Bocconi

Università Commerciale Luigi Bocconi

BAFFI CAREFIN
Centre for Applied Research on International Markets, Banking, Finance and Regulation

Working Paper

Angela Abbate and Massimiliano Marcellino

Macroeconomic activity and risk indicators: an unstable relationship

BAFFI CAREFIN Centre Research Paper Series No. 2017- 56

Macroeconomic activity and risk indicators: an unstable relationship *

Angela Abbate[†] Massimiliano Marcellino[‡]

Abstract

We assess to what extent indicators of financial conditions can be considered relevant determinants and predictors of macroeconomic aggregates. The main finding is that controlling for default risk and risk aversion measures improves the forecasts of output, employment and loans, but that this improvement is largely attributable to the recession periods of 2001 and 2008. A structural VAR analysis further reveals that financial condition indicators display significant real effects only after the Great Financial Crisis. In particular, an unexpected increase in the credit spread in 2010 causes an output contraction that lasts for about two years, with an annualised through of 4.8%, and explains up to 35% of the forecast error variance of industrial production.

J.E.L. Classification: C53, E02, E27, E37

Keywords: forecasting, credit spreads, SVAR, time-varying parameters

*We would like to thank the Baffi-Carefin center at Bocconi University for partial financial support. The views expressed here are those of the authors and do not represent those of the Deutsche Bundesbank. We would like to thank Dimitris Korobilis, colleagues at the Bundesbank and participants to the "New trends and developments in Econometrics" conference at the Bank of Portugal for useful comments on an earlier draft.

[†]Deutsche Bundesbank angela.abbate@bundesbank.de

[‡]Bocconi University, IGIER and CEPR, massimiliano.marcellino@unibocconi.it

1 Introduction

A vast range of theoretical models suggests that measures of default risk and of risk aversion should contain predictive content for macroeconomic aggregates. Variations in default risk, or in the risk propensity of intermediaries, cause in fact movements in the cost of external finance, proxied by credit spreads. This in turn may induce fluctuations in borrowing and output, according for instance to the financial-accelerator theory of Bernanke and Gertler [1989] and Bernanke et al. [1999]. In addition, Gilchrist et al. [2015] have built a model to rationalise the empirical finding that the pricing behaviour of firms depends on the extent of their reliance on external finance. Several empirical works have provided evidence to support these theoretical models, such as Gilchrist and Zakrajšek [2012], Faust et al. [2013], and De Pace and Weber [2016]. We revisit these findings using a time-varying methodology and a broad set of statistical evaluation criteria, and show that the predictive content of default risk and of risk aversion measures for key US macroeconomic quantities is limited to the recession periods of 2001 and 2008.

Our goal is to establish the marginal predictive content of three different risk indicators for US output, inflation and lending activity measures. The three risk indicators considered are two credit spreads commonly used in the empirical literature, and one measure of risk aversion. The two credit spreads are the difference between the Baa and Aaa-rated bond yields, and the measure recently developed by Gilchrist and Zakrajšek [2012] using prices of individual corporate bonds traded in the secondary market. Both measures capture variations in default risk only with a noise, as they are affected by other factors such as risk aversion and, to a lesser extent, liquidity premia and bond taxation conditions. In order to isolate the predictive content of the two main risk components, we consider a third predictor that captures precisely risk aversion: the excess bond premium, also developed by Gilchrist and Zakrajšek [2012]. To assess time variation in the predictive power of these indicators we employ a series of time-varying forgetting factor models recently proposed by Koop and Korobilis [2013]. These models, which can be considered a classical approximation of the time-varying Bayesian

VARs of Cogley and Sargent [2005] and Primiceri [2005], are considerably faster to estimate and, since they are not as densely parametrised, they have been often found to be good forecasting tools.[1]

Several results emerge from the forecasting analysis. Controlling for the credit spread developed by Gilchrist and Zakrajšek [2012] improves the point and density forecasts of both industrial production and employment after 2000. This result complements the findings in Gilchrist and Zakrajšek [2012] by showing that the enhanced predictive performance is limited to the recession periods of 2001 and 2008, when credit constraints were arguably binding. An additional result is that both default risk and risk aversion matter for improving the forecasts of the two activity indicators.

By contrast, financial indicators have no predictive content for the two inflation measures considered, with the marginal exception of personal consumption expenditure inflation. Financial indicators display instead predictive power only for loans (i.e. quantity of credit), while not for the loan interest rate spread (i.e. price of credit). To obtain accurate forecasts for the loan interest rate spread it is in fact sufficient to control for past values of the spread, as well as for the term structure and the real risk-free rate.

The conclusion on the time-varying leading properties of financial indicators for industrial production is strengthened by a structural VAR analysis which shows that credit spread shocks display significant real effects only after the latest recession. In particular, an unexpected increase in the credit spread causes in 2010 a contraction that lasts for about two years, with an annualised through of 4.8%, and explains up to 35% of the forecast error variance of industrial production. Moreover, both increased default risk and higher risk aversion lie behind the latest US recession and subsequent slow recovery.

There is a vast empirical literature that seeks to assess to what extent macroeconomic activity can be forecasted by credit spreads. Gilchrist and Zakrajšek [2012] find that their newly developed credit spread measure has a significantly higher predictive

[1] See for instance Koop and Korobilis [2013] and Abbate and Marcellino [2016].

power for macroeconomic activity indicators than other measures such as the Baa-Aaa corporate bond spread, or the paper-bill spread. They moreover show that most of the predictive content of their spread measure can be attributed to its risk aversion component (the excess bond premium). Using a different methodology we show that, once time variation is taken into account, the predictive power of credits spreads is limited to the 2001 and 2008 recession periods, and that the excess bond premium does not exhaust the informational content of the credit spread measure. A similar conclusion is reached by Faust et al. [2013], without however establishing the relative role of risk aversion versus default risk. In particular, the authors forecast macroeconomic activity[2] with a Bayesian model averaging (BMA) scheme that includes, alongside with a large set of macroeconomic and financial predictors, 20 credit spreads constructed from a confidential database of corporate fixed income securities, and differentiated by risk and maturity categories. The BMA models are found to deliver lower mean squared forecast errors than an autoregressive model, particularly during NBER-dated recessions. Disentangling their results, the authors find that the highest BMA weights are assigned to credit spreads, and it is their inclusion that significantly improves point forecasts relative to the autoregressive benchmark.

Time variation in the leading properties of credit spreads for the growth rates of US GDP and industrial production is examined in De Pace and Weber [2016]. The authors find that the leading properties of the two high-yield spreads analysed[3] decline after 2000 and disappear around the 2007-2009 downturn, which they attribute to the improved forecasting ability of the term spread. We find a contrasting result, confirmed by the structural analysis, by relying on a different modelling strategy, that controls for the term spread itself, as well as for other variables. Lastly, Alessandri et al. [2016] assess potential nonlinearities in the relationship between credit spreads and economic activity in the US and find through reduced-form predictive regressions that sharp

[2]The macroeconomic activity considered are the growth rates in real GDP, personal consumption expenditures, business fixed investment, industrial production, private payroll employment, civilian unemployment rate, exports and imports, over the period 1986:Q1- 2011:Q3.

[3]The two high-yield spreads considered are the differences between an interest-rate index associated with below investment grade corporate bonds and either the 10-year government bond yield, or the AAA-rated corporate bond yield.

increases in corporate bond spreads anticipate slowdowns in economic activity.

The paper is organised as follows. The choice of indicators and the VAR methodology are discussed in Section 2. The in-sample fit of the competing models is presented in Section 3, while the out-of-sample evaluation is carried out in Section 4. Sections 5 performs a structural VAR analysis that establishes whether the macroeconomic effect of shocks to default risk and risk aversion measures has changed over time. Section 6 concludes.

2 Macroeconomic activity prediction models

The empirical strategy followed in this work has two main dimensions. First, we would like to investigate what is the marginal contribution of financial condition indicators in forecasting key macroeconomic variables over a large evaluation period that spans several US recessions and goes from $1980m1$ to $2012m12$. We consider three indicators: two measures of default risk, and one measure of risk aversion. Default risk is captured, noisily, by two credit spread measures: the difference between Baa and Aaa-rated bond yields (henceforth BA spread), and the credit measure developed by Gilchrist and Zakrajšek [2012] (henceforth GZ spread). The latter is constructed by averaging credit spreads computed for twenty different maturity and credit risk categories from a confidential database of US firms, and is found to outperform the BA spread in forecasting output and employment. Gilchrist and Zakrajšek [2012] further decompose their spread measures in two parts: one capturing the countercyclical movements in expected default and another, the excess bond premium (henceforth EBP), that reflects deviations between the expected default risk of borrowers and the price of their debt claims. Gilchrist and Zakrajšek [2012] argue that this latter component signals the risk aversion of the financial sector, as it spikes during US recessions and is highly correlated with the leverage of highly leveraged (and high risk) agents (broker-dealers).

Standard financial accelerator theories such as in Bernanke et al. [1999], as well as recently proposed models such as Gilchrist et al. [2015] show that a shock that increases the default risk of borrowers or diminishes the risk-bearing capacity of financial intermediaries will be followed by a sharp contraction in lending and in output, and possibly by either an increase or by only a moderate decline in prices.[4] Based on these models we look at the predictive content of these three financial indicators for key US macroeconomic aggregates: the month-on-month change in the industrial manufacturing index and in civilian employment, producer price and personal consumption expenditure inflation. We further look at the predictive content for two banking sector variables: the month-on-month change in commercial and industrial loans and the loan interest rate spread measured as the difference between the bank prime loan rate and the federal funds rate. A summary of all the variables can be found in the lower panel of Table 1 below.

As a second goal, we would like to know whether the predictive power of the two credit spreads and of risk aversion has changed over time. In particular, we would like to assess whether default risk or risk aversion matter more in times of economic recessions, and especially during the latest financial crisis. To do so, we rely on a methodology recently proposed by Koop and Korobilis [2013]: the time-varying forgetting factor VAR. The latter can be considered a classical approximation of the time-varying parameter BVAR with stochastic volatility of Cogley and Sargent [2005] and Primiceri [2005], that exploits one-step ahead prediction errors to determine the degree of variation. This model, outlined in detail below, is estimated via the Kalman filter and it is parsimoniously parametrised. Hence, despite not taking sampling uncertainty into account, it has been shown to be a very useful forecasting tool.

To assess the marginal forecasting content of the three financial condition indicators, we first define a baseline time-varying VAR model that controls for lagged values of the macro variable being forecasted, as well as for the real interest rate (federal funds rate deflated through CPI) and for the slope of the Treasury yield curve, i.e. the difference between the three-month and the ten-year constant-maturity Treasury yields.

[4] In their model Gilchrist et al. [2015] show that an adverse financial shock is followed by an output contraction and an overall increase in inflation, as financially constrained firms are forced to raise prices.

These are the same control variables used in the forecasting exercise of Gilchrist and Zakrajšek [2012]. Moreover, the slope of the Treasury yield curve is typically found to be a good predictor for macroeconomic activity.[5] Note that we rely on a VAR rather than on a regression analysis such as in Gilchrist and Zakrajšek [2012] as we believe that there are important dynamic relationships across the variables that should be taken into account. We then augment the benchmark VAR with one indicator at a time and assess how and to what extent the in sample and out of sample fit of the model changes. In particular, we evaluate the point and density forecasts of the indicator-based models relative to the baseline, and to an autoregressive model estimated recursively with the same lag length of the VARs. A description of all forecast models used is found in Table 1 below.

The spreads, the term structure slope and real interest rate measures enter the VAR in levels. Models are estimated with a lag length of 6. The VARs are initialised by shrinking the model to random walk or to a white noise process, depending on the variable transformation, while prior variance is initialised through a training sample. The estimation period is $1975m1 - 1979m12$ while the forecast period is $1980m1 - 2012m12$. After $1980m1$ the estimation sample is progressively enlarged, one month at a time, and forecasts are computed up to one year ahead.

2.1 The time-varying parameter forgetting factor VAR

The time-varying parameter BVAR with stochastic volatility developed by the seminal works of Cogley and Sargent [2005] and Primiceri [2005] has become an important econometric tool in recent years. It is estimated via Bayesian methods with a 5-step Gibbs-sampler and requires a high computational time, especially to draw the covariance matrix of the innovations to the slope parameters, and the covariance matrix of the structural shocks. To reduce computational time, Koop and Korobilis [2013] have developed a procedure which approximates the Bayesian model by replacing the posterior draws of these two covariance matrices with empirical estimates. In

[5] For recent evidence see for instance Chinn and Kucko [2010].

Table 1: **Macroeconomic activity prediction models**: All VAR models control for past lags of the activity variable analysed, as well as for the real interest rate (R) (measured as the federal funds rate deflated through CPI), and for the difference between the three-month and the ten-year constant-maturity Treasury yields, i.e. the slope of the Treasury yield curve (TS). TVP-VAR models are estimated using the forgetting factor methodology described in Section 2.1, with a lag length of 6. The estimation period is $1975m1 - 1979m12$ while the forecast period is $1980m1 - 2012m12$. All variables are retrieved from the FRED database, with the exception of the GZ spread and the EBP, which are relieved from Simon Gilchrist's website.

#	MODELS	VARIABLES
0	autoregressive	activity variable
1	baseline TVP-VAR	activity variable, TS and R
2	TVP-VAR GZ SPREAD	activity variable, TS, R and GZ spread
3	TVP-VAR BA SPREAD	activity variable, TS, R and BA spread
4	TVP-VAR EBP	activity variable, TS, R and EBP

ALTERNATIVE ACTIVITY VARIABLES CONSIDERED	TRANSFORMATION	DESCRIPTION
INDUSTRIAL PRODUCTION (IP)	growth rates (period on period)	industrial manufacturing index
EMPLOYMENT (N)	growth rates (period on period)	index
PRODUCER PRICES (PPI)	growth rates (period on period)	index
PERSONAL CONSUMPTION EXPENDITURE (PCEPI)	growth rates (period on period)	index
LOANS (L)	growth rates (period on period)	deflated with CPI deflator
LOAN INTEREST RATE SPREAD (LS)	levels	bank prime loan rate - federal funds rate

particular, their time-varying parameter forgetting factor VAR is based on a state-space representation, just like its Bayesian counterpart. It is assumed that the $n \times 1$ vector y_t of observed variables is expressed as:

$$y_t = Z_t \beta_t + u_t \quad \text{with } E_t[u_t u_t'] = \Omega_t \ , \tag{1}$$

where Z_t is a $n \times k$ matrix of regressors, and u_t is a $n \times 1$ vector of innovations with covariance matrix Σ_t. Z_t contains a constant and p lags of each variable and is thus defined as $Z_t = I_n \otimes [1, y_{t-1}', \ldots y_{t-p}']$. The measurement equation (1) is complemented by the transition equation for the vector of time-varying slope coefficients β_t:

$$\beta_t = \beta_{t-1} + v_t \ , \tag{2}$$

implying that, given information up to $t-1$, the slope coefficients in t are draws from a normal distribution:

$$\beta_t \,|\, \mathcal{F}_{t-1} \sim N(\beta_{t|t-1},\, P_{t|t-1}) \ . \tag{3}$$

In the prediction step of the Kalman filter routine, the following approximation is used:

$$P_{t|t-1} = \frac{1}{\lambda} P_{t-1|t-1}, \quad \text{with } \lambda \in (0,\, 1]\,, \tag{4}$$

where the parameter λ is a forgetting factor which discounts past information. Throughout this work, a value of λ equal to 0.99 is used, implying, in the case of monthly data, that observations one year ago receive 89% as much weight as current observations.

The second modification to the otherwise standard Kalman filter routine pertains to the estimator of the covariance matrix of the non-structural innovations, Ω_t. The latter is a weighted average of its past value, and of its current estimate:[6]

$$\hat{\Omega}_t = \kappa \hat{\Omega}_{t-1} + (1-\kappa)\hat{u}_t \hat{u}_t'\,, \tag{5}$$

where $\hat{u}_t$ are the Kalman filter errors and the weight is represented by the decay factor κ, set in this application to 0.96 as in Koop and Korobilis [2013].

To summarise, the procedure developed by Koop and Korobilis [2013] is based on a modified version of the Kalman filter which relies on the parametrisation of equations 4 and 5, as well as on the choice of initial conditions for the covariance matrix Ω_0, for the slope coefficients β_0 and their variance P_0. Differently from its Bayesian counterpart, it does not disentangle the structural shocks and delivers filtered estimates, which should be better suited for forecasting purposes. More importantly, the model does not posit a proper law of motion for the covariance matrices of the slope parameters and of the non-structural errors. To circumvent this problem, the parameters are assumed to be fixed out of sample, when sampling from the predictive density. Overall, the time-varying forgetting factor model disregards sampling uncertainty but it is very parsimoniously parametrised, which could potentially yield efficiency gains big enough to compensate the bias induced by the approximations.

[6] This is the Exponentially Weighted Moving Average estimator, commonly used in the finance literature.

3 In-sample fit of financial condition indicators

As a preliminary step, we assess the marginal contribution of the three financial condition measures in improving the in-sample fit of the target variables. We plot in Figure 1 the in-sample log likelihood of the indicator-based models, relative to that of the baseline TVP-VAR model. Increases in the plotted statistics denote observations whose fit is improved by the inclusion of the risk indicator.

The three indicators of financial conditions do not seem to improve the in-sample fit of the baseline VARs in the 1980s and 1990s. Relevant exceptions are personal consumption expenditure (Figure 1d) and employment (Figure 1b). Particularly for the latter, the inclusion of the BA corporate bond spread marginally improves the in-sample likelihood between 1980 and 2000, though only to a small extent.

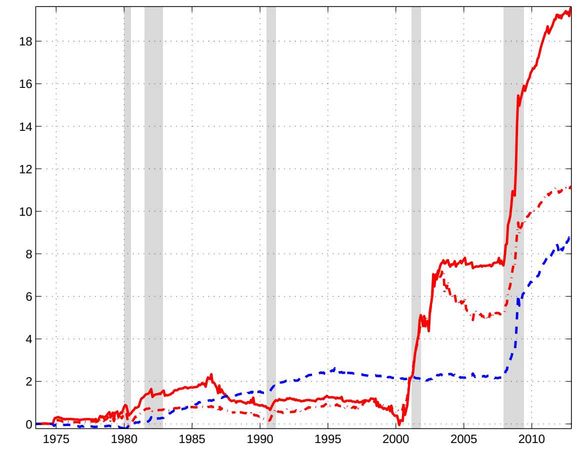
During the dot-com crisis in 2001 the pattern changes and controlling for the GZ spread and for the EBP improves the in-sample log likelihood of output, employment, producer price inflation and loans. Since the in-sample fit of the BA model does not likewise increase, it is possible to conclude that risk aversion appears to have been the main contributor of the better in sample fit of these macro variables during the 2001 crisis.

The great financial crisis in 2008 alludes to a different story. All indicator-based models improve the in-sample fit of the target variables, but this time the bulk of the improvement can be attributed to countercyclical variations in default premia. Exceptions are producer prices (Figure 1c) and loans (Figure 1e), for which risk aversion seems to have played a much bigger role.

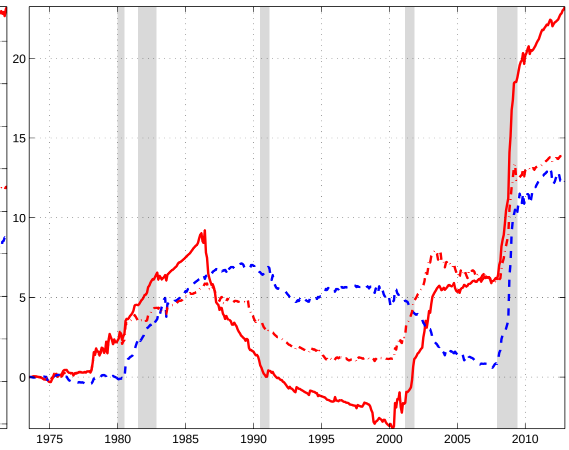
Since the in-sample and out-of-sample fit are not necessarily related, we now turn to discussing the out-of-sample forecasting performance of the three indicator-based models.

Figure 1: **In-sample fit of the indicator-based TVP-VAR models, relative to the baseline TVP-VAR**: Indicators are the GZ spread (solid red line), the BA spread (dashed blue line) and the EBP (dashed dotted red line).

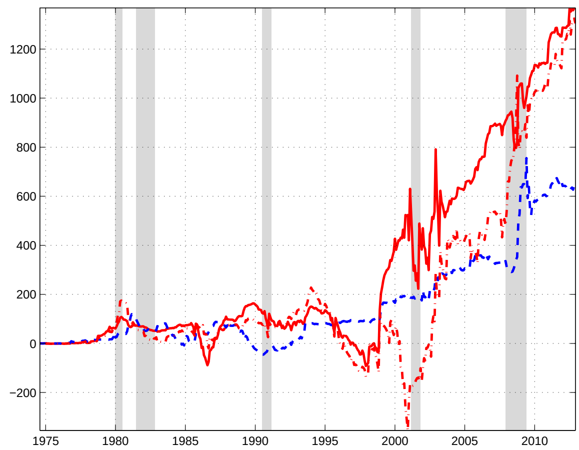
(a) **Industrial production**



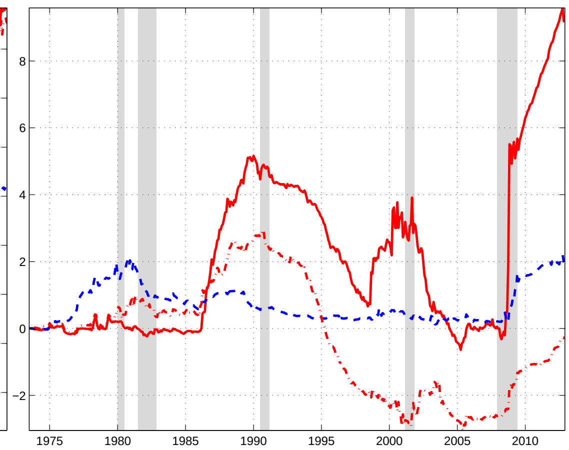
(b) **Employment**



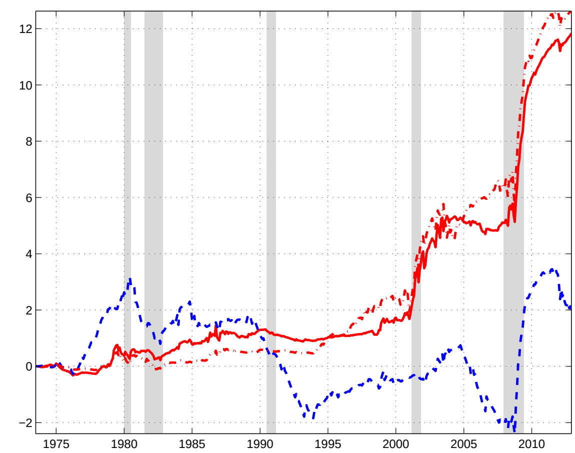
(c) **Producer prices**



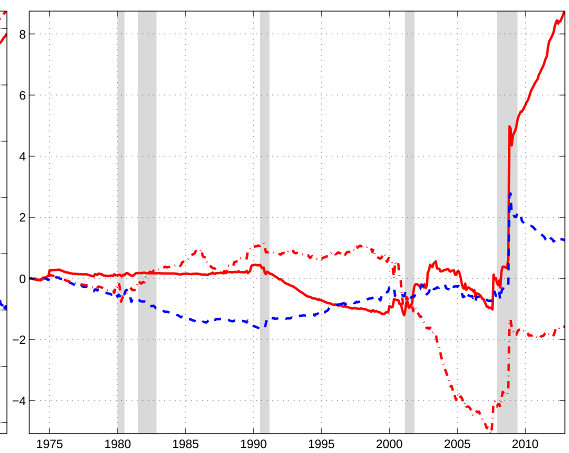
(d) **Personal consumption expenditure**



(e) **Loans**



(f) **Loan interest rate spread**



Note: Positive values denote that, on average, the indicator-based model provides a better in-sample fit. Increases in the statistics denote observations in which the fit of the spread-based model model is higher. Shaded areas denote NBER-dated recessions.

4 Forecasting performance

4.1 Point forecasts

The first step of the out-of-sample analysis is to verify to what extent the addition of default risk and risk aversion indicators improves the point forecasts of the target variables relative to two benchmarks: the baseline TVP-VAR model and an autoregressive model estimated recursively. To assess if the relative forecasting ability of the VAR models has changed during the great financial crisis period, we divide the forecast sample into two: one that excludes the crisis period, ranging from 1980m1 to 2008m8 such that the last forecasted period is always one month before the Lehman bankruptcy filing, and a crisis forecast sample that starts in 2007m8 and ends in 2012m12. The mean squared forecast error relative to the autoregressive model are reported in Table 2. Values smaller than one with one (two) star(s) denote that the model in the column index does significantly better than the autoregressive benchmark at a 10% (5%) significance level, according to a Diebold-Mariano test, modified using the small-sample correction of Harvey et al. [1998]. We find several interesting results.

First, TVP-VAR models perform better than the AR model at forecasting the change in industrial production for horizons of 1 to 3 months over the whole forecast sample. A similar result is found for employment, but only for the crisis subsample (and not for the core TVP-VAR). The best forecast model for these two activity variables is the one augmented with the GZ spread indicator, particularly in the crisis forecast subsample. Noticeably, the forecast errors of this model are smaller than those from the model augmented with the EBP, suggesting that default risk contains informational content for the point forecasts of both industrial production and employment.

Second, inflation appears by contrast to be much harder to forecast, as the autoregressive benchmark beats the VAR models over the whole forecast sample.[7] A relevant exception is personal consumption expenditure inflation, which is forecasted well in

[7] It is an established result in the empirical literature that multivariate models often produce worse inflation forecasts than simpler univariate models. See for instance Stock and Watson [2007] and Faust and Wright [2013].

Table 2: **Mean squared forecast errors of the competing models models relative to an AR model** for different forecast samples and horizons (months). Two (one) stars denote significantly different RMSFE at a 5% (10%) significance level according to a Diebold-Mariano test, modified using the small-sample correction of Harvey et al. [1998]. The forecasting models are described in table 1. The pre-crisis sample goes from 1980:m1 to 2008:m8. The crisis sample spans the period from 2008:m9 to 2012:m12.

	PRE-CRISIS SAMPLE				CRISIS SAMPLE				FULL SAMPLE			
	CORE	GZ	BA	EBP	CORE	GZ	BA	EBP	CORE	GZ	BA	EBP
IMI												
$h = 1$	0.94	0.91**	0.93	0.93*	1.08	0.76**	0.90*	0.91*	0.97	0.88**	0.92**	0.93**
$h = 3$	0.95	0.91	0.95	0.92	1.05	0.89	1.01	0.94	0.97	0.91	0.96	0.92
$h = 6$	1.05	0.99	1.05	0.99	1.03	1.04	1.09	1.03	1.05	1.00	1.06	1.00
$h = 12$	1.01	0.99	1.00	0.99	1.00	1.18	1.11*	1.10*	1.01	1.06	1.04	1.02
EMPLOYMENT												
$h = 1$	1.08	1.03	1.01	1.04	1.08	0.75**	0.82	0.92	1.08	1.00	0.99	1.03
$h = 3$	1.22	1.13	1.19	1.09	1.11*	0.79	0.88	0.91	1.20	1.07	1.14	1.06
$h = 6$	1.18	1.09	1.22	1.09	1.09**	0.86	1.02	0.92	1.16	1.04	1.17	1.05
$h = 12$	0.94	0.91	1.03	0.92	1.07*	1.08	1.13**	1.08	0.98	0.96	1.06	0.97
PPI												
$h = 1$	1.22*	1.25**	1.24**	1.29**	1.05	1.18	1.06	1.08	1.21*	1.25**	1.23**	1.27**
$h = 3$	1.38**	1.40**	1.40**	1.48**	1.10	1.29	1.11	1.12	1.36**	1.40**	1.38**	1.45**
$h = 6$	1.33**	1.35**	1.37**	1.46**	1.02	1.39	1.03	1.08	1.30**	1.36**	1.34**	1.43**
$h = 12$	1.49**	1.57**	1.51**	1.71**	0.87	1.58**	0.88	0.99	1.44**	1.57**	1.45**	1.65**
PCEPI												
$h = 1$	1.21**	1.23**	1.24**	1.23**	0.93	0.67	0.90*	0.89*	1.13**	1.08	1.15**	1.14**
$h = 3$	1.38**	1.41**	1.41**	1.42**	0.88	0.99	0.87	0.89	1.21**	1.26**	1.22**	1.23**
$h = 6$	1.56**	1.61**	1.61**	1.64**	0.82*	1.06	0.83**	0.87	1.28**	1.40**	1.31**	1.35**
$h = 12$	1.60**	1.64**	1.61**	1.75**	0.88**	1.26	0.94	0.96	1.40**	1.54**	1.42**	1.53**
LOANS												
$h = 1$	0.98	0.97	1.01	0.96	1.00	0.85	0.86	0.83	0.99	0.94*	0.98	0.93
$h = 3$	1.02	0.99	1.07	0.94	1.00	0.67	0.74	0.65	1.01	0.91	0.99	0.87
$h = 6$	1.09**	1.02	1.19**	0.93	0.99	0.52	0.66	0.49	1.06**	0.86	1.02	0.79
$h = 12$	1.09**	0.99	1.19**	0.85	0.96**	0.59	0.73	0.50**	1.04	0.84	1.02	0.72*
LOAN SPREAD												
$h = 1$	0.86	0.87	0.87	0.85	0.68**	0.53*	0.71	0.65**	0.86	0.86	0.87	0.85
$h = 3$	0.84	0.84	0.87	0.81*	0.53**	0.63	0.70	0.59*	0.84	0.84	0.87	0.81*
$h = 6$	0.78**	0.78**	0.84	0.73**	0.40**	0.48**	0.49**	0.47**	0.77**	0.77**	0.83	0.73**
$h = 12$	0.70**	0.71**	0.78	0.63**	0.22**	0.29**	0.25**	0.29**	0.68**	0.69**	0.76	0.61**

the crisis subsample by both the BA and EBP model at short horizons.

Lastly, both loans and the loan interest rate spread are better forecasted by the TVP-VAR models. However, only in the case of loans the addition of default risk and risk aversion indicators improves the performance of the core TVP-VAR model.

4.2 Density forecasts

An alternative way to assess the forecasting performance is by examining the entire forecast distribution, rather than just the point forecasts. The aim of this Section is to gauge to what extent the inclusion of default and risk aversion measures helps to better characterise the forecast density and to predict, not only the mean, but also low and high realisations of the variables of interest. The key statistic is the evolution over time of the differences in log-predictive likelihoods between one of the three indicator-based TVP-VAR models and a benchmark forecast model (either the autoregressive model or the core TVP-VAR):

$$\mathcal{S}_h = \sum_{t=1}^{T_f-h} \left[\log p_{h,t}(y_{t+h} \,|\, \mathcal{F}_{j,t-1}) - \log g_{h\,t}(y_{t+h}|\, \mathcal{F}_{g_i,t-1}) \right] \quad , \tag{6}$$

where $\log p_{h,t}(\cdot)$ and $\log g_{h,t}(\cdot)$ respectively denote the likelihood (in log scale) of observing the realisation of the variable y_{t+h} given an indicator-based model or a benchmark model (either the baseline TVP-VAR or the AR model), and given the information set of model j available at t, $\mathcal{F}_{t-1,j}$.

Figures 2 to 7 display the relevant $\mathcal{S}_h$ statistics, for different indicator-based models, where the benchmark model is either the autoregressive (left panels) or the baseline TVP-VAR (right panels), and where increases in the statistics denote observations in favour of the indicator-based TVP-VAR models. This exercise, similar to what is undertaken in Amisano and Geweke [2010], has a justification in terms of the Kullback-Leibler distance (KLIC). In particular, if one model has on average higher predictive likelihood than a competitor, then the forecast model is closest (in terms of the KLIC distance) to the true data generating process.

Figures 2 and 3 show that the marginal contribution of financial indicators in forecasting production and employment starts to be relevant after the 2001 recession but becomes substantial only after the great financial crisis, when credit constraints were arguably binding. Between 2000 and 2008, the EBP model does equally well or better than the GZ and BA spread models, indicating that risk aversion is the most important predictor for economic activity in this period. Default risk premia seem to have played a more important role during the 2008 financial crisis, particularly at a one-year ahead horizon. In this period in fact the GZ model yields on average the highest predictive likelihoods among the VARs. Note however that at large horizons all TVP-VAR models perform worse than the autoregressive benchmark after 2008.

By contrast, financial indicators do not seem to have any predictive content for inflation measures, as shown in Figures 4 and 5. Inflation is typically very hard to forecast and it is difficult to find a model that is able to beat an autoregressive benchmark. However, controlling for financial indicators and in particular for risk aversion, substantially worsens the forecast of producer price inflation with respect to the baseline TVP-VAR. In the case of personal consumption expenditure inflation controlling for the GZ spread or the excess bond premium only marginally improves the forecasts of the baseline TVP-VAR model between the mid 1980s and the mid 1990s, and during the 2008 financial crisis. The improvement is in any case much smaller than that witnessed for industrial production.

Lastly, financial indicators seem to have some predictive power only for loans and not for the loan interest rate spread. In the case of loans (Figure 6), the three indicator-based models deliver very similar predictive likelihoods to the baseline TVP-VAR and to the autoregressive models for most of the forecast subsample. After 2001 the EBP model scores progressively better, particularly at larger horizons, and the bulk of the increase is due once again to the 2008 financial crisis. The EBP offers a marginal improvement to forecasting the loan interest rate spread (Figure 7), but this is limited to longer forecast horizons and to the 1990-2000 period. Overall, controlling for past values of the spread, as well as for the term structure and the real risk free rate, is sufficient to deliver accurate forecasts for the loan interest rate spread.

To understand what drives the better performance of the indicator-based models, and in particular of the GZ model, during the 2008 crisis period, Figure 8 shows the one-period ahead point forecasts (dashed lines) and 68% forecast confidence intervals (shaded area) of the four TVP-VAR models, together with those of the autoregressive model (dashed-dotted lines). Across the wide range of variables considered, the point forecasts of the baseline TVP-VAR model are very similar to those of the autoregressive model, though the confidence intervals are narrower. Controlling for financial conditions, and in particular for the GZ spread, shifts the entire forecast distribution of industrial production and employment to lower values during the 2008 financial crisis period, thereby delivering smaller forecast errors. A similar result holds for personal consumption expenditure inflation, but not for producer price inflation or for loans. Conversely, controlling for the GZ spread helps to capture the increase in the compensation for lending risk in the aftermath of the 2008 financial crisis (see Figure 8f).

5 Has the importance of shocks to financial conditions changed over time?

In the previous Section we have documented time variation in the predictive ability of financial condition indicators for macroeconomic activity. In particular, credit spreads and the excess bond premium are found to be significant regressors for macroeconomic activity during the periods of financial distress of 2001 and 2008. In this Section we seek a causal interpretation of these results and examine whether the effect of credit spreads shocks on industrial production growth has changed over time. For this purpose, we rely on a modified version of the forecast models that includes industrial production, inflation, the measure of financial conditions, the nominal risk-free interest rate and the term structure slope.[8] The VARs are identified recursively, assuming that financial condition measures are not affected by monetary policy shocks, but that the monetary authority takes into account financial conditions when setting its rate. This identification scheme is the same as the one used in Gilchrist and Zakrajšek [2012] and in other works by the same authors.[9] While this identification scheme has many caveats, we choose it in order to show that the main structural result in Gilchrist and Zakrajšek [2012] appears to be driven by the recent financial crisis.

Figures 9, 11 and 13 display the effects of a one-standard deviation shock in one of the three financial condition indicators at different points in time. These plots clearly show how the importance of financial condition measures has changed over time. While a GZ credit spread shock has no significant effect on industrial production before 2000,

[8]In this way, no restriction on the coefficients of inflation and of the nominal interest rate are imposed. Moreover, the shock to the interest rate series can be interpreted as a pure monetary policy shock. Using the real interest rate does not alter the results.

[9]Differently from Gilchrist and Zakrajšek [2012], we employ a smaller set of variables (the same used in the forecast exercise). Results from an alternative variable ordering do not differ significantly from those reported here and are available upon request.

Figure 2: **Density forecasts for industrial production**

(a) **Relative to the AR model** (b) **Relative to baseline TVP-VAR model**

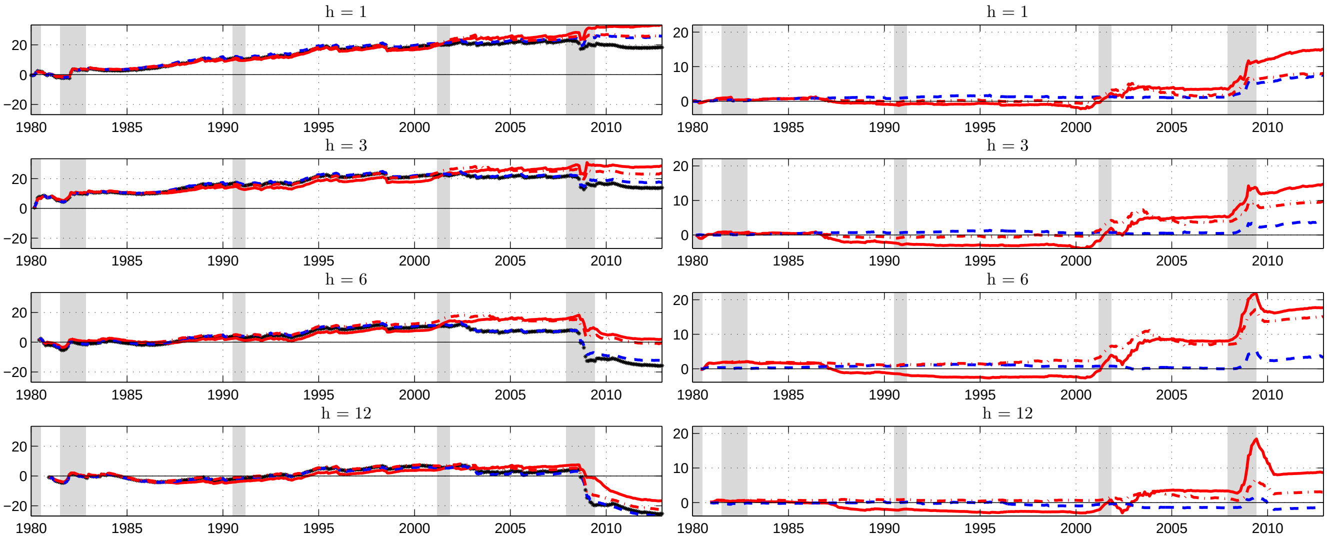
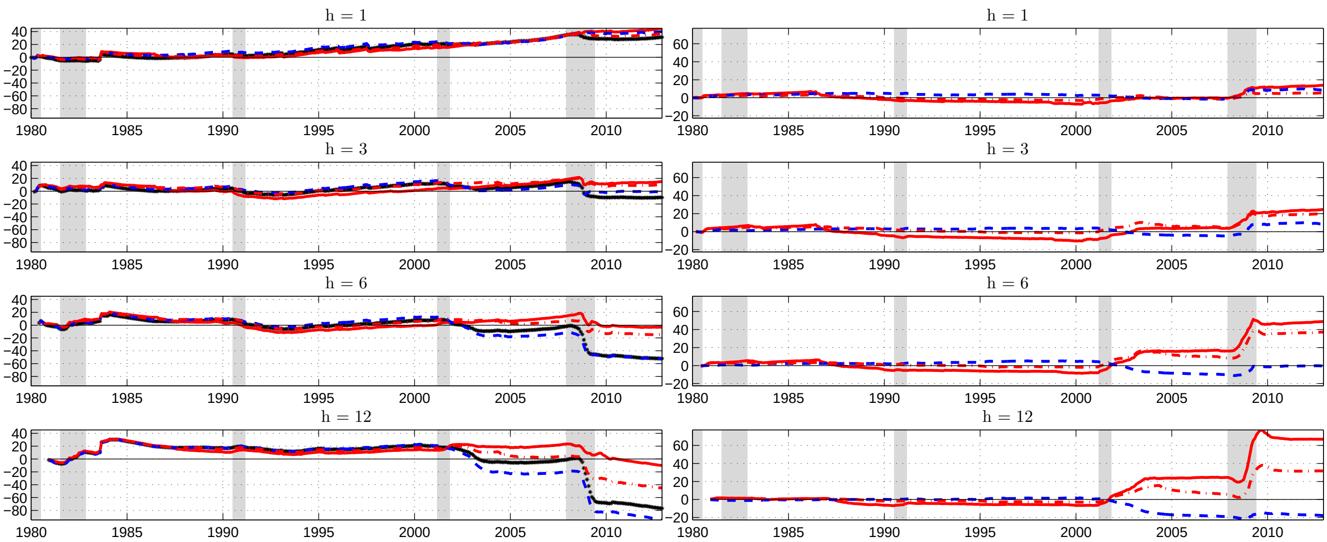


Figure 3: **Density forecasts for employment**

(a) **Relative to the AR model** (b) **Relative to baseline TVP-VAR model**



Note: Predictive log-likelihoods of the GZ model (solid red line), of the BA model (dashed blue line), and of the EBP model (dashed dotted red line) relative to the AR model (Figure a) and to the baseline TVP-VAR model (Figure b) at different forecast horizons (months). The starred black line in Figure a denotes the log-likelihood of the baseline TVP-VAR relative to the AR model. Positive values denote that, on average, the model provides a better out-of-sample fit than the benchmark. Increases in the statistics denote observations in which the fit of the TVP-VAR models is higher. Shaded areas denote NBER-dated recessions.

its importance gradually increases over time, and in 2010 a one-standard deviation increase in the GZ spread causes an output contraction that lasts for two years, with

Figure 4: **Density forecasts for producer price inflation**

(a) **Relative to the AR model** (b) **Relative to baseline TVP-VAR model**

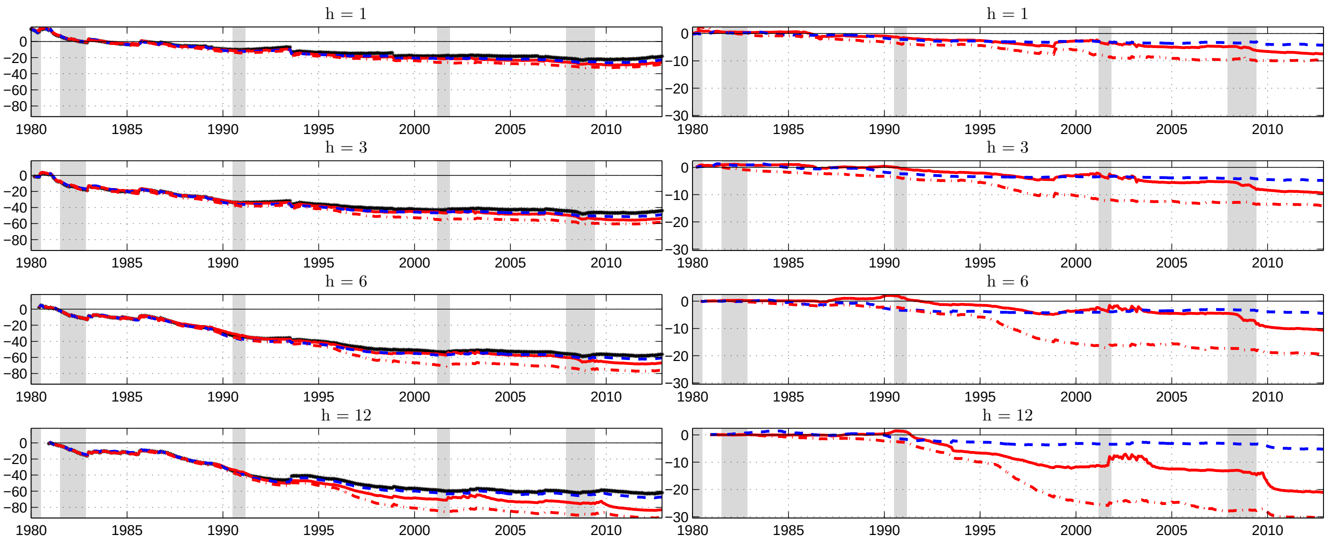
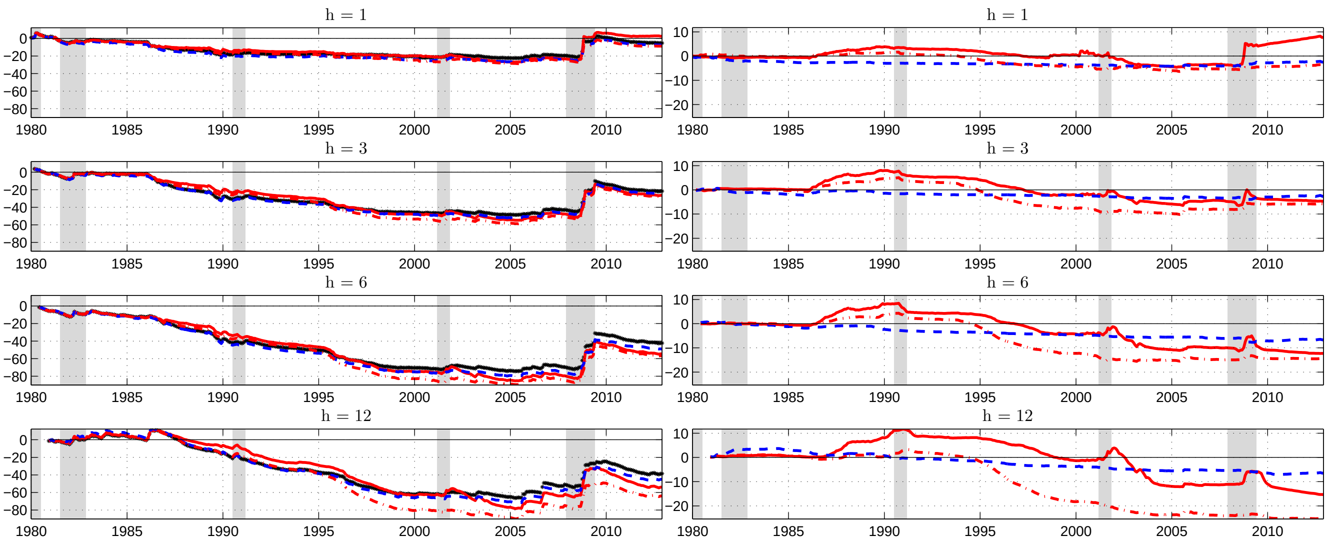


Figure 5: **Density forecasts for personal consumption expenditure inflation**

(a) **Relative to the AR model** (b) **Relative to baseline TVP-VAR model**



Note: Predictive log-likelihoods of the GZ model (solid red line), of the BA model (dashed blue line), and of the EBP model (dashed dotted red line) relative to the AR model (Figure a) and to the baseline TVP-VAR model (Figure b) at different forecast horizons (months). The starred black line in Figure a denotes the log-likelihood of the baseline TVP-VAR relative to the AR model. Positive values denote that, on average, the model provides a better out-of-sample fit than the benchmark. Increases in the statistics denote observations in which the fit of the TVP-VAR models is higher. Shaded areas denote NBER-dated recessions.

Figure 6: **Density forecasts for loans**

(a) **Relative to the AR model** (b) **Relative to baseline TVP-VAR model**

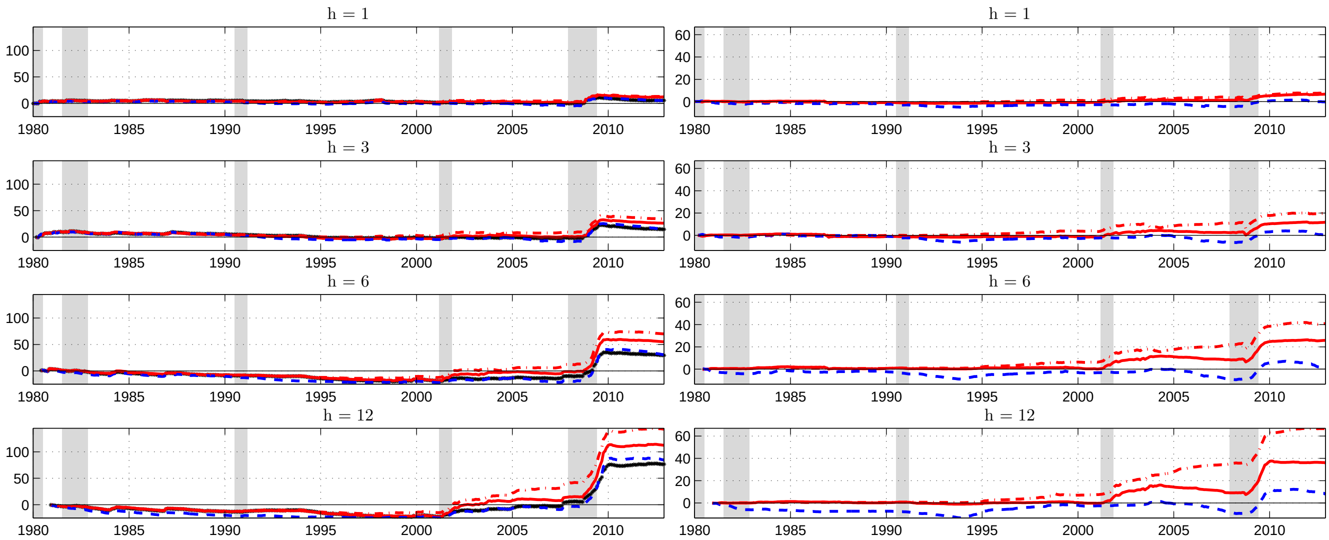
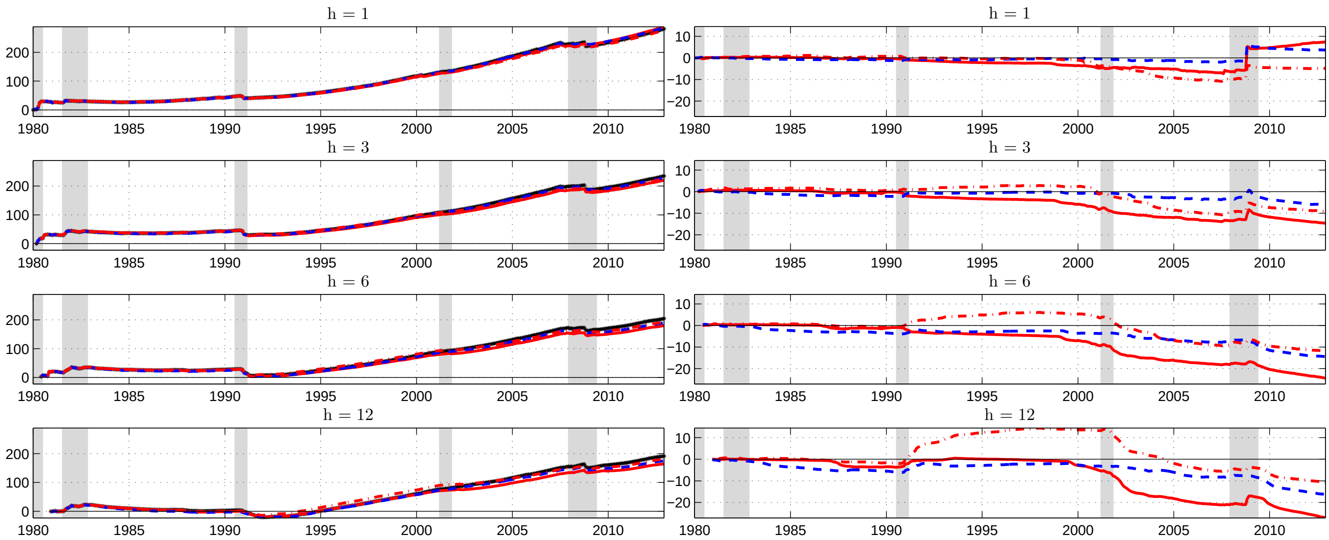


Figure 7: **Density forecasts for the loan interest rate spread**

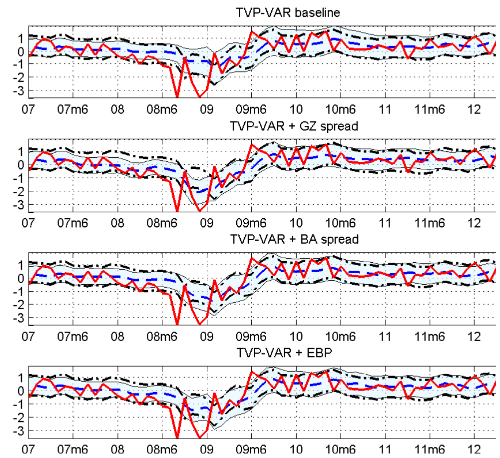
(a) **Relative to the AR model** (b) **Relative to baseline TVP-VAR model**



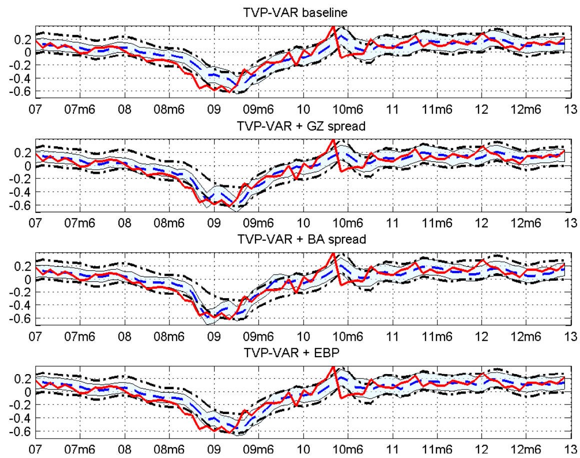
Note: Predictive log-likelihoods of the GZ model (solid red line), of the BA model (dashed blue line), and of the EBP model (dashed dotted red line) relative to the AR model (Figure a) and to the baseline TVP-VAR model (Figure b) at different forecast horizons (months). The starred black line in Figure a denotes the log-likelihood of the baseline TVP-VAR relative to the AR model. Positive values denote that, on average, the model provides a better out-of-sample fit than the benchmark. Increases in the statistics denote observations in which the fit of the TVP-VAR models is higher. Shaded areas denote NBER-dated recessions.

Figure 8: **One-month ahead forecast confidence intervals**

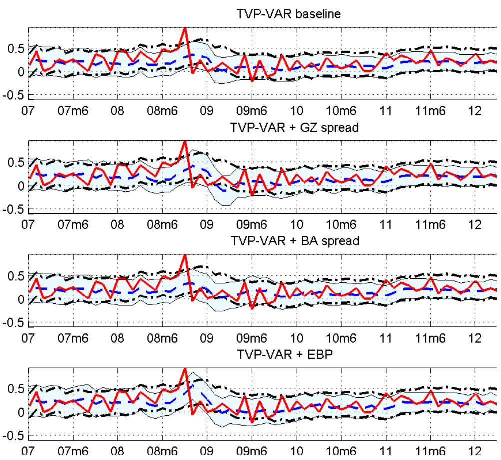
(a) **Industrial production**



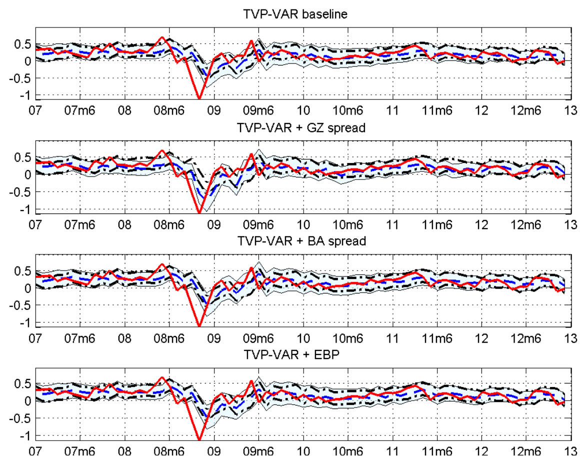
(b) **Employment**



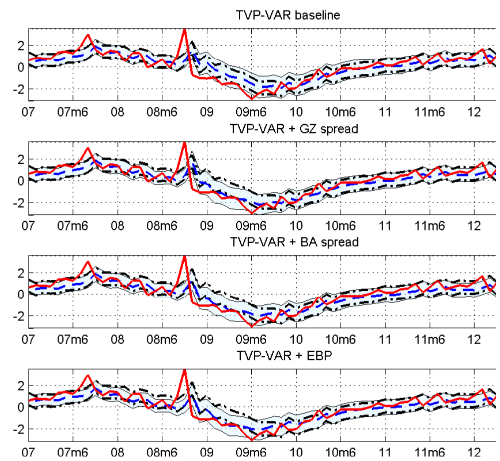
(c) **Producer prices**



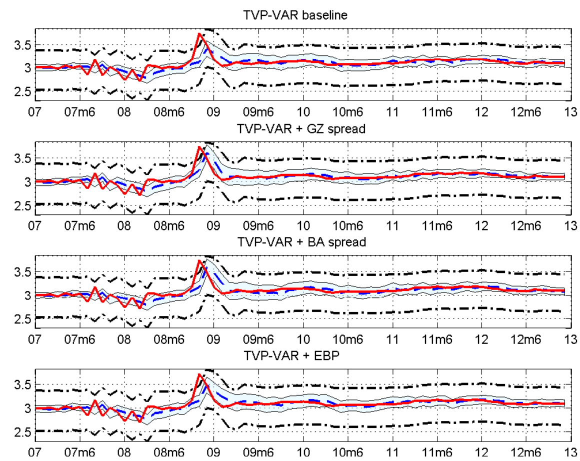
(d) **Personal consumption expenditure**



(e) **Loans**



(f) **Loan interest rate spread**



Note: Actual series (solid red lines) plotted together with the point forecasts (dashed lines) and 68% forecast confidence intervals (shaded area) of one of the four TVP-VAR models (reported in the plot title). The dashed-dotted lines are the 68% forecast confidence intervals of the benchmark autoregressive model.

an annualised through of 4.8%, despite an easing of monetary policy. The forecast error variance decomposition shown in Figure 10 reveals that GZ spread shocks account for a negligible fraction of industrial production variance until the end of the 1990s, but that in 2010 this fraction increases to 35% after two years. This result is in line with the existing empirical literature on the time-varying effects of financial shocks. For instance, Prieto et al. [2016] find that the explanatory power of financial shocks for US GDP growth is twice as high during the Great Recession than over more "normal" times.

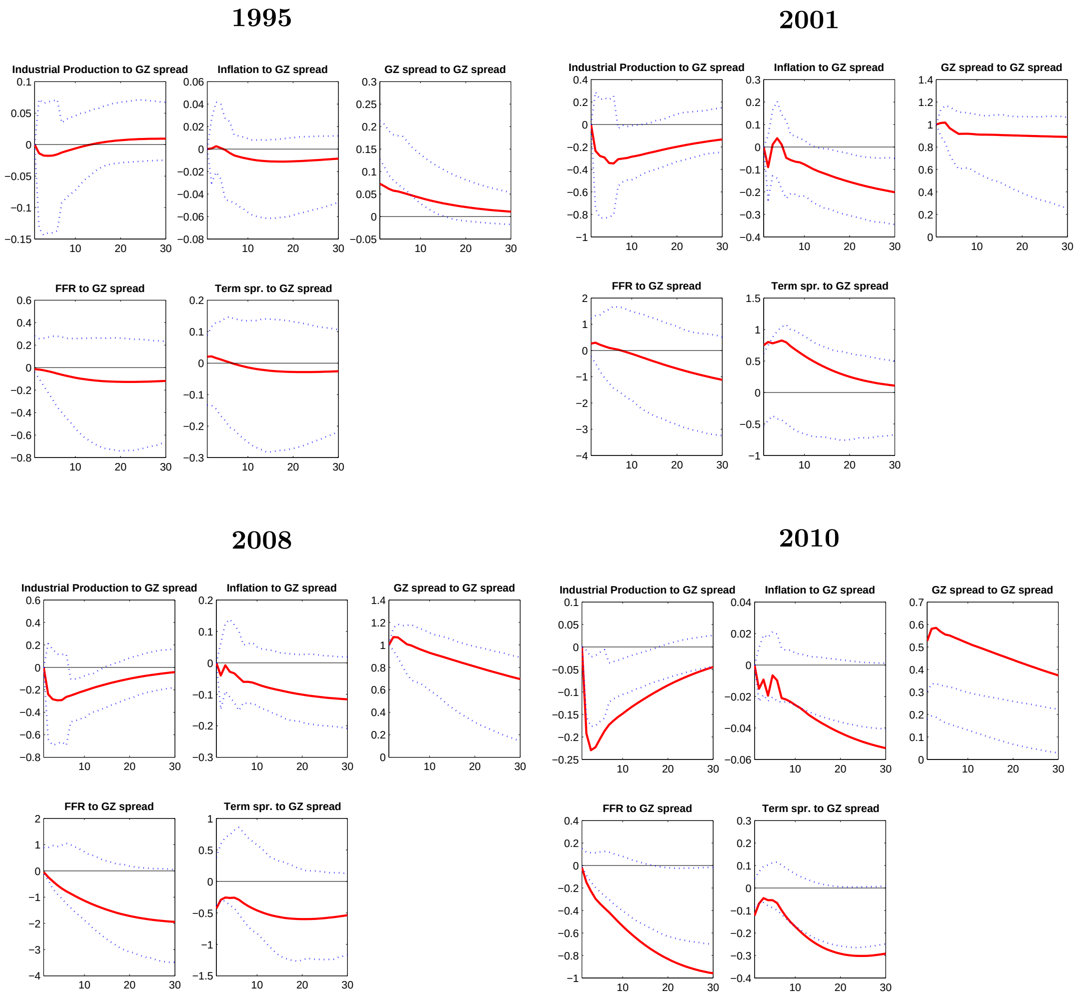
In Figure 13 financial conditions are instead measured with the excess bond premium, thus taking into account only risk aversion and not default risk. The same conclusions obtained with the GZ spread hold, but the magnitude of the effects are reduced. The decline in industrial production in 2010 has an annualised through of 2.4%, and the fraction of total variance explained by the financial condition shock is only 13% after 2 years (Figure10). Recalling that the GZ spread controls for two components, one reflecting countercyclical default premia and one capturing the extra compensation required by investors to accept default risk, these results show that default premia are just as important as risk aversion in explaining the decline in industrial production growth in the recent years. Hence, the fall in output and slow recovery seems to have been caused by an increased risk aversion of lenders as well as by worsened macro-fianancial fundamentals (i.e. higher default risk) .

Lastly, note that results depend on how financial conditions are measured. If the BA spread replaces the GZ spread in the VAR (Figure 11) the retrieved financial condition shocks are never found to affect industrial production significantly.

6 Conclusions

The recent financial crisis has illustrated how interrelated the financial and the real side of the economy are. Economic models tell us moreover that the existence of financial frictions strengthens the transmission of shocks, and that movements in default

Figure 9: **Responses to a one-standard deviation GZ-spread shock over the years**



Note: Median impulse response functions (solid lines) with 95% confidence band (area between dotted lines) obtained by bootstrap. SVAR estimated through the forgetting factor methodology outlined in Section 2.1, with a lag-length of 6 months, and identified through a recursive scheme, with the credit spread ordered third.

Figure 10: **Forecast error variance of industrial production** attributable to a shock in the GZ spread, for selected years.

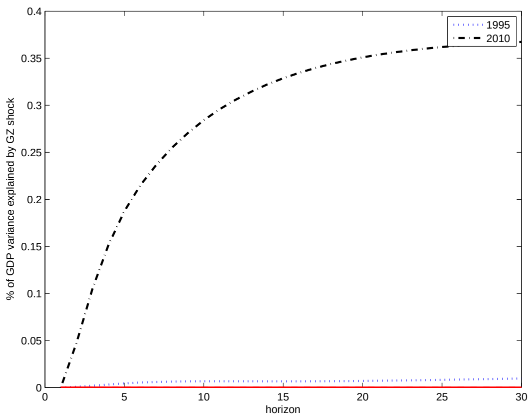
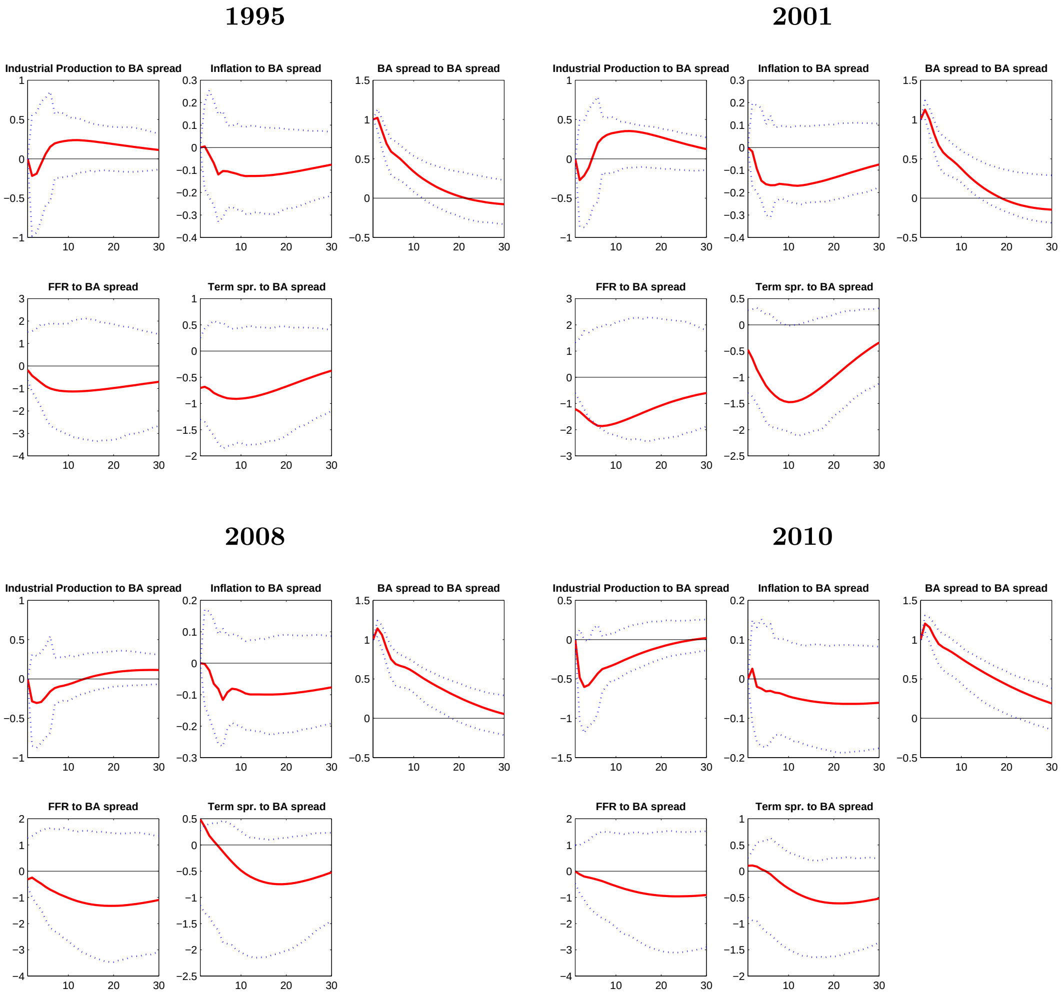


Figure 11: **Responses to a one-standard deviation BA-spread shock over the years**



Note: Median impulse response functions (solid lines) with 95% confidence band (area between dotted lines) obtained by bootstrap. SVAR estimated through the forgetting factor methodology outlined in Section 2.1, with a lag-length of 6 months, and identified through a recursive scheme, with the credit spread ordered third.

Figure 12: **Forecast error variance of industrial production** attributable to a shock in the BA spread, for selected years.

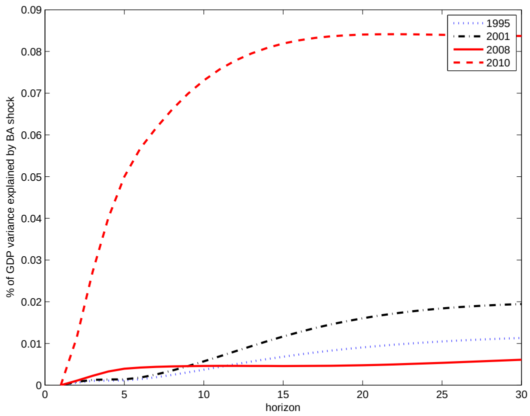
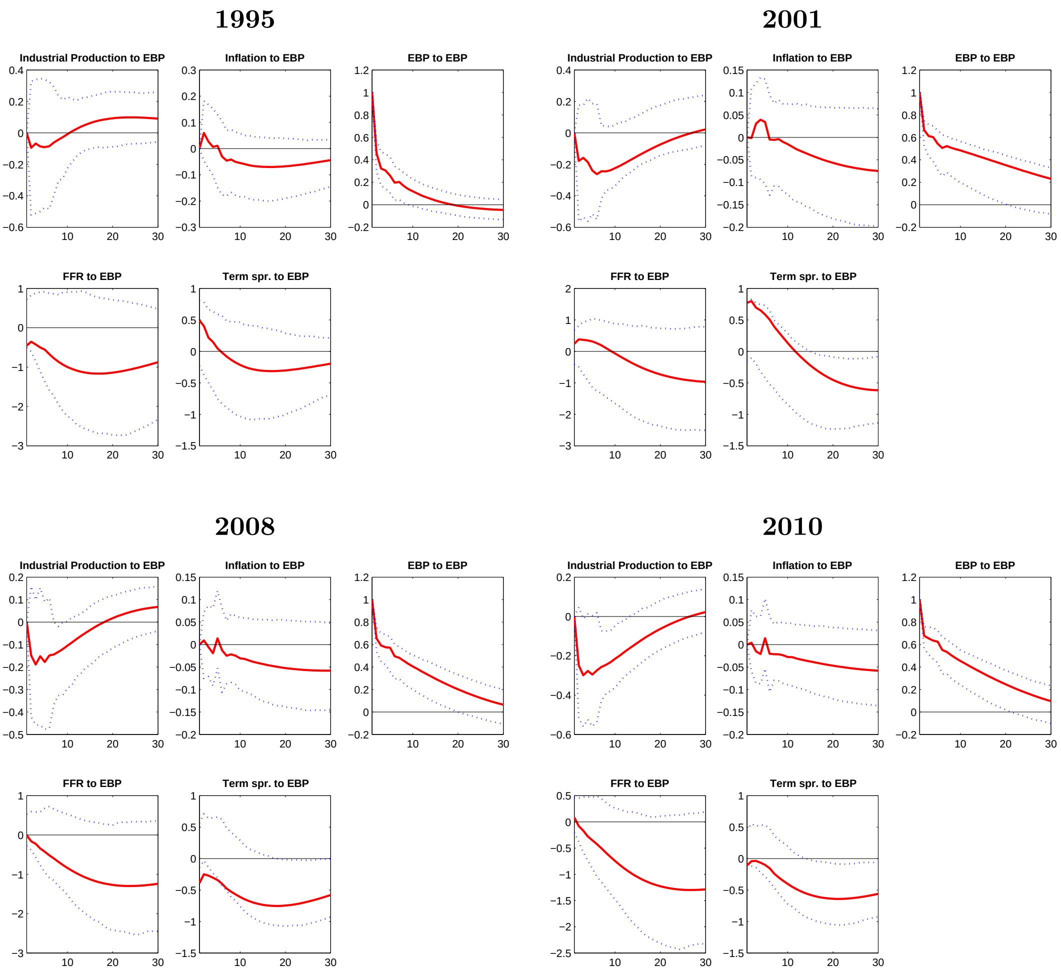
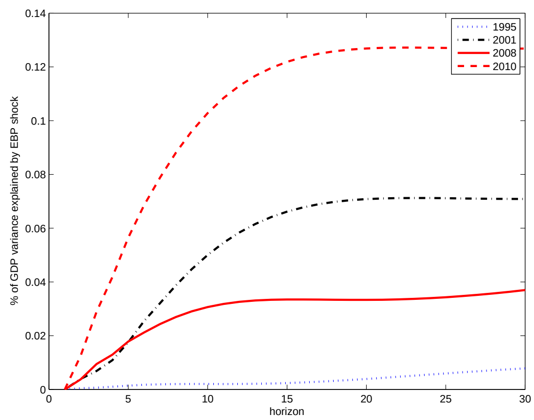


Figure 13: **Responses to a one-standard deviation EBP shock over the years**



Note: Median impulse response functions (solid lines) with 95% confidence band (area between dotted lines) obtained by bootstrap. SVAR estimated through the forgetting factor methodology outlined in Section 2.1, with a lag-length of 6 months, and identified through a recursive scheme, with the EBP ordered third.

Figure 14: **Forecast error variance of industrial production** attributable to a shock in the EBP measure, for selected years.



risk or in risk aversion induce changes in the cost and in the quantity of borrowing, which anticipate future movements in loans, output, and even prices. Based on these theoretical models, we verify to what extent the inclusion of financial condition indicators improves the point and density forecasts of key macroeconomic variables, relative to a baseline TVP-VAR model with control variables, and to a simple autoregressive model estimated recursively.

Three financial condition indicators are considered: the spread between the Baa and Aaa-rated corporate bond yields, a spread developed by Gilchrist and Zakrajšek [2012] using a richer corporate bond database, and the excess bond premium. The latter has been proposed by Gilchrist and Zakrajšek [2012] and is defined as the residual component of the GZ spread that is not correlated to default risk but captures overall risk aversion.

Several result emerge from the forecasting analysis. First, controlling for financial condition indicators improves the point forecasts of output and employment with respect to the baseline TVP-VAR model during the 2001 and 2008 recessions, when credit constraints are arguably binding. The GZ spread delivers more accurate point forecasts for output and employment than the model that controls for the excess bond premium. This suggests that default risk is an important predictor for aggregate activity indicators, especially in the forecast subsample that includes the 2008 financial crisis. Controlling solely for risk aversion suffices instead to improve the point forecasts of loans relative to the autoregressive benchmark.

Second, the analysis of the entire forecast distribution reveals that the marginal contribution of financial indicators in forecasting output, employment and loans starts to be relevant only after the 2001 recession and becomes substantial only after the great financial crisis, when credit constraints were arguably binding. While risk aversion is the most relevant predictor for loans, default risk premia have played an important role in forecasting output and employment during the 2008 financial crisis, particularly at a one-year ahead horizon. In particular, controlling fro the GZ spread helps to forecast the macroeconomic downturn seen at the end of 2008, shifting the entire distribution of industrial production towards lower values. By contrast, controlling for financial condition indicators does not improve the density forecasts of the two inflation measures considered, nor of the interest rate spread.

Lastly, we have conducted a structural analysis to assess whether shocks to financial conditions display different real effects over time. Following the same identification strategy as in Gilchrist and Zakrajšek [2012], we show that their results appear to be influenced by the latest financial crisis. An unexpected worsening in financial conditions displays in fact adverse real effects only after 2008. Moreover, the latest economic recession and subsequent slow recovery can be explained by both higher default risk and increased risk aversion.

Bibliography

Angela Abbate and Massimiliano Marcellino. Point, interval and density forecasts of exchange rates with time-varying parameter models. Discussion Papers 19/2016, Deutsche Bundesbank, Research Centre, 2016.

Piergiorgio Alessandri, Antonio M. Conti, and Fabrizio Venditti. The Financial Stability Dark Side of Monetary Policy. BCAM Working Papers 1601, Birkbeck Centre for Applied Macroeconomics, May 2016.

Gianni Amisano and John Geweke. Comparing and evaluating bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230, April 2010.

Ben Bernanke and Mark Gertler. Agency Costs, Net Worth, and Business Fluctuations. *American Economic Review*, 79(1):14–31, March 1989.

Ben S. Bernanke, Mark Gertler, and Simon Gilchrist. The financial accelerator in a quantitative business cycle framework. In J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 21, pages 1341–1393. Elsevier, 1999.

Menzie D. Chinn and Kavan J. Kucko. The Predictive Power of the Yield Curve across Countries and Time. NBER Working Papers 16398, National Bureau of Economic Research, Inc, September 2010.

Timothy Cogley and Thomas J. Sargent. Drift and volatilities: Monetary policies and outcomes in the post wwii u.s. *Review of Economic Dynamics*, 8(2):262–302, April 2005.

Pierangelo De Pace and Kyle D. Weber. The time-varying leading properties of the high yield spread in the United States. *International Journal of Forecasting*, 32(1): 203–230, 2016.

Jon Faust and Jonathan H. Wright. Forecasting inflation. In *Handbook of Economic Forecasting*, volume 2, chapter 1, pages 2–56. Elsevier, 2013.

Jon Faust, Simon Gilchrist, Jonathan H. Wright, and Egon Zakrajsek. Credit spreads as predictors of real-time economic activity: a Bayesian Model-Averaging approach. *Review of Economics and Statistics*, 95(5):1501–1519, 2013.

Simon Gilchrist and Egon Zakrajšek. Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4):1692–1720, 2012.

Simon Gilchrist, Raphael Schoenle, Jae Sim, and Egon Zakrajšek. Inflation dynamics during the financial crisis. Finance and Economics Discussion Series 2015-012, Board of Governors of the Federal Reserve System, February 2015.

David I Harvey, Stephen J Leybourne, and Paul Newbold. Tests for Forecast Encompassing. *Journal of Business & Economic Statistics*, 16(2):254–59, April 1998.

Gary Koop and Dimitris Korobilis. Large time-varying parameter vars. *Journal of Econometrics*, 177:185–198, February 2013.

Esteban Prieto, Sandra Eickmeier, and Massimiliano Marcellino. Time variation in macro-financial linkages. *Journal of Applied Econometrics*, 2016. ISSN 1099-1255.

Giorgio E. Primiceri. Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72(3):821–852, 2005.

James H. Stock and Mark W. Watson. Why Has U.S. Inflation Become Harder to Forecast? *Journal of Money, Credit and Banking*, 39(s1):3–33, 02 2007.