The Quarterly Review of Economics and Finance 71 (2019) 211–222

Contents lists available at ScienceDirect

The Quarterly Review of Economics and Finance

journal homepage: www.elsevier.com/locate/qref

Time-varying predictive content of financial variables in forecasting GDP growth in the G-7 countries

Petri Kuosmanen*, Juuso Vataja

Department of Economics, University of Vaasa, P.O. Box 700, 65101 Vaasa, Finland

ARTICLE INFO

Article history:
Received 23 February 2018
Received in revised form 3 July 2018
Accepted 7 August 2018
Available online 10 August 2018

JEL classification:
E37
E44
E47

Keywords:
Term spread
Short-term interest rates
Stock market
Forecasting
Macroeconomy

ABSTRACT

The predictive association between financial markets and the real economy has proven unstable and transitory over time. This study reexamines empirical evidence regarding the predictive content of financial variables for GDP growth in light of the changed economic circumstances in the G-7 countries in the 2000s. We explicitly address time variations in the predictive power of financial variables for GDP growth. The results indicate that the behavior of the forecasting ability contains a considerable amount of temporal dominance and time persistence, which often vary contemporaneously among the G-7 countries. The forecasting content is clearly connected to unsettled economic conditions.

1. Introduction

In many respects, the relationship between financial markets and the real economy is odd and puzzling: there are many well-based theoretical arguments and a substantial amount of empirical evidence that financial markets can be used to forecast the real economy. However, the causal and predictive links between financial markets and the real economy can be characterized as momentary. There are periods when some financial variables seem to be highly useful predictors for the real economy in certain countries or time periods, but soon thereafter, the same forecasting relation is revealed to be coincidental or nonexistent (Stock & Watson, 2003a). This study explicitly addresses time variation in financial variables' ability to predict GDP growth in the G-7 countries. Despite its importance, this issue has been largely overlooked in previous studies.

This study considers the predictive ability of three key financial indicators – the term spread, the real short-term interest rate and real stock returns – of GDP growth during the Great Moderation and financial crisis eras in the G-7 countries, i.e., Canada, France, Germany, Italy, Japan, the United Kingdom and the United States. It is of interest to reexamine the forecasting content of financial predictors in the 2000s, a period characterized by varying economic circumstances, e.g., the burst of the techno bubble, the end of the Great Moderation, the financial crisis and the subsequent sovereign debt and banking crises, and the implementation of unconventional monetary policy.

We propose that the time-varying forecasting content of financial predictors is better captured by analyzing the behavior of the actual forecast errors at each time point than by concentrating on the behavior of the average forecast errors, typically the root mean square errors. We suggest applying the forecast error spreads to uncover the time-varying predictive content of financial predictors. This is the main contribution of the study. Our empirical results reveal a considerable amount of temporal dominance and time persistence in the forecast errors, which often move contemporaneously across the G-7 countries. The forecasting content is obviously connected to unsettled economic conditions. These empirical results are novel.

Understanding the regularities and possible reasons behind time variations in predictive power is of vital importance for investors and economists. The stock market responds quite pow-

* Corresponding author.
E-mail addresses: pmk@uva.fi (P. Kuosmanen), jpv@uva.fi (J. Vataja).

https://doi.org/10.1016/j.qref.2018.08.002
1062-9769/

erfully to changes in economic activity; however, episodically, the markets and the economy are out of sync and the financial market may deliver false signals (Siegel, 2014). If investor can distinct trustworthy and false signals from financial markets, the rewards of investing stocks are large. In a similar vein, financial variables are useful in economists' toolbox because they are precisely measured in real time and are easy to use in forecasting economic activity; however, it is crucial to recognize when financial variables contain more – or less – trustworthy and useful information for future real activity.

The paper is organized as follows. Section 2 provides a literature review. Section 3 describes the modeling strategy. Section 4 introduces the data. The in-sample analysis and out-of-sample forecasts are presented in Section 5. Time variations of the forecast errors are analyzed in Section 6. Finally, Section 7 concludes.

2. Empirical regularities in forecasting economic activity with financial data

From the late 1980s, the slope of the yield curve or even the simple term spread (the difference between long-term and short-term interest rates) began to be recognized as the single most important predictor of economic activity in the developed countries (Harvey, 1988, 1989, 1991; Laurent, 1989; Estrella & Hardouvelis, 1991). However, its prevalence as the unambiguous leading indicator was short-lived; not long after it was introduced, numerous studies emerged suggesting that the forecasting power of the term spread for the real economy had diminished since the mid-1980s (e.g., Haubrich & Dombrosky, 1996; Dotsey, 1998; Estrella, Rodrigues, & Schich, 2003; Stock & Watson, 2003a; Giacomini & Rossi, 2006; D'Agostino, Giannone, & Surico, 2006; Wheelock & Wohar, 2009; Chinn & Kucko, 2015). The reasons for this deterioration have largely remained a conundrum. Another conflicting outcome emerged from Ang, Piazzesi, and Wei's, (2006) finding that the mere short end of the yield curve, the nominal short-term interest rate, performs better at forecasting GDP than any term spread in the U.S. economy. In addition, combining the term spread with the short-term interest rate improved forecasts more than using either variable alone for the U.S. over the 1875 to 1997 period (Bordo & Haubrich, 2008).

The periodically shifting predictive power of the term spread also holds true for the other G-7 countries. For example, Stock and Watson (2003a) found that the term spread was a useful predictor of output growth in Germany prior to the mid-1980s, but not after that period. Canada and Japan were the only exceptions among the G-7; in these two countries, the term spread continued to perform well as a predictor for economic growth after the mid-1980s. Chinn and Kucko (2015) confirmed that the predictive power of the term spread was weaker during the Great Moderation; however, they also suggested that the financial crisis and the increased volatility in economic activity may have once again strengthened the forecasting power of the term spread, at least in some European countries. Hännikäinen (2015) found similar results in the U.S. context. The alternative explanation for the association between financial markets and the real economy is connected to monetary regimes and monetary policy (e.g., Benati & Goodhart, 2008). Estrella and Mishkin (1997) noted that the predictive power of the term spread is connected to many determinants. However, the independence of the monetary policy is one of the most central factors, i.e., an independent monetary policy, as in the cases of the U.S. and Germany, is associated with a stronger predictive power of the term spread. However, it seems that this proposition does not hold true in small open economies, such as the Nordic economies in the 2000s (Kuosmanen, Nabulsi, & Vataja, 2015).

The objective function of central banks may play a decisive role in explaining the predictive power of term spreads, i.e., if the monetary authorities pay attention mainly to deviations between the actual and potential output growth and pay less attention to inflation, then the term structure is more informative in predicting future growth (Wheelock & Wohar, 2009). Others (e.g., Chinn & Kucko, 2015) have linked this predictive relation to the volatility of growth: during the Great Moderation, the predictive relationship was diluted; however, the financial crisis of 2008 seems to have once again strengthened this relation. These two explanations, monetary policy and macroeconomic volatility, may well be linked. D'Agostino et al., 2006 stated that the decrease in the predictability of output growth reflects improvements in monetary policy. Accordingly, successful monetary policy management and transparent communication lead to more stable economic growth, which, in turn, produces the somewhat counterintuitive consequence that the predictive power of the term spread weakens.

Estrella and Mishkin (1995) suggested that in addition to the term spread, stock price indices are the most useful financial indicators in macroeconomic predictions. Siegel (2014: 238) stressed that "Stock values are based on corporate earnings, and the business cycle is a prime determinant of these earnings." The main link between stock returns and output growth is attributable to the forward-looking characteristics of the stock market: news about future output growth is quickly reflected in stock prices (Mauro, 2003). Oddly enough, the weakening of the predictive content of stock returns in the U.S. since the 1980s nearly coincided with the diminished forecasting power of the term spread (Binswanger, 2000). Binswanger (2004) also detected similar breakdowns in the predictive power of stock returns in Japan, Canada and the G-7 European countries combined in the early and late 1980s; however, this phenomenon was not clearly observed in the four individual European G-7 countries. The breakdowns are explained by, among other factors, the speculative stock market bubbles in the 1980s and 1990s, which led to the decoupling of the stock market and the real economy in several developed economies (Binswanger, 2004).

Moreover, the nature of the relationship between stock returns and output growth may not be linear or symmetric. Henry, Olekalns, and Thong, (2004) found evidence from data on 27 countries that stock returns contain useful information for predicting economic growth only when the economy is contracting. The role of stock returns in forecasting economic activity may also be connected to the size of the stock market relative to GDP: a high stock market capitalization increases the predictive power of stock returns in advanced economies and in emerging markets (Mauro, 2003). In general, the predictive power of stock returns for output growth is even more time varying, controversial and murky than that for term spreads (e.g., Stock & Watson, 2003a). Accordingly, Samuelson (1966) mocked the U.S. case, stating that "*Wall Street indexes predicted nine out of the last five recessions!*" Despite many reservations and economics jokes, there is a long tradition (see, e.g., Mitchell & Burns, 1938) of considering stock market movements as a potential and serious way to anticipate cyclical movements in the U.S. economy and other advanced economies.

The optimal number of financial predictors has also remained an open question. The previous literature has focused primarily on the predictive power of a single financial variable rather than studying the importance of additional financial predictors (e.g., Harvey, 1989, 1991; Kozicki, 1997; Domian & Louton, 1997; Dotsey, 1998; Binswanger, 2004; Bordo & Haubrich, 2008; Tsouma, 2009). Stock and Watson (2003a) found that no clear systematic patterns of improvement in forecasting performance existed when additional asset indicator candidates were added to bivariate models for the G-7 countries. In contrast, multivariate forecasting models were found to be superior to bivariate models in forecasting GDP growth in the Nordic countries (Kuosmanen et al., 2015).

3. Modeling Strategy

We focus on a forecast horizon of four quarters because it is the horizon most often used in practice and has been found to be the most suitable for financial data[1] (Kozicki, 1997; Wheelock & Wohar, 2009). The linear autoregressive (AR) model (Model 1) constitutes a natural and often-used benchmark against which more versatile competing models are compared.

$$y_{t+4} = \alpha^1 + \sum_{j=1}^{p} \beta_{1j}^1 y_{t-j+1} + u_{t+4}^1 \tag{1}$$

where $y_{t+4} = \ln\left(\frac{Y_{t+4}}{Y_t}\right)$, $y_t = \ln\left(\frac{Y_t}{Y_{t-1}}\right)$, Y_t is the quarterly real GDP at quarter t, α is the constant term, β_{1j}^1 are the parameter estimates for the AR terms, u_{t+4}^1 is the forecast error, and the superscript refers to the model number.

We assess the marginal predictive content of key financial indicators other than lagged GDP growth. The financial predictor set constitutes the term spread (TS), real stock returns (R), and the real short-term interest rate (ir). We first specify bivariate models for each financial predictor, as in Stock and Watson (2003a) (Models 2–4). Next, the models are augmented one by one with additional financial indicators until all the combinations of financial predictors are implemented. This process generates the following model specifications:

$$y_{t+4} = \alpha^2 + \sum_{j=1}^{p} \beta_{1j}^2 y_{t-j+1} + \beta_2^2 TS_t + u_{t+4}^2 \tag{2}$$

$$y_{t+4} = \alpha^3 + \sum_{j=1}^{p} \beta_{1j}^3 y_{t-j+1} + \beta_2^3 R_t + u_{t+4}^3 \tag{3}$$

$$y_{t+4} = \alpha^4 + \sum_{j=1}^{p} \beta_{1j}^4 y_{t-j+1} + \beta_2^4 ir_t + u_{t+4}^4 \tag{4}$$

$$y_{t+4} = \alpha^5 + \sum_{j=1}^{p} \beta_{1j}^5 y_{t-j+1} + \beta_2^5 TS_t + \beta_3^5 R_t + u_{t+4}^5 \tag{5}$$

$$y_{t+4} = \alpha^6 + \sum_{j=1}^{p} \beta_{1j}^6 y_{t-j+1} + \beta_2^6 TS_t + \beta_4^6 ir_t + u_{t+4}^6 \tag{6}$$

$$y_{t+4} = \alpha^7 + \sum_{j=1}^{p} \beta_{1j}^7 y_{t-j+1} + \beta_3^7 R_t + \beta_4^7 ir_t + u_{t+4}^7 \tag{7}$$

$$y_{t+4} = \alpha^8 + \sum_{j=1}^{p} \beta_{1j}^8 y_{t-j+1} + \beta_2^8 TS_t + \beta_3^8 R_t + \beta_4^8 ir_t + u_{t+4}^8 \tag{8}$$

In addition to the autoregressive behavior of GDP, the models relate future GDP growth to current observations of the financial predictors. This approach is motivated by the conventional assumption that all relevant information about the future stance of the economy is incorporated in the most recent observation of a financial indicator, i.e., the models do not include lagged values of the financial indicators.

The forecasting period runs from 2000Q1 to 2016Q2. The first half of the period constitutes the Great Moderation era, which was characterized by stable economic conditions. The second half comprises the financial crisis (the Great Recession) and its aftermath, including the subsequent sovereign debt crisis, the zero lower bound (ZLB), and the unconventional monetary policy era. Hence, it is obvious that the out-of-sample period is not uniform. The entire sample of the G-7 countries' growth rates is depicted in Fig. 1. The forecasting period is shaded. The dotted vertical line denotes 2008Q3, the quarter during which the Lehman Brothers bankruptcy occurred.

4. The data

The dataset for the G-7 countries comprises quarterly data from 1980Q1 to 2016Q2. Notably, Germany's time series describes West Germany until 1990Q4, after which time, the data describe the reunified Germany. GDP growth rates are calculated as logarithmic changes in real GDP indices. The term spread is conventionally defined as the difference between the ten-year government bond yield and the three-month interest rate. Real stock returns are obtained using logarithmic changes in real stock prices, which are obtained by dividing the nominal stock price index by the consumer price index (CPI). The real short-term interest rate is derived by subtracting annual CPI inflation from the nominal interest rate. Annual inflation is calculated using annual logarithmic changes in the consumer price index. All data were obtained from OECD databases. Table 1 presents a detailed description of the data and the data transformations.

The global financial crisis is generally considered the end of the Great Moderation era. The financial crisis was followed by the sovereign debt crisis, the unconventional monetary policy and the ZLB, among other things. Thus, the forecasting period is divided into two sub-periods: the Great Moderation era and the financial crisis era. Although the official onset of the financial crisis is somewhat vague[2], we follow convention and regard the Lehman Brothers bankruptcy as the starting point of the global financial crisis. Accordingly, we report the descriptive statistics of the data for the three time frames to gain better insight into possible changes in the data-generation process (DGP): the in-sample period (1981Q1–1999Q4), the Great Moderation era (2000Q1–2008Q3), and the financial crisis and its aftermath (2008Q4–2016Q2). The descriptive statistics of the data are presented in Table 2.

The descriptive statistics clearly demonstrate the exceptional economic circumstances during the financial crisis era, i.e., a substantial decline in growth rates combined with a marked increase in volatility of economic activity. In many G-7 countries, GDP growth more than halved, and economic activity declined to the lowest figures during the entire sample period. In Italy, for example, the realized growth was actually negative during the crisis subsample.

Regarding the financial predictors, short-term real interest rates decreased significantly during the crisis period, and, consequently, the term spreads increased. The dip in real short-term interest rates was so marked that the real rates turned negative in all G-7 countries except Japan. The exceptionally low real short-term interest rates were due to the ZLB and negligible inflation rates. The burst

[1] We also tested the 2- and 8-quarter forecasting horizons and found that the financial variables improved the forecasting performance in the G-7 countries. The results are available upon request.

[2] In the U.S., for example, August 2007 has been suggested as the starting point of the financial crisis (Mishkin, 2011a, 2011b); however, the NBER business cycle committee officially announced a recession in December 2008 (Ng & Wright, 2013).

[7] $\Delta^4 y = \ln\left(\frac{Y_t}{Y_{t-4}}\right) \times 100$.

Table 1
Data description.

Raw Data	*Details and Source of the Data*
Y =Real GDP	Volume index of gross domestic product – expenditure approach. Seasonally adjusted. Source: *OECD Quarterly National Accounts*.
$i3$ = Nominal short-term interest rate	Three-month interbank offer rate or three-month treasury bill, certificate of deposit or comparable instrument rate. Percent per annum. Source: OECD Monthly Monetary and Financial Statistics (MEI).
$i10$ = Nominal long-term interest rate	Ten-year government bond rate. Percent per annum. Source: *OECD Monthly Monetary and Financial Statistics* (MEI).
P = Consumer price index	Consumer price index – all items. Source: *OECD Consumer Prices* (MEI).
S = Share price index	National all-share or broad share price index. Average of monthly figures, which are averages of daily quotations. Source: *OECD Monthly Monetary and Financial Statistics* (MEI).
Variable	*Variable Construction*
Annual future GDP growth	$y_{t+4} = ln\left(y_{t+4}/y_t\right) \times 100$
Quarterly GDP growth	$y_t = ln\left(y_t/y_{t-1}\right) \times 100$
Term spread	$TS_t = i10_t - i3_t$
Annual inflation	$Inf_t = ln\left(P_t/P_{t-4}\right) \times 100$
Real short-term interest rate	$ir_t = i3_t - Inf_t$
Real quarterly stock returns	$R_t = ln\left[(S_t/P_t)/(S_{t-1}/P_{t-1})\right] \times 100$

Table 2
Descriptive statistics for the data.

	$\Delta^4 \boldsymbol{y}$			*TS*			*R*			*ir*		
Period	A	B	C	A	B	C	A	B	C	A	B	C
Canada												
Mean	2.11	1.67	0.81	0.61	1.10	1.38	0.19	0.02	−0.19	4.81	1.32	−0.48
Std.dev.	2.21	0.82	1.15	1.80	1.12	0.87	5.67	2.64	2.30	1.84	1.17	0.73
Min	−4.11	0.56	−2.38	−4.27	−0.59	0.41	−16.23	−6.40	−8.73	1.04	−1.39	−2.09
Max	5.92	3.80	2.10	3.35	3.33	3.12	22.70	6.04	4.47	9.15	3.32	1.18
ρ_1	0.88	0.77	0.84	0.84	0.91	0.89	0.40	0.28	0.20	0.80	0.75	0.68
France												
Mean	1.84	1.33	0.32	0.88	0.99	1.71	0.66	−0.46	−0.04	4.54	1.48	−0.27
Std.dev.	1.08	0.73	1.03	1.36	0.82	0.73	4.93	2.16	2.08	1.86	1.11	0.89
Min	−0.91	−0.04	−2.48	−4.14	−0.50	−0.31	−19.04	−6.33	−5.88	0.04	−0.28	−1.60
Max	4.19	3.06	1.76	2.90	2.22	2.82	12.84	4.11	3.45	9.70	3.45	2.47
ρ_1	0.86	0.77	0.85	0.83	0.85	0.62	0.37	0.43	0.23	0.84	0.91	0.75
Germany												
Mean	1.78	1.12	0.56	0.86	0.90	1.15	0.87	−0.36	0.09	3.56	1.66	−0.41
Std.dev.	1.54	1.12	1.99	1.53	0.85	0.74	3.97	3.18	2.52	1.39	0.85	0.92
Min	−1.59	−0.65	−4.69	−2.83	−0.72	−0.72	−11.82	−7.98	−7.74	0.94	0.18	−1.80
Max	5.82	3.27	3.61	3.16	2.14	2.51	12.13	6.98	4.60	7.51	3.37	2.61
ρ_1	0.78	0.85	0.83	0.94	0.85	0.66	0.33	0.41	0.22	0.86	0.86	0.72
Italy												
Mean	1.62	0.94	−0.67	0.10	1.19	3.19	−0.53	−0.78	0.02	5.41	0.98	−0.61
Std.dev.	1.30	0.96	1.75	1.32	0.77	1.20	3.82	3.72	4.49	1.94	0.94	1.00
Min	−1.02	−0.86	−4.99	−2.89	−0.20	0.45	−8.36	−7.32	−13.08	0.57	−0.57	−2.76
Max	4.09	2.98	1.50	3.21	2.38	5.33	8.19	8.86	11.57	11.50	2.88	1.46
ρ_1	0.80	0.78	0.86	0.80	0.86	0.78	0.36	0.33	0.28	0.79	0.93	0.84
Japan												
Mean	2.29	0.86	0.11	0.66	1.17	0.61	0.75	−0.15	0.27	2.77	0.45	0.01
Std.dev.	1.99	0.69	2.00	1.19	0.33	0.30	3.91	3.70	4.41	1.90	0.55	1.40
Min	−1.60	−1.02	−5.77	−3.67	0.47	−0.07	−7.97	−6.13	−12.84	−1.57	−1.20	−3.40
Max	6.94	2.27	3.50	2.63	1.68	1.08	8.84	9.55	12.01	5.72	1.52	2.63
ρ_1	0.83	0.68	0.74	0.87	0.75	0.85	0.38	0.30	0.24	0.94	0.65	0.86
U.K.												
Mean	2.14	1.67	0.54	−0.23	−0.12	1.86	0.52	−0.58	−0.33	5.24	3.19	−1.42
Std.dev.	1.68	0.72	1.51	1.80	0.75	0.85	3.02	1.65	1.67	1.81	1.05	1.34
Min	−3.39	−0.84	−3.70	−4.57	−1.52	−0.44	−9.96	−5.36	−6.05	1.36	1.15	−3.67
Max	5.74	3.33	1.96	3.23	1.10	3.45	8.14	2.00	2.54	9.69	5.55	0.87
ρ_1	0.81	0.54	0.86	0.89	0.90	0.68	−0.03	0.31	0.14	0.82	0.75	0.83
U.S.												
Mean	2.60	1.42	0.69	0.75	1.08	1.97	0.61	−0.54	−0.01	3.85	0.66	−0.76
Std.dev.	1.80	0.77	1.09	1.67	1.49	0.74	3.83	1.72	1.81	2.03	1.66	1.42
Min	−2.63	−0.17	−2.27	−4.21	−1.06	−0.19	−10.80	−4.99	−5.65	0.07	−1.97	−3.29
Max	7.89	3.16	1.72	3.39	3.34	3.31	17.59	2.89	3.12	8.51	3.41	2.34
ρ_1	0.86	0.76	0.83	0.82	0.93	0.64	0.18	0.28	0.28	0.91	0.83	0.77

Notes: A = 1981Q1 — 1999Q4; B = 2000Q1 — 2008Q3; C = 2008Q4 — 2016Q2; $\Delta^4 y$ = annual GDP growth[7]; TS = term spread; *R* = quarterly real stock returns; *ir* = real short-term interest rate; *Std.dev.* = standard deviation; ρ_1 = first-order autocorrelation coefficient.

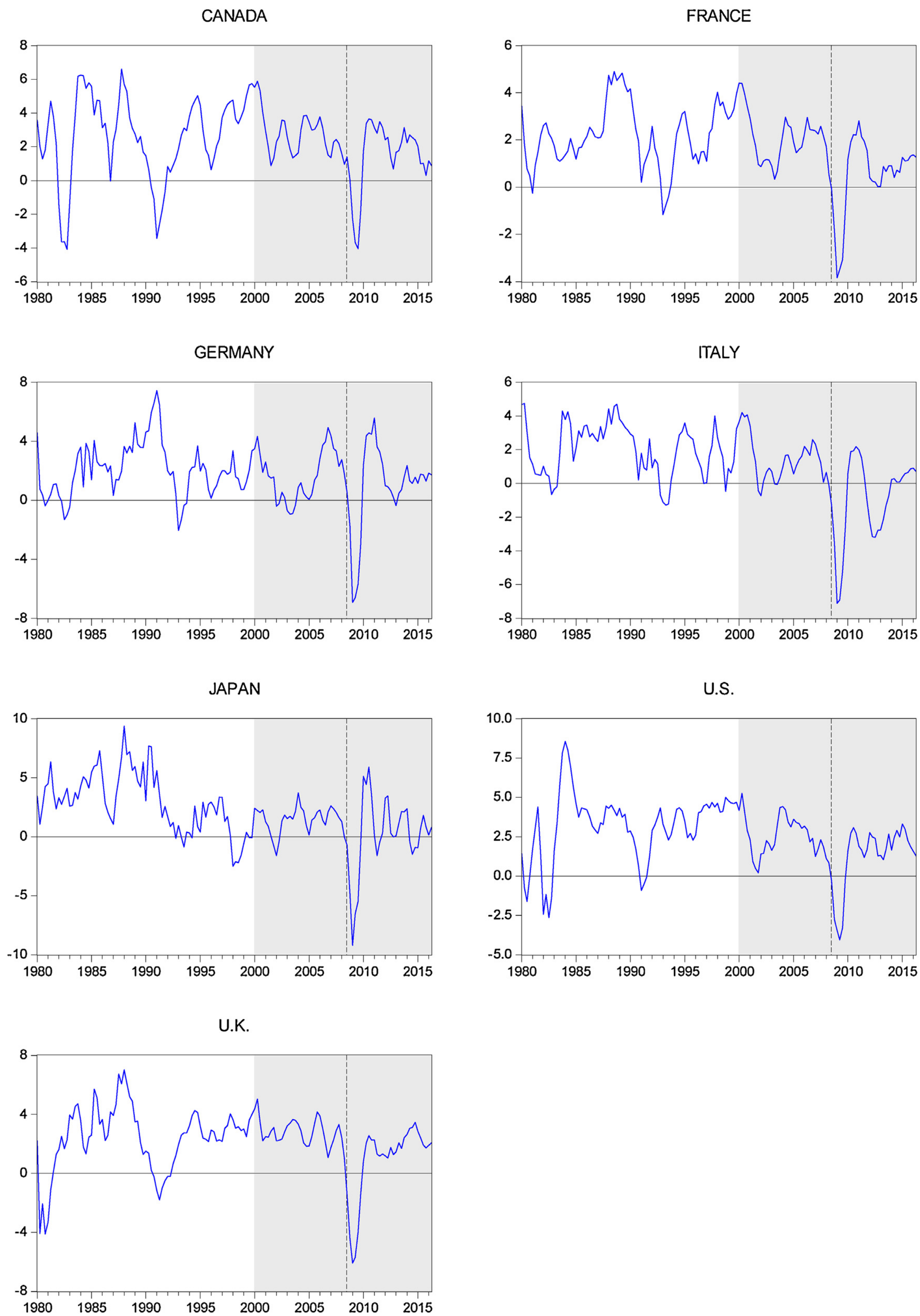

Fig. 1. Annual GDP growth in the G-7 countries. The forecasting period is shaded. The dotted vertical line indicates 2008Q3.

of the techno bubble at the beginning of the 2000s and the early stages of the financial crisis were reflected in low real stock returns during the first half of the forecasting period (2000Q1–2008Q3). During the second half (2008Q4–2016Q2), the stock markets generally bounced back, although deep dips due to the financial and sovereign debt crises pushed down the average returns.

The behavior of the term spreads is particularly interesting during the crisis period, given that the nominal short-term rates were stuck at the ZLB in the G-7 countries and that many central banks launched unconventional asset purchase programs (quantitative easing) to bring down long-term interest rates. Conventionally, as the term spread increases, the expected future GDP growth will increase. Alternatively, inversion of the term spread has traditionally preceded a recession (e.g., Wheelock & Wohar, 2009). The term spreads increased notably in France, Italy, U.K. and the U.S. during the crisis period. According to the conventional interpretation, this increase in the term spread is a precursor for a strengthening economy rather than a contracting one. Additionally, negative term spreads that traditionally precede recessions were not detected in Canada and Italy. This finding suggests that during unconventional monetary policy and quantitative easing programs, the term spread may deliver misleading signals regarding economic activity (Ng & Wright, 2013: 1137–1138). Therefore, it is reasonable to use a number of financial predictors, especially during periods in which an unconventional monetary policy is in place. Finally, relatively high first-order autocorrelation coefficients imply a high degree of persistence in the data-generation process for all the time series, excluding real stock returns. Moreover, all the coefficients are lower than unity, implying that the data are stationary[3].

5. In-sample fit and out-of-sample forecast performance

5.1. In-sample analysis

Before considering the out-of-sample forecasting results, it is interesting to scrutinize how well the models fit the data within the entire sample (1981Q1–2016Q2). Given the number of models and countries, it is not feasible to report all the parameter estimates and their significance. Therefore, we focus on the explanatory power and the overall significance of all the predictors. More specifically, we explicitly consider the joint significance of (a) all the predictors (AR terms and financial predictors) and (b) the joint significance of only the financial predictors. In this manner, we gain information about the role of the financial predictors in the models (in-sample). Throughout the study, the models are estimated using OLS with heteroscedasticity and autocorrelation-robust Newey-West standard errors. The number of AR terms is based on the Schwartz information criterion. The in-sample results are presented in Table 3.

Overall, the in-sample analysis suggests that the financial indicators improve the explanatory power of the models. Compared to the simple AR models, financial indicators improve the model performance the most in Canada, Japan, and the U.S., while the improvement is the lowest in Germany. In France, only two of the seven financial models are significant, based on *F*-tests. Finally, the explanatory power of the AR models is notably low in Canada, Germany, Japan, and the U.S., although all models are still significant.

Altogether, the in-sample results lend support for the use of financial predictors to forecast economic activity. However, it is well documented in the previous literature (e.g., Stock & Watson, 2003a) that a good in-sample performance does not guarantee a good out-of-sample forecast performance.

5.2. Out-of-sample forecasting analysis

The forecasting analysis is conducted recursively outside the initial estimation period (1981Q1–1999Q4): when a new observation is received, the model is re-estimated, which in turn produces a new forecast of the future GDP growth over four quarters. Hence, this pseudo out-of-sample analysis by Stock and Watson (2003a) resembles the actual forecasting situation in the sense that it uses all information when the actual forecast is calculated. The forecasting performance is conventionally evaluated based on the root mean square error (RMSE). As the model's RMSE decreases, the forecasting performance improves.

The fundamental question is whether financial predictors significantly lower forecast errors relative to the AR benchmark model (Model 1). If this is the case, the models are ranked according to the RMSEs, and finally, a test is performed to determine whether the RMSEs differ statistically from each other. This situation is formally tested using the Clark and McCracken (2001) test whereby the forecasting models are nested (e.g., Models 2–8 nest Model 1). If the compared models are not nested, the Diebold and Mariano (1995) test is applied. The results of the forecasting analysis are presented in Table 4.

The forecasting analysis yields several interesting outcomes. First, in all the G-7 countries, the forecasting performance is unambiguously improved by including financial predictors in the forecasting models. Compared to the AR benchmark, even a single financial predictor is sufficient to significantly improve the forecasting performance in all the G-7 countries except France. However, the optimal model choice is not the same for all the countries. Should a forecaster select a single financial predictor, the correct choice would be real stock returns for Canada and the U.K., the real short-term interest rate for Italy, Japan and the U.S., and the term spread for Germany. The predictive power of short-term interest rates is consistent with Ang et al. (2006). The relatively weak performance of the term spread appears noteworthy, given its previous dominance as the most useful financial predictor for economic activity (e.g., Stock & Watson, 2003a; Estrella, 2005).

Second, in contrast to the seminal results of Stock and Watson (2003a), the best forecasting results are consistently obtained using several financial predictors. In four out of seven countries (Canada, France, Germany, and the U.K.), the lowest RMSEs are captured by the model specification that contains all three key financial predictors (Model 8). In Italy and Japan, the combination of real stock returns and the real short-term interest rate is the best choice (Model 7), whereas in the U.S., the preferred selection contains the term spread and the real short-term interest rate (Model 6). Altogether, financial predictors yield significant improvements in forecasting accuracy in 38 out of 49 cases. In summary, the results demonstrate that the use of several financial predictors is favorable for forecasting GDP growth, although no single dominant model specification exists for all the G-7 countries. The lack of robustness in the predictive power of financial predictors is consistent with the classic results of Stock and Watson (2003a).

A closer inspection of the RMSE figures in Table 4 reveals that several RMSEs are very similar to each other. Time series econometrics typically favors less parameterized models; hence, it is prudent to formally test whether the RMSEs actually differ from each other. More specifically, the lowest RMSE among model specifications 2–8 is tested against the second-lowest RMSE among the less parameterized model specifications. The significant test statistic confirms that the RMSE is actually the lowest, whereas the insignificant test outcome suggests that the less parameterized model should be preferred. Table 5 presents the test results.

The test results suggest that the lowest RMSEs in Table 4 are actually the lowest in all cases except France. In France, the RMSEs of Models 6 and 8 do not differ statistically from each other,

[3] The time series properties of the data were also tested formally using unit root tests. The results suggested that all the data are stationary, excluding the short-term real interest rates. However, by allowing for breaks and shifts in the DPG process, the real short-term interest rates were stationary. Given the ambiguity of the unit root test results, we carried out all the forecasting analyses by considering short-term real interest rates both in levels and in first differences. The level specification unambiguously yielded the lowest RMSEs. Hence, we report only the level specification for the short rates (see also Cohrane, 1991). The full unit root test results are available upon request.

Table 3
Summary of in-sample results (1981Q1–2016Q2).

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Canada								
$\bar{R}^2$	0.082	0.274	0.245	0.090	0.384	0.402	0.262	0.515
$Prob_1$	0.003	0.000	0.000	0.004	0.000	0.000	0.000	0.000
$Prob_2$		0.000	0.000	0.832	0.000	0.000	0.000	0.000
France								
$\bar{R}^2$	0.177	0.213	0.206	0.227	0.235	0.392	0.237	0.392
$Prob_1$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$Prob_2$		0.112	0.201	0.181	0.131	0.002	0.172	0.004
Germany								
$\bar{R}^2$	0.031	0.126	0.088	0.043	0.152	0.215	0.109	0.245
$Prob_1$	0.024	0.000	0.021	0.059	0.001	0.001	0.035	0.001
$Prob_2$		0.025	0.029	0.298	0.009	0.002	0.042	0.002
Italy								
$\bar{R}^2$	0.139	0.183	0.211	0.226	0.275	0.333	0.221	0.328
$Prob_1$	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
$Prob_2$		0.033	0.000	0.008	0.000	0.000	0.020	0.000
Japan								
$\bar{R}^2$	0.037	0.078	0.102	0.402	0.149	0.413	0.462	0.468
$Prob_1$	0.013	0.002	0.006	0.000	0.000	0.000	0.000	0.000
$Prob_2$		0.028	0.006	0.000	0.000	0.000	0.000	0.000
U.K.								
$\bar{R}^2$	0.195	0.230	0.227	0.164	0.257	0.354	0.229	0.400
$Prob_1$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$Prob_2$		0.075	0.008	0.317	0.009	0.001	0.006	0.000
U.S.								
$\bar{R}^2$	0.083	0.129	0.257	0.206	0.293	0.514	0.379	0.615
$Prob_1$	0.029	0.015	0.000	0.000	0.000	0.000	0.000	0.000
$Prob_2$		0.064	0.000	0.079	0.000	0.000	0.000	0.000

Notes: $Prob_1$ = P-value for the F-test statistics (H_0: all parameter estimates excluding the constant term are zero). $Prob_2$ = P-value for the F-test statistics of the null hypothesis that all parameter estimates for the financial predictors are zero.

Table 4
Forecasting results for the entire forecasting period (2000Q1–2016Q2).

Model specification	Canada	France	Germany	Italy	Japan	U.K.	U.S.
(1) AR	1.210	1.096	1.697	1.730	1.972	1.352	1.373
(2) AR + TS	1.349	1.100	1.608***	1.660***	1.943**	1.429	1.467
(3) AR + R	1.146***	1.087	1.613***	1.686***	1.958***	1.310***	1.292***
(4) AR + ir	1.648	1.111	1.688*	1.588***	1.479***	1.395	1.133***
(5) AR + TS + R	1.266	1.095*	1.560***	1.577***	1.919***	1.383	1.369***
(6) AR + TS + ir	1.164***	0.882***	1.508***	1.647***	1.516***	1.132***	0.852***
(7) AR + R + ir	1.399	1.093*	1.578***	1.487***	1.406***	1.337**	1.109***
(8) AR + TS + R + ir	1.006***	0.876***	1.453***	1.516***	1.432***	1.098***	0.880***

Notes: Figures in the columns are the RMSEs of the corresponding forecasting model specifications given in column one. Asterisks refer to significance levels for the Clark and McCracken (2001) test: *** = 1%, ** = 5%, * = 10%. The null hypothesis is that the RMSE does not differ significantly from the RMSE of the benchmark AR model (Model 1).

Table 5
Test results for the equality of RMSEs for 2000Q1–2016Q2.

Country	Null Hypothesis	Test Statistic	Best Model
Canada	RMSE(3) = RMSE(8)	29.69***	8
France	RMSE(6) = RMSE(8)	0.51	6
	RMSE(3) = RMSE (6)	1.743**DM	
Germany	RMSE(6) = RMSE(8)	3.19***	8
Italy	RMSE(4) = RMSE(7)	10.42***	7
Japan	RMSE(4) = RMSE(7)	8.27***	7
U.K.	RMSE(6) = RMSE(8)	3.50***	8
U.S.	RMSE(4) = RMSE(6)	65.55***	6

Notes: The figures in parentheses in column 2 refer to the model specification. The test statistic in column three is the Clark and McCracken (2001) test statistic for nested models or the Diebold and Mariano (1995) test for non-nested models. The null hypothesis is that the RMSE of the more parsimonious model does not differ significantly from the RMSE of the less parsimonious nested model. Superscript DM refers to Diebold and Mariano test. The rejection of the null implies that the RMSE of the richly parameterized model is preferred. Significance levels: *** = 1%, ** = 5%, * = 10%. The model in column four refers to the preferred model specification.

although the RMSE for Model 8 (0.876) is marginally lower than that of Model 6 (0.882)[4]. Given the negligible difference between the RMSEs, the test outcome is expected. Hence, the more parsimonious Model 6 is preferred for France.

The best forecasting models are summarized in the last column of Table 5. It is notable that in all the G-7 countries, more than one financial predictor is required to achieve the lowest RMSEs when forecasting economic activity. The preferred models reveal that the real short-term interest rate is included in all the selected predictor sets for the G-7 countries. Previously, Ang et al. (2006) found that the predictive power of the short-term interest rate was greater than that of any interest spreads in the U.S. Hence, our results extend the importance of the short-term interest rate

[4] In this case, it was necessary to conduct an additional test for equality of RMSEs between Model 6 (0.779) and the next lowest RMSE (0.811 for Model 3). Note that the models are now no longer nested. Therefore, the Diebold and Mariano (1995) test should be applied in this case (cf. Table 5).

Table 6
Out-of-sample forecasting results for the pre-financial crisis and financial crisis and its aftermath periods.

(a) The pre-financial crisis era (2000Q1–2008Q3).

Model specification	Canada	France	Germany	Italy	Japan	U.K.	U.S.
(1) AR	0.919	0.810	1.201	1.070	1.283	1.008	1.211
(2) AR + TS	1.201	0.822	1.269	1.042**	1.114***	0.975***	1.250
(3) AR + R	0.957	0.811	1.146**	1.011***	1.360	0.964***	1.151***
(4) AR + ir	1.659	0.933	1.213	1.042**	0.857***	1.091	1.097***
(5) AR + TS + R	1.128	0.829	1.227	0.960**	1.184***	0.927***	1.142***
(6) AR + TS + ir	1.219	0.779***	1.239	1.118	0.905***	0.786***	0.809***
(7) AR + R + ir	1.414	0.918	1.126***	0.961***	0.828***	1.059	1.104***
(8) AR + TS + R + ir	1.028	0.772***	1.187***	0.995***	0.859***	0.772***	0.813***

(b) The financial crisis and its aftermath era (2008Q4–2016Q2).

Model specification	Canada	France	Germany	Italy	Japan	U.K.	U.S.
(1) AR	1.470	1.348	2.123	2.254	2.533	1.656	1.535
(2) AR + TS	1.500	1.346**	1.920***	2.155***	2.576	1.810	1.679
(3) AR + R	1.328***	1.331	2.014***	2.212***	2.464***	1.614**	1.434***
(4) AR + ir	1.637	1.283***	2.099*	2.036***	1.956***	1.673	1.172***
(5) AR + TS + R	1.405***	1.333**	1.865***	2.063***	2.501**	1.761	1.586
(6) AR + TS + ir	1.100***	0.986***	1.763***	2.089***	1.991***	1.424***	0.898***
(7) AR + R + ir	1.383***	1.261***	1.967***	1.915***	1.853***	1.594**	1.115***
(8) AR + TS + R + ir	0.981***	0.981***	1.703***	1.943***	1.879***	1.377***	0.951***

Note: Significance levels: *** = 1%, ** = 5%, * = 10%.

in forecasting economic activity beyond the U.S. Moreover, term spreads and (real) stock returns – representing a more traditional financial predictor set – were not selected as the preferred predictor combination.

Given that the entire forecasting period includes the Great Moderation and the turbulent financial crisis eras, it is of interest to evaluate the forecasting performance during both sub-periods. Moreover, previous studies suggest that the predictive content of financial indicators has increased since the end of the Great Moderation (e.g., Ng & Wright, 2013; Kuosmanen & Vataja, 2014; Chinn & Kucko, 2015; Kuosmanen et al., 2015). The forecasting results for the sub-periods are shown in Table 6.

Panels (a) and (b) in Table 6 present the RMSEs for the Great Moderation era and the turbulent era, respectively. Comparing the crisis subsample with the Great Moderation reveals that the marginal predictive content of the financial predictors increases markedly during the crisis period: in the Great Moderation subsample, the financial predictors improve the forecasting power in 28 out of 49 cases, whereas in the crisis subsample, the corresponding figure is 40 out of 49 cases. It is also noteworthy that the RMSEs of the models with a single financial predictor display more instability between the sub-periods compared to the more richly parameterized models.

Moreover, the forecasting content of financial predictors vanishes during the Great Moderation in Canada but plays a significant role during the crisis period. The lowest RMSEs are again obtained using several financial predictors; however, the preferred predictor sets are not uniform for all the countries. As expected, the forecast errors tended to increase during the crisis period; in particular, the RMSEs increased in Germany, Italy, Japan, and the U.K. Finally, the differences between the RMSEs for the sub-periods are tested in Table 7.

The test results for the Great Moderation period demonstrate that the best models vary between the countries; however, the richly parameterized models outperform the parsimonious ones. Note that in the case of France, the difference between the RMSEs of Models 8 and 6 is insignificant in both subsamples; hence, the less parameterized Model 6 is preferred. The results for Germany and Italy are different from those for the other countries in the sense that the preferred model changes between the sub-periods. In the other G-7 countries, the preferred models do not change between the Great Moderation and financial crisis periods.

Table 7
Test results for equality of RMSEs for the sub-periods.

(a) Great Moderation era (2000Q1–2008Q3).

Country	Null Hypothesis	Test Statistic	Best Model
Canada	–	–	–
France	RMSE(6) = RMSE(8)	0.350	6
Germany	RMSE(3) = RMSE(7)	1.615**	7
Italy	RMSE(3) = RMSE(5)	1.338**	5
Japan	RMSE(4) = RMSE(7)	7.357***	7
U.K.	RMSE(6) = RMSE(8)	2.233**	8
U.S.	RMSE(4) = RMSE(6)	44.406***	6

(b) Financial crisis and its aftermath (2008Q4–2016Q2).

Country	Null Hypothesis	Test Statistic	Best Model
Canada	RMSE(6) = RMSE(8)	6.037***	8
France	RMSE(6) = RMSE(8)	0.190	6
	RMSE(4) = RMSE(6)	14.041***	
Germany	RMSE(6) = RMSE(8)	1.346**	8
Italy	RMSE(4) = RMSE(7)	4.18***	7
Japan	RMSE(4) = RMSE(7)	3.288***	7
U.K.	RMSE(6) = RMSE(8)	1.528**	8
U.S.	RMSE(4) = RMSE(6)	22.970***	6

Note: Significance levels: *** = 1%, ** = 5%, * = 10%.

The improvement in forecasting performance relative to the AR benchmark is presented in Table 8. Overall, the forecasting performance improves markedly during the financial crisis era, with an average improvement of 25 percent. However, during the Great Moderation, the improvement is clearly smaller, i.e., 16 percent on average. During the Great Moderation, the forecasting performance is distinctly twofold: the improvement is clear in Japan, the U.K., and the U.S., whereas in France, Germany, and Italy, the improvement is modest.

6. Time variation in forecast errors

The RMSE is likely the most commonly used measure for forecast performance. However, the RMSE is not an appropriate measure to assess the development of forecast errors over time. The RMSE is, by

Table 8
Percentage improvements in forecasting performance.

	Canada	France	Germany	Italy	Japan	U.K.	U.S.	Mean
2000Q1 – 2016Q2	17 (8)	20 (6)	14 (8)	14 (7)	29 (7)	19 (8)	19 (6)	19
2000Q1 – 2008Q3	–	4 (6)	6 (7)	8 (5)	35 (7)	23 (8)	33 (6)	16
2008Q4 – 2016Q2	33 (8)	27 (6)	20 (8)	15 (7)	26 (7)	17 (8)	38 (6)	25

Notes: Columns 2–7 present the percentage improvement in RMSEs for the lowest RMSE and the RMSE of the benchmark model (Model 1) for each country. The best model specification in terms of RMSE is given in parentheses. Column 8 presents the mean improvement in forecasting performance for all the G-7 countries.

Table 9
Descriptive statistics for the forecast error spreads (2000Q1–2016Q2).

	Canada	France	Germany	Italy	Japan	U.K.	U.S.
Mean	0.17	0.15	0.10	0.25	0.42	0.16	0.39
Med	0.20	0.16	0.23	0.17	0.58	0.15	0.33
Max	1.79	0.96	1.61	1.38	2.14	1.13	2.26
Min	−1.46	−0.46	−1.22	−1.34	−2.54	−1.21	−1.45
Std. Dev.	0.71	0.32	0.74	0.52	1.02	0.56	0.86
J-B. (Prop.)	0.25 (0.88)	1.59 (0.45)	1.94 (0.38)	1.47 (0.48)	4.61 (0.10)	1.30 (0.52)	0.41 (0.82)
ρ_1	0.46	0.41	0.70	0.28	0.31	0.24	0.52

Notes: J-B = Jarque-Bera test; H_0: the variable is normally distributed. Prob. = Probability value for the Jarque-Bera test statistic. ρ_1 = first-order autocorrelation coefficient.

definition, constructed to express the *average* behavior of forecast errors during the forecasting period rather than the exact behavior of individual forecast errors in a distinct time frame. Therefore, to evaluate the behavior of forecast errors over time, it is more appropriate to consider the behavior of the absolute forecast errors.

The forecasting results from the entire forecasting period demonstrate that the financial indicators clearly improve forecasting performance compared to the AR benchmark model. To study the intertemporal behavior of the forecast errors over time, we use similar approach as Ng and Wright (2013) and Hännikäinen (2017). We define the forecast error spread (Q_t) as the difference between the root squared forecast errors of the benchmark model (r_1) and the best financial indicators model $\left(r_i^*\right)$.

$$Q_t = r_{1,t} - r_{i,t}^* \tag{9}$$

$$Q_t = \sqrt{\left(\Delta^4 lny_t - \Delta^4 ln\hat{y}_{Benchmark,t}\right)^2} - \sqrt{\left(\Delta^4 lny_t - \Delta^4 ln\hat{y}_{Model\,i,t}\right)^2} \tag{9.1}$$

The forecast error spreadsare straightforward to interpret: as the spread becomes more positive (negative), the financial model's forecast improves (worsens) compared to the AR benchmark. Fig. 2 plots the forecast error spreads. The vertical dotted line in 2008Q3 divides the entire forecasting period into the Great Moderation and the financial crisis sub-periods. Table 9 presents the summary statistics of the error spreads.

A visual examination of Fig. 2 clearly indicates that the error spreads contain a considerable amount of temporal dominance and time persistence. All first-order autocorrelation coefficients are positive (Table 9): the history dependence is highest for Germany (0.70) and lowest for the U.K. (0.24). Moreover, the majority (19/21) of the pairwise correlations of the forecast error spreads are positively associated (Table 10). One of the striking findings is the distinctive concentration of positive values of the error spreads during the financial crisis (2008–2010) in the G-7 countries, implying that the financial predictors systematically improved the forecasting performance during the crisis.

The positive means of the error spreads (Table 9) lend further support for the forecasting content of the financial predictors. To extend the analysis beyond the mean values, we now calculate the absolute sums of the positive and negative values of the spreads, as well as their relative shares of the total absolute sum. Furthermore, we summarize the forecast errors spreads of all the G-7 countries

Table 10
Correlations of the forecast error spreads (2000Q1–2016Q2).

	Canada	France	Germany	Italy	Japan	U.K.	U.S.
Canada	1.00						
France	0.45***	1.00					
Germany	0.14	0.17	1.00				
Italy	0.07	0.21*	−0.22*	1.00			
Japan	0.17	0.16	0.31**	0.04	1.00		
U.K.	0.23*	0.20	0.03	0.01	0.16	1.00	
U.S.	0.51***	0.34**	0.22*	−0.07	0.29**	0.37***	1.00

Note: Significance levels: *** = 1%, ** = 5%, * = 10%.

to construct an "aggregate G-7″ forecast error spread and thereby uncover the composite predictive power of the financial models (Table 11; Fig. 3).

Table 11 shows that the share of the positive sum of the error spreads is larger than 50% in all cases. The positive share is the highest for Italy (78%), France (77%) and the U.S. (77%) and the lowest for Germany (58%). The positive share for the aggregate G-7 forecast error spread in the last column of Table 11 (82%) is even higher. The remarkably high share for the aggregate forecast error spread is consistent with the uniform behavior of the country-specific forecast error spreads.

A visual inspection of the aggregate G-7 forecast error spread (Fig. 3) clarifies that the positive values of the errors spreads unambiguously dominate, providing clear support for the forecasting content of financial predictors for the G-7 countries as a whole. Moreover, at least three distinctive concentrations of predictive ability emerge[5]. The first positive concentration (2000Q2–2002Q3) connects to the burst of the techno bubble in the stock market and the impending 2001 recession in the U.S. and other Western countries (Stock & Watson, 2003b). The second (2007Q4–2010Q1) relates unambiguously to the global financial crisis. Finally, the third concentration (2011Q1–2013Q3) obviously coincides with the worsening of the sovereign debt and banking crises in the Eurozone (Davis, 2016). Similar concentrations of the positive forecast error spreads can also be detected from Fig. 4, which illustrates the stacked binary values[6] of positive forecast error spreads. Note that when the stacked binary variable is equal to or greater than four, the

[5] Distinctive concentration is defined here as at smallest four subsequent positive values of the forecast error spread.

[6] The binary variable is equal to one when the forecast error spread is positive, and zero otherwise.

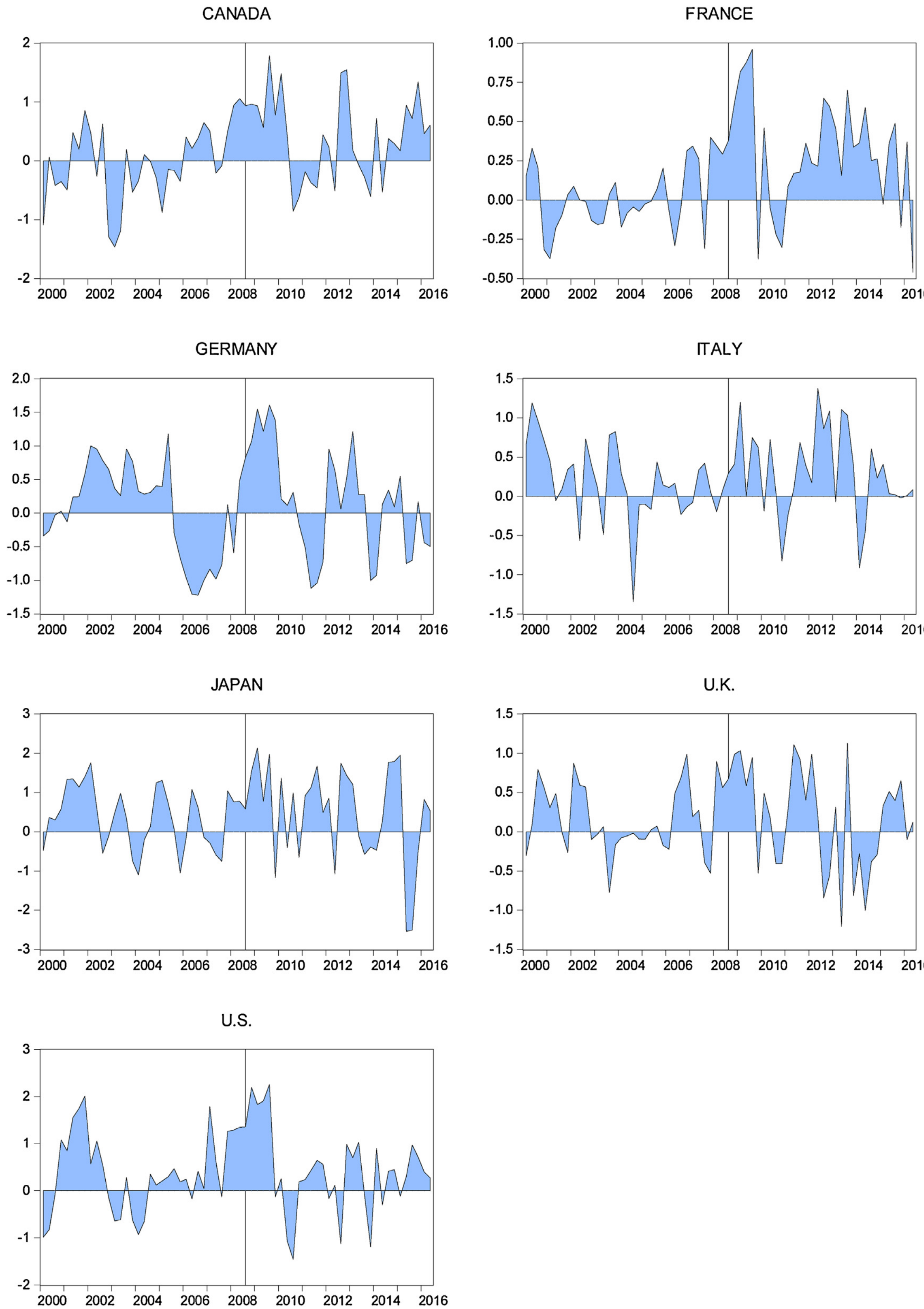

Fig. 2. Forecast error spreads for the entire forecasting period 2000Q1–2016Q2. The dotted vertical line is for 2008Q3.

financial predictors are useful in predicting economic activity in the majority of the G-7 countries. If forecast error spreads are simultaneously positive for all the G-7 countries, the stacked binary value is equal to seven. Clear similarities between the aggregated and stacked forecast errors spreads (Figs. 3 and 4) demonstrate that the contemporary forecasting content of financial predictors is based on several countries instead of only a few individual countries.

In summary, the behavior of the forecast error spreads demonstrates that there exists contemporaneous periods when financial predictors have significant predictive power in the G-7 countries. These periods are obviously connected to unsettled economic conditions. The results are obtained using linear models. The detected systematic temporal variation and time persistence in the predictive ability of financial variables indicate that we should

Table 11
Behavior of the forecast error spreads.

	Canada	France	Germany	Italy	Japan	U.K.	U.S.	G-7
Sum of positive	25.06	14.14	23.82	22.37	44.29	20.75	37.47	138.00
Sum of negative (absolute value)	14.06	4.14	17.18	6.16	16.55	10.11	11.44	29.77
Total sum	39.12	18.28	41.00	28.53	60.84	30.86	48.91	167.76
Share (%) of positive	64	77	58	78	73	67	77	82
Share (%) of negative	36	23	42	22	27	33	23	18

Notes: *Sum of positive* (*negative*) refers to the sum of positive (negative) values of the error spreads. *Share* (%) of *positive* (negative) refers to the percentage share of the positive and negative sums of the total sum of the error spread values in absolute terms.

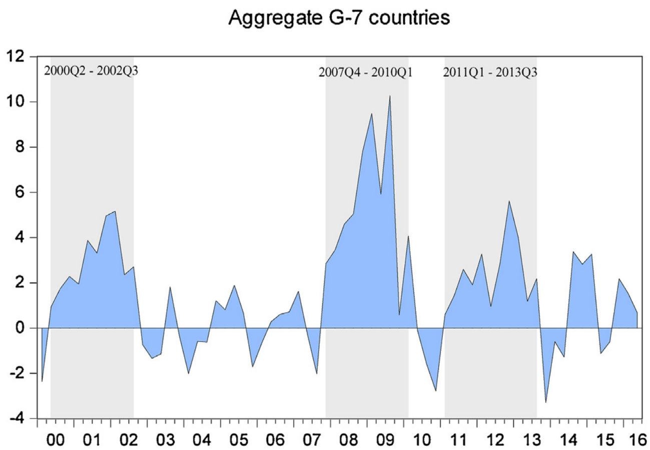

Fig. 3. Aggregated forecast error spread for the G-7 countries.

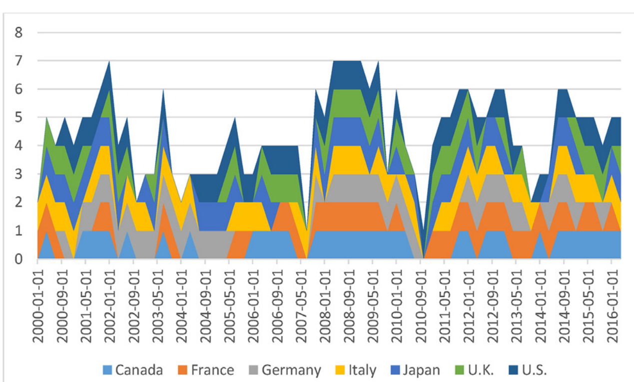

Fig. 4. Stacked binary values of positive forecast error spreads.

study the various economic circumstances that affect the predictive association between financial markets and real economy in greater depth. This implication indicates that future research should search for sources of systematic changes, asymmetric behavior and breakpoints in the predictive content of financial variables. In this context, the next step would be to define workable regime-switching signals and, ultimately, to construct nonlinear regime-switching models for forecasting economic activity. Several studies already suggest that the predictive relation between financial market and real activity is often nonlinear (e.g., Galbraith & Tkacz, 2000; Ahrens, 2002; Henry et al., 2004; Kuosmanen et al., 2015); however, previous literature on nonlinear models suggests that these models fit well in-sample but fail in forecasting out-of-sample (e.g., Teräsvirta, 2006). To date, the selection of regime-switching signals has been implemented largely on an ad hoc basis due to an inadequate understanding of the time-varying processes in the predictive content. Consequently, a cutting-edge topic for future research is to capture the nonlinear association between the financial markets and the real economy in order to improve the out-of-sample forecasting performance.

7. Conclusions

In contrast to the previous literature, the forecasting results of this study demonstrate that the key financial indicators – the term spread, real stock returns and the real short-term interest rate – have useful predictive content for economic activity under varying economic circumstances in the G-7 countries in the 2000s. The predictive content emerges both during the Great Moderation and during the financial crisis and its aftermath, even though the marginal predictive content of these financial predictors increases during the crisis. Moreover, it is generally preferable to use several financial predictors to forecast GDP growth.

The behavior of the forecast errors reveals that the predictive power of the financial indicators is history dependent and often varies similarly across the G-7 countries over time. The increased forecasting content is obviously connected to unsettled economic conditions. This finding is in line with Chinn and Kucko (2015), who suggested that the enhanced predictive content of financial indicators is related to the increased volatility of economic activity. Furthermore, our results stress the importance of considering actual, time-specific forecast errors. Paying attention to only the average behavior of the forecast errors, e.g., RMSEs, may mask useful information about predictive connections between the financial sector and the real economy. This aspect has been overlooked in the previous literature.

The results of our study are of important for economists because they provide guidance for understanding the workings of financial markets in developed economics and gradually confirm new stylized facts concerning the forecasting of economic activity. According to our results, it is evident that economists should have financial variables in their toolbox, especially during turbulent economic times when forecasting economic activity is overall difficult and the need for better forecasts is most evident. Correspondingly, as Siegel (2014: 230) stressed, if investors can predict the business cycle, they can easily beat a buy-and-hold strategy and, consequently, reap substantial gains by investing in stocks. Our study indicates that investors should not ignore the association between financial markets and real activity when planning their investment portfolios.

Conflicts of interest

None.

References

Ahrens, R. (2002). Predicting recessions with interest rate spreads: A multicountry regime-switching analysis. *Journal of International Money and Finance*, *21*(4), 519–537.

Ang, A., Piazzesi, M., & Wei, M. (2006). What does the Yield Curve tell us about GDP growth? *Journal of Econometrics*, *131*, 359–403.

Benati, L., & Goodhart, C. (2008). Investigating time-variation in the marginal predictive power of the yield spread. *Journal of Economic Dynamics & Control*, *32*, 1236–1272.
Binswanger, M. (2000). Stock market booms and real economic activity: Is this time different? *International Review of Economics & Finance*, *9*, 387–415.
Binswanger, M. (2004). Stock returns and real activity in the G-7 countries: did the relationship change during the 1980s? *The Quarterly Review of Economics and Finance*, *44*, 237–252.
Bordo, M. D., & Haubrich, J. G. (2008). Forecasting with the yield curve; level, slope and output 1875–1997. *Economics Letters*, *99*, 48–50.
Chinn, M. D., & Kucko, K. J. (2015). The predictive power of the yield curve across countries and time. *International Finance*, *18*(2), 129–156.
Clark, T. E., & McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, *105*, 85–110.
Cohrane, J. (1991). What macroeconomists should know about unit roots: Comment. In O. Blanchard, & S. Fisher (Eds.), *NBER macroeconomics annual, 6* (pp. 201–210). University of Chicago Press.
D'Agostino, A., Giannone, D., & Surico, P. (2006). *(Un)predictability and macroeconomic stability. (working paper series. No 605, European Central Bank)* April.
Davis, S. J. (2016). *An index of global economic policy uncertainty (NBER working paper No 22740)*.
Diebold, F., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*, 253–263.
Domian, D. L., & Louton, D. A. (1997). A threshold autoregressive analysis of stock returns and real economic activity. *International Review of Economics & Finance*, *6*(2), 167–179.
Dotsey, M. (1998). The predictive content of the interest rate term spread for future economic growth. *Federal Reserve Bank of Richmond Economic Quarterly*, *84*(3), 31–51.
Estrella, A. (2005). *The yield curve as a leading indicator: Frequently asked questions* Available online at:. http://www.ny.frb.org/research/capital_markets/ycfaq.pdf
Estrella, A., & Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *The Journal of Finance*, *46*(2), 555–576.
Estrella, A., & Mishkin, F. S. (1995). *Predicting U.S. recessions: Financial variables as leading indicators (NBER Working Paper 5379)* December.
Estrella, A., & Mishkin, F. S. (1997). The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review*, *41*, 1375–1401.
Estrella, A., Rodrigues, A. P., & Schich, S. (2003). How stable is the predictive power of the yield curve? Evidence from Germany and the United States. *The Review of Economics and Statistics*, *85*(3), 629–644.
Galbraith, J. W., & Tkacz, G. (2000). Testing for asymmetry in the link between the yield spread and output in the G-7 countries. *Journal of International Money and Finance*, *19*(5), 657–672.
Giacomini, R., & Rossi, B. (2006). How stable is the forecasting performance of the yield curve for output growth. *Oxford Bulletin of Economics and Statistic*, *68*, 783–795.
Hännikäinen, J. (2015). Zero lower bound, unconventional monetary policy and indicator properties of interest rate spreads. *Review of Financial Economics*, *26*, 47–54.
Hännikäinen, J. (2017). When does the yield curve contain predictive power? Evidence from a data-rich environment. *International Journal of Forecasting*, *33*, 1044–1064.
Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of Financial Economics*, *22*(December), 305–333.
Harvey, C. R. (1989). Forecasts of economic growth from the bond and stock markets. *Financial Analyst Journal*, (September/October), 38–45.
Harvey, C. R. (1991). The term structure and world economic growth. *The Journal of Fixed Income*, *1*(1), 7–19.
Haubrich, J. G., & Dombrosky, A. M. (1996). Predicting real growth using the yield curve. *Federal Reserve Bank of Cleveland. Economic Review*, *32*(1), 26–35.
Henry, Ó. T., Olekalns, N., & Thong, J. (2004). Do stock market returns predict changes to output? Evidence from a nonlinear panel data model. *Empirical Economics*, *29*, 527–540.
Kozicki, S. (1997). Predicting real growth and inflation with the yield spread. *The Federal Reserve Bank of Kansas City Economic Review*, *82*, 39–57.
Kuosmanen, P., & Vataja, J. (2014). Forecasting GDP growth with financial market data in Finland: Revisiting stylized facts in a small open economy during the financial crisis. *Review of Financial Economics*, *23*(2), 90–97.
Kuosmanen, P., Nabulsi, N., & Vataja, J. (2015). Financial variables and economic Activity in the Nordic countries. *International Review of Economics & Finance*, *37*, 368–379.
Laurent, R. (1989). Testing the spread. *Federal Reserve Bank of Chicago Economic Perspectives*, *13*, 22–34.
Mauro, P. (2003). Stock returns and output growth in emerging and advanced Economies. *Journal of Development Economics*, *41*, 129–153.
Mishkin, F. (2011a). Over the cliff: From the subprime to the global financial crisis. *The Journal of Economic Perspectives*, *25*(1), 49–70.
Mishkin, F. (2011b). Monetary policy strategy: Lessons from the crisis. In M. Jarociński, F. Smets, & C. Thimann (Eds.), *Approaches to monetary policy revisited – Lessons from the crisis* (pp. 67–118). Frankfurt: European Central Bank.
Mitchell, W. C., & Burns, A. F. (1938). Statistical indicators of cyclical revivals. NBER bulletin, 69, NY. In G. H. Moore (Ed.), *Reprinted in business cycle indicators* (p. 61). Princeton: Princeton University Press.
Ng, S., & Wright, J. H. (2013). Facts and challenges from the great recession for forecasting and macroeconomic modeling. *Journal of Economic Literature*, *51*, 1120–1154.
Samuelson, P. (1966). Science and stocks. *Newsweek*, *19*(September), 92.
Siegel, J. J. (2014). *Stocks for the long run* (fifth edition). McGraw-Hill.
Stock, J. H., & Watson, M. W. (2003a). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, *41*, 788–829.
Stock, J. H., & Watson, M. W. (2003b). How did leading indicator forecasts perform during the 2001 recession? *Economic Quarterly, Federal Reserve Bank of Richmond*, *89*(3), 71–90.
Teräsvirta, T. (2006). Forecasting economic variables with nonlinear models. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting* (Volume 1) (pp. 413–458). Amsterdam and Oxford: Elsevier, North-Holland.
Tsouma, E. (2009). Stock returns and economic activity in mature and emerging Markets. *The Quarterly Review of Economics and Finance*, *49*, 668–685.
Wheelock, D. C., & Wohar, M. E. (2009). Can the term spread predict output growth and recessions? A survey of the literature. *Federal Reserve Bank of St. Louis Review*, *91*(September/October), 419–440.