# Forecasting and Stress Testing with Quantile Vector Autoregression[*]

Sulkhan Chavleishvili[†] and Simone Manganelli[‡]

June 2021

**Abstract**

A quantile vector autoregressive (VAR) model, unlike standard VAR, traces the interaction among the endogenous random variables at any quantile. Forecasts of multivariate quantiles are obtained by factorizing the joint distribution in a recursive structure, but cannot be

obtained from reduced form estimation. Quantile impulse response functions are derived as direct generalization of standard VAR impulse response functions. The model is estimated using real and financial variables for the euro area. The dynamic properties of the system change across quantiles. This is relevant for stress testing exercises, whose goal is to forecast the tail behavior of the economy when hit by large financial and real shocks.

# 1 Introduction

Vector autoregressive (VAR) models are the empirical workhorse of macroeconomics. In their most basic formulation, they rely on constant coefficients and i.i.d. Gaussian innovations. There is, however, substantial empirical evidence that macroeconomic variables are characterized by nonlinearities and asymmetries (Perez-Quiros and Timmermann 2000, Hubrich and Tetlow 2015, Kilian and Vigfusson 2017, Adrian et al. 2019). This paper introduces a Quantile VAR (QVAR) model that captures such nonlinearities exploiting the semiparametric flexibility of regression quantiles. We spell out the links between traditional VAR and QVAR models, and show how to construct multistep forecasts with QVAR. We introduce the law of iterated quantiles,

which highlights the pitfall of relying on QVAR models in reduced form. By rewriting the QVAR model as a time-varying coefficient VAR model, quantile impulse response functions can be defined as a straightforward generalization of the standard ones. The methodology is applied to the euro area. QVAR provides the natural modeling framework to perform macro stress testing, by incorporating in an internally consistent fashion the robust empirical fact that financial shocks tend to have a strong and persistent asymmetric impact on the tails of the real economy. Our empirical findings reveal that the euro area economy exhibits different degrees of vulnerability to financial shocks at different points in time, with the months preceding Lehman's default standing out as particularly critical.

Quantile regression was introduced by Koenker and Bassett (1978) and has found many applications in economics (Koenker 2005, 2017). It is a semiparametric technique which allows different covariates to affect different parts of the distribution. Early applications to univariate time series include Engle and Manganelli (2004) and Koenker and Xiao (2006). White et al. (2015) have been the first to put forward the idea of a quantile VAR. Their focus is on high frequency financial variables, requiring an altogether different modeling strategy than for macroeconometric applications. One important issue White et al. (2015) have been grappling with is the concept of multivariate quantile. In fact, they refer to *pseudo* quantile impulse response functions, explicitly acknowledging the difficulty of extending the concept of impulse response functions to quantiles. Extensions to multivariate quantile

3

models are not straightforward. Hallin and Šiman (2017) discuss the theoretical difficulties of the concept of multivariate quantiles, showing how such extensions are not unique and still object of active research. In this paper, we follow the stratified modeling strategy of Wei (2009), because it provides the most direct link to the VAR literature. Wei's (2009) suggestion is to model multivariate quantiles by factorizing the contemporaneous joint distribution of random variables into marginal and conditional distributions. This corresponds to the recursive structural estimation strategy of VAR, which in turn is equivalent to the popular Cholesky decomposition of the reduced form VAR residuals. A similar triangular structure is adopted by Koenker and Ma (2006), although in a cross-sectional context. The extension of the numerous identification strategies proposed in the VAR literature (Ramey 2016) to the QVAR context constitutes an important area of research. The control variate approach by Koenker and Ma (2006) and the quantile instrumental variable estimation of Chernozhukov and Hansen (2005) are available tools to extend the research in this direction. Recent independent contributions deal with QVAR (Schüler 2014, Montes-Rojas 2019), quantile impulse response functions (Han, Jung and Lee 2019) and identification (Ruzicka 2020). Chavleishvili et al. (2021) present an application of QVAR for macroprudential policy.

We estimate a QVAR model on euro area data for industrial production growth and an indicator of financial distress, and perform three types of exercises. First, we estimate euro area growth at risk, defined as the 10%

4

quantile of industrial production growth. We find that severe financial shocks have an asymmetric impact on the distribution of the real variable. Modeling the conditional mean with a standard VAR seriously underestimates these macro-financial dynamics in times of stress, and underscores the potential of QVAR models for financial stability purposes. These results are broadly in line with those found by Adrian et al. (2019) for the U.S. economy. The empirical model estimated by Adrian et al. (2019) is equivalent to estimating only one equation of our QVAR model. Estimating the full QVAR allows us to perform impulse response analyses. We find that by hitting the system with a financial shock there is a strong, persistent and asymmetric impact on the distribution of industrial production, which takes about two years to be absorbed.

Second, one important policy application of QVAR is to stress testing. QVAR gives the flexibility to assess the impact of any future quantile realization. At its core, stress testing is a forecast of what happens to the system when hit by an arbitrary sequence of negative shocks. Stress scenarios can therefore be defined as a series (to be chosen by the policy maker or calibrated to past crises) of future quantile realizations within the QVAR system. We forecast euro area growth under alternative stress scenarios. If the euro area is hit by a sequence of six monthly consecutive financial and real 10% quantile realizations, its industrial production contracts by a maximum amount of about 4% if the stress scenario were applied in August 2008 and by less than 2% in July 2018. This contrasts with a median forecast

(that is, a sequence of median realizations of the endogenous variables) of industrial production hovering around 0% at the same points in time.

Third, we perform a counterfactual scenario analysis before Lehman Brothers' default and replay this scenario at each point in time. Using estimates up to August 2008, we find evidence of sizable and unprecedented downside risk to the euro area real economy already in mid 2007. Such counterfactual exercises can help policy makers to better understand the financial stability risks to the economy and put them in an historical perspective.

The paper is organized as follows. Section 2 introduces the QVAR model. It provides the links with standard OLS VAR and derives its forecasting properties. Section 3 estimates the QVAR model for the euro area, performs a stress testing exercise and estimates the counterfactual scenario of Lehman's bankruptcy at each point in time. Section 4 concludes.

## 2    Quantile vector autoregression

This section introduces and studies the properties of QVAR. It starts by defining the model in section 2.1, it shows how to construct forecasts (section 2.2) and discusses the implications of the law of iterated quantiles (section 2.3). Section 2.4 introduces the concept of quantile impulse response functions. Section 2.5 generalizes the results to any desired number of lags and Section 2.6 contains details about estimation and asymptotics.

## 2.1 Quantile VAR(1)

QVAR can be seen as a special case of White et al. (2015) and as a generalization of the univariate quantile autoregression model proposed by Koenker and Xiao (2006). We combine it with the triangular structure proposed by Wei (2009), to address the issue of multivariate quantile. We adopt Wei's approach because of its simplicity and because it provides the multivariate quantile counterpart of the Cholesky decomposition, the most commonly used identification strategy for structural VAR. Hallin and Šiman (2017) review the most recent literature on extensions of traditional single-output quantile regression methods to the multiple-output setting. This is an active area of research, which hopefully can bring further insights also for applications to QVAR.

Consider a time series vector $\{Y_t\} \equiv \{[Y_{1t}, \ldots, Y_{nt}]'\}$. For our purposes, it is important to define a recursive information set, which allows us to work with the stratified modeling strategy suggested by Wei (2009).

**Definition 1 *(Recursive information set)*** — *The recursive information set is defined as:*

$$\Omega_{1t} \equiv \{Y_{t-1}, Y_{t-2}, \ldots\}$$
$$\Omega_{it} \equiv \{Y_{i-1,t}, \Omega_{i-1,t}\} \quad i = 2, \ldots, n$$

According to this definition, the recursive information set $\Omega_{2t}$, say, contains

7

all the lagged values of $Y_t$ as well as the contemporaneous value of $Y_{1t}$.

We say that $\{Y_t\}$ follows a QVAR(1) process if the recursive $\theta_i$ quantile of $Y_{it}$ can be written as:

$$Q_{\theta_1}(Y_{1t}|\Omega_{1t}) = \omega_1(\theta_1) + a_{11}(\theta_1)Y_{1,t-1} + a_{12}(\theta_1)Y_{2,t-1} + \ldots + a_{1n}(\theta_1)Y_{n,t-1}$$

$$Q_{\theta_2}(Y_{2t}|\Omega_{2t}) = \omega_2(\theta_2) + a_{021}(\theta_2)Y_{1t}+$$

$$+ a_{21}(\theta_2)Y_{1,t-1} + a_{22}(\theta_2)Y_{2,t-1} + \ldots + a_{2n}(\theta_2)Y_{n,t-1}$$

$$\vdots$$

$$Q_{\theta_n}(Y_{nt}|\Omega_{nt}) = \omega_n(\theta_n) + a_{0n1}(\theta_n)Y_{1t} + \ldots + a_{0n,n-1}(\theta_n)Y_{n-1,t}+$$

$$+ a_{n1}(\theta_n)Y_{1,t-1} + a_{n2}(\theta_n)Y_{2,t-1} + \ldots + a_{nn}(\theta_n)Y_{n,t-1}$$

for any $\theta_i \in (0,1)$, $i \in \{1,\ldots,n\}$. When $n = 1$, this simplifies to the quantile autoregressive process of Koenker and Xiao (2006).

In more compact notation, we write:

$$\underbrace{Q_\theta(Y_t|\Omega_t)}_{n\times 1} = \underbrace{\omega(\theta)}_{n\times 1} + \underbrace{A_0(\theta)}_{n\times n}\underbrace{Y_t}_{n\times 1} + \underbrace{A_1(\theta)}_{n\times n}\underbrace{Y_{t-1}}_{n\times 1} \tag{1}$$

where $\theta \in (0,1)^n$ and the other elements stack the appropriate terms. The matrix $A_0(\theta)$ is a lower triangular $n \times n$ coefficient matrix, with zeros along the main diagonal. In the context of the VAR literature, this representation is equivalent to identification of the system by assuming a Cholesky decomposition of the variance covariance matrix of the residuals from a standard reduced form VAR (see, for instance, chapter 2 of Lütkepohl 2005).

8

The conditions for the QVAR process (1) to be covariance-stationary are given in Proposition 1, which is an extension of the result of Koenker and Xiao (2006). For a sequence of $n$-vectors of i.i.d. standard uniform random variables $\{U_t\}$, the QVAR process (1) can be written as:

$$Y_t = \omega_0(U_t) + A_0(U_t)Y_t + A_1(U_t)Y_{t-1}, \tag{2}$$

such that each $Y_{it}$, given the recursive information set of Definition 1, is monotonically increasing in $U_t$. Defining the terms $\nu(U_t) = [I - A_0(U_t)]^{-1}\omega_0(U_t)$ and $B(U_t) = [I - A_0(U_t)]^{-1}A_1(U_t)$, the QVAR process can equivalently be written as $Y_t = \nu(U_t) + B(U_t)Y_{t-1}$.

**Proposition 1** *(Stationarity of the QVAR process)* — *Assume that:*

1. *$\nu(U_t) - \mathrm{E}[\nu(U_t)]$ is a vector of i.i.d. random variables with zero mean, finite variance and continuous density function;*

2. *the matrix $\mathrm{E}[B(U_t) \otimes B(U_t)]$ has the largest eigenvalue less than one in absolute value.*

*Then the QVAR process (2) is covariance stationary and satisfies*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} (Y_t - \mu_Y) \sim \mathrm{N}\left(0, \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathrm{E}\,(Y_t - \mu_Y)(Y_t - \mu_Y)'\right),$$

*where $\mu_Y = [I - \mathrm{E}[B(U_t)]]^{-1}\,\mathrm{E}[\nu(U_t)]$.*

**Proof** — See appendix.

Equation (4) in Koenker and Xiao (2006) can be readily extended to QVAR, to give an example of how a QVAR process can be globally stationary, but display at the same time local, i.e. quantile specific, explosive behavior.

To further understand the link with the traditional VAR, the QVAR model (2) can be interpreted as a VAR model with time series dependence in its error structure:

$$Y_t = \omega_0 + A_0 Y_t + A_1 Y_{t-1} + \varepsilon_t \tag{3}$$

where $\omega_0 = E(\omega(U_t))$, $A_i = E(A_i(U_t))$ for $i = 0, 1$, $\varepsilon_t = \omega(U_t) - \omega_0 + (A_0(U_t) - A_0)Y_t + (A_1(U_t) - A_1)Y_{t-1}$. If the data generating process was a standard VAR with i.i.d. innovations, then the innovations would simplify to $\varepsilon_t = \omega(U_t) - \omega_0$, which in fact is an i.i.d. sequence. Under this assumption, the VAR and QVAR are characterized by identical dynamics. The more general model (1) allows for a richer structure in the data.

Model (1) implies that the quantile specification is monotonically increasing in $\theta$. If the quantile model is correctly specified, then the population quantiles are monotonic. This is obviously the case when the underlying DGP is a standard VAR with i.i.d. innovations. If the quantile model is misspecified, quantile monotonicity may not be satisfied. In that case, QVAR models can still be interpreted as useful local linear approximations of the data generating process, as discussed in Koenker and Xiao (2006) for univariate models. Relative to the standard VAR, the QVAR estimate can therefore

shed additional light into the dynamic properties of the random variables of interest. If quantile crossing is of concern, one can use techniques such as the monotonization method by Chernozhukov et al. (2010), the dynamic additive quantile specification of Gourieroux and Jasiak (2008), or the isotonization method suggested by Mammen (1991).

## 2.2  Forecasting with QVAR

Generic dynamic quantile forecasts can be produced following the stratified modeling approach described by Wei (2009). See also section 17.9 of Xiao (2017) for an analogous forecasting procedure for the univariate case.

According to Lemma 1 of Wei (2009), when the joint distribution of $Y_t$ is absolutely continuous, there is a one-to-one continuous mapping (also known as Rosenblatt's transformation) between the sample space $[Y_{1t}, \ldots, Y_{nt}]$ and the hypercube $(0,1)^n$. To generate a $p$-step ahead forecast, consider an $n$-vector $u_1^*$ whose elements are random draws from the i.i.d. uniform distribution with support on $(0,1)$. Then a draw from the one-step ahead forecast distribution of $Y_{T+1}$ is:

$$Y_{T+1}^* = (I_n - A_0(u_1^*))^{-1}(\omega(u_1^*) + A_1(u_1^*)Y_T)$$

where $I_n$ is the $n$-dimensional identity matrix.

Conditional on this draw, a draw from the two-step ahead forecast dis-

tribution of $Y_{T+2}$ is:

$$Y^*_{T+2} = (I_n - A_0(u^*_2))^{-1}(\omega(u^*_2) + A_1(u^*_2)Y^*_{T+1})$$

where $u^*_2$ is another $n$-vector with i.i.d. random draws from the standard uniform distribution. Iterating this process forward, it is possible to obtain a sample path of any desired $p$ length:

$$(Y^*_{T+1}, Y^*_{T+2}, \ldots, Y^*_{T+p}) \tag{4}$$

Under the regularity conditions ensuring consistent estimation of the parameters of the QVAR, with sufficiently large sample size, repeating the procedure provides a sequence of random draws from the forecast conditional distribution of $Y_{T+p}$.

To clarify the intuition behind the mechanics of quantile forecasting, consider forecasting the sequence of medians. Denoting (with slight abuse of notation) the respective matrices with $\omega(.5)$, $A_0(.5)$ and $A_1(.5)$, the sequence of median forecasts $p$-steps ahead is $\sum_{h=0}^{p-1} B(.5)^h \nu(.5) + B(.5)^p Y_T$, where $\nu(.5) \equiv (I_n - A_0(.5))^{-1}\omega(.5)$ and $B(.5) \equiv (I_n - A_0(.5))^{-1}A_1(.5)$. This is the median counterpart of the standard mean VAR forecast. However, as explained in the next subsection, this forecast does not coincide with the median of the $p$-step ahead forecast distribution. Such a forecast can be produced following the simulation procedure discussed in the previous paragraph.

Unlike the classical VAR, QVAR can forecast any desired sequence of

quantiles. This provides the natural environment to perform stress testing exercises. A policy maker interested in how the endogenous variables react to a given stress scenario can first define the scenario by choosing a series of future tail quantiles of interest (say, 10%), and then obtain the forecast of the endogenous variables conditional on the chosen scenario.

As discussed by Wei (2009), switching the order in which one constructs the stratified model provides another estimate of the joint distribution. With an $n$-variate QVAR, there are $n!$ possible orderings. If a particular ordering is justified by economic reasoning (as in a Cholesky decomposition), it can be used as primary order and be given a structural interpretation, as discussed in section 2.4. Otherwise, one can estimate the model using alternative orderings and combine the data for a more complete and robust covering of the sample space.

## 2.3 The Law of Iterated Quantiles

Unlike with traditional VAR, the reduced form of model (1) cannot be used for forecasting.

Let us work with a specific example to illustrate the intuition. Consider a bivariate quantile version of (1). The model can be rewritten as:

$$Y_{1t} = \omega_1(\theta_1) \qquad\qquad + a_{11}(\theta_1)Y_{1,t-1} + a_{12}(\theta_1)Y_{2,t-1} + \varepsilon_{1t}(\theta_1) \qquad (5)$$

$$Y_{2t} = \omega_2(\theta_2) + a_{021}(\theta_2)Y_{1t} + a_{21}(\theta_2)Y_{1,t-1} + a_{22}(\theta_2)Y_{2,t-1} + \varepsilon_{2t}(\theta_2) \qquad (6)$$

where the error terms satisfy the property $P(\varepsilon_{1t}(\theta_1) < 0|\Omega_{1t}) = \theta_1$ and $P(\varepsilon_{2t}(\theta_2) < 0|\Omega_{2t}) = \theta_2$.

The $\theta_2$ quantile of $Y_{2t}$ conditional on $\Omega_{2t} = \{Y_{1t}, Y_{t-1}\}$ is:

$$Q_{\theta_2}(Y_{2t}|\Omega_{2t}) = q_{\theta_2}^2(Y_{t-1}) + a_{021}(\theta_2)Y_{1t}$$

where $q_{\theta_2}^2(Y_{t-1}) \equiv \omega_2(\theta_2) + a_{21}(\theta_2)Y_{1,t-1} + a_{22}(\theta_2)Y_{2,t-1}$. This quantity is still a random variable at time $t-1$, because of the term $a_{021}(\theta_2)Y_{1t}$. Consider now the $\theta_1$ quantile of $Y_{1t}$, which in the current example is simply $Q_{\theta_1}(Y_{1t}|\Omega_{1t}) = q_{\theta_1}^1(Y_{t-1})$, where $q_{\theta_1}^1(Y_{t-1}) \equiv \omega_1(\theta_1) + a_{11}(\theta_1)Y_{1,t-1} + a_{12}(\theta_1)Y_{2,t-1}$. If $a_{021}(\theta_2) > 0$,[1] the $\theta_1$ quantile of the $\theta_2$ quantile of $Y_{2t}$ is:

$$Q_{\theta_1}((Q_{\theta_2}(Y_{2t}|\Omega_{2t})|\Omega_{1t}) = q_{\theta_2}^2(Y_{t-1}) + a_{021}(\theta_2)q_{\theta_1}^1(Y_{t-1})$$

The logic behind such iteration is the same as the simulation procedure described in the previous subsection. The following theorem reformulates the quantile forecasting iterating procedure in a way that allows us to draw comparisons with the law of iterated expectations used for forecasting expected values.

**Theorem 1 (Law of Iterated Quantiles)** — *Define $\varepsilon_{it}(\theta_i) \equiv Y_{it} - Q_{\theta_i}(Y_{it}|\Omega_{it})$, for $i \in [0, \ldots, n]$, where $Q_{\theta_i}(Y_{it}|\Omega_{it})$ is a generic element of*

---

[1]If $a_{021}(\theta_2) < 0$, the following computation gives the $1 - \theta_1$ quantile.

*the vector defined in* (1). *Then:*

$$Q_{\theta_1}\left(\ldots Q_{\theta_{i-1}}\left(Q_{\theta_i}\left(\varepsilon_{1t}\left(\theta_1\right)+\ldots+\varepsilon_{i-1,t}\left(\theta_{i-1}\right)+\varepsilon_{it}\left(\theta_i\right)|\Omega_{it}\right)|\Omega_{i-1,t}\right)\ldots|\Omega_{1t}\right)=0$$

**Proof** — See appendix.

A key difference with the law of iterated expectations is that the quantile of the sum of random variables is not necessarily equal to the quantile of the quantile of the sum. The important implication is that one should be careful in forecasting using parameter estimates from a QVAR in reduced form, as proposed for instance by Montes-Rojas (2019). Continuing with the example helps to clarify the issue. The reduced form version of model (5)-(6) gives for the second equation:

$$Y_{2t}=\left[\omega_2(\theta_2)+a_{021}(\theta_2)\omega_1(\theta_1)\right]+\left[a_{21}(\theta_2)+a_{021}(\theta_2)a_{11}(\theta_1)\right]Y_{1,t-1}+$$

$$\left[a_{22}(\theta_2)+a_{021}(\theta_2)a_{12}(\theta_1)\right]Y_{2,t-1}+\left[\varepsilon_{2t}(\theta_2)+a_{021}(\theta_2)\varepsilon_{1t}(\theta_1)\right]$$

Unless any contemporaneous structural relationship between the two endogenous variables is ruled out by setting $a_{021}(\theta_2)=0$, the residual does not satisfy the usual quantile regression restriction:

$$Q_{\theta_2}(\varepsilon_{2t}(\theta_2)+a_{021}(\theta_2)\varepsilon_{1t}(\theta_1)|\Omega_{1t})\neq 0$$

and the parameter estimates of the second reduced form equation would

depend also on $\theta_1$. The forecasting strategy of the previous subsection would therefore be no longer valid, because it is based on *independent* draws from the multivariate standard uniform distribution, for *given* parameter estimates of the two equations (5)-(6).[2]

The situation is different with standard VAR, as in that case the law of iterated expectations implies that $E(\varepsilon_{2t} + a_{021}\varepsilon_{1t}|\Omega_{1t}) = E(E(\varepsilon_{2t} + a_{021}\varepsilon_{1t}|\Omega_{2t})|\Omega_{1t}) = 0$, where $\varepsilon_{1t}$ and $\varepsilon_{2t}$ denote the residuals for the standard VAR and $a_{021}$ the corresponding contemporaneous coefficient.

## 2.4 Quantile impulse response functions

If the recursive model can be given a structural interpretation, it is possible to derive a structural quantile impulse response function.

Model (1) can be rewritten as a random coefficient model. Given a sequence of i.i.d. $n$-variate standard uniform random variables $\{U_t\}$, setting $\omega_0 = E(\omega(U_t))$ and $\epsilon_t = \omega(U_t) - \omega_0$ gives:

$$Y_t = \omega_0 + A_0(U_t)Y_t + A_1(U_t)Y_{t-1} + \epsilon_t \tag{7}$$

$$= \nu(U_t) + B(U_t)Y_{t-1} + (I_n - A_0(U_t))^{-1}\epsilon_t$$

where $\nu(U_t) = [I_n - A_0(U_t)]^{-1}\omega_0$ and $B(U_t) = [I_n - A_0(U_t)]^{-1}A_1(U_t)$.

---

[2]Another way to make the same point is that each equation of the reduced form QVAR can be used to recover the marginal distributions of the endogenous random variables. It is however impossible to recover a joint distribution from the marginal distributions without further assumptions, such as independence or a specific copula function.

16

In a standard VAR model, a shock to variable $i$ at time $t$ is affecting the forecast distribution only via the conditional expectations. Since it is a pure location model, a shift in the conditional expectation produces a parallel shift of the whole distribution and therefore of all its quantiles. The impulse response functions of the conditional expectation and the conditional quantiles would be identical.

In the case of QVAR, the shock is affecting all the quantiles in a potentially different way. Define the shock to the structural residuals of variable $i$, for $i = 1, \ldots, n$, as follows:

$$\ddot{\epsilon}_t^i = \epsilon_t + s^i \delta$$

where $\delta \in \mathbb{R}$ and $s^i$ is an $n$-vector of zeros with one in the $i^{th}$ position. Denoting with $\ddot{Y}_t$ the value of the dependent variables if the shock $\ddot{\epsilon}_t^i$ is applied, the impulse is affecting the entire future distributions. Such counterfactual distributions can be estimated using the same simulation procedure described in subsection 2.2. The quantile impulse response function can then be computed as the difference in forecast distributions, conditional on the applied shock. Denoting with $\ddot{Y}_{T+p}^*$ the corresponding draw conditional on $\ddot{Y}_T$, the

quantile impulse response function is:

$$\Delta_T^i \equiv \ddot{Y}_T - Y_T$$

$$= (I_n - A_0(\theta))^{-1} s^i \delta \tag{8}$$

$$\Delta_{T+p}^i = \ddot{Y}_{T+p}^* - Y_{T+p}^* \tag{9}$$

Notice that if one were to model only the median, this is again the median impulse response analogue of the standard mean impulse response function:

$$\Delta_{T+p}^i = B(.5)^p (I_n - A_0(.5))^{-1} s^i \delta \tag{10}$$

Quantile impulse response functions, however, will generally depend on the quantile paths which are considered.

If there are reasons to doubt the structural interpretation of the recursive system (1), it becomes necessary to develop alternative strategies to identify the structural shocks. The framework of Ma and Koenker (2006) allows for both the contemporaneous value of the endogenous variables and the structural residuals to affect the behavior of the following variables in a recursive fashion and uses a control variate estimation framework for identification. Another possibility is to introduce external instruments, following the approach of Chernozhukov and Hansen (2005). Recent independent contributions that adopt alternative definitions and identification strategies for quantile impulse response functions are Lee et al. (2021), Han et al. (2019),

Montes-Rojas (2019) and Ruzicka (2020).

## 2.5 General QVAR(q) model

Model (1) can be generalized to any desired lag order $q$ using the companion form of the time-varying coefficient model (7). Define the $pq$ vectors $\bar{\omega}(\theta) \equiv [\omega'(\theta), 0', \ldots, 0']'$, $\bar{Y}_{t+1} \equiv [Y'_{t+1}, Y'_t, \ldots, Y'_{t-q+2}]'$, $\varepsilon_{t+1}(\theta) \equiv [\epsilon'_{t+1}(\theta), 0', \ldots, 0']'$, and the $(pq \times pq)$ matrices

$$
A^0(\theta) = \begin{bmatrix} A_0(\theta), & \mathbf{0}, & \ldots, & \mathbf{0} \\ \mathbf{0}, & \mathbf{0}, & \ldots, & \mathbf{0} \\ \vdots & & \ddots & \\ \mathbf{0}, & \mathbf{0}, & \ldots, & \mathbf{0} \end{bmatrix} \text{ and } A^1(\theta) = \begin{bmatrix} A_1(\theta), & A_2(\theta), & \ldots, & A_q(\theta) \\ I_p, & \mathbf{0}, & \ldots, & \mathbf{0} \\ \vdots & & \ddots & \\ \mathbf{0}, & \ldots, & I_p, & \mathbf{0} \end{bmatrix}.
$$

Then the companion form of the QVAR(q) model is:

$$
\bar{Y}_{t+1} = \bar{\omega}(\theta) + A^0(\theta)\bar{Y}_{t+1} + A^1(\theta)\bar{Y}_t + \varepsilon_{t+1}(\theta) \tag{11}
$$

All the results of the previous sections extend to model (11).

## 2.6 Estimation and asymptotics

The recursive QVAR model (1) fits the framework of White et al. (2015), which can therefore be used for inference. Suppose the interest lies in estimating $p$ different quantiles. Let $q_t^j(\beta) \equiv \omega(\theta^j) + A_0(\theta^j)Y_t + A_1(\theta^j)Y_{t-1}$ and denote its $i^{th}$ element with $q_{it}^j(\beta)$, for $i = 1, \ldots, n$, where $\theta^j \in (0,1)^n$

for $j = 1, \ldots, p$ and we have made explicit the dependence on $\beta$, the vector containing all the unknown parameters in $[\omega(\theta^j)]_{j=1}^p$, $[A_0(\theta^j)]_{j=1}^p$, and $[A_1(\theta^j)]_{j=1}^p$. Define the regression quantile estimator $\hat{\beta}$ as:

$$\hat{\beta} = \arg\min_{\beta} T^{-1} \sum_{t=1}^{T} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{p} \rho_{\theta_i^j} \left( Y_{it} - q_{it}^j(\beta) \right) \right\}, \tag{12}$$

where $\rho_\theta(u) \equiv u(\theta - I(u < 0))$ is the standard check function and $\theta_i^j$ is the $i^{th}$ element of the $n$-vector $\theta^j$. The asymptotic distribution of the regression quantile estimator is provided by White et al. (2015), under the assumption that the process is globally stationary and ergodic:

$$\sqrt{T}(\hat{\beta} - \beta^*) \xrightarrow{d} N(0, Q^{-1}VQ^{-1}) \tag{13}$$

where

$$Q \equiv \sum_{i=1}^{n} \sum_{j=1}^{p} E[f_{it}^j(0) \nabla q_{it}^j(\beta^*) \nabla' q_{it}^j(\beta^*)]$$

$$V \equiv E[\eta_t \eta_t']$$

$$\eta_t \equiv \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{it}^j(\beta^*) \psi^j(\epsilon_{it}(\theta_i^j))$$

$$\psi^j(\epsilon_{it}(\theta_i^j)) \equiv \theta_i^j - I(\epsilon_{it}(\theta_i^j) \leq 0)$$

$$\epsilon_{it}(\theta_i^j) \equiv Y_{it} - q_{it}^j(\beta^*)$$

and $f_{it}^j(0)$ is the conditional density function of $\epsilon_{it}(\theta_i^j)$ evaluated at 0.

The asymptotic variance-covariance matrix can be consistently estimated as suggested in theorems 3 and 4 of White et al. (2015).[3]

The following corollary derives the standard errors of the forecasts.

**Corollary 1 (*Forecast standard errors*)** — *Let $Y_{t+p}^*(\hat{\beta})$ be the quantile forecast associated with a given path of the simulation procedure discussed in section 2.2, where it has been made explicit the dependence on the estimated model parameters $\hat{\beta}$. Then:*

$$\sqrt{T}(Y_{t+p}^*(\hat{\beta}) - Y_{t+p}^*(\beta^*)) \xrightarrow{d} N(0, \Phi(\beta^*)Q^{-1}VQ^{-1}\Phi'(\beta^*)) \qquad (14)$$

*where $\Phi(\beta^*) \equiv \partial Y_{t+p}^*(\beta^*)/\partial \beta'$.*

**Proof** — See appendix.

The standard errors associated with the impulse response functions (8)-(9) can be computed in a similar fashion.

# 3   Stress testing the euro area economy

We estimate a QVAR(1) to model the interaction between real and financial variables in Europe. We study the interrelationship between the euro area industrial production growth ($Y_{1t}$) and the composite indicator of systemic stress in the financial system (CISS, $Y_{2t}$) of Hollo, Kremer and Lo Duca

---

[3]Modern statistical software contain packages for regression quantile estimation and inference. This paper uses the interior point algorithm discussed by Koenker and Park (1996).

(2012). Our data sample is monthly and ranges from January 1999 to July 2018. We perform three exercises. First, we estimate short term euro area growth at risk (defined as the 10% quantile of $Y_{1t}$), as a function of financial conditions. Relative to univariate equations, estimation of the QVAR model permits us to construct quantile impulse response functions. Second, we forecast euro area growth under a severe stress scenario, where both the real and financial parts of the euro area economy are hit by a sequence of consecutive tail shocks. Third, we ask whether the QVAR methodology could have been helpful in detecting vulnerabilities in the months preceding Lehman Brothers' default.

## 3.1    Euro area growth at risk

Adrian et al. (2019) have shown that there are substantial asymmetries in the relationship between the US real GDP growth and financial conditions. In particular, they find that the estimated lower quantiles of the distribution of future GDP growth are significantly affected by financial conditions, while the upper quantiles appear to be more stable over time. The quantile model specification of Adrian et al. (2019) is the following:

$$Y_{1t} = \omega_1(\theta) + a_{11}(\theta)Y_{1,t-1} + a_{12}(\theta)Y_{2,t-1} + \varepsilon_{1t}(\theta)$$

They estimate this model for $\theta \in \{0.05, 0.25, 0.75, .95\}$. This corresponds to the first line of model (1) with two endogenous variables. We estimate,

instead, the full QVAR model and study its dynamic properties:

$$Y_{1t} = \omega_1(\theta) \qquad\qquad\quad +a_{11}(\theta)Y_{1,t-1} + a_{12}(\theta)Y_{2,t-1} + \varepsilon_{1t}(\theta) \qquad (15)$$

$$Y_{2t} = \omega_2(\theta) + a_{021}(\theta)Y_{1t}+a_{21}(\theta)Y_{1,t-1} + a_{22}(\theta)Y_{2,t-1} + \varepsilon_{2t}(\theta) \qquad (16)$$

By ordering CISS after industrial production, we impose the structural identification assumption that financial variables can react contemporaneously to real variables, but real variables react to financial developments only with a lag. This corresponds to a Cholesky identification where shocks to real economic variables can have an immediate impact on financial variables, while shocks to financial variables are allowed to affect real variables only with a lag. Given the speed at which financial markets react to news, this seems like a reasonable assumption, and it is relatively standard in the literature (see, for instance, Sims 1980, Christiano et al. 1999, Bloom 2009, Gilchrist and Zakrajšek 2012, and section 2.3.1 of Ramey 2016).

Figure 1 reports the estimated quantile coefficients of (15)-(16), together with 95% confidence intervals and the corresponding OLS estimates.[4] Checking condition 2 of Proposition 1 via simulation, results in the largest eigenvalue being 0.93. This implies that the system is globally stationary.

The interaction between real and financial variables can be tested by

---

[4]Estimation of the 5% and 95% quantiles may be stretching the limits of the available sample size, which consists of 235 monthly observations. In this case, the approach of Chernozhukov (2005) and Chernozhukov and Fernandez-Val (2011), based on extreme value theory, may provide a better approximation to the finite sample distribution of the quantile estimator than the one provided by White et al. (2015).

checking whether the off-diagonal coefficients are statistically different from zero. We observe the presence of substantial asymmetries, especially in the $a_{12}(\theta)$ coefficient, which cannot be detected with standard OLS models. The coefficient estimates of $a_{12}(\theta)$ are consistent with the findings of Adrian et al. (2019), whereby financial conditions significantly affect the left tail of the distribution of industrial production, but not the right tail.

Figure 2 shows that the impact of financial conditions is not only statistically significant, but also economically relevant. The figure reports the 10% quantile one step ahead forecast of industrial production, together with the 95% confidence intervals. As a comparison, the figure also shows the 10% quantile estimated indirectly from an OLS VAR, obtained as follows. We first estimated the OLS version of model (15)-(16). Second, we computed the 10% quantile of the OLS model residuals and added it to the estimated conditional VAR mean. This procedure would be consistent if model (15)-(16) were correctly specified for the mean and the residuals were i.i.d.

The comparison reveals the strong impact that worsening financial conditions have on the left tail of the forecast distribution. In relation to the OLS estimate, the estimated quantiles are quantitatively and statistically similar in tranquil times, but sharply different in crisis times. This highlights how modeling the interactions between real and financial variables with a standard OLS VAR could miss significant dynamics in the left tail, which are relevant from a financial stability perspective.
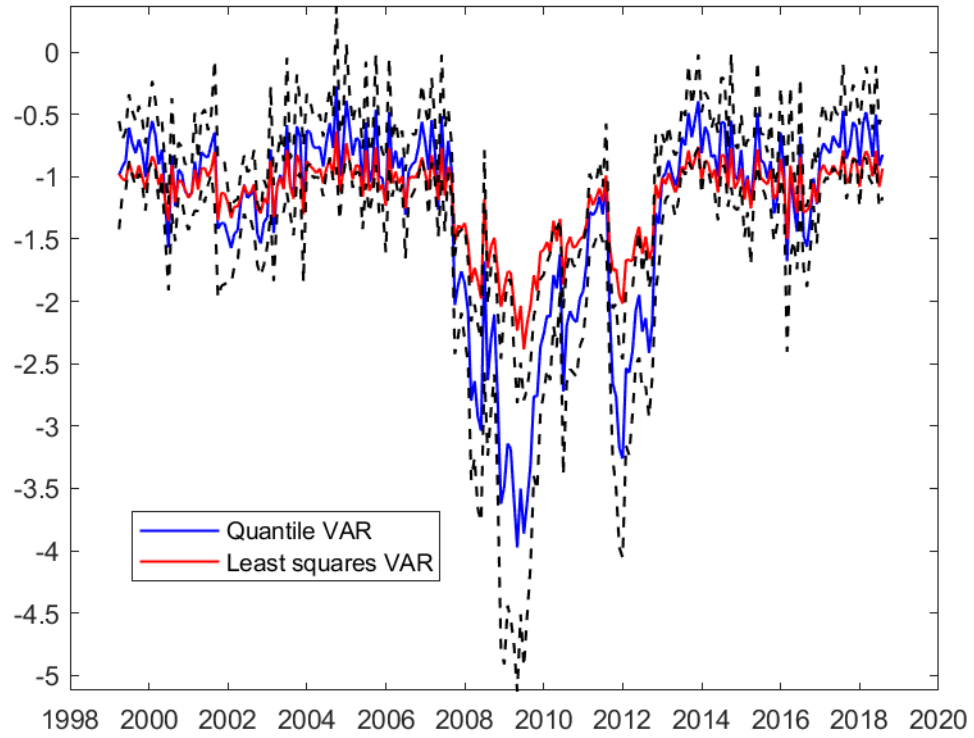
In figure 3, we compute the quantile impulse response function of indus-

Figure 1: Comparison of QVAR and VAR estimates
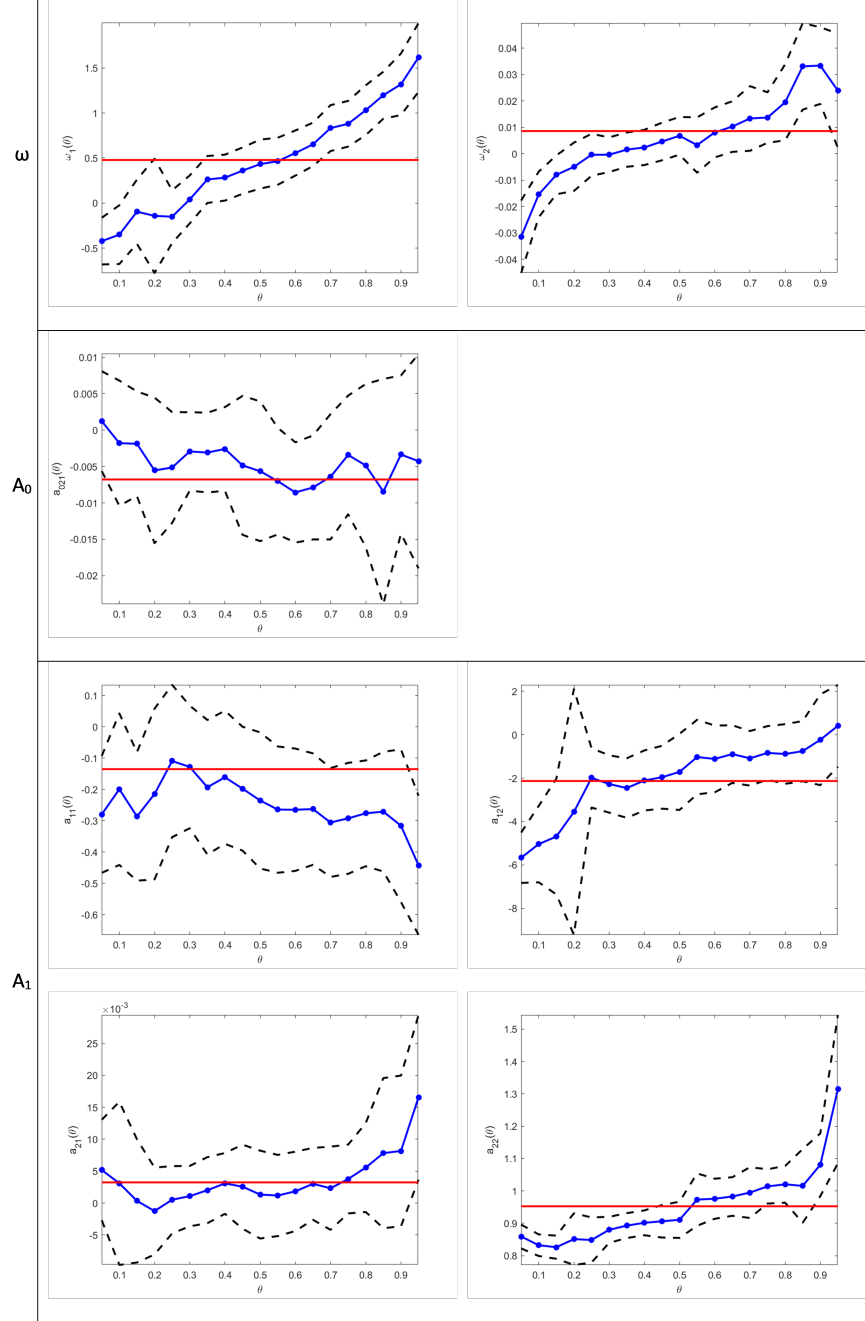
*Note:* Estimated coefficients of model (15)-(16) at different $\theta$ quantiles, with 95% confidence intervals. The flat lines represent the corresponding OLS estimates.

Figure 2: Euro area growth at risk



*Note:* Time series estimates of the 10% quantile of euro area industrial production, together with 95% confidence intervals. As a comparison, it is also reported the 10% quantile estimated by adding to the mean the 10% quantiles of the residuals from a standard OLS VAR. Under correct model specification, the two procedures would give consistent estimates of the 10% quantile. The OLS VAR procedure, however, is not able to capture the asymmetries between financial and real variables.

Figure 3: Quantile impulse response functions for the euro area industrial production
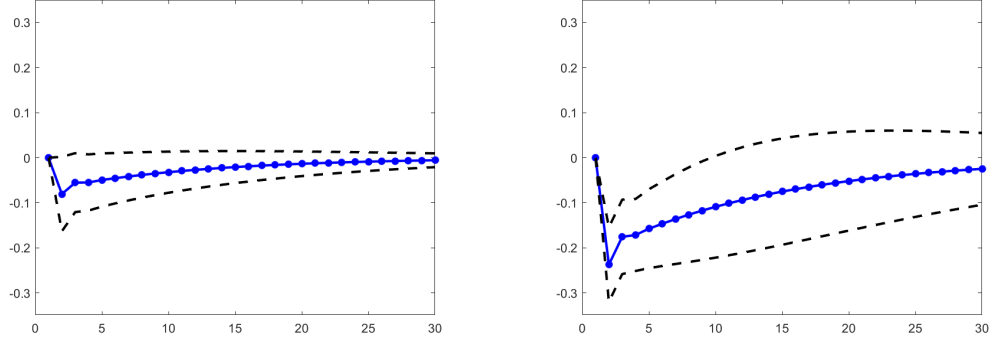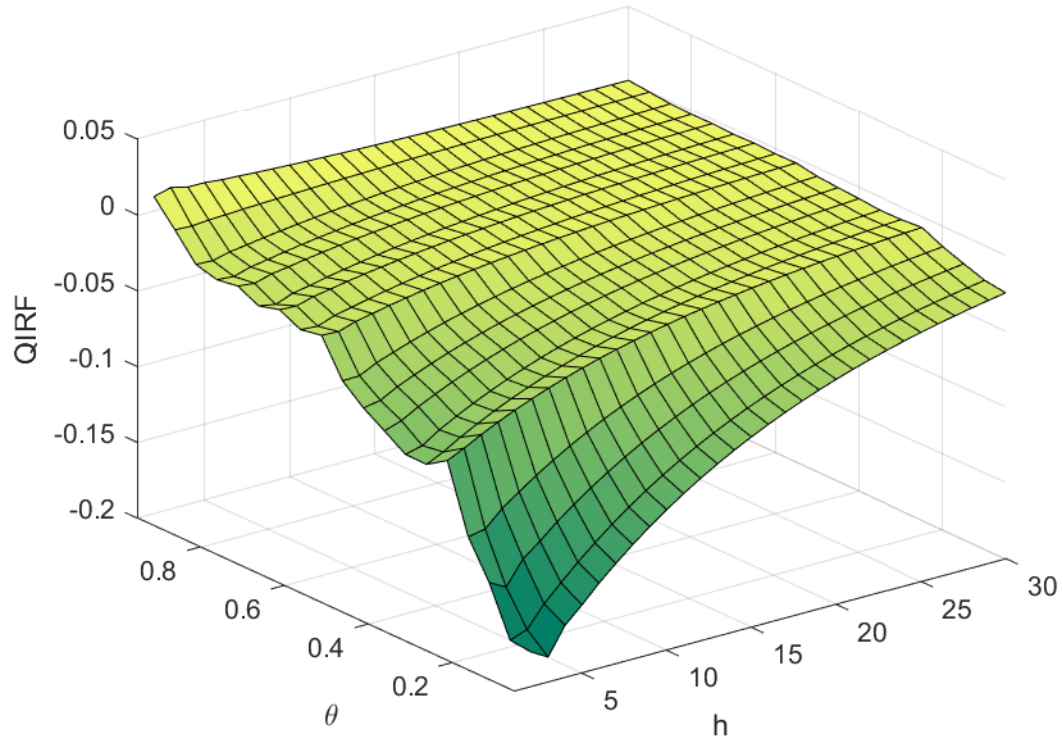


*Note:* The figure reports how a shock to the financial variable would affect the estimates of future median (left panel) and 10% (right panel) quantiles of euro area industrial production at different time horizons, conditional on a median forecast of the financial variable. 95% confidence intervals are also reported.

trial production corresponding to (8)-(9), following a one standard deviation shock to CISS structural median residuals. The thought experiment is the following: How different at any point in time the sequence of quantile forecasts would have been if we had observed a more severe realization in the financial conditions of the euro area economy? The left panel is the quantile impulse response function for the median forecasting path for both endogenous variables. It is the median equivalent to the standard OLS impulse response function for the mean. The QVAR model, however, allows us the flexibility to analyze any part of the forecast distribution, for any period ahead. The right panel of the figure reports the impulse response function for the sequence of 10% quantiles of industrial production and the median for CISS. It shows a stronger impact relative to the median.

In figure 4 we report a three dimensional quantile impulse response function. It is a concise way to visualize how each quantile of industrial production is responding to a shock to CISS. It is obtained by stacking next to each other all the panels of Figure 3, for the median forecasting path for CISS and different values of $\theta$ for industrial production. We did not report the confidence intervals to avoid cluttering the chart, but they can be readily constructed for each quantile as illustrated in Figure 3. The figure shows on the vertical axis the magnitude of the impulse responses, on the $h$ axis the number of periods for which the response is computed, and on the $\theta$ axis the quantile probabilities $\theta \in \{0.05, 0.1, \ldots, 0.9, 0.95\}$.

If the OLS VAR model were the correct representation of the dynamic interactions between real and financial variables, all elements of this three dimensional plot would shift in parallel and by the same magnitude across the different quantile probabilities: in an homoskedastic OLS VAR model, shifts in the forecast distribution are entirely driven by changes in the mean forecast. The fact that this does not happen is a further confirmation that OLS VAR may paint a misleading picture when the interest of the analysis is away from the central tendency of the distribution. Consistently with Figure 2, we continue to notice substantial asymmetric impacts in different parts of the distribution. In addition, the chart now reveals that the impact of the shock disappears for all quantiles considered after around 24 months. This analysis highlights one advantage of our framework. It is an internally consistent fully dynamic model of the real and financial variables of the euro

28

Figure 4: Three dimensional quantile impulse response functions



*Note:* The figure reports how a shock to the financial variable would affect the estimates of the different quantiles of euro area industrial production at different time horizons, conditional on a median forecast for CISS.

area economy, which allows us to study the propagation of shocks across the different parts of the distribution and through time.

## 3.2   Forecasting growth under stress scenarios

In Figure 5, we report the multi step QVAR forecasts of industrial production several months ahead, conditional on many different forecasting quantile

Figure 5: Forecasting and stress testing in the euro area
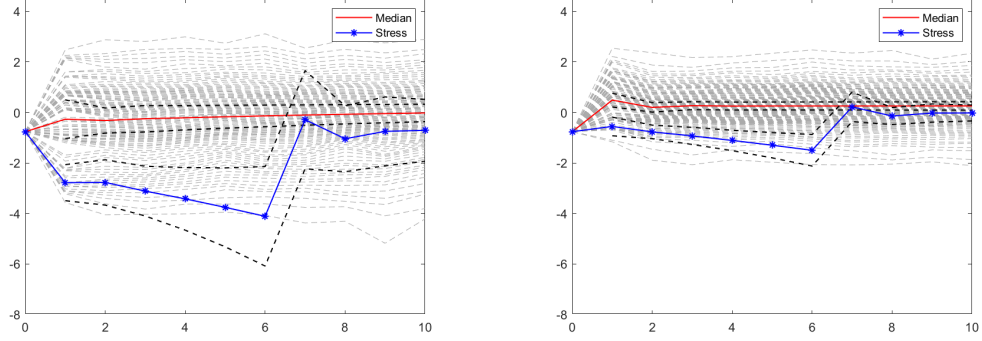


*Note:* The figure reports the forecasts of industrial production for the euro area associated with different scenarios. The path highlighted in blue corresponds to a scenario where both the real and financial variables evolve according to their median values. The path highlighted in red corresponds to the stress scenario with a 90% quantile forecast for the financial variable and and a 10% quantile forecast for the real variable for six consecutive months, followed by median forecasts afterwards. The panel on the left is the forecast as of August 2008, the panel on the right as of July 2018. 95% confidence intervals are reported around each scenario.

paths. The figure on the left reports the forecast as of September 2008 (the month of Lehman's default). The figure on the right is the forecast as of July 2018. Each dotted line corresponds to alternative forecasting sequences (4). The various dots at each point in time can be thought as possible realizations from the distribution of the future random variables, as discussed in Section 2.2.

We have highlighted two specific scenarios, both reported with the 95% confidence intervals. The red line corresponds to a situation where the sequence of future random variables are set to their median values. This roughly corresponds to the results that one would obtain from a standard OLS VAR analysis. Our framework, however, allows us also to create ar-

bitrary stress scenarios and to assess their impact. In the same figure, we have highlighted in blue with asterisk dots the forecast of the system associated with the following stress testing exercise. We assume that the euro area economy is hit by a series of six consecutive 90% quantile realizations to its financial system and 10% quantile realizations to its real economy. After that, we assume that the system is reverting to normal functioning, by imposing median realizations for all the variables. We notice that the median scenario is very similar at the two points in time considered in this exercise. The stress scenario, however, sees a much more severe contraction in industrial production in August 2008, peaking at about -4%, than in July 2018, where the peak is around -2%.

## 3.3 Counterfactual scenario analysis of Lehman Brothers' default

One year after the collapse of Lehman Brothers, Queen Elizabeth II famously asked: *Why did nobody notice it?* From the perspective of the methodology of this paper, predicting a crisis and its severity is like predicting that a certain sequence of adverse quantile realizations will hit the system. This is impossible. It is possible, however, to use the QVAR methodology to assess the resilience of an economy to alternative stress scenarios.

We estimate the model (15)-(16) using data only up to August 2008, one month before Lehman's default. For given parameter estimates, we use the

31

system to forecast industrial production six months ahead under the following sequences of quantile realizations to define alternative scenarios:

1. **Good financial scenario**: sequence of six 10% quantile realizations for both industrial production and CISS.

2. **Normal financial scenario**: sequence of six 10% quantile realizations for industrial production and median realizations for CISS.

3. **Bad financial scenario**: sequence of six 10% and 90% quantile realizations for industrial production and CISS, respectively.

We apply these scenarios at each month of our sample, and report in figure 6 the six month ahead forecasts for industrial production. It is evident that the good and normal financial scenarios were posing little risks to the euro area economy, since even after a sequence of six adverse quantile realizations of industrial production, growth at risk was quite contained. It is only under the combination of adverse real and financial quantile realizations that growth at risk is significantly affected. In fact, already in mid 2007, growth at risk under this adverse scenario had reached unprecedented magnitudes for the euro area, from an historical perspective. The large growth at risk under the bad financial scenario reveals the presence of a fat left tail in the distribution of the euro area industrial production, which would go unnoticed by simply estimating the 5% growth at risk using direct estimation techniques.

More generally, such counterfactual exercises are not feasible with the direct forecast approach. By directly quantile regressing industrial production

Figure 6: Growth at risk under alternative scenarios as of August 2008
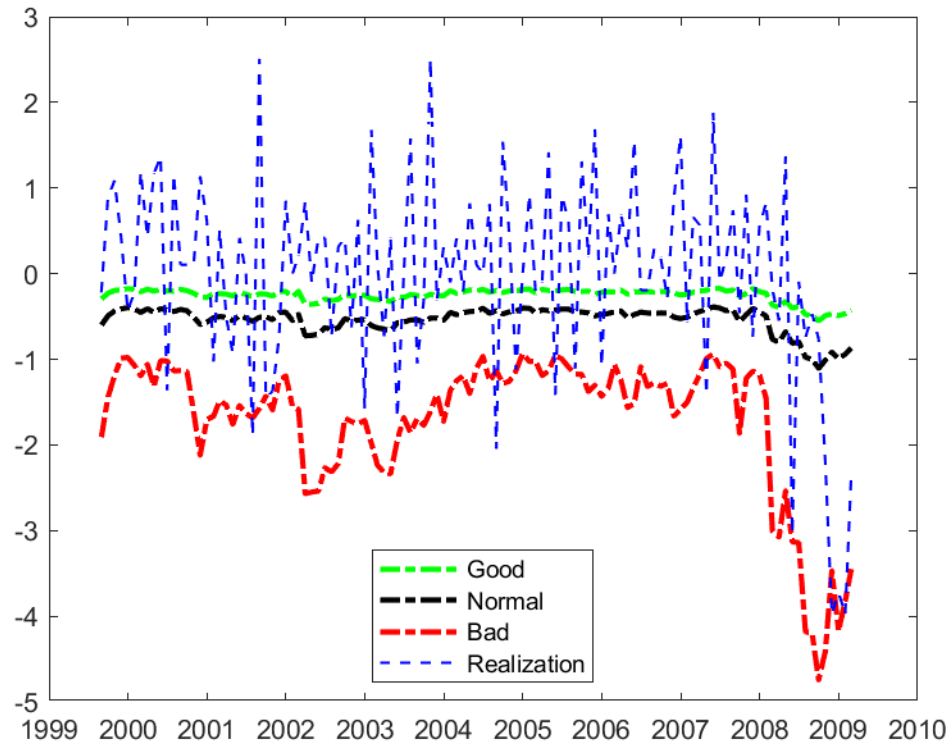


*Note:* Six month ahead forecast of euro area industrial production under three alternative scenarios. The good, normal and bad scenarios are defined by a sequence of six consecutive benign, normal and adverse quantile realizations. The parameter of QVAR are estimated using only observations up to August 2008.

six months ahead against current real and financial conditions, one implicitly imposes that the system evolves according to some average scenario during the intervening six months. While this may be a reasonable assumption if one is interested in modeling the conditional mean of the endogenous variables, it seems like an undesirable constraint to impose when modeling their tail behavior. Notice, however, that if one is interested in such unconditional scenario, this can be recovered from the empirical distribution obtained by simulating QVAR under all alternative quantile scenarios (similarly to all the possible dotted lines of figure 5) and then choosing the desired empirical quantile forecast.

# 4   Conclusion

We have developed a quantile VAR model and used it to forecast and stress test the interaction between real and financial variables in the euro area. Unlike OLS VAR, QVAR models each quantile of the distribution. This provides the natural modeling environment to design particular stress scenarios and test the impact that they have on the economy. A stress scenario is just a sequence of tail quantile realizations, which can be arbitrarily chosen by the policy maker or calibrated to mimic previous crisis episodes. We find the presence of strong asymmetries in the transmission of financial shocks in the euro area, with negative financial shocks being particularly harmful. By modeling the average interaction between the random variables, OLS VAR

34

models miss most of these detrimental interactions.

# Appendix — Proofs

**Proof of Proposition 1 (Stationarity of the QVAR process)** —
Following Koenker and Xiao (2006), the results exploits the arguments of
Nicholls (1982), chapter 2, and the central limit theorem given by Theorem A.1.4 discussed in details in Billingsley (1961). In particular, for
$\tilde{\nu}_t = \nu(U_t) - \mathrm{E}[\nu(U_t)] + (B(U_t) - \mathrm{E}[B(U_t)]) \mu_Y$ and $\bar{Y}_t = Y_t - \mu_Y$, recursive
substitution gives:

$$\bar{Y}_t = \prod_{j=0}^{h-1} B(U_{t-j})\bar{Y}_{t-h} + \sum_{j=0}^{h-1} C_j \tilde{\nu}_{t-j},$$

where $C_0 = I$ and $C_j = \prod_{i=0}^{j-1} B(U_{t-i})$ for $j \geq 1$. By the result of Nicholls
(1982), the stationary is implied by the convergence of the following term as
$h \to \infty$:

$$\mathrm{vec}\,\mathrm{E}\left(\sum_{j=0}^{h-1} C_j \tilde{\nu}_{t-j}\right)\left(\sum_{j=0}^{h-1} C_j \tilde{\nu}_{t-j}\right)' = \sum_{j=0}^{h-1} (\mathrm{E}\,B(U_t) \otimes B(U_t))^j \,\mathrm{vec}\,\mathrm{E}\,\tilde{\nu}_t\tilde{\nu}_t'.$$

The result follows from $\mathrm{E}[\tilde{\nu}_t\tilde{\nu}_s'] = 0$ for $t \neq s$, standard matrix manipulations
and the conditions of the proposition 1. □

**Proof of Theorem 1 (Law of iterated quantiles)** — Start from the

innermost expression:

$$Q_{\theta_i}(\epsilon_{1,t+1}(\theta_1)+\ldots+\epsilon_{i-1,t+1}(\theta_{i-1})+\epsilon_{i,t+1}(\theta_i)|\Omega_{it}) = \epsilon_{1,t+1}(\theta_1)+\ldots+\epsilon_{i-1,t+1}(\theta_{i-1})$$

because, by definition $Q_{\theta_i}(\epsilon_{i,t+1}(\theta_i)|\Omega_{it}) = 0$ and the other terms are not random, as they belong to the conditioning set. Repeating this reasoning for each of the remaining terms gives the result. $\square$

**Proof of Corollary 1 (Forecast standard errors)** — Consider the mean value expansion $Y_{T+h}(\hat{\beta}) = Y_{T+h}(\beta^*) + \Phi(\bar{\beta})(\hat{\beta} - \beta^*)$. The result follows from the asymptotic properties of $\hat{\beta}$. $\square$

# References

Adrian, T., N. Boyarchenko and D. Giannone (2019), Vulnerable growth, *American Economic Review*, 109 (4), 1262-1289.

Billingsley, P. (1961), The Lindeberg-Levy Theorem for Martingales, *Proceedings of the American Mathematical Society*, 12: 788-792.

Bloom, N. (2009), The impact of uncertainty shocks, *Econometrica*, 77(3): 623-685.

Chavleishvili, S., R.F. Engle, S. Fahr, M. Kremer, S. Manganelli and B. Schwaab (2021), The risk management approach to macro-prudential policy, ECB Working Paper, forthcoming.

Chernozhukov, V. (2005) Extremal quantile regression, *Annals of Statistics*, 33: 806–839.

Chernozhukov, V. and I. Fernandez-Val (2011) Inference for extremal conditional quantile models, with an application to market and birthweight risks, *Review of Economic Studies*, 78: 559–589.

Chernozhukov, V., I. Fernandez-Val and A. Galichon (2010), Quantile and probability curves without crossing, *Econometrica*, 78: 1093-1125.

Chernozhukov, V. and C. Hansen (2005), An IV model of quantile treatment effects, *Econometrica*, 73(1), 245–262.

Christiano, L., M. Eichenbaum and C. Evans (1999), The effects of monetary policy shocks: evidence from the flow of funds, *Review of Economics and Statistics*, 78: 16-34.

Engle, R. F. and S. Manganelli (2004), CAViaR: Conditional autoregressive Value at Risk by regression quantiles, *Journal of Business & Economic Statistics*, 22, 367-381.

Gilchrist, S. and E. Zakrajšek (2012), Credit Spreads and Business Cycle Fluctuations, *American Economic Review*, 102(4): 1692-1720.

Gourieroux, C. and J. Jasiak (2008), Dynamic Quantile Models, *Journal of Econometrics*, 147 (1): 198-205.

Hallin, M and M. Šiman (2017), Multiple-Output Quantile Regression, in *Handbook of Quantile Regression*, edited by Roger Koenker, et al., CRC Press LLC.

Han, H., Jung, W. and J.H. Lee (2019), Estimation and Inference of Quantile Impulse Response Functions by Local Projections: With Applications to VaR Dynamics, Working Paper Available at SSRN: https://ssrn.com/abstract=3466198.

Hollo, D., M. Kremer and M. Lo Duca (2012), CISS – A composite indicator of systemic stress in the financial system, ECB Working Paper No. 1426.

Hubrich, K. and R. J. Tetlow (2015), Financial stress and economic dynamics: The transmission of crises, *Journal of Monetary Economics*, 70 (C), 100-115.

Kilian, L. and R. J. Vigfusson (2017), The role of oil price shocks in causing U.S. recessions, *Journal of Money, Credit and Banking*, 49 (8), 1747-1776,

Koenker, R. (2005), *Quantile regression*, Econometric Society Monograph, Cambridge University Press.

Koenker, R. (2017), Quantile regression: 40 years on, *Annual Review of Economics*, 9, 155-176.

Koenker, R. and G. Bassett (1978), Regression quantiles, *Econometrica*, 46, 33-50.

Koenker, R. and L. Ma (2006), Quantile regression methods for recursive structural equation models, *Journal of Econometrics*, 134(2), 471–506.

Koenker, R. and B. J. Park (1994), An interior point algorithm for nonlinear quantile regression, *Journal of Econometrics*, 71, 265-283.

Koenker, R. and Z. Xiao (2006), Quantile autoregression, *Journal of the American Statistical Association*, 101 (475), 980-990.

Lütkepohl, H. (2005), *The New Introduction to Multiple Time Series Analysis*, Springer-Verlag Berlin.

Mammen, E. (1991), Nonparametric regression under qualitative smoothness assumptions, *The Annals of Statistics*, 19: 741-759.

Montes-Rojas, G. (2019) Multivariate Impulse Response Functions, *Journal of Time Series Analysis*, 40: 739-752.

Nicholls, D.F. and B.G. Quinn (1982), *Random Coefficient Autoregressive Models: An Introduction*, Springer-Verlag.

Perez-Quiros, G. and A. Timmermann (2000), Firm size and cyclical variation in stock returns, *The Journal of Finance*, LV (3), 1229-1262.

Ramey, V. (2016), Macroeconomic shocks and their propagation, in *Handbook of Macroeconomics*, edited by J.B. Taylor and H. Hulig.

Ruzicka, J. (2020), Quantile structural vector autoregression, mimeo, University Carlos III.

Sims, C.A. (1980), Macroeconomics and reality, *Econometrica*, 48: 1–48.

Schüler, Y. S. (2014), Asymmetric effects of uncertainty over the business cycle: A quantile structural vector autoregressive approach, University of Konstanz, Working Paper Series 2014-02.

White, H., T.H. Kim and S. Manganelli (2010) Modeling autoregressive conditional skewness and kurtosis with multi-quantile CAViaR, in *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, edited by T. Bollerslev, J. Russell and M. Watson, Oxford University Press.

White, H., T.H. Kim and S. Manganelli (2015), VAR for VaR: Measuring tail dependence using multivariate regression quantiles, *Journal of Econometrics*, 187, 169-188.