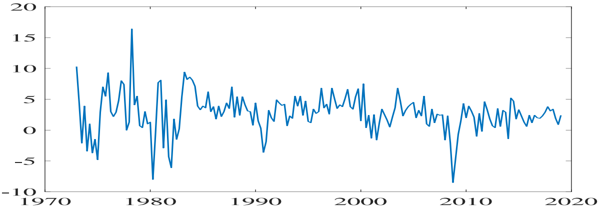
Distribution Vector Autoregression: Eliciting Macro and Financial Dependence

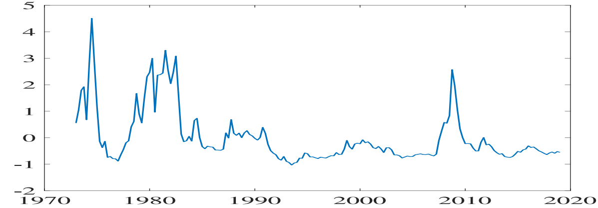
Yunyun Wang, Tatsushi Oka and Dan Zhu

August 24, 2023

GDP and NFCI from 1973 to 2019



(a) GDP growth



(b) Chicago Fed's NFCI

Observations

1. A nonlinear relationship between future GDP growth and financial conditions.
2. GDP growth is, on average, more volatile than NFCI
3. Extreme negative outcomes in GDP growth tend to coincide with extreme positive outcomes of the NFCI

NFCI

The Chicago Fed's National Financial Conditions Index (NFCI) is a weighted average of 105 indicators of risk, credit, and leverage in the financial system–each expressed relative to its sample average and scaled by its sample standard deviation.

$$\text{NFCI} > 0 \leftrightarrow \text{ tighter than on average}$$

The methodology for the NFCI is described in Brave and Butters (2012) and is based on the quasi-maximum likelihood estimators for large dynamic factor models developed by Doz, Giannone, and Reichlin (2012).

Some Existing Empirical Results

1) Adrian et.al(2019, AER) studied

$$GDP_{t+h} \Big| GDP_t, NFCI_t$$

and suggest deteriorating financial conditions are associated with an increase in the conditional volatility, in particular downside risks to GDP increase. Mitchell et.al(2022) finds that this increased skewness is coming from the conditional distribution exhibits multi-modality.

2) Adrian et.al(2021, IER) studied

$$GDP_{t+h}, NFCI_{t+h} \Big| GDP_t, NFCI_t$$

and suggested deteriorating financial conditions are associated with multi-modality of the joint conditional distribution.

Multivariate Time series

Suppose that we observe a stationary time series $\{(Y_t, Z_t)\}_{t=1}^{T}$ with a sample size of T, where $Y_t = (Y_{1t}, \ldots, Y_{Jt})^\top$ is a J-dimensional outcome variables and Z_t is a $k \times 1$ vector of conditioning variables.

Given the multivariate time series, the objective is to estimate the conditional joint distribution

$$F_{Y_t|Z_t}(y|z)$$

of Y_t given $Z_t = z$, for $(y, z) \in \mathscr{Y} \times \mathscr{Z}$.

Vector Autoregression

Since the seminal work of Sims (1980), vector autoregression (VAR)

$$y_t = b + \sum_{i=1}^{p} B_i y_{t-i} + \varepsilon_t$$

where $\varepsilon_t \sim N(\mathbf{0}_n, \Sigma)$, has become an workhorse model in empirical macroeconomics and finance, for basic quantitative description, forecasting and structural analysis of multivariate time series (see Litterman, 1986; Stock and Watson, 2001).

$$Y_t | z_t \sim N(z_t' \beta, \Sigma)$$

where $z_t = [1, y_{t-1}', ..., y_{t-p}]'$ and $\beta = [b', vec(B_1)', ..., vec(B_p)']'$.

The extension of the model to incorporate

- time-varying coefficient,

$$\beta_t = [b_t', vec(B_{t1})', ..., vec(B_{tp})']$$

and

$$\beta_t = \beta_{t-1} + v_t, v_t \sim N(\mathbf{0}, \Omega);$$

- stochastic volatility, i.e., common

$$\Sigma_t = \exp(h_t)\Sigma,$$

Cholasky

$$\Sigma_t^{-1} = B_0' diag(exp(h_{t1}), ..., exp(h_{tn}))B_0$$

and Factor

$$\varepsilon_t = Lf_t + u_t.$$

General characterisation of the Conditional Distribution

First, we apply distribution factorization,

$$X_{1t} := (1, Z_t^\top)^\top \text{ and } X_{jt} := (1, Z_t^\top, Y_{1t}, \ldots, Y_{j-1,t})^\top \text{ for } j = 2, \ldots, J,$$

and let $F_{Y_{jt}|X_{jt}}$ be the marginal conditional distribution of Y_{jt} given X_{jt}. Subsequently, the distribution factorization yields a transformation $\rho : \ell^\infty(\mathscr{Y} \times \mathscr{Z}) \to \times_{j=1}^{J} \ell^\infty(\mathscr{Y}_j \times \mathscr{X}_j)$ such that

$$F_{Y_t|Z_t} = \rho(F_{Y_{1t}|X_{1t}}, \ldots, F_{Y_{Jt}|X_{Jt}}). \tag{1}$$

Issue: Order...

IRF

An impulse response function in the VAR literature has been long interested in

$$\mathbb{E}^*[Y_{j,t+h}|Z_t] - \mathbb{E}[Y_{j,t+h}|Z_t],$$

- Forecast error impulse response
- Orthogonal impulse responses: Cholasky
- Structural impulse responses: Structural VAR identification
- Generalised impulse responses

DIRF: Baseline

We consider a local projection approach by integrating the conditional distribution of observable variables with respect to a counterfactual distribution to develop the DIRFs. Given a non-negative integer h, first, the (baseline) joint distribution $F_{Y_{t+h}|Z_t}$ of h-ahead outcomes Y_{t+h} conditional on Z_t is written as

$$F_{Y_{t+h}|Z_t} = \int F_{Y_{t+h}|Y_t,Z_t} dF_{Y_t|Z_t},$$

where $F_{Y_t|Z_t}$ and $F_{Y_{t+h}|Y_t,Z_t}$ are two different conditional distributions of the observed variables that are identified from the data and characterized in a manner similar to the proposed semiparametric approach.

DIRF: An Alternative Scenario

We consider a scenario in which the conditional distribution at time t is a conditional distribution $G_{Y_t|Z_t}$ instead of $F_{Y_t|Z_t}$.

Under the scenario with the distribution $G_{Y_t|Z_t}$, the counterfactual conditional joint distribution is defined as

$$F^*_{Y_{t+h}|Z_t} := \int F_{Y_{t+h}|Y_t,Z_t} dG_{Y_t|Z_t}.$$

DIRF

We formally define the distribution impulse response function (DIRF) as follows.

Definition

Let $F_{Y_{j,t+h}|Z_t}$ and $F^*_{Y_{j,t+h}|Z_t}$ be the marginal conditional distributions of the j-th outcome variable for the joint distributions $F_{Y_{t+h}|Z_t}$ and $F^*_{Y_{t+h}|Z_t}$ conditional on Z_t, respectively. The distribution impulse response function of the j-th variable after h periods is defined as

$$DIR_{j,h} := F^*_{Y_{j,t+h}|Z_t} - F_{Y_{j,t+h}|Z_t}. \tag{2}$$

SVAR: Orthogonal shocks

When we consider a linear system, the reduce form shock ε_t can be written as a linear combination of structural form shocks z_t

$$\varepsilon_t = Az_t$$

where

$$\mathscr{A} := \left\{ A : AA' = \mathbb{E}[e_t e_t'] \right\}.$$

The problem is that $\dim(\mathscr{A}) = \infty$, even when n is very small. Thus,

- the cholasky decomposition as a point identification scheme
- the sign restriction as a set identification scheme

are typically used to identify the shocks.

Uniform Shocks

Under our DVAR case, we have a nonlinear transformation

$$\varepsilon_t = T(u_t)$$

where $u_t \in [0,1]^n$ is the vector of independent uniforms. The Rosenblat transformation implies that the combinatory set $\mathscr{C}_n$ is of finite dimension, $n!$ such that

$$\begin{aligned}
\varepsilon_{t,k_1} &= F_{k_1}^{-1}(u_{t1}) \\
\varepsilon_{t,k_2} &= F_{k_2}^{-1}(u_{t2}|\varepsilon_{t,k_1}) \\
&\dots \\
\varepsilon_{t,k_n} &= F_{k_n}^{-1}(u_{tn}|\varepsilon_{t,k_1},...,\varepsilon_{t,k_{n-1}})
\end{aligned}$$

where $[k_1,...k_n] \in \mathscr{C}_n$. Our definition of DIRF,

$$\int F_{t+h|t+1} d\left(F^*_{t+1|t} - F_{t+1|t}\right)$$

where $F^*_{t+1|t} - F_{t+1|t}$ is the distributional shock, that its identification implies identifying an element in $\mathscr{C}_n$.

Obtain other IRF

The mean IRF for the j-th variable is given by

$$MIRF_{j,h} := \int y_{j,t+h} dF^*_{Y_{j,t+h}|Z_t}(y_{j,t+h}) - \int y_{j,t+h} dF_{Y_{j,t+h}|Z_t}(y_{j,t+h}),$$

and the τ-th quantile IRF of the j-th element for $\tau \in (0,1)$ is given by

$$QIRF_{j,h}(\tau) := F^{*-1}_{Y_{j,t+h}|Z_t}(\tau) - F^{-1}_{Y_{j,t+h}|Z_t}(\tau),$$

where $F^{*-1}_{Y_{j,t+h}|Z_t}(\cdot)$ and $F^{-1}_{Y_{j,t+h}|Z_t}(\cdot)$ are the quantile functions as inverse of the j-th variable's distribution functions $F^*_{Y_{j,t+h}|Z_t}(\cdot)$ and $F_{Y_{j,t+h}|Z_t}(\cdot)$, respectively.

Univariate Distributional Regression

Univariate Distributional Regression was introduced by Foresi & Peracchi (1995) and further developed in Chernozhukov et al. (2013), and has become a comprehensive and flexible tool for modeling and estimating the entire conditional distribution

$$F(z|x) = \Lambda(T(x)'\gamma(z)) \quad \forall z \in \mathscr{Z}.$$

Typical choices of the link function is logit and probit.

Apply Univariate DR to each component

For the estimation of the j-th conditional distribution $F_{Y_{jt}|X_{jt}}$, we consider, for any $(y_j, x_j) \in \mathscr{Y}_j \times \mathscr{X}_j$,

$$F_{Y_{jt}|X_{jt}}(y_j|x_j) = \Lambda\big(\phi_j(x_j)^\top \theta_j(y_j)\big), \tag{3}$$

where $\Lambda : \mathbb{R} \to [0,1]$ is a known link function such as a logistic or probit function, $\phi_j : \mathscr{X}_j \mapsto \mathbb{R}^{d_j}$ is a transformation and $\theta_j(y_j)$ is a $d_j \times 1$ vector of unknown parameters specific to the location y_j.
The proposed model is sufficiently general given its flexibility in the manner of incorporating covariates are incorporated and the choice of link functions.

Estimating a collection of DR

In this study, we estimate Model (3) using a binary choice model for the outcome $\mathbf{1}\{Y_{jt} \leq y_j\}$ under the maximum likelihood framework, where $\mathbf{1}\{\cdot\}$ is the indicator function. Suppose that the sample size is T, and the estimators of the unknown parameters are defined as the maximizer of a log-likelihood function as follows:

$$\widehat{\theta}_j(y_j) = \arg\max_{\theta_j \in \Theta_j} \widehat{\ell}_{y,j}(\theta_j), \tag{4}$$

where $\Theta_j \subset \mathbb{R}^{d_j}$ is the parameter space and

$$\widehat{\ell}_{y,j}(\theta_j) := \frac{1}{T}\sum_{t=1}^{T}\left[\mathbf{1}\{Y_{jt} \leq y_j\} \ln \Lambda(\phi_j(X_{jt})^\top \theta_j) + \mathbf{1}\{Y_{jt} > y_j\} \ln\left(1 - \Lambda(\phi_j(X_{jt})^\top \theta\right.\right.$$

The conditional distribution estimator of Y_{jt} given $X_{jt} = x_j$ is given by

$$\widehat{F}_{Y_{jt}|X_{jt}}(y_j|x_j) := \Lambda\left(\phi_j(x_j)^\top \widehat{\theta}_j(y_j)\right). \tag{5}$$

Theorem

Suppose that Assumptions A1-A6 hold. Then, we have

$$\sqrt{T}\left(\widehat{\theta}(\cdot)-\theta(\cdot)\right) \rightsquigarrow \mathbb{B}(\cdot) \quad \text{in} \quad \times_{j=1}^{J} \ell^{\infty}(\mathscr{Y}_j)^{d_j}$$

where $\mathbb{B}(\cdot)$ *is a m-dimensional tight mean-zero Gaussian process over* $\mathscr{Y}$*. For any* $y, y' \in \mathscr{Y}$*, the covariance kernel of* $\mathbb{B}(\cdot)$ *is given by* $H(y)^{-1}\Sigma(y,y')H(y')^{-1}$*, where* $\Sigma(y,y') := \lim_{T\to\infty}\mathbb{E}[\widehat{\Psi}_y(\theta(y))\widehat{\Psi}_{y'}(\theta(y'))^{\top}]$ *and* $H(y) := \operatorname{diag}\left(\{H_j(y_j)\}_{j=1}^{J}\right)$. [1]

[1] We denote $H_j(y_j) := \nabla^2\ell_{y,j}(\theta_j(y_j))$ and $\widehat{\Psi}_y(\theta) := \left[\widehat{\Psi}_{y,1}(\theta_1)^{\top}\ldots,\widehat{\Psi}_{y,J}(\theta_J)^{\top}\right]^{\top}$ with $\widehat{\Psi}_{y,j}(\theta_j) := \sqrt{T}\nabla\widehat{\ell}_{y,j}(\theta_j)$.

Theorem

Suppose that Assumptions A1-A6 hold. Then,

(a) *we have*

$$\sqrt{T}\begin{pmatrix} \widehat{F}_{Y_{1t}|X_{1t}} - F_{Y_{1t}|X_{1t}} \\ \vdots \\ \widehat{F}_{Y_{Jt}|X_{Jt}} - F_{Y_{Jt}|X_{Jt}} \end{pmatrix} \rightsquigarrow \varphi'_{\theta(\cdot)}(\mathbb{B}) \quad \text{in} \quad \times_{j=1}^{J} \ell^{\infty}(\mathscr{Y}_j \times \mathscr{X}_j),$$

where $\mathbb{B}$ is the tight mean-zero Gaussian process.

(b) *if a map $\nu : \mathbb{S}_{\varphi} \mapsto \ell^{\infty}(\mathscr{Z} \times \mathscr{Y})$ is Hadamard differentiable at $(F_{Y_{1t}|X_{1t}}, \ldots, F_{Y_{Jt}|X_{Jt}})$ tangentially to $\varphi'_{\theta(\cdot)}(\mathbb{D})$ with the Hadamard derivative $\nu'_{F_{Y_{1t}|X_{1t}}, \ldots, F_{Y_{Jt}|X_{Jt}}}$, then*

$$\sqrt{T}\left\{\nu(\widehat{F}_{Y_{1t}|X_{1t}}, \ldots, \widehat{F}_{Y_{Jt}|X_{Jt}}) - \nu(F_{Y_{1t}|X_{1t}}, \ldots, F_{Y_{Jt}|X_{Jt}})\right\} \rightsquigarrow \nu'_{F_{Y_{1t}|X_{1t}}, \ldots, F_{Y_{Jt}|X_{Jt}}} \circ \varphi'_{\theta(\cdot)}(\mathbb{B})$$

in $\ell^{\infty}(\mathscr{Z} \times \mathscr{Y})$.

Alternative Nonparametric Method

1. Kernel density

$$f(y,z|x)=\frac{\sum_{i=1}^{n}K_{h_1}([y,z]-[y_i,z_i])K_{h_2}(x-x_i)}{\sum_{i=1}^{n}K_{h_2}(x-x_i)}$$

   - bandwidth
   - dimensionality

2. Quantile regression

$$Q_p^y(x,z)=x'\beta_p+z\gamma_p$$

   - less natural
   - computational costly.

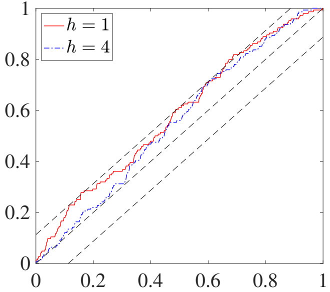
Macro: Empirical Questions

We use real GDP growth and National Financial Condition Index (NFCI) to gauge the economic and financial conditions. That is we elicit the joint distribution of $Y_t = [NFCI_t, GDP_t]$ given the past lagged observations $Z_t = [Y_{t-1}, Y_{t-2}, ..., Y_{t-p}]$

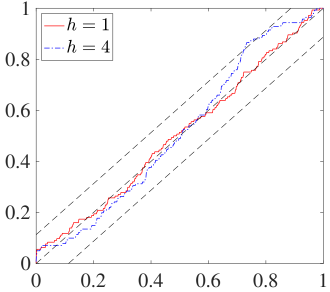
We answer two main questions in this section.

1. Do the joint distribution of GDP growth and NFCI conditional on past lagged information change during financial stress?
2. How does their joint distribution respond to a distributional shock in macro and financial conditions?

How does DR perform



(c) GDP PITS

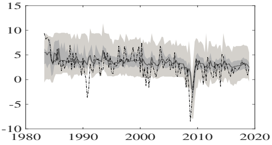


(d) NFCI PITS

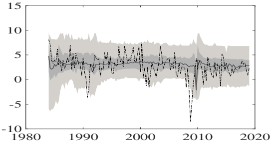
This figure reports the empirical CDF of the PITs by the DR approach for one-quarter-ahead (red solid line), and one-year-ahead (blue dotdash line), plus the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line) and the 5% confidence bands (dashed line) of the Rossi et. al(2019) PITs test.

How does DR perform

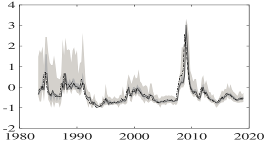
Figure: Out-of-sample Predicted Distributions: The predicted 5%, 25%, 50%, 75% and 95% quantiles (gray shadow) of the marginal distributions together with the realizations (dotdash lines) are plotted.



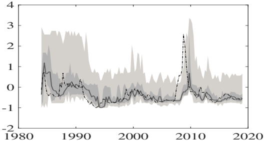
(a) GDP: one-quarter-ahead



(b) GDP: one-year-ahead



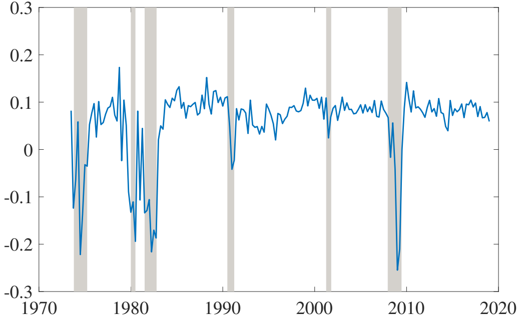
(c) NFCI: one-quarter-ahead



(d) NFCI: one-year-ahead

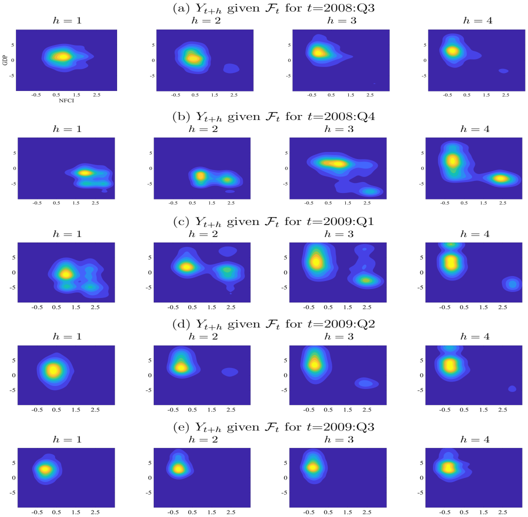
Correlation between the two variables

Figure: In-sample Contempreneous Correlation between NFCI and real GDP Growth



Notes: The figure plots the correlation coefficients based on the one-step-ahead forecasting distributions. Shaded areas indicate U.S. recessions.

Figure 3: Contour Plots of the Joint Distribution during the Great Recession



Notes: Each contour plot is constructed using bivariate kernel density estimation with a bandwidth of $(0.25, 0.7)$ based on 10,000 samples of the real GDP growth and NFCI generated from the corresponding multistep forecasting distribution. Different columns correspond to forecast horizons from one (leftmost column, $h = 1$) to four (rightmost column, $h = 4$) quarters, and different rows correspond to different conditioning information from 2008:Q2-Q3 (top row) to 2009:Q2-Q3 (bottom row).

Macro: Counterfactual DIRF

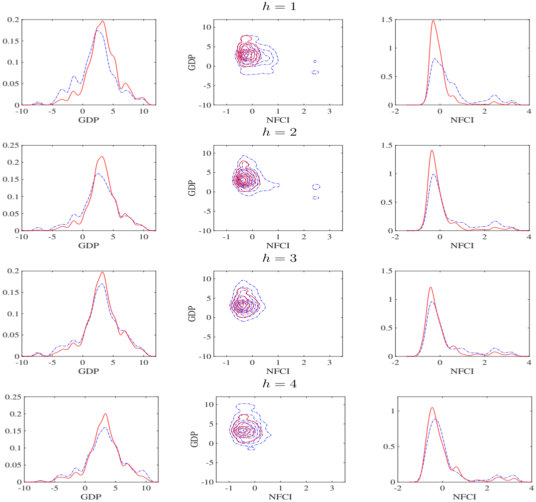
We first explore the effect that if the policy in 2008:Q3 were able to limit the possibility of tightening financial conditions during 2008:Q4, that is to perturb the conditional distribution of

$$NFCI_{2008:Q4} | \mathscr{F}_{2008:Q3}$$

to mean 0 and standard deviation 0.2 on $(-1.5, 2)$.

Figure 5: Distributional Response to the NFCI Impulse

Distributions of Y_{t+h} given $\mathcal{F}_t$ for t=2008:Q4



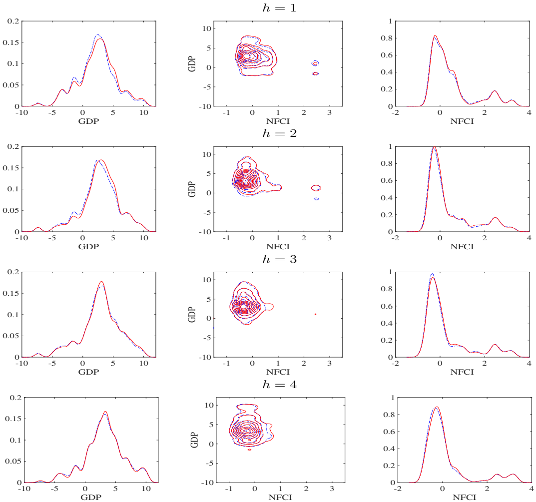
Notes: Different rows corresponding to forecasting distributions for forecast horizons from one (first row, $h = 1$) to four (last raw, $h = 4$) quarters, conditional on 2008:Q3-Q4. Refer to Figure 4 for details of the plots in each row.

Macro: Counterfactual DIRF

Now, we explore the effect that if the policy in 2008:Q4 were able to limit the possibility of low economic activity during 2008:Q4. Specifically, we keep the conditional distribution of the NFCI in 2008:Q4 as it is, and a counterfactual variable Y_{2t}^{*} following the truncated gamma distribution on $(0, 11)$ with scale parameter 6 and shape parameter 0.6 is considered.

Figure 7: Distributional Response to the GDP Impulse

Distributions of Y_{t+h} given $\mathcal{F}_t$ for t=2008:Q4



Notes: Refer to Figure 5.

Conclusion

We provide a semiparametric formulation for

$$Y_t | Z_t$$

for vector Y_t and Z_t's, that

- avoid imposing too rigid parametric assumptions
- the covariates are incorporated to influence the whole distribution instead of only affecting the distributions' location parameters
- by including the discrete outcomes as a covariate for the whole distribution of continuous ones, our model can malleable to capture intricate dependence structure.