Scenario Design for Macro-Financial Stress Testing*

Emanuele De Meo†

November 26, 2019

DRAFT PAPER

Abstract

The goal of this paper is to provide a possible approach to Scenario Design for selecting a stress scenario on economic growth, inflation and long-term interest rates in Italy. The Scenario Design framework belongs to the class of Second Generation Stress Tests and is composed of a few building blocks. First, multiple scenarios on the risk factors are generated simulating a Large Bayesian VAR for the Italian economy. Second, we take the perspective of a representative investor who aims to select a *severe yet plausible* scenario on the systematic risk factors follwing a factor investing strategy. Moreover, we compare the stress scenarios selected under two different approaches to measure plausibility: the Mahalanobis distance and Entropy pooling under three alternative subjective views with a clear economic narrative. We give evidence that our framework is suitable for the selection of a proper forward-looking *severe yet plausible* stress scenario.

Keywords: Scenario Analysis, Scenario Design, Large Bayesian VAR, Factor-mimicking portfolios, Entropy pooling.

JEL: C11, C53, G11

---

*The views expressed in this paper are mine and do not necessarily represent the opinions of my organization.

†Regulation and Economic Studies, UnipolSai Assicurazioni SpA; mailto: emanuele.demeo@unipolsai.it.
I am grateful to Renzo Avesani for encouraging me to develop the Scenario Analysis framework. I thank Giuseppe Cinquemani, Chiara Guerello and Andrea Tafuro for interactions on model specification, forecasting and Scenario Analysis. I am also grateful to Lea Zicchino, Rita Romeo, Paolo Barbieri and Federica Orsini for helpful discussions on an earlier specification of the Large Bayesian VAR for Italy. Any errors are mine.

1 Introduction

Scenario Analysis (from now on, SA) can be defined as the analysis of possible future states of the world. The Financial and Sovereign Debt Crisis has given an unprecedented boost to Stress Testing methodologies, leading to significant advances in the evaluation of the impact of adverse financial scenarios and tail risks (see Baudino et al [4]). Surprisingly enough, Scenario Design (SD), i.e. the selection criteria of the possibile scenarios to be evaluated in the SA, has deserved relatively limited attention in the applied literature.

The goal of this paper is to provide a possible operational framework in terms of Systematic Scenario Choice (SSC). The methodology outlined in this paper potentially applies to the analysis of both bad and good states of the world, providing a unified framework to SSC regarldess of the ultimate scope of SA.

In general terms, a Scenario is a possible realization for a set of random economic and/or financial variables, the Risk Factors (RFs). Picking a meaningful scenario is crucial for the evaluation of the behaviour of the object of interest - e.g. a portfolio of assets - in a particular state of the world. SD concerns the choice of an internally consistent scenario in two main steps (see Henry [13]).

The Risk Factors Selection. This stage concerns the identification of the RFs domain and the stylized representation of their materialization and interaction with less significant factors. At this step, SD requires a combined use of quantitative tools and judgement, providing a mapping of RFs to the object of SA and a narrative of the possibile propagation channels.

The Scenario Choice. This stage concerns the selection of a possibile future realization of the RFs. At this step the analysis is merely quantitative, providing a numerical path for (shocks to) the identified RFs. Any Scenario can be qualified in terms of three key characteristics (see Borio [6]):

- the plausibility, i.e. the likelihood that the scenario can actually occur in the future;
- the severity, i.e. the impact of the scenario on the object of interest;

- the horizon, i.e. the future time window in which the scenario must occur to be considered plausible and severe.

The strategy to address a plausible, severe and timely scenario should acknowledge the *trade-off* nature of SSC. An alternative scenario should therefore be sufficiently plausible, i.e. it must hold a significant probability to occur in a given future horizon, but also meaningful, i.e. its impact on the object of interest should be significantly different from what would be observed in the baseline scenario.

The distinction between First and Second Generation Stress Testing lies at the heart of modern SA (Breuer and Krenn [8]). A First Generation Stress Test evaluates a portfolio as a function of a set of the RFs under a limited number of scenarios, with no formal role for the distribution of the RFs and, as a consequence, without any quantitative assessment of the plausibility of the scenario. The most common example of First Generation Stress Test is the historical approach, i.e. the selection of a scenario already observed in history. Transparency makes this strategy appealing and popular among practictioners. The main advantage is that the plausibility requirement is satisfied: a scenario that already happened in the past can also be considered plausible in the future. Moreover, a scenario already observed in history is easier to implement and communicate. On the other hand, the principal drawback of First Generation Stress Tests is that they do not allow the evaluation of relative plausibility of different scenarios. Secondly, not all plausible scenarios have already occurred in the past. Third, the historical approach is subject to the identification and selection of specific stress episodes, which can be ambiguous[1].

In contrast, Second Generation Stress Tests consider a handful of scenarios which are sufficiently plausible. This approach has also pros and cons. The main advantage of this approach is the possibility to evaluate the plausibility of forward-looking scenarios. This requires a probabilistic modelling of the risk factors distribution: different risk factors distribution lead to different plausibility evaluations of a given scenario.

Even though the objective of the experiment largely determines the approach (see Borio et al. [6]), there are generally two ways to deal with SSC (see Cihak [10]):

[1]The Recession Approach to SD by the Federal Reserve to regulatory Stress Testing is one of the practical implementation of the historical approach.

1. the worst (best)-case approach: for a given level of plausibility, the most (least) severe scenario is selected, i.e. the scenario with the largest impact on the object of interest;
2. the threshold approach: for a given severity, we select the most likely combination of shocks that are needed to have that impact on the object of interest.

The goal of this paper is to provide a possible approach to SD. In particular, we will take the case of selecting a stress scenario for economic growth, inflation and long-term interest rates in Italy with a worst-case approach. The SD framework belongs to the family of Second Generation Stress Tests and is composed of a few building blocks. First, multiple scenarios on the risk factors are generated simulating a Large Bayesian VAR for the Italian economy, estimated as in Banbura et al. [2]. Large Bayesian VARs present several advantages for scenario generation purposes, i.e. they are able to generate forecast density paths in a natural way, they are well suited for large information sets and they have been proved to provide superior out-of-sample forecasting performances, ensuring that SD does not overvalue the plausibility of the scenario.

Second, we will take the general perspective of a representative investor who aims to select a *severe yet plausible* scenario on the systematic risk factors. As in Breuer et al. [7], severity is a portfolio-specific concept, acknowledging that a stress scenario could be harmful for a portfolio but relatively safe for another. Since in this paper the goal is to select a stress scenario on a given set of Macro-Financial risk factors, it is reasonable to assume that the Scenario must be severe for a representative investor who is exposed to those risk-factors.

More precisely, we assume that the investor adopts a factor investing strategy, allocating wealth in the three diversified portfolios mostly correlated with the systematic risk factors. This portfolio allocation strategy maps portfolio losses to adverse scenarios through the channel of portfolios mimicking the dynamics of Macro-Financial risk factors. Reverse engineering, i.e. selecting the worst-case scenario in terms of investor's welfare, is the solution to the SD problem.

Moreover, we compare the stress scenarios selected under two different plausibility measures: the Mahalanobis distance proposed by Breuer et al. [7], which expresses plausibility as an inverse function of the distance of a scenario from the first moment, and posterior probabilities obtained with the Entropy pooling approach by Meucci [14], which takes subjective views on scenarios into account.

In particular, we compare selected scenarios under three possible views. The first view is a view on lower-for-long 3-month Euribor, which is expected to be lower than or equal to current (negative) values minus 10 basis points in the forecast horizon. This view is consistent with a low inflation expectations environment, which require a highly accomodative monetary policy stance by the ECB. In the second view, expectations of an inverted US yield curve in the forecast horizon can be interpreted as a view on a future US recession or economic downturn. Finally, in the third view the Italian stock market is expected under stress, i.e. market returns are expected to be lower or equal to the forward-looking Italian stock market Conditional Value-at-Risk (CVaR).

To anticipate some of the results, the main findings can be summarized as follows. Under the lower-for-long view, the Frequentist approach based on the Mahalanobis distance is robust in terms of scenario selection with respect to Entropy pooling views, both with full and partial confidence. The stress scenario is consistent with a low-inflation or deflationary environment, with a deep and prolonged recession and higher long-term bond yields.

The view on an inverted US yield curve implies a proper adverse scenario only with a 50% confidence. The scenario embedding a view on a US downturn appears in any case less severe than the one selected under the Frequentist approach, which is consistent with a prolonged low-inflation environment in Italy. Real GDP drops to a lower extent compared to the Frequentist scenario, showing a gradual recovery towards a no-growth situation in the second half of the forecast horizon.

Finally, the view on a bearish Italian stock market has a similar narrative under both full and partial confidence, albeit with some notably differences. First, with full confidence in the view, Real GDP decline is more gradual and prolonged with respect to the partial confidence case, in which GDP immediately drops by 2% on a yearly basis and recovers afterwards. Second, long-term bond yields experience a steep increase in the full-confidence case, with 10-year interest rate rising above 6% in few quarters, while in the partial confidence view bond risk premia hike is less dramatic. Since the scenarios are generated with a Large Bayesian VAR for Italy, spillover and second-round effects of a systemic event in Italy could have been underestimated.

The paper is structured as follows. Section 2 contains a general presentation of the Scenario Design methodology. In Section 3 we present the empirical application of our framework to the selection of a stress scenario for the Italian economy. In Section 4 we present and discuss the results of the analysis. Section 5 contains the conclusions and proposals for future developments. Finally, in the Appendix we give full technical disclosure on the econometric tools.

2 The methodology

The goal of SA in a Second-Generation Stress Tests is to optimally select a Scenario on systematic risk factors in a multiple scenario framework. The optimal Scenario should be sufficiently *severe*, i.e. harmful enough to be considered a stress scenario, yet *plausible*, i.e. holding a non-negligible probability to occur in the forecast horizon (see Breuer et al. [7]).

In what follows, we will take the general perspective of a representative investor who aims to select a *severe yet plausible* Scenario on a set of Macro-Financial risk factors. As in Breuer et al. [7], severity is a portfolio-specific concept, acknowledging that a stress scenario could be harmful for a portfolio but relatively safe for another. Since in this paper the goal is to select a stress scenario on a given set of Macro-Financial risk factors, it is reasonable to assume that the Scenario must be severe for a representative investor who is exposed to those risk-factors.

As in Breuer et al. [7], I also adopt a *partial scenario* perspective. While it is recognized that a portfolio may depend, directly or indirectly, on a broad set of risk factors, the modeler can be interested in stressing only a few factors. Formally, let y be the set of risk factors which can be partitioned in two subsets of systematic and non-systematic risk factors:

$$y := \{\phi, \theta\}$$

where ϕ is the set of systematic risk factors and θ is the set of other risk factors. In the empirical application I will assume that ϕ includes the following risk factors:

$$\phi := \{\gamma, \pi, i\}$$

where γ represents economic growth, π denotes inflation rates, and i is the long-term (10-year) government interest rate for Italy.

SA involves selecting a *severe yet plausible* scenario with respect to the forecast path of ϕ in the joint distribution of all factors. In principle, the scenario on risk factors θ could also be chosen from their conditional density, *given that* risk factors ϕ follow a stress scenario. This would involve a causality structure in the scenario generation process from the set of stressed systematic risk factors to other risk factors. In many cases this could represent an advantage, for example when the causality structure of the stress test experiment is unambiguous or when the number of systematic risk factors is large enough to require a convenient partition of the factors that need to be stressed. However, in the empirical application we will not follow this strategy, as the number of systematic risk factors is intentionally kept sufficiently low for the sake of exposition.

Multiple scenarios on y are generated simulating a Large Bayesian VAR for the Italian economy, estimated as in Banbura et al. [2]. Large Bayesian VARs present several advantages for scenario generation purposes:

- they are able to generate forecast density paths in a natural way, as Bayesian models explicitly take uncertainty into account;
- large Bayesian VARs are also well suited for large information sets, generating forecast densities on systematic and other risk factors in a coherent setting, handling correlation structures without incurring in a *curse of dimensionality* issue;
- they provide superior out-of-sample forecasting performances compared to other econometric tools, like frequentist VARs, factor models or structural models; this ensures that SA does not overvalue plausibility, i.e. selecting severe stress scenarios in a density forecast which is biased by a poor out-of-sample predicting power; this would lead to *ex-ante apparently plausible* stress scenarios which are *ex-post hardly plausible.*

Let observations on all risk factors y at time t be denoted with $y_t = \left[y'_{1,t}\ y'_{2,t}\ ...\ y'_{M,t}\right]' = \left[\phi'_t\ \theta'_t\right]'$. A scenario s on y_t over the forecast horizon $t+1, ..., t+h$ is defined as

$$y^s_{t,h} = \left[y^s_{t+1}\ y^s_{t+2}\ ...\ y^s_{t+h}\right]'$$

and the cube of all possible scenarios as

$$\boldsymbol{y_{t,h}} = \left[y^1_{t,h}\ y^2_{t,h}\ ...\ y^S_{t,h}\right]$$

where S is a possibly large number of scenarios. As mentioned earlier, $\boldsymbol{y_{t,h}}$ is generated by the Large Bayesian VAR[2].

For every scenario s, the matrix $y^s_{t,h}$ is mapped into the vector $\widetilde{y}^s_{t,h}$ by a vector of weights $\omega = [\omega_1\ \omega_2\ ...\ \omega_h]$. The element $\omega_\tau, \tau = 1, ..., h$ synthesizes the relative importance of period τ in the evaluation of the plausibility and severity of the scenario[3]. Formally:

$$\widetilde{y}^s_{t,h} = \omega y^s_{t,h}.$$

The cube of all possible $\widetilde{y}^s_{t,h}$ is accordingly denoted with $\boldsymbol{\widetilde{y}_{t,h}}$.

Given a portfolio $P = [w_{p,1}\ w_{p,2}\ ...\ w_{p,N}]$ of weights over N assets, with $\sum_i w_{p,i} = 1$, the investor is assumed to have preferences defined over the value of her portfolio according to an exponential loss utility function penalizing more heavily extreme portfolio losses:

$$U(P) = -\exp\left[-\left(P - E(P)\right)\right]. \tag{1}$$

We assume that the investor adopts a factor investing strategy, allocating wealth in the three diversified portfolios mostly correlated with the systematic risk factors ϕ. More precisely, we assume that portfolio choice involves allocation on sectoral stock indexes with the highest β to factor mimicking portfolios $f = \{f_\gamma, f_\pi, f_i\}$ that replicate systematic risk factors ϕ[4]. This portfolio allocation strategy maps portfolio losses to adverse scenarios through the channel of portfolios mimicking the dynamics of Macro-Financial risk factors (see Balduzzi and Robotti [3] for an overview on mimicking portfolios). Reverse engineering, i.e. selecting

[2]See the Appendix for details on the specification of the Large Bayesian VAR.

[3]A similar parametrization of the scenario over the forecast horizon is the one proposed by Mokinski [15], who however follows a threshold approach and a different specification of the severity function.

[4]Equivalent results derive from the allocation over an investment universe represented by all sectoral indexes. Therefore, we opted for a lower degree of diversification for the sake of exposition.

the worst-case scenario in terms of investor's welfare, is the solution to the SD problem[5].

Formally, define the ellipsoid Λ_α of plausible scenarios at the confidence level α as:

$$\Lambda_\alpha := \left\{ \widetilde{y}^s_{t,h} \in \boldsymbol{\widetilde{y}_{t,h}} : \Pi(\widetilde{y}^s_{t,h}) \geq \alpha \right\} \tag{2}$$

where $\Pi(.)$ is the plausibility of the scenario. The scenario is selected solving the following optimization problem:

$$\begin{aligned} \widetilde{y}^*_{t,h} &= \arg\min U[P(\widetilde{y}^s_{t,h})] \\ s.t.\ \widetilde{y}^s_{t,h} &\in \Lambda_\alpha \end{aligned}$$

In the following sections, we will explore two alternative measures of plausibility under which the ellipsoid of plausible scenarios is defined. The first measure of plausibility, originally proposed by Breuer et al. [7] for SA purposes, is based on the Mahalanobis distance, which reflects a frequentist approach. The second measure of plausibility is based on posterior probabilities obtained by the Entropy Pooling approach by Meucci [14], which reflects a subjective approach to plausibility measurement.

2.1 Frequentist plausibility: the Mahalanobis distance

An intuitive approach to the measurement of plausibility is to compare the extreme realization of a risk factor with its expected value: the further away from the expected value, the less plausible the scenario can be considered.

The Mahalanobis distance is the statistical translation of this intuition (Breuer et al. [7]). It is defined as the distance of a given scenario $\widetilde{y}^s_{t,h}$ from the expected scenario μ divided by the width of the ellipsoid in the direction of the test point. Intuitively, the Mahlanobis distance can be interpreted as the number of standard deviations of the multivariate move from μ to $\widetilde{y}^s_{t,h}$. In symbols:

$$Maha(\widetilde{y}^s_{t,h}) := \sqrt{(\widetilde{y}^s_{t,h} - \mu)' \Sigma^{-1} (\widetilde{y}^s_{t,h} - \mu)}$$

where $\mu = E(\boldsymbol{\widetilde{y}_{t,h}})$ and $\Sigma^{-1} = Cov(\boldsymbol{\widetilde{y}_{t,h}})$. In this case, plausibility can be defined as the inverse of the Mahalanobis distance, i.e.:

[5] See the appendix on Portfolio Construction and Allocation for details on the methodology.

$$\underline{p} = \left[\frac{1}{Maha(\widetilde{y}^1_{t,h})} \; \frac{1}{Maha(\widetilde{y}^2_{t,h})} \cdots \frac{1}{Maha(\widetilde{y}^S_{t,h})} \right]$$

and normalized as:

$$p = [p_1 \; p_2 \; ... \; p_S] \,, \; p_s = \frac{\underline{p}_s}{\sum_{j=1}^{S} \underline{p}_j}, \; s = 1, ..., S.$$

2.2 Embedding Subjective Views: Entropy pooling

The second option to measure plausibility is based on the Entropy Pooling approach. In this case, given a vector of prior probabilities p, plausibility is defined as the vector of joint posterior probability distribution $\tilde{p}$ solving the following optimization problem:

$$\tilde{p} = \arg \min_{\substack{Fx \leq f \\ Hx \equiv h}} \mathscr{E}(f, p)$$

where $\mathscr{E}(f, p)$ is defined as the relative entropy between a generic distribution f and the reference distribution p and subjective views are expressed in the form of equality and inequality constraints, $Fx \leq f, Hx \equiv h$. More precisely, the relative entropy $\mathscr{E}(f, p)$ is equal to:

$$\mathscr{E}(f, p) = \sum_{s=1}^{S} f_s [\log f_s - \log p_s]$$

where S is the number of scenarios[6].

The opinion-pooling, confidence-weighted posterior probabilities p_c are then obtained as a weighted average of the prior and the posterior distributions, with weights depending on the confidence on subjective views:

$$p_c = c\tilde{p} + (1 - c)p.$$

In most applications $p = 1/S$, i.e. the prior is non-informative. In this paper we will explore the case of prior probabilities derived from the Mahalanobis distance.

A few aspects of the Entropy Pooling approach are worth mentioning.

[6] See Meucci [14] for the numerical solution to this optimization problem.

- Equality and inequality views on the risk factors can be specified in the most general form as: expectations (absolute and relative); volatilities; correlations; behaviour in the tail of the distribution; tail co-dependence; etc. This flexibility expands the options in terms of SA with respect to standard Conditional Forecast experiments, in which views are generally specified in terms of expectations of one or more variable over the forecast horizon.

- The opinion-pooling approach is particularly appealing in a multi-manager environment, in which managers have different opinions about the future realization of one or more risk factors.

3 Empirical application: Designing Stress Scenarios for Italy

In this section we will apply our SA framework to the selection of a stress scenario for Italy. The experiment is conducted under a few assumptions that are outlined below.

First, we assume that the representative investor is exposed to three portfolios mimicking the dynamics local systematic risk factors, namely Italian economic growth, inflation and long-term yields. The exposure to the systematic risk factors is obtained through an investable universe represented by Italian sectoral stock indexes. This assumption should be general enough to design a Macro-Financial stress scenario for Italy. Scenarios on the other risk factors can then be obtained with the same Large Bayesian VAR or alternatively with a structural macro-econometric model, conditionally to the stress scenario on the systematic risk factors.

Second, we assume that the exposure to the mimicking portfolios is decided by the investor solving a portfolio optimization problem. This is a convenient assumption for the goal of this paper but can be relaxed in a real-world experiment, using weights known *ex-ante* to the investor (e.g. from her balance sheet exposures).

Third, we assume that the representative investor does not rebalance her portfolio when the stress scenario materializes. This is restrictive in a real-world experiment, but it is convenient for the scope of the paper. Indeed, in the case of portfolio rebalancing, the investor could change weights in order to mitigate the severity of the scenario. In that case, lower scenario severity is the result of optimal portfolio allocation and not of an ineffective SD methodology.

Fourth, we assume that the investor equally weights each quarter over the forecast horizon. Assuming that each period in the future has the same weight is unrealistic, as in many cases it appears reasonable to give more weights to the near future. However, the assumption is sufficiently agnostic to allow an analysis of the results under very general assumptions about the intertemporal preferences of the representative investor or firm in charge of designing the scenario. In any case, the choice of a given weighting scheme is dependent on the specific goals of the SA and any other weighting schemes is applicable.

Fifth, the entropy pooling approach is conducted both with a full confidence and with a 50% confidence on subjective views. Again, this assumption is convenient to evaluate the impact of different confidence degrees on the posterior probabilities of the scenarios.

Finally, the plausibility domain is evaluated at the 10% confidence degree, i.e. the 10% least plausible scenarios are not included in the plausibility ellipse.

The scenario is selected in the joint density forecasts of the endogenous variables of the Large Bayesian for Italy, conditional to economic policy expectations. A pre-determined path for 3-month Euribor interest rate (a proxy for ECB monetary policy) and for the Italian primary balance as a percentage of Real GDP (a proxy for fiscal policy) is embedded in the Conditional Forecast experiment. For Conditional Forecasting, we adopt the Waggoner and Zha [16] algorithm[7].

4 Results

The SD framework relies on a portfolio construction and allocation strategy on Italian sectoral indexes, which present some degree of exposure to the Macro-Financial risk factors of interest. The results of portfolio construction and allocation are the following.

[7]We follow the procedure on hard conditioning described by Dieppe et al. [11] in the BEAR Toolbox published by the ECB. See the Appendix for details.

First, in Figure 1 factor-mimicking portfolios are plotted against their respective underlying Macro-Financial factors. Portfolios present a sufficient ability to mimick the dynamics of Macro-Financial factors, in particular during specific periods (e.g. the Sovereign Debt Crisis), with the growth-mimicking portfolio showing the highest correlation with the underlying factor. Second, Sectoral stock indexes are heterogeneously correlated to the mimicking portfolios (Figure 2). Concerning the growth portfolio, Italian sectors show a β ranging from 0.05 to 0.35. In the case of inflation and long-term interest rate portfolios, the variance is larger. In the case of inflation, the sign of the exposure is different across sectors. In the case of bond yields, the βs range from 0.2 to 0.6. From this evidence, we can draw the conclusion that our SD framework will likely present a higher ability to select a severe scenario on economic growth, while SD for inflation appears more problematic and subject to instability.

In Figure 3 worst-case scenarios at the 5% of probability for each Macro-financial risk factors are plotted[8]. In terms of the probability mass associated with the worst-case scenarios, 5%-tail realizations present a joint plausibility around 4%. In other words, scenarios that are *marginally adverse* tend also to be *jointly less plausible.*

In this section we will compare the outcome of SA experiments under three different approaches:

- the Systematic Scenario Choice under the Frequentist approach, in which the plausibility of a scenario is equal to the inverse of its Mahalanobis distance from the central scenario;
- the Entropy-pooling approach with full confidence in a given subjective view, with the inverse of Mahalanobis distances as prior probabilities of the scenarios;
- the Entropy-pooling approach with a 50% confidence in the subjective view.

As mentioned in the previous section, in the Entropy-pooling methodology subjective views can be specified in a very general form. Here we will consider three alternative subjective views, described by a specific economic rationale.

[8] A worst-case scenario for economic growth is selected in the left tail of its distribution, whereas for inflation and 10-year interest rate worst-case scenarios are selected in the right tail.

Lower for long. This view is specified in terms of expectations on interbank interest rates. In particular, the 3-month Euribor is expected to be lower than or equal to current (negative) values minus 10 basis points in the forecast horizon. In other words, the view is of a further ECB policy interest rate (i.e. deposit rate) cut by 10 basis points. This view is consistent with a low inflation expectations environment, which require a highly accomodative monetary policy stance by the ECB. Moreover, this can be interpreted as a stress scenario for the banking sector, as negative interest rates are harmful for banks' profitability.

Inverted US yield curve. In the empirical literature, the slope of the yield curve has been proved to be an accurate leading indicator of economic activity, in particular in the US (Bauer and Mertens [5]). Therefore, a view on the spread between long-term and short-term US government interest rates can be interpreted as a view on the state of US business cycle and, to a larger extent, of the world economy. In particular, an inverted yield curve - a negative term spread - is often followed by a recession. We specificy a subjective view on a future US recession expecting a spread between 10-year and 3-month US government bond yields to be lower or equal to zero.

Bearish Italian stock market. A view on the Italian stock market under stress is specified as a view that average Italian stock market returns will be lower than or equal to the forward-looking Conditional Value-at-Risk (CVaR, or Expected Shortfall) at the 99.9% confidence level. Formally, the view is specified as:

$$\widetilde{r^s}_{m,t,h} \leq E(\widetilde{r^\iota}_{m,t,h} | \widetilde{r^\iota}_{m,t,h} \leq VaR(\widetilde{r^s}_{m,t,h}, 0.001)$$

where $r^s_{m,t,h}$ is the stock market return over the forecast horizon and $VaR(.)$ denotes the Value-at-Risk, i.e. the 0.1% quantile of returns. This view is consistent with a systemic stress narrative in Italy.

4.1 Lower for long 3-month Euribor

Under the view of lower interest rates for a long period, the selected stress scenario is the same across the three approaches. The scenario is consistent with a low-inflation or deflationary environment. Real GDP drops immediately, entering a deep and prolonged recession. The downturn of the business cycle is quantitatively comparable with the contractionary path experienced during the Sovereign Debt Crisis (-4% on a yearly basis at trough), but more prolonged compared to the previous recession episodes (Figure 4). Consumer prices grow at a very slow pace, lower than 1% on a yearly basis, and long term bond yields rapidly increase above 4%, a value experienced during the Sovereign Debt Crisis.

Overall, the Frequentist SD is robust to the specification of a subjective view on persistently low (negative) interest rates. In other words, and unsurprisingly, this subjective view does not change the relative plausibility ranking of different scenarios, as they already discount an accomodative monetary policy environment (recall that the joint density forecasts are obtained as Conditional Forecasts with respect to a predetermined path of 3-month Euribor, which is expected to be negative for the entire forecast horizon). Interestingly enough, the SD framework outlined in this paper allows the selection of a stress scenario that should be considered hardly plausible on the basis of recent recession episodes, at least in terms of the length of Real GDP recession. In other words, the scenario would not have been selected under a historical approach.

It is worth mentioning that the same scenario is selected under two alternative subjective views: a view on high long-term Italian government bond yields and a view on high risk aversion (as proxied by the Credit Spread for US Corporates). In other words, with low inflation risks, an increasingly accomodative monetary policy view is equivalent to a view on high long-term bond yields driven by an increase of investors' risk aversion.

4.2 Inverted US yield curve

As mentioned earlier, a view on inverted US yield curve can be interpreted as a subjective view on an economic downturn in the US. Both Entropy pooling approaches, with full and 50% confidence, return a different scenario from the Frequentist approach, but only partial confidence allows the selection of a proper stress scenario (Figure 5).

Indeed, with full confidence in the view, Real GDP grows as in the central scenario and inflation rises above the monetary policy target (close to but below 2%). On the other hand, Italian long term interest rates experience a volatile path, falling to very low values in the first quarters and returning close to the median path afterwards.

Alternatively, under a 50% confidence, a proper recession scenario with higher interest rates is selected. Not surprisingly, the scenario embedding a view on a US downturn appears in any case less severe than the one selected under the Frequentist approach, which, as mentioned above, is consistent with a persistently expansionary ECB monetary policy driven by a low inflation and/or high bond risk premia environment. Real GDP drops to a lower extent, showing a gradual recovery towards no-growth in the second half of the forecast horizon. Italian long-term interest rates rapidly increase above 4%, while the path for inflation remains somewhat surprisingly above the monetary policy target. To overcome this counterintuitive behaviour of consumer prices, a stress scenario on inflation could be obtained in a Conditional Forecast framework, simulating the path of inflation *conditional to* the stress scenario on Real GDP and long-term interest rates.

4.3 Bearish Italian Stock Market

The narrative of the scenario under both Entropy pooling approaches is qualitatively similar, albeit with some notably differences. First, with full confidence in the view, Real GDP decline is more gradual and prolonged with respect to the partial confidence case, in which GDP immediately drops by 2% on a yearly basis and recovers afterwards. Second, long-term bond yields experience a steep increase in the full-confidence case, with 10-year interest rate rising above 6% in few quarters, while in the partial confidence view bond risk premia hike is less dramatic. Finally, and most importantly, in the full-confidence case the dynamics of inflation is hardly plausible. While higher inflation is a possible outcome in a stress episode, given the systemic relevance of Italy and possible effects in terms of nominal Euro exchange rates depreciation with respect to non-Euro currencies, in a low inflation environment a rapid increase above the monetary policy target is hardly plausible. In contrast, a 50% confidence allows the selection of a stress scenario lying midway between the Frequentist and the Full confidence scenarios and characterized by a recessive economic activity, high bond risk premia and low inflation.

It is worth mentioning that scenarios are generated with a large Bayesian VAR for the Italian economy, with a limited role for international Macro-Financial factors. Therefore, the spillover effects of a financial shock in Italy to the rest of the World Economy, and in particular to other Euro countries, is not completely taken into account. Hence, we can expect that the effects under the subjective view are updward biased. In other words, if the shock would have been modeled in a proper global setting, spillover effects could have likely been larger and second-round effects on Italian Real GDP more pronounced.

5 Conclusions

In this paper we presented a possible framework for Macro-Financial Scenario Design. Our approach belongs to the class of Second Generation Stress Tests, i.e. the scenario is selected in the set of plausible realizations in order to maximize a portfolio-specific severity function. Moreover, we allow the scenario designer to have subjective views on the future with a certain confidence. We apply our framework to the selection of a stress scenario for the Italian Economy and compare the results obtained under the Frequentist and the Entropy pooling approach with different subjective views and confidence levels. We give evidence that our framework is suitable for the selection of a proper forward-looking *severe yet plausible* stress scenario.

A few possible improvements can be applied to the Scenario Design framework presented in the paper. First, scenarios might be generated in an integrated international setting, properly modelling spillover and second-round effects of adverse scenarios. Second, more sophisticated identification schemes might help to identify underlying structural shocks in the forecast densities generated by Conditional Forecasting. For example, Sign restrictions (see for example Fry and Pagan [12]) or Narrative Sign Restrictions (Antolìn-Diaz et al. [1]) might be suitable identification schemes.

We leave all these improvements for possible future research.

References

[1] Antolín-Díaz, J., and Rubio-Ramírez, J. F. Narrative sign restrictions for svars. *American Economic Review 108*, 10 (October 2018), 2802–29.

[2] Bańbura, M., Giannone, D., and Reichlin, L. Large bayesian vector auto regressions. *Journal of Applied Econometrics 25*, 1 (2010), 71–92.

[3] Balduzzi, P., and Robotti, C. Mimicking Portfolios, Economic Risk Premia, and Tests of Multi-Beta Models. *Journal of Business & Economic Statistics 26* (2008), 354–368.

[4] Baudino, P., Goetschmann, R., Henry, J., Taniguchi, K., and Zhu, W. Stress-testing banks – a comparative analysis. Financial Stability Institute Insights 12, Bank for International Settlements, Nov. 2018.

[5] Bauer, M. D., and Mertens, T. M. Economic Forecasts with the Yield Curve. *FRBSF Economic Letter* (2018).

[6] Borio, C., Drehmann, M., and Tsatsaronis, K. Stress-testing macro stress testing: Does it live up to expectations? *Journal of Financial Stability 12*, C (2014), 3–15.

[7] Breuer, T., Jandacka, M., Rheinberger, K., and Summer, M. How to Find Plausible, Severe and Useful Stress Scenarios. *International Journal of Central Banking 5*, 3 (September 2009), 205–224.

[8] Breuer, T., and Summer, M. Solvency stress testing of banks: Current practice and novel options. Tech. rep., Sveriges Riksbank and Finansinspektionen, 2017.

[9] Chan, J. C. C. Large bayesian vars: A flexible kronecker error covariance structure. *Journal of Business & Economic Statistics 0*, 0 (2018), 1–12.

[10] Cihak, M. Introduction to Applied Stress Testing. IMF Working Papers 07/59, International Monetary Fund, Mar. 2007.

[11] Dieppe, A., van Roye, B., and Legrand, R. The BEAR toolbox. Working Paper Series 1934, European Central Bank, July 2016.

[12] Fry, R., and Pagan, A. Sign restrictions in structural vector autoregressions: A critical review. *Journal of Economic Literature 49*, 4 (December 2011), 938–60.

[13] Henry, J. Macrofinancial scenarios for system-wide stress tests: Process and challenges. In *Europe's New Supervisory Toolkit* (2015), M. Quagliariello, Ed., Risk Books.

[14] Meucci, A. Fully flexible views: Theory and practice, 2010.

[15] Mokinski, F. A severity function approach to scenario selection. Discussion Papers 34/2017, Deutsche Bundesbank, 2017.

[16] Waggoner, D., and Zha, T. Conditional forecasts in dynamic multivariate models. *The Review of Economics and Statistics 81*, 4 (1999), 639–651.

6 Appendix

6.1 Scenario Generation

The Macro-Financial model used for scenario generation is a Vector-Autoregression (VAR) specified as

$$y_t = A_0 + A_1 y_{t-1} + ... + A_p y_{t-p} + e_t, e_t \sim N(0, \Sigma)$$

and estimated in a Large Bayesian VAR framework (see Banbura et al. [2]), by shrinking the coefficients with a Minnesota-type prior distribution, which is equivalent to shrinking the dynamics of the system towards a random walk for integrated variables or a white noise for stationary variables. Formally:

$$E[(A_k)_{ij}] = \begin{cases} \delta_i \text{ if } j = i, k = 1 \\ 0 \text{ otherwise} \end{cases}$$

and

$$Var[(A_k)_{ij}] = \begin{cases} \lambda^2/k^2 \text{ if } j = i \\ \frac{\lambda^2/k^2}{\sigma_i^2/\sigma_j^2} \text{ otherwise.} \end{cases}$$

Two features of the VAR are worth mentioning. First, density forecasts are homoscedastic Gaussian. This assumption, although admittedly restrictive, is convenient to reduce the computational burden of the SD experiment [9].

Second, yhe hyperparameter λ captures the tightness of the prior: lower values of λ imply a more precise prior. The posterior distribution of the parameters are derived implementing the prior the dummy observations as in Banbura et al. [2]. The prior is calibrated out-of-sample minimizing the Relative Mean Squared Forecast Error on a given set of variables (the variables on which SD should be performed), i.e. the ratio between the Mean Squared Forecast Error of the Large Bayesian VAR and the Mean Squared Forecast Error of a benchmark model (a random walk with a drift) on a given forecast horizon H:

[9] See for example Chan [9] for Large Bayesian VARs with non-Gaussian and heteroscedastic error terms.

$$\lambda^* = \arg\min RMSFE_{\phi,H}^{\lambda,b} = \frac{MSFE_{\phi,H}^{\lambda}}{MSFE_{\phi,H}^{\lambda,b}} \quad (5)$$

The forecast density is simulated conditional to a predetermined path for the 3-month Euribor and the primary balance for Italy, in line with expectations on monetary and fiscal policy. The Conditional Forecast algorithm is due to Waggoner and Zha [16], who derive a Gibbs sampling algorithm under the hypothesis of normally-distributed error terms. We follow the algorithm as outlined in Dieppe et al. [?]. To obtain the forecast density, we also identified the VAR with a Choleski recursive structure, assuming that market variables react instantaneously to shocks to real (slow moving) variables, while the opposite is not true.

The model is estimated on quarterly data from 1980Q1 to 2019Q2 on the following 21 endogenous variables:

- US Corporate Credit Spread, as a proxy for global risk aversion;
- 3-month and US 10-year government interest rates;
- global stock market returns;
- 3-month Euribor and German 10-year government interest rates;
- world year-on-year Real GDP growth;
- global year-on-year import growth;
- Italian year-on-year employment growth;
- Italian unemployment rate;
- Italian year-on-year wage earnings growth;
- Italian year-on-year disposable income growth;
- Italian primary balance as a percentage of nominal GDP;
- Italian year-on-year Real GDP growth;
- Italian year-on-year Real consumtpion growth;
- Italian Consumer price core inflation;

- Italian Producer price inflation;
- Italian year-on-year House prices growth;
- Italian 10-year government interest rates;
- Italian capacity utilization rate;
- Italian stock market quarterly returns.

.

Full details on the model are available upon request. The datasource is Haver Analytics as provided by Oxford Economics.

6.2 Portfolio Construction and Allocation

The representative investor is assumed to allocate wealth in the three investable instruments mostly exposed to the mimicking-portfolios replicating the dynamics of the Macro-Financial factors of interest, i.e. economic growth, inflation and long-term interest rates. The factor-mimicking portfolios are constructed following a 2-stage regression procedure (see for example Balduzzi and Robotti [3] for an overview on mimicking portfolios).

In the first stage, we estimate a time-series regression of stock returns on the Macro-Financial factors:

$$r_{it} = \alpha_i + \beta_i \phi_t + \epsilon_{it} \tag{6}$$

where $\phi_t' = [\Delta\gamma_t \ \Delta\pi_t \ \Delta i_t]$. The estimated $\widehat{\beta}$s summarize the correlation of each stock returns to the Macro-Financial factors. In the second stage, we solve the following minimum portfolio variance optimization problem:

$$\begin{aligned} P^* &= \arg\min \frac{1}{2} P' D P \\ s.t.\ P'\beta &= 1 \end{aligned}$$

where $D = diag(\Sigma)$ and Σ is the variance-covariance matrix of the time-series regression residuals.

The optimization problem has the following solution, which is equivalent to the outcome of a Generalized Least Squares cross-sectional regression:

$$P^* = \left(\beta' D^{-1} \beta\right)^{-1} \beta' D^{-1}$$

which describes the factor-mimicking portfolios of factors ϕ.

Finally, wealth is optimally allocated among the three Sectoral Indexes mostly correlated with the factor-mimicking portfolios. As a portfolio optimization tool we adopted the Portfolio Analytics package in R in a mean-variance framework, with a diversification constraint that avoids concentration over only two of the three assets (weights are constrained to be lower or equal to 0.4).

The datasource for quarterly stock returns from 1980Q1 to 2019Q2 is Thomson Reuters.

Factor-mimicking portfolios and Macro-Financial factors

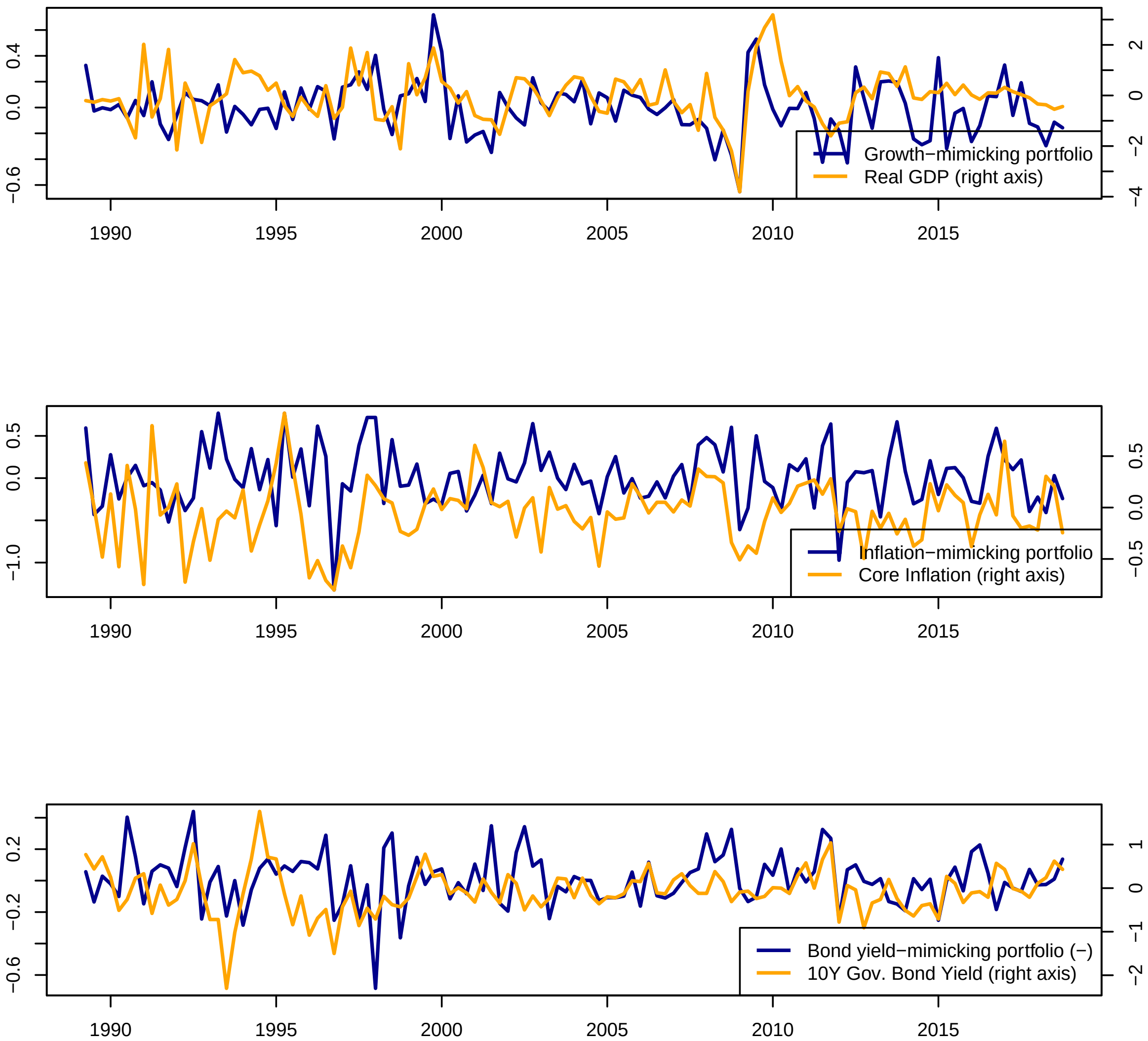


Figure 1: Factor mimicking portfolios values are expressed as quarterly % returns. Macroeconomic factors are in first differences; YoY growth rates for Real GDP and Core inflation, in % for 10-year Government Bond Yield. The Bond yield mimicking portfolio is expressed in reciprocal terms.

Sectoral Stock Index Betas to Factor-mimicking portfolios

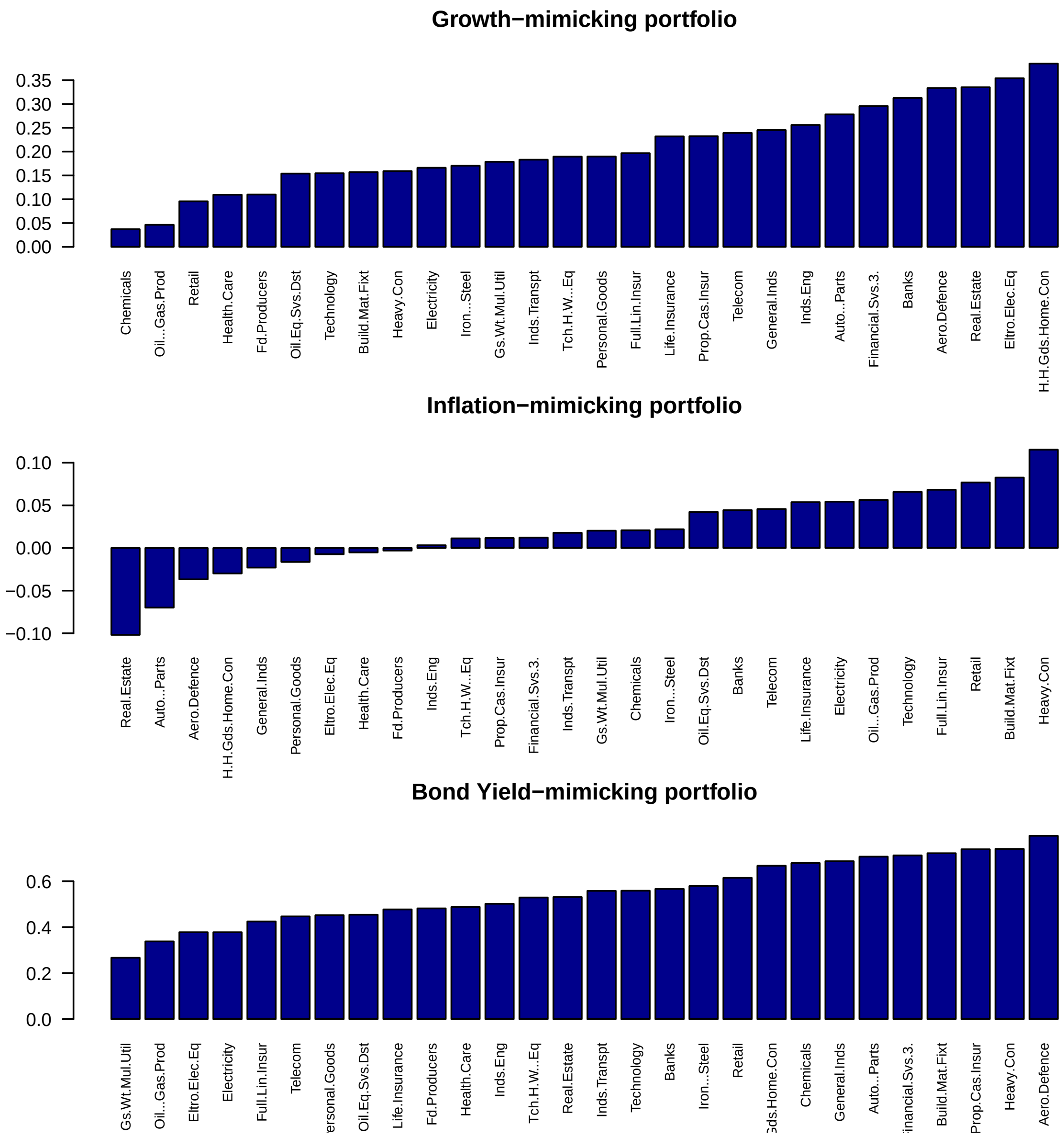


Figure 2: Betas are derived from the regression of Sectoral Stock Index returns on Factor-mimicking portfolios returns.

Worst-case Scenarios for Italian Macro-Financial Variables

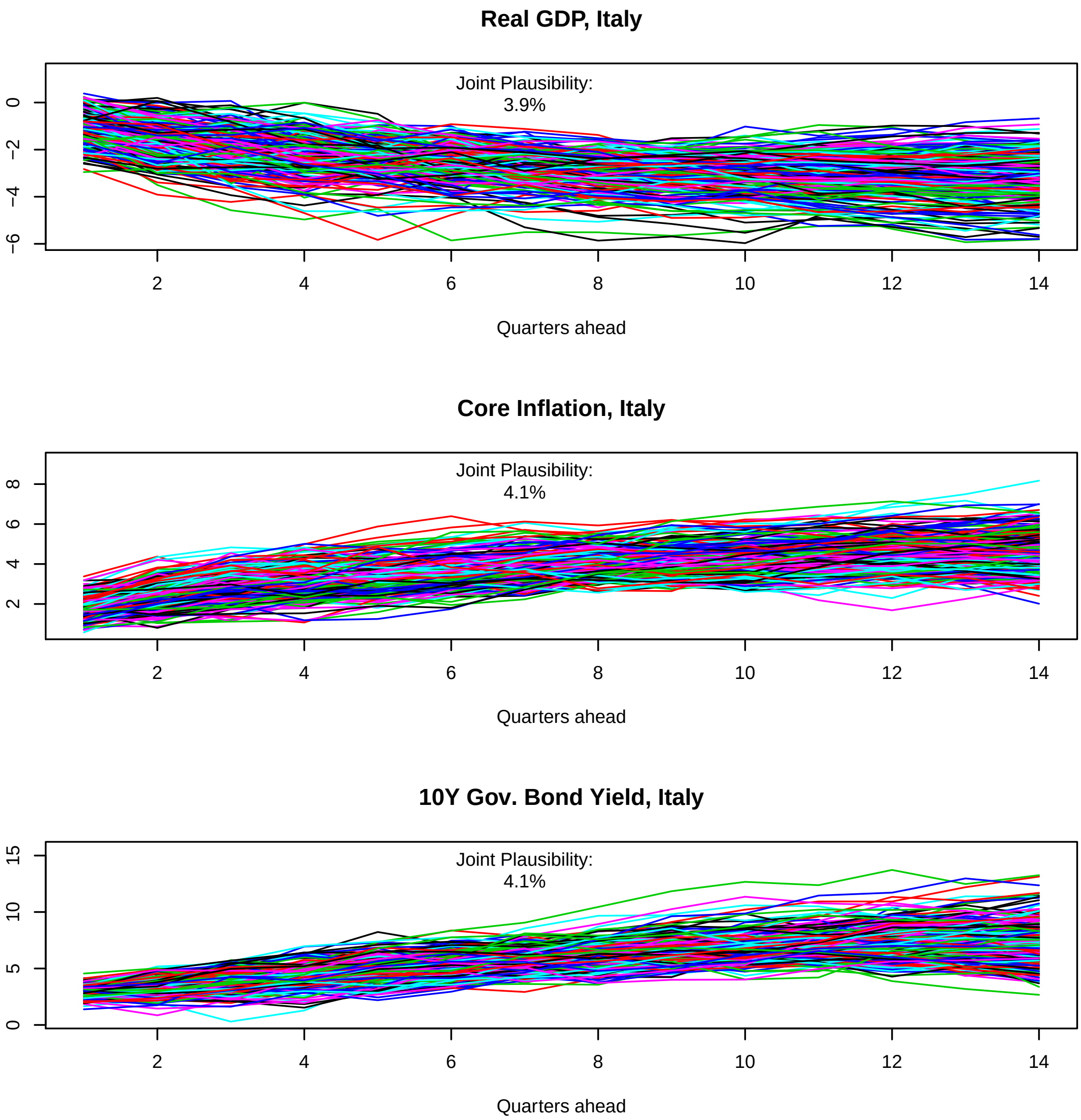


Figure 3: Worst-case Scenarios are selected in the respective 5% tail of the density forecast generated with 5000 simulations of the Large Bayesian VAR for Italy. The Joint plausibility is equal to the joint cumulative plausibility of the scenarios calculated with the Mahalanobis distance. Macroeconomic variables are in YoY growth rates for Real GDP and Core inflation, in % for 10-year Government Bond Yield.

Scenario Design with a view on lower-for-long 3-month Euribor interest rate

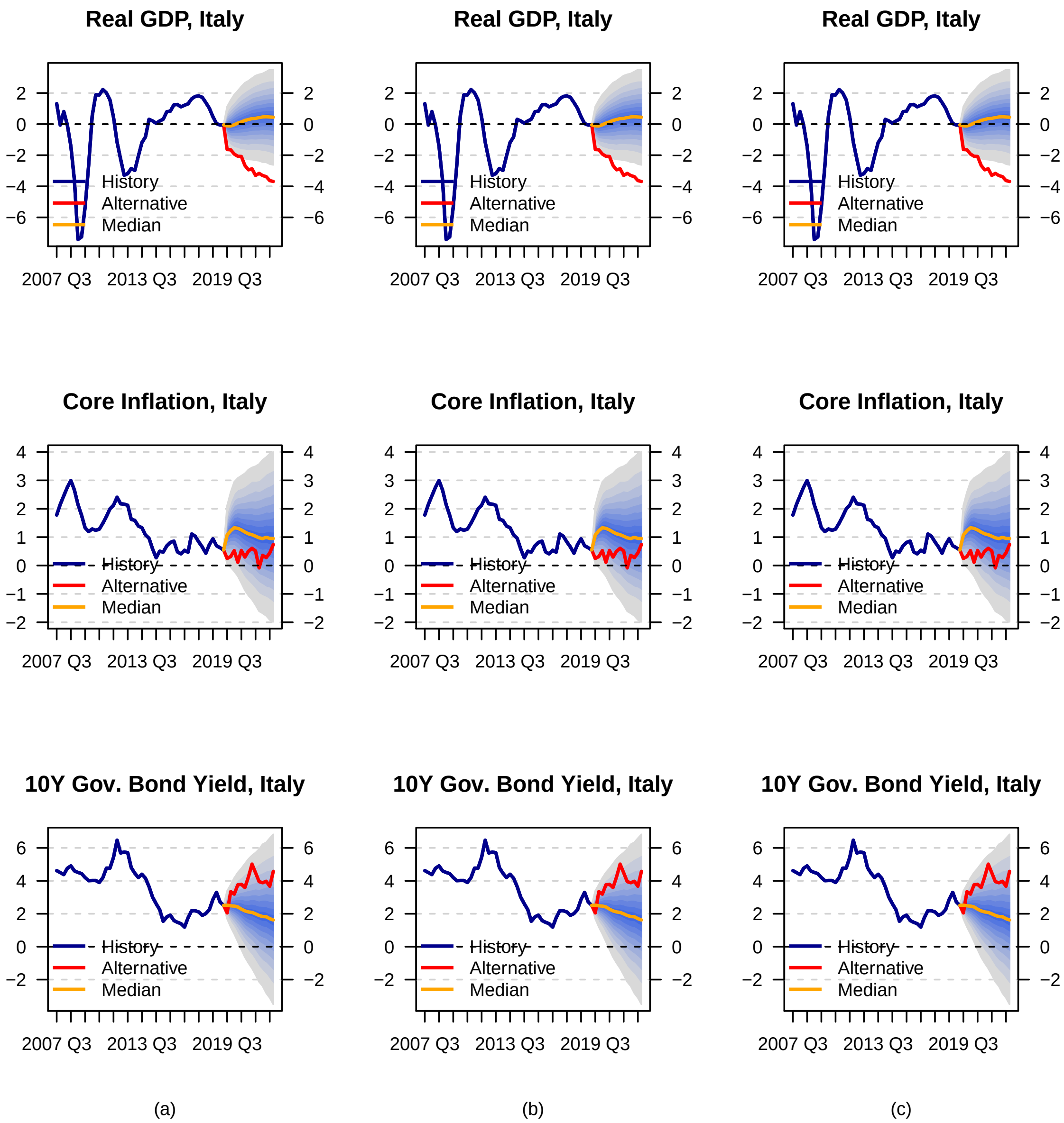
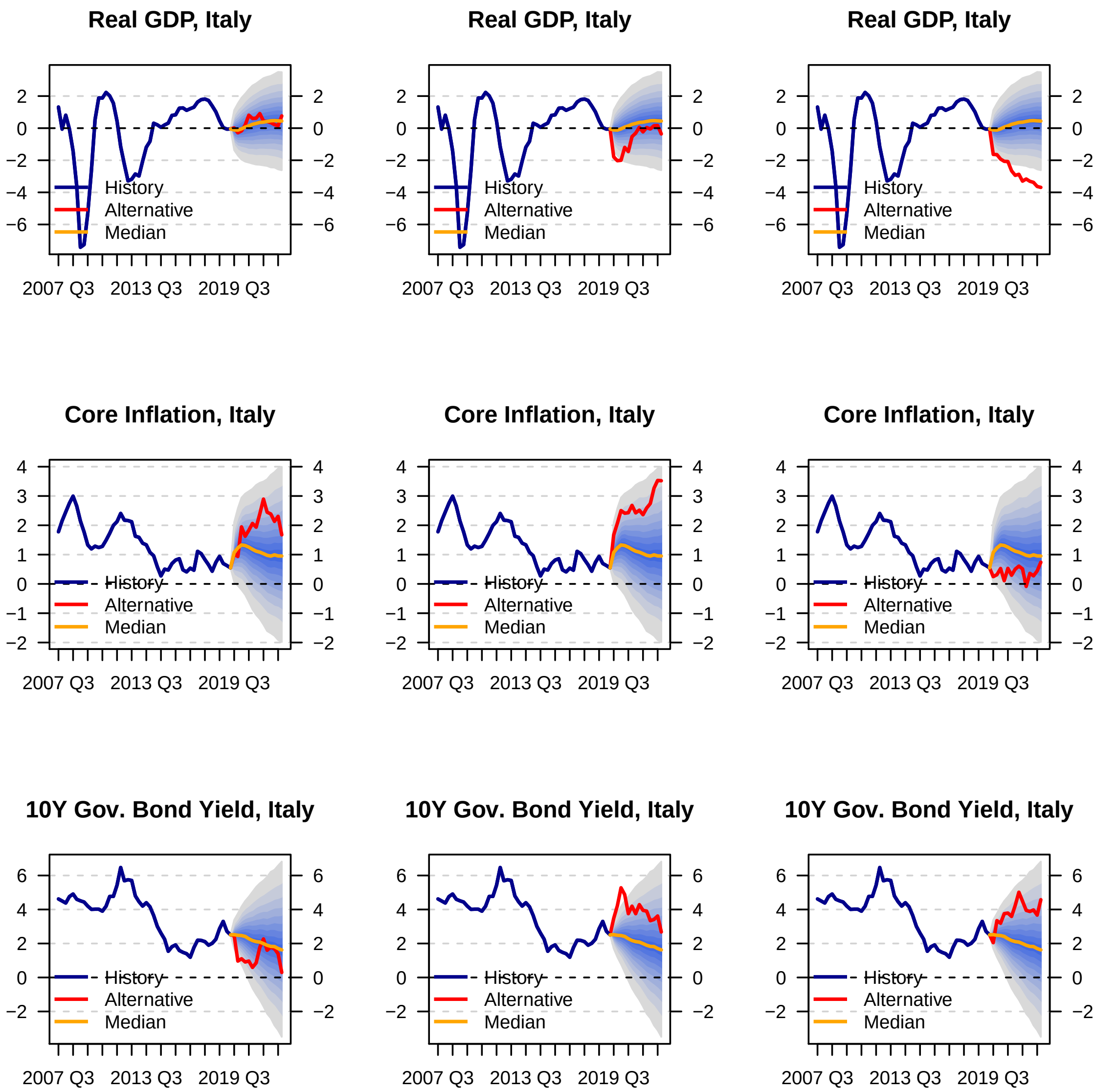


Figure 4: Fan-Charts describe the 10-90% confidence intervals around the median forecast. The forecast horizon is from 2019Q3 to 2022Q4. Panel (a): Entropy pooling with full confidence in the subjective view. Panel (b): Entropy pooling with 50% confidence in the subjective view. Panel (c): Frequentist approach. Real GDP and Core Inflation are in % YoY growth rate. 10-year Government Bond Yield is in %.

Scenario Design with a view on inverted US yield curve



(a) (b) (c)

Figure 5: Fan-Charts describe the 10-90% confidence intervals around the median forecast. The forecast horizon is from 2019Q3 to 2022Q4. Panel (a): Entropy pooling with full confidence in the subjective view. Panel (b): Entropy pooling with 50% confidence in the subjective view. Panel (c): Frequentist approach. Real GDP and Core Inflation are in % YoY growth rate. 10-year Government Bond Yield is in %.

Scenario Design with a view on a Bearish Italian Stock Market

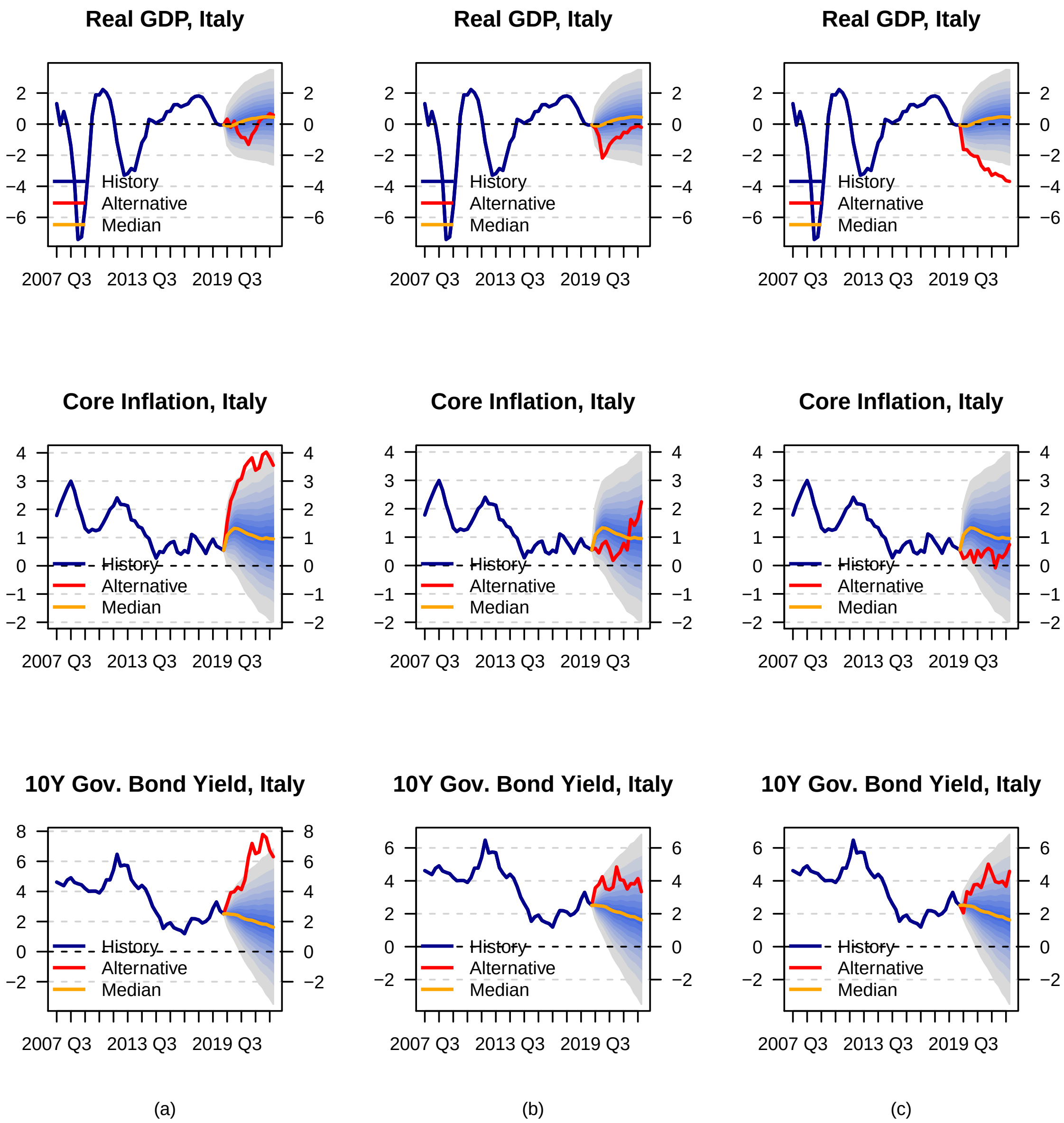


Figure 6: Fan-Charts describe the 10-90% confidence intervals around the median forecast. The forecast horizon is from 2019Q3 to 2022Q4. Panel (a): Entropy pooling with full confidence in the subjective view. Panel (b): Entropy pooling with 50% confidence in the subjective view. Panel (c): Frequentist approach. Real GDP and Core Inflation are in % YoY growth rate. 10-year Government Bond Yield is in %.