Modeling and Forecasting Macroeconomic Downside Risk*

Davide Delle Monache [†] Andrea De Polis [‡] Ivan Petrella [§]

This draft: June, 2020

Abstract

We investigate the relation between downside risk to the economy and the financial markets within a fully parametric model. We characterize the complete predictive distribution of GDP growth employing a Skew-t distribution with time-varying location, scale, and shape, for which we model both secular trends and cyclical changes. Episodes of downside risk are characterized by increasing negative asymmetry, which emerges as a clear feature of the data. Negatively skewed predictive distributions arise ahead, and during, recessions, and tend to be anticipated by tightening of financial conditions. Indicators of excess leverage and household credit outstanding are found to be significant drivers of changes in the asymmetry of the forecast densities. Moreover, the great recession marks a neat shift in the unconditional distribution of GDP growth, which has featured a distinct negative skewness since then. The model delivers competitive out-of-sample (point and density) forecasts, improving upon standard benchmarks, especially due to financial conditions providing a strong signal of increasing downside risk.

Keywords: business cycle, downside risk, skewness, score driven model, financial conditions.

*The views expressed in this presentation are those of the authors and do not necessarily reflect the views of the Bank of Italy. Any errors and omissions are the sole responsibility of the authors.

[†]Bank of Italy. DG Economics, Statistics and Research. davide.dellemonache@bancaditalia.it

[‡]University of Warwick. Corresponding author: andrea.depolis.17@mail.wbs.ac.uk

[§]University of Warwick, CEPR. ivan.petrella@wbs.ac.uk

1 Introduction

The recent Great Financial Crisis (GFC) and the subsequent recession left policymakers with several new challenges to face. Amongst these, low interest rates and slow growth, as well as the downside risk to the economy have become priorities in Central Banks' agendas. Policy makers pursuing a "*plan for the worst, hope for the best*" approach rely on downside risk measures to quantify the optimality of the distribution of risk around modal forecasts, as measured by the balance of risk (see Kilian & Manganelli 2007). To this extent, the tail behaviour of predictive distributions that exhibit non-Gaussian features (e.g. non-zero skewness of leptokurtosis) can potentially provide leading signals for policymaking. Prominently, Adrian et al. (2019) uncover a novel negative correlation between financial conditions and the lower quantiles of the distribution of future real economic growth, suggesting that financial conditions may provide a better signal of non-Gaussian shocks to the economy.

In this paper, we introduce a novel flexible methodology to characterize and forecast the full conditional distribution of GDP growth, with a focus on the dynamics of downside risk, defined as the risk of observing substantially weaker growth levels relative to the historical growth average. To this purpose, we characterize the conditional distribution of economic growth with a versatile parametric Skew-t distribution with time-varying location, scale and shape parameters. As in Creal et al. (2013) and Harvey (2013), the dynamics of the time-varying parameters hinges upon the scaled score of the predictive likelihood function. The score-driven setting posits the distributional assumption to explicitly drive the time variation of the parameters. Furthermore, by augmenting the law of motion of the time-varying parameters with financial indicators, we also explore to what extent the predictive distribution of real economic growth responds to imbalances arising in the financial sector. Within this setting, we provide clear evidence that financial conditions consistently predict downside risk episodes. Indicators of household financial distress, as well as equity prices and credit spreads emerge as reliable predictors of increasing downside risk to economic growth, while conveying little information about future growth level or its uncertainty.

Assessing the degree of asymmetry of business cycle fluctuations can be challenging, as standard tests often reject the null hypothesis of zero skewness (see Carriero et al. 2020). In contrast, the score-driven setting provides a suitable ground to formally test for the significant time variation of the full distribution of GDP growth. We find a strong evidence in support of the presence of conditional asymmetry, despite it not being unconditionally different from zero.

Unlike alternative frameworks, the versatile setting we introduce allows to fully capture business cycle fluctuations by accounting for both secular and transitory changes in the distribution of GDP growth. Therefore, the model is able to account for the well documented changes in the long-run GDP growth (see, e.g. Antolin-Diaz et al. 2017, Cette et al. 2016), as well as capturing the volatility reduction observed starting from the mid-80s (see, e.g. McConnell & Perez-Quiros 2000, Stock & Watson 2002). Furthermore, in line with Jensen et al. (2020), we provide evidence of a decreasing trend in the skewness of GDP growth since the early 90s, highlighting that GDP fluctuations are characterized by a marked negative skewness since the early 2000s. At the onset of recessions, business cycle swings exhibits significant negative skewness, while normal times are characterize by closer-to-Gaussian and positively skewed GDP growth distributions. We also document to what extent changes in this latter property of business cycle fluctuations have contributed to shape the expected value and variance of GDP growth. Specifically, we highlight that the deepening negative asymmetry observed since the early 2000s accounts for a large share of the fall in long-run growth over the same period. Moreover, we relate spikes in volatility observed during recessions (see, e.g. Jurado et al. 2015) to increasing downside volatility not being entirely matched by upside volatility.

When we introduce indicators of financial conditions as drivers of the parameters, we recover predictive densities that better approximate future output levels (both in-sample and out-of-sample). Financial indicators contribute in enhancing the timeliness and accuracy of the predicted distribution, by providing a better signal of non-Gaussian shocks to the economy.

Evaluating and measuring financial stress has always been a daunting task. Traditional spread measures (see Bernanke & Gertler 1995) as well as the slow building up of leverage have recently been identified as important risk factor to take into account (Drehmann et al. 2010, Jordà et al. 2013). Brave & Butters (2011) introduce a measure of aggregate financial distress using data relative to money, debt and equity markets and the banking sector. This index, the Chicago Fed National Financial Condition Index (NFCI), arises as the first factor out of 105 individual predictors, which track risk, leverage and credit sector imbalances (see Brave & Butters 2012, for further details).

We rely on Bayesian methods to estimate the models. This allows us to fully account for the underlying parameter uncertainty when producing forecasts, as well as providing great flexibility in dealing with a potentially large set of predictors. We first evaluate the gains associated with using broad subindices of financial conditions: credit, risk, leverage and nonfinancial leverage. The latter emerges as a leading indicators of the dispersion, whereas skewness relates to risk

and credit spreads. Overall, financial conditions significantly contribute to increase the forecast accuracy of our model, in particular around turning points, allowing to better predict the odds of forthcoming recessions.

Whereas these aggregate measures succeed in summarising a large amount of data, concerns that information relevant for assessing risk can remain unaccounted for is a novel issue, recently discussed in Galvão & Owyang (2018) and Plagborg-Møller et al. (2020). This opens the question of whether using an even larger number of disaggregated financial indicators can further improve our understanding of shifts in downside risk. While some indicators might provide a better signal for the central tendency of economic growth, or its overall uncertainty, others instead might convey relevant guidance for assessing downside risk developments. To answer this question we perform a variable selection exercise based on the "*shrink-then-sparsify*" approach of Hahn & Carvalho (2015) using the full set of NFCI indicators. While Giannone et al. (2018) warn against the use of sparse modelling without critical judgement, as sparsity can be obtained as a result of strong *a priori* beliefs (parameter setting) while not being a robust feature of the data, we highlight the purely data-driven nature of the approach we take. Specifically, shrinkage is achieved through Horseshoe (HS) priors (Carvalho et al. 2010), whereas sparsity is achieved via the Signal Adaptive Variable Selector (SAVS) of Ray & Bhattacharya (2018). Variables tracking developments in credit and leverage markets emerge as recurrent predictors of downside risk. In particular, indicators of the amount of traded security over GDP and corporate bond spread provide useful to gauge the increasing downside risk associated with the period. The ability of the model of processing the signal from the large panel of financial predictors proves particularly useful in providing gains in both density and point forecasts in particular for the latest financial crisis and the subsequent recession.[1]

This paper builds on recent literature exploring the relation between real economic activity and financial conditions (Jordà et al. 2013, Gertler & Gilchrist 2018, Jensen et al. 2020, among others). Giglio et al. (2016) and Adrian et al. (2019) show that measures of systemic risk are more informative about future economic downturns, as they better predict lower quantiles of the conditional distribution of real output.[2] Our framework is similar in spirit to the one in Adrian

[1]This is in line with the observation in Alessi et al. (2014) which have argued that the Central Banks' failure to predict the 2007-2009 downturn can be ascribed to the inability of reading the signals from deteriorating financial conditions in part of the financial markets. What we highlight is that these additional information would have been particularly useful in picking up the build up in downside risk ahead of the recession.

[2]Similarly, De Nicolò & Lucchetta (2017) uphold the superior forecast accuracy of factor-augmented quantile projections to predict economic growth tail risk.

et al. (2019), with the key difference that, using a score driven framework, we directly model the parameters of a skew-Student-t distribution. In contrast to the two-step quantile approach, that further requires to fit a distribution to the estimates, we propose a parametric model which allows great flexibility. Unlike in Adrian et al. (2019) and Giglio et al. (2016), our framework allows us to parsimoniously account for the low frequency variation in GDP growth distribution, to capture features as the Great moderation period, the fall in long-run GDP growth and the increasingly negatively skewed business cycle. Differently from Plagborg-Møller et al. (2020), which impose a linear relation between financial predictors and past GDP growth, our score driven filter updates the parameters according to a complex moving average of all the lags of financial predictors, as well as to a nonlinear function of the prediction error. While their specification remains nested into our framework, we are able to capture smoother variation of the scale and shape parameters of the distribution, characteristic of the cyclical fluctuations of economic growth.

From a methodological standpoint, we propose an adaptive algorithm to fit and forecast GDP growth conditional moments. In order to capture downside risk episodes, we allow for Skew-t innovations à la Gómez et al. (2007).[3] Therefore we allow for, but do not impose, skewness and fat tails of the conditional density of GDP growth. In fact, our setting encompasses, as limiting cases, the symmetric Student-t (see, e.g. Delle Monache & Petrella 2017), the skew-Normal and the Gaussian densities (as in Creal et al. 2013). Most importantly, we put forward a two component specification for the time varying parameters of the model which allows to take into account both, secular and cyclical changes of the underlying distribution. Within the score-drive setting, we are thus able to generate a wide spectrum of adaptive weighting schemes suitable for the precise characterization of GDP growth downside risk dynamics over time. Furthermore, we build upon the work of Zhang & Schwaab (2016) and Harvey (2013) by allowing exogenous predictors in the updating equations of the time-varying parameters. Moreover, relaying on Bayesian estimation methods we put forward a simple approach to jointly tackle parameters' proliferation, and increasing estimation uncertainty when producing forecasts.

The remainder of the paper is organized as follows. The next Section provides the motivation for our modelling framework, later detailed in Section 3. Section 4 provides a description of the data. The estimation methodology, the forecasting procedure and the shrinkage approach are reported in Section 5, along with in-sample results. Section 6 reports the out-of-sample forecast

[3]Lucas & Zhang (2016) use the same distribution in the score-drive setting to estimate and forecast the value at risk of stock and exchange rate returns

and downside risk prediction evaluation. In Section 7 we investigate the the predictive ability of the large set of financial indicators. Section 8 concludes.

2 Motivating evidence

Assessing the degree of skewness of GDP growth is a notoriously challenging task (Neftci 1984, Morley & Piger 2012). When measured over the sample 1973-2018, we obtain a negative sample skewness of -0.36. According to the Bai & Ng (2005) test, the low precision of this skewness estimate suggests that one cannot reject the null of symmetry.[4] However, the absence of skewness in the unconditional distribution does not necessarily imply the conditional distribution being symmetric as well (Carriero et al. 2020). In fact, the difficulty of getting significant skewness estimates can potentially reflect the dynamic evolution of business cycles properties over time. For instance, Jensen et al. (2020) note that the reduction in volatility during the Great Moderation period is associated with a more pronounced asymmetry in business cycle fluctuations. Moreover, using Bai & Ng (2005) test over different rolling windows of 1 to 5 years, the test often rejects the null of symmetry. The erratic behaviour of the test statistics points out at substantial movements in the skewness of GDP growth over the sample, suggesting that prolonged expansions are characterized by significant positive skewness, while contractions being associated with significant negative skewness.

In this Section, we report a number of tests for the presence of time varying asymmetry. Harvey (2013, Section 2.5) and Harvey & Thiele (2016) highlight that test for the time variation of parameters can be constructed from the score of the conditional likelihood estimated under the null hypothesis of no time variation. Testing for time varying parameters in this framework requires that we start from a benchmark model under the null of no time variation of the parameter of interest. Here, we assume that GDP growth can be characterized by a Skew-t distribution, where a shape parameter pins down the degree of asymmetry in the distribution. We consider two alternatives. In the first case, we only allow time variation of the location, whereas in the second case, we let both the location and volatility to vary over time. With the latter case, we make sure that the presence of time varying asymmetry does not simply stem from the presence of countercyclical volatility, or the downward trend in volatility over the sample considered. The location is specified as an AR(2) with a stochastic (random walk)

[4]The results are in line with the ones in Carriero et al. (2020). Other measures of skewness, such as Galton (quantile) or Kelly (decile) skewness coefficients, deliver similar evidence.

Table 1: Time-varying parameters tests

	time-varying location			*time-varying location & scale*		
	Q	Q^*	N	Q	Q^*	N
$Scale^2$	70.12^{***}	71.22^{***}	1.91^{***}			
$Shape$	46.95^{***}	47.68^{***}	0.41^{*}	10.76^{***}	10.93^{***}	0.97^{***}

Note: Specification: $y_t = skt_\nu(\mu_t, \sigma^2, \varrho),\ \mu_t \sim RW + AR(2)$.
Q is the portmanteau test, Q^* is the Ljung-Box extension and N corresponds to the Nyblom test. The first two tests are distributed as a χ^2 with 1 degree of freedom, the Nyblom test statistics is distributed as a Cramer von-Mises distribution with 1 degree of freedom. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$

trend. Both processes are driven by the score of the skew-t distribution.[5] Appendix A provides additional details on the specifics of the tests. In Table 1 we report the test statistics for the score-based portmanteau (Q), Ljung-Box (Q^*) and Nyblom (1989) (N) tests.[6] In both cases the null hypothesis of constant shape is clearly rejected against the alternative that the shape parameters varies over time for all three tests. Note, howerver, that the rejection of the Nyblom test, which under the alternative hypothesis assumes that the shape parameter follows a martingale process, suggests that the parameter is likely to be highly persistent. Starting from this evidence, in the next Section we introduce an observation driven model, where the conditional distribution of GDP growth features time-varying asymmetry.

3 A time varying Skew-t model for GDP growth

Let y_t denote the annualized quarter-on-quarter GDP growth at time t. We assume its conditional distribution can be characterized by the Skew-t of Gómez et al. (2007), with time-varying location μ_t, scale σ_t, and shape ϱ_t parameters, and constant degrees of freedom ν. Specifically,

$$y_t = \mu_t + \sigma_t \varepsilon_t, \qquad \varepsilon_t \sim Skt_\nu(0, 1, \varrho_t), \tag{1}$$

[5]The score-driven updating, based on the score of the log-likelihood function, is introduced in the next Section.

[6]The lag length for the portmanteau and Ljung-Box tests are selected following Escanciano & Lobato (2009).

with the scale parameter $\sigma_t > 0$, the degrees of freedom $\nu > 2$, and the asymmetry parameter $\varrho_t \in [-1, 1]$, such that when it attains negative (positive) values, the distribution features positive (negative) skewness. The conditional log-likelihood function of the observation at time t is:

$$\ell_t = \log p(y_t|\theta, Y_{t-1}) = \log \mathcal{C}(\eta) - \frac{1}{2}\log \sigma_t^2 - \frac{1+\eta}{2\eta}\log\left[1 + \frac{\eta\varepsilon_t^2}{(1 - sgn(\varepsilon_t)\varrho_t)^2\sigma_t^2}\right], \qquad (2)$$

with $\eta = \frac{1}{\nu}$ being the inverse of the degrees of freedom, $\mathcal{C}(\eta) = \frac{\Gamma\left(\frac{1+\eta}{2\eta}\right)}{\sqrt{\frac{\pi}{\eta}}\Gamma\left(\frac{1}{2\eta}\right)}$, $\Gamma(\cdot)$ is the Gamma function, and $sgn(.)$ is the sign function. The vector θ collects all the static parameters, and $Y_{t-1} = \{y_j\}_{j=1}^{t-1}$ is the information set up to time t. For $\varrho_t = 0$ we have the symmetric Student-t distribution, for $\eta \to 0$ we retrieve the epsilon-skew-Gaussian distribution of Mudholkar & Hutson (2000), while when both conditions hold, the distribution collapses to a Gaussian density.[7]

Following the score-driven framework of Creal et al. (2013) and Harvey (2013), we postulate the driving force for the time-varying parameters to be the score of the predictive likelihood.[8] Let f_t collect the vector of time-varying parameters of interest, which follows the score-driven law of motion:

$$f_{t+1} = c + Af_t + BX_t + Cs_t, \qquad (3)$$

where A contains autoregressive parameters governing the persistence of the updating, B collects loadings on explanatory variables X_t, and C collects the score loadings adjusting the speed of the updating, given the prediction error. The scaled score, s_t is defined as $s_t = \mathcal{S}_t\nabla_t$, where:

$$\nabla_t = \frac{\partial \ell_t}{\partial f_t}, \qquad \mathcal{S}_t = \mathcal{I}_t^{-\frac{1}{2}} = \mathbb{E}\left(-\frac{\partial^2 \ell_t}{\partial f_t \partial f_t'}\right)^{-\frac{1}{2}}, \qquad (4)$$

with ∇_t being a vector of scores, namely the gradient of the likelihood function ℓ_t with respect to the dynamic parameters, while the scaling matrix $\mathcal{S}_t$ is proportional to the square root of the Moore-Penrose pseudo-inverse of the Information matrix, $\mathcal{I}_t$. The resulting scaled score is a martingale difference sequence with conditional variance equal to the identity matrix.[9] The

[7] As opposed to the Skew-t distribution of Azzalini & Capitanio (2003), adopted by Adrian et al. (2019) in the context of GDP growth, the Skew-t distribution of Gómez et al. (2007) retrieve an information matrix which is always non-singular and can be inverted, provided that $|\varrho_t| < 1$.

[8] Blasques et al. (2015) formalize the theoretical optimality of the score-drive class; parameter updating based on the score of the predictive likelihood always reduces the local Kullback-Leibler divergence between the true conditional density and the model implied one, even under severe misspecification.

[9] Creal et al. (2013) suggest alternative scaling matrix as the identity matrix or $\mathcal{I}_t^{-1}$. The former choice leaves the score unscaled, with $Var(s_t) = \mathcal{I}_t$, whereas the latter scaling matrix sets the variance

parameter updating is defined by the steepest ascent direction for improving the model's local fit, in terms of its likelihood. In fact, the direction and magnitude of the updating are dictated by the steepness and curvature of the likelihood function relative to the position of the parameters.[10] Therefore, given equations (1)-(4), the resulting model belongs to the observation-driven class, for which the trajectories of the time-varying parameters are perfectly predictable given past information and the log-likelihood function being available in closed form (Cox 1981). The following Proposition provides the closed form expressions for the gradient and the associated information matrix.

Proposition 1. *Given the model specification* (1) *and the likelihood in* (2)*, the elements of the gradient* ∇_t*, with respect to location, scale and asymmetry, are:*

$$\frac{\partial \ell_t}{\partial \mu_t} = \frac{1}{\sigma_t^2} w_t \varepsilon_t, \qquad \frac{\partial \ell_t}{\partial \sigma_t^2} = \frac{1}{2\sigma_t^4}(w_t \varepsilon_t^2 - \sigma_t^2), \qquad \frac{\partial \ell_t}{\partial \varrho_t} = -\frac{1}{\sigma_t^2} \frac{sgn(\varepsilon_t)}{(1 - sgn(\varepsilon_t)\varrho_t)} w_t \varepsilon_t^2, \tag{5}$$

where $w_t = \frac{(1+\eta)}{(1-sgn(\varepsilon_t)\varrho_t)^2+\eta\zeta_t^2}$ *and* ζ_t *denotes the standardized innovation,* $\zeta_t = \varepsilon_t/\sigma_t$*. The associated information matrix reads as follows:*

$$\mathcal{I}_t = \begin{bmatrix} \frac{(1+\eta)}{(1+3\eta)(1-\varrho_t^2)\sigma_t^2} & 0 & -\frac{4c(1+\eta)}{\sigma_t(1-\varrho_t^2)(1+3\eta)} \\ 0 & \frac{1}{2(1+3\eta)\sigma_t^4} & 0 \\ -\frac{4c(1+\eta)}{\sigma_t(1-\varrho_t^2)(1+3\eta)} & 0 & \frac{3(1+\eta)}{(1-\varrho_t^2)(1+3\eta)} \end{bmatrix} \tag{6}$$

Proof. See Appendix A. □

Proposition 1 highlights the central role of the re-weighting of the standardized prediction errors (and its square values) for the updating of the time varying parameters. The weights w_t depend on the number of degrees of freedom and on the vector f_t, estimated conditional on the information available at time $t-1$. Thus, extreme standardized innovations are penalized depending upon the thickness of the tails, as well as by conditional asymmetry and volatility. The bottom right panel of Figure 1 illustrates the reweighing of the prediction error implied by the scaled scores, for a fixed value of the scale, 5 degrees of freedom and three alternative levels of the shape parameter. When the asymmetry parameter is zero (red lines), the weights display the classic outlier-discounting typical of the observation driven models with Student-t distributions (see, e.g., Harvey & Luati 2014, Delle Monache & Petrella 2017). When the distribution is

of the scaled score to the inverse of Information matrix.

[10] The resulting filter can be thought of as a stochastic counterpart of the Newton-Raphson maximization algorithm.

positively (negatively) skewed, i.e. for $\varrho_t < 0$ ($\varrho_t > 0$), negative (positive) prediction errors, less likely in expectation, command a larger update of parameters of the model when observed. This asymmetric treatment of the signal of the prediction error is more pronounced as the skewness of the distribution grows larger (i.e. $|\varrho_t| \to 1$).

In order to ensure the scale σ_t to be positive and the shape ϱ_t to lie within the unit circle, we apply time-invariant, invertible and differentiable *"link functions"* to these parameters. In practice, we model $\gamma_t = \log \sigma_t$ and $\delta_t = \operatorname{arctanh} \varrho_t$, so that the vector of time-varying parameters becomes $f_t = (\mu_t, \gamma_t, \delta_t)'$. Moreover, we follow Lucas & Zhang (2016) and scale the score only using the diagonal elements of the information matrix.[11] Therefore, the associated scaled score vector is

$$s_t = (J_t' diag(\mathcal{I}_t) J_t)^{-\frac{1}{2}} J_t' \nabla_t = \begin{bmatrix} s_{\mu t} \\ s_{\gamma t} \\ s_{\delta t} \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{(1+3\eta)(1-\varrho_t^2)}{(1+\eta)}} w_t \zeta_t \\ \sqrt{\frac{(1+3\eta)}{2}} (w_t \zeta_t^2 - 1) \\ -sgn(\varepsilon_t) \sqrt{\frac{(1+sgn(\varepsilon_t)\varrho_t)(1+3\eta)}{3(1-sgn(\varepsilon_t)\varrho_t)(1+\eta)}} w_t \zeta_t^2 \end{bmatrix}, \tag{7}$$

where $J_t = \frac{\partial(\mu_t, \sigma_t^2, \varrho_t)}{\partial(\mu_t, \gamma_t, \delta_t)'} = diag\big(1; 2\sigma_t^2; (1-\varrho_t^2)\big)'$ is the Jacobian matrix associated to the link functions. Equation (7) highlights that the updating of the shape parameter is substantially different than the ad-hoc specification chosen in Hansen (1994).[12]

Figure 1 plots the scaled scores against the standardized innovations, for the location, scale and shape parameters, respectively. Negative prediction errors are associated with a negative update in the location and a positive update in the shape (i.e. the distribution becoming more left skewed). The opposite is true for positive prediction errors, whereas in both cases the scale is updated upward. For extreme values of the prediction errors, parameters' update becomes inelastic to the standardized innovations, as the scaled scores converge to their limiting values. Most importantly, the asymmetry of the distribution plays a key role in the translation of the standardized prediction error into a signal for the update of the distribution's parameters. When the distribution is left skewed (i.e. $\varrho_t > 0$) a positive (negative) prediction error leads to a more (less) pronounced update for the parameters, whereas the opposite is true in the case of a

[11] Specifically, we do not allow the score associated to the location parameter to affect the shape parameter and vice-versa. This increases the computational stability of the scaled score vector (see Lucas & Zhang 2016).

[12] Note also that the skewness of the process is implicitly updated using information from the standardized prediction error and its squared value, without relying on higher powers, which in practice can be quite volatile and produce unreliable estimates. This stands in contrast with the specification of conditional skewness in Harvey & Siddique (1999), which instead uses the cube of the prediction error to update the parameter.

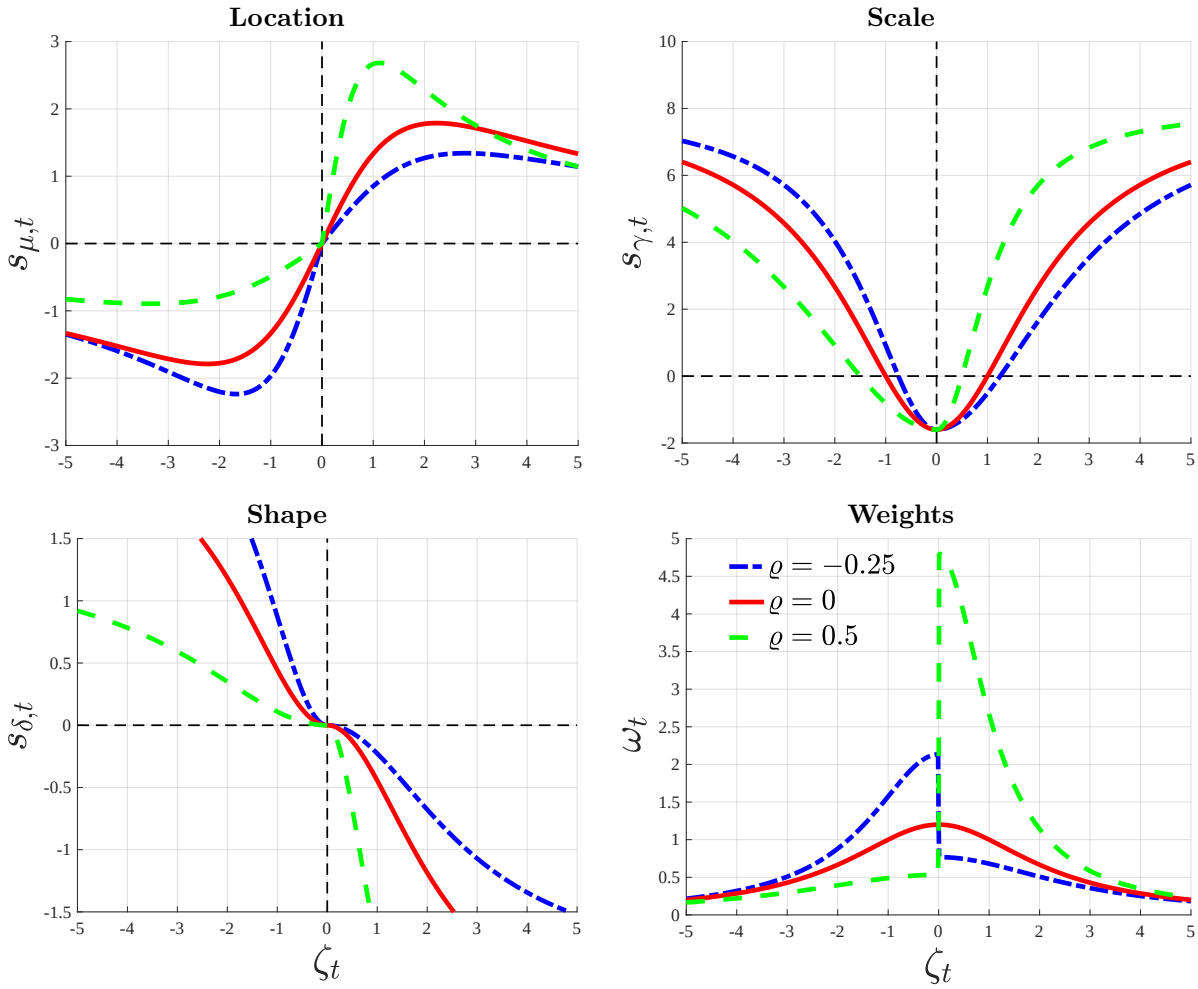

Figure 1: Prediction error signal re-weighting

Note: The figures plot the weighting scheme implied by the scaled scores for different values of the prediction error ε_t. From left to right we plot the influence function for the location, scale and shape parameters, for three different values of the asymmetry parameter: -0.5 (blue), 0 (red) and 0.7 (green). The scale parameter is set to 1 with 5 degrees of freedom.

positively skew distribution. This property of the updating function is a direct consequence of the stark asymmetry in the weights, and makes the model particularly reactive to turning points in GDP growth.

3.1 Secular and cyclical variation of GDP growth's distribution

When modelling the conditional distribution of GDP growth it is important to allow for both cyclical and secular movements of the central moments. In fact, a number of paper have documented that over the sample under analysis, GDP growth have experienced significant changes in the long run mean (see, e.g. Antolin-Diaz et al. 2017, Cette et al. 2016), as well

as shifts in the volatility (McConnell & Perez-Quiros 2000, Stock & Watson 2002), and in the skewness of the distribution (Jensen et al. 2020) for the post '80s data (i.e. the Great Moderation). At the same time, others have documented that the volatility of GDP growth is countercyclical (Jurado et al. 2015). To account for these feature of the data, we postulate a two-component specification for the time-varying parameters, in the spirit of Ding & Granger (1996). Specifically, each parameter consists of a (stationary) short-run component, identified with a $\tilde{\cdot}$, tracking transitory movements of the parameters' dynamics, and a permanent long-term component, marked by a $\bar{\cdot}$. We posit a random walk updating for the latter, as it allows higher flexibility in the presence of breaks in the level of the parameter. Moreover, we allow a set of predictors, X_t, to affect the dynamics of the short run components.

The location parameter is a linear combination of a trend component, $\bar{\mu}_t$, and a stationary component, $\tilde{\mu}_t$:

$$\mu_t = \bar{\mu}_t + \tilde{\mu}_t \tag{8}$$

$$\bar{\mu}_{t+1} = \bar{\mu}_t + \varsigma_\mu s_{\mu t} \tag{9}$$

$$\tilde{\mu}_{t+1} = \phi_{\mu,1}\tilde{\mu}_t + \phi_{\mu,2}\tilde{\mu}_{t-1} + \beta'_\mu X_t + \kappa_\mu s_{\mu t} \tag{10}$$

where the AR(2) specification for $\tilde{\mu}_t$ aims to recover the characteristic humped shape of the data (see, e.g. Chauvet & Potter 2013, Morley et al. 2013).

For the log-scale parameter, $\gamma_t = \log \sigma_t$, we posit the following additive two-component model:

$$\gamma_t = \bar{\gamma}_t + \tilde{\gamma}_t \tag{11}$$

$$\bar{\gamma}_{t+1} = \bar{\gamma}_t + \varsigma_\gamma s_{\gamma t} \tag{12}$$

$$\tilde{\gamma}_{t+1} = \phi_\gamma \tilde{\gamma}_t + \beta'_\gamma X_t + \kappa_\gamma s_{\gamma t} \tag{13}$$

that implies a multiplicative relation, $\sigma_t = \bar{\sigma}_t \tilde{\sigma}_t$ (with $\bar{\sigma}_t = \exp \bar{\gamma}_t$, $\tilde{\sigma}_t = \exp \tilde{\gamma}_t$), as put forward by Engle & Rangel (2008). With this specification, the unconditional volatility corresponds to the long-run component $\lim_{h\to\infty} \mathbb{E}[\gamma_{t+h}] = \bar{\gamma}_t$. Similarly, for the transformed shape parameter, $\delta_t = \operatorname{arctanh} \varrho_t$, we posit an additive two-component model, in which the permanent component $\bar{\delta}_t$ follows a random walk, while the transitory component, $\tilde{\delta}_t$, follows and autoregressive process

of the first order:

$$\begin{aligned} \delta_t &= \bar{\delta}_{t+1} + \tilde{\delta}_{t+1} && (14) \\ \bar{\delta}_{t+1} &= \bar{\delta}_t + \varsigma_\delta s_{\delta t} && (15) \\ \tilde{\delta}_{t+1} &= \phi_\delta \tilde{\delta}_t + \beta'_\delta X_t + \kappa_\delta s_{\delta t}. && (16) \end{aligned}$$

that implies $\varrho_t = \tanh(\bar{\delta}_t + \tilde{\delta}_t)$. Therefore, the resulting vector of time-varying parameters is equal to $f_t = (\bar{\mu}_t, \tilde{\mu}_t, \bar{\gamma}_t, \tilde{\gamma}_t, \bar{\delta}_t, \tilde{\delta}_t)'$.

Plagborg-Møller et al. (2020) consider a time varying Skew-t specification for GDP growth as in Adrian et al. (2019), but they specify the time varying parameters as a linear function of the set of predictors including lagged GDP growth.[13] A specification similar to their can be retrieved assuming $\varsigma_\mu = \phi_{\mu,1} = \phi_{\mu,2} = \kappa_\mu = \varsigma_\gamma = \phi_\gamma = \kappa_\gamma = \varsigma_\delta = \phi_\delta = \kappa_\delta = 0$ (and including an intercept and past GDP growth among the predictors X_t, for each parameter). In contrast our specification allows for both cyclical and secular shifts in the parameters. It is worth noting that the autoregressive structure of the cyclical components makes them functions of the discounted values of all past predictors and (scaled) scores, where these latter are themselves nonlinear functions of past data.

3.2 Estimation

The parameters of the model and the associated conditional distribution of GDP growth are estimated using Bayesian methods. Maximum likelihood estimates are used to initialize an adaptive Random-Walk Metropolis-Hastings (ARWMH) algorithm (Haario et al. 1999), combined with rejection sampling.[14] Credible sets for both static and time-varying parameters are obtained from the empirical distribution functions arising from the resampling.

We set Minnesota-type Normal priors for the AR coefficients of the cyclical parameters, centered around a high persistence value.[15] For the location AR parameters, we also introduce a prior on the sum of coefficients, in line with Doan et al. (1984) and Sims & Zha (1998). Similarly, we assume Normal prior for the loadings associated to the predictors. These coefficients are centered around zero, with tight scales in order to avoid overfitting of the parameters, in the

[13] To be precise, they consider the Skew-t in Azzalini & Capitanio (2003) and also do not impose any constraint on the shape parameter.

[14] Appendix C provides an extensive description of the sampling algorithm as well as the details on the exact prior specification for the parameters.

[15] In addition, we reject draws associated with explosive behavior for the cyclical components.

fashion of L_2 (Ridge) regularization. The prior distribution of the score loadings follow inverse gamma distributions, as we expect these parameters to be positive. Lastly, we also assume an inverse gamma prior for η.[16]

3.3 Forecast

For any draw of the model parameters, θ, the last step of the the observation driven filter (3) provides the optimal one-step-ahead prediction of the parameters of interest (Cox 1981). These values can then be used to retrieve the one-step-ahead prediction density for GDP growth, $p(y_{T+1}|\theta) = skt_\nu(f_{T+1}(\theta))$, which allows us to draw the forecast of interest as $p(y_{T+1}) = \int p(y_{T+1}|\theta)d\theta$.

For multiple steps ahead forecasts additional complications arise from the necessity to sample the score, and the dependence of the forecasts on the predicted values of the conditioning variables. In Section 6, we produce forecasts keeping the conditioning variables fixed to their last observations, akin to assuming a random walk specification for their law of motion.[17] As for the the score vector, we follow Koopman et al. (2018) and adopt a *"bootcasting"* algorithm to sample multiple $h-1$ dimensional vectors of the scores from the (scaled) score vector obtained in the estimation. Therefore, for a given (bootstraped) draw of the score, and assuming $X_{T+h} = X_T$, the score filter (3) can be used to obtain f_{T+h} and thus compute $p(y_{T+h}|\theta, X_{T+h} = X_T) = skt_\nu(f_{T+h}(\theta, X_{T+h} = X_T))$. The h-step ahead forecast reads $p(y_{T+h}) = \int p(y_{T+h}|\theta, X_{T+h} = X_T)d\theta$.

4 Data

The model presented in the previous section is used to investigate downside risk in economic activity, and how prevailing financial conditions improve its prediction. In particular, we focus on (real) GDP growth in the US, and we measure financial conditions using the Chicago FED National Financial Condition Index (NFCI), and its subcomponents (Brave & Butters 2011). The NFCI is a weekly index tracking the status of money markets, debt and equity markets, and the banking sector (comprehensive of the *"shadow"* banking sector). The contribution of

[16] Juárez & Steel (2010) recommend $\nu \sim \mathcal{G}(2, 0.1)$ as a general prior specification for the degrees of freedom parameter of a Student-t distribution. This specification covers a wide range of plausible values, featuring a prior mean of 20, median of 10 and variance of 200.

[17] Therefore taking advantage of the fact that the underlying predictors in our applications are persistent enough, so that holding them fixed over the forecast window results in small loss of overall predicting ability.

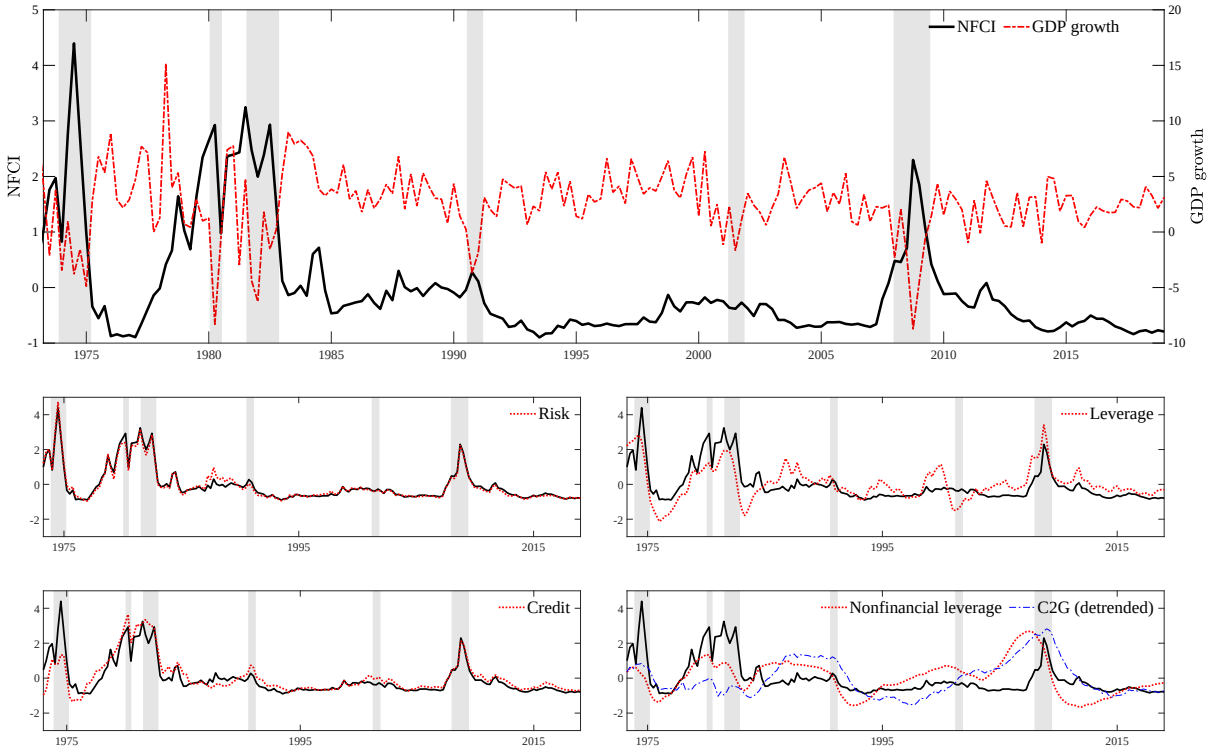

Figure 2: National Financial Condition Index and subcomponents

Note: The top panel illustrate the evolution of the NFCI; bottom panels plots the risk, leverage and credit subcomponents, respectively. Shaded bands represent recessions.

these sectors maps into three subindices, each of which gauges financial conditions in terms of risk, leverage, and credit (refer to Brave & Butters 2012, for further details). We use quarterly data on economic activity and financial conditions over the period 1973Q1 to 2019Q1, where the starting date for the analysis is dictated by the availability of the NFCI. As new information is available, these indices are standardized. Thus, when an index takes a value of 1 (-1), it signals that the financial conditions are one standard deviation tighter (looser) than their historical average. Consistently with the concept of financial tightening, the measure relating to the risk subcomponent positively contributes to the NFCI, reflecting increasing risk premia. In contrast, measures of leverage and credit receive negative weights in response to lower asset values and declining liquidity in the financial sector. The top panel of Figure 2 plots the NFCI against GDP growth, whereas the bottom panels report the NFCI subcomponents. The NFC index (and the three subcomponents) spikes during recessions, as indicated by the gray-shaded bands, and displays a clear negative correlation with GDP growth, suggesting that periods of financial overheating coincide with severe economic troughs, as documented by Adrian et al. (2019). The NFCI closely tracks the risk subcomponent, due to high weight being attached to some of the risk variables such as the VIX index.This is consistent with the observation that volatility of the

stock market is an accurate predictor of financial instability, as argued by (see, e.g. Bekaert & Hoerova 2014). The leverage component shows strong procyclicality and tends to pick up on high values before other indices, consistently with the findings of Adrian & Shin (2010) and Jordà et al. (2013). Lastly, measures of credit spread and credit risk are contained in the risk and credit components, as identified by Krishnamurthy & Muir (2017).

Alongside these subindices, the Chicago FED also produces the nonfinancial leverage index (bottom right panel of Figure 2). This series tracks developments in the nonfinancial credit market using data on household and nonfinancial business leverage. Household data receive weights roughly 1.5 times higher than the latter. Mian & Sufi (2010) argue that the build up of financial instability leading to the Great Financial Crisis was mainly due to an "over-leveraged" household sector. Similarly, Jensen et al. (2020) observe that increasing households and firms leverage anticipates a deepening of business cycle skewness. A closer examination of the panel suggests that the massive deleveraging that took place starting in early 2008 coincides with the beginning of the recession in that same year. This measure can thus be intended as an "early-warning" signal for economic downturns. In addition, this index is a clear leading indicator of the HP-filtered credit-to-GDP ratio measure (blue line in the bottom right panel of Figure 2), put forward as a leading signal of financial distress (Drehmann et al. 2010, Jordà et al. 2017).[18]

Real time data: In sections 6 and 7 we evaluate the performance of our models and compare to a number of alternatives in an out-of-sample forecasting exercise. For this exercise, we use real time data on GDP, provided by the Philadelphia FED's Real Time Data Research Center. This dataset contains GDP observations available at a particular vintage date, that is, GDP values available in the second week of the middle month of each quarter, as described in Croushore & Stark (2001). GDP growth undergoes substantial revisions, in particular over the first releases, and even more so around turning points (see, e.g. Croushore 2011). Despite there is no consensus on what releases to use as "actual" data for forecast evaluation (Croushore 2006), we test out models against the toughest benchmark, the latest release.[19]

While real time data are available for the GDP, the same is not true for the NFCI and its subcomponents. Those are common factors in a broad panel of financial indicators and are extracted using state space methods. As such, the indicators are always revised at any point in

[18]Reichlin et al. (2020) extensively investigate the relation between the Nonfinancial leverage index and the Credit-to-GDP ratio index for forecastin purposes.

[19]This choice implies that the forecast error produced by our models accounts also for a "measurement error" component, induced by possible GDP measurement redefinitions, that took place over the considered forecast sample (for further discussion, see Stark & Croushore 2002).

time, when new information arrives. To this extent, we can only use a pseudo real-time dataset for the NFCI and its subcomponents. In Section 7, we use the disaggregated financial indicators which form the basis of the NFCI. Those data are available at weekly frequencies and, most importantly, they are not subject to revision.[20] As GDP data are released about 45 days after the end of the reference quarter, we assume that it is at that point in time that the forecaster estimates the model and produces the forecasts. This means that we can use all the information that is made available within the reference quarter. The weekly nature of these data allows us to include all the available information, by averaging all the weeks within the reference quarter. Therefore, in this case, the information set is fully real-time for the conditioning variables.

5 Time variation in the distribution of GDP growth

Our model allows us to study the characteristics of the conditional distribution of GDP growth. We place particular emphasis on the asymmetry (skewness) parameter, as we seek to provide a timely and accurate description of the downside risk to GDP growth. In this Section, we consider a baseline specification including (two lags of) all the four subindices of the NFCI. In fact, while the risk and credit components closely track the dynamics of the NFCI, it is clear from the discussion above that the leverage indicator, as well as the nonfinancial leverage index provide additional information. In particular, the NFL index is starkly smoother and shows a clear build up ahead of contractions. In Section 6 we will show that predictive densities benefit from fat tails, time-varying skewness, and financial information. Moreover, we show that the specification used in this section is associated with substantial improvements over the one that only includes the NFCI, or the one that does not feature any financial indicator as a predictor of the distribution of GDP growth.

Figure 3 reports the time varying moments of the distribution of GDP growth, together with the implied density. The mean moves along the business cycle and displays sharp contractions during recessions. Volatility features a marked counter-cyclical behaviour, with peaks occurring during recessions, whereas skewness is distinctively pro-cyclical. While the sample skewness value of -0.36 is mainly due to the large drawdowns occurring at the onset of recessions, sharp rebounds of the coefficient towards positive values accompany economic recoveries. Interestingly, skewness tends to decrease in anticipation of recessions, a feature which we show to be related

[20]The NFCI and the its subcomponents are extracted from a state space mode, and are therefore subject to revision.

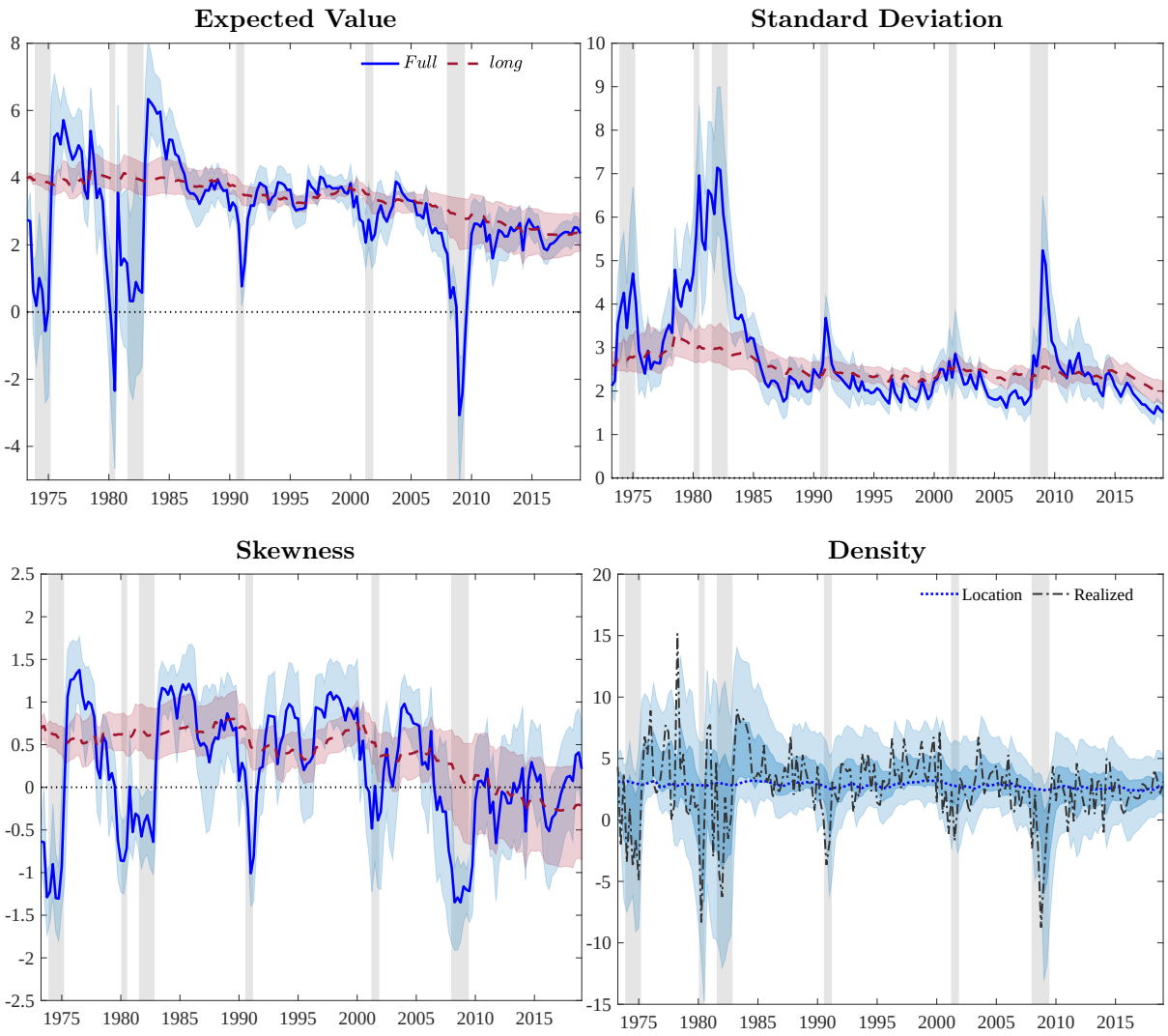

Figure 3: Time-varying moments

Note: The plot illustrate the estimated time-varying moments (black) along with the respective long-run components in red.

to the information contained in the financial indicators affecting the shape of the conditional distribution.

On the same chart we also report, in red, the low-frequency component of each of the moments. That is, the moments of the distribution which would prevail in the absence of cyclical variations of the model's parameters.[21] The model neatly captures a fall in long run growth, with the expected value falling from roughly 4% in the '70s, to roughly 2.3% over the sample, in line with the evidence in Antolin-Diaz et al. (2017). The model also captures the substantial decline in variance, starting from the mid 80s. The Great Moderation is reflected by

[21] This corresponds to setting $\mu_t = \bar{\mu}_t$, $\gamma_t = \bar{\gamma}_t$ and $\delta_t = \bar{\delta}_t$.

the long-run volatility being revised downward by roughly 30%.[22] In the most recent times, such lower trend level suggest that the GM period might not be over (Gadea et al. 2018). Similarly, the long-run skewness displays a downward trend, starting in the early '90s and more markedly in the post 2000 sample. As a result, business cycle fluctuations are characterized by decreasing, but positive, trend-skewness until the onset of the financial crisis in 2007. In the aftermath of the subsequent recession, this long-term trend turns to negative values, implying negatively skewed long-term conditional distributions.

Overall, the conditional distribution of GDP growth features mild positive skew over expansions, whereas downside risk clearly dominates (ahead of, and) during downturns. Notwithstanding increasing volatility, during recessions we observe marked downward movements in the lower quantiles of the distribution, not being matched by increases in the higher quantiles.[23]

Expected value and variance decomposition The estimated model suggests that conditional skewness fluctuations are a prominent feature of the the predictive distribution of GDP growth. Here we illustrate how these shifts affect the central moments of the distribution, besides shaping the behavior of the tails. In particular, for the Skew-t model, one can easily isolate the contributions of the variation in the asymmetry of the distribution as follows:

$$\mathbb{E}(y_t|Y_{t-1}) = \mu_t - g(\eta)\sigma_t^2\varrho_t, \qquad g(\eta) = \frac{4\mathcal{C}(\eta)}{1-\eta} \tag{17}$$

$$Var(y_t|Y_{t-1}) = \sigma_t^2\left(\frac{1}{1-2\eta} + h(\eta)\varrho_t^2\right), \qquad h(\eta) = \frac{3}{1-2\eta} - g(\eta)^2 \tag{18}$$

Therefore, the expected value and variance are equal to the location and (scaled) scale of a standard Student-t distribution, plus a component which is a function of the shape parameter of of the Skew-t density. This latter is magnified by larger values of the scale σ_t, while it disappears when the shape ϱ_t is equal to 0 (i.e., $\varrho_t = 0$). Figure 5 isolates the contribution of the asymmetry for both the first and second moment.

The expected value decomposition highlights how the location parameter (red line) is remarkably stable over the sample, while most of the fluctuations reflect shifts in the shape of the distribution, with expansions characterized by positive skew and contractions associated with

[22] According to the decomposition in Figure 3 the high volatility in the early part of the sample is also due to a large cyclical component over that sample.

[23] In contrast, when the model is estimated imposing a Gaussian conditional distribution, the widely documented countercyclicality of GDP growth (see, e.g. Jurado et al. 2015) implies that higher and lower quantiles of the distribution move in opposite direction, so that the fall in the expected value is necessarily met by a fall in the location.

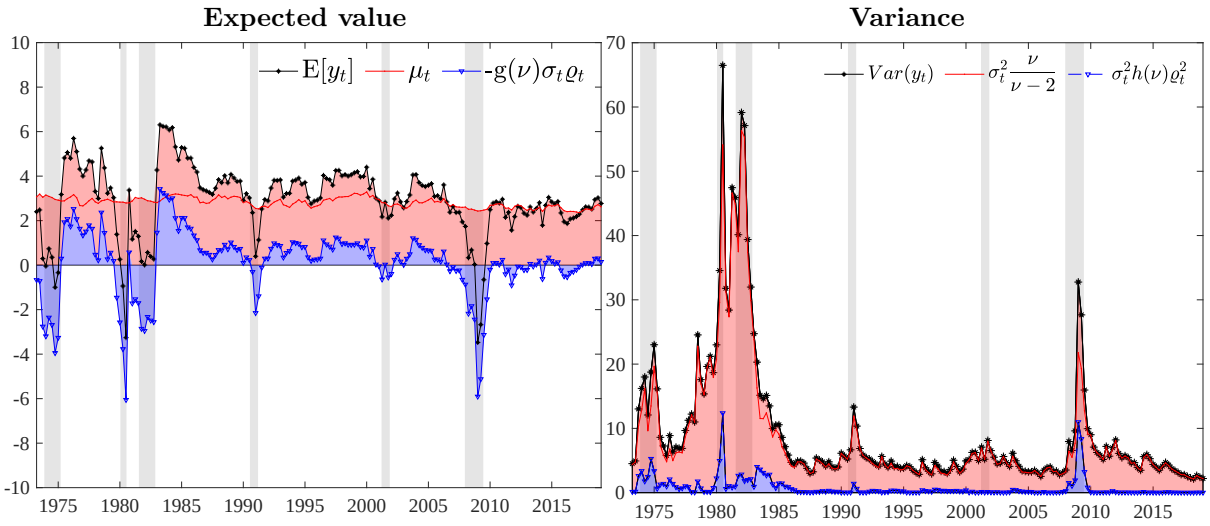

Figure 4: Expected value and variance decomposition

Note: The plot shows the decomposition of the expected value and variance of GDP growth. We report in red the contribution of the moment estimator (e.g. location and (scaled) scale), whereas the blue area identifies the contribution of higher order moments. Central moments (black lines) are computed as per Equation (17)-(18).

negative skew. Interestingly, the contribution of the asymmetry for positive expected values becomes more muted during the Great Moderation, whereas the negative drag from the asymmetry remains of substantial importance during recessions.[24] In contrast, the effect of the shape parameter on the second moment is less pervasive. Nevertheless, it accounts for a non-trivial share of the increase in variance during recessions, as recently documented by Salgado et al. (2019) in the cross-section of firms, for several countries.

Accounting for the fall in long-run growth. The upper left panel of Figure 3 shows that the model picks up a substantial fall in the underlying long-run GDP growth.[25] Following the decomposition in (17), we are able to disentangle the contribution of increasing risk to the decline in long run growth. Starting during the Great Moderation, the distribution of GDP growth has featured decreasing positive skewness, particularly over the post-2000 sample. Notwithstanding two transitory upward corrections, due to the oil crisis of the early '90s and the dot-com bubble of the first '00s, increasing downside risk is exerts a pervasive effect on the decline of the long-run growth, as shown in Figure 5. We find that this gradual reduction of the upside risk, eventually replaced by permanent downside risk, accounts for roughly 90% of the overall variation in long-run GDP growth.

[24] This is overall consistent with the findings in Jensen et al. (2020) and Gadea et al. (2018).
[25] The decomposition we report is akin to taking $\lim_{h\to\infty}\mathbb{E}[y_{t+h}]$

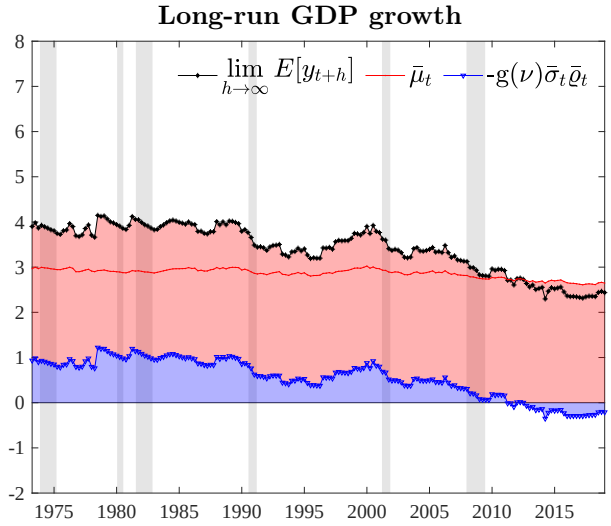

Figure 5: Long-run GDP growth

Note: The plot shows the decomposition of the long-run expected value (black lines), akin to taking $\lim_{h\to\infty}\mathbb{E}[y_{t+h}]$. We report in red the contribution of the long-run location component $\bar{\mu}_t$, whereas the blue area identifies the contribution of higher order moments. $\bar{\sigma}_t$ and $\bar{\varrho}_t$ refers to the secular components of scale and shape, respectively.

How important are financial predictors? Financial indicators are important predictors of the scale and shape parameters, whereas they do not affect much the overall location. To gauge the contribution of financial indicators int the overall variation of these parameters, we exploit the the moving average representation of the cyclical components of the (log) scale and the (arctanh) shape (eq. 13 and 16)

$$\tilde{\gamma}_{t+1} = \sum_{j=0}^{\infty} \phi_{\gamma}^{-j} \left(\beta_{\gamma}' X_{t-j} + \kappa_{\gamma} s_{\gamma t-j}\right) \tag{19}$$

$$\tilde{\delta}_{t+1} = \sum_{j=0}^{\infty} \phi_{\delta}^{-j} \left(\beta_{\delta}' X_{t-j} + \kappa_{\delta} s_{\delta t-j}\right). \tag{20}$$

In Figure 6 we decompose $\tilde{\gamma}_t$ and $\tilde{\delta}_t$ into a "score-driven" component ($\kappa_{\gamma}\sum_{j=0}^{\infty}\phi_{\gamma}^{-j}s_{\gamma t-j}$ and $\kappa_{\gamma}\sum_{j=0}^{\infty}\phi_{\delta}^{-j}s_{\delta t-j}$, respectively) and a component reflecting the share of variation driven by the predictors, for which we highlight the contribution of each financial index we consider. This decomposition highlights the importance of the predictors. Looser financial conditions are associated with lower than average volatility and a distribution that features mildly positive skew. Nonfinancial leverage is by far the largest contributor to the variation in the scale parameter, followed by the leverage indicator, whereas both credit and risk play a limited role. On the contrary, indicators heterogeneously affect the shape parameter. Specifically, during the severe recessions of the early 80s and 2008, the marked increase of the parameter reflects the increasing

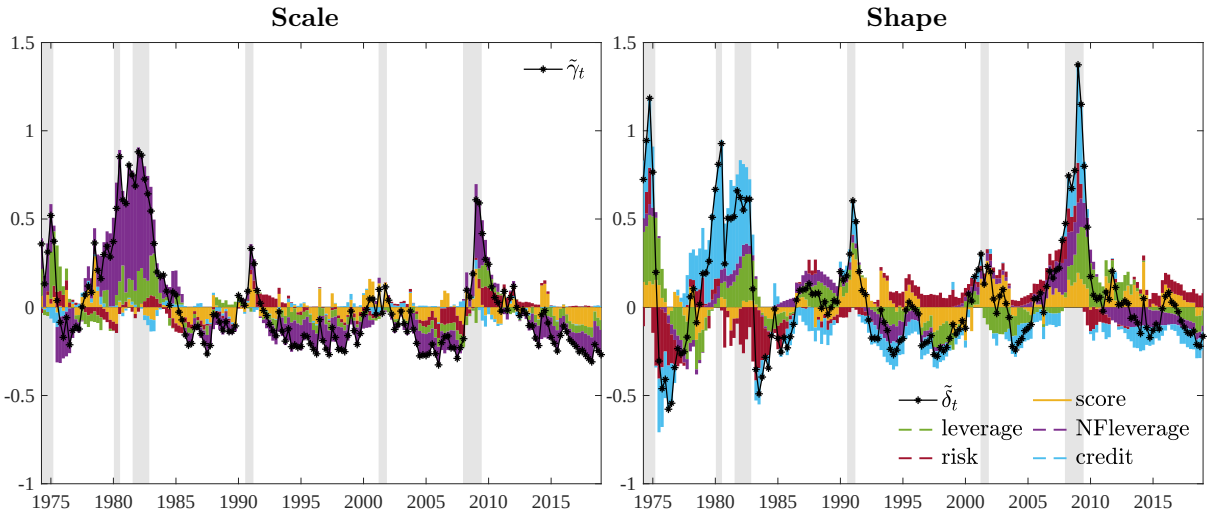

Figure 6: Predictive financial conditions

Note: The figures plot the decomposition of the "untransformed" short-term parameters (in black) into a "Score-driven" (yellow) and a "Predictors-driven" component, for which we highlight all the subcomponents.

of imbalances in the financial markets. Risk and nonfinancial leverage tend to positively feed into the dynamics ahead of recessions (see, e.g. Bekaert & Hoerova 2014, Reichlin et al. 2020, respectively), while credit and leverage indices contribute to determine the amount of asymmetry recessions are characterized by. Interestingly, the contribution of financial indicators is muted during the dot-com recession, in line with the weak link between this recession and financial predictors, as highlighted by Stock & Watson (2003).

6 Out-of-sample forecast

In this section, we evaluate the out-of-sample forecast performance of the different model specifications, over the period 1993Q1 to 2018Q4. We consider both one-quarter and one-year-ahead forecasts. In the latter case, the results are reported in terms cumulated output growth over the next four quarters. Forecasts are obtained from real-time data, that is, we apply our methodology to the data that were available to the forecaster at each point in time before producing the forecast, whereas predictive scores are based on the most recent release.

We assess the predictive accuracy of the models with respect to their point forecast accuracy, via the mean square forecast error (MSFE). Density forecasts accuracy is evaluated via the predictive log-score and quantile scores. In particular, we consider two versions of the weighted

Quantile Score measure of Gneiting & Ranjan (2011)

$$wQS_{t+h} = \int_0^1 QS(\alpha)\omega(\alpha)d\alpha \tag{21}$$

where $QS(\alpha) = 2(I(y_{t+h} < F^{-1}(\alpha)) - \alpha)(F^{-1}(\alpha) - y_{t+h})$, with $F^{-1}(\alpha)$ being the empirical quantile function of the density forecast, for $\alpha \in [0, 1]$ and $\omega(\alpha)$ is a weighting function: the Continuously Ranked Probability Score (CRPS) which assigns equal weight to each quantile of the empirical distribution function (Gneiting & Raftery 2007), as well as a scoring rule with $\omega(\alpha) = (1 - \alpha)^2$ that assigns higher weights to the lower quantiles of the distribution function, thus emphasising the accuracy in predicting the left tail and downside risk. Similarly, when we evaluate the correct calibration of the predictive densities by means of the probability integral transforms (PITs) (Diebold et al. 1998), we explicitly consider the calibration of the the left side of the distribution. For this purpose, we rely on the Crámer-von Mises (CvM) test specification of Rossi & Sekhposyan (2019). Last, we investigate the use of the model in producing measures of downside risk and predicting recessions.

6.1 Point and density forecasts

Asymmetry and the value of financial predictors. As a first step into our analysis, we aim at evaluating the importance of accounting for the skew of the distribution of GDP growth, as well as assessing whether conditioning on financial predictors leads to improved forecasts. Tables 2 and 3 report the forecast scores for the two horizons, for alternative skew-t models against the forecast produced by a Gaussian model with time varying volatility.[26] In particular, we produce forecasts from (a) a skew-t model without any financial predictor ($no - X$), (b) a skew-t model that includes as predictor only the NFCI ($NFCI$) and (c) a skew-t model including the four disaggregate financial indices ($4DFI$). Scores are reported as the ratio with respect to the Gaussian benchmark, a part from the log-score for which we report the difference. Values in parentheses report the p-values of the Giacomini & White (2006) test for conditional predictive accuracy. Moreover, we compare the performance of the model for the entire out-of-sample

[26] A number of recent contributions, including Brownlees & Souza (2019) and Carriero et al. (2020), highlight that simple time series models with time-varying volatility and Gaussian innovations provide accurate density forecasts for GDP growth. Therefore, as a benchmark we choose an AR(2) model with (score-driven) time varying volatility and Gaussian innovations. This means that the log-volatility follows a GARCH-type model (see Creal et al. 2013).

Table 2: Forecast scores - One quarter ahead

	Skt no-X	*Skt NFCI*	*Skt 4DFI*	*Skt no-X*	*Skt NFCI*	*Skt 4DFI*
	MSFE			**logS**		
Full	**0.827** (0.021)	**0.769** (0.025)	**0.777** (0.054)	0.106 (0.139)	**0.154** (0.066)	0.105 (0.203)
Rec.	**0.764** (0.038)	**0.674** (0.057)	**0.651** (0.067)	0.408 (0.101)	0.539 (0.102)	**0.545** (0.088)
GFC	0.767 (0.129)	0.621 (0.107)	**0.569** (0.085)	0.581 (0.244)	0.897 (0.142)	0.955 (0.103)
	CRPS			**wQS**		
Full	**0.916** (0.029)	**0.890** (0.019)	**0.903** (0.075)	0.910 (0.121)	**0.870** (0.056)	0.888 (0.124)
Rec.	**0.841** (0.027)	**0.791** (0.045)	**0.775** (0.063)	**0.796** (0.031)	**0.732** (0.054)	**0.719** (0.049)
GFC	0.825 (0.117)	**0.720** (0.088)	**0.666** (0.057)	0.798 (0.156)	0.679 (0.142)	**0.642** (0.096)

Note: The table reports the average forecast scores, expressed in terms of the Gaussian model's scores; the p-value for Giacomini & White (2006) test (in parentheses) compares the models with respect to the benchmark model. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

period, as well as for the two recessions in our sample and the great financial recession [27]

Simply introducing fat tails and asymmetry improves the forecast accuracy of the model with respect to the baseline Gaussian specification. However, further predictive accuracy is gained when conditioning on financial information. Using the four subcomponents is associated with additional gains during downside episodes. These conclusions are valid for both the one-quarter-ahead forecasts and for the one-year-ahead forecasts (where the improvements are in generally larger than for the short term forecast), and irrespective of the specific forecast score considered. The gains over the benchmark Gaussian model are quite substantial, for instance, the benchmark model produces roughly 25% (35%) improvement in RMSE, as well as 10% (15%) and 12% (23%) improvements in the CRPS and wQS respectively for the one-quarter (one-year) ahead forecasts. The gains become even larger if one focuses on recessions, or the 2007-09 recession.

[27] Test statistics for the Recession and the GFC periods are to be taken as guidelines given the considered small sample.

Table 3: Forecast scores - One year ahead

	Skt no-X	*Skt NFCI*	*Skt 4DFI*	*Skt no-X*	*Skt NFCI*	*Skt 4DFI*
	MSFE			**logS**		
Full	**0.759** (0.046)	**0.660** (0.052)	0.667 (0.162)	0.179 (0.208)	**0.306** (0.008)	**0.265** (0.096)
Rec.	**0.853** (0.035)	**0.715** (0.064)	**0.671** (0.003)	0.346 (0.437)	**0.652** (0.007)	**0.713** (0.007)
GFC	**0.866** (0.019)	**0.690** (0.036)	**0.583** (0.002)	0.072 (0.300)	**0.911** (0.000)	**1.192** (0.000)
	CRPS			**wQS**		
Full	**0.878** (0.023)	**0.841** (0.009)	**0.846** (0.045)	**0.837** (0.044)	**0.770** (0.039)	0.774 (0.153)
Rec.	**0.871** (0.009)	**0.821** (0.028)	**0.789** (0.008)	**0.796** (0.001)	**0.731** (0.008)	**0.682** (0.001)
GFC	**0.864** (0.004)	**0.793** (0.015)	**0.734** (0.003)	**0.807** (0.000)	**0.700** (0.018)	**0.597** (0.001)

Note: The table reports the average forecast scores, expressed in terms of the Gaussian model's scores; the p-value for Giacomini & White (2006) test (in parentheses) compares the models with respect to the benchmark model. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

Comparison against Adrian et al. (2019) In Table 4 we report the comparison of the benchmark specification against the model of Adrian et al. (2019).[28] Also in this case, the baseline model specification is associated with better point and density forecasts and with significant values arising especially in the recession subperiods. In particular, it is worth noticing that the forecast gains that we document during recessions stem from the adaptiveness of the score filter. The shape parameter we estimate promptly reacts to turning points, thus implying timely fluctuations of the skewness of the predictive distributions. As a consequence, forecast densities are characterized by longer negative tails during recessions, as compared to those implied by the skew-t model of Adrian et al. (2019).

Density calibration Table 5 evaluates the calibration of the density forecast looking at the properties of the probability integral transform (PITs). In particular, we report the CvM test statistic of Rossi & Sekhposyan (2019) for the Gaussian model, the model of Adrian et al. (2019) and our skew-t model with the four disaggregated financial indices. The CvM test rejects the null hypothesis that the cumulative distribution function of the PIT lies within the 10% critical

[28]Specifically, we re-estimated the model of Adrian et al. (2019) using real time data for GDP growth in an expanding window exercise, consistently with our approach.

Table 4: Forecast scores with respect to Adrian et al. (2019)

	one-quarter-ahead				*one-year-ahead*			
	MSFE	**logS**	**CRPS**	**wQS**	**MSFE**	**logS**	**CRPS**	**wQS**
Full	0.977 (0.775)	0.031 (0.623)	0.975 (0.774)	0.986 (0.943)	0.873 (0.625)	0.113 (0.151)	0.964 (0.654)	0.962 (0.481)
Rec.	0.970 (0.799)	0.120 (0.248)	**0.946** (0.077)	0.930 (0.236)	0.772 (0.329)	0.212 (0.131)	**0.779** (0.002)	0.833 (0.278)
GFC	0.965 (0.936)	0.274 (0.334)	**0.907** (0.096)	0.917 (0.538)	**0.652** (0.000)	**0.802** (0.000)	**0.698** (0.000)	**0.692** (0.000)

Note: The table reports the average forecast scores of the baseline model, expressed in terms of Adrian et al. (2019) model's scores; the p-value for Giacomini & White (2006) test (in parentheses) compares the models with respect to the benchmark model. Values in **bold** are significant at the 10% level

Table 5: Density calibration tests

	AR(2)	*ABG*	*Skt 4DFI*	*AR(2)*	*ABG*	*Skt 4DFI*
	One-quarter ahead			**One-year ahead**		
Dist.	**0.561**	**1.401**	0.167	**1.688**	**1.524**	0.414
Left tail	**0.510**	**0.567**	0.195	**3.244**	**2.295**	0.782

Note: The table reports the test statistics for the Rossi & Sekhposyan (2019) tests, based on the Cramér-von Mises type tests. Values in **bold** indicate the rejection of the null hypothesis of correct specification of the density forecast, at the 10% confidence level. Critical values are obtained by 1000 bootstrap simulations. Gray shaded cells indicate the best score. The left tail score is computed over the support [0, 0.25].

value bands for both the Gaussian benchmark and for the model of Adrian et al. (2019), for the one-quarter and one-year ahead forecasts. Whereas, the null is not rejected for our model, neither for the calibration of the entire density, nor for the calibration of the 'left tail' of the distribution.

6.2 Downside Risk predictions for the Great Recession

In this Section, we assess downside risk predictions, placing particular focus on the ability of the model to anticipate the build up in downside risk ahead of the 2008 financial crisis, and its reduction during the subsequent recovery. Measures such as Value at Risk (VaR), as well as

the Expected Shortfall (ES) are readily obtained within our framework. ES^{α}_{t+h} describes the expected growth level for $y_{t+h} < VaR^{\alpha}_{t+h}$, corresponding to the $(100\alpha)^{\text{th}}$ percentile of the h-step ahead predictive distribution, whereas the Expected Longrise ($EL^{1-\alpha}_{t+h}$) is the upper counterpart of the ES. Specifically,

$$ES^{\alpha}_{t+h} = \alpha^{-1} \int_0^{\alpha} VaR^{a}_{t+h|t} da, \qquad EL^{1-\alpha}_{t+h} = \alpha^{-1} \int_{1-\alpha}^{1} VaR^{a}_{t+h|t} da \tag{22}$$

The left hand panel of Figure 7 contrasts the 5% expected shortfall and the 95% expected longrise for the Gaussian model, the score-driven skew-t model without financial predictors and our baseline model with the four subindices of the NFCI.

The Gaussian model fails to capture the building-up of risk ahead of the recession, predicting an expected shortfall around zero as the economy is entering the recession. In addition, assuming a symmetric distribution implies that falls in the ES are often associated with peaks in the EL, as clearly visible during the latest recession, where the minimum ES corresponds to the maximum EL in 2009Q2. On the other hand, allowing for Skew-t innovations alleviates both problems and delivers more conservative risk measures with less erratic longrise figures. Moreover, skew-t models anticipate the build-up of downside risk ahead of the recession. Conditioning the forecasts on the available subindices of financial conditions increases the timeliness of the prediction of risk, due to the prompt discounting of financial overheating. The prediction of the ES falls to roughly -5% in the first quarter of the recessions, and decreases consistently until the first quarter of recovery, as indicated by a sharp upward revision. Moreover, within the recession, the model delivers a downwarding longrise, predicting modest gains even for the most optimistic scenario. Likely to the ES, the longrise is sharply revised upward already for the first post-trough quarter.

Brownlees & Souza (2019) argue that the time-varying dynamics of the lower quantiles can be better captured by means of off-the-shelf volatility models. On the contrary, we document substantial gains in the assessment of downside risk associated to our class of skew-t models. Evaluating ES accuracy using the score metric proposed by Taylor (2019) highlights that the baseline model produces gains of up to 80% with respect to the GARCH-type Gaussian model, and 25% with respect to the Skew-t model with no predictors, for the one-quarter ahead forecast during the crisis period; even larger gains are found for the one-year ahead forecast. Using different scores, such as the Fissler et al. (2016) loss function, or the tick loss function for the

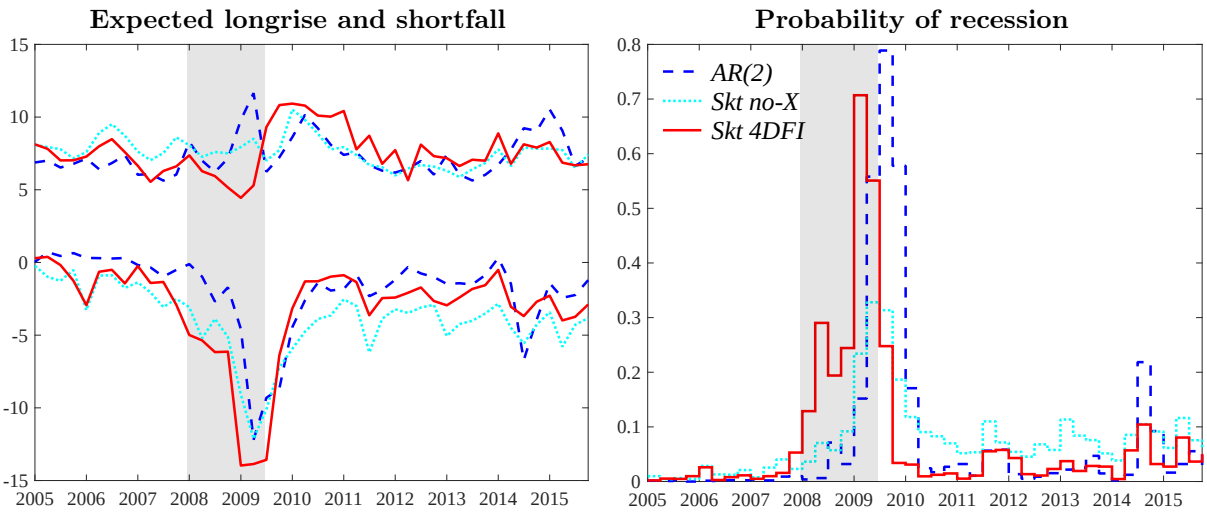

Figure 7: Expected Shortfall and Expected Longrise

Note: We report the ES and EL for $\alpha = 0.5$.

5% VaR proposed by Giacomini & Komunjer (2005) we document similar gains.[29]

Last, we investigate the ability of the model to predict recessions. The NBER Business Cycle Dating Committee (BCDC) defines a recession as *"[...] a significant decline in activity spread across the economy, lasting more than a few months [...]"*. Within our forecast distributions of GDP growth up to a one year horizon, we retrieve the probability of observing any two consecutive negative forecasts for the next four quarters. In line with previous results, combining the information on financial conditions and allowing for asymmetry in the forecast densities produces a realistic assessment of recession risk. The implied probability of recession stars picking up earlier as compared to the other measures, warning against an imminent output contraction. Moreover, the probability of observing a recession within the forthcoming year recedes sharply when the recession ends and is already below 5% just a quarter after the end of the recession as dated by the BCDC. In contrast, the Gaussian model (as well as the Skew-t model without conditioning information) starts to produce a reasonable probability of recession only toward the end of the recession period, whereas they continue to perceive a substantial treat of recession many quarters after the formal end of it. Evaluating the ability of the model to time recessions over the sample 1993-2018, using the Brier score, produces (significant) gains of more than 30% with respect to the Gaussian model, whereas the addition of financial predictors is responsible for a 10% in the Brier score.

Overall, the results suggest that financial conditions significantly contribute to the prediction

[29]The full set of results is available in Table D6, in Appendix D.

of downside risk to growth. The higher frequency at which financial data are sampled makes financial condition indicators accurate signals of financial shocks hitting the real economy sector. Thus, exploiting this information improves the prediction of downside risk episodes, both in terms of magnitude and timing.

7 Dissecting the Financial Condition Index: large data evidence

In the previous Section, we have highlighted that financial conditions are important predictors of downside risk. In fact, using the disaggregated indices proves particularly useful in anticipating the vulnerability of GDP growth. A natural question at this stage is to what extent the predictive power of the model can be further improved considering the entire panel of data that feeds into the NFCI, and what are the indicators that provide useful in predicting downside risk.

In this Section, we consider the full set of 105 indicators of financial activity that constitute the NFCI.[30] Exploiting the full set of financial indicators also allows us to produce "true" real-time forecasts. Specifically, to obtain predictors in real-time, we assume that a time t, the set of predictors corresponds to the quarterly average of the financial indicators from the third week of the previous quarter to the second week of the current quarter. This approach mimics the information set available to the econometrician who produces real-time forecasts and avoids dealing with overlapping quarters.[31] As indicators enter the predictors' set at different points in time, for each each time t, we only consider predictors available for at least four years. This implies that the first forecast produced in 1992Q4 includes less than 50% of the 105 financial indices. As the availability of the predictors steadily increases over the sample, the first forecast of the 2001 recession includes 70% of the total predictors, and by the time we produce the first forecast for the 2007-2009 recession we include 85%.

[30] Figure B1 in Appendix B for a comprehensive picture of data availability. Brave & Butters (2012) illustrate the miscellaneous nature of the series that feed into the NFCI at different times and frequencies. We are grateful to the authors and the Chicago FED for making the full panel of weighted indicators available for this work.

[31] This approach is consistent with the aggregation approach used by the FRED database.

7.1 Variables selection: *"shrink-then-sparsify"*

A potential concern of this exercise lies in the steep increase in the number of parameters our model needs to accommodate. In fact, when all indicators (and their lags) are included at the same time, the model features more than 600 coefficients. We tackle this dimensionality problem through a *"shrink-then-sparsify"* strategy (see Hahn & Carvalho 2015). Specifically, the shrinkage of the predictor loadings is induced by means of hierarchical priors and, in particular, we rely on the Horseshoe (HS) prior specification of Carvalho et al. (2010): $b^j \sim \mathcal{N}(0, \lambda^j \tau)$, where the hyperparameters λ^j and τ control the local (coefficient specific) and the global shrinkage, respectively. Therefore, $\lambda^j \sim HC^+(0,1)$ and $\tau \sim HC^+(0,1)$, where $HC^+(0,1)$ denotes the standard Half-Cauchy distribution. Unlike other common shrinkage priors (e.g., Ridge, Lasso), the HS priors are free of exogenous inputs, implying a fully adaptive shrinkage procedure. We add a second sparsification step to reduce the estimation uncertainty arising from the near-zero shrinkage coefficient.[32] We approach sparsification through the Signal Adaptive Variable Selector (SAVS) algorithm of Ray & Bhattacharya (2018). This data-driven procedure specifies the sparsification tuning parameter as $m_j = |\hat{b}_j|^{-2}$ such that each of the j variables receives a penalization *"ranked in inverse-squared order of magnitude of the corresponding coefficient"* (Ray & Bhattacharya 2018). The sparsified coefficients are computed as

$$b_j^* = sgn(\hat{b}_j)||X_j||^{-2} \max\left\{|\hat{b}_j| \cdot ||X_j||^2 - m_j, 0\right\} \tag{23}$$

where $||\cdot||$ represents the Euclidean norm of the vector X_j. We apply the sparsification step for each draw of the MCMC algorithm to further account for model uncertainty. As noted by Huber et al. (2020), this procedure is akin to the idea of Bayesian model averaging.

7.2 Results

The sparsification strategy is highly effective in reducing the number of coefficients we need to estimate when producing forecasts from a large panel of financial indicators. Figure 8 plots the evolution over time of the number of financial indicators available that are selected as predictors for each of the three time varying parameters. These values are computed as the number of predictors that receive a loading of at least 1% after the SAVS algorithm has been applied to

[32] Giannone et al. (2018) warn against the use of sparse modelling without critical judgement, as sparsity can be obtained as a result of strong *a priori* beliefs (parameter setting) while not being a robust feature of the data.

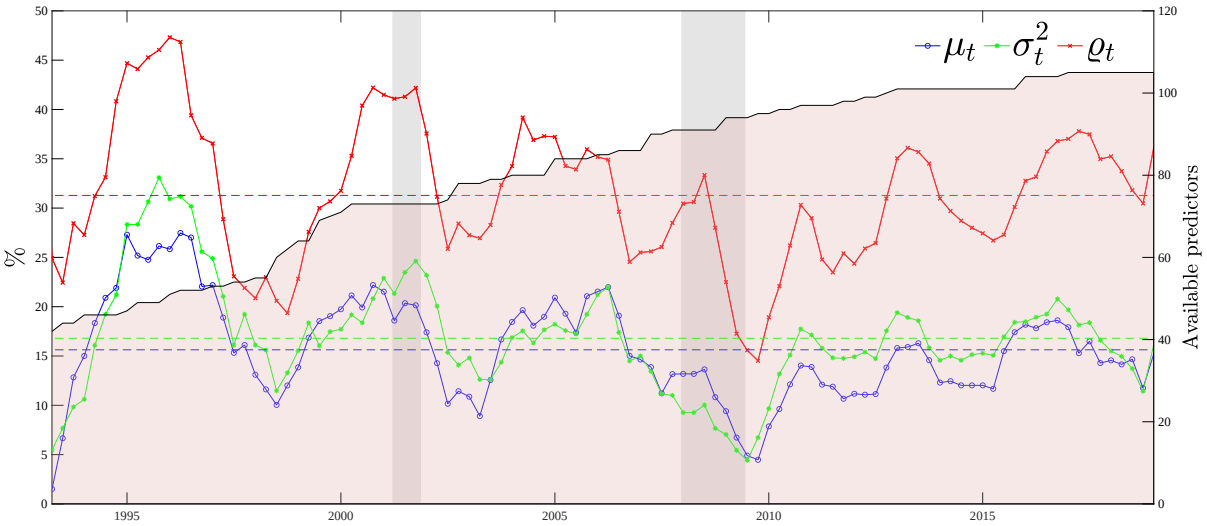

Figure 8: Number of predictors for the time-varying parameters over time

Note: For each period, the number of predictors corresponds to amount of predictors that receive a loading of at least 0.01 after the SAVS algorithm has been applied.

the shrinkage parameters.[33] The number of selected coefficients increases over time, also as a reflection of the higher number of indicators that become available over the sample. Interestingly, the number of predictors the model selects on average twice as many for the shape parameter than for the location and scale (roughly 31% vs. 15% and 16% of the available predictors). This highlights the importance of predicting the asymmetry of the distribution, to appropriately reflect the underlying uncertainty in GDP growth. Moreover, the number of predictors of the shape parameters increases ahead of the financial crisis, roughly around the time when the model starts to predict an increasing downside risk in the one year ahead forecast. In fact, the information in the financial indicators maps into substantial gains in prediction accuracy (as is highlighted in Table 6).

The upper panel of Figure 9 reports ten predictors for which the average posterior probability of inclusion is the highest for the three time-varying parameters. Overall, leverage and credit indicators are the most relevant categories of predictors. In particular, corporate equity issuances (STKGR) and household leverage (HH) show up as the most important predictors of the asymmetry of the distribution, whereas commercial property prices (CPH) and the 10-2 year Treasury spread are relevant predictors for the scale.[34] The predictors that are selected do not remain constant over time. The lower panel of Figure 9 reports the most relevant predictors during the Great Recession. In this case, risk indicators play a more relevant role also in

[33] Data has been smoothed using a 2-years moving average filter for comparison purposes.

[34] Appendix D reports the full ranking of the predictors for the three parameters.

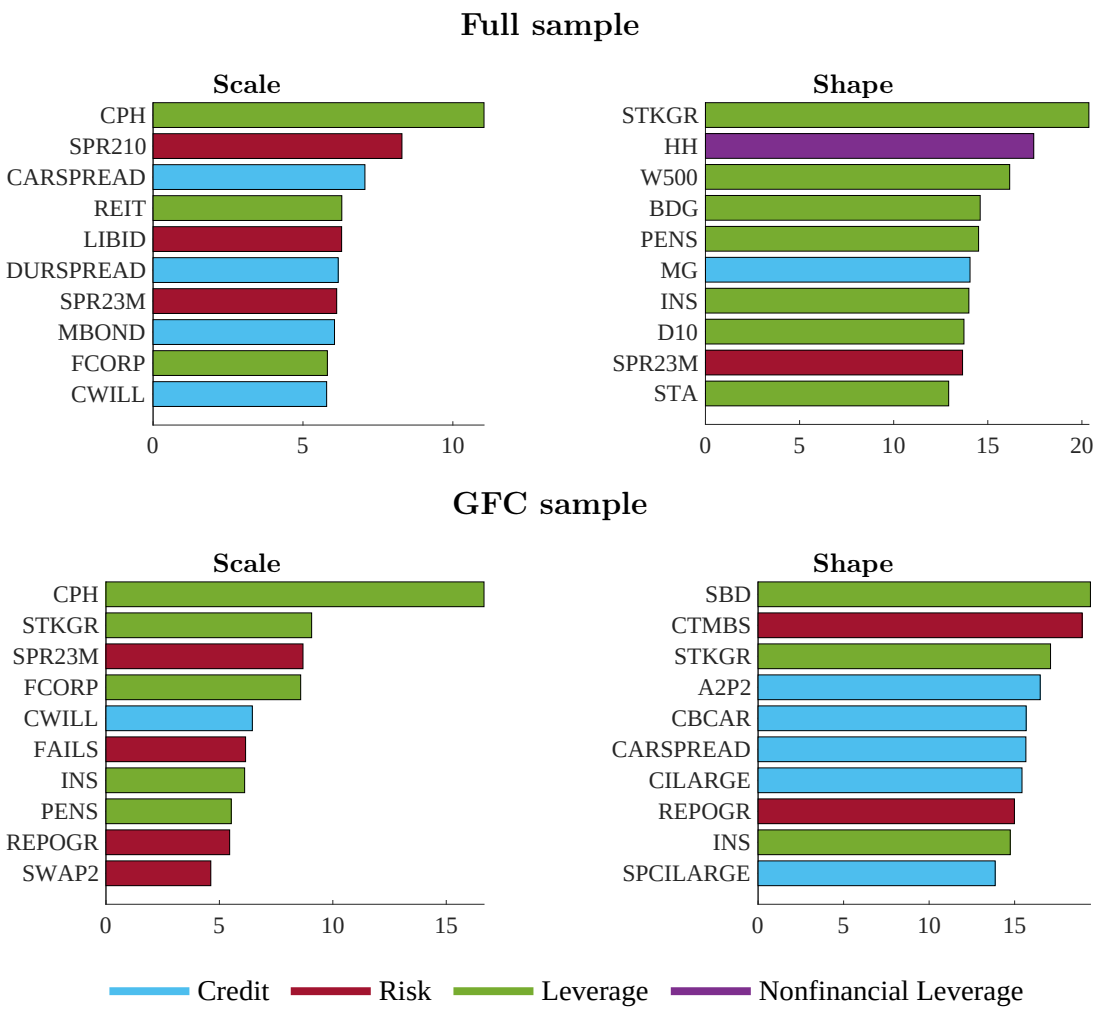

Figure 9: Top 10 predictors for σ_t and ϱ_t

Note: The bar plots report the top 10 predictors for each parameter. The *x-axis* reports average posterior probability of inclusion, expressed in percentage terms.

predicting the shape of the distribution. The (\$)amount of traded security over GDP (SBD) and a measure of corporate bond spread provide useful to gauge the increasing downside risk associated with the period.

We assess the gains associated with using the disaggregated data in Table 6. Evaluated over the full sample, using a large dataset produces (point and density) forecasts which are in line with the baseline model using the 4 major subindices of the NFCI.[35] However, the performance of the model varies substantially over the sample. Using a large information set is particularly helpful during recessions. In this case, for predictions one quarter ahead we observe (significant) gains in density forecasts, in particular for the downside of the distribution, where we report

[35]It is worth stressing however, that using the raw financial indicators the model uses only full-real time predictors.

Table 6: Forecast score - Large data model

	one-quarter-ahead				*one-year-ahead*			
wrt	*ABG*	*Skt 4DFI*	*ABG*	*Skt 4DFI*	*ABG*	*Skt 4DFI*	*ABG*	*Skt 4DFI*
	MSFE		**logS**		**MSFE**		**logS**	
Full	1.022 (0.254)	1.047 (0.335)	**−0.015** (0.100)	−0.047 (0.985)	0.789 (0.583)	0.904 (0.165)	0.089 (0.350)	−0.024 (0.644)
Rec.	**0.937** (0.090)	**0.949** (0.002)	**0.203** (0.078)	0.116 (0.128)	0.570 (0.104)	**0.739** (0.067)	**0.638** (0.014)	**0.426** (0.017)
GFC	**0.969** (0.090)	**1.036** (0.022)	**0.379** (0.002)	**0.031** (0.004)	**0.501** (0.096)	**0.768** (0.001)	**1.152** (0.000)	**0.350** (0.000)
	CRPS		**wQS**		**CRPS**		wQS	
Full	1.019 (0.133)	1.045 (0.801)	1.035 (0.386)	1.050 (0.564)	0.978 (0.693)	1.015 (0.892)	0.930 (0.863)	0.967 (0.760)
Rec.	0.915 (0.135)	**0.960** (0.005)	**0.866** (0.075)	**0.940** (0.035)	0.746 (0.238)	**0.958** (0.013)	**0.639** (0.077)	**0.767** (0.033)
GFC	0.888 (0.103)	**1.027** (0.000)	**0.851** (0.045)	**0.990** (0.075)	0.652 (0.302)	**0.934** (0.000)	**0.540** (0.018)	**0.779** (0.000)

Note: The table reports the average forecast scores of the shrinkage model, in terms of the competing specifications' scores; the p-value for Giacomini & White (2006) test (in parentheses) compares the models with respect to the benchmark model. Values in **bold** are significant at the 10% level

gains of up to 15% if compared against the model of Adrian et al. (2019). The gains become more substantial for one-year hear predictions. We document 10/20% improvements (with respect to the baseline model and Adrian et al. (2019)'s, respectively) in point forecast during the Great recession, and up to almost 50% improvements (with respect to Adrian et al. (2019), 20% with respect to the baseline model) in the evaluation of the downside of the density. These results further reinforce the suitability of financial predictors in predicting the downside risk to the economy. A comprehensive dataset provides several advantages in nailing down the implied risk in the forecast of GDP growth, in particular for medium term forecasts.

8 Conclusions

The severity of the latest financial crisis and the following recession has spurred the interest of both academics and practitioners in the downside risk to economic growth, and how it relates to financial sector's turmoil. In this paper, we introduce a class of parametric models to fully characterize the conditionals distribution of GDP growth. The flexibility of these models allows us to inspect the dynamics and the drivers of the parameters relating to the level, uncertainty and asymmetry of business cycle fluctuations. We document asymmetry as a strong feature of the

data. This feature becomes more prominent after we condition the dynamics of the parameters of the predictive density of economic growth to the tightness of financial conditions. When financial markets are overheating, future economic growth becomes more uncertain not only due to increasing risk, but also due to deepening negative skewness of the predictive distributions.

Our findings highlight a significant contribution of financial indicators in predicting downturns, especially the extreme realizations of the 2007-2009 period. We document that indicators measuring the level of leverage and credit provide a better signal to track developments of real output growth. Moreover, an in-deep examination of the predictive information coming from the financial markets emphasizes that financial, and most importantly nonfinancial leverage measures, anticipate plummeting economic growth, driving the rising negative asymmetry characteristic of the 2007-2009 recession. Moreover, we find that since the 2007 recession the unconditional distribution of GDP growth display a marked negative skewness, indicating that the weak recovery observed over the last 10 years stands on very weak ground, and that a (severe) contraction in economic activity is more likely than in previous expansionary phases of the cycle. These findings lead to relevant implications for the forecasting of future output growth and for policy planning. When we evaluate risk measures and the implied probability of future recessions, models augmented with financial conditions outperform their competing specifications and deliver more realistic measures, both in terms of magnitude and timing.

References

Adrian, T., Boyarchenko, N. & Giannone, D. (2019), 'Vulnerable growth', *American Economic Review* **109**(4), 1263–89.

Adrian, T. & Shin, H. S. (2010), 'Liquidity and leverage', *Journal of Financial Intermediation* **19**(3), 418–437.

Alessi, L., Ghysels, E., Onorante, L., Peach, R. & Potter, S. (2014), 'Central bank macroeconomic forecasting during the global financial crisis: the european central bank and federal reserve bank of new york experiences', *Journal of Business & Economic Statistics* **32**(4), 483–500.

Antolin-Diaz, J., Drechsel, T. & Petrella, I. (2017), 'Tracking the slowdown in long-run gdp growth', *Review of Economics and Statistics* **99**(2), 343–356.

Azzalini, A. & Capitanio, A. (2003), 'Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65**(2), 367–389.

Bai, J. & Ng, S. (2005), 'Tests for skewness, kurtosis, and normality for time series data', *Journal of Business & Economic Statistics* **23**(1), 49–60.

Bekaert, G. & Hoerova, M. (2014), 'The vix, the variance premium and stock market volatility', *Journal of Econometrics* **183**(2), 181–192.

Bernanke, B. S. & Gertler, M. (1995), 'Inside the black box: the credit channel of monetary policy transmission', *Journal of Economic perspectives* **9**(4), 27–48.

Blasques, F., Koopman, S. J. & Lucas, A. (2015), 'Information-theoretic optimality of observation-driven time series models for continuous responses', *Biometrika* **102**(2), 325–343.

Brave, S. A. & Butters, R. (2011), 'Monitoring financial stability: A financial conditions index approach', *Economic Perspectives* **35**(1), 22.

Brave, S. & Butters, R. A. (2012), 'Diagnosing the financial system: Financial conditions and financial stress', *International Journal of Central Banking* **8**(2), 191–239.

Brownlees, C. T. & Souza, A. (2019), Backtesting global growth-at-risk, Working paper, Available at SSRN 3461214.

Carriero, A., Clark, T. E. & Marcellino, M. (2020), Capturing macroeconomic tail risks with bayesian vector autoregressions, Working Paper 20-02, Federal Reserve Bank of Cleveland.

Carvalho, C. M., Polson, N. G. & Scott, J. G. (2010), 'The horseshoe estimator for sparse signals', *Biometrika* **97**(2), 465–480.

Cette, G., Fernald, J. & Mojon, B. (2016), 'The pre-Great Recession slowdown in productivity', *European Economic Review* **88**(C), 3–20.

Chauvet, M. & Potter, S. (2013), Forecasting output, *in* 'Handbook of Economic Forecasting', Vol. 2, Elsevier, pp. 141–194.

Cox, D. R. (1981), 'Statistical analysis of time series: Some recent developments', *Scandinavian Journal of Statistics* pp. 93–115.

Creal, D., Koopman, S. J. & Lucas, A. (2013), 'Generalized autoregressive score models with applications', *Journal of Applied Econometrics* **28**(5), 777–795.

Croushore, D. (2006), 'Forecasting with real-time macroeconomic data', *Handbook of Economic Forecasting* **1**, 961–982.

Croushore, D. (2011), 'Frontiers of real-time data analysis', *Journal of Economic Literature* **49**(1), 72–100.

Croushore, D. & Stark, T. (2001), 'A real-time data set for macroeconomists', *Journal of Econometrics* **105**(1), 111–130.

De Nicolò, G. & Lucchetta, M. (2017), 'Forecasting tail risks', *Journal of Applied Econometrics* **32**(1), 159–170.

Delle Monache, D. & Petrella, I. (2017), 'Adaptive models and heavy tails with an application to inflation forecasting', *International Journal of Forecasting* **33**(2), 482–501.

Diebold, F. X., Gunther, T. A. & Tay, A. (1998), 'Evaluating density forecasts, with applications to financial risk management', *International Economic Review* **39**, 863–883.

Ding, Z. & Granger, C. W. (1996), 'Modeling volatility persistence of speculative returns: a new approach', *Journal of Econometrics* **73**(1), 185–215.

Doan, T., Litterman, R. & Sims, C. (1984), 'Forecasting and conditional projection using realistic prior distributions', *Econometric reviews* **3**(1), 1–100.

Drehmann, M., Borio, C. E., Gambacorta, L., Jimenez, G. & Trucharte, C. (2010), Countercyclical capital buffers: exploring options, Working Paper 317, Bank for International Settlement.

Engle, R. F. & Rangel, J. G. (2008), 'The spline-garch model for low-frequency volatility and its global macroeconomic causes', *The Review of Financial Studies* **21**(3), 1187–1222.

Escanciano, J. C. & Lobato, I. N. (2009), 'An automatic portmanteau test for serial correlation', *Journal of Econometrics* **151**(2), 140–149.

Fissler, T., Ziegel, J. F. & Gneiting, T. (2016), 'Expected shortfall is jointly elicitable with value-at-risk: implications for backtesting.', *Risk.net (www. risk. net/ 2439862)* .

Gadea, M. D., Gómez-Loscos, A. & Pérez-Quirós, G. (2018), 'Great moderation and great recession: From plain sailing to stormy seas?', *International Economic Review* **59**(4), 2297–2321.

Galvão, A. B. & Owyang, M. T. (2018), 'Financial stress regimes and the macroeconomy', *Journal of Money, Credit and Banking* **50**(7), 1479–1505.

Gertler, M. & Gilchrist, S. (2018), 'What happened: Financial factors in the great recession', *Journal of Economic Perspectives* **32**(3), 3–30.

Giacomini, R. & Komunjer, I. (2005), 'Evaluation and combination of conditional quantile forecasts', *Journal of Business & Economic Statistics* **23**(4), 416–431.

Giacomini, R. & White, H. (2006), 'Tests of conditional predictive ability', *Econometrica* **74**(6), 1545–1578.

Giannone, D., Lenza, M. & Primiceri, G. E. (2018), 'Economic predictions with big data: The illusion of sparsity', *Federal Reserve Bank of New York staff report* **847**.

Giglio, S., Kelly, B. & Pruitt, S. (2016), 'Systemic risk and the macroeconomy: An empirical evaluation', *Journal of Financial Economics* **119**(3), 457–471.

Gneiting, T. & Raftery, A. E. (2007), 'Strictly proper scoring rules, prediction, and estimation', *Journal of the American Statistical Association* **102**(477), 359–378.

Gneiting, T. & Ranjan, R. (2011), 'Comparing density forecasts using threshold-and quantile-weighted scoring rules', *Journal of Business & Economic Statistics* **29**(3), 411–422.

Gómez, H. W., Torres, F. J. & Bolfarine, H. (2007), 'Large-sample inference for the epsilon-skew-t distribution', *Communications in Statistics—Theory and Methods* **36**(1), 73–81.

Haario, H., Saksman, E. & Tamminen, J. (1999), 'Adaptive proposal distribution for random walk metropolis algorithm', *Computational Statistics* **14**(3), 375–396.

Hahn, P. R. & Carvalho, C. M. (2015), 'Decoupling shrinkage and selection in bayesian linear models: a posterior summary perspective', *Journal of the American Statistical Association* **110**(509), 435–448.

Hansen, B. E. (1994), 'Autoregressive Conditional Density Estimation', *International Economic Review* **35**(3), 705–730.

Harvey, A. C. (2013), *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, Vol. 52, Cambridge University Press.

Harvey, A. & Luati, A. (2014), 'Filtering With Heavy Tails', *Journal of the American Statistical Association* **109**(507), 1112–1122.

Harvey, A. & Thiele, S. (2016), 'Testing against changing correlation', *Journal of Empirical Finance* **38**, 575–589.

Harvey, C. R. & Siddique, A. (1999), 'Autoregressive conditional skewness', *Journal of Financial and Quantitative Analysis* **34**(4), 465–487.

Huber, F., Koop, G. & Onorante, L. (2020), 'Inducing sparsity and shrinkage in time-varying parameter models', *Journal of Business & Economic Statistics* **0**(0), 1–15.

Jensen, H., Petrella, I., Ravn, S. H. & Santoro, E. (2020), 'Leverage and deepening business-cycle skewness', *American Economic Journal: Macroeconomics* **12**(1), 245–81.

Jordà, Ò., Schularick, M. & Taylor, A. M. (2013), 'When credit bites back', *Journal of Money, Credit and Banking* **45**(s2), 3–28.

Jordà, Ò., Schularick, M. & Taylor, A. M. (2017), 'Macrofinancial history and the new business cycle facts', *NBER Macroeconomics Annual Report* **31**(1), 213–263.

Juárez, M. A. & Steel, M. F. (2010), 'Model-based clustering of non-gaussian panel data based on skew-t distributions', *Journal of Business & Economic Statistics* **28**(1), 52–66.

Jurado, K., Ludvigson, S. C. & Ng, S. (2015), 'Measuring uncertainty', *American Economic Review* **105**(3), 1177–1216.

Kilian, L. & Manganelli, S. (2007), 'Quantifying the risk of deflation', *Journal of Money, Credit and Banking* **39**(2-3), 561–590.

Koopman, S. J., Lucas, A. & Zamojski, M. (2018), 'Dynamic term structure models with score-driven time-varying parameters: estimation and forecasting', *Vrije Universiteit Amsterdam, working paper* .

Krishnamurthy, A. & Muir, T. (2017), How credit cycles across a financial crisis, Working Paper 23850, National Bureau of Economic Research.

Lucas, A. & Zhang, X. (2016), 'Score-driven exponentially weighted moving averages and value-at-risk forecasting', *International Journal of Forecasting* **32**(2), 293–302.

McConnell, M. M. & Perez-Quiros, G. (2000), 'Output fluctuations in the united states: What has changed since the early 1980's?', *American Economic Review* **90**(5), 1464–1476.

Mian, A. & Sufi, A. (2010), 'Household leverage and the recession of 2007–09', *IMF Economic Review* **58**(1), 74–117.

Morley, J. & Piger, J. (2012), 'The asymmetric business cycle', *Review of Economics and Statistics* **94**(1), 208–221.

Morley, J., Piger, J. & Tien, P.-L. (2013), 'Reproducing business cycle features: are nonlinear dynamics a proxy for multivariate information?', *Studies in Nonlinear Dynamics & Econometrics* **17**(5), 483–498.

Mudholkar, G. S. & Hutson, A. D. (2000), 'The epsilon–skew–normal distribution for analyzing near-normal data', *Journal of Statistical Planning and Inference* **83**(2), 291–309.

Neftci, S. N. (1984), 'Are economic time series asymmetric over the business cycle?', *Journal of Political Economy* **92**(2), 307–328.

Nyblom, J. (1989), 'Testing for the constancy of parameters over time', *Journal of the American Statistical Association* **84**(405), 223–230.

Plagborg-Møller, M., Reichlin, L., Ricco, G. & Hasenzagl, T. (2020), When is growth at risk?, Conference draft, Brooking Paper on Economic Activity.

Ray, P. & Bhattacharya, A. (2018), 'Signal adaptive variable selector for the horseshoe prior', *arXiv preprint arXiv:1810.09004* .

Reichlin, L., Ricco, G. & Hasenzagl, T. (2020), Financial variables as predictors of real growth vulnerability, Discussion Paper DP14322, CEPR.

Rossi, B. & Sekhposyan, T. (2019), 'Alternative tests for correct specification of conditional predictive densities', *Journal of Econometrics* **208**(2), 638–657.

Salgado, S., Guvenen, F. & Bloom, N. (2019), Skewed business cycles, Working Paper 26565, National Bureau of Economic Research.

Sims, C. A. & Zha, T. (1998), 'Bayesian methods for dynamic multivariate models', *International Economic Review* pp. 949–968.

Stark, T. & Croushore, D. (2002), 'Forecasting with a real-time data set for macroeconomists', *Journal of Macroeconomics* **24**(4), 507–531.

Stock, J. H. & Watson, M. W. (2002), 'Has the business cycle changed and why?', *NBER Macroeconomics Annual Report* **17**, 159–218.

Stock, J. H. & Watson, M. W. (2003), 'How did leading indicator forecasts perform during the 2001 recession?', *FRB Richmond Economic Quarterly* **89**(3), 71–90.

Taylor, J. W. (2019), 'Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric laplace distribution', *Journal of Business & Economic Statistics* **37**(1), 121–133.

Zhang, X. & Schwaab, B. (2016), 'Tail risk in government bond markets and ecb unconventional policies'. http://www.berndschwaab.eu/papers/ScoreTailRisk.pdf.

Modeling and forecasting macroeconomic downside risk Supplementary material

Davide Delle Monache [*] Andrea De Polis [†] Ivan Petrella [‡]

[*]Bank of Italy. DG Economics, Statistics and Research. davide.dellemonache@bancaditalia.it
[†]University of Warwick. Corresponding author: andrea.depolis.17@mail.wbs.ac.uk
[‡]University of Warwick, CEPR. ivan.petrella@wbs.ac.uk

A Score-driven setting

A.1 Score-based tests for time varying asymmetry

The Lagrange multiplier principle can be employed to formally test for the time variation of the parameter of interest (Harvey 2013, Section 2.5). When the conditional distribution is available in closed form, we can use the score of the log-likelihood function with respect to the parameter under consideration to investigate its autocorrelation properties. In fact, the scores incorporates information about the level of time variation of the respective parameter, and local power arguments indicate that the resulting test can be expected to be more powerful as the unconditional parameter moves away from zero. This is not the case with the standard moment-based portmanteau test, which simply uses the cross-product of standardized residuals (see, e.g. Bollerslev 1990).

Assume the conditional distribution of the GDP growth being a skew-t $y_t|Y_{t-1} \sim skt_\nu(\mu_t, \gamma, \delta)$, with log-likelihood

$$\ell_t(y_t|\theta, Y_{t-1}) = \log \mathcal{C}(\eta) - \frac{1}{2}\log\sigma^2 - \frac{1+\eta}{2\eta}\log\left[1 + \frac{\eta\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho)^2\sigma^2}\right], \tag{A1}$$
$$\log \mathcal{C}(\eta) = \log\Gamma\left(\frac{\eta+1}{2\eta}\right) - \log\Gamma\left(\frac{1}{2\eta}\right) - \frac{1}{2}\log\left(\frac{1}{\eta}\right) - \frac{1}{2}\log\pi$$

where $\Gamma(\cdot)$ is the Gamma function, $sgn(.)$ is the sign function, and $\eta = 1/\nu$ is the inverse of the degrees of freedom. To ensure positive scale and bounded shape, we model $\gamma = \log\sigma$, $\delta = \operatorname{arctanh}\varrho$, and the score vector with respect to μ, γ, and δ (scaled by inverse square root of information matrix as described later in details) will be equal to:

$$\begin{bmatrix} s_\mu \\ s_\gamma \\ s_\delta \end{bmatrix} = \sqrt{\frac{(1+3\eta)}{(1+\eta)}} \begin{bmatrix} \sqrt{(1-\varrho_t^2)}w_t\zeta_t \\ \sqrt{\frac{(1+\eta)}{2}}(w_t\zeta_t^2 - 1) \\ -\sqrt{\frac{(1-\varrho^2)}{3}}\frac{sgn(\varepsilon_t)}{(1-sgn(\varepsilon_t)\varrho)}w_t\zeta_t^2 \end{bmatrix}. \tag{A2}$$

where

$$w_t = \frac{(1+\eta)}{(1-sgn(\varepsilon_t)\varrho_t)^2 + \eta\zeta_t^2}, \quad \zeta_t = \frac{\varepsilon_t}{\sigma}. \tag{A3}$$

The scores in Equation (A2) can thus be used to test for the presence of time variation of the parameters. First, we fix the scale and shape parameters to their maximum likelihood estimates, $\hat{\gamma}$ and $\hat{\delta}$, respectively, and let μ_t being the only time-varying parameter. Then, using the estimated parameters we recover the time series of of the scores s_γ and s_δ, which we use to carry out the tests. Specifically, following Harvey & Thiele (2016), which apply this methodology to test against the time variation of correlations, we employ three different test specifications: the Portmanteau test, Q, the Ljung-Box test, Q^*, and the Nyblom (1989) test. We replace the k$^{\text{th}}$ order autocorrelation with the raw scores, under the assumption that parameters' dynamics

can be expressed as a moving average of past scores.

We apply the testing procedure to three different specifications for the dynamics of μ_t:

S1: $\mu_t \sim GAS(2,1)$ with stationary roots (see Section 2)

S2: $\mu_t \sim GAS(1,1)$ with a unit root.

S3: $\mu_t = 0$ and $y_t \sim AR(2)$ with skew-t innovations $skt_\nu(0, \sigma, \varrho)$.

For each case, we consider a specification with constant $\gamma = \log \sigma$, and another with time-varying $\gamma_t = \log \sigma_t$, following a non-stationary $GAS(1,1)$ process. Tests' results are reported in Table 1, Section 2, for S1, and in Table A2 and A1 for S2 and S3, respectively. For all the cases and specifications we are able to reject the null hypothesis of constant shape of Q and Q^*, at the 1% confidence level. Similarly, the Nyblom specification rejects the null, but the test provide less precise evidence. In fact, while the first two tests leave the alternative hypothesis unspecified, the latter tests for parameter constancy against a random walk law of motion. Therefore, we ascribe this lack of precision to the stationary nature of the scale parameter.

Table A1: Time-varying parameters tests

	time-varying location			*time-varying location & scale*		
	Q	Q^*	N	Q	Q^*	N
$Scale^2$	82.97^{***}	84.27^{***}	2.09^{***}			
$Shape$	44.30^{***}	45.00^{***}	0.46^{*}	11.30^{***}	11.48^{***}	0.58^{**}

Note: Specification: $y_t \sim skt_\nu(\mu_t, \sigma^2, \varrho)$, $\mu_t \sim RW$.
Q is the portmanteau test, Q^* is the Ljung-Box extension and N corresponds to the Nyblom test. The first two tests are distributed as a χ^2 with 1 degree of freedom, the Nyblom test statistics is instead distributed as a Cramer von-Mises distribution with 1 degree of freedom. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$

Table A2: Time-varying parameters tests

	time-varying location			*time-varying location & scale*		
	Q	Q^*	N	Q	Q^*	N
$Scale^2$	70.28^{***}	71.39^{***}	1.95^{***}			
$Shape$	46.4^{***}	47.13^{***}	0.41^{*}	21.44^{***}	21.78^{***}	0.78^{***}

Note: Specification: $y_t \sim AR(2),\ skt_\nu(0, \sigma^2, \varrho)$.
Q is the portmanteau test, Q^* is the Ljung-Box extension and N corresponds to the Nyblom test. The first two tests are distributed as a χ^2 with 1 degree of freedom, the Nyblom test statistics is instead distributed as a Cramer von-Mises distribution with 1 degree of freedom. $^{*}\ p < 10\%,\ ^{**}\ p < 5\%,\ ^{***}\ p < 1\%$

A.2 Scaled scores

Differentiating the log-likelihood function in Equation (A1) with respect to location, scale and asymmetry we obtain the gradient vector $\nabla_t = \left[\frac{\partial \ell_t}{\partial \mu_t}, \frac{\partial \ell_t}{\partial \sigma_t^2}, \frac{\partial \ell_t}{\partial \varrho_t}\right]'$, with elements:

$$\frac{\partial \ell_t}{\partial \mu_t} = \frac{1}{\sigma_t^2}\left[\frac{(1+\eta)\sigma_t^2\varepsilon_t}{(1-sgn(\varepsilon_t)\varrho_t)^2\sigma_t^2+\eta\varepsilon_t^2}\right] = \frac{1}{\sigma_t^2} w_t \varepsilon_t; \tag{A4}$$

$$\frac{\partial \ell_t}{\partial \sigma_t^2} = \frac{1}{2\sigma_t^2}\left[\frac{(1+\eta)\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho_t))^2\sigma_t^2+\eta\varepsilon_t^2} - 1\right] = \frac{1}{2\sigma_t^4}(w_t\varepsilon_t^2 - \sigma_t^2); \tag{A5}$$

$$\frac{\partial \ell_t}{\partial \varrho_t} = -\frac{sgn(\varepsilon_t)}{(1-sgn(\varepsilon_t)\varrho_t)}\frac{(1+\eta)}{(1-sgn(\varepsilon_t)\varrho_t))^2\sigma_t^2+\eta\varepsilon_t^2}\varepsilon_t^2 = -\frac{1}{\sigma_t^2}\frac{sgn(\varepsilon_t)}{(1-sgn(\varepsilon_t)\varrho_t)} w_t\varepsilon_t^2 \tag{A6}$$

where

$$w_t = \frac{(1+\eta)}{(1-sgn(\varepsilon_t)\varrho_t)^2+\eta\zeta_t^2} \tag{A7}$$

and $\zeta_t = \frac{\varepsilon_t}{\sigma_t}$ is the standardized innovations.

The Fisher information matrix is computed as the expected values of outer product of the gradient vector, for a fixed number of degrees of freedom:

$$\mathcal{I}_{t|t-1} = \mathbb{E}_{t-1}[\nabla_t \nabla_t'] = \begin{bmatrix} \frac{(1+\eta)}{(1+3\eta)(1-\varrho_t^2)\sigma_t^2} & 0 & -\frac{4c(1+\eta)}{\sigma_t(1-\varrho_t^2)(1+3\eta)} \\ 0 & \frac{1}{2(1+3\eta)\sigma_t^4} & 0 \\ -\frac{4c(1+\eta)}{\sigma_t(1-\varrho_t^2)(1+3\eta)} & 0 & \frac{3(1+\eta)}{(1-\varrho_t^2)(1+3\eta)} \end{bmatrix} \tag{A8}$$

Given we model $\gamma_t = \ln \sigma_t$ and $\delta_t = \operatorname{arctanh}(\varrho_t)$, for the chain rule we have:

$$\frac{\partial \ell_t}{\partial \gamma_t} = \frac{\partial \ell_t}{\partial \sigma_t^2}\frac{\partial \sigma_t^2}{\partial \gamma_t}, \quad \frac{\partial \ell_t}{\partial \delta_t} = \frac{\partial \ell_t}{\partial \varrho_t}\frac{\partial \varrho_t}{\partial \delta_t} \tag{A9}$$

where $\frac{\partial \sigma_t^2}{\partial \gamma_t} = 2\sigma_t^2$ and $\frac{\partial \varrho_t}{\partial \delta_t} = (1-\varrho_t^2)$. We can thus define the vector of interest as $f_t = (\mu_t, \gamma_t, \delta_t)'$

with the associated Jacobian matrix

$$J_t = \frac{\partial(\mu_t, \sigma_t^2, \varrho_t)}{\partial f_t'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2\sigma_t^2 & 0 \\ 0 & 0 & 1-\varrho_t^2 \end{bmatrix} \tag{A10}$$

As such the scaled score reads:

$$\mathbf{s}_t = (J_t' \text{diag}(\mathcal{I}_t) J_t)^{-\frac{1}{2}} J_t' \nabla_t = \begin{bmatrix} s_{\mu t} \\ s_{\sigma t} \\ s_{\varrho t} \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{(1+3\eta)(1-\varrho_t^2)}{(1+\eta)}} w_t \varepsilon_t \\ \sqrt{\frac{(1+3\eta)}{2}} (w_t \zeta_t^2 - 1) \\ -sgn(\varepsilon_t) \sqrt{\frac{(1+sgn(\varepsilon_t)\varrho_t)(1+3\eta)}{3(1-sgn(\varepsilon_t)\varrho_t)(1+\eta)}} w_t \zeta_t^2 \end{bmatrix}. \tag{A11}$$

To prevent numerical instability, it is often the case that we replace the scaling matrix $\dot{\mathcal{S}}_t = (J_t' \text{diag}(\mathcal{I}_t) J_t)^{-\frac{1}{2}}$ by its smoothed estimator, $\ddot{\mathcal{S}}_t = (1-\lambda)\ddot{\mathcal{S}}_{t-1} + \lambda \dot{\mathcal{S}}_t, 0 < \lambda < 1$.

A.3 Transformed parameters' scores

Here we provide a proof of the equivalence between the the score vector arising from the restriction imposed on the scale and shape parameters and that arising by imposing constraints on the two-component specification. We will drop the time subscript for the sake of clarity.

Proof. Let consider $\gamma = \log \sigma$ being the log-scale, it follows that $\sigma = \exp \gamma$, and the gradient is

$$\frac{\partial \ell}{\partial \gamma} = \frac{\partial \ell}{\partial \sigma} \frac{\partial \sigma}{\partial \gamma} = \frac{\partial \ell}{\partial \sigma} \sigma$$

Let now consider the multiplicative two-component counterpart $\sigma = \bar{\sigma}\tilde{\sigma}$, such that $\gamma = \log \sigma = \log \bar{\sigma} + \log \tilde{\sigma} = \bar{\gamma} + \tilde{\gamma}$. The gradient with respect to the first component reads

$$\frac{\partial \ell}{\partial \bar{\gamma}} = \frac{\partial \ell}{\partial \sigma} \frac{\partial \sigma}{\partial \gamma} \frac{\partial \gamma}{\partial \bar{\gamma}} = \frac{\partial \ell}{\partial \sigma} \sigma$$

The same applies for the second component. □

Proof. For the shape parameter $\varrho = \tanh \delta$, we model $\delta = \text{arctanh}\, \varrho$ and the gradient is

$$\frac{\partial \ell}{\partial \delta} = \frac{\partial \ell}{\partial \varrho} \frac{\partial \varrho}{\partial \delta} = \frac{\partial \ell}{\partial \varrho} (1 - \varrho^2)$$

Consider the two-component counterpart $\delta = (\bar{\delta} + \tilde{\delta})$, so that $\varrho = \tanh(\bar{\delta} + \tilde{\delta})$, it is easy to see that that

$$\frac{\partial \ell}{\partial \bar{\delta}} = \frac{\partial \ell}{\partial \varrho} \frac{\partial \varrho}{\partial \delta} \frac{\partial \delta}{\partial \bar{\delta}} = \frac{\partial \ell}{\partial \varrho} (1 - \varrho^2)$$

The same applies for the second component. □

B Data

In this Appendix we provide additional details on the data sources.

B.1 Main data sources and mnemonics

Data on quarterly real economic activity come from the Federal Reserve Bank of St. Louis FRED dataset (mnemonic; GDPC1). The NFCI and the relative risk, leverage, nonfinancial leverage and credit subcomponents are downloaded from the same source (mnemonic: NFCI, NFCIRISK, NFCILEVERAGE, NFCINONFINLEVERAGE and NFCICREDIT). These latter are available at the weekly frequency, and they are converted into quarterly figure by taking the quarterly average. Overlapping weeks are accounted for in the averaging process: weekly values are computed in only one of the quarterly variables.[1]

B.2 Additional details on the disaggregated financial indicators

We are grateful to Scot Brave and the Federal Reserve Bank of Chicago for sharing the weighted individual contribution to the NFCI. The dataset contains weekly time series of financial indicators that feed into the factor model generating both the NFCI and its subcomponens, as illustrated by Brave & Butters (2012). In Tables B1, B2, B3 and B4 we report all the indicators, their mnemonic, and the date they become available for the index. Figure B1 provides a graphical illustration of the time schedule of the availability of the predictors. The y-axis reports the years considered in our work in an ascending order (e.g. most recent times are closer to the origins).

[1] See https://fredhelp.stlouisfed.org/fred/data/understanding-the-data/how-are-data-aggregated-when-periods-overlap/

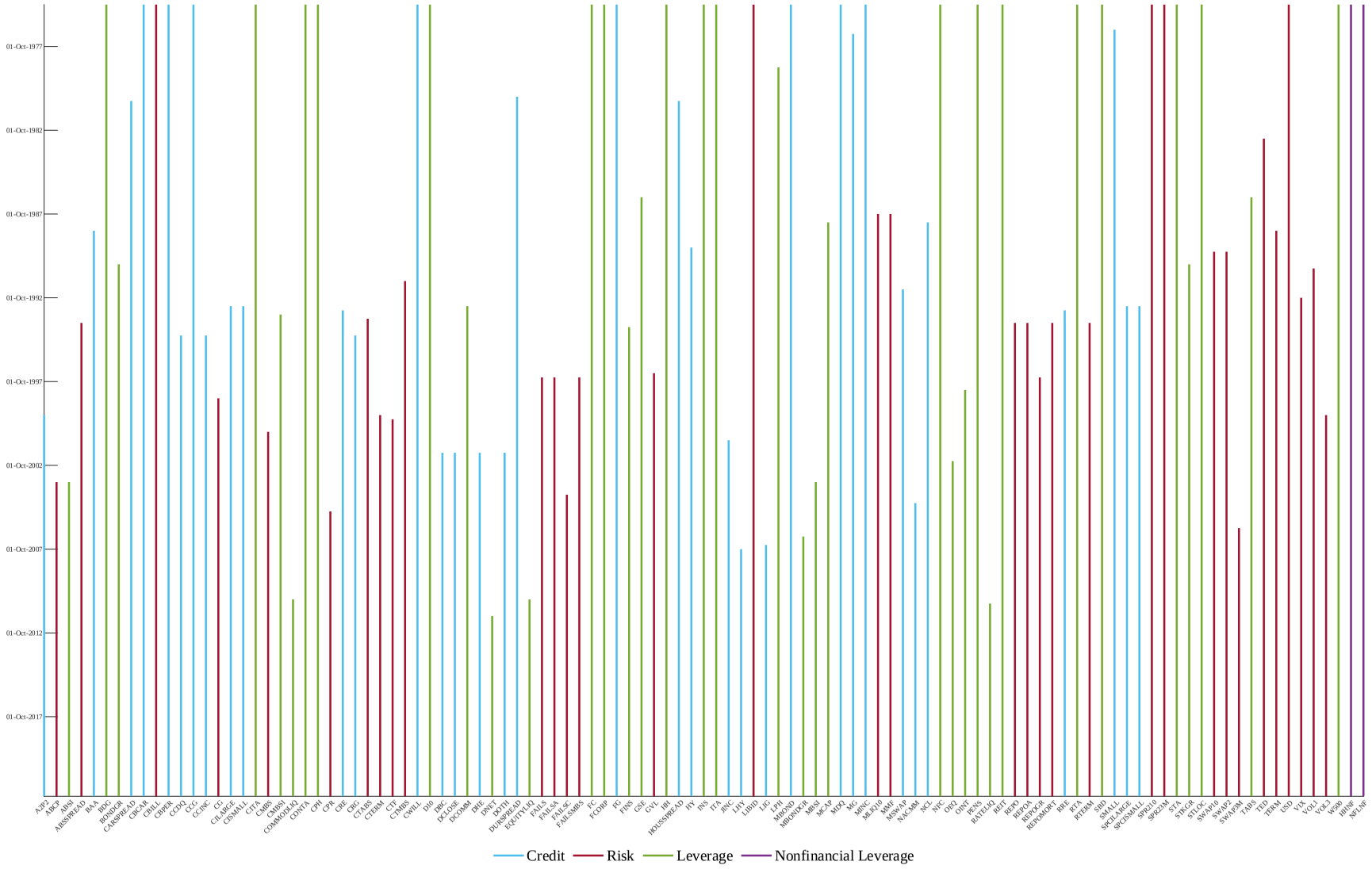

Figure B1: Timetable of predictors

Note: The chart report the availability of the 105 predictors over time, during the period 1973Q1 to 2018Q4.

Table B1: Risk subindex components

Mnemonic	Financial Indicator	Starting date
ABCP	1-mo. Asset-backed/Financial commercial paper spread	05/01/2001
ABSSPREAD	BofAML Home Equity ABS/MBS yield spread	05/07/1991
CBILL	3-mo. Financial commercial paper/Treasury bill spread	08/01/1971
CG	Commercial Paper Outstanding	10/11/1995
CMBS	BofAML 3-5 yr AAA CMBS OAS spread	02/01/1998
CPR	Counterparty Risk Index (formerly maintained by Credit Derivatives Research)	13/09/2002
CTABS	FTSE Russell US Global Markets ABS/5-yr Treasury yield spread	01/02/1991
CTERM	3-mo./1-wk AA Financial commercial paper spread	10/01/1997
CTF	FTSE Russell US Global Markets Financial/Corporate Credit bond spread	31/01/1997
CTMBS	FTSE Russell US Global Markets MBS/10-year Treasury yield spread	27/01/1989
FAILS	Treasury Repo Delivery Fails Rate	07/10/1994
FAILSA	Agency Repo Delivery Failures Rate	07/10/1994
FAILSC	Corporate Securities Repo Delivery Failures Rate	05/10/2001
FAILSMBS	Agency MBS Repo Delivery Failures Rate	07/10/1994
GVL	FDIC Volatile Bank Liabilities	01/07/1994
LIBID	3-mo. Eurodollar spread (LIBID-Treasury)	08/01/1971
MLIQ10	On-the-run vs. Off-the-run 10-yr Treasury liquidity premium	04/01/1985
MMF	Total Money Market Mutual Fund Assets/Total Long-term Fund Assets	28/12/1984
REPO	Fed Funds/Overnight Treasury Repo rate spread	24/05/1991
REPOA	Fed Funds/Overnight Agency Repo rate spread	24/05/1991
REPOGR	Repo Market Volume (Repurchases+Reverse Repurchases of primary dealers)	07/10/1994
REPOMORT	Fed Funds/Overnight MBS Repo rate spread	24/05/1991
RTERM	3-mo./1-wk Treasury Repo spread	24/05/1991
SPR210	10-yr/2-yr Treasury yield spread	20/08/1971
SPR23M	2-yr/3-mo. Treasury yield spread	08/01/1971
SWAP10	10-yr Interest Rate Swap/Treasury yield spread	03/04/1987
SWAP2	2-yr Interest Rate Swap/Treasury yield spread	03/04/1987
SWAP3M	3-mo. Overnight Indexed Swap (OIS)/Treasury yield spread	19/09/2003
TED	3-mo. TED spread (LIBOR-Treasury)	06/06/1980
TERM	1-yr/1-mo. LIBOR spread	10/01/1986
USD	Advanced Foreign Economies Trade-weighted US Dollar Value Index	12/01/1973
VIX	CBOE Market Volatility Index VIX	05/01/1990
VOL1	1-mo. BofAML Option Volatility Estimate Index	08/04/1988
VOL3	3-mo. BofAML Swaption Volatility Estimate Index	06/12/1996

Table B2: Leverage subindex components

Mnemonic	Financial Indicator	Starting date
ABSI	Nonmortgage ABS Issuance (Relative to 12-mo. MA)	29/12/2000
BDG	Broker-dealer Debit Balances in Margin Accounts	29/01/1971
BONDGR	New US Corporate Debt Issuance (Relative to 12-mo. MA)	01/01/1988
CITA	Commercial Bank C&I Loans/Total Assets	02/03/1973
CMBSI	CMBS Issuance (Relative to 12-mo. MA)	28/12/1990
COMMODLIQ	COMEX Gold/NYMEX WTI Futures Market Depth	04/01/2008
CONTA	Commercial Bank Consumer Loans/Total Assets	02/03/1973
CPH	FRB Commercial Property Price Index	02/04/1971
D10	10-yr Constant Maturity Treasury yield	08/01/1971
DCOMM	Commercial Bank Total Unused C&I Loan Commitments/Total Assets	29/06/1990
DNET	Net Notional Value of Credit Derivatives	07/11/2008
EQUITYLIQ	CME E-mini S&P Futures Market Depth	04/01/2008
FC	Total Assets of Finance Companies/GDP	02/04/1971
FCORP	Total Assets of Funding Corporations/GDP	02/04/1971
FINS	S&P 500 Financials/S&P 500 Price Index (Relative to 2-yr MA)	06/09/1991
GSE	Total Agency and GSE Assets/GDP	30/12/1983
INS	Total Assets of Insurance Companies/GDP	02/04/1971
ITA	Fed funds and Reverse Repurchase Agreements/Total Assets of Commercial Banks	30/03/1973
LPH	CoreLogic National House Price Index	02/04/1976
MBONDGR	New State & Local Government Debt Issues (Relative to 12-mo.h MA)	27/02/2004
MBSI	Total MBS Issuance (Relative to 12-mo. MA)	29/12/2000
MCAP	S&P 500, NASDAQ, and NYSE Market Capitalization/GDP	28/06/1985
OEQ	S&P 500, S&P 500 mini, NASDAQ 100, NASDAQ mini Open Interest	24/09/1999
OINT	3-mo. Eurodollar, 10-yr/3-mo. swap, 2-yr and 10-yr Treasury Open Interest	23/06/1995
PENS	Total Assets of Pension Funds/GDP	02/04/1971
RATELIQ	CME Eurodollar/CBOT T-Note Futures Market Depth	01/02/2008
REIT	Total REIT Assets/GDP	02/04/1971
RTA	Commercial Bank Real Estate Loans/Total Assets	02/03/1973
SBD	Total Assets of Broker-dealers/GDP	02/04/1971
STA	Commercial Bank Securities in Bank Credit/Total Assets	02/03/1973
STKGR	New US Corporate Equity Issuance (Relative to 12-mo. MA)	01/01/1988
STLOC	Federal, state, and local debt outstanding/GDP	02/04/1971
TABS	Total Assets of ABS issuers/GDP	30/12/1983
W500	Wilshire 5000 Stock Price Index	29/01/1971

Table B3: Nonfinancial Leverage subindex components

Mnemonic	Financial Indicator	Starting date
HH	Household debt outstanding/PCE Durables and Residential Investment	02/04/1971
NFC	Nonfinancial business debt outstanding/GDP	02/04/1971

Table B4: Credit subindex components

Mnemonic	Financial Indicator	Starting date
A2P2	1-mo. Nonfinancial commercial paper A2P2/AA credit spread	10/01/1997
BAA	Moody's Baa corporate bond/10-yr Treasury yield spread	03/01/1986
CARSPREAD	UM Household Survey: Auto Credit Conditions Good/Bad spread	24/02/1978
CBCAR	Commercial Bank 48-mo. New Car Loan/2-yr Treasury yield spread	05/05/1972
CBPER	Commercial Bank 24-mo. Personal Loan/2-yr Treasury yield spread	05/05/1972
CCDQ	S&P US Bankcard Credit Card: 3-mo. Delinquency Rate	28/02/1992
CCG	Consumer Credit Outstanding	29/01/1971
CCINC	S&P US Bankcard Credit Card: Excess Rate Spread	31/01/1992
CILARGE	FRB Senior Loan Officer Survey: Tightening Standards on Large C&I Loans	13/07/1990
CISMALL	FRB Senior Loan Officer Survey: Tightening Standards on Small C&I Loans	13/07/1990
CRE	FRB Senior Loan Officer Survey: Tightening Standards on CRE Loans	12/10/1990
CRG	S&P US Bankcard Credit Card: Receivables Outstanding	28/02/1992
CWILL	FRB Senior Loan Officer Survey: Willingness to Lend to Consumers	15/01/1971
DBC	ABA Value of Delinquent Bank Card Credit Loans/Total Loans	26/02/1999
DCLOSE	ABA Value of Delinquent Consumer Loans/Total Loans	26/02/1999
DHE	ABA Value of Delinquent Home Equity Loans/Total Loans	26/02/1999
DOTH	ABA Value of Delinquent Noncard Revolving Credit Loans/Total Loans	26/02/1999
DURSPREAD	UM Household Survey: Durable Goods Credit Conditions Good/Bad spread	27/01/1978
FG	Finance Company Owned & Managed Receivables	29/01/1971
HOUSSPREAD	UM Household Survey: Mortgage Credit Conditions Good/Bad spread	24/02/1978
HY	BofAML High Yield/Moody's Baa corporate bond yield spread	07/11/1986
JINC	30-yr Jumbo/Conforming fixed rate mortgage spread	12/06/1998
LHY	Markit High Yield (HY) 5-yr Senior CDS Index	07/01/2005
LIG	Markit Investment Grade (IG) 5-yr Senior CDS Index	01/10/2004
MBOND	20-yr Treasury/State & Local Government 20-yr GO bond spread	08/01/1971
MDQ	MBA Serious Delinquencies	30/06/1972
MG	Money Stock: MZM	01/03/1974
MINC	30-yr Conforming Mortgage/10-yr Treasury yield spread	02/04/1971
MSWAP	Bond Market Association Municipal Swap/20-yr Treasury yield spread	07/07/1989
NACMM	NACM Survey of Credit Managers: Credit Manager's Index	15/02/2002
NCL	Commercial Bank Noncurrent/Total Loans	28/06/1985
RRE	FRB Senior Loan Officer Survey: Tightening Standards on RRE Loans	12/10/1990
SMALL	NFIB Survey: Credit Harder to Get	02/11/1973
SPCILARGE	FRB Senior Loan Officer Survey: Increasing spreads on Large C&I Loans	13/07/1990
SPCISMALL	FRB Senior Loan Officer Survey: Increasing spreads on Small C&I Loans	13/07/1990

C Bayesian estimation

In this section we outline the Bayesian approach for the estimation of the models. We estimate the vector of static parameters θ, and the time-varying parameters f_t, through the following steps:

(i) Estimate θ^* via ML and initialize the algorithm with θ^*

(ii) Sample θ and f_t with the Adaptive Random-Walk Metropolis-Hastings (ARWMH)

(iii) Compute the statistics of interest, such as the percentiles of the empirical distribution function of the densities of interest.

In the following subsections, we detail the prior specifications and the ARWMH algorithm.

C.1 Prior specification

General prior specifications for the parameters are:

$$\phi \sim NID(\mu_\phi, \sigma_\phi) \cdot I_{(\phi \in \Phi)} \qquad \text{(C1)} \qquad \beta_i \sim NID(\mu_\beta, \sigma_\beta^i),\ i = \mu, \sigma, \varrho \qquad \text{(C2)}$$

$$\varsigma, \kappa \sim \mathcal{G}^{-1}(a_j, b_j),\ j = \varsigma, \kappa \qquad \text{(C3)} \qquad \eta \sim \mathcal{G}^{-1}(a_\eta, b_\eta) \cdot I_{(\eta \in H)} \qquad \text{(C4)}$$

Prior specification C1 sets the prior distribution for the first autroregressive parameters: we target high persistence, with $\mu_\phi \simeq 1$, with a standard tightness of $\sigma_\mu = 0.2$, in line with Bayesian Vector Autoregressive models (see, e.g., Doan et al. 1984, Sims & Zha 1998). We restrict the prior distribution to only span the stationary region, Φ, by truncating the support, which gives rise to an improper prior distribution as in Cogley & Sargent (2005). Predictor loadings are drawn from a Normal distribution with $\mu_\beta = 0$, and standard deviations σ_β are set to small values of 0.05, 0.2 and 0.1 for the location, scale and shape parameters, respectively. These priors aim at preventing model overfitting, applying an L_2 regularization, akin to the shrinkage induced by a Ridge-type regression. We set inverse gamma priors for the score loadings, such that they reflect the a priori expectation of small, but positive coefficients, in line with the properties of the score-driven filters (for further discussion, see Blasques et al. 2014). We set $a_\varsigma = a_\kappa = 4$, and $b_\varsigma = 3$ and $b_\kappa = 10$, to reflect the propertied of the long- and the short-run components. Eventually, we use an inverse gamma prior for η, the inverse of the degrees of freedom parameter, with $a_\eta = 2$ and $b_\eta = 10$. In line with Juárez & Steel (2010), these values allow the distribution to explore a wide range of feasible values, with a mean of 20 and a median of 10. In order to ensure the existence of, at least, the first three moments, we restrict the support to the $H = [0, 0.34]$ set.

We also target the stationarity of the cyclical component of the location through a prior on the sum of the autoregressive coefficients (Sims & Zha 1998). Let ϕ^S be the sum of two generic

autoregressive coefficients, $\phi^S = \phi^1 + \phi^2$, we assume

$$\phi^S \sim \mathcal{N}(\mu_{\phi^S}, \sigma_{\phi^S}) \cdot I_{(\phi^S \in \Phi)}$$

with $\mu_{\phi^2} \simeq 1$ and and $\sigma^s_\phi = 0.2$.

C.2 Horseshoe priors

Consider the regression model:

$$y_t = b' X_t + e_t, \quad e_t \sim iid(0, \sigma_e^2 I_p) \tag{C5}$$

the Horseshoe (HS) priors of (Carvalho et al. 2010) posit a Ridge-type priors for the coefficients: $b^j \sim \mathcal{N}(0, \lambda^j \tau)$, where the hyperparameters λ^j and τ control the local (coefficient specific) and the global shrinkage, respectively. We follow the original paper in setting the hyperpriors as:[2]

$$\lambda^j \sim HC^+(0,1), \quad \tau \sim HC^+(0,1) \tag{C6}$$

where $HC^+(0,1)$ denotes the standard Half-Cauchy distributions with density function

$$p(z) = \frac{2}{\pi(1+z^2)}, \quad z \in \mathbb{R}^+ \tag{C7}$$

Unlike other common shrinkage priors (e.g. Ridge, Lasso), the HS priors is free of exogenous inputs, implying a fully adaptive shrinkage procedure. Despite expecting several coefficients to be shrunk to zero, a known pitfall of shrinkage operators is that such figures are not exactly set to zero, thus increasing parameter uncertainty. A common remedy to this pathology is the sparsification of the near-zero elements. Let p be the number of possible predictors, the sparsification problem can be cast as:

$$b^* = \underset{b}{\operatorname{argmin}} \left[\frac{1}{2} ||X\hat{b} - Xb||_2^2 + \sum_{j=1}^{p} m_j |b_j| \right] \tag{C8}$$

where the sparse coefficients b^* are obtained minimizing the Euclidean distance between Xb and the model fit obtained from the shrinkage operator. Additional penalty for the non-zero parameters is controlled by m_j, which represents a variable specific tuning parameter, generally chosen on the basis of computationally expensive methods (e.g. cross-validation). In their recent contribution, (Ray & Bhattacharya 2018) introduce a purely data-driven and less burdensome solution to the choice of the parameter m_j. They suggest to specify the tuning parameter as $m_j = |\hat{b}_j|^{-2}$ such that the j^{th} variable receives a penalization *"ranked in inverse-squared order of*

[2](Makalic & Schmidt 2015) exploit the link between the scale mixture of Inverse Gamma distributions to approximate the Half-Cauchy distributions to propose a tractable framework for posterior simulation via Gibbs sampling. However, this hierarchy features twice as may parameters (auxiliary variables) compared to the original specification.

the magnitude of the corresponding coefficient" (Ray & Bhattacharya 2018). The Signal Adaptive Variable Selector (SAVS) algorithm can then be expressed as:

$$b_j^* = sgn(\hat{b}_j)||X_j||^{-2} \max\left\{|\hat{b}_j| \cdot ||X_j||^2 - m_j, 0\right\} \tag{C9}$$

where $||\cdot||$ represents the Euclidean norm of the vector X_j.

C.3 Adaptive Metropolis-Hastings

Posterior estimates of the parameters are obtained via simulation by means of the Adaptive Metropolis-Hastings algorithm proposed by Haario et al. (1999). Given that estimated parameters lie in bounded regions of the parameter space, we augment the algorithm with a rejection step to prevent numerical instability due to invalid parameter draws. The algorithm works as follows: given the $(d \times 1)$ vector of static parameters $\theta = (\kappa', \phi', \eta)'$, as reported in Section 3, we define θ^j as the draw of the parameters at the j^{th} iteration of the sampler, generated from the random walk kernel:

$$\theta^j = \theta^{j-1} + \sigma_{j,s}^2 \epsilon \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \Sigma_H^j) \tag{C10}$$

where $\sigma_{j,s}$ is a *"heuristical"* scale parameter that contributes to the stability of the adaptive algorithm. Σ_H^j is the covariance matrix of the last H draws, that is, the covariance matrix is updated to reflect the variability of the last H iterations. Given these two elements, we postulate the following adaptive scheme: we compute the local acceptance rate $\tilde{\alpha}^s$ each s iterations; if the acceptance rate is significantly greater (smaller) than the 25-35%target rate, we adjust the scale parameter according to a rescaling function $r(\tilde{\alpha}^s)$

$$r(\tilde{\alpha}^s) = \sigma_{j,s} \times \begin{cases} \tanh(4\tilde{\alpha}^s + 0.24), & \tilde{\alpha}^s < 0.25 \\ 1, & 0.25 \leq \tilde{\alpha}^s \leq 0.35 \\ 0.88 + \tilde{\alpha}^{s2}, & \tilde{\alpha}^s > 0.35 \end{cases} \tag{C11}$$

Figure C1 plots the shape of the rescaling function against the admissible range of values for $\tilde{\alpha}^s$. Due to the asymmetry of the intervals before and after the 25-35% acceptance region, the two branches of the reshaping function features different slopes and curvatures.

In addition, every $U \leq H$ iterations, we use the last H draws to recalibrate the covariance matrix of the proposal distribution: $\Sigma_H^j = \frac{\tilde{K}}{\sqrt{H-1}}$, where $\tilde{K}$ is the $(H \times d)$ centered matrix of the last H draws, obtained as $K - \mathbb{E}[K]$; in this process we also reinitialize the scale parameter at the value of $\frac{2.38}{\sqrt{d}}$, as suggested by Gelman et al. (1996). According to the Metropolis-Hasting procedure, we accept θ^j with probability $p = \min\left\{1, \exp(f(\theta^j) - f(\theta^{j-1}))\right\}$ where $f(\theta) =$ log-likelihood $\times$ priors. Discarded draws are replaced by the latest accepted draw. In order to prevent failures of the algorithm, we discard draws that do not lie in the bounded region of the parameter space, replacing them as previously specified. We set the number of iterations

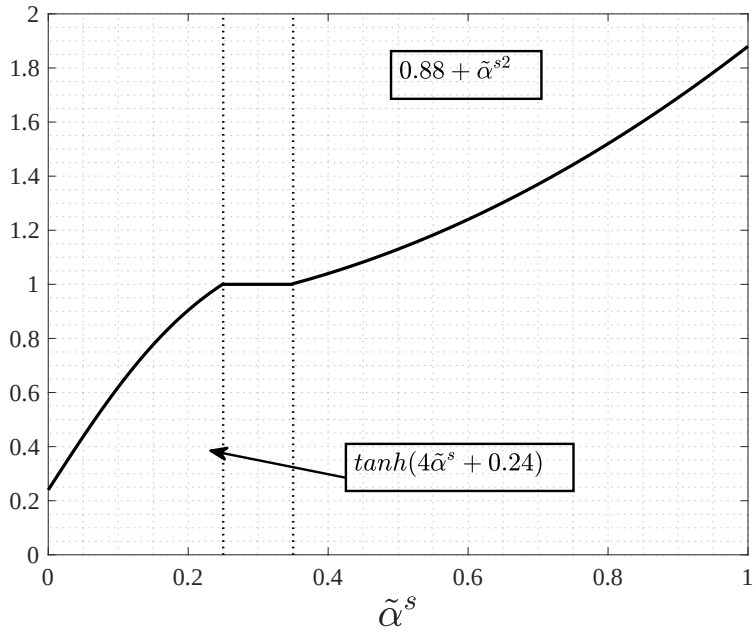

Figure C1: $r(\tilde{\alpha}^s)$ rescaling function

Note: The figure plots the rescaling function we apply to the variance of the candidate distribution to target an acceptance rate between 25% and 35%.

to 20000, of which we keep the last 40%. We reduce dependence between consecutive draws by *"thinnig"* each 2 realizations. Posterior distributions are eventually obtained from a sample of 4000 iterations. We adjust the scale parameter each hundred iterations ($s = 100$), while the adaptive step takes place every $U = 1500$ iterations, considering $H = 2500$ past iterations.[3]

[3]The adaptive process starts when $j \geq H$, while the scale updating takes place from the 100^{th} iteration

Algorithm 1: Adaptive Random-Walk Metropolis-Hastings

for *j=1,2,...,N* **do**

Generate: $\theta^* \sim \mathcal{N}(\theta^{j-1}, \sigma_{j,s}^2 \Sigma_H^j)$;

if $\theta^* \in \Theta$ **then**

Evaluate: $p = \min\left\{1, \exp(f(\theta^j) - f(\theta^{j-1}))\right\}$;

Draw: $u \sim U[0,1]$;

if $u < p$ **then**

$\theta^j = \theta^*, \quad \alpha^j = 1$;

else

$\theta^j = \theta^{j-1}, \quad \alpha^j = 0;$

end

else

$\theta^j = \theta^{j-1}, \quad \alpha^j = 0;$

end

if $j(\mod s) = 0$ **then**

$\sigma_{j,s} = \sigma_{j,s} * r(\tilde{\alpha}^s)$

end

if $j(\mod U) = 0$ **then**

$\Sigma_H^j = Cov(\{\theta^i\}_{i=j-H}^{H}), \quad \sigma_{j,s} = \frac{2.38}{\sqrt{d}}$

end

end

Note: the notation "$a(\mod b)$" represents the modulo operation, while "$-s$" indicates the previous s draws.

D Additional plots and tables

Table D1: Static parameters

	μ_y	μ_σ	$\phi_{y,1}$	$\phi_{y,2}$	$\phi_{\mu,1}$	$\phi_{\mu,2}$	$\phi_{\tilde{\gamma}}$	$\phi_{\tilde{\delta}}$
AR(2)	1.504 [1.400,1.622]	0.014 [0.013,0.016]	0.337 [0.303,0.383]	0.210 [0.171,0.249]			0.958 [0.954,0.960]	
Skt 1					0.984 [0.892,1.073]	−0.192 [−0.293,−0.100]	0.833 [0.778,0.877]	0.785 [0.738,0.828]
Skt 2					0.949 [0.862,1.034]	−0.176 [−0.271,−0.070]	0.675 [0.618,0.728]	0.703 [0.627,0.772]
Skt 3					1.048 [1.002,1.092]	−0.206 [−0.257,−0.156]	0.501 [0.434,0.569]	0.773 [0.720,0.813]

	ς_μ	κ_μ	ς_γ	κ_γ	ς_δ	κ_δ	χ	η
AR(2)				0.108 [0.095,0.121]			0.210 [0.123,0.303]	
Skt 1	0.023 [0.019,0.028]	0.077 [0.062,0.097]	0.029 [0.023,0.036]	0.096 [0.080,0.118]	0.022 [0.018,0.026]	0.095 [0.081,0.110]	0.265 [0.169,0.354]	0.201 [0.171,0.233]
Skt 2	0.025 [0.020,0.033]	0.074 [0.061,0.091]	0.026 [0.020,0.034]	0.071 [0.058,0.086]	0.026 [0.022,0.032]	0.067 [0.057,0.079]	0.293 [0.215,0.366]	0.173 [0.146,0.203]
Skt 3	0.023 [0.019,0.028]	0.072 [0.059,0.085]	0.025 [0.020,0.031]	0.066 [0.056,0.075]	0.021 [0.018,0.024]	0.075 [0.063,0.086]	0.109 [0.068,0.148]	0.151 [0.130,0.175]

Note: The table reports the estimates of the static parameters for all the models: the Gaussian *AR(2)*, the *Skt 1* without predictors, with lags of the NFCI (*Skt 2*) and with lags of the subcomponents of the index (*Skt 3*). Confidence intervals at the 68% level are in brackets.

Table D2: Predictor loadings

		Skt NFCI	*Skt 4DFI*			
		NFCI	Leverage	NF Leverage	Risk	Credit
t	μ_t	0.000 [−0.004,0.005]	−0.002 [−0.007,0.003]	−0.001 [−0.005,0.004]	0.000 [−0.004,0.004]	0.001 [−0.003,0.005]
	σ_t^2	0.124 [0.078,0.187]	0.059 [0.030,0.085]	−0.041 [−0.082,0.003]	−0.069 [−0.106,−0.036]	0.203 [0.165,0.249]
	ϱ_t	0.308 [0.259,0.348]	0.001 [−0.045,0.039]	−0.002 [−0.034,0.029]	0.140 [0.097,0.180]	0.274 [0.235,0.313]
$t-1$	μ_t	−0.001 [−0.005,0.003]	0.000 [−0.004,0.005]	0.000 [−0.004,0.004]	0.003 [−0.003,0.007]	0.000 [−0.005,0.006]
	σ_t^2	−0.042 [−0.088,0.018]	−0.051 [−0.081,−0.021]	0.018 [−0.023,0.061]	0.086 [0.055,0.119]	−0.089 [−0.132,−0.038]
	ϱ_t	−0.231 [−0.276,−0.185]	0.050 [0.003,0.085]	0.012 [−0.020,0.044]	−0.153 [−0.194,−0.109]	−0.249 [−0.290,−0.208]

Note: 68% confidence intervals are reported in brackets.

Table D3: Forecast scores - One quarter ahead

	AR(2)	*Skt NFCI*	*Skt 4DFI*	*AR(2)*	*Skt NFCI*	*Skt 4DFI*
	MSFE			**logS**		
Full	1.258 (0.120)	0.968 (0.850)	0.977 (0.775)	−0.073 (0.503)	0.081 (0.195)	0.031 (0.623)
Rec.	1.491 (0.122)	1.004 (0.298)	0.970 (0.799)	−0.425 (0.139)	0.114 (0.612)	0.120 (0.248)
GFC	1.696 (0.135)	1.053 (0.514)	0.965 (0.936)	−0.681 (0.171)	0.216 (0.473)	0.274 (0.334)
	CRPS			**wQS**		
Full	1.080 (0.254)	0.961 (0.546)	0.975 (0.774)	1.110 (0.180)	0.966 (0.664)	0.986 (0.943)
Rec.	1.221 (0.164)	0.966 (0.215)	**0.946** (0.077)	1.294 (0.125)	0.948 (0.565)	0.930 (0.236)
GFC	1.363 (0.141)	0.981 (0.278)	**0.907** (0.096)	1.428 (0.114)	0.969 (0.966)	0.917 (0.538)

Note: The table reports the average forecast scores, expressed in terms of the Adrian et al. (2019) model's scores; the p-value for Giacomini & White (2006) test (in parentheses) compares the models with respect to the benchmark model. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

Table D4: Forecast scores - One year ahead

	AR(2)	*Skt NFCI*	*Skt 4DFI*	*AR(2)*	*Skt NFCI*	*Skt 4DFI*
	MSFE			**logS**		
Full	1.308 (0.230)	**0.864** (0.094)	0.873 (0.625)	−0.152 (0.553)	0.154 (0.166)	0.113 (0.151)
Rec.	1.151 (0.286)	0.823 (0.161)	0.772 (0.329)	−0.500 (0.128)	0.151 (0.827)	0.212 (0.131)
GFC	1.118 (0.303)	**0.771** (0.000)	**0.652** (0.000)	**−0.390** (0.086)	**0.521** (0.005)	**0.802** (0.000)
	CRPS			**wQS**		
Full	1.140 (0.272)	0.959 (0.638)	0.964 (0.654)	1.242 (0.103)	0.957 (0.553)	0.962 (0.481)
Rec.	0.987 (0.676)	**0.810** (0.003)	**0.779** (0.002)	**1.222** (0.041)	0.893 (0.534)	0.833 (0.278)
GFC	0.951 (0.840)	**0.754** (0.000)	**0.698** (0.000)	1.160 (0.282)	**0.811** (0.003)	**0.692** (0.000)

Note: The table reports the average forecast scores, expressed in terms of the Adrian et al. (2019) model's scores; the p-value for Giacomini & White (2006) test (in parentheses) compares the models with respect to the benchmark model. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

Table D5: Density calibration tests

	Kolmogorov-Smirnov					
	AR(2)	*ABG*	*Skt 4DFI*	*AR(2)*	*ABG*	*Skt 4DFI*
	One-quarter ahead			**One-year ahead**		
Dist.	1.178	**1.911**	0.906	**2.399**	**2.050**	1.300
Left tail	**1.018**	**1.439**	0.866	**2.399**	**2.050**	**1.300**

Note: The table reports the test statistics for the Rossi & Sekhposyan (2019) tests, based on the Kolmogorov-Smirnov type tests. Values in **bold** indicate the rejection of the null hypothesis of correct specification of the density forecast, at the 10% confidence level. Critical values are obtained by 1000 bootstrap simulations. Gray shaded cells indicate the best score. The left tail score is computed over the support [0, 0.25].

Table D6: Downside risk scores

	Skt no-X	*Skt 4DFI*	*Skt no-X*	*Skt 4DFI*	*Skt no-X*	*Skt 4DFI*
	FZG		**ALS**		**TLF**	
	one-quarter-ahead					
Full	0.502 (0.356)	0.530 (0.323)	0.820 (0.206)	0.813 (0.174)	0.825 (0.293)	0.782 (0.229)
Rec.	0.297 (0.155)	0.317 (0.176)	0.566 (0.116)	0.541 (0.125)	0.640 (0.111)	0.562 (0.142)
GFC	0.254 (0.330)	0.192 (0.337)	0.558 (0.283)	0.464 (0.288)	0.654 (0.263)	0.507 (0.229)
	one-year-ahead					
Full	0.394 (0.335)	0.121 (0.303)	0.674 (0.240)	0.596 (0.302)	0.828 (0.307)	0.596 (0.300)
Rec.	**0.225** (0.027)	**0.286** (0.076)	**0.542** (0.082)	0.553 (0.448)	**0.589** (0.015)	**0.436** (0.019)
GFC	**0.394** (0.092)	0.121 (0.172)	**0.474** (0.056)	0.294 (0.187)	0.618 (0.122)	0.297 (0.198)

Note: The table reports the average downside risk test scores, expressed in terms of the Gaussian model's scores; the p-value for Giacomini & White (2006) test (in parentheses) compares the models with respect to the benchmark model. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score. FZG: Fissler et al. (2016) loss function; ALS: Taylor (2019) loss function; TLF: Giacomini & Komunjer (2005) tick loss function.

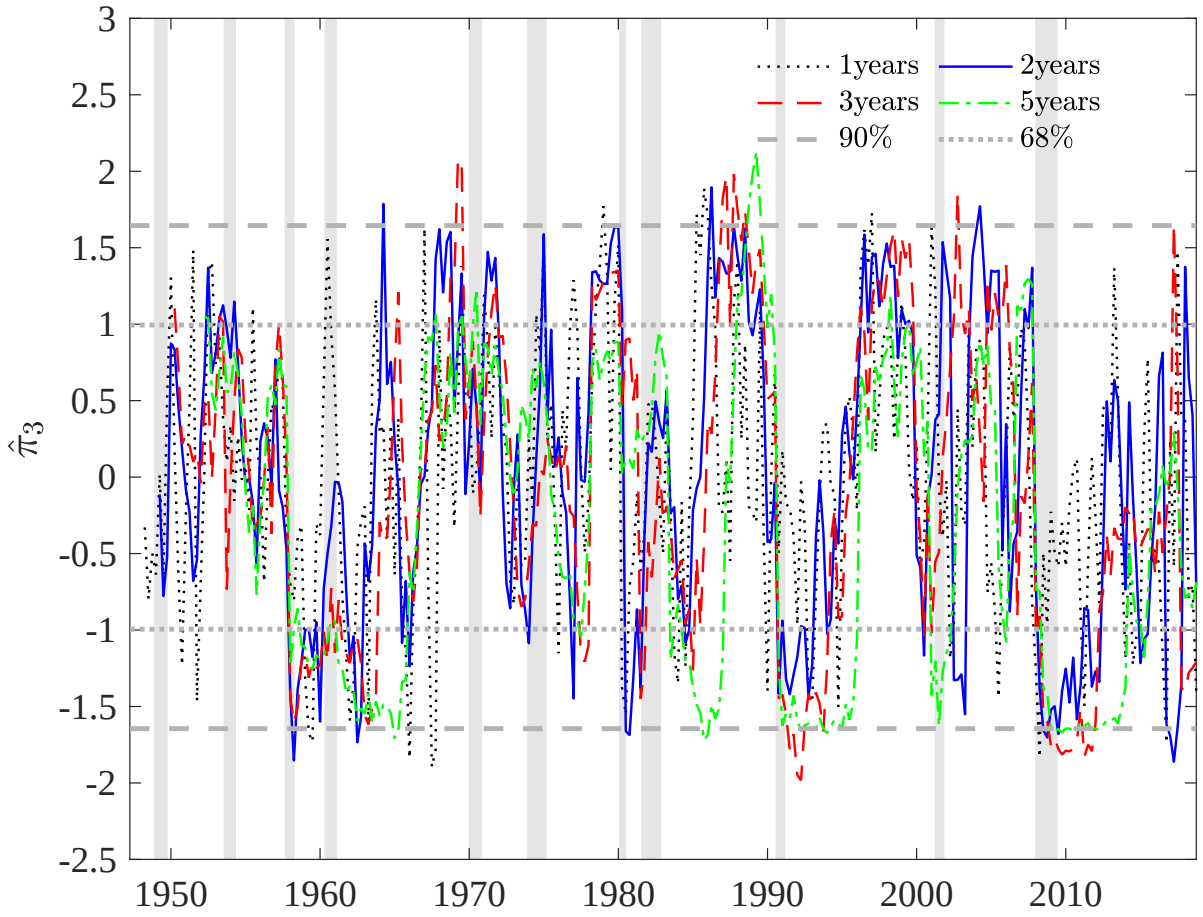

Figure D1: Bai & Ng (2005) test for skewness

Note: We report the Bai & Ng (2005) test for skewness' test statistic, for different rolling windows sizes of 2, 3 and 5 years. The gray lines represent the 68% (dashed) and 90% (dotted) critical values.

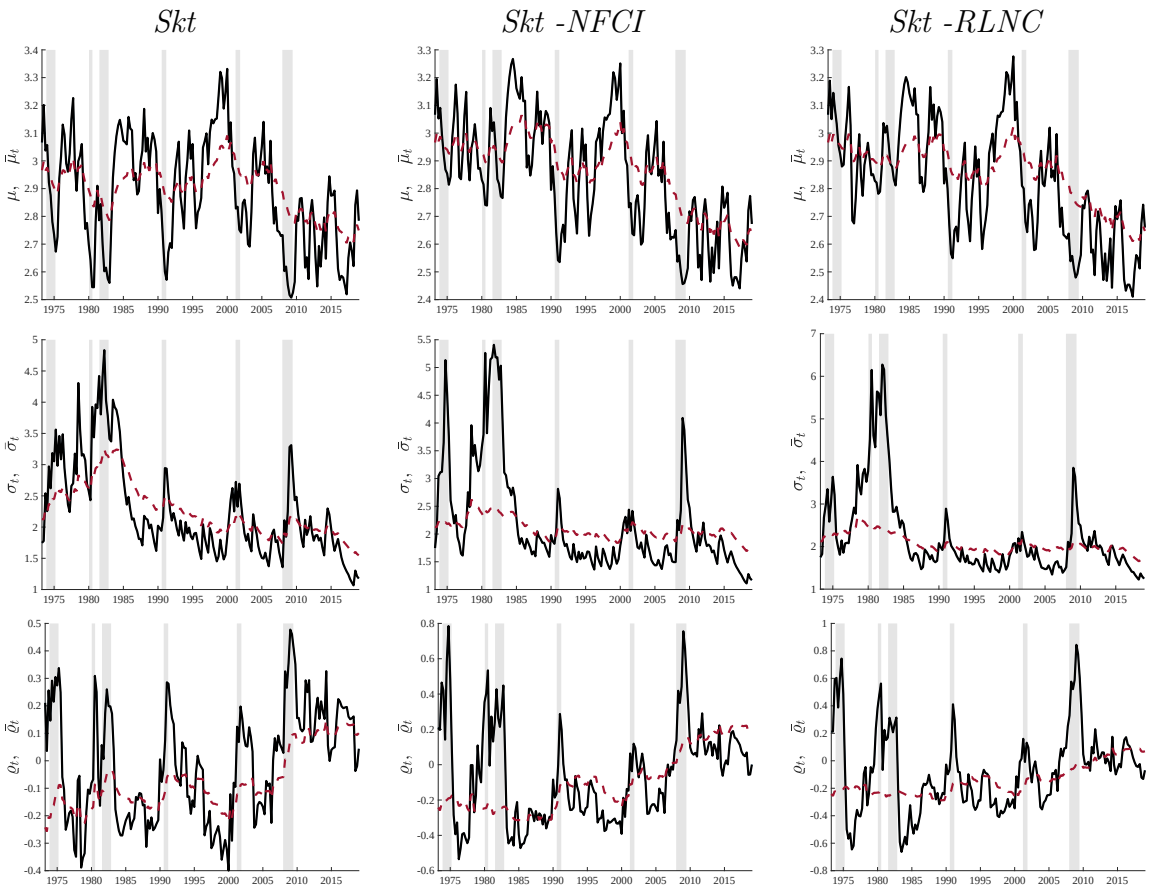

Figure D2: Time-varying parameters

Note: The plot illustrate the estimated time-varying parameters (black) for the three model specifications. Long-run components are reported in red (right scale).

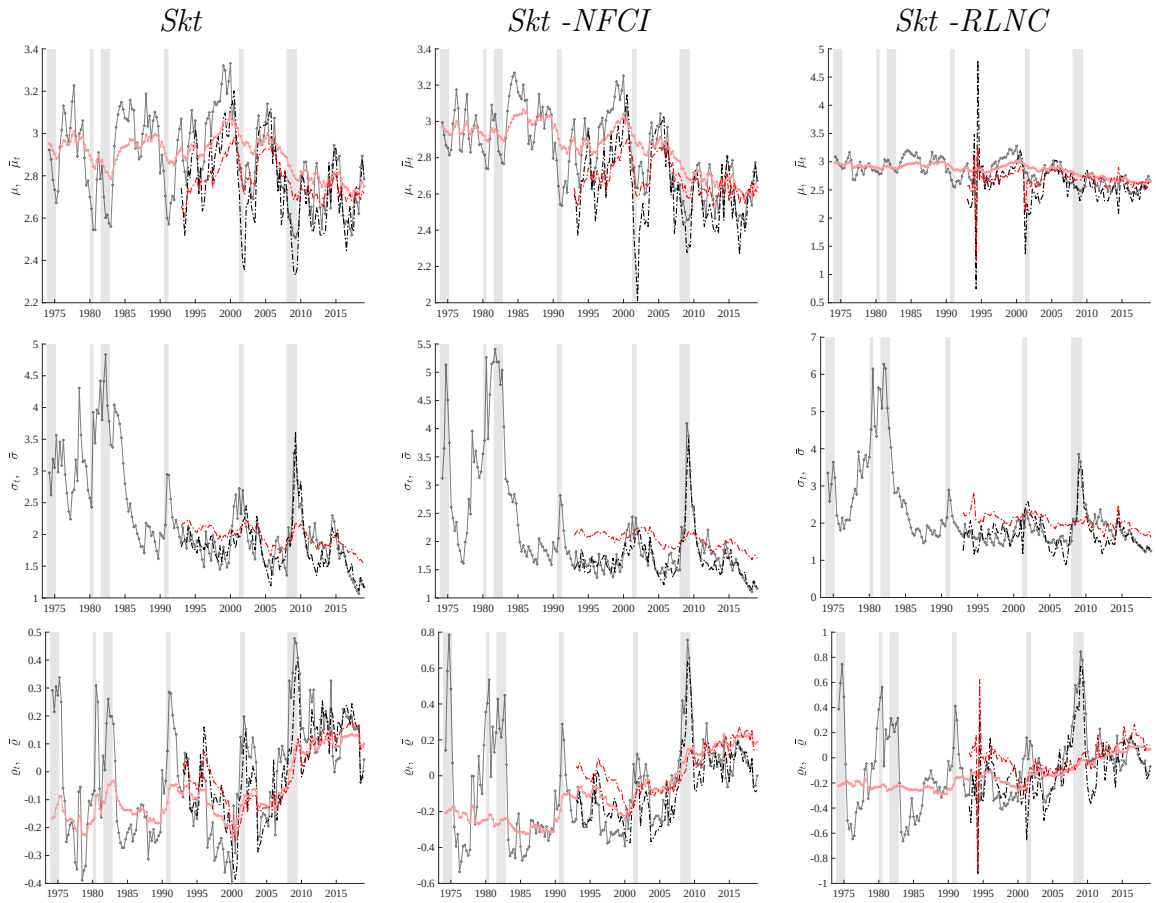

Figure D3: Time-varying parameters

Note: The plot compares the in sample (solid gray) and out of sample (dotted black) parameters. Long-run components are reported in scales of red (right scale).

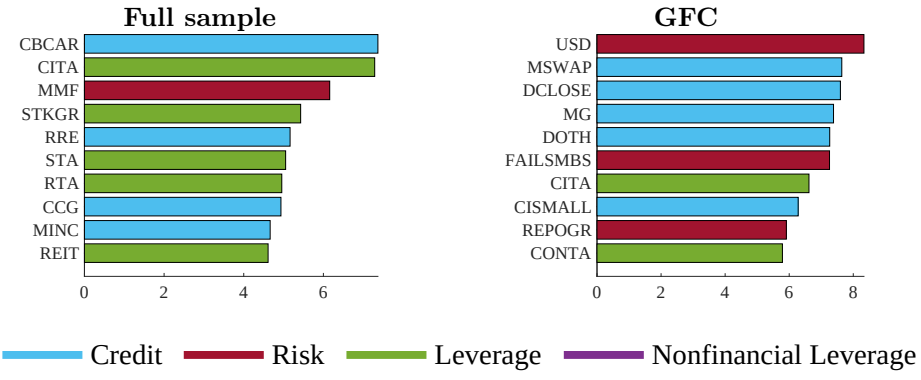

Figure D4: Top 10 predictors for μ_t

Note: The bar plots report the top 10 predictors for the location parameter. The *x-axis* reports average posterior probability of inclusion, expressed in percentage terms.

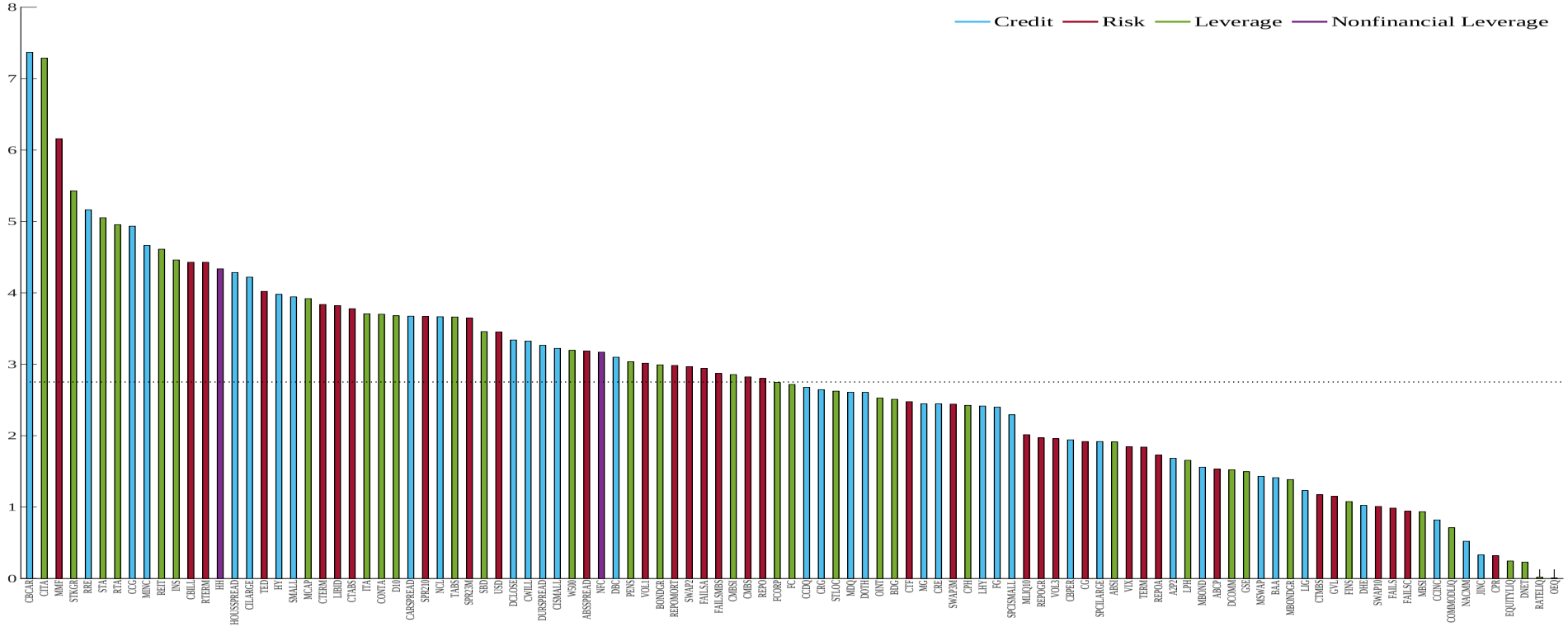

Figure D5: Predictors - Location

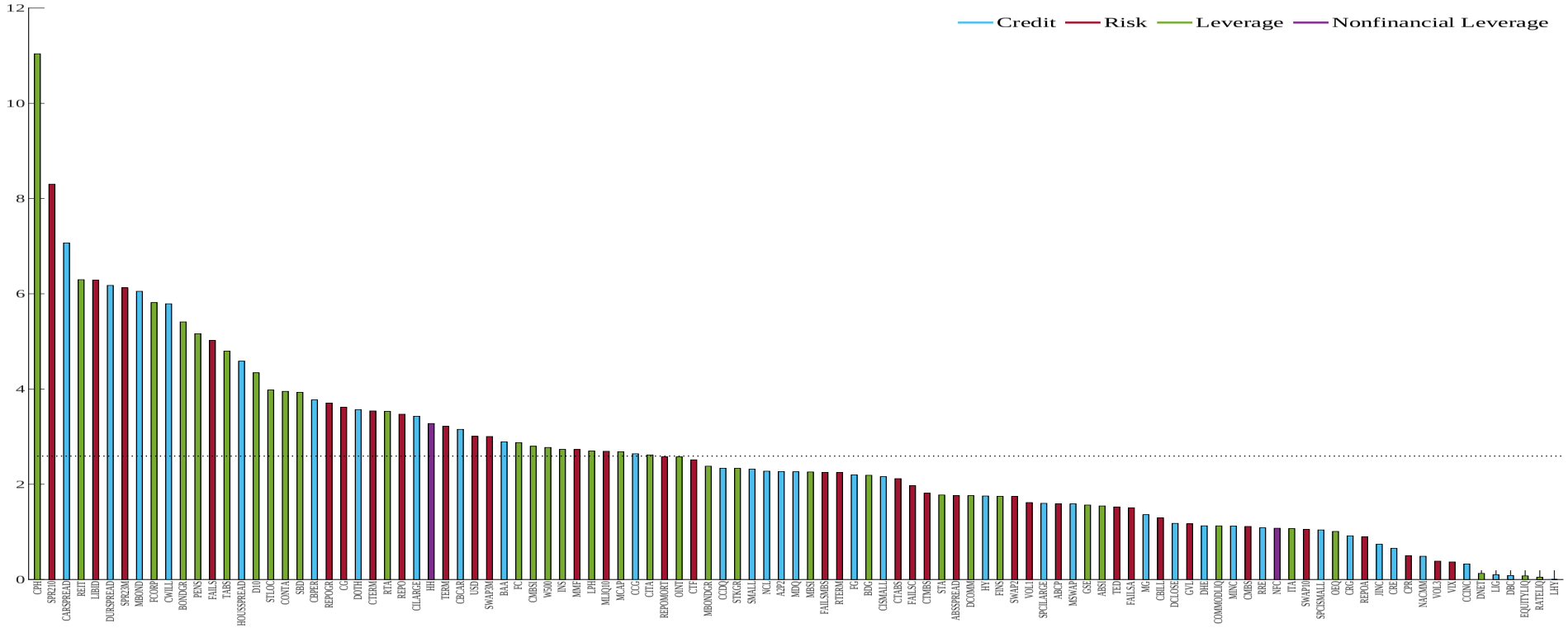

Figure D6: Predictors - Scale

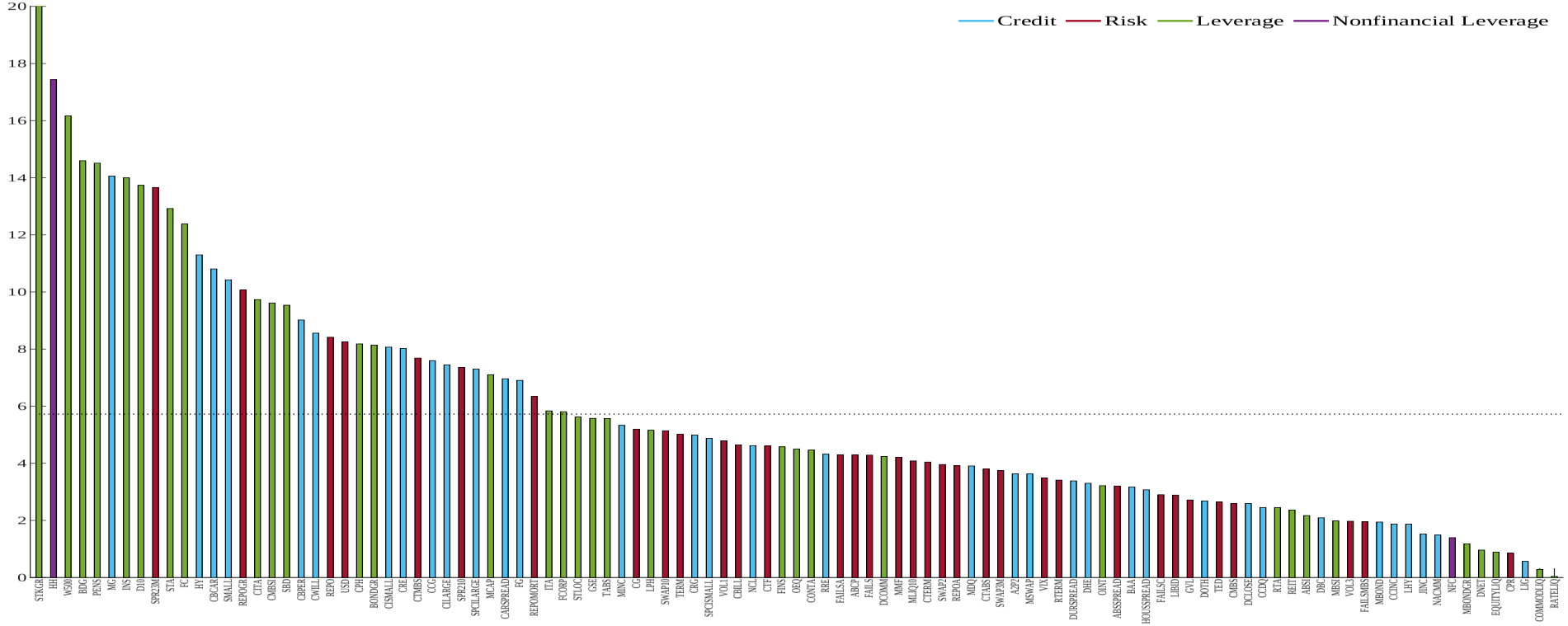

Figure D7: Predictors - Asymmetry

References

Blasques, F., Koopman, S. J., Lucas, A. et al. (2014), 'Stationarity and ergodicity of univariate generalized autoregressive score processes', *Electronic Journal of Statistics* **8**(1), 1088–1112.

Bollerslev, T. (1990), 'Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model', *The review of economics and statistics* pp. 498–505.

Carvalho, C. M., Polson, N. G. & Scott, J. G. (2010), 'The horseshoe estimator for sparse signals', *Biometrika* **97**(2), 465–480.

Cogley, T. & Sargent, T. J. (2005), 'Drifts and volatilities: monetary policies and outcomes in the post wwii us', *Review of Economic Dynamics* **8**(2), 262–302.

Doan, T., Litterman, R. & Sims, C. (1984), 'Forecasting and conditional projection using realistic prior distributions', *Econometric reviews* **3**(1), 1–100.

Gelman, A., Roberts, G. O., Gilks, W. R. et al. (1996), 'Efficient metropolis jumping rules', *Bayesian Statistics* **5**(599-608), 42.

Haario, H., Saksman, E. & Tamminen, J. (1999), 'Adaptive proposal distribution for random walk metropolis algorithm', *Computational Statistics* **14**(3), 375–396.

Harvey, A. C. (2013), *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, Vol. 52, Cambridge University Press.

Harvey, A. & Thiele, S. (2016), 'Testing against changing correlation', *Journal of Empirical Finance* **38**, 575–589.

Juárez, M. A. & Steel, M. F. (2010), 'Model-based clustering of non-gaussian panel data based on skew-t distributions', *Journal of Business & Economic Statistics* **28**(1), 52–66.

Makalic, E. & Schmidt, D. F. (2015), 'A simple sampler for the horseshoe estimator', *IEEE Signal Processing Letters* **23**(1), 179–182.

Nyblom, J. (1989), 'Testing for the constancy of parameters over time', *Journal of the American Statistical Association* **84**(405), 223–230.

Ray, P. & Bhattacharya, A. (2018), 'Signal adaptive variable selector for the horseshoe prior', *arXiv preprint arXiv:1810.09004* .

Sims, C. A. & Zha, T. (1998), 'Bayesian methods for dynamic multivariate models', *International Economic Review* pp. 949–968.