Financial conditions and density forecasts for US output and inflation.

January 4, 2017

Abstract

If the links between credit markets and real economy tighten in a crisis, financial indicators might be particularly useful in forecasting the macroeconomic outcomes associated with episodes of financial distress. We examine this conjecture by using a range of linear and nonlinear VAR models to generate predictive distributions for US inflation and industrial production growth. Financial variables display significant predictive power over the Great Recession period, particularly if used within a threshold model that captures the structural break associated to the crisis. However, the Great Recession is unique: financial information and thresholds make little difference for forecasting prior to 2008.

Keywords: forecasting, financial crises, Great Recession, Threshold VAR, stochastic volatility. *JEL classification*: C53, E32, E44, G01.

1 Introduction

The economic slowdown that followed the financial crisis of 2008 suggests that the link between financial markets and the real economy might be nonlinear and that severe financial shocks might have disproportionately large costs in terms of economic activity. If that is the case, financial market information might be particularly useful in forecasting the macroeconomic outcomes associated with episodes of financial distress. In order to test this conjecture, we use monthly observations on the USA between 1973 and 2012 to analyze the predictive power of financial market indicators in vector autoregression models that allow for a range of nonlinearities, including stochastic volatility and finance-driven shifts in regimes. We find that financial variables improve output forecasts over the Great Recession period. Their predictive power emerges clearly in specifications with and without stochastic volatility, and it turns out to be stronger in a threshold model that captures changes in the size and transmission of financial shocks over time. However, the Great Recession is a *unicum*. Unlike stochastic volatility, threshold effects do not yield reliable forecast improvements prior to 2008. Furthermore, the good performance of the threshold model over the crisis would have been hard to anticipate based on the model's track record, so the warnings issued by the model in 2007 would have been very likely to go unheeded.

The question of whether financial markets predict economic activity has a long history in economics. The conclusion by Stock and Watson (2003) that *"some asset prices predict inflation or output growth in some countries in some periods"* epitomizes the common view among econometricians that financial indicators are too noisy and erratic to be exploited for macroeconomic forecasting. Yet macroeconomists have got to the conclusion that financial shocks are an important source of business cycles (Jermann and Quadrini (2012), Gilchrist and Zakrajsek (2012), Liu, Wang, and Zha (2013)), which implies that financial information should, in the right circumstances, be useful in predicting macroeconomic fluctuations. In this paper we offer two contributions to the debate. First, we focus on distributions rather than just point forecasts. Density forecasts have been studied extensively in finance and

macroeconomics.[1] However, they have not been used to study the relation between financial variables and macroeconomic aggregates, which has so far been investigated almost exclusively using point forecasts and linear models (see Stock and Watson (2003) for an earlier survey; more recent discussions can be found in Stock and Watson (2012) and Ng and Wright (2013)). Second, we examine nonlinear VAR models that are capable of capturing two potentially crucial features of the data: changes in aggregate volatility and structural breaks associated to financial crises.

Our claim that nonlinearities may be important in this context rests on two simple considerations. The first consideration is that "volatility matters". Heteroscedasticity is a pervasive feature in US data and it is known to play a significant role for forecasting (Sims and Zha (2006), Justiniano and Primiceri (2008); Clark (2011), Carriero, Clark, and Marcellino (2015)). Accounting for it is in our case particularly important. Financial markets price aggregate risk. If changes in the volatility of the fundamentals are one of the reasons why asset prices move in the first place, then exploiting their fluctuations to predict the fundamentals might be intrinsically very difficult. The second consideration is that perhaps "crises are different". Sims (2012) and Ng and Wright (2013) suggest that structural breaks played an important role during and after the financial crisis of 2008. Macroeconomic models with financial frictions provide a natural way to formalize this possibility. Firms and households are subject to borrowing constraints that limit their access to credit markets and financial crises might constitute episodes where these constraints bind at the aggregate level, amplifying the propagation of real and financial shocks with potentially dramatic implications for the dynamics of the economy.[2] This mechanism has first-order implications for forecasting: financial variables might become more informative if and when borrowing constraints tighten and these amplification effects are acti-

[1] A non exclusive list of density forecasting applications includes Clements and Smith (2000), Cogley, Morozov, and Sargent (2005), Geweke and Amisano (2010), Jore, Mitchell, and Vahey (2010), Clark (2011), De Nicolo and Lucchetta (2012), Carriero, Clark, and Marcellino (2015).

[2] See e.g. Bianchi and Mendoza (2010), Bianchi (2011), He and Krishnamurthy (2011), Brunnermeier and Sannikov (2014). Empirical evidence in support of this possibility is provided for instance by McCallum (1991), Balke (2000), Li and Dressler (2011), Guerrieri and Iacoviello (2013) and Hubrich and Tetlow (2015).

vated.[3] By studying threshold vector autoregressions (TARs) where the dynamics of the economy change at times of financial distress we can allow for this possibility and test its relevance for forecasting. In a similar spirit, Del Negro, Hasegawa, and Schorfheide (2016) compare DSGE models with and without financial frictions, showing that the former delivers better forecasts in periods of financial turmoil but not in normal times.

Our analysis confirms that heteroscedasticity helps a great deal in forecasting output and inflation. Thresholds may help too, but not in a systematic way. Although there is clear evidence of two distinct financial regimes in US history, the 2008-2009 period is the only one where this knowledge turns out to be useful for forecasting. Furthermore, since the defining characteristic of the 'crisis' regime is an increase in the variance of the shocks rather than a change in their transmission mechanism, the line between TAR and heteroscedastic VAR models is thinner than one could expect. The warning issued by the TAR at the end of 2007 is fairly forceful: the model predicts a 20% recession probability for 2008, compared to only 5% for a heteroscedastic VAR based on the same data. But the rarity of the event, combined with the impossibility to foresee the improvement in the relative performance of the TAR, implies that it would have been extremely hard for policy makers to act upon this signal.

The structure of the paper is the following. Sections 2 and 3 describe respectively our data and forecasting models. Section 4 presents empirical evidence on the existence of finance-driven regimes in the USA. In Section 5 we document the results of the forecasting exercises and discuss the accuracy of the models before and after the Great Recession. Section 6 examines a number of robustness issues and extensions. Section 7 concludes.

[3] In the online appendix to the paper we flesh out the link between financial frictions and forecasting using a stylized partial equilibrium model with an occasionally binding borrowing constraint. When agents are close to their borrowing limits credit shocks have a stronger impact on their consumption-saving decisions and consumption is both lower and more volatile, illustrating why studying distributions as well as point forecasts may be important.

2 Data and forecasting methodology

We use monthly data covering the period between March 1973 and August 2012. Industrial production index, consumer price index and the effective federal funds rate (an average of daily figures) are taken from the Federal Reserve Bank of St. Louis (FRED) Database. Industrial production and prices are transformed into annualized log changes between month t-1 and month t, while the monetary policy rate is used in levels. We do not use real-time data. Hence, our forecasts and statistics are not comparable to those presented e.g. in Clark (2011). To capture the state of financial markets we use the Financial Condition Index (FCI) constructed and maintained by the Chicago Fed. FCI is constructed using dynamic factor analysis from a set of 120 series that relate to money, debt and equity markets, as well as the leverage of financial intermediaries, and it represents an extremely broad indicator of aggregate financial conditions in the USA (Brave and Butters (2012)). This has two key advantages. First, by including FCI we effectively turn our VARs into factor models based on a larger information set, thus minimizing the risk that the small number of series we consider biases the results in favour of nonlinear models. Second, since the predictive power of different financial variables can change over time, using a broad indicator reduces the risk of obtaining results that are too heavily affected by the idiosyncratic behavior of specific variables in specific periods.[4]

The forecasting models we consider include a linear VAR, a VAR with stochastic volatility and a Threshold VAR. Models and estimation are described in Section 3. All models are estimated recursively over an expanding data window. Starting from an initial 1973.03–1983.04 window, this gives a set of 354 out-of-sample forecasts. We examine horizons of one, three, six and twelve months. Forecasts at horizons greater than one month are obtained recursively. For output and inflation, which are modelled in first differences, we look at cumulative growth rates. All predictive densities are estimated using kernel methods rather than parametric approximations as in *e.g.* Clark (2011) in order to take into account the nonlinear nature of the

[4]The choice of the financial indicator and the information set are obviously important for our analysis. We discuss them in greater detail in Section 6.

models.[5] Point forecasts are calculated as the arithmetic means of the predictive densities and evaluated in terms of root mean square errors (RMSE). The accuracy of the densities is evaluated mainly through the models' log-scores (LS), which measure the (log) likelihood that the model assigns to the actual observations based on lagged information (see Mitchell and Wallis (2011) and references therein). RMSE and LS are routinely used to compare the average performance of a set of models over a given period. Another important issue is how a forecaster would choose among the models on the basis of the information available at various points within the sample period. We investigate this problem in two ways. First, we calculate log-predictive Bayes factors that summarize the differences between the models' cumulative log-scores at every date t (Mitchell and Wallis (2011); Geweke and Amisano (2010)). Second, following Giacomini and White (2006), we test whether the differences in accuracy among models can themselves be predicted out of sample. These tests are based on the principle that, given a pair of models $\{A, B\}$ and some accuracy criterion X_t, one can define a decision criterion $C_t^X = (X_t^A - X_t^B)$, regress C_t on a set of time-t covariates, and then select the model that is expected to work better in the future (that is, prefer A to B whenever $E_t X_{t+1} > 0$). [6]

3 Forecasting models

3.1 Linear VAR

The benchmark model that we use is the following Bayesian VAR(13) model:

$$Y_t = c + \sum_{j=1}^{P} B_j Y_{t-j} + \Omega^{1/2} e_t, e_t \tilde{} N(0,1) \tag{1}$$

[5] Although the densities generated by linear VARs are Gaussian by construction, the densities from the threshold VARs are generally mixtures of Gaussians, the mixing being caused by endogenous changes in regime occurring over the forecasting horizon. Kernel methods are thus more reliable than parametric approximations in our case (see the online appendix for details).

[6] We refer the reader to the Technical Appendix for a more detailed description of the evaluation criteria based on Giacomini and White (2006).

where Y_t denotes the $T \times N$ data matrix of endogenous variables described below. We introduce a natural conjugate prior for the VAR parameters à la Sims and Zha (1998) (see also Bańbura, Giannone, and Reichlin (2010)):

$$Y_{D,1} = \begin{pmatrix} \frac{diag(\gamma_1\sigma_1...\gamma_N\sigma_N)}{\tau} \\ 0_{N\times(P-1)\times N} \\ ..... \\ diag\left(\sigma_1...\sigma_N\right) \\ ..... \\ 0_{1\times N} \end{pmatrix}, and\ X_{D,1} = \begin{pmatrix} \frac{J_P\otimes diag(\sigma_1...\sigma_N)}{\tau} & 0_{NP\times 1} \\ 0_{N\times NP} & 0_{N\times 1} \\ ..... & \\ 0_{1\times NP} & c \end{pmatrix} \tag{2}$$

where γ_1 to γ_N denotes the prior mean for the coefficients on the first lag, τ is the tightness of the prior on the VAR coefficients and c is the tightness of the prior on the constant terms. In our application, the prior means are chosen as the OLS estimates of the coefficients of an AR(1) regression estimated for each endogenous variable using a training sample. As is standard for US data, we set $\tau = 0.1$. The scaling factors σ_i are set using the standard deviation of the error terms from these preliminary AR(1) regressions. Finally we set $c = 1/10000$ in our implementation indicating a flat prior on the constant. We also introduce a prior on the sum of the lagged dependent variables by adding the following dummy observations:

$$Y_{D,2} = \frac{diag\left(\gamma_1\mu_1...\gamma_N\mu_N\right)}{\lambda},\ X_{D,2} = \begin{pmatrix} \frac{(1_{1\times P})\otimes diag(\gamma_1\mu_1...\gamma_N\mu_N)}{\lambda} & 0_{N\times 1} \end{pmatrix} \tag{3}$$

where μ_i denotes the sample means of the endogenous variables calculated using the training sample. As in Bańbura, Giannone, and Reichlin (2010), the tightness of this sum of coefficients prior is set as $\lambda = 10\tau$. Given the natural conjugate prior, the conditional posterior distributions of the VAR parameters $B = vec([c, B_1; B_2..; B_j]\ )$ and Ω take a simple form and are defined as

$$G\left(B\left|\Omega\right.\right) \sim N(B^*, \Omega \otimes \left(X^{*\prime}X^*\right)^{-1}) \tag{4}$$

$$G(\Omega\,|B\,) \sim IW(S^*, T^*). \tag{5}$$

The posterior means are given by $B^*=(X^{*\prime}X^*)^{-1}(X^{*\prime}Y^*)$ and $S^*=(Y^*-X^*)\tilde{B}'(Y^*-X^*\tilde{B})$, where $Y^*=[Y;Y_{D,1};Y_{D,2}]$, $X^*=[X;X_{D,1};X_{D,2}]$ and $\tilde{B}$ is the draw of the VAR coefficients B reshaped to be conformable with X^*. T^* denotes the number of rows of Y^*. A Gibbs sampler offers a convenient method to simulate the posterior distribution of B and Ω by drawing successively from these conditional posteriors. We employ 20,000 iterations using the last 5000 for inference. In particular, these 5000 draws are used to produce the forecast density:

$$G(Y_{t+K}\,|Y_t\,) = \int G(Y_{t+K}\,|Y_t, \Gamma\,) \times G(\Gamma\,|Y_t\,)\, d\Gamma \tag{6}$$

where $K = 1, 2, ..12$ and $\Gamma = \{B, \Omega\}$. The forecast density can be easily obtained by simulating Y_t forward using the Gibbs draws for B and Ω.We use two versions of this model: the first one, henceforth labelled VAR, only contains our macroeconomic variables ($Y_t = \{y_t, r_t, \pi_t\}$), while the second one, labelled VAR_f, also includes the Financial Condition Index ($Y_t = \{y_t, r_t, \pi_t, f_t\}$).

3.2 VAR with stochastic volatility

The VAR with stochastic volatility (VAR^σ) is defined by the following equation:

$$Y_t = c + \sum_{j=1}^{P} B_j Y_{t-j} + \Omega_t^{1/2} e_t \tag{7}$$

Following Cogley and Sargent (2005), the time-varying covariance matrix is factored as $\Omega_t = A^{-1}H_tA^{-1'}$ where A is lower triangular and $H_t = diag([h_{1t}, h_{2t}, .., h_{Nt}])$. The log stochastic volatility associated with the i$^{\text{th}}$ orthogonal error, $\ln h_{it}$ follows a random walk:

$$\ln h_{it} = \ln h_{it-1} + G_{it}, G_{it} \tilde{} N(0, g_i) \tag{8}$$

The priors for the VAR coefficients remain as in the linear VAR model. The prior for the free elements of A are assumed to be normal with the prior mean equal to zero and the variance set at 1000. The prior for g_i is inverse Gamma with a scale parameter equal to $1e-4$ and degrees freedom set at 1. In short, we employ uninformative priors for these parameters. As shown in Cogley and Sargent (2005), a Metropolis within Gibbs algorithm can be used to approximate the marginal posterior distributions. The independence Metropolis step of Jacquier, Polson, and Rossi (1994) is used to draw from the conditional posterior distribution of h_{it}. Given the stochastic volatilities, the model collapses to a heteroscedastic VAR and standard Gibbs steps can be used to draw the remaining parameters. As in the case of the linear VAR, we estimate a version of this model that only contains macroeconomic variables (VAR^{σ}) and one that also includes the financial indicator (VAR_f^{σ}).

3.3 Threshold VAR

The Threshold VAR model (TAR_f) is defined as follows:

$$Y_t = \left[c_1 + \sum_{j=1}^{P} B_{1,j} Y_{t-j} + \Omega_1^{1/2} e_t \right] S_t + \left[c_2 + \sum_{j=1}^{P} B_{2,j} Y_{t-j} + \Omega_2^{1/2} e_t \right] (1 - S_t), \quad (9)$$

where

$$S_t = 1 \Longleftrightarrow z_{t-d} \leq z^*. \quad (10)$$

The vector of endogenous variables is $Y_t = \{y_t, r_t, \pi_t, f_t\}$. As in the case of VAR_f and VAR_f^{σ}, a financial shock is thus implicitly added to the set of fundamental shocks driving the dynamics of the economy. This model, however, allows for the possibility of two distinct regimes depending on the level taken by some variable z_{t-d} relative to an unknown threshold z^*. In our application the threshold variable is assumed to be the d^{th} lag of the financial conditions indicator f_{t-d}, and both the delay d and the threshold z^* are assumed to be unknown parameters. This means that the size and propagation mechanisms of both real and financial shocks are allowed to change when the economy is experiencing financial distress. This formulation is appealing because

the parameters $\{c_s, B_{s,j}, \Omega_s\}$ can be regarded as the reduced-form counterparts of the two sets of first-order conditions that arise in a general equilibrium model with an occasionally binding credit constraint. As such, they should capture the behavior of the US economy in periods when the constraint binds ($s = 0$) and when it does not ($s = 1$).

We again impose a natural conjugate prior on the VAR parameters that appear in equation (9), and we set the prior tightness in an identical fashion to the linear case discussed in Section 3.1. Importantly, we choose identical priors for the two regimes. This assumption is clearly counterintuitive – the model is motivated precisely by the presumption that the dynamics differ across regimes, particularly after a financial shock – but it is also conservative in the sense that it 'lets the data speak' on how relevant such differences are in practice. In principle one could exploit the restrictions implied by the models discussed in Section **??** to parameterize the regimes in a different way; this extension is left to future work. As far as equation (10) is concerned, we assume a flat prior on the delay d, limiting its maximum value to 12, and we assume a normal prior for $z^* \sim N(\bar{z}, \bar{v})$, where $\bar{z} = 1/T\sum_{i=1}^{T} z_t$ and $\bar{v} = 10$. Given the scale of the Financial Condition Index this represents a fairly loose prior. We employ the Gibbs sampler introduced in Chen and Lee (1995) to simulate the posterior distribution of the unknown parameters. Given an initial value for z^* and d, the conditional posterior for the VAR parameters in the two regimes is standard and given by equations 4 and 5. Given a draw for the VAR parameters and a value for d, a random walk Metropolis Hastings step can be employed to sample z^*. We draw candidate value of z^*_{new} from $z^*_{new} = z^*_{old} + \Psi^{1/2}\epsilon$, $\epsilon \sim N(0,1)$. The acceptance probability is given by $f\left(Y_t \left| z^*_{new}, \Xi\right.\right) / f\left(Y_t \left| z^*_{old}, \Xi\right.\right)$, where $f(.)$ denotes the posterior density and Ξ represents all other parameters in the model. We choose the scaling factor Ψ to ensure that the acceptance rate remains between 20% and 40%. Chen and Lee (1995) show that the conditional posterior for d is a multinomial distribution with probability $L\left(Y_t \left| d, \Xi\right.\right) / \sum_{d=1}^{12} L\left(Y_t \left| d, \Xi\right.\right)$, where $L(.)$ denotes the likelihood function. We employ 20,000 iterations of the Gibbs sampler discarding the first 15,000 as burn-in. The forecast density for TAR_f is defined as in equation (6),

but in this case the set of parameters is given by $\Gamma = \{B_1, \Omega_1, B_2, \Omega_2, z^*, d\}$. Given draws from the Gibbs sampler, the density can be easily computed by iterating equations (9) and (10) K periods in the future. We also consider a variant of this model where the level dynamics are driven by a threshold structure, as in equations (9)-(10) while the residual covariance matrix follows the smoothly-changing process described in equation (8). This model, labelled TAR_f^σ, allows us to check whether the combination of regime switching and stochastic volatility yields any improvement over models that only account for one of the two nonlinearities.

4 Financial regimes in the US

Before moving to forecasting we discuss the features of the financial regimes identified by the TAR_f model described in Section 3.3. The regimes are displayed in Figure 1.[7] The shaded area represents the median estimate of $1 - S_t$, which is equal to one when the Financial Condition Index is above the estimated critical threshold (see equation (10)). For the sake of brevity we refer to this as the "crisis" regime. The US economy enters this regime in 1974-1975, in the early 1980s and around 1987, after which the economy remains in the normal regime for about two decades. The last crisis breaks out in 2008 and lasts roughly two years, covering the period from the peak in the FCI to the end of the contraction in industrial production. Given the structure of the model, the sub-periods identified as crises are by construction characterized by high financial volatility and tight credit markets. With the exception of 1987, they also turn out to be periods of weak or negative growth. This association, however, does not say anything on causality. To check whether financial shocks slow down economic activity and whether their impact is larger during crises we resort to structural impulse-response analysis. The impulse responses are calculated using Monte Carlo integration as described in Koop, Pesaran, and Potter (1996). In

[7] All results in this section are based on full-sample estimates of the TAR model. The lag length is set to 24 to insure well-behaved residuals.

particular, the responses are based on the following definition:

$$IRF_t^S = E\left(Y_{t+k} \left| \Psi_t, Y_{t-1}^s, \mu \right.\right) - E\left(Y_{t+k} \left| \Psi_t, Y_{t-1}^s \right.\right), \tag{11}$$

where Ψ_t denotes all the parameters and hyperparameters of the VAR model, k is the horizon under consideration, $S = 0, 1$ denotes the regime and μ denotes the shock. Equation (11) states that the impulse-response functions are calculated as differences between two conditional expectations representing, respectively, a forecast of the endogenous variables conditioned on one of the structural shocks μ, and a baseline forecast where all shock equal zero. These conditional expectations can be approximated via a stochastic simulation of the model. We condition the responses on observations in each regime. The impulse-response corresponding to regime $S = 0$, for instance, is obtained simulating the model for all possible starting values in that regime, Y_{t-1}^0, and then calculating the average of the responses. To identify the shocks we adopt a simple recursive scheme where y_t, π_t and r_t appear in this order, reflecting as customary the relative sluggishness of output and prices in responding to exogenous disturbances, and the financial indicator f_t is ordered last. This assumption is consistent with financial variables moving quickly in response to any news on the macroeconomic outlook, including changes in the monetary policy stance. It is also conservative from our perspective because by placing f_t last we minimize the risk of overestimating the role played by genuine financial shocks in explaining the dynamics of the system.

Figure 2 displays the estimated impact of an adverse financial shock, i.e. an increase in FCI. To save space we focus on the responses of industrial production (left column) and the FCI itself (right column).[8] The dynamics associated to normal times and crises are shown respectively in black and in red. For each regime we report median responses and 68% confidence bands. The model allows for changes in both the volatility and the transmission of the shocks across regimes, and both can play

[8]The online appendix to the paper provides impulse-responses for all variables included in the model, confidence bands for the differences across regimes, and a set of responses calculated on a pre-crisis sample ending in December 2007 (these are qualitatively similar to the full-sample IRFs discussed below).

a role in a financial crisis. To disentangle these two factors we run two experiments. In the first case (row 1) the shock is defined as a one standard deviation increase in FCI, so its absolute size is allowed to change across regimes. In the second case (row 2) the shock is normalized to 0.1 units in both regimes. The responses generally resemble those generated by a negative demand shock, with a contraction in output and inflation and a fall in the policy rate. In the first experiment, the quantitative difference between regimes is stark. The drop in output and prices is deeper and more abrupt in a crisis: output falls by up to 3% on an annual basis, against roughly 0.5% in normal times. The size of the shock plays an important role in generating the asymmetry: as the right panel shows, the standard deviation of the financial shock is estimated to be roughly three times larger during a crisis (0.3 *versus* 0.1). In the second experiment we shut down this mechanism by simulating a shock of the same size (0.1 units) in both regimes. The median responses are again more pronounced in the crisis regime, with a trough in output that is roughly twice as deep as in normal times. However, the confidence bands are wide and largely overlap across regimes; the differences turn out to be statistically significant for inflation and the policy rate but not for output (see online appendix). This suggests that, although the transmission mechanism contributes to the overall asymmetry of the IRFs across regimes, its contribution is smaller and statistically more uncertain than that of the variance of the shock. From a density forecasting perspective these mechanisms are complementary: a combination of larger variance and stronger amplification, whatever the relative weight of the two factors, implies that the predictive power of FCI should indeed be higher when the economy enters the crisis regime.[9] At the same time, the results suggest that the evidence on the role of borrowing constraints is far from being overwhelming, and that the wedge between threshold and stochastic volatility specifications might be smaller than expected.

[9]We find that an alternative TAR specification where volatilities change over time but the level dynamics are kept constant has a worse forecasting performance (see Section 6). A similar combination of changes in volatilities and transmission mechanisms is documented by Hubrich and Tetlow (2015).

5 Forecast analysis

5.1 Average forecasting accuracy

Table 1 shows the average root mean square errors (RMSE) and log-scores (LS) produced by the models over the entire evaluation period, which runs from April 1983 to August 2012.[10] We adopt as benchmark the 4-variable linear VAR that includes industrial production (y), prices (π), fed funds rate (r) and the Financial Condition Index (f_t). The model is labelled VAR_f. Using it as a benchmark is convenient because it is the simplest model that generates predictions for all variables including FCI. For VAR_f the RMSEs and LS are reported in levels, while for all alternative specifications the RMSEs are ratios and the LS differences relative to benchmark, as in e.g. Clark (2011). Moving down the table, the alternative models are a linear VAR without the financial indicator (VAR), two stochastic volatility VARs with and without the financial indicator (VAR_f^σ, VAR^σ), the threshold model with finance-driven regimes (TAR_f), and a threshold model with stochastic volatility (TAR_f^σ). The subscript f thus identifies models whose information sets includes the financial indicator and the superscript σ those that incorporate stochastic volatility.

The RMSE for industrial production are broadly similar across specifications, although VAR_f^σ tends to be marginally more accurate than the benchmark. For all other variables, introducing stochastic volatility significantly improves the accuracy of the forecasts. The improvement is of the order of 3% to 5% for π_t, it can reach 10-20% for the two financial variables r_t and f_t, and is fairly stable across horizons. In the case of inflation and interest rates, these gains emerge irrespective of whether FCI is included in the model or not. TAR_f delivers lower RSMEs for r_t and f_t but not for π_t. In terms of LS, VAR_f^σ emerges clearly as the best model of the pool. It beats the benchmark for all variables and horizons, with the only exception of the one-month ahead output projections, and it performs at least as well as the two threshold models

[10] Calibration diagnostics such as probability integral transforms and probability coverage ratios are similar across models and thus not particularly informative from our perspective. Details are available upon request.

if not better. The improvements relative to benchmark are quantitatively significant. In the case of the 12-month ahead output forecasts, for instance, the log-scores increase by roughly 20% thanks to the inclusion of heteroscedasticity. VAR^{σ} works well for for r_t and π_t but not for y_t: in this case the model is far less accurate than the benchmark. Like VAR_f^{σ}, TAR_f improves over the benchmark for all variables and horizons. The improvements are of the same order of magnitude than those delivered by VAR_f^{σ} in the case of r_t and f_t but tend to be smaller for y_t and π_t. TAR_f^{σ} beats the benchmark in forecasting r_t but tends to perform poorly for all remaining variables, and particularly for y_t. The online appendix provides some evidence on the statistical significance of the figures reported in Table 1 based on pairwise tests of the null hypothesis of equal unconditional or conditional predictive ability across models. Since the models are estimated recursively, rather than using a fixed moving window, the tests are only indicative. Subject to this important caveat, though, the tests indicate that most of the differences that emerge from Table 1 are significant, particularly in conditional terms, suggesting that the relative accuracy of the models varies widely over time and that some of this variation is predictable.

Figure 3 displays a set of industrial production forecasts from the three main models of interest, namely VAR^{σ}, VAR_f^{σ} and TAR_f. For each model the figure shows 50% and 80% prediction intervals at the three-month horizon. Although VAR^{σ} and VAR_f^{σ} generate broadly similar predictions for most of the sample period, financial information turns out to be important in the GR. From 2007 onwards VAR^{σ} becomes less accurate because it captures a large increase in variance that spreads the density symmetrically, suggesting that large positive outcomes are also more likely than before. Thanks to the presence of FCI, VAR_f^{σ} combines instead an increase in the variance of the distribution with a downward revision of the central forecast. Compared to VAR_f^{σ}, TAR_f generates wider densities throughout the sample. This larger dispersion is the reason why the model assigns a higher probability for instance to the 2001 recessions. The GR episode is different. In this case the model also captures a switch into the "crisis" regime, and this allows it to predict a trough of about -7%, compared to -5% for VAR_f^{σ}, and to generate prediction intervals that lie

entirely below the zero line. TAR_f^σ (not shown) performs even better in the GR, but this result is of limited interest because the distributions generated by this model are very wide throughout the evaluation sample. During the GR we observe that (i) VAR^σ underperforms the benchmark, (ii) both VAR_f^σ and TAR_f outperform it in terms of both RMSE and LS, and (iii) TAR_f has the best overall performance.[11] However, figure 3 shows clearly that the GR is an exception. If one focuses instead on the average performance of the models in 'bad times', defined as occurrences of the crisis regime according to the TAR_f model, then the statistics turn out to be roughly similar to those reported in table 1. In particular, in this case the models (TAR_f included) beat the benchmark in LS space but not in RMSE space (see table D.1 in the online appendix for details). The timeliness and economic significance of the warnings issued by the models ahead of the GR are discussed further in Sections 5.2 and 5.3.

A first clear message from Table 1 is that heteroscedasticity is a clear asset for density forecasting across models, variables and evaluation criteria. This result corroborates the evidence in Clark (2011), where heteroscedasticity is also found to improve point as well as density forecasts. A second message is that the financial indicator helps on average in predicting output growth, as adding f_t to VAR^σ leads to a (small) reduction in RMSE and a (large) increase in LS. Indeed, the ranking among the models' average log-scores ($VAR^\sigma < VAR_f < VAR_f^\sigma$) points to the twofold conclusion that (i) accounting for the role of FCI as well as heteroscedasticity might be important, and (ii) the advantages of a stochastic volatility specification become more substantial if the model also includes financial information. However, figure 3 suggests that the relevance of f_t might hinge heavily on the GR episode: the gap between VAR^σ and the finance-augmented models VAR_f^σ and TAR_f effectively opens up only after 2007. Furthermore, the presence of f_t in the information set does generally improve the forecasts for inflation and interest rates.

[11] In particular, between December 2007 and June 2009 (the GR as dated by the NBER), the average RMSE ratios of VAR^σ, VAR_f^σ, TAR_f, TAR_f^σ for annual output growth are 1.03, 0.97, 0.96, 0.98, and the corresponding log-score differences are -28.1, 0.6, 1.2 and 2.2.

The evidence on threshold effects is more mixed. The regime structure embedded in TAR_f improves both RMSE and LS for r_t and f_t. In the case of y_t and π_t, however, the model only beats the benchmark along the LS dimension; the only exception is the GR period, where its RMSEs are also lower. This suggests that switches in level dynamics might help in capturing the behavior of financial markets and/or the monetary policy rule rather than that of output and prices. It might be that the changes in transmission mechanisms associated to financial distress episodes are not powerful enough to be exploited from a forecasting perspective (see Section 4). Whatever the interpretation, this result suggests that in general the value of the threshold model originates mostly from its ability to capture the variance of the densities, that is, to quantify risks around the underlying central forecasts. It is only in exceptional circumstances – in our data, the GR – that this is complemented by a more accurate prediction of the level dynamics.

5.2 Model selection over time

In figure 4 we plot the evolution over time of the root mean square errors and log-scores for the 12-month ahead industrial production forecasts generated by $VAR^{\sigma}, VAR_f^{\sigma}$ and TAR_f.[12] The RMSE ratios (top panel) vary widely over time and their instability makes it hard to identify periods of clear dominance of either model over the benchmark and/or its other competitors. TAR_f is penalized by large errors in the earlier part of the sample, possibly on account of larger fluctuations in the estimated parameters. There are timid signals that the model works better around the 2001 and 2008-2009 recessions, but the improvements are small and disappear or reverse quickly after the recession. The comparison between VAR^{σ} and VAR_f^{σ} shows that the GR is one of the very few periods when omitting FCI carries a cost in terms of accuracy (the RMSE of VAR^{σ}, which does not include the indicator, exceed those of VAR_f^{σ} by roughly 10% on average between 2008 and 2010).

[12] In all cases RMSEs (LS) are calculated as ratios (differences) relative to the linear benchmark VAR_f, as in Table 1. All models become less accurate in absolute terms during recessions (see online appendix for details).

The log-scores (figure 4, bottom panel) are less volatile. Before 2008 VAR^{σ} and VAR_f^{σ} turn out to be virtually identical, confirming that the presence of FCI in the information set makes little difference. TAR_f appears to be very similar to the benchmark VAR_f. Except for a few observations at the end of the 1980s, the Great Moderation decades are indeed entirely classified as 'normal times' so these models have the same structur, though of course not exactly the same posterior. VAR^{σ} and VAR_f^{σ} consistently outperform the benchmark (and hence TAR_f too) as long as output growth is positive, but their relative performance deteriorates in both the 1991 and the 2001 recession. The subsequent outbreak of the GR has three major consequences: a dramatic and persistent drop in the score of VAR^{σ}; a large swing in that of VAR_f^{σ}, which drops in 2008 and then rebounds, leading to an average performance analogous to that of the benchmark; and a rise in the score of TAR_f relative to both the benchmark and VAR_f^{σ}. In short, as suggested by figure 3, the omission of FCI turns out to be a significant liability and the threshold model works better than the stochastic volatility models. Unfortunately these changes were largely unpredictable prior to the beginning of the recession. The log-Bayes factors show that the evidence in favor of the stochastic volatility VARs builds up consistently over time, whereas the evidence in favor of using FCI and/or allowing for a financial threshold when predicting output growth is limited to the GR episode (see online appendix for details). Furthermore, the Giacomini-White selection criteria confirm that a forecasters would have had no reason to abandon VAR_f^{σ} in favor of TAR_f before the recession: both in 2001 and in 2008/9 the log-score of TAR_f for output growth is predicted to exceed that of VAR_f only after the recession has begun. In the case of inflation, VAR_f^{σ} is (rightly) predicted to consistently beat TAR_f. Hence, the case for dropping VAR_f^{σ} would have been even weaker for someone aiming to jointly forecast output and inflation.

5.3 The Great Recession

In the last five years central banks around the world have been grappling with the challenge of designing and setting up new "macroprudential" policy regimes with

the aim of monitoring and possibly mitigating the consequences of financial crises. Predictive densities can in principle be exploited to construct indicators that can be used in this process. In particular, a macroprudential authority could use the distributions to estimate the *ex ante* probability of a tail event, such as a large output loss, and intervene if and when this exceeds a predefined threshold. How useful would our VAR models be in this respect? Figure 5 displays a set of predicted tail probabilities for the Great Recession period. We focus on the probability of a contraction in industrial production of 5% or more over a six-month horizon, and compare the predictions of the two stochastic volatility models with and without financial indicator (VAR^{σ} and VAR^{σ}_{f}) and those of the threshold model (TAR_{f}).[13] All models fail to anticipate the beginning of the slowdown. In the case of VAR^{σ}, the probability remains low and approximately constant for most of the relevant time window, reaching a maximum of around 30% only in 2009. The presence of the financial indicator allows both VAR^{σ}_{f} and TAR_{f} to do relatively better. Among these, TAR_{f} performs better in terms of both timing and likelihood: it assigns a 20% probability to a contraction in output in the first half of 2008 (against 5% for VAR^{σ}_{f}), and this rises to 80% in correspondence with the actual trough of the recession in 2009. The records show that on August 7th 2007 the Federal Reserve Board voted to hold the target for the federal funds rate constant at 5.25 percent. On August 17th the primary credit rate was cut by 50 basis points, to 5.75 percent, noting that *"downside risks to growth [had] increased appreciably"*. Further cuts to both federal funds rate and primary credit rate were decided on September 18th. Given this timeline, the estimates could have been a useful input for the Board, at least in principle. A first *caveat* though is that we do not use real time data. A second and perhaps more important one is that, as of September 2007, the track record of the models gave no indication of the upcoming improvement in the accuracy of TAR_{f} (see Section 5.2). Neither formal statistical criteria nor a qualitative assessment

[13] Formally, the plot shows $P_{t|t-h} \equiv \Pr\left(\Sigma_{i=1}^{h} y_{t-h+i} < c \,|Y_{t-h}\right)$, where $h = 6$ and $c = -5\%$. The $c = -5\%$ threshold is meant to mimic the beginning of a "great" recession. The timing convention is such that $P_{t|t-h}$ indicates the probability of a contraction between $t-h$ and t estimated using information dated $t-h$. This should ideally increase exactly when the recession begins, facilitating the interpretation of the figure.

of its performance in previous recessions suggested that the threshold model would outperform the (generally more accurate) stochastic volatility specification. This suggests that, barring an extremely high degree of risk aversion, the Board would have been likely to pay little attention to the warning signals coming from the TAR.

6 Discussion

Two important concerns raised by our analysis relate to the information set on which the models are based and the possible occurrence of structural changes that have little to do with financial frictions and might consequently confound our model comparisons. We discuss them in turn below, and conclude by examining the possibility of combining forecasts from different models. More details on the models discussed in this section can be found in the online appendix to the paper. Replacing the Financial Condition Indicator with the Excess Bond Premium of Gilchrist and Zakrajsek (2012), which is constructed using exclusively spreads on non-financial corporate bonds, does not alter the key results. In particular, TAR_f generates better predictive densities than VAR_f for all variables including output, particularly during the Great Recession. Since the small size of the models may leave room to spurious nonlinearities, we also consider as a linear benchmark a larger VAR (VAR^{Large}) where the baseline specification is augmented to include employment and unemployment rate, hours worked, money stock, housing starts, the Reuters/Jefferies CRB spot commodity price index and the NAPM-ISM Purchasing Managers Index. VAR^{Large} predicts output growth better than VAR_f in terms of both point and density forecasts, confirming that there are gains from expanding the information set, but it performs generally worse than TAR_f in terms of log-scores. This suggests that neither the choice of the indicator nor the small number of series used in our baseline analyses influence our results in a significant way.

In order to probe the nature of the regimes in the data, we study a Markov-switching VAR with endogenous transition probabilities ($MS\text{-}VAR$). This model includes output, prices and interest rates and, like TAR_f, it assumes a double-regime

structure. However, it does not mechanically link the transitions across regimes to changes in the financial indicator. The transitions are driven instead by a latent variable whose connection with the financial indicator is estimated from the data. The posterior estimates show that the financial indicator (either FCI or EBP) loads very significantly on the latent variable, implying that financial conditions are a key driver of changes in economic regimes. In terms of accuracy, $MS\text{-}VAR$ produces better forecasts than VAR_f but it turns out to be on average more or less accurate than TAR_f depending on whether EBP or FCI is used as a financial indicator, suggesting that using a more flexible specification does not necessarily pay off in terms of forecasting accuracy. To shed more light on the role of volatility we also introduce an additional threshold model where the covariance matrix of the shocks changes but the conditional mean parameters are assumed to be constant across regimes (TAR^{vol}). This model is obtained by simply restricting the coefficients of the TAR_f model in equation (9) and it allows us to isolate the contribution given by switching volatilities to the results documented in Section 5. We find that TAR^{vol} is on average more accurate than TAR_f in terms of root mean square error but less accurate in terms of log-scores. In other words, the restriction embedded in TAR^{vol} is useful for point forecasts (which is consistent with the success of linear VARs on this front) but detrimental for density forecasts.

Since the models tend to work well in different periods and their calibration is less than perfect, it is also natural to ask what the forecaster could gain by combining them in an opinion pool (Hall and Mitchell (2007), Geweke and Amisano (2011)). Furthermore, the pooling weights can provide useful information on the evolution of the economy. Del Negro, Hasegawa, and Schorfheide (2016) follow this route to investigate the role of credit markets, and document that a structural model with financial frictions generates better forecasts than its frictionless counterpart (and is thus given a larger weight in an optimal pool) only during periods of financial distress. We examine two simple pools: the first one (P_1) includes VAR and VAR_f, while the second one (P_2) also includes TAR_f. In each case we consider a naïve weighting scheme that attaches constant equal weights to all models or alternatively

the optimal weighting scheme of Geweke and Amisano (2011), where the weights are updated recursively over time. Naïve pools often outperform the VAR specification but struggle to do better than VAR_f. Optimal pools are more accurate than both VAR and VAR_f, but do not generally beat TAR_f. In the case of P_1, the optimal weight on VAR_f increases in periods of financial distress and it is equal to one in the Great Recession, consistent with Del Negro, Hasegawa, and Schorfheide (2016), but the weights do not change as quickly as the TAR_f regimes. More importantly, linear VARs estimate the size of the shocks and the strength of the transmission mechanisms on average over the entire sample: insofar as either of these changes in a crisis, a pool of linear models may still be dominated by a threshold specification. Indeed, it is only in predicting the depth of the GR that P_1 looses ground relative to TAR_f (the two are virtually indistinguishable until 2008).

7 Conclusions

Is it possible to exploit financial market information to predict the macroeconomic outcomes associated with financial distress episodes, such as the Great Recession of 2008-2009? To answer this question, we estimate a set of linear and nonlinear VAR models using US data for the 1973-2012 period and check to what extent financial variables improve point and density forecasts for inflation and industrial production growth. The analysis delivers three main results. First, stochastic volatility models perform generally better than linear models, confirming that heteroscedasticity is a first-order feature of the data. Second, adding a financial indicator to the information set improves the forecasts for output growth irrespective of whether or not the model includes stochastic volatility. This result is however heavily influenced by the Great Recession: prior to 2008 a forecaster would have had little reason to add financial variables to a small heteroscedastic VAR. Third, the Great Recession is also pivotal in building up evidence on financial threshold effects. The data clearly point to the coexistence of two distinct regimes in US history, where periods of financial distress have been typically associated with the occurrence of larger financial shocks and, to a minor extent, changes in their transmission mechanism. However, the 2008-2009

period is the only one where these forms of non-linearity turn out to be useful for forecasting. At the end of 2007 a Threshold VAR model would have predicted a 20% probability of a sharp output contraction in 2008, compared to only 5% for a stochastic volatility VAR based on the same data. But the rarity of the event, combined with the impossibility to anticipate the forthcoming improvement in the TAR's forecasting performance, implies that it would have been very hard for policy makers to act upon this signal.

Acknowledgement 1 *The authors would like to acknowledge helpful feedback received from Gianni Amisano, Valentina Aprigliano, Marco Del Negro, Gianni De Nicolò, James Mitchell, Harald Uhlig, Shaun Vahey, two anonymous referees and an associate editor of the Review of Economic Dynamics, as well as seminar participants at Banca d'Italia, the Einaudi Institute for Economics and Finance, the European Central Bank, the International Monetary Fund and Queen Mary University of London. The paper was written while Piergiorgio Alessandri was hosted as a Research Fellow by Banca d'Italia. The views expressed in this paper are those of the authors and not necessarily those of the Banca d'Italia.*

References

Balke, N. S. (2000): "Credit and economic activity: credit regimes and nonlinear propagation of shocks," *The Review of Economics and Statistics*, 82(2), 344–9.

Bańbura, M., D. Giannone, and L. Reichlin (2010): "Large Bayesian vector auto regressions," *Journal of Applied Econometrics*, 25(1), 71–92.

Bianchi, J. (2011): "Overborrowing and systemic externalities in the business cycle," *The American Economic Review*, 101(7), 3400–26.

Bianchi, J., and E. G. Mendoza (2010): "Overborrowing, financial crises and macroprudential taxes," NBER Working Paper n.16091.

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

Brunnermeier, M. K., and Y. Sannikov (2014): "A macroeconomic model with a financial sector," *The American Economic Review*, 104(2), 379–421.

Carriero, A., T. E. Clark, and M. Marcellino (2015): "Bayesian VARs: specification choices and forecast accuracy," *Journal of Applied Econometrics*, 30(1), 46–73.

Chen, C. W., and J. C. Lee (1995): "Bayesian inference of threshold autoregressive models," *Journal of Time Series Analysis*, 16(5), 483–492.

Clark, T. E. (2011): "Real-time density forecasts from Bayesian Vector Autoregressions with stochastic volatility," *Journal of Business and Economic Statistics*, 29(3), 327–41.

Clements, M. P., and J. Smith (2000): "Evaluating the forecast densities of linear and non-linear models: applications to output growth and unemployment," *Journal of Forecasting*, 19, 255–76.

Cogley, T., S. Morozov, and T. Sargent (2005): "Bayesian fan charts for UK inflation: forecasting and sources of uncertainty in an evolving monetary system," *Journal of Economic Dynamics and Control*, 29(11), 1891–1925.

Cogley, T., and T. J. Sargent (2005): "Monetary Policy and Learning Drifts and volatilities: monetary policies and outcomes in the post WWII US," *Review of Economic Dynamics*, 8(2), 262 – 302.

De Nicolo, G., and M. Lucchetta (2012): "Systemic real and financial risks: measurement, forecasting, and stress testing," IMF Working Paper 12/58.

Del Negro, M., R. B. Hasegawa, and F. Schorfheide (2016): "Dynamic prediction pools: an investigation of financial frictions and forecasting performance," *Journal of Econometrics*, 192(2), 391–405.

Geweke, J., and G. Amisano (2010): "Comparing and evaluating Bayesian predictive distributions of asset returns," *International Journal of Forecasting*, 26, 216–30.

——— (2011): "Optimal prediction pools," *Journal of Econometrics*, 164(1), 130–141.

Giacomini, R., and H. White (2006): "Tests of conditional predictive ability," *Econometrica*, 74(6), 1545–78.

Gilchrist, S., and E. Zakrajsek (2012): "Credit spreads and business cycle fluctuations," *The American Economic Review*, 102(4), 1692–720.

Guerrieri, L., and M. Iacoviello (2013): "Collateral constraints and macroeconomic asymmetries," International Finance Discussion Paper.

Hall, S. G., and J. Mitchell (2007): "Combining density forecasts," *International Journal of Forecasting*, 23(1), 1–13.

He, Z., and A. Krishnamurthy (2011): "A model of capital and crises," *The Review of Economic Studies*, p. rdr036.

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

Jacquier, E., N. G. Polson, and P. E. Rossi (1994): "Bayesian Analysis of Stochastic Volatility Models," .

Jermann, U., and V. Quadrini (2012): "Macroeconomic effects of financial shocks," *The American Economic Review*, 102(1), 238–71.

Jore, A. S., J. Mitchell, and S. P. Vahey (2010): "Combining density forecasts from vars with uncertain instabilities," *Journal of Applied Econometrics*, 25, 621–34.

Justiniano, A., and G. E. Primiceri (2008): "The time-varying volatility of macroeconomic fluctuations," *The American Economic Review*, 98(3), 604–641.

Koop, G., M. H. Pesaran, and S. M. Potter (1996): "Impulse response analysis in nonlinear multivariate models," *Journal of econometrics*, 74(1), 119–147.

Li, S. M., and S. Dressler (2011): "Business cycle asymmetry via occasionally binding international borrowing constraints," *Journal of Macroeconomics*, 33, 33–41.

Liu, Z., P. Wang, and T. Zha (2013): "Land-price dynamics and macroeconomic fluctuations," *Econometrica*, 81(3), 1147–84.

McCallum, J. (1991): "Credit rationing and the monetary transmission mechanism," *The American Economic Review*, 81(4), 946–51.

Mitchell, J., and K. F. Wallis (2011): "Evaluating density forecasts: forecast combinations, model mixtures, calibration and sharpness," *Journal of Applied Econometrics*, 26, 1023–40.

Ng, S., and J. H. Wright (2013): "Facts and challenges from the Great Recession for forecasting and macroeconomic modeling," *Journal of Economic Literature*, 51(4), 1120–54.

Sims, C. A. (2012): "Disentangling the channels of the 2007-2009 recession: a comment," Brookings Papers on Economic Activity.

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

Sims, C. A., and T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy?," *American Economic Review*, 96(1), 54–81.

Stock, J. H., and M. H. Watson (2003): "Forecasting output and inflation: the role of asset prices," *Journal of Economic Literature*, 41(3), 788–829.

Stock, J. H., and M. W. Watson (2012): "Disentangling the channels of the 2007-2009 recession," NBER Working Paper n.18094.

		RMSE				LS			
		1M	3M	6M	12M	1M	3M	6M	12M
VAR_f	y	*0.454*	*0.514*	*0.547*	*0.576*	*-3.553*	*-3.156*	*-3.032*	*-2.964*
	r	*0.177*	*0.365*	*0.602*	*0.989*	*-0.645*	*-1.357*	*-1.723*	*-2.101*
	π	*0.172*	*0.218*	*0.237*	*0.260*	*-2.583*	*-2.550*	*-2.339*	*-2.171*
	f	*0.102*	*0.198*	*0.289*	*0.386*	*0.135*	*-0.649*	*-0.957*	*-1.130*
VAR	y	1.029	1.048	1.038	1.015	-0.121	-0.182	-0.386	-0.984
	r	0.944	0.978	0.993	0.996	-0.030	-0.023	-0.031	-0.017
	π	1.005	0.995	0.990	0.988	0.000	-0.108	0.073	0.034
	f	–	–	–	–	–	–	–	–
VAR_σ	y	0.987	1.011	1.009	0.997	-0.004	-0.742	-0.937	-1.167
	r	0.825	0.792	0.802	0.856	0.956	0.702	0.392	0.122
	π	0.974	0.958	0.945	0.933	0.238	0.023	-0.295	-0.466
	f	–	–	–	–	–	–	–	–
VAR_f^σ	y	0.950	0.986	0.998	0.996	-0.307	0.202	0.107	0.199
	r	0.808	0.789	0.804	0.899	0.899	0.722	0.405	0.085
	π	0.970	0.956	0.946	0.938	0.237	0.371	0.285	0.171
	f	0.971	0.899	0.872	0.907	0.601	0.449	0.485	0.320
TAR_f	y	1.008	1.003	1.005	1.003	0.062	0.004	0.027	0.079
	r	0.944	0.926	0.922	0.953	0.667	0.579	0.359	0.102
	π	1.023	1.018	1.009	1.000	0.081	0.135	0.144	0.091
	f	1.020	0.960	0.938	0.951	0.361	0.527	0.526	0.413
TAR_f^σ	y	0.948	0.971	0.983	0.988	-0.101	-0.448	-0.689	-0.562
	r	0.825	0.784	0.786	0.797	0.696	0.772	0.480	0.270
	π	0.977	0.966	0.954	0.926	0.210	0.036	-0.239	-0.477
	f	1.000	0.894	0.834	0.832	0.517	0.565	0.609	0.470

Table 1: Forecast evaluation. The table shows the average root mean square errors (RMSE) and predictive log-scores (LS) generated by six alternative forecasting models over the period April 1983 – August 2012. The variables are US industrial production growth (y), federal funds rate (r), consumer price inflation (π) and the Financial Condition Index (f). For the benchmark model VAR_f, a linear VAR in the four variables, the statistics are reported in levels (top panel). For all remaining models, RMSEs (LS) are reported as ratios (differences) relative to the benchmark. VAR is a three-variable linear VAR without f_t. VAR_f^σ and VAR^σ are two stochastic volatility VAR models with and without f_t. TAR_f is a threshold VAR model where the regime is determined by f_t. TAR_f^σ is an analogous threshold model with stochastic volatility.

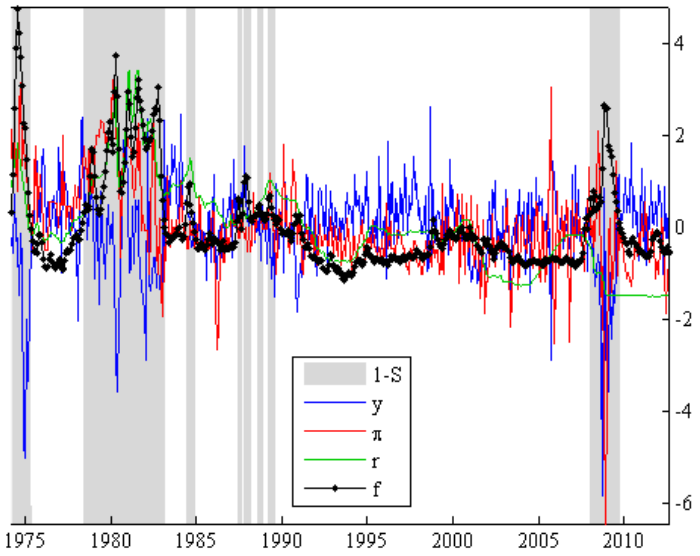

Figure 1: Financial regimes in the USA. Gray bands identify periods when the US economy is estimated to be in financial distress by the TAR described in equations (9)-(10).The lines represent the monthly series used to estimate the model, namely industrial production growth (y), consumer price inflation (π), the federal funds rate $(r)$and the Chicago Fed Financial Condition Index (f).

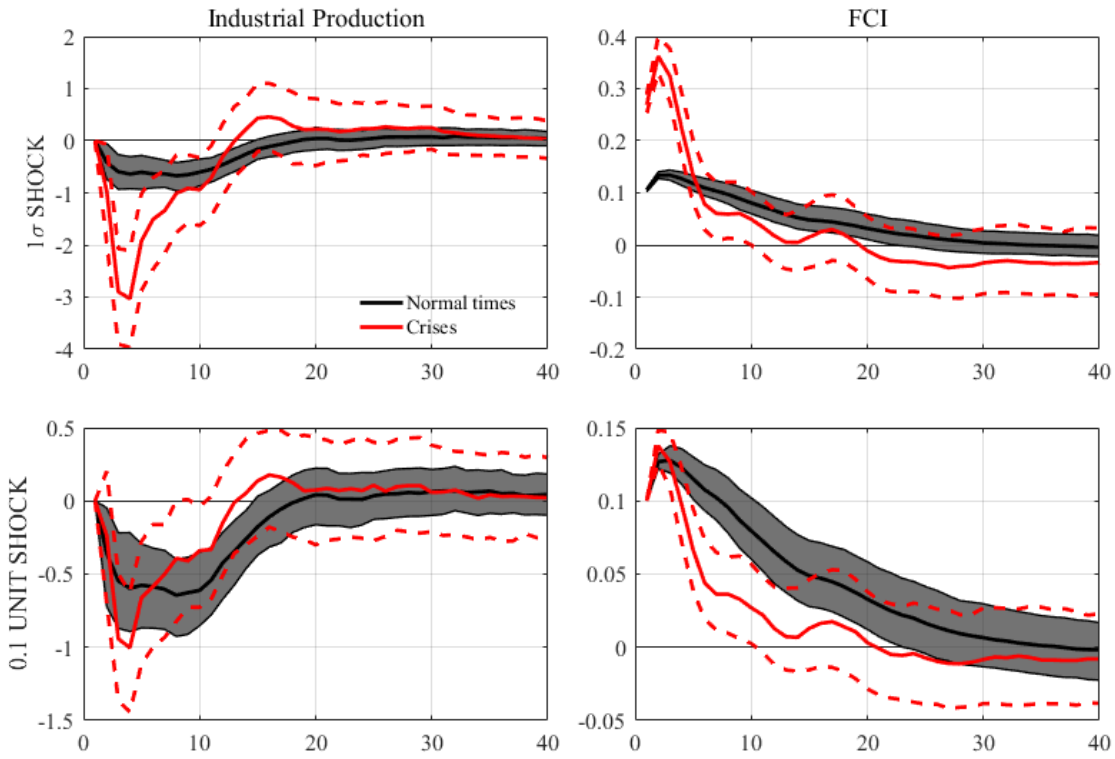

Figure 2: Impact of financial shocks in good and bad times. The figure shows the responses of industrial production (column one) and the Chicago Fed Financial Condition Index (column two) to an adverse financial shock. The responses are estimated using the TAR model of equations (9)-(10) and a recursive identification scheme where FCI is ordered last, and they are simulated conditioning separately on normal times (black) and financial distress episodes (red). For each regime the figure reports the median responses with a 68% confidence band. In the top row the shock is defined as a one standard deviation increase in FCI and its size varies across regimes (see top right panel). In the bottom row it is normalized to a 0.1 units increase in FCI and held constant across regimes.

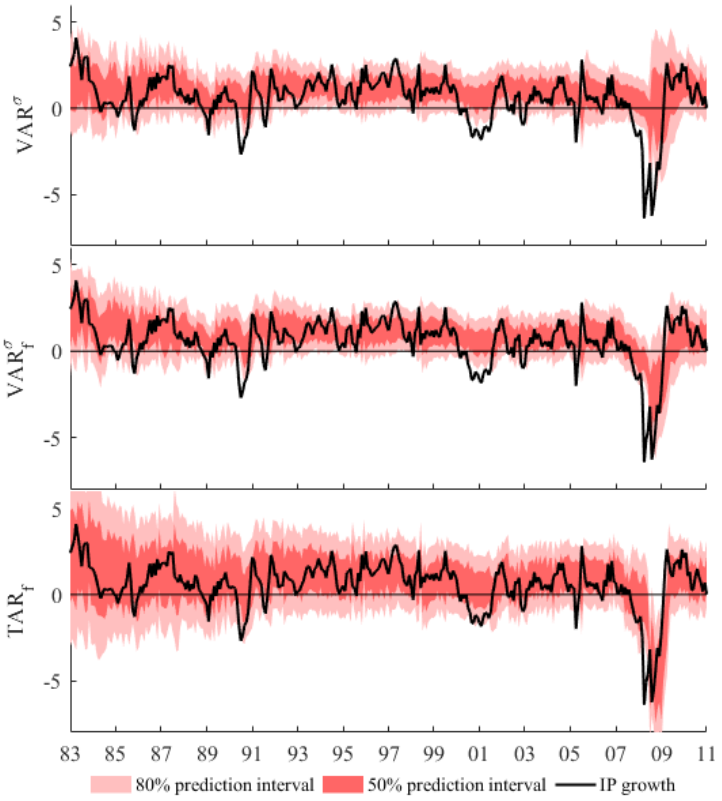

Figure 3: prediction intervals. The panels show the three-month ahead prediction intervals for industrial production growth generated by three alternative forecasting models. All models are estimated recursively using monthly data on industrial production, consumer price inflation, the nominal federal funds rate and the Chicago Fed Financial Condition Index (FCI). VAR_f^{σ} and VAR^{σ} are two stochastic volatility VAR models that, respectively, do and do not include the FCI. $TAR_f$is a threshold VAR model where the switch across regimes is determined by FCI. The sample begins in March 1973 and the forecasts are calculated starting from April 1983.

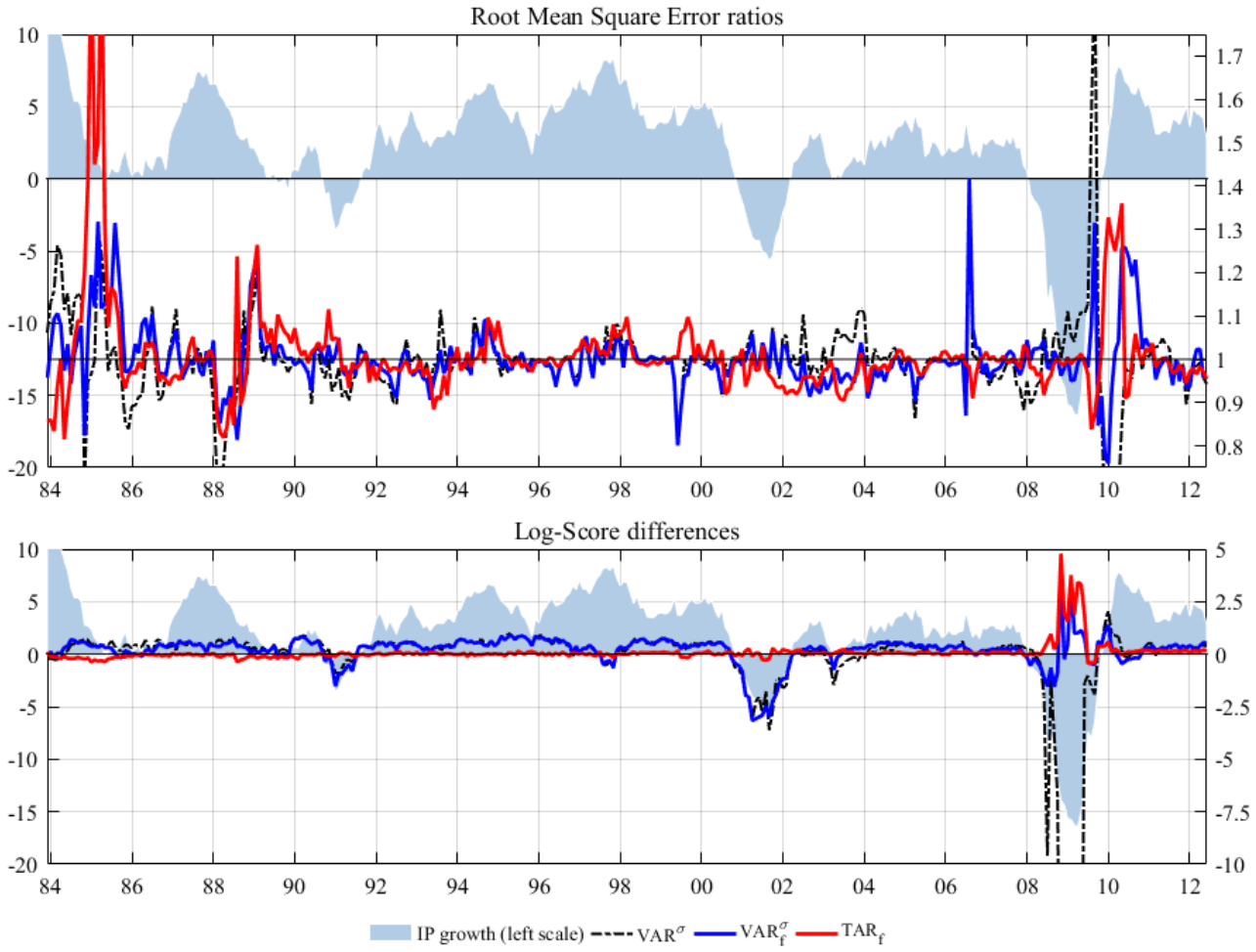

Figure 4: root mean square error and log-score patterns. The top panel shows the RMSE associated to the 12-month ahead industrial production growth forecasts generated by three alternative models. VAR^{σ} is a stochastic volatility VAR model estimated using data on industrial production, inflation and the nominal federal funds rate.. VAR^{σ}_{f} is an analogous model that also includes the Chicago Fed Financial Condition Index. TAR_{f} is a threshold VAR model based on the same data, where FCI determines the regime. All RMSEs are expressed as ratios relative to a linear benchmark VAR. The top panel shows the log-scores associated to the same set of predictions. These are displayed as differences relative to the benchmark model.

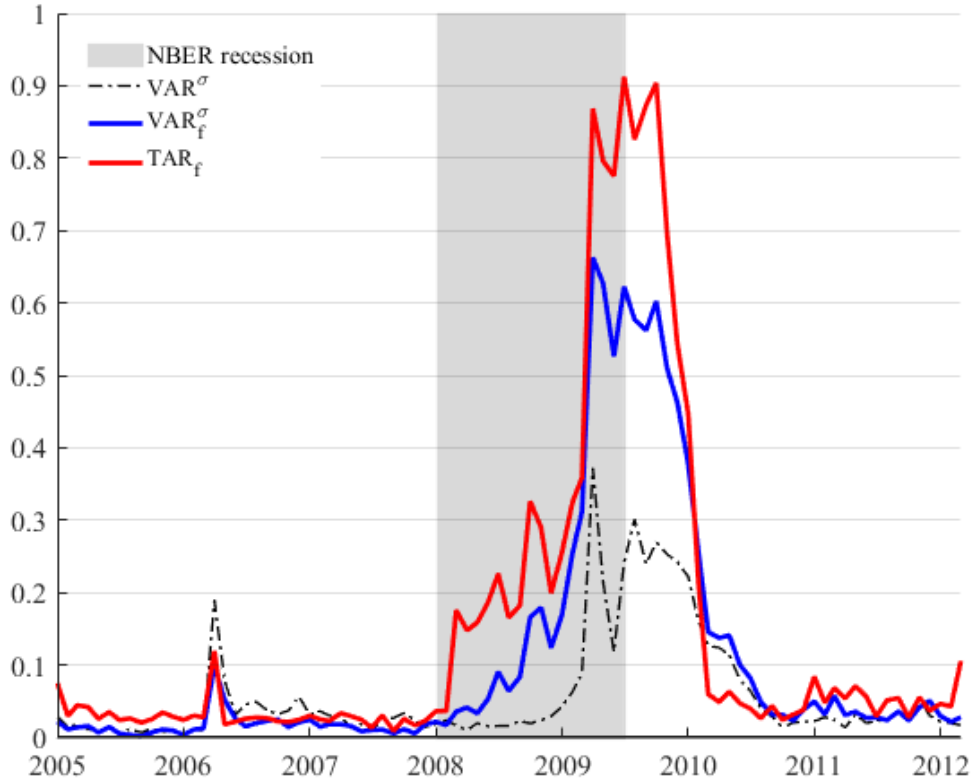

Figure 5: Recession probabilities. Model-implied probabilities of observing a cumulative contraction in industrial production of 5% or more over a six-month window ending at the date shown on the horizontal axis. $VAR^{\sigma}$is a stochastic volatility VAR that includes industrial production growth, inflation and the nominal federal funds rate. VAR^{σ}_{f}also includes the Financial Condition Index. TAR_f is a regime-switching model that includes the same four series and allows for a regime change when the Financial Condition Index exceeds a critical threshold.

Supplementary Material for on-line publication only

Click here to download Supplementary Material for on-line publication only: Technical_Appendix_aug_29.pdf

LaTeX Source Files

Click here to download LaTeX Source Files: LaTex_source_files.zip