An Early Warning System for Tail Financial Risks

Gianni De Nicolo'
Johns Hopkins University
Carey Business School

ASSA 2021

Motivation

- Renewed efforts in constructing early warning systems for systemic risk in the aftermath of the financial crisis of 2007-2009
- Several Central Banks and the IMF conduct early warning exercises, often embedded in stress testing
- Financial institutions do the same for internal risk management and regulatory compliance
- Yet, there is no standardized forecasting procedure that maximizes forecasting performance of tail risk measures and provides vulnerability signals based on these forecasts
- This paper proposes such a procedure.

The Early Warning System (EWS)

- EWS based on based on real-time multi-period forecast *combinations* of Value-at-Risk (VaR) and Expected Shortfalls (ES) of portfolio returns of non-financial firms and banks.
- Forecast combinations include *baseline* (VaR,ES) forecasts conditional on a domestic risk factor, as well as *stress* (sVaR,sES) forecasts conditional on CoVaRs of the risk factor
- Implementated Using monthly data of the G-7 economies for the period 1975:01-2018:12,

Three novel features

1. Weight selection: determined by maximization of an average of a scoring function over a set of evaluation windows at each forecasting date.

2. Integrating stress testing with forecasting
   - The forecast combination includes forecasts conditional on risk factors (volatilities), called baseline forecasts, and forecasts conditional on the VaR of risk factors, called stress forecasts, and denoted by (sVaR,sES)
   - The sVaR measure is a forecasting version of the CoVaR (Adrian and Brunnermeier (2016). The sES measure is the ES conditional on the sVaR.
   - The value added of a stress test measured by the weights assigned to stress forecasts in the forecast combination.

3. A vulnerability index ES forecasts are used as predictors of a binary (Logit) model of the probability of the occurrence of VaR violations,

Forecasting Methods

- Forecast methods are specifications of models' forecasts that vary according to the length of the estimation window and the forecast evaluation window.
- Three basic models with an aggregate risk factor (log volatility) as a predictor:
  1. simple linear model with variance independent of the risk factor;
  2. Same as the first model, except that the variance of a return has the risk factor as predictor
  3. A quantile model with the risk factor as predictor
- The scoring function is the FZ0 function derived by Patton, Ziegel and Chen (2019),
- Tests of equal forecasting performance at each forecasting date and for a range of significance levels using the Diebold and Mariano (1995) tests.
- Zero weights are assigned to forecasts found inferior to at least one competing forecast at a given significance level, called *dominated* forecasts.

Results

- Significant out-of-sample tail financial risk forecasts and reliable vulnerability signals up to a 12-month forecasting horizon
- Stress forecasts have a significant role in improving performance, since they receive sizable weights in the forecast combinations.
- No ”forecast combination puzzle”: the equally weighted forecast combination does not dominate any forecast combination

The EWS set-up

1. Baseline and stress forecasts
2. The FZ0 scoring function
3. ”Optimal” forecast combinations
4. A vulnerability index

Baseline forecasts (1 of 3)

Model 1

$$R^{i,j}_{t+h} = \alpha^{i,j}_h + \beta^{i,j}_h V^i_t + \sigma^{i,j}_{t+h}\eta^{i,j}_{t+h} \tag{1}$$

The baseline forecasts (projections) of the h-month-ahead expected return and (VaR_τ, ES_τ) are:

$$E_t(\hat{R}^{i,j}_{t+h}) \equiv \hat{\alpha}^{i,j}_h + \hat{\beta}^{i,j}_h V^i_t \tag{2}$$

$$VaR_\tau(\hat{R}^{i,j}_{t+h}) = E_t(\hat{R}^{i,j}_{t+h}) + \hat{\sigma}^{i,j}_{t+h} G(\tau) \tag{3}$$

$$ES_\tau(\hat{R}^{i,j}_{t+h}) = E_t(\hat{R}^{i,j}_{t+h}) - \hat{\sigma}^{i,j}_{,t+h} H(\tau) \tag{4}$$

Baseline forecasts (2 of 3)

Model 2

Model 2's projection of the h-month-ahead return is the same as that of Model 1, but the variance depends on the risk factor:

$$\sigma_{2t+h} = \exp(\phi_0 + \phi_1 V_t) \tag{5}$$

The h-month-ahead baseline (VaR, ES) forecasts of Model 2 are therefore:

$$VaR_\tau(\bar{R}_{t+h}) = E_t(\hat{R}^{i,j}_{t+h}) + \sqrt{\exp(\bar{\phi}_0 + \bar{\phi}_1 V_t)}G(\tau) \tag{6}$$

$$ES_\tau(\bar{R}_{t+h}) = E_t(\hat{R}^{i,j}_{t+h}) - \sqrt{\exp(\bar{\phi}_0 + \bar{\phi}_1 V_t)}H(\tau) \tag{7}$$

where $G(\tau)$ and $H(\tau)$ are defined as above.

Baseline forecasts (3 of 3)

Model 3 (quantile model)

$$VaR_\tau(\hat{R}^{i,j}_{t+h}) = \hat{\alpha}^{i,j}_h(\tau) + \hat{\beta}^{i,j}_h(\tau)V^i_t \tag{8}$$

Conditional h-month-ahead ES forecast:

$$ES_\tau(\hat{R}^{i,j}_{t+h}) = E_t R^{i,j}_{t+h} - \tau^{-1}\hat{\sigma}^{i,j}_{t+h} \tag{9}$$

Gourieroux and Li (2012):

$$E_t R^{i,j}_{t+h} - \tau^{-1}\hat{\sigma}^{i,j}_{t+h} = L^h_{ij}(\tau) VaR_\tau(\hat{R}^{i,j}_{t+h}) \tag{10}$$

$$L^h_{ij}(\tau) = c^h_{ij,1}(\tau) I_{(VaR_\tau(\hat{R}^{ij}_{t+h})<0)} + c^h_{ij,2}(\tau) I_{(VaR_\tau(\hat{R}^{ij}_{t+h})>0)} \tag{11}$$

$$ES_\tau(\bar{R}^{ij}_{t+h}) = [\hat{c}^h_{ij,1}(\tau) I_{VaR_\tau(\hat{R}^{ij}_{t+h})<0} + \hat{c}^h_{ij,2}(\tau) I_{VaR_\tau(\hat{R}^{ij}_{t+h}>0)}] VaR_\tau(\hat{R}^j_{ij,t+h}) \tag{12}$$

Stress forecasts

- Stress forecasts are (VaR,ES) return forecasts conditional on CoVaRs of risk factors.
- CoVaRs of the risk factors that capture domestic and external tail risk shocks in reduced-form.
- The VaR of the risk factor V_t^i in country i;, and, the VaR of the leave-one-out average of risk factors across countries, defined by $V_t^{-i} \equiv \sum_{k \neq i}^{N} \frac{V_t^k}{N-1}$, for quantile levels $\tau' \leq \tau$:

$$VaR_{\tau'}(V_t^i) = a^i(\tau') + b^i(\tau')V_{t-1}^{-i} + c^i(\tau')V_{t-1}^i \tag{13}$$

$$VaR_{\tau'}(V_t^{-i}) = a^{-i}(\tau') + b^{-i}(\tau')V_{t-1}^{-i} \tag{14}$$

.

- Two stress scenarios defined by the following CoVaRs:

$$co_1 VaR_{\tau'}(V_t^i) = \hat{a}^i(\tau') + \hat{b}^i(\tau')V_{t-1}^{-i} + \hat{c}^i(\tau')VaR_{\tau'}(V_{t-1}^i) \tag{15}$$

$$co_2 VaR_{\tau'}(V_t^i) = \hat{a}^i(\tau') + \hat{b}^i(\tau')VaR_{\tau'}(V_{t-1}^{-i}) + \hat{c}^i(\tau')V_{t-1}^i \tag{16}$$

The FZ0 scoring function

- I use the following (strictly consistent) FZ0 scoring function derived by Patton, Ziegel and Chen (2019, Proposition 1), which applies to strictly negative values of VaR and ES:

$$FZ0(VaR_{t+h}, ES_{t+h}) \equiv -\frac{1}{\tau ES_{t+h}} I(R_{t+h} \leq VaR_{t+h})(VaR_{t+h} - R_{t+h}) + \frac{VaR_{t+h}}{ES_{t+h}} + \log(-VaR_{t+h}) - 1 \tag{17}$$

- The FZ0 statistics has negative orientation, that is, lower values indicate higher scores.

- The FZ0 scoring function applies to strictly negative values of VaR and ES (details in the paper)

"Optimal" forecast combinations (1 of 3)

- $\Delta f_{m,m'}(t,h)$ is the difference between the *FZ0* scores of methods m and m' in M.
- The performance of forecasting method m relative to m' at forecasting date t is tracked by the average of $\Delta f_{m,m',t}$ over a *rolling evaluation* window of the last w periods, given by:

$$\mu_t(m,m'|w) = \frac{1}{w}\sum_{t-w+1}^{t} \Delta f_{m,m'}(t,h) \qquad (18)$$

- α_j the j'th confidence level in the discrete set $A \equiv \{0.05,..,0.95\}$, and with W a set of evaluation windows of different lenght.
- The h-month-ahead forecast combination of (VaR, ES) at forecasting date t is given by:

$$(VaR_\tau(\hat{R}_{t+h}), ES_\tau(\hat{R}_{t+h})) = (\sum_{m=1}^{M} w_t^m VaR_m(\hat{R}_{t+h}), \sum_{m=1}^{M} w_t^m ES_m(\hat{R}_{t+h})) \qquad (19)$$

- The weights depend on the confidence level and the length of an evaluation window.

”Optimal” forecast combinations (2 of 3)

Optimal weights are determined in three steps

1. The inclusion of a forecast in a combination is determined by pairwise DM tests of equal forecasting performance at confidence level $\alpha_j \in A$ for any given evaluation window $w \in W$. *Dominated* forecasts are assigned zero weight.

2. Forecast combinations are compared for every confidence level in A and evaluation data window in W. The weights of each forecast at confidence level $\alpha_j \in A$ are computed as the fraction of the instances a forecast is non-dominated for all confidence levels preceding and including α_j.

3. The weights of the best forecast combination are obtained by selecting the confidence level α_j and the evaluation window w that minimize the average FZ0 score defined below.

"Optimal" forecast combinations (3 of 3)

$I^m(\alpha_j, w)$ is an indicator function of forecast m: 0 if forecast m is dominated, and 1 otherwise.

1. For all $\alpha_j \in A$ and $w \in W$, $I^m(\alpha_j, w) = 0$ if there exists a forecast m' such that: (a) $\mu(m, m'|w) > 0$; and, (b) the null hypothesis $\mu_(m, m'|w) = 0$ is rejected according to a DM test at a significance level $\alpha_j \in A$. $I^m(\alpha_j, w) = 1$ otherwise.
2. The weights of a forecast combination evaluated at the pair (α_j, w) are given by:

$$w_t^m(\alpha_j, w) = \frac{\sum_{h=1}^{j} I^m(\alpha_h, w)}{\sum_{m=1}^{M} \sum_{h=1}^{j} I^m(\alpha_h, w)} \tag{20}$$

3. The optimal weights are those associated with the pair (α_j, w) that minimizes the average FZ0 score defined by:

$$aFZ0(\alpha_j, w) \equiv \frac{1}{w} \sum_{i=t-w+1}^{t} FZ0(\sum_{m=1}^{M} w_i^m(\alpha_j, w) VaR^m(\hat{R}_{t+h}), \sum_{m=1}^{M} w_i^m(\alpha_j, w) ES^m(\hat{R}_{t+h})) \tag{21}$$

A vulnerability index

- Forecasts are used to generate signals of forthcoming increases in tail risks.
- A prediction exceeding a threshold determined by minimization of the sum of forecast errors provides a signal of future realizations of VaR violations.
- The binary model of the probability of a violation estimated with the available data up to the forecasting date t is a Logistic regression given by:

$$P(I(R_t)) = Logit\left(\sum_{h=0}^{12} a_h ES^*(\hat{R}_{t-h})\right) \tag{22}$$

- The prediction of Equation (22) is used to identify the threshold value $\hat{P}(I(R_t))$ corresponding to the minimization of a weighted sum of false alrqms and missed violations
- Thee vulnerability index is defined by:

$$VI(R_T) = \max\{0, \hat{P}(I(R_t)) - P^*(I(R_t))\} \tag{23}$$

Implementation

- See paper