots, \delta_t^{\alpha_m})'$. In the following three competing statistics for VaR diagnosis are listed that rely on the centered *hit* process δ_t^α or its vector valued counterpart δ_t.

1. Binary regression (I): Engle and Manganelli (2004) suggest the dynamic quantile test building upon a regression design for binary variables. Conditional on presample values this regression reads as

$$\begin{aligned} \delta_t^\alpha &= \beta_0 + \sum_{i=1}^{p} \beta_i \delta_{t-i}^\alpha + u_t, \\ &= \mathbf{x}_{t-1}^{\alpha'} \beta + u_t,\ t = 1, \ldots, T, \end{aligned} \tag{3}$$

   where $\mathbf{x}_{t-1}^\alpha = (1, \delta_{t-1}^\alpha, \delta_{t-2}^\alpha, \ldots, \delta_{t-p}^\alpha)'$. For backtesting the conventional Wald statistic,

$$W^\alpha = \frac{\hat\beta'(\sum_{t-1} \mathbf{x}_{t-1}^\alpha \mathbf{x}_{t-1}^{\alpha'})\hat\beta}{\alpha(1-\alpha)} \xrightarrow{d} \chi^2(p+1), \tag{4}$$

   is used, where $p+1$ is the row dimension of $\mathbf{x}_{t-1}^\alpha$ and $\hat\beta$ contains OLS estimates from (3). The asymptotic $\chi^2(p+1)$-distribution holds if the VaR model is well specified (H_0).

2. Binary regression (II): Analogous to (3), diagnosing $\mathrm{VaR}_t(\alpha)$ may also exploit lagged *hits* measured for $\mathrm{VaR}_t(a)$, $a \neq \alpha$. Relying on lag one *hits*, for instance, obtains

$$\begin{aligned} \delta_t^{(\alpha)} &= \beta_0 + \sum_{i=2}^{p} \beta_i \delta_{t-i}^{(\alpha)} + \sum_{a=1}^{m} \gamma_a \delta_{t-1}^{(a)} + u_t, \\ &= \mathbf{x}_{t-1}' \beta + u_t, \end{aligned} \tag{5}$$

   with a redefined vector of explanatory variables $\mathbf{x}_{t-1}$ and $\beta = (\beta_0, \beta_2, \ldots, \beta_p, \gamma_1, \ldots, \gamma_m)$. To test $H_0 : \beta = 0$ the Wald statistic in (4) applies with a $\chi^2(p+m)$ asymptotic distribution under H_0.

3. A multivariate portmanteau statistic: Although the regression in (5) takes 'cross dependence' of VaR violations into account, the test is specific for nominal level α. For the vector *hit* process δ_t joint accuracy of VaR estimates implies

$$\mathrm{Cov}\,(\delta_t, \delta_{t-j}) = \begin{cases} C_0 & \text{for } j = 0 \\ 0 & \text{else.} \end{cases} \tag{6}$$

Covariance and correlation estimates are, respectively,

$$\hat{C}_j = \frac{1}{T}\sum_{t=j+1}^{T}\delta_t\delta'_{t-j},\ \hat{R}_j = D^{1/2}\hat{C}_j D^{1/2} \text{ and } D = I_m \odot \hat{C}_0,$$

where I_m is the m-dimensional identity matrix and '$\odot$' signifies element-by-element multiplication. Hurlin and Tokpavi (2006) propose a portmanteau statistic with asymptotic distribution applying under the null hypothesis of a well specified VaR model

$$Q_m(J) \quad = \quad T\sum_{j=1}^{J}\left(\text{vec}\hat{R}_j\right)'\left(\hat{R}_0^{-1}\otimes\hat{R}_0^{-1}\right)\left(\text{vec}\hat{R}_j\right) \xrightarrow{d} \chi^2(Jm^2). \tag{7}$$

Both approaches the dynamic quantile regression, in (3) or (5) and the portmanteau statistic in (7) adopt (auto)regression tools that originated in the framework of modelling continous random variables. Since *hit* statistics defined in (2) are binary, the postulated asymptotic distributions may only hold in very large samples. Thus, with access to finite dimensional sample information, an analyst runs the risk of biased inferential conclusions when founding test decisions on standard, $\chi^2(\bullet)$, critical values.

3 Monte Carlo test

The test statistics (4) and (7) build upon processes that can be very easily constructed by simulation under the null hypothesis of a correct (dynamic) VaR specification. Thus, Monte Carlo critical values (Dufour 2006) offer exact significance levels for VaR diagnosis. This section illustrates the merits of Monte Carlo based VaR diagnosis. Size violations invoked by inference based on asymptotic critical values are illustrated. Moreover, alternative diagnostic tools are compared in terms of power estimates implied by Monte Carlo critical values.

3.1 Design

To imitate a well specified VaR model, δ_t^α processes are determined by means of iid Gaussian variates. Let $\Phi(\bullet)$ denote the Gaussian distribution function. Then,

$$\delta_t^\alpha = I(\xi_t < \Phi^{-1}(\alpha)) \text{ with } \xi_t \sim iidN(0,1),\ t = 1,\ldots,T. \tag{8}$$

Vector *hit* processes are obtained by stacking δ_t^α for $\alpha \in \{.005, .01, .015, .02, .025, .03, .035\}$ ($m = 7$). To investigate the performance of VaR diagnostics in case of misspecification, ξ_t in (8) is generated as a first order autoregressive process, $\xi_t = 0.1\xi_{t-1} + \zeta_t$, $\zeta_t \sim iidN(0, 0.99)$. The adjusted variance for drawing ζ_t ensures that the unconditional variance of ξ_t is one. To specify the regression models in (3) or (5) the lag order is $p = 5$. For the portmanteau statistic in (7) order parameters $J = 5, 15$ are considered alternatively. Thus, the asymptotic distributions of alternative diagnostic tools are $\chi^2(q)$, $q = 6, 12, 245, 735$. Simulated sample sizes vary from

T	W^{α}, $\chi^2(p+1)$			W^{α}, $\chi^2(p+m)$			$Q_m(J)$	
	.005	.01	.025	.005	.01	.025	$J=5$	15
	Empirical size ($\chi^2(\bullet)$ 5% critical values)							
2	.223	.101	.062	.201	.149	.102	.265	.250
4	.118	.079	.059	.147	.125	.084	.224	.215
8	.094	.074	.054	.134	.099	.069	.170	.167
20	.083	.063	.049	.094	.077	.059	.113	.116
50	.067	.051	.049	.075	.063	.056	.080	.078
	Empirical power (Monte Carlo 5% critical values)							
2	.068	.080	.098	.100	.120	.143	.106	.085
4	.076	.082	.122	.121	.150	.190	.134	.090
8	.085	.109	.187	.176	.200	.280	.200	.126
20	.112	.174	.397	.259	.343	.491	.324	.176
50	.165	.327	.766	.434	.606	.823	.580	.320

Table 1: Top panel: Empirical size estimates for alternative asymptotic tests diagnosing $\text{VaR}_t(\alpha)$ with 5% nominal significance level. Bottom panel: Power estimates implied by applying 5% Monte Carlo critical values. Considered sample sizes are $T \cdot 1000$.

conventional magnitudes to extreme time dimensions, i.e. $T \in \{2000, 4000, 8000, 20000, 50000\}$. 10000 Monte Carlo replications are used for all experiments.

An often raised caveat of Monte Carlo results is their dependence on the data generating process used for the simulation. For simulating VaR *hits* under the null hypothesis of a well specified risk model, it is worthwhile to point out that the design in (8) matches the null hypothesis in a one-to-one manner. Employing competing, well specified VaR models for imitation of the null hypothesis would not deliver systematically different empirical size features. The data generating mechanism employed to imitate poor VaR specifications lacks parametric rigor. It is interesting, though, to investigate how alternative tests cope with a 'nonparametric' alternative.

3.2 Results

Table 1 provides inferential results for diagnosing VaR specifications with 5% significance. The top panel documents rejection frequencies of alternative tests based on the first order asymptotic approximations. Rejection frequencies under misspecification of the VaR model as implied by Monte Carlo critical values are displayed below. Performing analogous experiments for nominal significance levels 1% or 10% yields qualitatively identical results which are not displayed for space considerations.

Marked size distortions feature all diagnostic tools for conventional sample sizes ($T = 2000, 4000$) if the nominal coverage level α is small, 1% or less. Throughout, overrejections under the null hypothesis are higher for the Portmanteau statistic $Q_m(J)$ as for the Wald statistics W^α. For instance, empirical rejection frequencies for diagnosing $\mathrm{VaR}_t(.01)$ estimates exceed twice the nominal level for both regression models (3) and (5) conditional on $T = 2000$. Testing joint accuracy of the 7-dimensional *hit* process ($T = 2000$) obtains empirical size estimates of 26.5% ($J = 5$) and 25.0% ($J = 15$), respectively.

Convergence to the asymptotic pivotal distributions is effectively very slow. The speed of convergence of Wald test rejection frequencies depends on the investigated nominal VaR coverage. For instance, employing test regression (3) for $\alpha = .01$ and $\alpha = .025$ the empirical significance level comes close to the nominal level for $T = 50000$ (with rejection frequency 5.1%) and $T = 20000$ (with rejection frequency 4.9%), respectively. With respect to the augmented test regression (5) significant oversizing features $\mathrm{VaR}_t(\alpha)$ diagnosis for $\alpha = .005, .01$ even conditional on an 'extreme' sample size $T = 50000$.

Using Monte Carlo 5% critical values to diagnose VaR misspecification uncovers marked power differentials. Irrespective of the considered sample size $Q_m(5)$ turns out more powerful in comparison with $Q_m(15)$. The former is less powerful than a Wald statistic based on the regression (5) and testing $\mathrm{VaR}_t(\alpha)$ specifications with $\alpha \geq 0.1$. Diagnosing $\mathrm{VaR}_t(.005)$ appears most challenging in terms of power. Obviously, the power of the Wald test in (4) is improved by conditioning on lagged *hit* processes measured for nominal VaR coverage $a \neq \alpha$.

4 Conclusion

In diagnosing conditional VaR estimates the dynamic quantile test (Engle and Manganelli 2004) and a Portmanteau approach (Hurlin and Tokpavi 2006) are shown to suffer from marked size distortions prevailing in even rather large (finite) samples. With regard to the dynamic quantile test oversizing is particularly likely if low VaR coverage levels are subjected to diagnostic testing. Monte Carlo inference offers exact empirical significance levels. In terms of power the dynamic quantile regression augmented with lagged hits measured for distinct VaR nominal coverages is most effective.

References

[1] Dufour, J.M. (2006) 'Monte Carlo Tests with Nuisance Parameters: A General Approach to Finite Sample Inference and Nonstandard Asymptotics in Econometrics, *Journal of Econometrics*, **133**, 443-477.

[2] Engle, R.F., S. Manganelli (2004), 'CAViaR: Conditional Autoregressive Value-at-Risk by Regression Quantiles, *Journal of Business and Economic Statistics*, **22**, 367-381.

[3] Hurlin, C., S, Tokpavi (2006), 'Backtesting VaR Accuracy: A New Simple Test', *Working Paper, LEO, University of Orléans.*