Modeling and Forecasting Macroeconomic Downside Risk*

Davide Delle Monache† Andrea De Polis‡ Ivan Petrella§

This draft: September, 2021

Abstract

We model secular trends and cyclical changes of the predictive density of US GDP growth. A substantial increase in downside risk to US economic growth emerges over the last 30 years, associated with the long-run growth slowdown started in the early 2000s. Conditional skewness moves procyclically, implying negatively skewed predictive densities ahead and during recessions, often anticipated by deteriorating financial conditions, while positively skewed distributions characterize expansions. The modelling framework ensures robustness to tail events, allows for either dense or sparse predictor designs, and delivers competitive out-of-sample (point, density and tail) forecasts, improving upon standard benchmarks.

Keywords: Business cycle, financial conditions, downside risk, skewness, score driven models.

JEL codes: C53, E32, E44

*We thank Jesús Fernández-Villaverde for the thoughtful discussion and Domenico Giannone for extensive comments on an earlier draft of the paper. We also thank Anastasia Allayioti, Daniele Bianchi, Scott Brave, Christian Brownlees, Andrew Butters, Andrea Carriero, Ana Galvao, Martin Iseringhausen, Francesco Saverio Gaudio, Gary Koop, Francesca Loria, André Lucas, Elmar Mertens, Massimiliano Marcellino, James Mitchell, Bernd Schwaab, Andreas Tryphonides, Fabrizio Venditti, the participants at the Conference on Score-Driven and Nonlinear Time Series Models at University of Cambridge, the participants at the workshop "Vulnerable Growth in the Euro Area" at European Central Bank, the participants at CFE-CMStatistics 2019 conference, the participants at the Conference on Real-Time Data Analysis, Methods, and Applications at the Philadelphia FED, and the seminar participants at University of Warwick and the University of Cyprus for valuable feedback and suggestions. We are grateful to the Chicago FED for making the full panel of weighted contribution of the financial indicators available for this work. The views expressed in this manuscript 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. davide.dellemonache@bancaditalia.it

‡University of Warwick. andrea.depolis.17@mail.wbs.ac.uk

§University of Warwick & CEPR. ivan.petrella@wbs.ac.uk

1 Introduction

The Global Financial Crisis and the subsequent recession left policymakers with several new challenges. In a world of persistently sluggish growth, subject to infrequent but deep recessions, the idea of central bankers as 'risk managers' gained renewed popularity (see, e.g., Cecchetti, 2008). In this environment, policy makers pursuing a '*plan for the worst, hope for the best*' approach rely on downside risk measures to assess the distribution of risk around modal forecasts. Recently, the sharp contraction, and the subsequent recovery in 2020 provided a sound reminder of the importance of accounting for tail events when assessing macroeconomic risk. Yet, assessing the degree of asymmetry of business cycle fluctuations remains a challenging task, and even more so it is to reliably gauge the time variation of downside risk. In addition, sound economic policy should consider the evolution of secular macroeconomic trends in pursuing the long-run goals of price stability and sustainable economic growth. Failure to account for permanent shifts in the properties of GDP growth can, in turn, lead to ineffective policy (see, e.g., Edge et al., 2007). In this paper, we introduce a generalised, comprehensive framework fit to provide policy guidance on the developments of downside risks, based on secular movement and cyclical variations of business cycle fluctuations.

We provide novel evidence in support of the presence of *conditional* asymmetry of GDP growth's distribution. Despite *unconditional* asymmetry remains unsupported by the data, conditional skewness, and thus downside risk to economic growth, exhibits significant variation over time. Motivated by this evidence, we introduce a novel, flexible methodology to model and forecast the conditional distribution of real economic growth with a parametric skewed Student-t (Skew-t) distribution with time-varying location, scale, and shape parameters. In order to fully capture the evolution of GDP growth's predictive densities, our framework accommodates secular trends and cyclical components for each parameter, with time variation being driven by the scaled score of the predictive likelihood function (Creal et al., 2013; Harvey, 2013), as well as by financial indicators. The latter allow us to explore to what extent downside risk to economic growth reflects imbalances arising in financial markets (Giglio et al., 2016; Adrian et al., 2019).

The conditional distribution of GDP growth is characterized by procyclical skewness fluctuations around a declining trend-skewness. At the onset of downturns, business cycle swings exhibit significant negative skewness, while expansions are characterized by positively skewed distributions. In turn, the well-documented counter-cyclicality of GDP growth's volatility (see Jurado et al., 2015, among others) largely reflects increasing downside volatility during recessions. We document that the fall in macroeconomic volatility since the mid-1980s, the so called Great Moderation (McConnell and Perez-Quiros, 2000; Stock and Watson, 2002), mainly reflects a significant reduction of upside volatility, with downside volatility remaining stable over the same period. During the last 30 years, skewness has decreased steadily implying a higher exposure to downside risk, which partially accounts for the slowdown in long-run growth observed since the early 2000s (see, e.g., Antolin-Diaz et al., 2017; Cette et al., 2016). Within our framework, the extreme realizations of the pandemic quarters are captured through movements in volatility and skewness,

allowing the model to remain remarkably stable, and suggesting that such outcomes were, to some extent, tail events. When assessed on its out-of-sample performance, our model delivers well calibrated predictive densities, improving upon competitive benchmarks in terms of point, density and tail forecasts, as well as leading to timely predictions of the odds of forthcoming recessions.

Several measures of financial distress have been identified as relevant indicators for predicting economic downturns.[1] We show that the four subcomponents of the National Financial Condition Index (NFCI, Brave and Butters, 2012), capturing risk, credit, leverage and nonfinancial leverage developments, help predict different features of the GDP growth distribution. The slow building up of leverage emerges as a key determinant of the scale of distribution, whereas skewness relates to all the subindices. We document that financial deepening during the expansionary phase of the cycle is associated with positive GDP growth's skewness, whereas tightening of financial conditions, especially the build-up of household debt, consistently predicts downside risk episodes. Similarly, financial conditions significantly improve the out-of-sample forecasting accuracy of our model, in particular during recessions.

Although aggregate measures succeed in summarizing a large amount of data, concerns that information relevant for assessing risk can remain undetected persist (see, e.g., Galvão and Owyang, 2018; Carriero et al., 2020b). To address this question, we perform an out-of-sample variable selection exercise based on the '*shrink-then-sparsify*' approach of Hahn and Carvalho (2015) using the full set of predictors feeding into the NFCI.[2] Indicators pertaining to the build up of leverage of financial institutions and households, as well as credit conditions, receive the least shrinkage over the full sample. During the latest financial crisis, the model selects indicators relating to developments in the shadow banking sector, credit spreads and mortgage-backed securities issuance. This highlights that that information in the balance sheet of the intermediary sector (Adrian and Shin, 2008), and the growing imbalances in the housing market (see, e.g., Gertler and Gilchrist, 2018) already provided considerable information about rising vulnerabilities in the economy, and thus about increasing downside risks ahead and during the financial crisis. We find that processing the signal from a large panel of financial predictors leads to substantial improvements in short term predictions, with sizable gains associated with the ability of the model to capture the increase in the downside risk ahead of the 2007-2009 Great Recession,[3] and especially during the most recent pandemic quarters.

Our results highlight the importance of accounting for asymmetric business cycle fluctuations. These can emerge through nonlinearities in the transmission of Gaussian shocks (see, e.g., Fernández-Villaverde and Guerrón-Quintana, 2020a), or alternatively reflect conditionally skewed shocks hitting the economy

[1] Among others, traditional spread measures (see, e.g., Rudebusch and Williams, 2009), credit and leverage growth (Drehmann et al., 2010; Jordà et al., 2013), as well as the leverage position of financial intermediaries (Adrian and Shin, 2008) received particular attention.

[2] Giannone et al. (2018) warn that sparsity can arise as an artefact of strong *a priori* beliefs. We approach sparsity within the purely data-driven framework of Ray and Bhattacharya (2018) which is robust to this concern, as also argued in Huber et al. (2020).

[3] The latter finding corroborates the evidence in Alessi et al. (2014), that the central banks' failure to predict the 2007-2009 downturn can be ascribed to their inability to read the signals from deteriorating financial conditions.

(as in Bekaert and Engstrom, 2017; Salgado et al., 2019). The procyclical skewness dynamics we document is found to be associated with a dynamic correlation between the first and second moments, being positive in expansions, and turning negative in recessions (in line with the empirical evidence in Carriero et al., 2020a). This further reinforces the necessity of distinguishing between "good" and "bad" uncertainty, which can potentially impact economic activity in opposite directions (Segal et al., 2015). Our findings also emphasise the need to account for the nonlinear relationship between financial conditions and credit availability, and the distribution of GDP growth (see, e.g., Fernández-Villaverde et al., 2019) for policy monitoring and stabilization policy design (Gadea Rivas et al., 2020; Jordà et al., 2020). Lastly, over the last three decades the fall in trend-skewness of economic fluctuations and the associated increase of downside risk emerges as a salient feature of the data which needs to be accounted by theoretical macroeconomic models (see, e.g., Jensen et al., 2020).

Related literature: This paper builds on the growing literature exploring the asymmetry characterizing business cycle fluctuations, and the relationship between real economic activity and financial conditions. Giglio et al. (2016) and Adrian et al. (2019) uncover a significant negative correlation between financial conditions and the lower quantiles of the conditional distribution of future economic growth by means of quantile regressions. We introduce a novel approach based on the modeling of the parameters of a Skew-t distribution. Out approach is based on a rich, yet parsimonious structure, and avoids the burdensome necessity to fit the estimated quantiles to a distribution. In particular, our model features persistence in the skewness of the distribution of GDP growth, in line with the term structure of the growth-at-risk displaying stronger asymmetry for the short- than for the medium-run (Adrian et al., 2018), and consistent with the pronounced skewness displayed by the Survey of Professional Forecasters' short-term predictions (Ganics et al., 2020). Differently to other contributions (see, e.g., Plagborg-Møller et al., 2020), we model both secular and cyclical changes of the distribution of GDP growth. This is essential in order to recover well-known stylized facts, such as the Great Moderation period (McConnell and Perez-Quiros, 2000; Stock and Watson, 2002) and the fall in long-run GDP growth (Antolin-Diaz et al., 2017), as well as to uncover an increasingly negatively skewed business cycle in the last part of the sample (Jensen et al., 2020).

A number of recent contributions have called into question the presence of asymmetry in business cycle fluctuations, and have suggested alternative ways of capturing time variation in downside risk (see, e.g., Brownlees and Souza, 2020; Carriero et al., 2020a). Our approach allows for, but does not impose, skewness in GDP growth. Yet, we document significant variation in the asymmetry of the conditional distribution of GDP growth, which is associated with substantial gains in out-of-sample forecasts and downside risk predictions over standard volatility models whose competitiveness has recently been highlighted by Brownlees and Souza (2020).

Existing models for conditional skewness rely on *ad hoc* law of motion for the time-varying parameters, and the asymmetry is updated as a function of higher-order powers of the residuals (Hansen, 1994; Harvey

and Siddique, 1999). In stead, we rely on the score-driven framework put forward by Creal et al. (2013) and Harvey (2013), which has proven to be particularly suitable for accommodating parameters' time variation under different distributional assumptions (Koopman et al., 2016). Within our setting, parameters update according to (highly) nonlinear functions of past prediction errors, depending, among other, on the shape of the conditional distribution at the time of the update. Thus, not only the updating mechanism adapts to the local properties of the data, but it is also robust to the presence of extreme realizations, contrary to model updates based on higher-order powers of the residuals. Within the score-driven setting, to the best of our knowledge, we are the first to rely on Bayesian estimation methods. This allows us to jointly tackle parameters' proliferation and incorporate estimation uncertainty when assessing the predictions of the model.

Structure: The remainder of the paper is organized as follows. Section 2 provides evidence of time-varying business cycle asymmetry. Section 3 presents the model, the estimation methodology and the forecasting procedure. Section 4 illustrates the characteristics of the conditional distribution of GDP growth, and how these relate to financial predictors. Section 5 reports the out-of-sample forecast and downside risk prediction evaluation. In Section 6 we investigate the predictive ability of the large set of financial indicators. Section 7 concludes.

2 Motivating evidence

Assessing the degree of skewness of GDP growth is notoriously challenging. When measured over the 1973-2020 (1973-2019) sample, we obtain a negative sample skewness of -2.58 (-0.42). Yet, due to the low precision of this estimate over the full sample, we cannot reject the null of symmetry using the Bai and Ng (2005) test. However, the absence of skewness in the unconditional distribution does not necessarily imply the conditional distribution being symmetric as well (Carriero et al., 2020a). The difficulty in obtaining significant skewness estimates can potentially reflect the dynamic nature of the asymmetry of economic fluctuations. Using the Bai and Ng (2005) test over different rolling windows of 4 to 6 years, the test often rejects the null of symmetry, with periods of significant negative and positive skewness detected over the sample.[4]

Harvey (2013, Section 2.5) highlights that the Lagrange multiplier principle can be employed to construct appropriate test statistics for the time variation of parameters. Within this framework, testing for parameters' variation requires starting from a benchmark model, under the null of no time variation of the parameter(s) of interest. Here, we assume that GDP growth follows an AR(2) process with Skew-t innovations, with a shape parameter pinning down the degree of asymmetry. We test for the time variation of this parameter considering both the case of constant volatility and the more realistic case of time-varying

[4]See Figure H1 in Appendix H, for a summary of these results.

Table 1: Score-based tests for time variation

	time-varying location			*time-varying location & scale*		
	Q	Q^*	N	Q	Q^*	N
$Scale^2$	7.187***	7.296***	0.979***			
Shape	8.497***	8.626***	0.603**	22.608***	22.951***	1.053***

Note: Q is the portmanteau test, Q^* is the Ljung-Box extension and N corresponds to the Nyblom test. The lag length for the Portmanteau and Ljung-Box tests are selected following Escanciano and Lobato (2009). 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. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$.

volatility.[5] Table 1 reports the statistics for the Portmanteau (Q), Ljung-Box (Q^*) and Nyblom (1989) (N) tests. The null hypothesis of a constant shape parameter is strongly rejected, against the alternative of time variation. Moreover, the rejection of the Nyblom test, which under the alternative hypothesis assumes the parameter to follow a martingale process, suggests that the shape parameter is likely to be highly persistent.

Starting from this novel evidence, in the next Section we introduce a modeling framework that accommodates, but do not impose, the presence of time-varying asymmetry in the conditional distribution of GDP growth. We will further show that, despite conditional skewness largely varying over the sample, this does not prevent the model from generating unconditional symmetry in the sample distribution of economic growth, consistently with what we find in the data.

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 a Skew-t (Arellano-Valle et al., 2005; Gómez et al., 2007), with time-varying location μ_t, scale σ_t, and shape ϱ_t parameters, and constant degrees of freedom ν:

$$y_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim skt_\nu(0, \sigma_t^2, \varrho_t), \tag{1}$$

with $\nu > 3$, $\sigma_t > 0$, and $\varrho_t \in [-1, 1]$. The shape parameter fully characterizes the asymmetry of the distribution, with $\frac{1-\varrho_t}{1+\varrho_t}$ defining the ratio of the probability mass on the right, over the probability mass on the left of the mode, μ_t. Therefore, positive (negative) values of ϱ_t imply negatively (positively) skewed distributions. The conditional log-likelihood function of the observation at time t is:

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

[5] Appendix A provides additional details on the specifics of the tests, as well as alternative specifications for the location and scale of GDP growth under the null Hypothesis, confirming the robustness of these results.

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 $sgn(\cdot)$ is the sign function. The vector θ collects all the static parameters of the model, and $Y_{t-1} = \{y_j\}_{j=1}^{t-1}$ is the information set, up to time $t-1$. For $\varrho_t = 0$ we have the symmetric Student-t distribution, for $\eta \to 0$ we retrieve the epsilon-Skew-Gaussian distribution of Mudholkar and Hutson (2000), whereas the distribution collapses to a Gaussian density when both conditions hold.[6] Thus, we allow for, but do not impose, asymmetric innovation terms.

We model the time-variation of the parameters within the score driven framework of Creal et al. (2013) and Harvey (2013). Let f_t collect the vector of time-varying parameters of interest, the updating mechanism is given by the autoregressive process

$$f_{t+1} = Af_t + BX_t + Cs_t, \tag{3}$$

where A contains autoregressive parameters governing the persistence of the updating, B collects loadings on a potential set of financial covariates X_t, and C collects the smoothing parameters adjusting the speed of the updating, driven by an appropriate function of the prediction error. Specifically, the scaled score, s_t, is defined as $s_t = \mathcal{S}_t \nabla_t$, where:

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

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. Hence, the direction and magnitude of the updating are dictated by the steepness and curvature of the likelihood function, to improve the local fit of the model. The resulting model belongs to the class of observation-driven models, for which the trajectories of the time-varying parameters are perfectly predictable given past information and the log-likelihood function is 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 in Equation (1) and the likelihood in Equation (2), the elements of the gradient ∇_t, with respect to location, scale and asymmetry, are:*

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

where $w_t = \frac{(1+\eta)}{(1-sgn(\varepsilon_t)\varrho_t)^2 + \eta\zeta_t^2}$ and ζ_t denotes the scaled prediction error, $\zeta_t = \frac{\varepsilon_t}{\sigma_t}$. The associated information

[6] As opposed to the Skew-t distribution of Azzalini and Capitanio (2003), the Skew-t distribution of Gómez et al. (2007) retrieve an information matrix which is always non-singular, and can thus be inverted, provided that $|\varrho_t| < 1$.

matrix reads as follows:

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

Proof. See Appendix B. □

Proposition 1 highlights the central role of re-weighting the standardized prediction error (and its square values) for updating of the time-varying parameters. Weights, w_t, penalize extreme standardized innovations depending on the thickness of the tails, as well as volatility and asymmetry estimated conditional to time $t-1$. The top left panel of Figure 1 displays the weights associated to the prediction error, for alternative model parametrizations. In a Gaussian setting (black line) weights are constant and equal to unity, implying no discounting. On the other hand, when the asymmetry parameter is zero (red line), the weights display the classic outlier-discounting typical of the Student-t distributions (see, e.g. Delle Monache and Petrella, 2017). When the distribution is 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 the parameters. 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 twice differentiable *"link functions"* to these parameters. In practice, we model $\gamma_t = \log(\sigma_t)$ and $\delta_t = \text{arctanh}(\varrho_t)$, such that the vector of time-varying parameters becomes $f_t = (\mu_t, \gamma_t, \delta_t)'$. Moreover, we follow Lucas and Zhang (2016) in scaling the score only using the diagonal elements of the information matrix. 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 $J_t = \frac{\partial(\mu_t, \sigma_t^2, \varrho_t)}{\partial(\mu_t, \gamma_t, \delta_t)'}$ is the Jacobian matrix associated to the link functions.

The remainder of Figure 1 plots the scaled scores against the standardized innovations, for the location, scale and shape parameters, respectively.[7] The location updates in the direction of the prediction error. The outlier discounting implied by the t distribution gives rise to the typical S-shaped influence function (see, e.g. Harvey and Luati, 2014), which in our case adapts to the asymmetry of the conditional distribution. The shape updates in the opposite direction of the prediction error, such that for negative prediction errors (i.e. $\zeta_t < 0$) the conditional distribution becomes more left skewed. On the contrary, updates of the

[7] These are akin to the news impact curves popularized by Engle and Ng (1993) in the context of the volatility models. In our case, the "news" are measured in terms of scaled prediction errors, as compared to the raw errors.

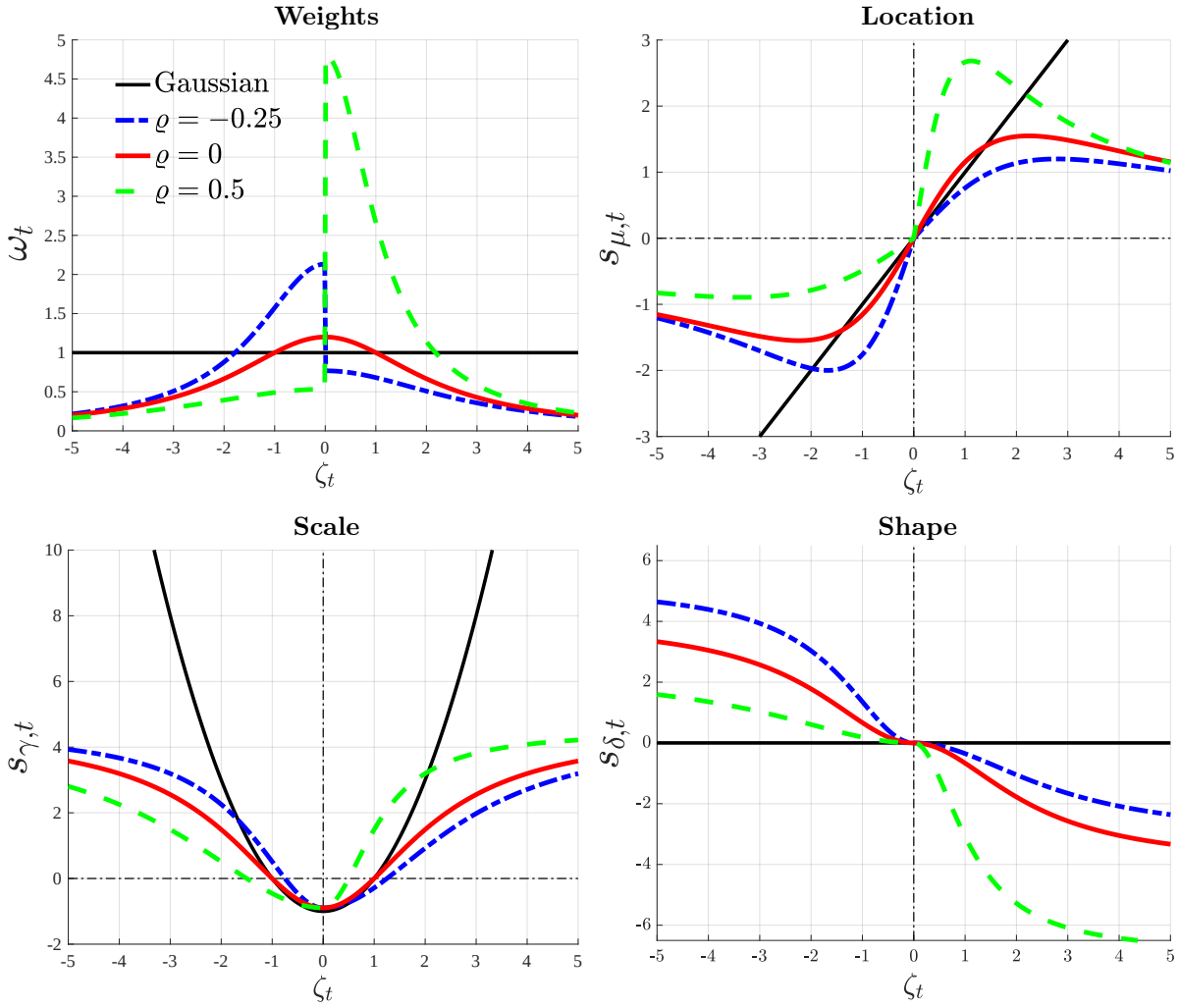

Figure 1: Prediction error and parameters' updating

Note: The figures plot the weighting scheme implied by ω_t, and the scaled scores, for different values of the scaled prediction error $\zeta_t = \varepsilon_t/\sigma_t$. We consider three values of the asymmetry parameter: -0.5 (blue), 0 (red) and 0.7 (green). The Gaussian case is reported in black. The scale parameter is set to 1, with 5 degrees of freedom.

scale do not depend on the sign of the innovations, but on their magnitude: the parameter increases for weighted squared prediction errors greater than unity, $\omega_t\zeta_t^2 > 1$, and decreases otherwise. Notice that while the (scaled) scores for the location and shape parameters are negatively correlated, updates of the scale are (unconditionally) uncorrelated with revisions of the other parameters, as suggested by the Information matrix in Equation (6). However, updates of the scale and shape parameters positively comove at the onset of recession. Business cycle turning points are typically marked by large, negative prediction errors, such that the conditional distribution of GDP growth features negative shifts in the location, increasing dispersion and deepening negative skewness.

The updating mechanism associated with the score function in Equation (7) depends on the estimated parameters at time t, and thus varies over time. For a given prediction error, the magnitude of the updates are smaller when large errors are expected, i.e. when the conditional distribution features a large scale. Most importantly, the asymmetry of the distribution plays a key role in the translation of the standardized prediction error into a signal for the updates. 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, while the opposite is true in the case of a positively skew distribution. This property of the updating function is a direct consequence of the stark asymmetry in the weights. Hence, the model is faster to update the size and sign of the asymmetry of the distribution whenever there is evidence for a change of sign, like the occurrence of a large, negative prediction error. In our application, this allows the model to promptly detect shifts in the skewness of GDP growth around business cycle turning points. In addition, the updating mechanism is robust to the presence of outliers. For extreme values of the (scaled) prediction errors, parameters' updates become inelastic to the standardized innovations (see Appendix B.3). As a consequence, our model remains well-behaved even in the face of extremely large realizations of the prediction error, as observed in the first quarters of 2020.

Existing models that account for time-varying conditional skewness and asymmetric t innovations are based on *ad hoc* updating mechanisms based on third-order powers of the prediction errors (see, e.g., Hansen, 1994; Harvey and Siddique, 1999). This type of updating presents two main drawbacks. First, the implied mapping between prediction errors and time-varying parameters does not depend on current estimates of the conditional distribution and remains constant over time. For instance, this framework does not account for the higher probability of a negative prediction error when the conditional distribution of the data is negatively skewed. Instead, it updates the parameters regardless of the local properties of the distribution. Second, the introduction of higher-order powers of the innovations makes the time-varying parameters inherently sensitive to large prediction errors, thus becoming unstable in the presence of outliers. On the contrary, the score driven framework provides outlier robust and information theoretic optimal updates. In fact, Blasques et al. (2015) show that score-driven updates reduce the local Kullback-Leibler divergence between the true and the model-implied conditional density, even under severe misspecification.

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. Several papers have documented that over the sample under analysis GDP growth has 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 and Perez-Quiros, 2000; Stock and Watson, 2002), and skewness of the distribution (Jensen et al., 2020) from the late 1980s. At the same time, Jurado et al. (2015) show that the volatility of GDP growth is countercyclical, while Giglio et al. (2016) and Adrian et al. (2019) argue that the skewness of the cycle falls sharply during recessions. To account for these features of the data, we postulate a two-component specification for the time-varying parameters, in the spirit of Engle and Lee (1999). We posit a random walk updating for the permanent components, where these are able to track both smooth variations and sudden breaks in the level of the parameters. Moreover, we allow a set of predictors, X_t, to have a transitory impact on the parameters of the distribution.

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

$$\mu_{t+1} = \bar{\mu}_{t+1} + \tilde{\mu}_{t+1}, \tag{8}$$

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

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

where the AR(2) specification for the cyclical component is able to recover the characteristic hump shaped impulse response of the data (see, e.g. Chauvet and Potter, 2013). Following Engle and Rangel (2008), we assume a multiplicative specification for the volatility, i.e. $\gamma_t = \log(\sigma_t)$ where:

$$\gamma_{t+1} = \bar{\gamma}_{t+1} + \tilde{\gamma}_{t+1}, \tag{11}$$

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

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

Similarly, we posit an additive specification for the unrestricted shape parameter, $\delta_t = \text{arctanh}(\varrho_t)$:

$$\delta_{t+1} = \bar{\delta}_{t+1} + \tilde{\delta}_{t+1}, \tag{14}$$

$$\bar{\delta}_{t+1} = \bar{\delta}_t + \varsigma_\delta s_{\delta t}, \tag{15}$$

$$\tilde{\delta}_{t+1} = \phi_\delta \tilde{\delta}_t + \beta'_\delta X_t + \kappa_\delta s_{\delta t}, \tag{16}$$

that implies $\varrho_t = \tanh(\bar{\delta}_t + \tilde{\delta}_t)$. Therefore, the resulting vector of time-varying parameters is equal to $f_t = (\bar{\mu}_t, \tilde{\mu}_t, \bar{\gamma}_t, \tilde{\gamma}_t, \bar{\delta}_t, \tilde{\delta}_t)'$, whose law of motion is described by a restricted specification of Equation (3), as we show in Appendix B.

Plagborg-Møller et al. (2020) consider a time-varying Skew-t specification for GDP growth and specify the time-varying parameters (location, log-scale and shape) as linear functions of a set of predictors. In this case - which remains nested within our setting - the sole source of parameters' variation stems from the dynamics of the predictors. This modelling choice generates substantial variability in the underlying parameters, and thus uncertainty around the estimates. In contrast, our specification allows for both secular and transitory shifts in the parameters, where the autoregressive structure of the cyclical components makes them functions of discounted values of all past predictors and scores (where these latter are themselves nonlinear functions of past data). Specifically,

$$\mu_{t+1} = \bar{\mu}_{t+1} + \sum_{j=0}^{t-1} \psi_{\mu,j} \left(\beta'_\mu X_{t-j} + \kappa_\mu s_{\mu t-j}\right) \tag{17}$$

$$\gamma_{t+1} = \bar{\gamma}_{t+1} + \sum_{j=0}^{t-1} \phi_\gamma^j \left(\beta'_\gamma X_{t-j} + \kappa_\gamma s_{\gamma t-j}\right), \tag{18}$$

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

where $\psi_{\mu,j}$ is a convolution of the autoregressive parameters $\phi_{\mu,1}$ and $\phi_{\mu,2}$, which decays to zero for $j \to \infty$, and the long-run components are proportional to the cumulative sum of past scores. As a result, the time-varying parameters we estimate are smoother and less affected by the noise of data.[8]

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). 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 high persistence values. For the location's AR parameters, we also introduce a prior on the sum of coefficients, in line with Doan et al. (1984). We assume Normal priors 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 fashion of L_2 (Ridge) regularization. The prior distribution of the score loadings is inverse gamma, as we expect these parameters to be positive. Lastly, we assume an inverse gamma prior for η. Appendix D provides an extensive description of the sampling algorithm as well as the details on the exact prior specification for the parameters.

Monte Carlo exercise Appendix E investigates the small sample properties of the model through a Monte Carlo analysis. The model successfully tracks parameters time variation for a variety of different data generating processes. When the distribution is symmetric throughout the entire sample, the model quickly estimates a null shape parameter, with limited variability over time. In particular, the model does not confound any correlation between the time variation of the location and the scale (known to generate unconditional skewness) for the presence of conditional asymmetry. We also simulated a model with a one-time break in the shape parameter: the model correctly captures the break in the asymmetry, and the two-component specification properly disentangles long- and short-lived fluctuations, so that the break is tracked by the long-run component whereas the short-run component features low variability.

[8] In addition, our framework allows for non-linearities through the non-linear mapping of the predictors into the scaled scores, and the non-linear link functions applied to the restricted parameters. This provides an automatic down-weighting of the extreme fluctuations in the data. In Appendix C we highlight that these additional features of the model turn out to be important to recover salient features of the distribution of GDP growth, such as the Great Moderation, and significant cyclical variation of the conditional volatility and skewness.

3.3 Forecasts

For any draw of the model parameters, θ, the last step of the the observation driven filter in Equation (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 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 5, due to the high persistence of the predictors we produce forecasts keeping these fixed to their last observations, akin to assuming a random walk specification for their law of motion.[9] As for 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, thus avoiding any assumption on the distribution of the score. 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 $p(y_{T+h}|\theta, X_{T+h} = X_T)$. The h-step ahead forecast reads $p(y_{T+h}) = \int p(y_{T+h}|\theta, X_{T+h} = X_T)d\theta$.

3.4 Data and alternative model specifications

We use US quarterly data over the period 1973Q1 to 2020Q4 on economic activity and financial conditions. For the latter we consider the broad NFCI and its four subindices, tracking developments in the credit, risk, leverage and non-financial leverage markets (Brave and Butters, 2012). We will consider alternative models with either the NFCI or the disaggregated components. While the risk and credit components closely track the dynamics of the NFCI, the leverage indicator, as well as the nonfinancial leverage (NFL) index, provide important additional information. In particular, the is often regarded as an "early warning" signal for economic downturns (Mian and Sufi, 2010).[10]

Table 2 investigates the goodness of fit of alternative models using the Deviance Information Criterion (DIC, Spiegelhalter et al., 2002). We consider models of increasing complexity, featuring different specifications for the time varying parameters. A simple Gaussian AR(2) model with stochastic volatility underperforms with respect to models featuring skewness and fat tails. When we introduce time variation in skewness, DIC scores substantially improve, especially around recessions, highlighting the relevance of time-varying business cycle asymmetry. Allowing for two lags of the financial predictors, as well as slow-moving, long-run components produces lower DIC values relative to simpler mean reverting specification of the time varying parameters. Specifically, including the subcomponents of the NFCI improves the fit of the model compared to using the aggregate index, in particular around recession.[11] Based on these results,

[9] As an alternative, one could feed predictions for the explanatory variables into the model. The latter approach produces results very similar to the one reported here.

[10] See Appendix F for a detailed discussion on the data.

[11] We do not explicitly account for lagged GDP growth among the predictors as the conditional score vector already summarizes the relevant information of current GDP growth for updating the parameters. A model which includes

Table 2: Deviance Information Criterion for alternative specifications

Model	DIC	DIC^{Rec}
Gaussian AR(2) + GARCH(1,1)	4.763	13.570
Skew-t AR(2) + GAS(1,1) with ϱ fixed	4.630	12.784
One-Component Skew-t Model	4.692	7.951
One-Component Skew-t Model with NFCI	4.992	8.469
One-Component Skew-t Model with 4DFI	4.929	7.921
Two-Component Skew-t Model	4.703	7.894
Two-Component Skew-t Model with NFCI	4.627	7.406
Two-Component Skew-t Model with 4DFI	4.612	7.137

Note: The table reports the Deviance Information Criterion of Spiegelhalter et al. (2002) for different model specifications. We consider two AR(2) specifications, one with Gaussian GARCH(1,1) volatility and one with Skew-t GARCH(1,1) volatility with constant asymmetry. Next, we include three different specifications of the time-varying parameters Skew-t model with time-invariant long-run components and no predictors, lags of the NFCI and lags of the four disaggregated financial indices. The last three models allow for time-varying long-run components. In the column DIC^{Rec} we evaluate the DIC measure during recession periods.

we now focus on the baseline specification including the four subindices of the NFCI.

4 Time variation in the distribution of GDP growth

Our model allows us to study the characteristics of the conditional distribution of GDP growth. Figure 2 reports the time-varying first and second moments of the distribution of GDP growth.[12] The mean moves along the business cycle and displays sharp contractions during recessions. Volatility features a marked countercyclical behaviour, with peaks occurring during recessions. Focusing on the last year of the sample, the model features a sharp increase in volatility and a quickly rebounding mean. Yet, these movements are not as sharp as the ones observed over the Great Recession, suggesting that through the lenses of the model Covid-quarters are, at least partially, characterized as tail events. On the same charts, we also report in red the low-frequency components. That is, the moments of the distribution that would prevail in the absence of cyclical variations of the parameters. The model neatly captures a fall in long-run growth, with the expected value falling from roughly 4% in the 1970s, to roughly 2.3% at the end of the sample, in line with the evidence in Antolin-Diaz et al. (2017). The Great Moderation is reflected in a visible reduction in GDP growth's volatility: starting in the mid-1980s, cyclical volatility fluctuations dampen down as the impact of the consecutive recessions of the 1970s and 1980s fades away, and the long-run volatility is revised downward by roughly 30%.

Time-varying skewness is reported in the left panel of Figure 3. This evolves in a distinctively procyclical pattern, such that substantially negative skewness characterizes recessions, whereas expansions are marked by positively skewed distributions. Interestingly, skewness tends to decrease in anticipation of recessions,

lagged GDP growth and financial indicators as predictors delivers DIC scores which are only marginally better than the baseline model's over the full sample, but are substantially worse when focusing on recessions. In Appendix H we show that this specification does not improve the out-of-sample performance of the model.

[12] See Figure 2, in Appendix H for estimated time-varying parameters.

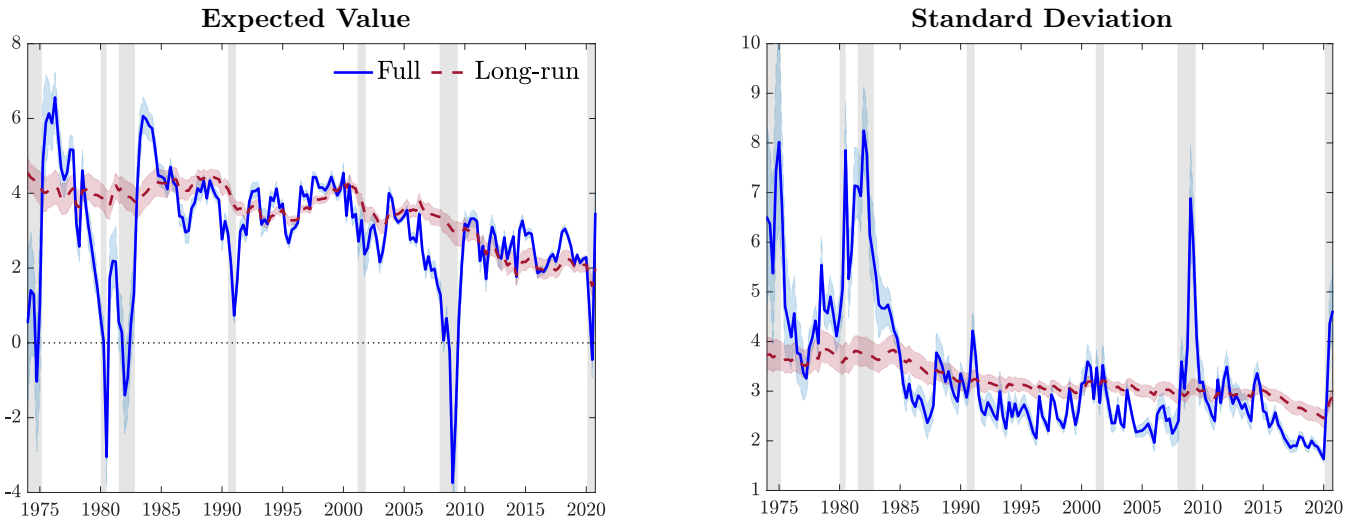

Figure 2: Time-varying mean and variance

Note: The plots illustrate the estimated time-varying mean and standard deviation (blue), along with the respective long-run components, in red, and 90% confidence bands. Shaded bands represent NBER recessions.

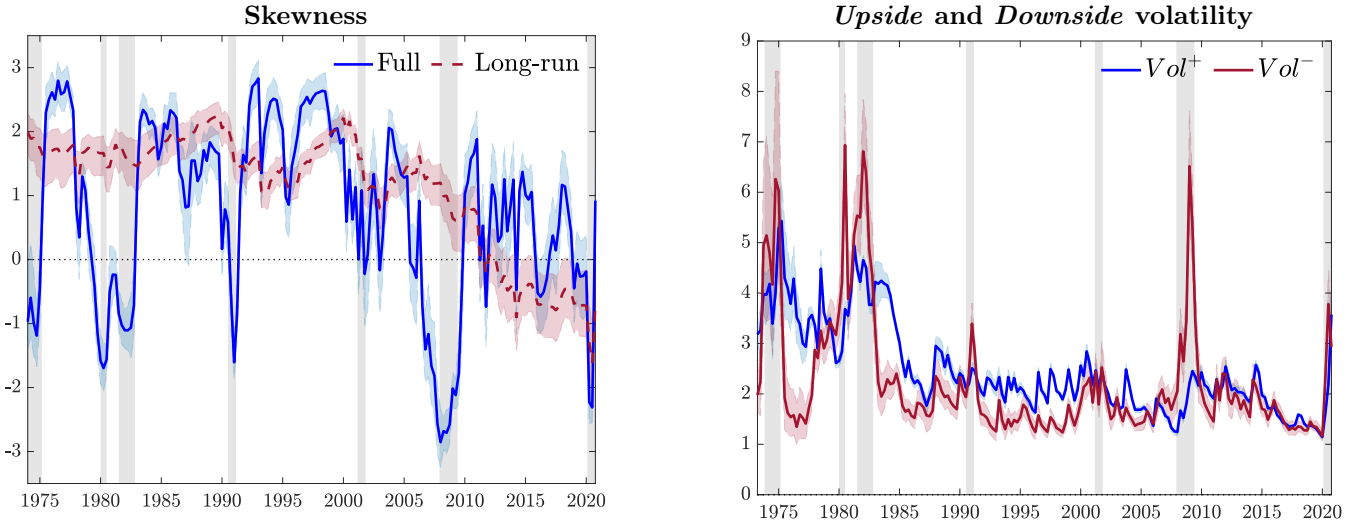

Figure 3: Time-varying asymmetry

Note: The left panel illustrates the estimated time-varying skewness (blue), along with its long-run component (red). The right panel reports the upside and downside volatilities, in blue and red, respectively. Shadings correspond to 90% confidence bands. Shaded bands represent NBER recessions.

a feature which we show to be related to the information contained in the financial indicators, suggesting that downside risk dominates ahead of, and during downturns. Over the long-run, skewness displays a downward trend starting in the late 1980s, and falling 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-run conditional distributions. This signals the build-up of vulnerabilities, resulting in the economy being increasingly exposed to downside risk episodes.

***Upside* and *downside* volatility** Arellano-Valle et al. (2005) highlight that the Skew-t distribution

can be easily rewritten as a two-piece Skew-t density by reweighting the scale by a function of the asymmetry parameter. In line with the "good" and "bad" volatility decomposition proposed by Bekaert and Engstrom (2017), one can define *"upside"* ($Vol^{+} = \sqrt{Var(y_t|y_t \geq \mu_t)}$) and *"downside"* ($Vol^{-} = \sqrt{Var(y_t|y_t < \mu_t)}$) volatility as:

$$Vol^{+} = \frac{1-\varrho_t}{2}\sqrt{Var(y_t|Y_{t-1})}, \quad Vol^{-} = \frac{1+\varrho_t}{2}\sqrt{Var(y_t|Y_{t-1})}. \tag{20}$$

These two volatility components are reported in the right panel of Figure 3. Downside volatility spikes during recessions, whereas upside volatility displays only modest (pro-)cyclicality. Therefore, the marked countercyclicality of aggregate volatility (see also Jurado et al., 2015) is largely a reflection of countercyclical downside risk. Whereas the financial crisis of 2007-2008 appears as an episode of purely downside volatility, the more recent Covid-recession is characterized by a spike in downside volatility in the first half of 2020 quickly receding in favor of upside volatility in the second half. Over the long-run, upside volatility features a steep decline starting in the mid-1980s with small cyclical variations. On the other hand, downside volatility has remained stable at around the same values observed in the aftermath of mid-1970s recession, with sudden spikes occurring during recessions.

Expected value and variance decomposition The estimated model suggests that conditional skewness fluctuations are a prominent feature of the predictive distribution of GDP growth, playing an important role in determining the dynamics of the first and second moment.[13] To see that, it is worth noting that in our model:

$$\mathbb{E}(y_t|Y_{t-1}) = \mu_t - g(\eta)\sigma_t\varrho_t, \qquad g(\eta) = \frac{4\mathcal{C}(\eta)}{1-\eta}, \tag{21}$$

$$Var(y_t|Y_{t-1}) = \sigma_t^2\left(\frac{1}{1-2\eta} + h(\eta)\varrho_t^2\right), \qquad h(\eta) = \frac{3}{1-2\eta} - g(\eta)^2, \tag{22}$$

Thus, the (time-varying) expected value and variance are equal to the location and scale of a standard Student-t distribution, plus a component which is a function of the shape parameter. The latter is magnified by larger values of σ_t, while it disappears for both moments when ϱ_t is equal to 0.

Figure 4 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 (blue line), with expansions characterized by positive skewness and contractions associated with negative skewness. 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.

[13] Although an expression for the skewness is not available in closed form, a Taylor expansion of the skewness function attributes almost all of the variation to movements of the shape parameter, while remaining relatively insensitive to location and scale.

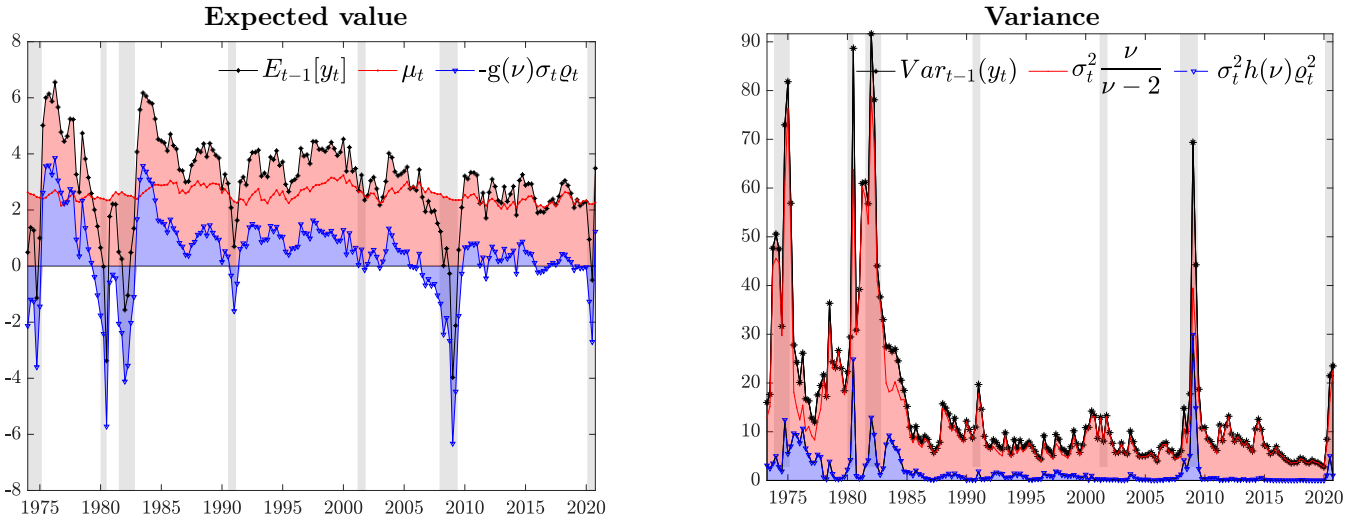

Figure 4: Expected value and variance decomposition

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 scale), whereas the blue area identifies the contribution of higher order moments. Central moments (black lines) are computed as per Equation (21)-(22). Shaded bands represent NBER recessions.

In contrast, the effect of the shape parameter on the second moment is less pervasive, despite deepening skewness during recessions accounts for a non-trivial share of the increase in variance.[14]

Equations (21) and (22) also highlight that procyclical variation in the skewness is reflected into a time-varying correlation between the mean and the volatility of GDP growth. While the mean is always negatively affected by shifts of the shape parameter (i.e. $\frac{\partial \mathbb{E}(y_t|Y_{t-1})}{\partial \varrho_t} < 0, \forall \varrho_t$), an increase of the latter is associated with an increase in the variance when the distribution is negatively skewed, and a decrease in the variance when the distribution is positively skewed, in that $\frac{\partial Var(y_t|Y_{t-1})}{\partial \varrho_t} = 2h(\eta)\sigma_t^2\varrho_t$. Since $h(\eta) > 0$ for $\nu > 3$, changes in the asymmetry are associated with changes in the variance of the same sign as the level of the shape parameter (and thus of opposite sign to the level of conditional skewness). The procyclicality of skewness reduces volatility during expansions, whereas dispersion increases as skewness plummets to negative values. This is reflected into a procyclicality in the correlation between mean and variance, which is in line with findings in Carriero et al. (2020a). These nonlinearities in the interaction between uncertainty and aggregate economic activity are consistent with findings in Segal et al. (2015), which highlight how "positive uncertainty" (i.e. volatility in the procyclical phases of the business cycle) tends to be associated with positive expected growth, whereas this correlation turns negative during contractionary phases of the cycle.[15]

Accounting for the slowdown in long-run growth and the Great Moderation. Making use of the expected value decomposition in Equation (21), we can assess to what extent shifts in long-run

[14] This is in line with the findings of Salgado et al. (2019) that a substantial fraction of the increase of the cross-sectional dispersion for firms' growth is associated with increasing negative dispersion.

[15] Fernández-Villaverde and Guerrón-Quintana (2020b) discuss the rationale for the presence of an expansionary impact of increases in in uncertainty.

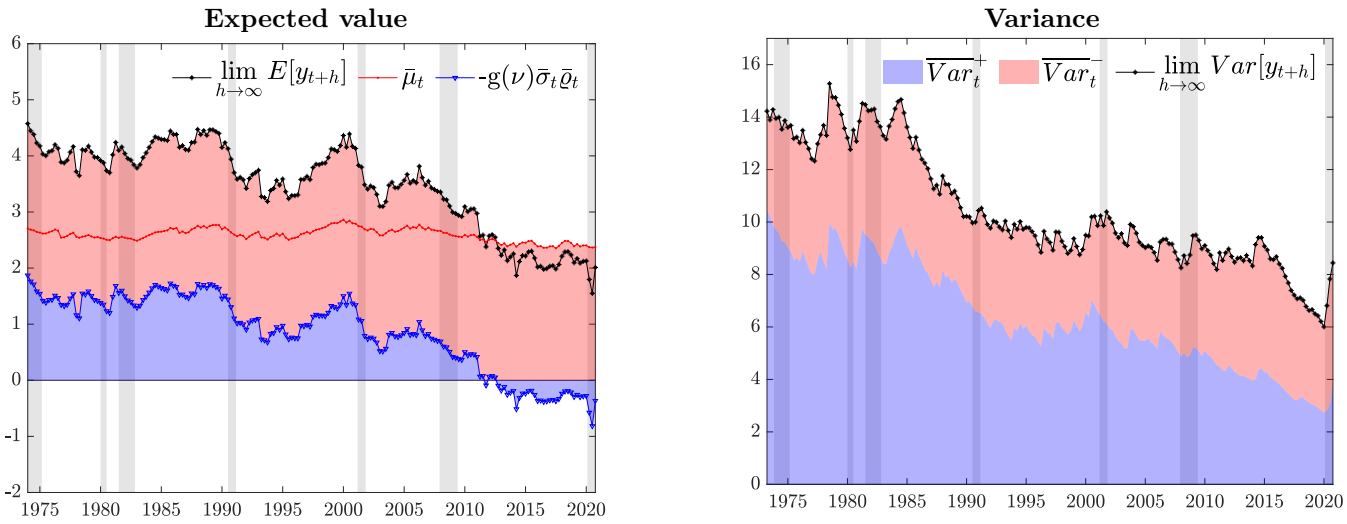

Figure 5: Long-run GDP growth and volatility

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. Shaded bands represent NBER recessions.

growth that we observe over the sample reflect a reassessment of (long-run) risk in GDP growth. Starting in the late 1980s, the distribution of GDP growth featured decreasing positive skewness. This increase in downside risk maps into a decline of the long-run growth, as shown in Figure 5, so that roughly two thirds of the slowdown reflect a reassessment of risk. The downward trend in long-run growth is temporarily reversed in correspondence of the IT productivity boom of the mid-1990s, when long-run growth is revised upward by roughly 0.5%. Interestingly, we find that this upward revision reflects, to a large extent, a shift in the risk of the upside, rather than the central tendency, of GDP growth's distributions. Over the post-2000 sample, the slowdown in long-run growth accelerates, and it is associated with a rebalancing of risks towards the downside. In the aftermath of the Great Recession, the long-run component displays a pronounced left tail, thus becoming a negative drag to long-term growth.

Similarly, we decompose long-run variance into the contributions of long-run upside (blue) and downside (red) variance in the right panel of Figure 5. Upside variance decreases quite markedly over the Great Moderation period, and by the end of the sample upside variance is half the level estimated for the 70s. On the contrary, the downside component of the long-run variance has remained quite stable throughout the sample. This evidence supports the findings in Jensen et al. (2020) and highlights that the Great Moderation reflects a reduction of the upside risk not matched by an equal fall in downside risk.

Conditional vs. Unconditional skewness Figure 3 highlights that the skewness of the conditional distribution displays a marked procyclicality: expansions are characterized by a right-skewed distributions, whereas contractions are associated with deepening asymmetry, leading to negatively-skewed conditional distributions during recessions. What does this mean for the unconditional distribution of GDP growth? To answer this question, we use the estimated model to draw inference of the unconditional (a)symmetry

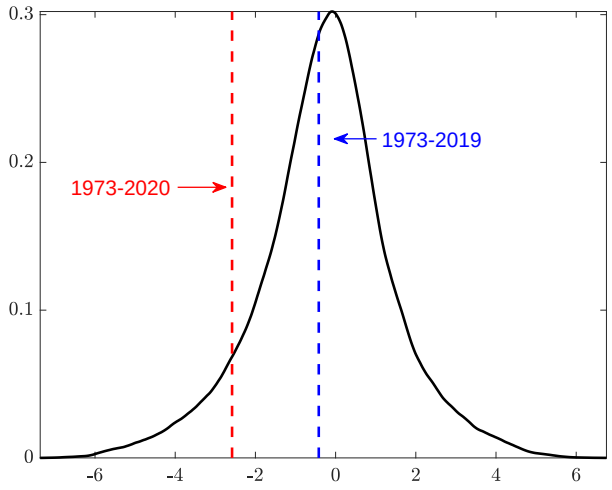

Figure 6: *Unconditional* skewness

Note: The figure reports the distribution of unconditional skewness, along with its sample value, in red, and the sample value excluding the last year of data, in blue

of the data. Specifically, we simulate 10000 alternative path of GDP growth from the model, and for each of them we compute the associated (unconditional) skewness. The results of this exercise are summarized in Figure 6. Negative skewness estimates of the unconditional distribution turn out to be 20% more likely than positive estimates, despite conditional distributions displaying positive skewness for a large part of the sample. As shown in the left panel of Figure 3, during expansions upside volatility is, on average, 15% higher than downside volatility, while downside volatility is almost double the upside volatility during recessions. Thus, despite expansions being typically characterized by right-skewed conditional distributions, the occurrence of tail events is impaired by lower dispersion. On the contrary, recessions are characterized by more downside uncertainty, resulting in large negative observations being most likely. Excluding the last year of observations, the sample skewness value of -0.42 lies close to the expected value of the empirical distribution. However, the sample skewness of -2.58 due to the Pandemic-recession still lies well within the 95% interval.[16] Performing the Bai and Ng (2005) test for unconditional skewness on simulated data fails to find evidence of any degree of asymmetry, not rejecting the null of symmetry about 90% of the times.

The contribution of financial predictors To gauge the contribution of financial indicators to the overall variation of the parameters, we exploit the moving average representation of the cyclical components in Equations (18) and (19). In Figure 7, we decompose $\tilde{\gamma}_t$ and $\tilde{\delta}_t$ into a "score-driven" component, $\kappa_\gamma \sum_{j=0}^{t-1} \phi_\gamma^j s_{\gamma t-j}$ and $\kappa_\delta \sum_{j=0}^{t-1} \phi_\delta^j s_{\delta t-j}$, respectively, and a component reflecting the share of variation driven by the predictors $\beta_\gamma' \sum_{j=0}^{t-1} \phi_\gamma^j X_{t-j}$ and $\beta_\delta' \sum_{j=0}^{t-1} \phi_\delta^j X_{t-j}$. These latter are further decomposed so as to highlight the contribution of each financial index, separately. Leverage is the largest contributor to the variation of the scale parameter. Increases in leverage during expansions are associated with a reduction of the overall dispersion of the conditional distribution of GDP growth, whereas deleveraging during recessions maps

[16] In fact, repeating the same exercise leaving out the post 2019 sample would leave the distribution of the unconditional skewness unchanged.

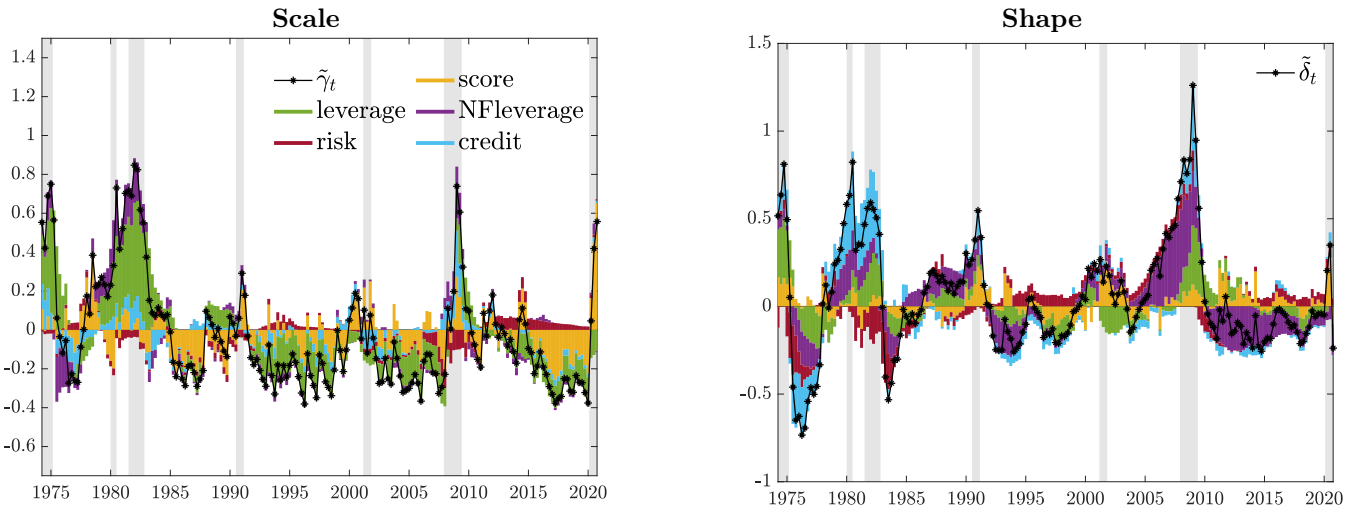

Figure 7: Predictive financial conditions

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

into higher uncertainty. Leverage is also an important determinant of the dynamics of the shape parameter, and thus of the skewness of the conditional distribution. However, in this case it is nonfinancial leverage that drives most of asymmetry's variation. Consistently with the leverage-cycle narrative of Mian and Sufi (2010) and Jordà et al. (2013), the build-up of the household leverage is identified as the main contributor to the increase in downside risk in the fist half of the 2000s, and the deleveraging associated with a substantial fall in downside risk. Indicators of credit spread and credit risk instead mainly seem to be a good indicator of the sharp increase in downside risk at the height of major recessions. Lower credit spreads before the crash impaired a sharp increase of the asymmetry parameter, while as the crisis exploded their higher values pushed the parameter towards further positive values, as argued by Krishnamurthy and Muir (2017), contributing to determine the amount of asymmetry recessions are characterized by. 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 and Watson (2003).

5 Out-of-sample evaluation

We further explore the importance of taking into account downside risk time variation, and its relation to financial variables evaluating the (out-of-sample) forecasting performance of the model over the period 1980Q1-2020Q4, at the one-quarter and one-year horizon. Forecasts are obtained from real-time GDP vintages and evaluated using the latest available release. For the one-year ahead forecasts, results are reported in cumulated output growth over the next four quarters. We compare the performance of the models for the entire out-of-sample period, as well as for the post-2000s, and for the recessive periods in the forecasting sample. We do not include the pandemic-quarters within the recession sample to avoid

forecast statistics being dominated by a single event.[17]

We assess the point forecast accuracy via the mean square forecast error (MSFE). Density forecasts accuracy is evaluated via the predictive log-score and quantile scores. The latter, put forward by Gneiting and Ranjan (2011), reads as $wQS_{t+h} = \int_0^1 QS(\alpha)\omega(\alpha)d\alpha$, where α represents the quantiles, $QS(\alpha) = 2(I(y_{t+h} < F^{-1}(\alpha)) - \alpha)(F^{-1}(\alpha) - y_{t+h})$, with $F^{-1}(\alpha)$ being the empirical quantile function of the density forecast, and $\omega(\alpha)$ is a weighting function. We consider two versions of this measure: i) the Continuously Ranked Probability Score (CRPS, Gneiting and Raftery, 2007), which assigns equal weight to each quantile of the empirical distribution function (i.e. $\omega(\alpha) = 1$) and ii) 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 of the distribution. The latter assesses the ability of the model to correctly characterize downside risk predictions. Similarly, when we evaluate the 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. Last, we investigate the use of the model in producing measures of downside risk and predicting recessions.

5.1 Point and density and downside risk 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 skewness of the distribution of GDP growth, as well as assessing whether conditioning on financial predictors leads to forecast improvements. We consider a Gaussian autoregressive model with GARCH innovations as a benchmark, which has been shown to deliver competitive out-of-sample forecasts for GDP growth-at-risk (Brownlees and Souza, 2020). Table 3 reports the predictive performance of alternative Skew-t models against the Gaussian benchmark, for one-quarter and one-year ahead predictions. In particular, we produce forecasts from (a) a Skew-t model without any financial predictors, (b) a Skew-t model that includes as predictor only the NFCI and (c) our baseline Skew-t model, including the four disaggregate financial indices (*4DFI*). For all the measures, we report ratios with respect to the Gaussian benchmark, except for the log-score (logS) for which we report differences. Values in parentheses report the p-values of the Diebold and Mariano (1995) test for equal predictive accuracy with the small sample correction proposed by Harvey et al. (1997). Simply introducing fat tails and asymmetry improves the forecast accuracy of the model with respect to the benchmark Gaussian specification. However, further predictive accuracy is gained when conditioning on financial information. Using the four subcomponents leads to additional gains during recessions. This is true for both the one-quarter and one-year ahead forecasts, for which the improvements are in generally larger, irrespective of the loss function. The gains over the benchmark Gaussian model are quite substantial, for instance, the baseline *Skt-4DFI* model produces roughly 25% (35%) improvement in MSFE, and 10% (15%) and 12%

[17] In Section 5.3 we show that the results of the out-of-sample evaluations are robust to the exclusion of the data from 2020, for both the full, and the post-2000s sample.

Table 3: Forecasting performance

	Skt	*Skt NFCI*	*Skt 4DFI*	*Skt*	*Skt NFCI*	*Skt 4DFI*
	One-quarter ahead					
	MSFE			**logS**		
Full	**0.842** (0.000)	**0.817** (0.000)	**0.812** (0.000)	**0.122** (0.000)	**0.140** (0.000)	**0.060** (0.084)
Post '00	**0.809** (0.000)	**0.804** (0.000)	**0.793** (0.000)	**0.181** (0.000)	**0.211** (0.000)	**0.167** (0.001)
Rec.	0.955 (0.315)	**0.822** (0.067)	**0.813** (0.067)	**0.349** (0.006)	**0.380** (0.007)	**0.270** (0.081)
	CRPS			**wQS**		
Full	**0.964** (0.047)	**0.941** (0.005)	**0.952** (0.025)	**0.960** (0.064)	**0.926** (0.006)	**0.926** (0.009)
Post '00	**0.934** (0.000)	**0.912** (0.000)	**0.918** (0.000)	**0.919** (0.000)	**0.894** (0.000)	**0.891** (0.002)
Rec.	0.962 (0.265)	0.934 (0.183)	0.928 (0.156)	0.948 (0.189)	0.914 (0.104)	**0.858** (0.025)
	One-year ahead					
	MSFE			**logS**		
Full	**0.720** (0.000)	**0.716** (0.002)	**0.694** (0.003)	**0.486** (0.000)	**0.585** (0.000)	**0.518** (0.001)
Post '00	**0.723** (0.000)	**0.699** (0.000)	**0.731** (0.004)	**0.814** (0.000)	**0.934** (0.000)	**0.895** (0.000)
Rec.	**0.574** (0.000)	**0.620** (0.030)	**0.545** (0.005)	**1.464** (0.000)	**1.572** (0.001)	**1.777** (0.001)
	CRPS			**wQS**		
Full	**0.912** (0.003)	**0.902** (0.002)	**0.883** (0.003)	**0.778** (0.001)	**0.747** (0.001)	**0.766** (0.005)
Post '00	**0.846** (0.000)	**0.831** (0.000)	**0.831** (0.000)	**0.731** (0.000)	**0.709** (0.000)	**0.726** (0.001)
Rec.	**0.855** (0.003)	**0.860** (0.006)	**0.782** (0.004)	**0.620** (0.000)	**0.651** (0.002)	**0.583** (0.002)

Note: The table reports the average forecast metrics relative to the Gaussian model. We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the column-specific model performs better than the Gaussian benchmark. The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

(23%) improvements in the CRPS and wQS, respectively, for the one-quarter (one-year) ahead forecasts. Gains become even larger if one focuses on the post-2000s sample, or only considers recessions.

Comparison with Adrian et al. (2019) In Table 4 we report the comparison of the baseline specification against the model of Adrian et al. (2019).[18] Our baseline model specification is associated with better point and density forecasts, and with significant values arising especially in the post-2000s sample. In particular, it is worth noticing that the forecast gains that we document during recessions stem from the adaptiveness of the score filter. This, as explained in Section 3, is due to the shape parameter reacting promptly to turning points, thus implying timely and marked movements of the skewness of the predictive distributions. As a consequence, forecast densities are characterized by longer left 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 forecasts looking at the properties of the PITs. In particular, we report the test statistic of Rossi and Sekhposyan (2019) for the Gaussian AR(2) model, the model of Adrian et al. (2019) (*ABG*), and our Skew-t model with the four disaggregated

[18]For comparability, we follow exactly the procedure of Adrian et al. (2019), but re-estimating the model using real-time vintages of GDP growth.

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

	One-quarter ahead				*One-year ahead*			
	MSFE	**logS**	**CRPS**	**wQS**	**MSFE**	**logS**	**CRPS**	**wQS**
Full	**0.890** (0.000)	**2.473** (0.000)	0.983 (0.221)	1.006 (0.599)	1.014 (0.561)	**0.571** (0.000)	0.989 (0.426)	1.026 (0.670)
Post '00	**0.837** (0.000)	**4.499** (0.000)	**0.920** (0.000)	**0.941** (0.006)	0.906 (0.133)	**0.394** (0.002)	**0.914** (0.073)	0.954 (0.269)
Rec.	1.110 (0.828)	**0.841** (0.000)	1.030 (0.689)	1.005 (0.534)	1.048 (0.581)	**1.387** (0.017)	0.900 (0.239)	0.943 (0.358)

Note: The table reports the average forecast metrics from the *Skt*-4DFI model relative to Adrian et al. (2019). We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the *Skt*4DFI model performs better than Adrian et al. (2019). The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. 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.	**2.102**	**1.925**	0.883	**4.865**	**2.306**	1.162
Left tail	**1.074**	**1.166**	0.501	**4.757**	**2.306**	1.162

Note: The table reports the test statistics for the Rossi and Sekhposyan (2019) tests, based on the Kolmogorov-Smirnov type tests. The left tail score is computed over the support [0, 0.25]. 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 lowest value of the statistic.

financial indices. The test rejects the null hypothesis of correct specification of the density forecast at the 10% critical value for both the Gaussian benchmark and for the model of Adrian et al. (2019), for the one-quarter and one-year ahead forecasts. In contrast, the test fails to reject the null of well calibrated forecasts for the entire density, as well as for the 'left tail' of the predictive distributions, capturing movements in downside risk, for our baseline model.

5.2 Tail risk predictions

We now turn to the assessment of tail 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. $ES^{\alpha}_{t+h} = \alpha^{-1} \int_0^{\alpha} VaR^{a}_{t+h|t} da$ 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 ($EL^{1-\alpha}_{t+h} = \alpha^{-1} \int_{1-\alpha}^{1} VaR^{a}_{t+h|t} da$) is the upper counterpart of the ES. In order to highlight the different risk assessments associated to the models, the left hand panel of Figure 8 contrasts the 5% expected shortfall and the 95% expected longrise for the Gaussian model, the Skew-t model without financial predictors and our baseline model, considering 10 years around the financial crisis.

The Gaussian model fails to capture the building-up of risk ahead of the Great Recession, predicting

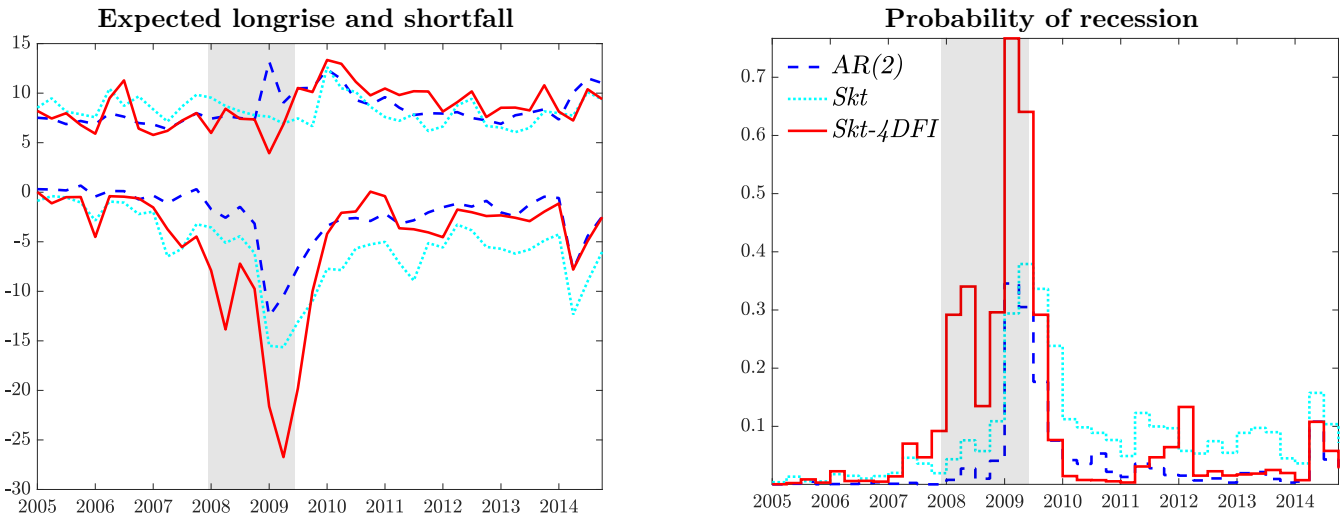

Figure 8: Expected Shortfall and Expected Longrise

Note: We report the ES and EL for $\alpha = 0.05$. Shaded bands represent NBER recessions.

an expected shortfall around zero as the economy enters the recession. In addition, assuming a symmetric distribution implies that a fall in the ES is often associated with peaks in the EL, in fact the minimum ES corresponds to the maximum EL in 2009Q2. Allowing for Skew-t innovations alleviates both problems, delivering more conservative risk measures with less erratic longrise figures, and anticipating 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, when it is sharply revised upward. Moreover, within the recession, the model delivers a downwarding longrise, predicting modest gains even for the most optimistic scenario. The longrise is sharply revised upward already for the first post-trough quarter. These results are mainly due to the fast updating of the asymmetry parameter, especially when the economy hits a turning point. This induces a reduction in the mean and an increase in perspective downside risk.[19]

Brownlees and Souza (2020) argue that a GARCH model provides competitive out-of-sample forecasts for the lower quantiles of the GDP growth distribution. Evaluating ES accuracy using the score metric proposed by Taylor (2019) highlights that allowing for time varying asymmetry in the conditional distribution of GDP growth produces large and significant gains with respect to the Gaussian model both in the full sample and during recessions. For instance, the baseline model produces 25% improvements in one-quarter ahead forecast during recessions, whereas more than 70% 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 5% VaR proposed by Giacomini and Komunjer (2005), in Table 6 we document significant gains in tail risk

[19] In Figure H3, in Appendix H, we report the predictive densities generated by the Gaussian model and the *Skt-4DFI* specification, in- and out-of-sample, for the quarters around the trough in Q4 of 2008. In Q1 of 2009, the *Skt*model predicts a negative mean and substantial downside risk, whereas the Gaussian model predict only a slightly negative growth, with a roughly symmetric assessment of the risk surrounding this prediction.

Table 6: Tail risk scores

	Skt no-X	*Skt 4DFI*	*Skt no-X*	*Skt 4DFI*	*Skt no-X*	*Skt 4DFI*
	FZG		**ALS**		**TLF**	
	One-quarter ahead					
Full	**0.831** (0.040)	**0.819** (0.028)	0.959 (0.111)	**0.948** (0.028)	0.978 (0.246)	**0.900** (0.003)
Post '00	**0.693** (0.002)	**0.720** (0.003)	**0.912** (0.012)	**0.926** (0.038)	**0.943** (0.066)	**0.915** (0.040)
Rec.	**0.641** (0.012)	**0.526** (0.012)	**0.846** (0.033)	**0.749** (0.016)	**0.858** (0.048)	**0.666** (0.010)
	One-year ahead					
Full	**0.241** (0.000)	**0.310** (0.001)	**0.397** (0.000)	**0.424** (0.000)	**0.692** (0.001)	**0.622** (0.004)
Post '00	**0.208** (0.002)	**0.170** (0.006)	**0.435** (0.001)	**0.395** (0.003)	**0.690** (0.003)	**0.475** (0.010)
Rec.	**0.137** (0.000)	**0.128** (0.001)	**0.256** (0.000)	**0.237** (0.001)	**0.441** (0.001)	**0.328** (0.005)

Note: The table reports the average downside tail risk scores, expressed as ratios relative to the Gaussian model. Ratios smaller than 1 indicate that the column-specific model performs better than the benchmark. The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are reported in parentheses. 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 and Komunjer (2005) tick loss function.

predictions.

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 [...]"*. Therefore, we define the one-year-ahead probability of recession as the probability of observing any two consecutive negative forecasts in the next four quarters. The right panel of Figure 8 highlights that 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, produce a reasonable probability of recession only toward the end of the recession period, and they continue to perceive a substantial treat of recession many quarters after the formal end of it. We evaluate the ability of the model to time recessions over the sample 1980-2020, using the Brier score. Deviating from the Gaussian assumption provides gains of around 10%, and an additional 20% gain can be directly ascribed to the inclusion of financial predictors.

Overall, these results underline the importance of allowing for time variation in the skewness of the conditional distribution of GDP growth for predicting downside risk, both in terms of magnitude and timing.

5.3 Extensions

In this Section, we provide a summary of additional analysis regarding the forecasting performances of the model. Detailed results are provided in Appendix G.

Lagged GDP growth as additional predictor Comparing the forecast performance of the baseline model with the same version of the model augmented with past lags of GDP growth as predictors, we find a generally weaker predictive ability in terms of point and density forecasts (see Table G1). These results confirm that the scaled score already appropriately summarizes the necessary information supplied by past realizations of the dependent variable, thus making the explicit use of lag of GDP growth redundant to update the parameters.

How important is parameters' uncertainty? To answer this question we produce forecasts from our baseline model fixing the parameters to the (recursively re-estimated) modal estimates, and we compare their performance with the baseline model's (which integrate over the estimated parameters). We find that the accuracy and calibration of density forecasts, especially over the medium horizon, are positively affected by the presence of parameters' uncertainty. In particular, explicitly accounting for parameter uncertainty leads to sizable gains in terms of log-scores and weighted quantile scores, in particular during recessions. Moreover, forecasts produced with the modal estimates of the model are found to understate downside risk, in particular for one year ahead forecasts (see Tables G2 and G3).

Excluding 2020 from the forecast evaluation sample We also evaluate the forecasting performances of the specifications taking out the last year of data and find that the results are generally in line with those presented above. This highlights that the gains in forecasting performance are associated with (a) the time variation of conditional asymmetry, and (b) the information of financial indicators (see Tables G4–G6).

Conditional vs. unconditional skewness We also compare the predictive ability of our models with respect to an AR(2) model with stochastic volatility and Skew-t innovations, where skewness is considered as fixed. Model performances of the latter model are largely in line with those of the Gaussian benchmark, reflecting the weak evidence in favour of unconditional skewness over the sample. As a result, the baseline model, allowing for time-varying conditional skewness, is associated with substantial gains, both in point and density forecasts, when compared to a specification with time varying mean and volatility, but constant conditional skewness (see Tables G7 and G8).

NFCI: last vintage vs. real time data The pseudo-out-of-sample forecasting exercise presented above relies on the latest vintages of the NFCI and its subcomponents. Real time vintages of these financial indicators are only available from 2013, and have historically only experienced few, limited revisions.

To evaluate whether the latest vintages provide an unfair advantage to the models, we compare these forecasting performances to those from a pure real time out-of-sample exercise, starting in 2013. Results point at some marginal improvements associated to the use of real time vintages, suggesting that the evaluation based on the latest releases does not advantage our models (see Table G9).

6 Dissecting the Financial Condition Index

So far, we have documented the importance of financial condition indicators to predict future economic growth. In this Section we explore the extent to which the predictive power of the model can be improved considering the whole set of data that feeds into the NFCI. We, furthermore, identify which indicators are better suited to predict downside risk.

We consider the full set of 105 (smoothed) indicators of financial activity that constitute the NFCI.[20] As indicators enter the predictors' set at different points in time, for each time t we only consider predictors available for at least four years, and we start our forecasting exercise at the beginning of the 2000s. This implies that the first forecast produced relies on about 70% of the total available indicators. As predictors' availability steadily increases over the sample, forecasts of the 2007-2009 recession exploit 85% of the full set of indicators.

6.1 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. We tackle this dimensionality problem through a *"shrink-then-sparsify"* strategy (see Hahn and Carvalho, 2015).[21] Predictor loading shrinkage is achieved by means of Horseshoe (HS) priors (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. Specifically, $\lambda^j \sim HC^+(0,1)$ 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. To reduce the estimation uncertainty associated with the near-zero shrinkage coefficients, we apply the Signal Adaptive Variable Selector (SAVS) algorithm of Ray and 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 and Bhattacharya, 2018). The sparsified coefficients are computed as

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

[20] While these predictors are not available in real-time, we assume that at 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.

[21] Further details on this approach are provided in Appendix D.2.

Table 7: *Big Data* forecast performance

	One-quarter ahead				*One-year ahead*			
	MSFE	**logS**	**CRPS**	**wQS**	**MSFE**	**logS**	**CRPS**	**wQS**
Full	**0.185** (0.000)	0.223 (0.163)	**0.655** (0.000)	**0.657** (0.000)	0.434 (0.112)	0.222 (0.562)	0.833 (0.259)	0.776 (0.250)
Pre-2020	**1.290** (0.087)	−0.108 (0.351)	1.109 (0.286)	**1.191** (0.071)	1.157 (0.637)	−0.065 (0.785)	1.051 (0.736)	1.083 (0.706)
Rec.	**0.408** (0.003)	0.357 (0.228)	**0.679** (0.035)	**0.708** (0.067)	0.649 (0.171)	0.664 (0.395)	0.591 (0.143)	0.449 (0.102)

Note: The table reports the average forecast metrics from the big data model relative to *Skt*4DFI. We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the big data model performs better than *Skt*4DFI. The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. Values in **bold** are significant at the 10% level.

where $||\cdot||$ represents the Euclidean norm of the vector X_j. By applying the sparsification step at each draw of the MCMC algorithm, this approach fully accounts for model uncertainty, akin to the idea of Bayesian model averaging (Huber et al., 2020).

6.2 On the importance of financial indicators

The forecasting exercise highlights large improvements associated with the sparse modeling of the predictors' design. Table 7 reports the forecasting performance of the *"sparse"* model against our baseline specification (*Skt*-4DFI), over three different samples: the full sample (2000-2020), the same sample excluding the quarters of 2020, and the recession periods, excluding the pandemic recession.[22] Over the full sample, the sparse model realizes gains of up to 80% (35%) in point (density) forecast accuracy, at the one-quarter-ahead. However, excluding 2020 shows that this model does not generally outperforms the baseline. Therefore, the idiosyncratic indicators offers drastic improvements for the forecasts of the pandemic-quarters. More generally, the sparse model substantially improves the forecasting performance during recession periods. Over the short-term, the loss in point forecast is more than halved, and the model delivers significant gains of roughly 30% in both density forecast and downside risk predictions, evaluated by the CRPS and the wQS respectively. Despite the sparse model might still be prone to overfit the data during expansions, closely monitoring signals of distress in financial markets can drastically affect the assessment of macroeconomic downside risk ahead and during times of crisis, in line with the insights of Alessi et al. (2014).

We illustrate the evolution over time of sparsity in the financial information set in Figure 9. Sparsity tends to slightly decrease over time as financial indicators appear to be greatly informative for the asymmetry parameter, while most of the information does not prove useful for predicting the scale parameter. On average, about 7.5% of the predictors feed into the prediction of the asymmetry parameter, while only about 2% contribute to the scale; towards the end of the sample, more than 10% of the indicators inform

[22]Therefore, the latter sample includes the 2001 recession, and the Great Financial crisis and the subsequent recession.

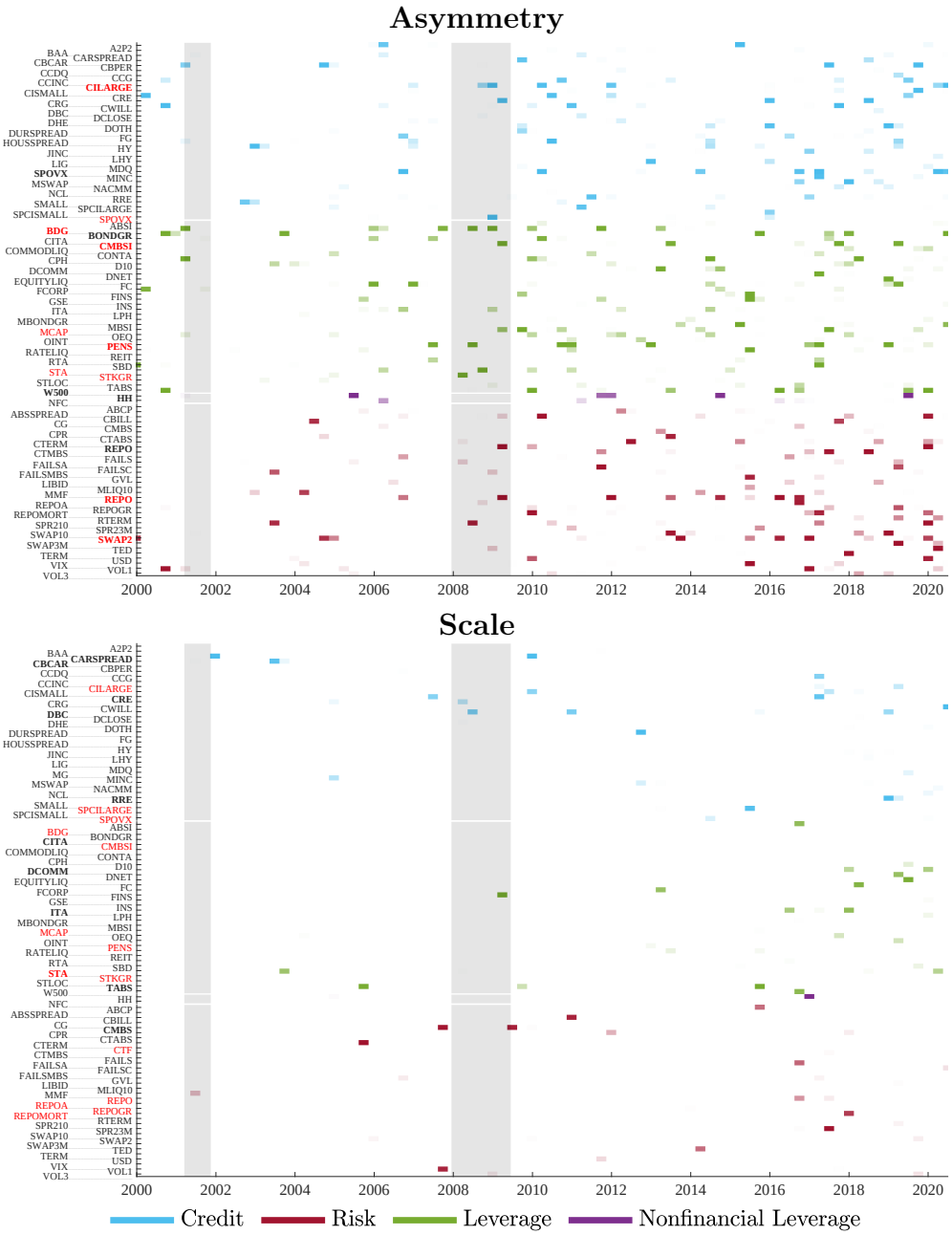

Figure 9: Sparsity

Note: The top panel reports the evolution of the financial predictors selected for the asymmetry parameter; the bottom panel reports those for the scale parameter. Names in red indicate the 10 predictors with the highers posterior probability (*pip*) of inclusion; names in bold indicates predictors with the highest *pip* around the Global Financial Crisis.

the asymmetry parameter, while 5% relate to the scale. Moreover, the number of predictors of the shape parameters increases during the financial crisis, highlighting the importance of monitoring developments in financial markets to gauge the severity of the Great Recession.

By ranking predictors by their average posterior probability of inclusion, it emerges that indicators pertaining to leverage markets provide most of the information relevant to predict the evolution of the asymmetry parameter, along with credit conditions indicator and household debt.[23] Credit spreads, instead, appear most informative for the scale parameter. If we focus on the latest Great Recession, the

[23]Best predictors and full ranks are reported in Appendix H.

size of the shadow-banking sector (as highlighted by Adrian and Shin, 2008), as well as the issuance of commercial mortgage-backed securities, prove useful to gauge the increasing downside risk associated with the crisis outburst.

7 Conclusions

The severity of the latest financial crisis and the ensuing recession has spurred the interest of both academics and practitioners in developing models that allow us to better understand and predict downside risk to economic growth. In this paper, we introduce a class of parametric models to fully characterize the conditional distribution of GDP growth. Procyclical asymmetry is a strong feature of the data, with business cycle fluctuations characterized by positive skewness in expansions and negative skewness in recessions. As a result, the correlation between the mean and the variance of the conditional distribution varies over business cycle phases, turning negative during recessions. These features become even more prominent after we allow measures of financial conditions to drive time variation in the parameters. When financial markets are overheating, future economic growth becomes more uncertain, and downside risk arises reflecting the negative skewness of the predictive distributions of GDP growth. Leverage and credit are important signals when assessing downside risk in GDP growth. In particular, non financial leverage measures anticipate a fall in economic growth, driving the rising negative asymmetry of the 2007-2009 recession. Overall, our results point to the paramount importance of nonlinearities and non-Gaussianity in capturing the dynamics of downside risk to economic growth, and for improving prediction accuracy. Conditioning on large financial information sets further improves forecasts, accurately characterizing the left tail of the predictive distribution, both at short- and medium-term projections.

Our model highlights how the statistical properties of GDP growth have dramatically changed over the last 50 years. The fall in volatility observed since the mid-1980s is reflected into a substantial fall in upside volatility, with downside volatility remaining relatively stable over the entire sample. The slowdown in long-run growth observed over the last two decades largely reflects a reassessment of risk, pointing to an increase in downside risk in the conditional distribution. We document that, since the Great Recession, the long-run predictions of the distribution of GDP growth displays a marked negative skewness. This suggests that in the last 10 years, the slow economic recovery took place in an environment characterized by increasing downside risk to economic activity, well before the COVID-19 crisis.

References

Adrian, T., N. Boyarchenko, and D. Giannone (2019): "Vulnerable Growth," *American Economic Review*, 109, 1263–89.

Adrian, T., F. Grinberg, N. Liang, and S. Malik (2018): "The Term Structure of Growth-at-Risk," Discussion Papers 13349, Centre for Economic Policy Research.

Adrian, T. and H. S. Shin (2008): "Financial intermediaries, financial stability and monetary policy," *Proceedings - Economic Policy Symposium - Jackson Hole*, 287–334.

Alessi, L., E. Ghysels, L. Onorante, R. Peach, and S. Potter (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, 483–500.

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2017): "Tracking the slowdown in long-run GDP growth," *Review of Economics and Statistics*, 99, 343–356.

Arellano-Valle, R. B., H. W. Gómez, and F. A. Quintana (2005): "Statistical inference for a general class of asymmetric distributions," *Journal of Statistical Planning and Inference*, 128, 427–443.

Azzalini, A. and A. Capitanio (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, 367–389.

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

Bekaert, G. and E. Engstrom (2017): "Asset Return Dynamics under Habits and Bad Environment-Good Environment Fundamentals," *Journal of Political Economy*, 125, 713–760.

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

Brave, S. and R. A. Butters (2012): "Diagnosing the Financial System: Financial Conditions and Financial Stress," *International Journal of Central Banking*, 8, 191–239.

Brownlees, C. and A. B. Souza (2020): "Backtesting Global Growth-at-Risk," *Journal of Monetary Economics*, Forthcoming.

Carriero, A., T. E. Clark, and M. Marcellino (2020a): "Capturing Macroeconomic Tail Risks with Bayesian Vector Autoregressions," Working Paper 20-02, Federal Reserve Bank of Cleveland.

——— (2020b): "Nowcasting Tail Risks to Economic Activity with Many Indicators," Working Papers WP 20-13, Federal Reserve Bank of Cleveland.

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

Cecchetti, S. G. (2008): "Measuring the Macroeconomic Risks Posed by Asset Price Booms," in *Asset Prices and Monetary Policy*, University of Chicago Press, NBER Chapters, 9–43.

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

Chauvet, M. and S. Potter (2013): "Forecasting output," in *Handbook of Economic Forecasting*, Elsevier, vol. 2, 141–194.

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

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

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

Diebold, F. X., T. A. Gunther, and A. Tay (1998): "Evaluating density forecasts, with Applications to Financial Risk Management," *International Economic Review*, 39, 863–883.

Diebold, F. X. and R. S. Mariano (1995): "Comparing predictive accuracy," *Journal of Business & Economic Statistics*, 20, 134–144.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and conditional projection using realistic prior distributions," *Econometric Reviews*, 3, 1–100.

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

Edge, R. M., T. Laubach, and J. C. Williams (2007): "Learning and shifts in long-run productivity growth," *Journal of Monetary Economics*, 54, 2421–2438.

Engle, R. F. and G. Lee (1999): "A permanent and transitory component model of stock return volatility," in *Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*, ed. by R. F. Engle and H. White, Oxford: Oxford University Press, chap. 20, 475–497.

Engle, R. F. and V. K. Ng (1993): "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, 1749–1778.

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

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

Fernández-Villaverde, J. and P. Guerrón-Quintana (2020a): "Estimating DSGE Models: Recent Advances and Future Challenges," NBER Working Papers 27715, National Bureau of Economic Research.

——— (2020b): "Uncertainty Shocks and Business Cycle Research," *Review of Economic Dynamics*, 37, 118–166.

Fernández-Villaverde, J., S. Hurtado, and G. Nuño (2019): "Financial Frictions and the Wealth Distribution," NBER Working Papers 26302, National Bureau of Economic Research.

Fissler, T., J. F. Ziegel, and T. Gneiting (2016): "Expected shortfall is jointly elicitable with value-at-risk: implications for backtesting." *Risk.net (www.risk.net/2439862)*.

Gadea Rivas, M. D., L. Laeven, and G. Pérez-Quirós (2020): "Growth-and-Risk Trade-off," CEPR Discussion Papers 14492, Center for Economic Policy and Research.

Galvão, A. B. and M. T. Owyang (2018): "Financial stress regimes and the macroeconomy," *Journal of Money, Credit and Banking*, 50, 1479–1505.

Ganics, G., B. Rossi, and T. Sekhposyan (2020): "From Fixed-event to Fixed-horizon Density Forecasts: Obtaining Measures of Multi-horizon Uncertainty from Survey Density Forecasts," CEPR Discussion Papers 14267, Centre for Economic Policy Research.

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

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

Giannone, D., M. Lenza, and G. E. Primiceri (2018): "Economic predictions with big data: The illusion of sparsity," Staff Report 847, Federal Reserve Bank of New York.

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

Gneiting, T. and A. E. Raftery (2007): "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, 102, 359–378.

Gneiting, T. and R. Ranjan (2011): "Comparing density forecasts using threshold-and quantile-weighted scoring rules," *Journal of Business & Economic Statistics*, 29, 411–422.

Gómez, H. W., F. J. Torres, and H. Bolfarine (2007): "Large-sample inference for the epsilon-skew-t distribution," *Communications in Statistics—Theory and Methods*, 36, 73–81.

Haario, H., E. Saksman, and J. Tamminen (1999): "Adaptive proposal distribution for random walk Metropolis algorithm," *Computational Statistics*, 14, 375–396.

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

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

Harvey, A. and A. Luati (2014): "Filtering With Heavy Tails," *Journal of the American Statistical Association*, 109, 1112–1122.

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, C. R. and A. Siddique (1999): "Autoregressive conditional skewness," *Journal of Financial and Quantitative Analysis*, 34, 465–487.

Harvey, D., S. Leybourne, and P. Newbold (1997): "Testing the equality of prediction mean squared errors," *International Journal of forecasting*, 13, 281–291.

Huber, F., G. Koop, and L. Onorante (2020): "Inducing Sparsity and Shrinkage in Time-Varying Parameter Models," *Journal of Business & Economic Statistics*, Forthcoming.

Jensen, H., I. Petrella, S. H. Ravn, and E. Santoro (2020): "Leverage and Deepening Business-Cycle Skewness," *American Economic Journal: Macroeconomics*, 12, 245–81.

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

Jordà, Ò., M. Schularick, and A. M. Taylor (2020): "Disasters Everywhere: The Costs of Business Cycles Reconsidered," NBER Working Papers 26962, National Bureau of Economic Research.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring uncertainty," *American Economic Review*, 105, 1177–1216.

Koopman, S. J., A. Lucas, and M. Scharth (2016): "Predicting time-varying parameters with parameter-driven and observation-driven models," *Review of Economics and Statistics*, 98, 97–110.

Koopman, S. J., A. Lucas, and M. Zamojski (2018): "Dynamic term structure models with score-driven time-varying parameters: estimation and forecasting," Working Papers 258, Narodowy Bank Polski.

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

Lucas, A. and X. Zhang (2016): "Score-driven exponentially weighted moving averages and Value-at-Risk forecasting," *International Journal of Forecasting*, 32, 293–302.

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

Mian, A. and A. Sufi (2010): "Household Leverage and the Recession of 2007–09," *IMF Economic Review*, 58, 74–117.

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

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

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020): "When is Growth at Risk?" Conference draft, Brooking Paper on Economic Activity.

Ray, P. and A. Bhattacharya (2018): "Signal Adaptive Variable Selector for the Horseshoe Prior," *arXiv preprint arXiv:1810.09004*.

Rossi, B. and T. Sekhposyan (2019): "Alternative tests for correct specification of conditional predictive densities," *Journal of Econometrics*, 208, 638–657.

Rudebusch, G. D. and J. C. Williams (2009): "Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve," *Journal of Business & Economic Statistics*, 27, 492–503.

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

Segal, G., I. Shaliastovich, and A. Yaron (2015): "Good and bad uncertainty: Macroeconomic and financial market implications," *Journal of Financial Economics*, 117, 369–397.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde (2002): "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 583–639.

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

——— (2003): "How did leading indicator forecasts perform during the 2001 recession?" *FRB Richmond Economic Quarterly*, 89, 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, 121–133.

Modeling and forecasting macroeconomic downside risk: Supplementary material

Davide Delle Monache* Andrea De Polis† Ivan Petrella‡

*Bank of Italy. davide.dellemonache@bancaditalia.it
†University of Warwick.andrea.depolis.17@mail.wbs.ac.uk
‡University of Warwick & CEPR. ivan.petrella@wbs.ac.uk

Contents

A Score-driven setting

Assume that the variable y_t is generated by the observation density $\mathcal{D}(\theta, f_t)$, with θ collecting the static parameters of the distribution. The score-driven setting postulates the dynamics of the time-varying parameters, f_t, being:

$$f_{t+1} = \varpi + \sum_{i=0}^{p-1} \alpha_i s_{t-i} + \sum_{j=0}^{q-1} \beta_j f_{t-j}, \tag{A1}$$

which we refer to as $GAS(p, q)$ dynamics.[1] The scaled score s_t is a non-linear function of past observations and past parameters' values. For $\ell_t = \log \mathcal{D}(\theta, f_t)$, we define:

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

where ∇_t corresponds to the gradient vector of the log-likelihood function, ℓ_t, and the scaling matrix $\mathcal{S}_{t-1}$ is proportional to the square-root generalized inverse of the Information matrix $\mathcal{I}_{t-1}$.[2] Within this framework, the parameters are updated in the direction of the steepest ascent, in order to maximize the local fit of the model.

A.1 Score-based tests for time varying asymmetry

In Section 2 we exploited properties of the score function to motivate the modeling of time-varying parameters. The Lagrange multiplier principle can be employed to formally test for the time variation of the parameter of interest (Harvey, 2013, Section 2.5).[3] 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 score 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).

Given the law of motion in Equation (A1), and provided $\{|\beta_j| < 1\}_{j=0}^{q-1}$, we can recover the (infinite) moving average representation of the filter:

$$f_{t+1} = \varpi^* + \sum_{i=0}^{\infty} \psi_i s_{t-i}, \tag{A2}$$

equivalent to a $GAS(\infty, 0)$, with $\varpi^* = \frac{\varpi}{(1-\beta_1-\beta_2-\ldots-\beta_{q-1})}$. Tests can then be carried out using the score's autocorrelation function, $r_s(\tau)$.

Assume the conditional distribution of y_t being the Skew-t of Gómez et al. (2007), $y_t|Y_{t-1} \sim skt_\nu(\mu_t, \gamma, \delta)$, with log-likelihood

$$\ell_t(y_t|\theta, Y_{t-1}) = \log \mathcal{C}(\eta) - \frac{1}{2}\log\sigma^2 - \frac{1+\eta}{2\eta}\log\left[1 + \frac{\eta\varepsilon_t^2}{(1 - sgn(\varepsilon_t)\varrho)^2\sigma^2}\right], \tag{A3}$$

[1] Creal et al. (2013) refer to the Generalized Autoregressive Score (GAS) model, while Harvey (2013) names it Dynamics Conditional Score (DCS) model. We stick to the former notation throughout this appendix.

[2] Refer to Creal et al. (2013) for additional details on this choice.

[3] A similar testing approach is put forward by Calvori et al. (2017).

$$\log \mathcal{C}(\eta) = \log \Gamma\left(\frac{\eta+1}{2\eta}\right) - \log \Gamma\left(\frac{1}{2\eta}\right) - \frac{1}{2}\log\left(\frac{1}{\eta}\right) - \frac{1}{2}\log \pi,$$

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 $\gamma = \log \sigma$, $\delta = \text{arctanh}(\varrho)$, 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:

$$\begin{bmatrix} s_{\mu,t} \\ s_{\gamma,t} \\ s_{\delta,t} \end{bmatrix} = \sqrt{\frac{(1+3\eta)}{(1+\eta)}} \begin{bmatrix} \sqrt{(1-\varrho^2)} w_t \zeta_t \\ \sqrt{\frac{(1+\eta)}{2}} (w_t \zeta_t^2 - 1) \\ -\sqrt{\frac{(1-\varrho^2)}{3}} \frac{sgn(\varepsilon_t)}{(1-sgn(\varepsilon_t)\varrho)} w_t \zeta_t^2 \end{bmatrix}, \tag{A4}$$

where

$$w_t = \frac{(1+\eta)}{(1-sgn(\varepsilon_t)\varrho)^2 + \eta\zeta_t^2}, \quad \zeta_t = \frac{\varepsilon_t}{\sigma}. \tag{A5}$$

The scores in Equation (A4) 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 and 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. For all these test, the optimal lag-length is selected following the methodology of Escanciano and Lobato (2009).

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

S1: $\mu_t = 0$ and $y_t \sim AR(2)$ with skew-t innovations $skt_\nu(0, \sigma, \varrho)$ (see Section 2).

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

S3: $\mu_t \sim RW + GAS(1,2)$ with stationary roots;

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, using data on GDP growth from 1973 to 2020, are reported in Table 1 (Section 2) for S1, while Table A1 reports the results for S2 and S3.[4]

For all the cases and specifications we reject the null hypothesis of constant shape at the 1% confidence level, for Q and Q^*. Similarly, the Nyblom specification rejects the null. While the first two tests leave the alternative hypothesis unspecified, the latter tests for parameter constancy, against a random walk law of motion.[5] To make sure that statistical significance is not induced by the Covid-related observations, we also report the same test over the 1973-2019 sample. Table A2 confirms that the evidence in favor of the shape parameter being time-varying is robust to the inclusion of outliers.

[4] Horswell and Looney (1993) show that, when testing for skewness, omitting the presence of kurtosis might deliver misleading results. Repeating the tests with Skew-Normal innovations delivers similar evidence.

[5] Harvey and Streibel (1998) argue that the Nyblom (1989) test can be used to test against stationary, but highly persistent, alternative hypotheses.

Table A1: Score-based tests for time variation

	time-varying location			*time-varying location & scale*		
	Q	Q^*	N	Q	Q^*	N
			S2			
$Scale^2$	33.221***	33.724***	1.189***			
Shape	24.238***	24.60***6	0.803***	11.565***	11.740***	1.468***
			S3			
$Scale^2$	48.028***	48.755***	0.982***			
Shape	38.668***	39.254***	1.009***	22.761***	23.106***	1.529***

Note: Q is the portmanteau test, Q^* is the Ljung-Box extension and N corresponds to the Nyblom test. The lag length for the Portmanteau and Ljung-Box tests are selected following Escanciano and Lobato (2009). 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. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$.

Table A2: Score-based tests for time variation, 1973-2019 sample

	time-varying location			*time-varying location & scale*		
	Q	Q^*	N	Q	Q^*	N
			S1			
$Scale^2$	3.432***	3.485***	1.489			
Shape	5.199**	5.280**	0.559**	31.705***	32.196***	0.800***
			S2			
$Scale^2$	37.416***	37.995***	1.796***			
Shape	31.987***	32.482***	0.827***	33.737***	34.258***	1.642***
			S3			
$Scale^2$	48.937***	49.694***	1.536***			
Shape	52.172***	52.979***	0.727**	32.743***	33.249***	1.458***

Note: Test scores relate to the 1973-2019 sample. Q is the portmanteau test, Q^* is the Ljung-Box extension and N corresponds to the Nyblom test. The lag length for the Portmanteau and Ljung-Box tests are selected following Escanciano and Lobato (2009). 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. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$.

B Derivation of the score-driven Skew-t model

In this appendix we provide details on the scaled scores for the Skew-t distribution of Gómez et al. (2007), the use of the link functions and the matrix representation of the model.

B.1 Scaled scores

Differentiating the log-likelihood function in Equation (B11) 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:

$$\frac{\partial \ell_t}{\partial \mu_t} = \frac{1}{\sigma_t^2}\left[\frac{(1+\eta)\sigma_t^2\varepsilon_t}{(1-sgn(\varepsilon_t)\varrho_t)^2\sigma_t^2+\eta\varepsilon_t^2}\right] = w_t\frac{\zeta_t}{\sigma_t}, \tag{B1}$$

$$\frac{\partial \ell_t}{\partial \sigma_t^2} = \frac{1}{2\sigma_t^2}\left[\frac{(1+\eta)\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho_t))^2\sigma_t^2+\eta\varepsilon_t^2}-1\right] = \frac{(w_t\zeta_t^2-1)}{2\sigma_t^2}, \tag{B2}$$

$$\frac{\partial \ell_t}{\partial \varrho_t} = -\frac{sgn(\varepsilon_t)}{(1-sgn(\varepsilon_t)\varrho_t)}\frac{(1+\eta)\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho_t))^2\sigma_t^2+\eta\varepsilon_t^2} = -\frac{sgn(\varepsilon_t)}{(1-sgn(\varepsilon_t)\varrho_t)}w_t\zeta_t^2. \tag{B3}$$

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:

$$\mathcal{I}_{t|t-1} = \mathbb{E}_{t-1}[\nabla_t\nabla_t'] = \begin{bmatrix} \frac{(1+\eta)}{(1+3\eta)(1-\varrho_t^2)\sigma_t^2} & 0 & -\frac{4c(1+\eta)}{\sigma_t(1-\varrho_t^2)(1+3\eta)} \\ 0 & \frac{1}{2(1+3\eta)\sigma_t^4} & 0 \\ -\frac{4c(1+\eta)}{\sigma_t(1-\varrho_t^2)(1+3\eta)} & 0 & \frac{3(1+\eta)}{(1-\varrho_t^2)(1+3\eta)} \end{bmatrix}. \tag{B4}$$

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

$$\frac{\partial \ell_t}{\partial \gamma_t} = \frac{\partial \ell_t}{\partial \sigma_t^2}\frac{\partial \sigma_t^2}{\partial \gamma_t}, \quad \frac{\partial \ell_t}{\partial \delta_t} = \frac{\partial \ell_t}{\partial \varrho_t}\frac{\partial \varrho_t}{\partial \delta_t}, \tag{B5}$$

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

$$J_t = \frac{\partial(\mu_t, \sigma_t^2, \varrho_t)}{\partial f_t'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2\sigma_t^2 & 0 \\ 0 & 0 & 1-\varrho_t^2 \end{bmatrix}. \tag{B6}$$

As such the scaled score reads:

$$\mathbf{s}_t = (J_t'\text{diag}(\mathcal{I}_t)J_t)^{-\frac{1}{2}}J_t'\nabla_t = \begin{bmatrix} s_{\mu t} \\ s_{\sigma t} \\ s_{\varrho t} \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{(1+3\eta)(1-\varrho_t^2)}{(1+\eta)}}w_t\zeta_t \\ \sqrt{\frac{(1+3\eta)}{2}}(w_t\zeta_t^2-1) \\ -sgn(\varepsilon_t)\sqrt{\frac{(1+sgn(\varepsilon_t)\varrho_t)(1+3\eta)}{3(1-sgn(\varepsilon_t)\varrho_t)(1+\eta)}}w_t\zeta_t^2 \end{bmatrix}. \tag{B7}$$

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

B.2 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

$$\frac{\partial \ell}{\partial \gamma} = \frac{\partial \ell}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial \gamma} = \frac{\partial \ell}{\partial \sigma^2}2\sigma^2.$$

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

$$\frac{\partial \ell}{\partial \bar{\gamma}} = \frac{\partial \ell}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial \gamma}\frac{\partial \gamma}{\partial \bar{\gamma}} = \frac{\partial \ell}{\partial \sigma^2}2\sigma^2.$$

The same applies for the second component. □

Proof. For the shape parameter $\varrho = \tanh \delta$, we model $\delta = \operatorname{arctanh} \varrho$ 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.3 Robustness to outliers

Here we consider the limiting behaviour of the *Skt* scaled scores. We show that for (standardized) forecast errors approaching positive (negative) infinity, the scaled scores either converge to zero, implying a trimming of the outliers, or converge to a positive (negative) constant, akin to Winsorizing extreme observations (see, e.g., Caivano and Harvey, 2014). Specifically,

$$\lim_{\zeta \to \pm\infty} s_{\mu t} = 0, \tag{B8}$$

$$\lim_{\zeta \to \pm\infty} s_{\sigma t} = \frac{1+\eta}{\eta}\sqrt{\frac{(1+3\eta)}{2}}, \tag{B9}$$

$$\lim_{\zeta \to \pm\infty} s_{\varrho t} = \mp\frac{1+\eta}{\eta}\sqrt{\frac{(1+3\eta)(1+\varrho_t)}{3(1+\eta)(1-\varrho_t)}}. \tag{B10}$$

In line with results for the t distribution, the scaled score for the location trims outliers, preventing any update of the parameter (Maronna et al., 2019). The limits of the scaled scores for the scale and shape parameter, on the other hand, converge to constant factors. These are functions of the degrees of freedom parameter, as well as the conditional asymmetry parameter at time t for the shape.

B.4 Matrix representation

Given the vector of time-varying parameters, f_t, we can express the score-driven law of motion of Equation (3) in matrix form.

$$\underbrace{\begin{bmatrix} \bar{\mu}_t \\ \tilde{\mu}_t \\ \tilde{\mu}_{t-1} \\ \bar{\gamma}_t \\ \tilde{\gamma}_t \\ \bar{\delta}_t \\ \tilde{\delta}_t \end{bmatrix}}_{f_{t+1}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi_{\mu,1} & \phi_{\mu,2} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi_\gamma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi_\delta \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \bar{\mu}_{t-1} \\ \tilde{\mu}_{t-1} \\ \tilde{\mu}_{t-2} \\ \bar{\gamma}_{t-1} \\ \tilde{\gamma}_{t-1} \\ \bar{\delta}_{t-1} \\ \tilde{\delta}_{t-1} \end{bmatrix}}_{f_t} + \underbrace{\begin{bmatrix} 0 & 0 \ldots 0 \\ \beta_\mu^1 & \beta_\mu^2 \ldots \beta_\mu^k \\ 0 & 0 \ldots 0 \\ 0 & 0 \ldots 0 \\ \beta_\gamma^1 & \beta_\gamma^2 \ldots \beta_\gamma^k \\ 0 & 0 \ldots 0 \\ \beta_\delta^1 & \beta_\delta^2 \ldots \beta_\delta^k \end{bmatrix}}_{B} X_t + \underbrace{\begin{bmatrix} \varsigma_\mu & 0 & 0 \\ \kappa_\mu & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \varsigma_\gamma & 0 \\ 0 & \kappa_\gamma & 0 \\ 0 & 0 & \varsigma_\delta \\ 0 & 0 & \kappa_\delta \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} s_{\mu t} \\ s_{\gamma t} \\ s_{\delta t} \end{bmatrix}}_{\boldsymbol{s}_t},$$

with X_t being a k-dimensional vector of predictors, at time t.

B.5 Derivations

Consider the log-likelihood function

$$\ell_t(y_t|\theta, Y_{t-1}) = \log \mathcal{C}(\eta) - \frac{1}{2}\log\sigma^2 - \frac{1+\eta}{2\eta}\log\left[1 + \frac{\eta\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho)^2\sigma^2}\right], \tag{B11}$$
$$\log \mathcal{C}(\eta) = \log\Gamma\left(\frac{\eta+1}{2\eta}\right) - \log\Gamma\left(\frac{1}{2\eta}\right) - \frac{1}{2}\log\left(\frac{1}{\eta}\right) - \frac{1}{2}\log\pi,$$

where $\Gamma(\cdot)$ is the Gamma function, $sgn(.)$ is the sign function, and $\eta = 1/\nu$ is the inverse of the degrees of freedom.

Differentiating Equation (B11) with respect to location, scale and asymmetry we obtain the gradient vector $\nabla_t = \left[\frac{\partial \ell_t}{\partial \mu_t}, \frac{\partial \ell_t}{\partial \sigma_t^2}, \frac{\partial \ell_t}{\partial \varrho_t}\right]'$. Recall that $\varepsilon_t = y_t - \mu_t$, $\zeta_t = \frac{\varepsilon_t}{\sigma_t}$ and $\omega_t = \frac{(1+\eta)}{(1-sgn(\varepsilon_t)\varrho_t)^2+\eta\zeta_t^2}$, and let

$$\begin{aligned} f(\mu_t, \sigma_t^2, \varrho_t) &= 1 + \frac{\eta\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho)^2\sigma^2} \\ &= \frac{(1-sgn(\varepsilon_t)\varrho_t)^2)\sigma_t^2 + \eta\varepsilon_t^2}{(1-sgn(\varepsilon_t^2)\varrho)^2\sigma^2}. \end{aligned} \tag{B12}$$

To avoid overburdening the notation, in what follows $\frac{\partial f(x)}{\partial x} = f'_x$ and $a = -\frac{1+\eta}{2\eta}$.
The score with respect to the location parameter reads

$$\frac{\partial \ell_t}{\partial \mu_t} = w_t \frac{\zeta_t}{\sigma_t}.$$

Proof. Define

$$g(\mu_t) = a \log f(\mu_t, \sigma_t^2, \varrho_t),$$

such that $\frac{\partial \ell_t}{\partial \mu_t} = \frac{\partial g(\mu_t)}{\partial \mu_t} = a\frac{f'_{\mu_t}}{f(\mu_t, \sigma_t^2, \varrho_t)}$. For

$$f'_{\mu_t} = \frac{2\eta}{(1-sgn(\varepsilon_t)\varrho)^2\sigma^2}\varepsilon_t,$$

it follows:

$$\begin{aligned} \frac{\partial \ell_t}{\partial \mu_t} &= -\frac{1+\eta}{2\eta}\frac{2\eta}{(1-sgn(\varepsilon_t)\varrho)^2\sigma^2} \cdot \varepsilon_t \cdot \frac{(1-sgn(\varepsilon_t)\varrho)^2\sigma^2}{(1-sgn(\varepsilon_t)\varrho_t)^2)\sigma_t^2 + \eta\varepsilon_t^2} \\ &= \frac{(1+\eta)}{(1-sgn(\varepsilon_t)\varrho_t)^2)\sigma_t^2 + \eta\varepsilon_t^2}\varepsilon_t \\ &= \omega_t \frac{\zeta_t}{\sigma_t} \end{aligned}$$

. □

The score with respect to the squared scale parameter reads

$$\frac{\partial \ell_t}{\partial \sigma_t^2} = \frac{(w_t\zeta_t^2 - 1)}{2\sigma_t^2}.$$

Proof. Define

$$g(\sigma_t^2) = -\frac{\log\sigma^2}{2} + a\log f(\mu_t, \sigma_t^2, \varrho_t),$$

such that $\frac{\partial \ell_t}{\partial \sigma_t^2} = \frac{\partial g(\sigma_t^2)}{\partial \sigma_t^2} = -\frac{1}{2\sigma_t^2} + b\frac{f'_{\sigma_t^2}}{f(\mu_t, \sigma_t^2, \varrho_t)}$.

For

$$f'_{\sigma_t^2} = -\frac{\eta\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho_t)^2\sigma_t^4},$$

it follows:

$$\begin{aligned}\frac{\partial \ell_t}{\partial \sigma_t^2} &= -\frac{1}{2\sigma_t^2} - \frac{\eta+1}{2\eta}\cdot\left[-\frac{\eta\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho_t)^2\sigma_t^4}\cdot\frac{(1-sgn(\varepsilon_t^2)\varrho)^2\sigma^2}{(1-sgn(\varepsilon_t)\varrho_t)^2)\sigma_t^2+\eta\varepsilon_t^2}\right] \\ &= \frac{1}{2\sigma_t^2}\left(\frac{(\eta+1)\varepsilon_t^2}{(1-sgn(\varepsilon_t^2)\varrho)^2\sigma^2+\eta\varepsilon_t^2}-1\right) \\ &= \frac{(w_t\zeta_t^2-1)}{2\sigma_t^2}\end{aligned}$$

. □

The score with respect to the shape parameter reads

$$\frac{\partial \ell_t}{\partial \varrho_t} = -\frac{sgn(\varepsilon_t)}{(1-sgn(\varepsilon_t)\varrho_t)}w_t\zeta_t^2.$$

Proof. Define

$$g(\varrho_t) = a\log f(\mu_t, \sigma_t^2, \varrho_t),$$

such that $\frac{\partial \ell_t}{\partial \varrho_t} = \frac{\partial g(\varrho_t)}{\partial \sigma_t^2} = a\frac{f'_{\varrho_t}}{f(\mu_t,\sigma_t^2,\varrho_t)}$.
For

$$f'_{\varrho_t} = \frac{2(sgn(\varepsilon_t)-\varrho_t)\eta\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho_t)^4\sigma_t^2},$$

it follows:

$$\begin{aligned}\frac{\partial \ell_t}{\partial \varrho_t} &= -\frac{\eta+1}{2\eta}\cdot\frac{2(sgn(\varepsilon_t)-\varrho_t)\eta\varepsilon_t^2}{(1-sgn(\varepsilon_t)\varrho_t)^4\sigma_t^2}\cdot\frac{(1-sgn(\varepsilon_t^2)\varrho)^2\sigma^2}{(1-sgn(\varepsilon_t)\varrho_t)^2)\sigma_t^2+\eta\varepsilon_t^2} \\ &= \frac{\eta+1}{(1-sgn(\varepsilon_t)\varrho_t)^2)+\eta\zeta_t^2}\cdot\frac{(sgn(\varepsilon_t)-\varrho_t)\zeta_t^2}{(1-sgn(\varepsilon_t)\varrho_t)^2} \\ &= -\frac{sgn(\varepsilon_t)}{(1-sgn(\varepsilon_t)\varrho_t)}\omega_t\zeta_t^2\end{aligned}$$

. □

Notes:

$$\frac{\partial sgn(x)}{\partial x} = \begin{cases}0, & x\neq 0\\ \infty, & x=0\end{cases}$$

$$\mathbb{E}[sgn(X)] = 1\cdot P(X>0) + (-1)\cdot P(X<0)$$

C On the importance of dynamic conditional moments

In this Section, we consider a restricted specification of the model in line with the work of Plagborg-Møller et al. (2020). As noted in Section 3, when the autoregressive features of the model is dropped out, and the information content of past prediction errors (condensed into the score) is ignored, the model collapses to a parametric model for the distribution of GDP growth akin to the one in Plagborg-Møller et al. (2020).[6] Specifically, setting $\varsigma_\mu = \phi_{\mu,1} = \phi_{\mu,2} = \kappa_\mu = \varsigma_\gamma = \phi_\gamma = \kappa_\gamma = \varsigma_\delta = \phi_\delta = \kappa_\delta = 0$, we obtain

$$\mu_{t+1} = \bar{\mu} + \beta'_\mu X_t, \tag{C1}$$

$$\gamma_{t+1} = \bar{\gamma} + \beta'_\gamma X_t, \tag{C2}$$

$$\delta_{t+1} = \bar{\delta} + \beta'_\delta X_t, \tag{C3}$$

so that the parameters' laws of motion become linear functions of predictors only.

We estimate the model above using as predictors the 4 disaggregated components of the NFCI, as well as the factors used in Plagborg-Møller et al. (2020) (including in both cases lagged values of GDP growth).[7] Table C1 reports the Deviance Information Criterion associated to these two specifications. Our baseline model strongly outperforms the two restricted specifications, both over the whole sample and during recessions. Figure C1 reports the estimated volatility and skewness from the baseline model (in blue) and the restricted specification (in red). For comparability purposes, both models are estimated using the disaggregated components of the NFCI as predictors.[8] The results of the restricted specification are broadly in line with those reported by Plagborg-Møller et al. (2020). The estimated skewness from their model is cyclical but displays large uncertainty, so that the 90% confidence bands only rarely excludes the zero. Moreover, the simplified specification of the time-varying model does not display the well-known reduction of the volatility occurred during the Great Moderation period.

Overall, comparing the restricted model to the baseline specification highlights the benefits associated with a richer framework, including the score of the predictive likelihood, long-run components, and persistent cyclical variations. In terms of volatility estimates, the presence of long-run components, as well as the addition of the score (which for the volatility is a function of the squared prediction errors) as signal for the updating, turn out to be of key importance to track the well-established stylized fact of a substantial decline in volatility from the mid-1980s (McConnell and Perez-Quiros, 2000; Stock and Watson, 2002). Persistent cyclical components, on the other hand, allow the model to recover significant variations of the skewness coefficients. This is not surprising, as the unrestricted model implies that the time variation in the parameters reflects current and past discounted variation of the predictors (as well as the information in the scaled score). This implies that high frequency noisy variations of the predictors are averaged out when incorporated into the time-varying dynamics. Moreover, while the skewness from both models set on similar levels at the beginning of the sample and during the 1990s, the asymmetry implied by the restricted specification becomes lower over the mid 1980s, while attaining higher values after the Global Financial Crisis. This pattern supports the addition of secular components, able to track the declining trend of business cycle skewness, as documented by Jensen et al. (2020).

[6] Note that Plagborg-Møller et al. (2020) consider the Skew-t of Azzalini and Capitanio (2003).

[7] For comparability purposes, we restrict the estimation sample to the 1973-2018 periods, as in Plagborg-Møller et al. (2020).

[8] The mean GDP growth extracted from the two models displays less noticeable differences.

Table C1: Deviance Information Criterion

Model	DIC	DIC^{Rec}
*Skt*4DFI	4.367	2.347
Plagborg-Møller et al. (2020) with 4DFI	4.589	2.387
Plagborg-Møller et al. (2020) Financial Factor	4.503	2.376

Note: The table reports the Deviance Information Criterion of Spiegelhalter et al. (2002) for the *Skt*-4DFI specification, and the nested version of Plagborg-Møller et al. (2020), with the four disaggregated financial indices and with the global and financial factors. In the column DIC^{Rec} we evaluate the DIC measure during recession periods.

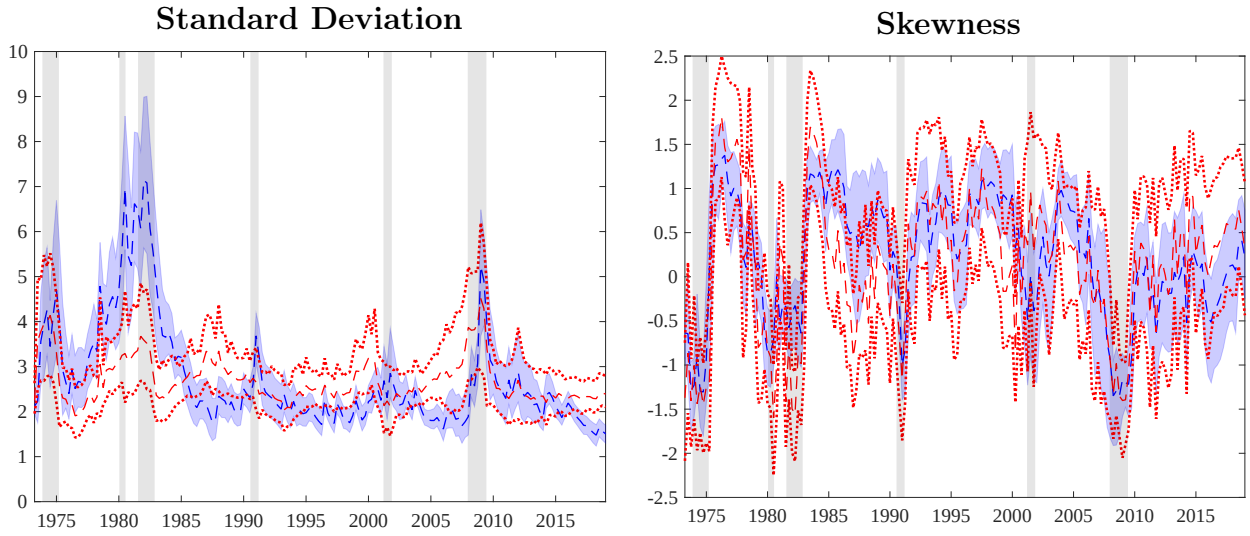

Figure C1: Time-varying moments: *Skt*4DFI vs. Plagborg-Møller et al. (2020)

Note: The plots illustrate the estimated time-varying volatility (left panel) and skewness (right panel) with 95% confidence bands. Results for the *Skt*-4DFI are reported in blue, while red lines report results for the nested version of Plagborg-Møller et al. (2020)'s model. Shaded bands represent NBER recessions.

Furthermore, we investigate to what extent our results are affected by the different choice of financial predictors. Re-estimating our baseline model using the two factors employed by Plagborg-Møller et al. (2020), we recover very similar time-varying second and third moments to the one reported for our baseline specification. Volatility displays a clear countercyclical dynamics, as well as a sharp fall in the mid-1980s. Skewness exhibits pronounced cyclical variation, with recessions characterized by negatively skewed conditional distributions, and expansions associated with positively skewed distributions. Again, one should not be surprised by these broad similarities. The inclusion of the scaled score as a driver of the time-varying parameters introduces a self-correction mechanism to the model, which allows the parameters to closely track the most pronounced features of the data, regardless of the predictors included in the model. In terms of DIC, the baseline specification with the disaggregated components of the NFCI continues to be preferred with respect to the alternative factors included by Plagborg-Møller et al. (2020).

D 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 sampling algorithm with θ^*

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

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

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

D.1 Prior specification

Starting from Equations (9) and (16), let ϕ be a generic first-lag partial autocorrelation, while ς and κ being generic score loadings for the long-run and short-run equations, respectively. Let $\beta_i \in B$, be a generic predictor loading for the parameter i, $i = \mu, \sigma, \varrho$, and let η be the inverse of the Skew-t degrees of freedom. Prior specifications for these parameters read:

$$\phi \sim NID(\mu_\phi, \sigma_\phi) \cdot I_{(\phi \in \Phi)}; \qquad \text{(D1)} \qquad \beta_i \sim NID(\mu_\beta, \sigma_\beta^i),\ i = \mu, \sigma, \varrho; \qquad \text{(D2)}$$

$$\varsigma, \kappa \sim \mathcal{G}^{-1}(a_j, b_j),\ j = \varsigma, \kappa; \qquad \text{(D3)} \qquad \eta \sim \mathcal{G}^{-1}(a_\eta, b_\eta) \cdot I_{(\eta \in H)}. \qquad \text{(D4)}$$

Prior specification D1 sets the prior distribution for the first autroregressive parameters: we target high persistence, with $\mu_\phi \lesssim 1$, with a standard tightness of $\sigma_\mu = 0.2$, in line with Bayesian Vector Autoregressive models (see, e.g., Doan et al., 1984; Sims and 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 and 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 properties 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 and 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 $H = [0, 0.33]$ set.

We also target the stationarity of the cyclical component of the location through a prior on the sum of the autoregressive coefficients (Sims and 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 $\mu_{\phi^2} \lesssim 1$ and and $\sigma_\phi^s = 0.2$.

D.2 Horseshoe priors and SAVS

Consider the regression model:

$$y_t = b' X_t + e_t, \quad e_t \sim iid(0, \sigma_e^2 I_p) \tag{D5}$$

the Horseshoe (HS) priors of (Carvalho et al., 2010) posit 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:[9]

$$\lambda^j \sim HC^+(0,1), \quad \tau \sim HC^+(0,1), \tag{D6}$$

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

$$p(z) = \frac{2}{\pi(1+z^2)}, \quad z \in \mathbb{R}^+. \tag{D7}$$

Unlike other common shrinkage priors (e.g. Ridge, Lasso), the HS priors is free of exogenous inputs, implying a fully adaptive shrinkage procedure.

Known pitfalls of shrinkage operators account for parameters not being exactly set to zero (e.g. Ridge), thus increasing parameter uncertainty, or weak selection performances when predictors are correlated (e.g. Lasso). A common remedy to the first pathology is the sparsification of the near-zero elements. Let p be the number of possible predictors, the sparsification problem can be cast as:

$$b^* = \underset{b}{\operatorname{argmin}} \left[\frac{1}{2} ||X\hat{b} - Xb||_2^2 + \sum_{j=1}^{p} m_j |b_j| \right], \tag{D8}$$

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 and Bhattacharya (2018) introduce a purely data-driven, and less burdensome solution to the choice of the parameter m_j, proven to be robust to correlated designs. 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 the magnitude of the corresponding coefficient"* (Ray and Bhattacharya, 2018), an approach similar to the Adaptive Lasso of Zou (2006). The Signal Adaptive Variable Selector (SAVS) algorithm can then be expressed as:

$$b_j^* = sgn(\hat{b}_j) ||X_j||^{-2} \max\left\{ |\hat{b}_j| \cdot ||X_j||^2 - m_j, 0 \right\}, \tag{D9}$$

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

D.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

[9] Makalic and 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.

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 $\theta = (\varsigma', \kappa', \phi', \eta)'$, we define θ^j as the draw of the parameters at the j$^{\text{th}}$ iteration of the sampler, generated from the random walk kernel:

$$\theta^j = \theta^{j-1} + \sigma_{j,s}\epsilon \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \Sigma_H^j), \tag{D10}$$

where $\sigma_{j,s}$ is a 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 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 $r(\tilde{\alpha}^s)$

$$r(\tilde{\alpha}^s) = \begin{cases} \tanh(4\tilde{\alpha}^s + 0.24), & \tilde{\alpha}^s < 0.25 \\ 1, & 0.25 \leq \tilde{\alpha}^s \leq 0.35 \\ 0.88 + \tilde{\alpha}^{s2}, & \tilde{\alpha}^s > 0.35 \end{cases} \tag{D11}$$

such that $\sigma_{j+1,s} = r(\tilde{\alpha}^s)\sigma_{j,s}$. Figure D1 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

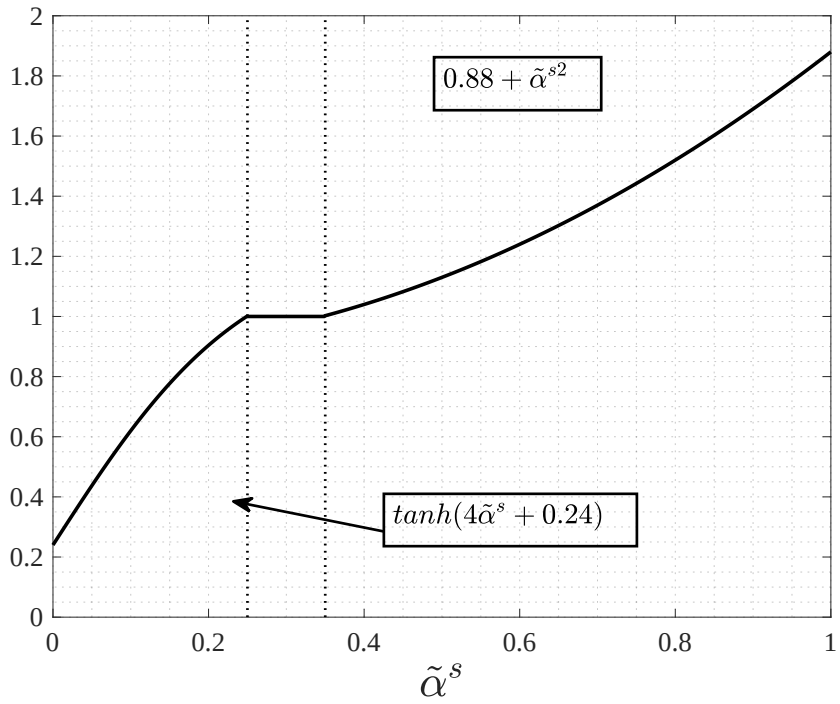

Figure D1: $r(\tilde{\alpha}^s)$ rescaling function

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%.

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 to 20000, of which we keep the last 40%. We reduce dependence between consecutive draws by *"thinning"* 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.[10] Markov Chain properties have been checked by visual inspection, as suggested by Cowles and Carlin (1996). The algorithm results to be rather efficient, and a complete chain from the baseline model in Section 4 can be obtained in less than 3 minutes, whereas for the large data model, a complete chain can be obtained in about 7 minutes.

Algorithm 1: Adaptive Random-Walk Metropolis-Hastings

for *j=1,2,...,N* **do**
- Generate: $\theta^* \sim \mathcal{N}(\theta^{j-1}, \sigma^2_{j,s}\Sigma^j_H)$;
- **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**
  - $\sigma_{j+1,s} = \sigma_{j,s} * r(\tilde{\alpha}^s)$
- **end**
- **if** $j(\mod U) = 0$ **then**
  - $\Sigma^{j+1}_H = Cov(\{\theta^i\}^H_{i=j-H}), \quad \sigma_{j,s} = \frac{2.38}{\sqrt{d}}$
- **end**

end

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

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

E Monte Carlo Exercise

In this Section we outline a Monte Carlo exercise with the aim of illustrating the small-sample properties of the Skew-t model described in Section 3. We are interested in the reliability of the estimates of the asymmetry parameter and, in particular, whether the model is able to correctly disentangle the movements in the asymmetry parameter from the movements in the location and scale parameters.

We simulate T=250 observations for the parameters of location, μ_t, scale, σ_t, and asymmetry, ϱ_t, for every period t from the $Skt_\nu(\mu_t, \sigma_t, \varrho_t)$ to obtain the series $\{Y_t\}_{t=1}^T$, then use to estimate the model discussed in Section 3 (excluding explanatory variables). Therefore, regardless of the DGP considered we always estimate a two component specification for all the three time-varying parameters. This approach mimics the true estimation process carried out on real data, to approximate the conditional distribution of the underlying GDP growth. We first consider ad-hoc specifications for the location and scale parameters (AR(2) and AR(1), respectively). Then, we turn to DGPs that replicate closely the dynamic properties of the time-varying parameters we estimate from the the data (and in particular replicate the correlation between location and scale parameter that we observe in the data). For the shape parameter, we consider a number of different DGPs, including no asymmetry, constant asymmetry, a single break in the asymmetry, and cyclical asymmetry.

In the first DGP, data are simulated from a specification featuring symmetric distributions, with independent, time-varying location and volatility, therefore aiming at checking that the model does not pick up any asymmetry when this is not a feature of the data. The second DGP considers the case in which the asymmetry parameter experiences a jump, from 0 to 0.5 after 100 observations, hence implying negatively skewed distributions in the second part of the sample. This case sheds light on the adaptiveness of the model, and provides a clear framework to illustrate the distinctive role played by the secular and transitory components. The third and fourth DGPs consider a specification for the location and scale parameter that replicate the basic features of the estimated location and scale parameters, e.g. a mildly negative correlation, whereas the shape parameter is constant, either at zero or at a positive value.[11] In the last DGP, the asymmetry parameter varies over time, with cyclical fluctuations. To ensure that we replicate the cyclical features of the parameters, we simulate the three parameters from a VAR(1), with parameters estimated on the parameters obtained from the *Skt*-4DFI specification. For each DGP we perform N=500 replications and we estimate the parameters of the two-component model using the estimation approach introduced in the main text, and detailed in Appendix D.

We simulate μ_t, $\gamma_t = \log(\sigma_t)$, and $\delta_t = \text{arctanh}(\varrho_t)$ from the following processes:

DGP1: (No asymmetry)

$$\mu_t = 0.125 + 1.2\mu_{t-1} - 0.25\mu_{t-2} + \epsilon_t, \quad \epsilon_t \sim N(0, 0.05);$$

$$\gamma_t = 0.95\gamma_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, 0.025);$$

$$\delta_t = 0.$$

[11] Specifically, we fit a VAR(1) to the location and log-scale parameters estimated from the data, and use these estimates for the DGPs.

DGP2: (Single break)

$$\mu_t = 0.125 + 1.2\mu_{t-1} - 0.25\mu_{t-2} + \epsilon_t, \quad \epsilon_t \sim N(0, 0.05);$$

$$\gamma_t = 0.95\gamma_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, 0.025);$$

$$\delta_t = \begin{cases} 0, & t \leq 100 \\ 0.5, & t > 100 \end{cases}.$$

DGP3: (Correlated location, scale and no asymmetry)

$$\begin{bmatrix} \mu_t \\ \gamma_t \end{bmatrix} = \begin{bmatrix} 0.29 \\ -0.12 \end{bmatrix} + \begin{bmatrix} 0.89 & 0.01 \\ 0.06 & 0.94] \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \gamma_{t-1} \end{bmatrix} + \epsilon_t, \quad \epsilon_t \sim N(\mathbf{0}, \Omega),$$

where

$$\Omega = \begin{bmatrix} 0.008 & -0.001 \\ -0.001 & 0.015 \end{bmatrix}$$

and $\delta_t = 0$, $\forall t$.

DGP4: (Correlated location, scale and constant asymmetry)

$$\begin{bmatrix} \mu_t \\ \gamma_t \end{bmatrix} = \begin{bmatrix} 0.29 \\ -0.12 \end{bmatrix} + \begin{bmatrix} 0.89 & 0.01 \\ 0.06 & 0.94] \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \gamma_{t-1} \end{bmatrix} + \epsilon_t, \quad \epsilon_t \sim N(\mathbf{0}, \Omega),$$

where

$$\Omega = \begin{bmatrix} 0.008 & -0.001 \\ -0.001 & 0.015 \end{bmatrix}$$

and $\delta_t = 0.5$, $\forall t$.

DGP5: (Cyclical asymmetry)

$$\begin{bmatrix} \mu_t \\ \gamma_t \\ \delta_t \end{bmatrix} = \begin{bmatrix} 0.40 \\ -0.38 \\ -0.33 \end{bmatrix} + \begin{bmatrix} 0.85 & 0.03 & -0.06 \\ 0.17 & 0.88 & 0.13 \\ 0.14 & -0.09 & 0.95 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \gamma_{t-1} \\ \delta_{t-1} \end{bmatrix} + \epsilon_t, \quad \epsilon_t \sim N(\mathbf{0}, \Omega),$$

where

$$\Omega = \begin{bmatrix} 0.008 & -0.001 & -0.007 \\ -0.001 & 0.015 & 0.008 \\ -0.007 & 0.008 & 0.025 \end{bmatrix}$$

For all the DGPs, the model correctly picks up the time variation in the location and scale with great precision, and the degree of freedom are estimated without bias. These results are not reported in the interest of brevity, but are available upon request. Hence, in what follow we comment on the ability of the model to track the underlying asymmetry of the distribution. For the first four DGPs, we report the true asymmetry parameter and the 99, 95, 90 and 75 confidence bands for the estimated counterpart. For the fifth DGP we consider the median difference between the estimated and the simulated parameters.

Figures E1 and E4 show that when the distribution is symmetric, the estimated median parameter, and the associated confidence bands, remain close to zero with small variability. Therefore, the model clearly

captures the symmetry of the underlying data generating process, and correctly disentangles the time variations of the volatility for changes in the asymmetry of the distribution in simple settings (DGP1), as well as when location and scale are correlated (DGP3). The filter is also able to detect the case with constant negative skewness (DGP4), in spite of the dynamic correlation between location and scale. The model does not confound variations of the first two moments of the data with variation of the asymmetry (Figure E5). Interestingly, for all the cases where the asymmetry parameter is fixed at a positive value, the cyclical components are always estimated to be approximately zero, suggesting that no variation at all is captured. Whereas, the secular components show minimal variation, always around the true value.[12] For this case we also investigate what happens when the initialization of the filter is away from the true parameter. The filter neatly discovers the true level of asymmetry relatively quickly (i.e. after roughly 20 observations), and the parameter sets around the true value.

In Figure E2 we report the results for the case of a single break in the asymmetry parameter. Until the jump takes place, the estimated asymmetry remains centered around zero, while it timely reacts to the jump occurring on the 100^{th} observation. Furthermore, due to the presence of a slow-moving trend component, the model fully captures the extent of the asymmetry in the second part of the sample in just 50 observations. Within this case, we also investigate to what extent the model is able to recognize the persistent nature of the shift in the asymmetry. Figure E3 reports the estimated secular and the cyclical components, respectively. The long-run components picks up the jump in a timely manner, and contributes to most of the variability of the estimated parameter. In addition, the first few observations after the jump are partly captured by the model as short-lived movements, revering to zero soon after.

Lastly, we report in Figure E6 the difference, and the associate confidence set, between the simulated and the estimated parameters. Despite co-movements in the location, scale and asymmetry parameters, the model clearly captures the cyclical variations of the latter, reported in red. The estimated parameter converges to the correct values after few observations, with small deviations, as suggested by the narrow bounds. Key for the identification of parameters' variation is that, as discussed in Section 4, shifts in the shape parameter map into a time-varying correlation between the first and second moments of the data. Therefore, in spite of the (dynamic) interaction between the key parameters of the distribution, the model correctly identifies, in a timely manner, movements in the asymmetry of the distribution.

[12] These results are not reported but available upon request.

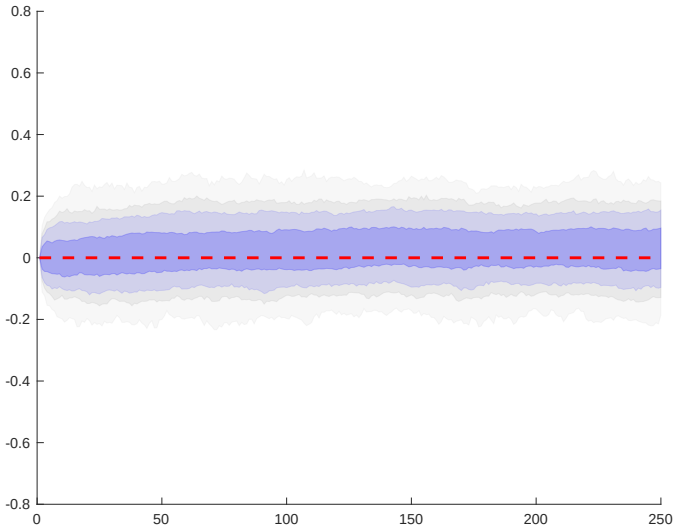

Figure E1: DGP1 - No asymmetry.

Note: The median simulated parameter is reported in red, while shades of blue represent the 99^{th}, 95^{th}, 90^{th}, 75^{th}, 25^{th}, 10^{th}, 5^{th} and 1^{st} percentiles of the empirical distribution of the estimated parameter.

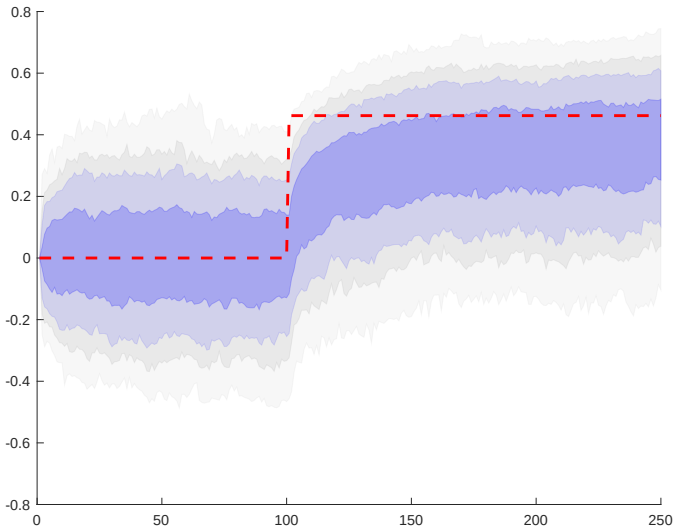

Figure E2: DGP2 - Single break in asymmetry.

Note: The median simulated parameter is reported in red, while shades of blue represent the 99^{th}, 95^{th}, 90^{th}, 75^{th}, 25^{th}, 10^{th}, 5^{th} and 1^{st} percentiles of the empirical distribution of the estimated parameter.

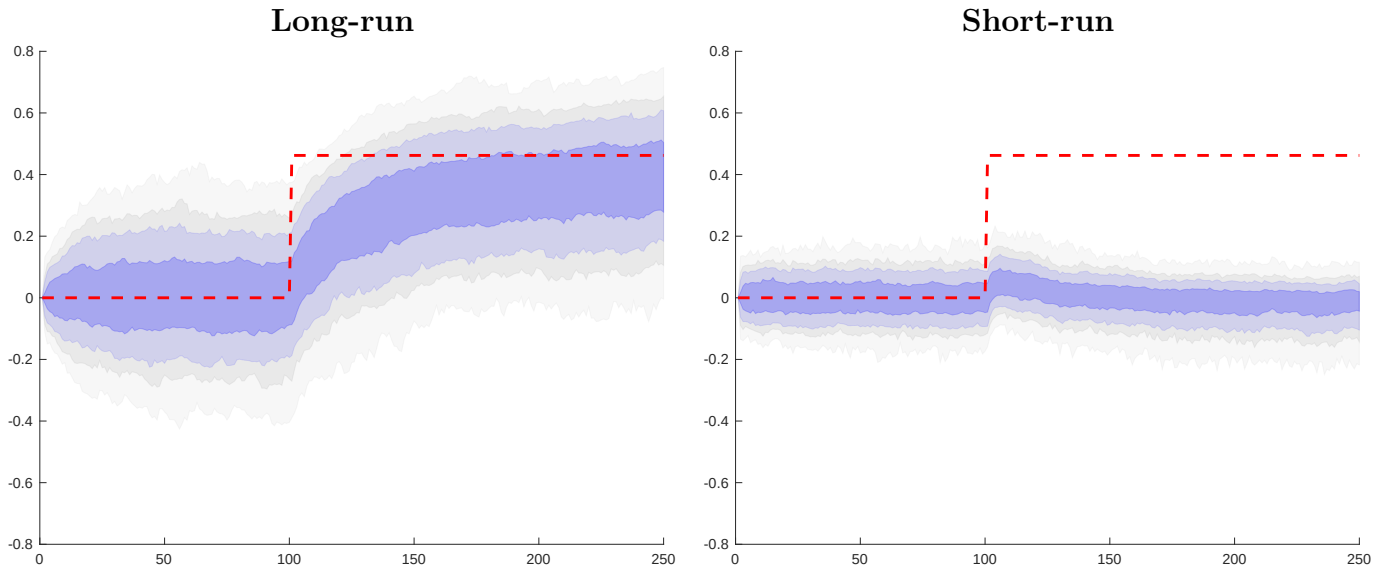

Figure E3: Secular and cyclical components for the **DGP2** case.

Note: The median simulated parameter is reported in red, while shades of blue represent the 99$^{\text{th}}$, 95$^{\text{th}}$, 90$^{\text{th}}$, 75$^{\text{th}}$, 25$^{\text{th}}$, 10$^{\text{th}}$, 5$^{\text{th}}$ and 1$^{\text{st}}$ percentiles of the empirical distribution of the estimated parameter.

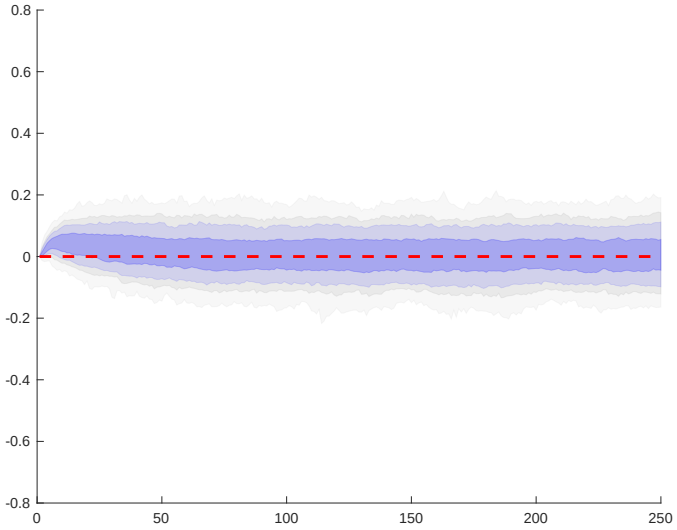

Figure E4: DGP3 - Correlated location, scale and no asymmetry.

Note: The median simulated parameter is reported in red, while shades of blue represent the 99$^{\text{th}}$, 95$^{\text{th}}$, 90$^{\text{th}}$, 75$^{\text{th}}$, 25$^{\text{th}}$, 10$^{\text{th}}$, 5$^{\text{th}}$ and 1$^{\text{st}}$ percentiles of the empirical distribution of the estimated parameter.

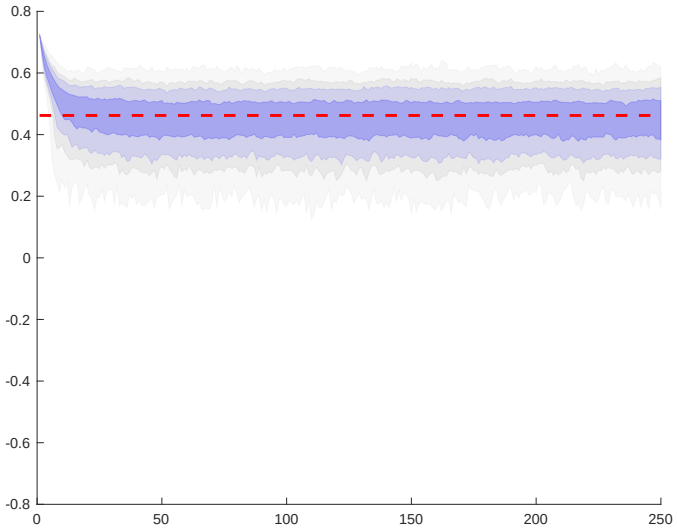

Figure E5: DGP4 - Correlated location, scale and constant negative skewness.

Note: The original parameter is reported in red, while shades of blue represent the 99$^{\text{th}}$, 95$^{\text{th}}$, 90$^{\text{th}}$, 75$^{\text{th}}$, 25$^{\text{th}}$, 10$^{\text{th}}$, 5$^{\text{th}}$ and 1$^{\text{st}}$ percentiles of the empirical distribution of the estimated parameter. The initial point for the estimation is set to 0.75.

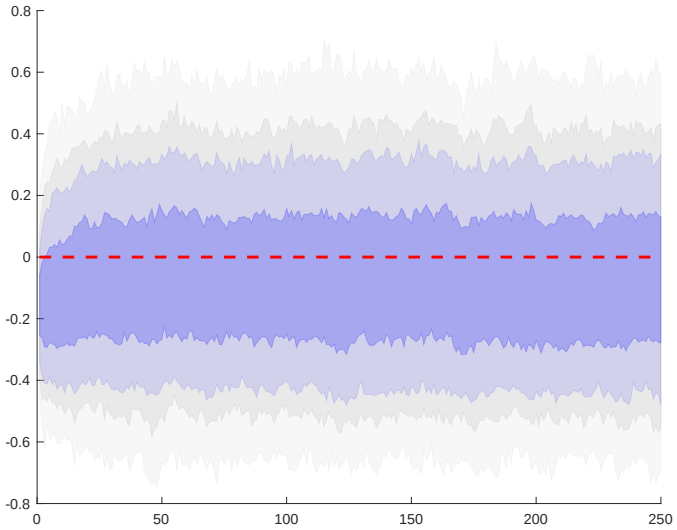

Figure E6: DGP5 - Cyclical asymmetry.

Note: The plot reports the dispersion round the difference between the simulated and estimated skewness. The red dashed line represents the perfect match while shades of blue represent the 99$^{\text{th}}$, 95$^{\text{th}}$, 90$^{\text{th}}$, 75$^{\text{th}}$, 25$^{\text{th}}$, 10$^{\text{th}}$, 5$^{\text{th}}$ and 1$^{\text{st}}$ percentiles of the empirical distribution of such difference.

F Data

In this Appendix we provide additional details on the data sources. 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 and 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 these sectors maps into three subindices, each of which gauges financial conditions in terms of risk, leverage, and credit (refer to Brave and Butters, 2012, for further details). We use quarterly data on economic activity and financial conditions over the period 1973Q1 to 2018Q4, 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 F1 plots the NFCI against GDP growth, whereas the bottom panels report the NFCI subcomponents. The NFCI (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 weights 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 (see, e.g. Bekaert and 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 and Shin (2010) and Jordà et al. (2013). Lastly, measures of credit spread and credit risk are contained in the risk and credit components (Krishnamurthy and Muir, 2017).

Alongside these subindices, the Chicago FED also produces the nonfinancial leverage (NFL) index (bottom right panel of Figure F1). 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 and Sufi (2010) argue that the build up of financial instability leading to the Great Recession 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 started in the early 2008 coincides with the beginning of the recession in the 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, put forward as a leading signal of financial distress (Drehmann et al., 2010; Jordà et al., 2017).[13]

Real-time data. In Sections 5 and 6 we evaluate the performance of our models, and compare them to a number of alternatives, both in an in- an out-of-sample forecasting exercise. For these exercises, we use real time data on GDP, provided by the Philadelphia FED's Real Time Data Research Center. This

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

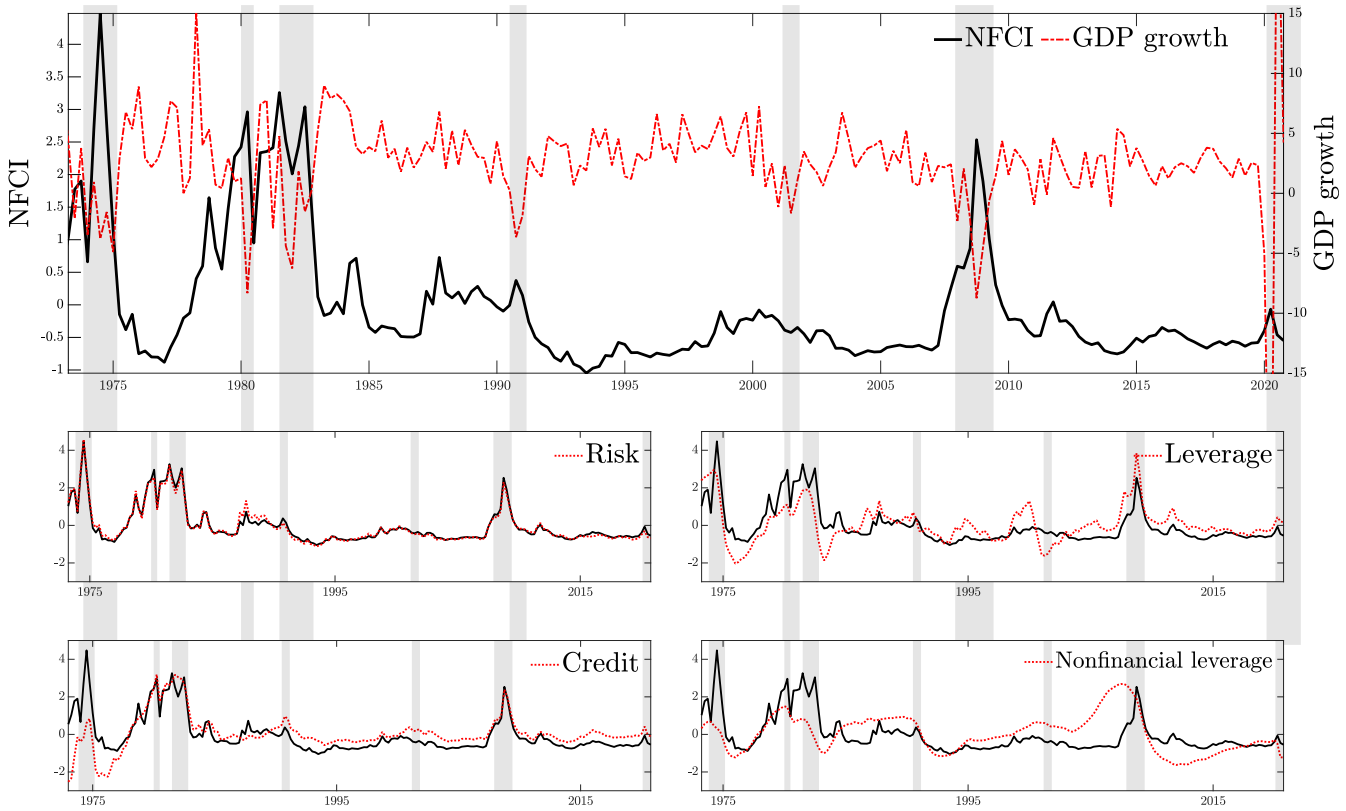

Figure F1: National Financial Condition Index and subcomponents

Note: The top panel illustrates the evolution of the NFCI and of GDP growth. Lower panels plot the risk, leverage, credit and nonfinancial leverage subcomponents, respectively. Shaded bands represent NBER recessions.

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 and 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 evaluate models using the latest release.[14]

While real time data are available for GDP, the same is not true for the NFCI and its subcomponents. These are common factors in a broad panel of financial indicators, extracted using state-space methods. As such, these indicators are always revised at any point in 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 6, we use the disaggregated contributions of all the financial indicators which form the basis of the NFCI, as illustrated in Brave and Butters (2012).[15] 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. The weekly nature of these data allows us to include all the available information, by averaging all the weeks within the reference quarter, meaning that we can use all the information made available within the reference quarter. Despite still not being real-time, this method comes the closest to mimic the available information set at the time of the forecasts.

[14]This choice implies that the forecast error produced by our models also accounts for a "measurement error" component, induced by possible GDP measurement redefinitions, that took place over the considered forecast sample (for further discussion, see Stark and Croushore, 2002). The results of our forecast exercise would be qualitatively similar to the one reported if we were to target the 'advanced' (i.e. third) revision of GDP growth.

[15]We are grateful to the authors and the Chicago FED for making the full panel of weighted contribution of the financial indicators available for this work.

F.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.[16]

F.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 and Butters (2012). In Tables F1, F2, F3 and F4 we report all the indicators, their mnemonic, and the date they become available for the index. Figure F2 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).

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

Table F1: 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 F2: 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 F3: 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 F4: 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

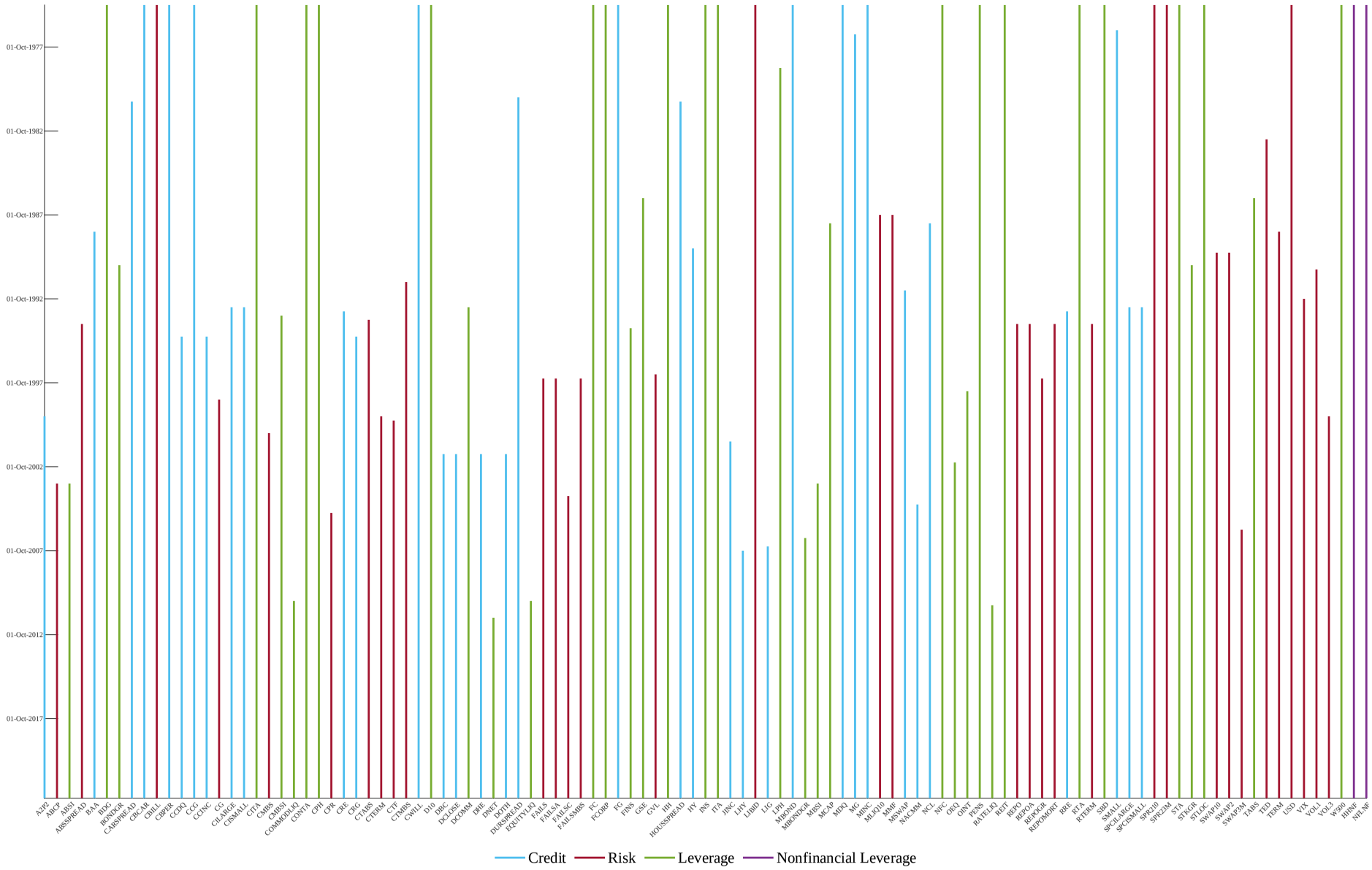

Figure F2: Timetable of predictors

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

G Additional forecast results

In this Appendix we report a number of additional results related to the forecast exercise we described in Section 5.

G.1 Lagged GDP growth as additional predictor

In this Section we show that adding past lags of the observable to the *Skt*specifications does not lead to any substantial gains in the forecasting accuracy of the models.

Table G1: Forecast performance with respect to 4DFI,Y

	One-quarter ahead				*One-year ahead*			
	MSFE	**logS**	**CRPS**	**wQS**	**MSFE**	**logS**	**CRPS**	**wQS**
Full	**0.957** (0.004)	−0.018 (0.784)	**0.971** (0.032)	0.987 (0.243)	**0.740** (0.000)	**0.104** (0.010)	0.988 (0.267)	**0.938** (0.049)
Post '00	**0.959** (0.000)	−0.008 (0.617)	**0.972** (0.006)	**0.983** (0.097)	**0.921** (0.012)	0.014 (0.378)	0.986 (0.205)	1.001 (0.513)
Rec.	0.994 (0.456)	−0.021 (0.639)	0.975 (0.256)	0.986 (0.376)	**0.756** (0.001)	−0.023 (0.611)	1.015 (0.642)	1.007 (0.546)

Note: The table reports the average forecast metrics from the *Skt*-4DFI model relative to *Skt*-4DFI,Y . We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the *Skt*4DFI model performs better than *Skt*-4DFI,Y . The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

G.2 How important is parameters' uncertainty?

The following Tables we compare forecasts made with, and without account for parameter uncertainty.

Table G2: Forecast performance - Parameter uncertainty

	One-quarter ahead		*One-year ahead*	
	MSFE	**logS**	**MSFE**	**logS**
Full	**0.965** (0.000)	−0.001 (0.510)	**0.845** (0.000)	**0.105** (0.017)
Post '00.	**0.850** (0.004)	0.011 (0.445)	**0.680** (0.006)	**0.314** (0.052)
Rec.	**0.959** (0.000)	**0.046** (0.058)	**0.770** (0.000)	**0.073** (0.079)
	CRPS	**wQS**	**CRSP**	**wQS**
Full	**0.983** (0.039)	**0.986** (0.095)	1.010 (0.753)	1.023 (0.859)
Post '00.	**0.942** (0.048)	**0.931** (0.064)	1.005 (0.552)	0.941 (0.109)
Rec.	**0.964** (0.000)	**0.955** (0.000)	1.005 (0.599)	**0.950** (0.014)

Note: In this table we compare the forecast metrics of the baseline model, *Skt*4FDI, against a version of the same model that does not account for parameter uncertainty (*w/o P.U.*). The table reports the average forecast metrics relative to the *w/o P.U.* model. We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the baseline model (which allows for uncertainty) performs better than the model without parameters uncertainty. The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses.

Table G3: Density calibration tests - Parameter uncertainty

	w/ PU	*w/o PU*	*w/ PU*	*w/o PU*
	One-quarter ahead		***One-year ahead***	
Dist.	0.883	0.763	1.162	1.217
Left tail	0.501	0.615	1.162	1.217

Note: The table reports the test statistics for the Rossi and Sekhposyan (2019) tests, based on the Cramér-von Mises type tests. The left tail score is computed over the support [0, 0.25]. 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 lowest value of the statistic.

G.3 Excluding 2020 from the forecast evaluation sample

In Sections 5.1 and 5.2, we excluded the 2020 quarters from the forecast evaluation of the recession subsample. In this Appendix we report the results for the full and the post-2000s samples excluding the pandemic period.

Table G4: Forecasting performance

	Skt	*Skt NFCI*	*Skt 4DFI*	*Skt*	*Skt NFCI*	*Skt 4DFI*
	One-quarter ahead					
	MSFE			**logS**		
Full	0.951 (0.120)	**0.847** (0.002)	**0.862** (0.004)	**0.123** (0.000)	**0.148** (0.000)	**0.074** (0.032)
Post '00	**0.881** (0.000)	**0.795** (0.000)	**0.795** (0.000)	**0.186** (0.000)	**0.231** (0.000)	**0.200** (0.000)
Rec.	0.955 (0.315)	**0.822** (0.067)	**0.813** (0.067)	**0.349** (0.006)	**0.380** (0.007)	**0.270** (0.081)
	CRPS			**wQS**		
Full	**0.964** (0.088)	**0.934** (0.009)	**0.952** (0.043)	0.984 (0.290)	**0.936** (0.016)	**0.941** (0.028)
Post '00	**0.913** (0.001)	**0.875** (0.000)	**0.890** (0.001)	**0.935** (0.037)	**0.888** (0.001)	**0.894** (0.006)
Rec.	0.962 (0.265)	0.934 (0.183)	0.928 (0.156)	0.948 (0.189)	0.914 (0.104)	**0.858** (0.025)
	One-year ahead					
	MSFE			**logS**		
Full	**0.720** (0.000)	**0.716** (0.002)	**0.693** (0.004)	**0.463** (0.000)	**0.561** (0.000)	**0.502** (0.002)
Post '00	**0.722** (0.000)	**0.697** (0.000)	**0.732** (0.005)	**0.774** (0.000)	**0.894** (0.000)	**0.867** (0.000)
	CRPS			**wQS**		
Full	**0.913** (0.003)	**0.902** (0.002)	**0.883** (0.003)	**0.779** (0.001)	**0.746** (0.001)	**0.765** (0.006)
Post '00	**0.844** (0.000)	**0.829** (0.000)	**0.829** (0.000)	**0.729** (0.000)	**0.705** (0.000)	**0.723** (0.001)

Note: The table reports the average forecast metrics relative to the Gaussian model over the 1973-2019 sample. We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the column-specific model performs better than the Gaussian benchmark. The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

Table G5: Forecast performance with respect to Adrian et al. (2019)

	One-quarter ahead				*One-year ahead*			
	MSFE	**logS**	**CRPS**	**wQS**	**MSFE**	**logS**	**CRPS**	**wQS**
Full	1.073	**0.212**	1.021	1.043	1.024	**0.560**	0.994	1.032
	(0.944)	(0.000)	(0.766)	(0.940)	(0.604)	(0.000)	(0.456)	(0.697)
Post '00	**0.905**	**0.078**	**0.951**	**0.962**	0.917	**0.371**	**0.918**	0.959
	(0.004)	(0.027)	(0.065)	(0.095)	(0.160)	(0.003)	(0.088)	(0.300)

Note: The table reports the average forecast metrics from the *Skt*-4DFI model relative to Adrian et al. (2019) over the 1973-2019 sample. We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the *Skt*4DFI model performs better than Adrian et al. (2019). The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

Table G6: Forecast performance with respect to 4DFI,Y

	One-quarter ahead				*One-year ahead*			
	MSFE	**logS**	**CRPS**	**wQS**	**MSFE**	**logS**	**CRPS**	**wQS**
Full	0.958	0.002	**0.974**	0.991	**0.737**	**0.124**	0.985	**0.931**
	(0.104)	(0.467)	(0.085)	(0.319)	(0.000)	(0.004)	(0.235)	(0.035)
Post '00	0.977	**0.032**	**0.979**	0.989	**0.926**	0.051	0.981	0.990
	(0.202)	(0.085)	(0.093)	(0.271)	(0.018)	(0.135)	(0.138)	(0.358)

Note: The table reports the average forecast metrics from the *Skt*-4DFI model relative to *Skt*-4DFI,Y over the 1973-2019 sample. We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the *Skt*4DFI model performs better than *Skt*-4DFI,Y . The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

G.4 Conditional vs. unconditional skewness

In this Section we assess the impact of accounting for unconditional skewness in the benchmark model. Specifically, we assume that this is an AR(2) model for GDP growth with Skt_ν innovations with fixed shape parameter.

Table G7: Forecasting performance against *Skt*-fix

	$\mathcal{N}$ *AR*	*Skt*	*Skt NFCI*	*Skt 4DFI*	$\mathcal{N}$ *AR*	*Skt*	*Skt NFCI*	*Skt 4DFI*
	One-quarter ahead							
	MSFE				**logS**			
Full	**0.958** (0.000)	**0.807** (0.000)	**0.783** (0.000)	**0.778** (0.000)	−0.119 (1.000)	0.002 (0.464)	0.021 (0.230)	−0.059 (0.937)
Post '00	**0.951** (0.000)	**0.770** (0.000)	**0.764** (0.000)	**0.754** (0.000)	−0.125 (1.000)	**0.056** (0.054)	**0.087** (0.004)	0.042 (0.155)
Rec.	1.003 (0.521)	0.957 (0.286)	**0.824** (0.027)	**0.815** (0.052)	−0.354 (0.992)	−0.005 (0.534)	0.026 (0.327)	−0.084 (0.799)
	CRPS				**wQS**			
Full	**0.982** (0.032)	**0.947** (0.001)	**0.924** (0.000)	**0.935** (0.001)	1.010 (0.820)	**0.969** (0.052)	**0.935** (0.000)	**0.935** (0.002)
Post '00	**0.977** (0.002)	**0.913** (0.000)	**0.891** (0.000)	**0.897** (0.000)	1.034 (0.999)	**0.951** (0.001)	**0.925** (0.000)	**0.921** (0.002)
Rec.	1.007 (0.590)	0.969 (0.212)	0.940 (0.120)	0.935 (0.182)	1.026 (0.838)	0.973 (0.271)	**0.938** (0.088)	**0.880** (0.045)
	One-year ahead							
	MSFE				**logS**			
Full	1.075 (0.999)	**0.774** (0.003)	**0.770** (0.012)	**0.746** (0.009)	−0.184 (1.000)	**0.294** (0.005)	**0.390** (0.000)	**0.325** (0.009)
Post '00	1.022 (0.971)	**0.739** (0.000)	**0.714** (0.000)	**0.747** (0.009)	−0.307 (1.000)	**0.507** (0.000)	**0.627** (0.000)	**0.587** (0.000)
Rec.	1.154 (0.922)	**0.662** (0.000)	**0.716** (0.029)	**0.629** (0.008)	−0.783 (0.997)	**0.681** (0.000)	**0.789** (0.001)	**0.994** (0.003)
	CRPS				**wQS**			
Full	0.998 (0.853)	**0.910** (0.003)	**0.900** (0.001)	**0.881** (0.003)	1.061 (1.000)	**0.826** (0.008)	**0.792** (0.002)	**0.812** (0.014)
Post '00	1.044 (1.000)	**0.883** (0.000)	**0.868** (0.000)	**0.868** (0.001)	1.070 (1.000)	**0.782** (0.000)	**0.759** (0.000)	**0.777** (0.007)
Rec.	1.033 (0.827)	**0.884** (0.002)	**0.889** (0.024)	**0.808** (0.006)	1.120 (0.997)	**0.695** (0.000)	**0.729** (0.008)	**0.653** (0.011)

Note: The table reports the average forecast metrics relative to the Gaussian model. We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that the column-specific model performs better than the Gaussian benchmark. The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. Values in **bold** are significant at the 10% level; gray shaded cells highlight the best score.

Table G8: Downside risk scores against *Skt*-fix

	Skt no-X	*Skt 4DFI*	*Skt no-X*	*Skt 4DFI*	*Skt no-X*	*Skt 4DFI*
	FZG		**ALS**		**TLF**	
	One-quarter ahead					
Full	0.831 (1.000)	0.819 (0.913)	**0.959** (0.000)	**0.948** (0.002)	0.978 (0.702)	**0.900** (0.001)
Post '00	0.693 (0.994)	0.720 (0.934)	**0.912** (0.000)	**0.926** (0.001)	0.943 (0.508)	0.915 (0.164)
Rec.	0.641 (0.998)	0.526 (0.135)	0.846 (0.941)	**0.749** (0.080)	0.858 (0.382)	**0.666** (0.017)
	One-year ahead					
Full	**0.241** (0.001)	**0.310** (0.002)	**0.397** (0.000)	**0.424** (0.002)	**0.692** (0.010)	**0.622** (0.010)
Post '00	**0.208** (0.016)	**0.170** (0.019)	**0.435** (0.012)	**0.395** (0.031)	**0.690** (0.052)	**0.475** (0.020)
Rec.	**0.137** (0.002)	**0.128** (0.007)	**0.256** (0.002)	**0.237** (0.011)	**0.441** (0.000)	**0.328** (0.011)

Note: The table reports the average downside risk test scores, expressed as ratios relative to the *Skt*-fix model. Ratios smaller than 1 indicate that the column-specific model performs better than the benchmark. The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are reported in parentheses. 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 and Komunjer (2005) tick loss function.

G.5 NFCI: last vintage vs. real time data

Here, we assess the difference in forecasting performance associated with the availability of real-time vintages of the NFCI and its four subcomponents.

Table G9: Forecast comparison with real-time NFCI

	Skt NFCI	*Skt 4DFI*	*Skt NFCI*	*Skt 4DFI*	*Skt NFCI*	*Skt 4DFI*	*Skt NFCI*	*Skt 4DFI*
	One-quarter ahead				***One-year ahead***			
	MSFE		**logS**		**MSFE**		**logS**	
2013-2020	**1.010** (0.019)	**0.984** (0.000)	0.005 (0.945)	−0.085 (0.241)	**0.942** (0.032)	**1.030** (0.000)	0.056 (0.336)	−0.058 (0.250)
2013-2019	0.900 (0.365)	0.985 (0.664)	0.067 (0.165)	−0.006 (0.782)	**0.956** (0.025)	**1.034** (0.000)	0.059 (0.378)	−0.041 (0.458)
	CRPS		**wQS**		**CRPS**		**wQS**	
2013-2020	0.979 (0.149)	**0.986** (0.014)	**0.968** (0.023)	**0.976** (0.000)	**0.949** (0.056)	1.012 (0.327)	**0.937** (0.013)	1.011 (0.260)
2013-2019	0.921 (0.182)	0.991 (0.557)	0.962 (0.327)	0.973 (0.146)	0.960 (0.112)	1.015 (0.256)	**0.943** (0.069)	1.017 (0.232)

Note: The table compares forecast produced with real time vintages of the NFCI and its subindices. Average forecast metrics are reported relative to the forecasts using the last release of the NFCI. The sample spans the period from 2013Q1 to 2020Q4 to match the availability of real-time observations. We evaluate forecasts for the full sample ('13-'20) and for a subsample excluding the pandemic ('13-'19). We use ratios for the MSFE, CRSP and wQS, and differences for the logS. Ratios smaller than 1, and positive values of the log-score differences indicate that using real-time NFCI observations improves forecasts' accuracy. The p-values for Diebold and Mariano (1995) tests, augmented with the small sample correction of Harvey et al. (1997), are in parentheses. Values in **bold** are significant at the 10% level.

H Additional material

H.1 Additional tables

Table H1: 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)	2.150 [2.024,2.282]	0.015 [0.013,0.017]	0.386 [0.330,0.440]	0.066 [0.019,0.120]			0.977 [0.971,0.983]	
Skt-No-X					1.027 [0.914,1.155]	−0.227 [−0.366,−0.096]	0.728 [0.641,0.811]	0.770 [0.689,0.834]
Skt-NFCI					0.960 [0.846,1.085]	−0.127 [−0.257,0.001]	0.662 [0.571,0.744]	0.719 [0.612,0.806]
Skt-4DFI					0.885 [0.840,0.927]	−0.032 [−0.079,0.011]	0.588 [0.521,0.654]	0.621 [0.565,0.684]

	ς_μ	κ_μ	ς_γ	κ_γ	ς_δ	κ_δ	χ	η
AR(2)				0.108 [0.099,0.121]			0.267 [0.166,0.362]	
Skt-No-X	0.025 [0.019,0.034]	0.081 [0.058,0.111]	0.027 [0.020,0.039]	0.148 [0.110,0.198]	0.027 [0.021,0.037]	0.091 [0.072,0.118]	0.283 [0.147,0.395]	0.301 [0.273,0.321]
Skt-NFCI	0.025 [0.018,0.035]	0.089 [0.063,0.131]	0.028 [0.020,0.040]	0.117 [0.084,0.158]	0.026 [0.020,0.032]	0.064 [0.051,0.080]	0.229 [0.109,0.372]	0.300 [0.269,0.321]
Skt-4DFI	0.032 [0.021,0.046]	0.083 [0.063,0.111]	0.019 [0.015,0.024]	0.110 [0.086,0.136]	0.029 [0.023,0.036]	0.070 [0.055,0.091]	0.031 [0.014,0.064]	0.312 [0.294,0.327]

Note: The table reports the estimates of the static parameters for all the models: the Gaussian *AR(2)*, the *Skt*no-X without predictors, with lags of the NFCI (*Skt*NFCI) and with lags of the subcomponents of the index (*Skt*4DFI). Confidence intervals at the 68% level are in brackets.

Table H2: Predictor loadings

		Skt NFCI	*Skt 4DFI*			
		NFCI	Leverage	NF Leverage	Risk	Credit
t	μ_t	0.001 [−0.006,0.007]	−0.001 [−0.008,0.005]	−0.000 [−0.006,0.005]	0.000 [−0.005,0.006]	0.003 [−0.005,0.010]
	σ_t^2	0.118 [0.041,0.178]	0.051 [−0.001,0.098]	0.013 [−0.047,0.074]	−0.002 [−0.060,0.068]	0.126 [0.074,0.178]
	ϱ_t	0.260 [0.198,0.324]	0.079 [0.037,0.125]	0.141 [0.087,0.195]	0.040 [−0.016,0.100]	0.206 [0.163,0.249]
$t-1$	μ_t	0.003 [−0.004,0.010]	−0.002 [−0.008,0.003]	0.002 [−0.003,0.007]	−0.000 [−0.006,0.005]	0.002 [−0.005,0.008]
	σ_t^2	−0.028 [−0.103,0.043]	−0.040 [−0.088,0.003]	−0.024 [−0.088,0.033]	0.106 [0.050,0.160]	−0.116 [−0.163,−0.068]
	ϱ_t	−0.209 [−0.278,−0.137]	−0.035 [−0.090,0.021]	−0.081 [−0.147,−0.028]	−0.054 [−0.110,0.006]	−0.163 [−0.206,−0.114]

Note: 90% confidence intervals are reported in brackets.

H.2 Additional plots

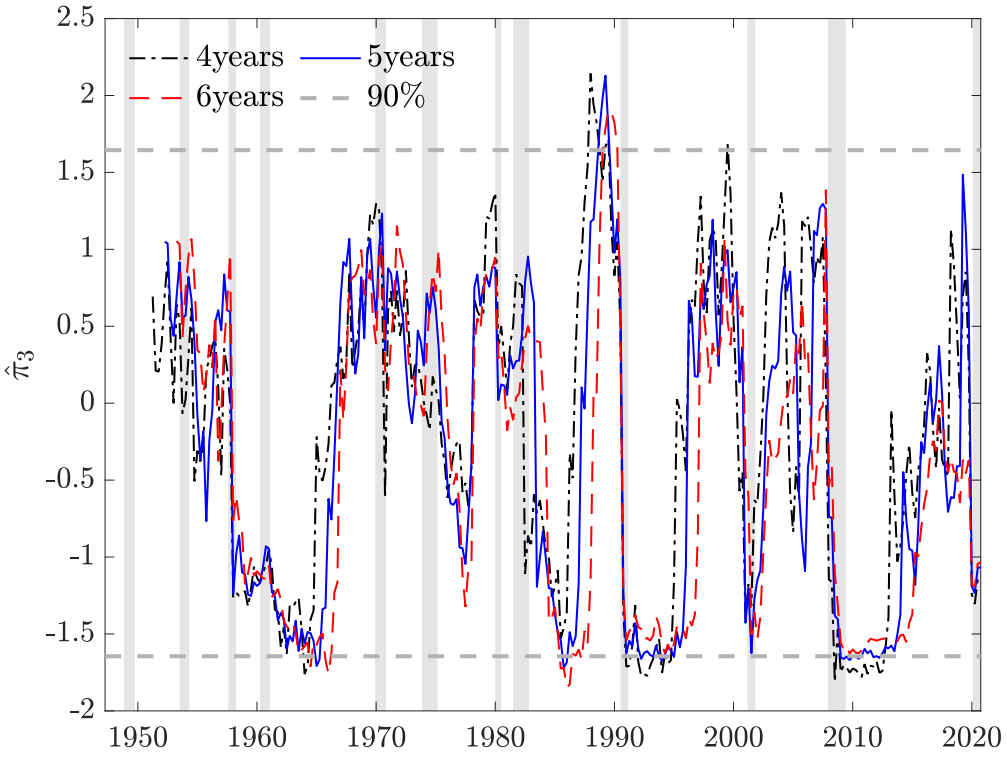

Figure H1: Bai and Ng (2005) test for skewness

Note: We report the Bai and Ng (2005) test for skewness' test statistic, for different rolling windows sizes of 1, 2, 3 and 5 years. The gray lines represent 90% (dotted) critical values. Shaded bands represent NBER recessions.

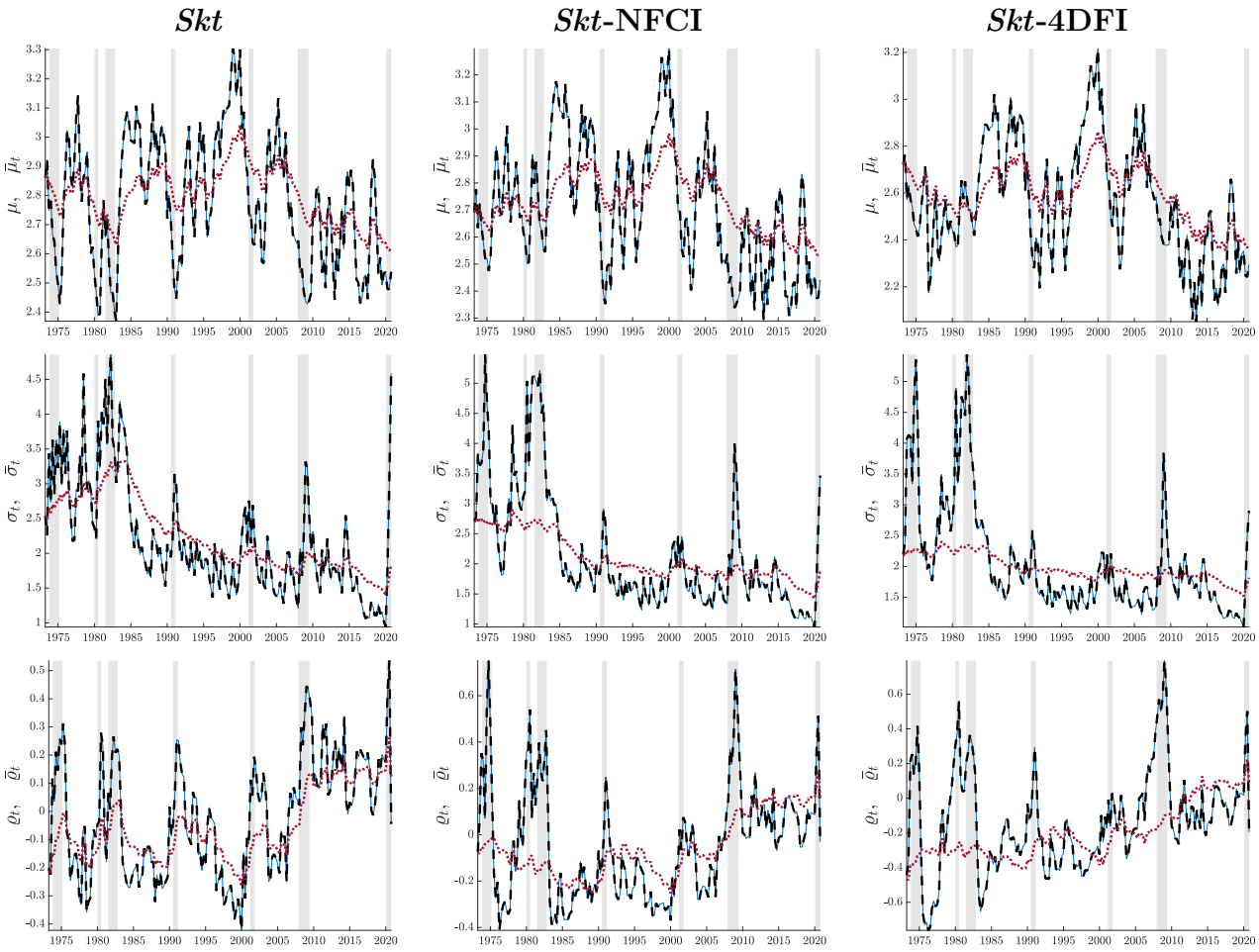

Figure H2: Time-varying parameters

Note: The plots illustrate the estimated time-varying parameters (black) for the three model specifications. Long-run components are reported in red (right scale). Shaded bands represent NBER recessions.

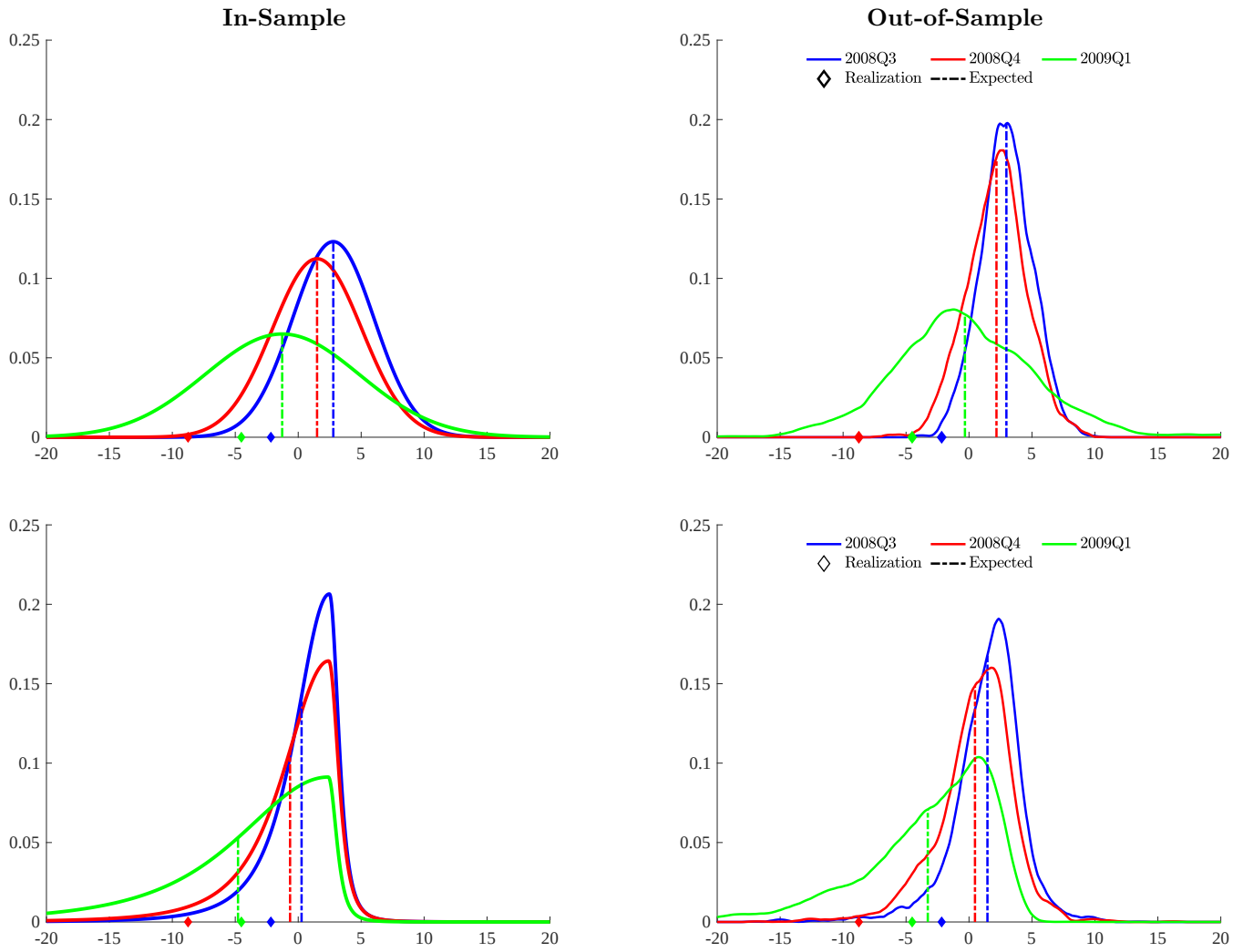

Figure H3: Predictive densities around the GFC

Note: The panels show the predictive densities for 2008Q3 (blue), 2008Q4 (red) and 2009Q1 (green) from the Gaussian AR(2) model (top panels) and from the *Skt*-4DFI specification (bottom panel). In-sample densities are in the right panels, out-of-sample densities are in the left ones. Realizations of economic growth are denote by diamonds, dot-dashed lines report expected values.

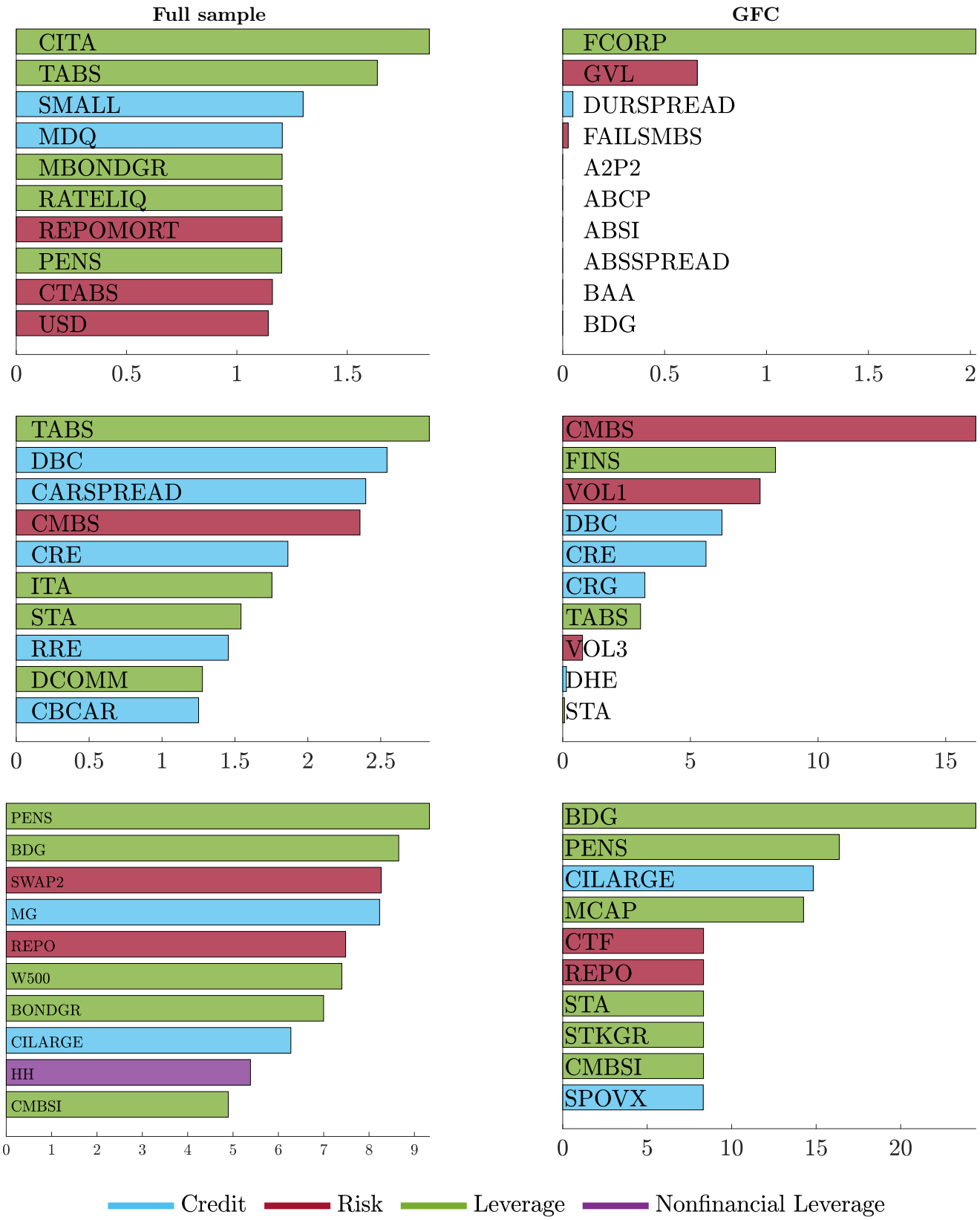

Figure H4: Top 10 predictors

Note: The bar plots report the top 10 predictors for the location, scale and shape parameters, over the forecasting sample and during the quarters of the Great Recession of 2008. The *x-axis* reports average posterior probability of inclusion, expressed in percentage terms.

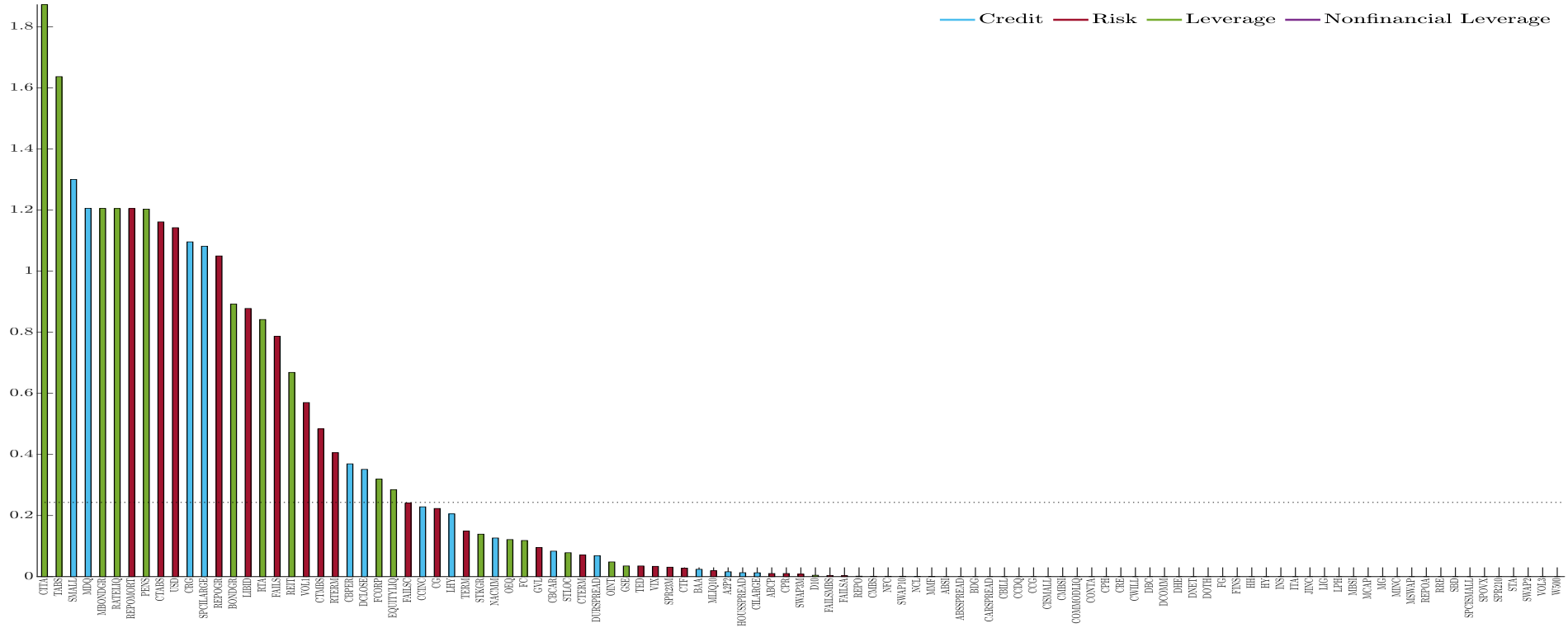

Figure H5: Predictors - Location

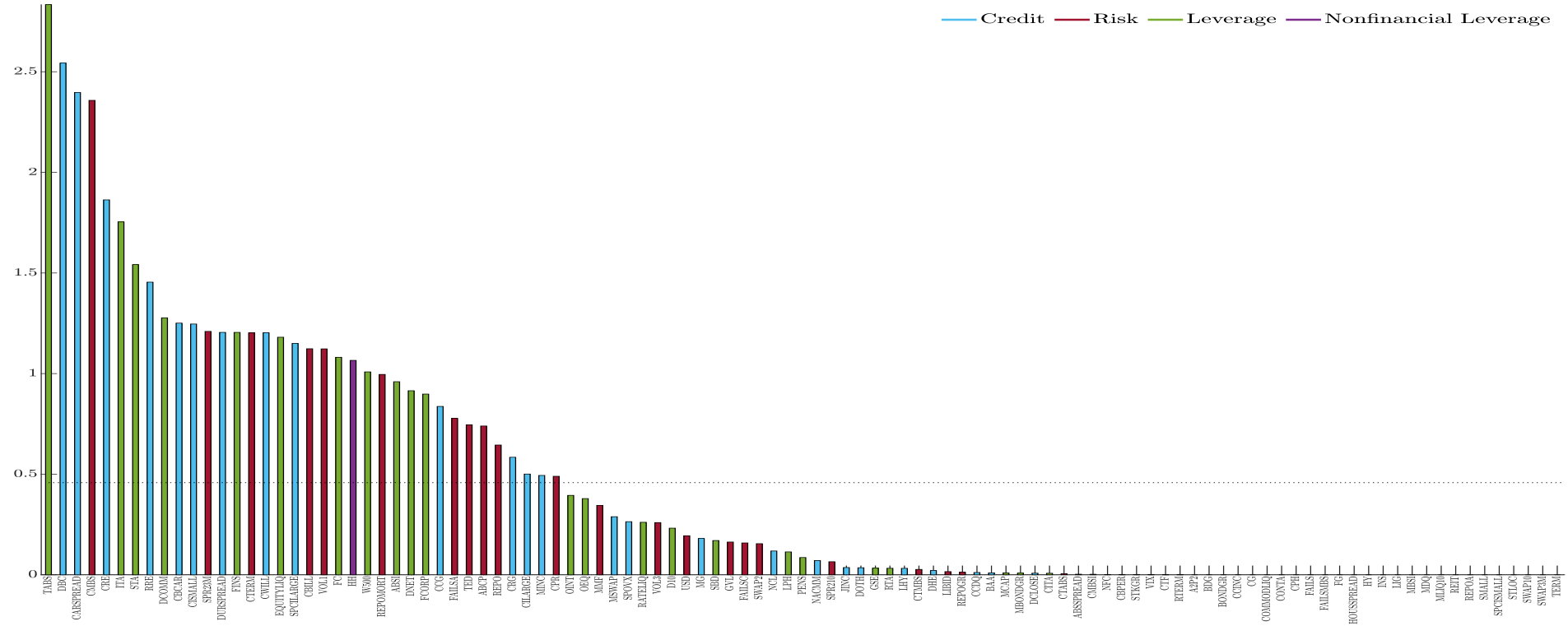

Figure H6: Predictors - Scale

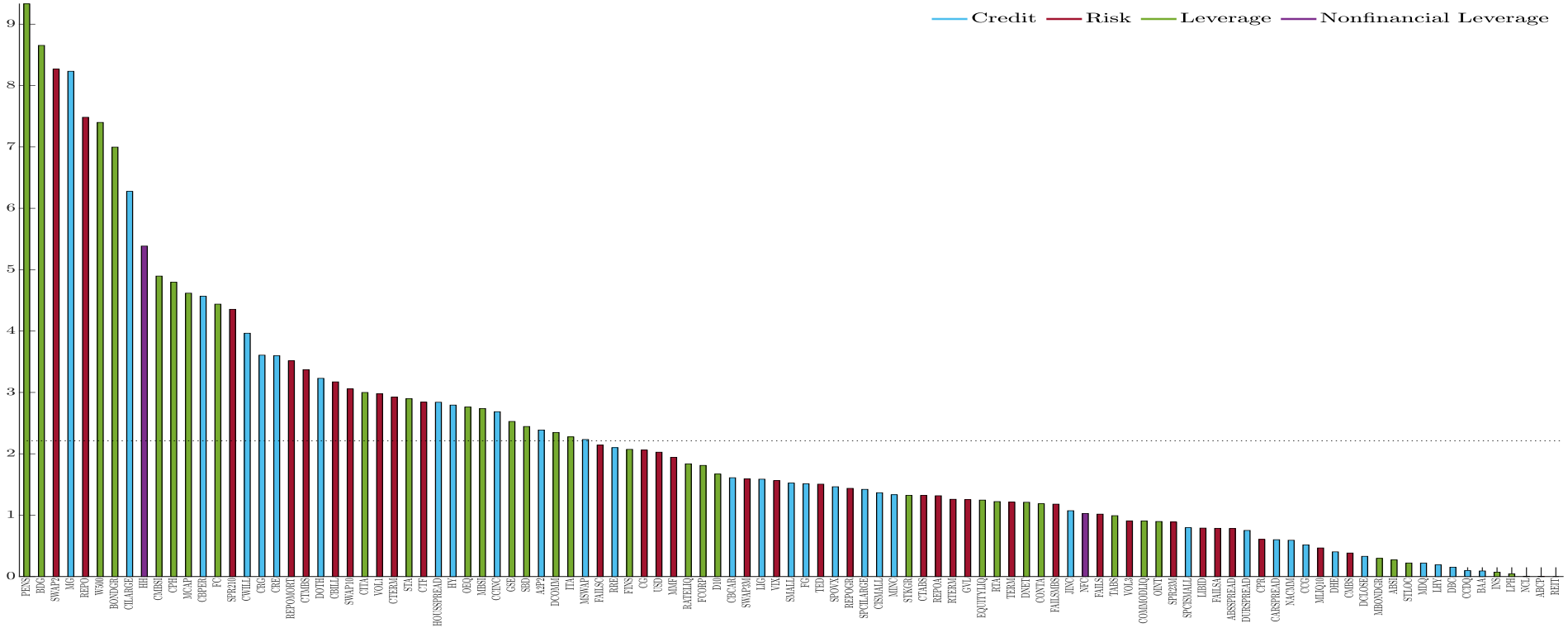

Figure H7: Predictors - Asymmetry

References

Adrian, T., N. Boyarchenko, and D. Giannone (2019): "Vulnerable Growth," *American Economic Review*, 109, 1263–89.

Adrian, T. and H. S. Shin (2010): "Liquidity and leverage," *Journal of Financial Intermediation*, 19, 418–437.

Azzalini, A. and A. Capitanio (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, 367–389.

Bekaert, G. and M. Hoerova (2014): "The VIX, the variance premium and stock market volatility," *Journal of Econometrics*, 183, 181–192.

Blasques, F., S. J. Koopman, and A. Lucas (2014): "Stationarity and ergodicity of univariate generalized autoregressive score processes," *Electronic Journal of Statistics*, 8, 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*, 498–505.

Brave, S. and R. A. Butters (2012): "Diagnosing the Financial System: Financial Conditions and Financial Stress," *International Journal of Central Banking*, 8, 191–239.

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

Caivano, M. and A. Harvey (2014): "Time-series models with an EGB2 conditional distribution," *Journal of Time Series Analysis*, 35, 558–571.

Calvori, F., D. Creal, S. J. Koopman, and A. Lucas (2017): "Testing for parameter instability across different modeling frameworks," *Journal of Financial Econometrics*, 15, 223–246.

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

Cogley, T. and T. J. Sargent (2005): "Drifts and volatilities: monetary policies and outcomes in the post WWII US," *Review of Economic Dynamics*, 8, 262–302.

Cowles, M. K. and B. P. Carlin (1996): "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review," *Journal of the American Statistical Association*, 91, 883–904.

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

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

——— (2011): "Frontiers of real-time data analysis," *Journal of Economic Literature*, 49, 72–100.

Croushore, D. and T. Stark (2001): "A real-time data set for macroeconomists," *Journal of Econometrics*, 105, 111–130.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and conditional projection using realistic prior distributions," *Econometric Reviews*, 3, 1–100.

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

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

Gelman, A., G. O. Roberts, W. R. Gilks, et al. (1996): "Efficient Metropolis jumping rules," *Bayesian Statistics*, 5, 42.

Gómez, H. W., F. J. Torres, and H. Bolfarine (2007): "Large-sample inference for the epsilon-skew-t distribution," *Communications in Statistics—Theory and Methods*, 36, 73–81.

Haario, H., E. Saksman, and J. Tamminen (1999): "Adaptive proposal distribution for random walk Metropolis algorithm," *Computational Statistics*, 14, 375–396.

Harvey, A. and M. Streibel (1998): "Testing for a slowly changing level with special reference to stochastic volatility," *Journal of Econometrics*, 87, 167–189.

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

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

Hasenzagl, T., L. Reichlin, and G. Ricco (2020): "Financial Variables as Predictors of Real Growth Vulnerability," Discussion Paper DP14322, CEPR.

Horswell, R. and S. Looney (1993): "Diagnostic limitations of skewness coefficients in assessing departures from univariate and multivariate normality: Diagnostic limitations of skewness coefficients," *Communications in statistics-simulation and computation*, 22, 437–459.

Jensen, H., I. Petrella, S. H. Ravn, and E. Santoro (2020): "Leverage and Deepening Business-Cycle Skewness," *American Economic Journal: Macroeconomics*, 12, 245–81.

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

——— (2017): "Macrofinancial history and the new business cycle facts," *NBER Macroeconomics Annual Report*, 31, 213–263.

Juárez, M. A. and M. F. Steel (2010): "Model-based clustering of non-Gaussian panel data based on skew-t distributions," *Journal of Business & Economic Statistics*, 28, 52–66.

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

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

Maronna, R. A., R. D. Martin, V. J. Yohai, and M. Salibián-Barrera (2019): *Robust statistics: theory and methods (with R)*, John Wiley & Sons.

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

Mian, A. and A. Sufi (2010): "Household Leverage and the Recession of 2007–09," *IMF Economic Review*, 58, 74–117.

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

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020): "When is Growth at Risk?" Conference draft, Brooking Paper on Economic Activity.

Ray, P. and A. Bhattacharya (2018): "Signal Adaptive Variable Selector for the Horseshoe Prior," *arXiv preprint arXiv:1810.09004*.

Sims, C. A. and T. Zha (1998): "Bayesian methods for dynamic multivariate models," *International Economic Review*, 949–968.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde (2002): "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 583–639.

Stark, T. and D. Croushore (2002): "Forecasting with a real-time data set for macroeconomists," *Journal of Macroeconomics*, 24, 507–531.

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

Zou, H. (2006): "The adaptive lasso and its oracle properties," *Journal of the American Statistical Association*, 101, 1418–1429.