Towson University
Department of Economics
Working Paper Series

Working Paper No. 2020-01

Nowcasting in Real Time Using Popularity Priors

By George Monokroussos and Yongchen Zhao

February 2020

Nowcasting in Real Time Using Popularity Priors[*]

George Monokroussos[†] Yongchen Zhao[‡]

February 2020

Abstract: We construct a "Google Recession Index" (GRI) using Google Trends data on internet search popularity, which tracks the public's attention to recession-related keywords in real time. We then compare nowcasts made with and without this index using both a standard dynamic factor model and a Bayesian approach with alternative prior setups. Our results indicate that using the Bayesian model with GRI-based "popularity priors" we could identify the 2008Q3 turning point in real time, without sacrificing the accuracy of the nowcasts over the rest of the sample periods.

JEL Classification: C11, C22, C53, E37, E52

Keywords: Gibbs Sampling, Factor Models, Kalman Filter, Real-Time Data, Google Trends, Monetary Policy, Great Recession.

---

[*] The authors thank our discussant Jonathan Wright and seminar participants at the 2016 Conference on Real-Time Data, Analysis, Methods and Applications (Federal Reserve Bank of Philadelphia), as well as the participants of the 3rd Seattle-Vancouver Econometrics Conference, the 5th European Seminar on Bayesian Econometrics (ESOBE 2014), the 34th International Symposium on Forecasting, and the Joint Research Centre of the European Commission for useful comments and discussions. Furthermore, we are grateful to David Small for making available his real-time data.

[†] Amazon - Seattle, WA, USA. Email address: monokroussos@gmail.com. This research was conducted while the author was with the Joint Research Centre of the European Commission. The content of this study does not reflect the official opinion of the European Union. Responsibility for the information and views expressed in the study lies entirely with the author.

[‡] Department of Economics, Towson University, Towson, MD, USA. Email address: yzhao@towson.edu.

Nowcasting in Real Time Using Popularity Priors

1. Introduction

Obtaining accurate and timely forecasts of real GDP growth, in particular its turning points, is well known to be challenging.[1] While recently developed nowcasting techniques, such as Giannone, Reichlin and Small (2008, henceforth GRS), are widely used to address the timeliness aspect of this challenge, successfully identifying turning points in real time remains difficult. In this paper, we demonstrate that we can better identify the 2008Q3 turning point using the nowcasts produced in the framework of GRS, when we adopt a Bayesian approach with priors guided by data on internet search volumes.

A turn in the GDP growth rate from positive to negative is usually accompanied by a widespread slowdown in economic activities. Workers, investors, employers, etc., who experience a change in their economic conditions due to the slowdown are more likely to conduct internet searches on recession-related topics than they do at other times. Internet-based services such as Google Trends publish time series indexes of search volumes. These indicators provide a valuable gauge of current economic conditions in real time due to the enormous amount of their underlying raw data, the short publication lag, and the high level of initial accuracy (i.e., no need for revisions).

We construct a "Google Recession Index" (GRI) using Google Trends data on search volumes of a selected set of recession-related terms. We show that one cannot effectively utilize this information by simply adding the index to a baseline data set of many macroeconomic

[1] See, inter alia, Stekler (1972), Zarnowitz (1986), Loungani (2001), and Lahiri and Wang (2013).

indicators. Instead, we propose a new Bayesian dynamic factor model, where we use "popularity priors" guided by the Google Recession Index.[2] We then carefully compare the nowcasts produced in real time by this model and alternative models with and without using the index. We find that for a time period centered around the Great Recession, our Bayesian nowcasting approach with "popularity priors" allows more timely detection of the 2008Q3 turning point than all of the alternative models.[3] Furthermore, this improvement does not come at the cost of having a lower nowcasting accuracy over the rest of our sample period from 2005 to 2014.

In the next section, we set up our Bayesian model. Specifics of the Google Recession Index and the baseline data set are described in Section 3. We present our empirical exercises and the results in Section 4. Section 5 concludes.

2. The Bayesian Dynamic Factor Model

Our objective is to nowcast US real GDP growth rate y. We use the Bayesian approach to estimate the dynamic factor model of GRS. This model allows us to forecast y at any given day of any given month using the most up-to-date information. But since the Google Trend data are monthly, we will focus on making forecasts on only the first business day of each month t. Note that due to varying data release schedules and publication lags, in month t's data set, not all of the N variables $x_{is}, i = 1, \ldots, N$, are non-missing for all $s \leq t$. As a result, the data set containing all

[2] Our approach is somewhat similar to Wright (2013), where a Bayesian VAR model was shown to produce good macroeconomic forecasts when the prior is constructed using survey responses. In a fashion similar to this "democratic prior," we refer to our prior as "popularity prior," highlighting the fact that our prior is based on the popularity of search terms. We follow GRS in setting up the dynamic factor model. Other contributions employing similar Bayesian approaches include Kim and Nelson (1998), D' Agostino et al. (2015), and Kose, Otrok and Whiteman (2003).
[3] This quarter is particularly consequential during the 2008 recession. It includes major events such as the collapse of Lehman Brothers, and it is the starting point of four consecutive quarters of negative GDP growth.

the xs is unbalanced, i.e., with a jagged-edge. Moreover, due to data revisions, any value x_{it} may change over time.

To capture the collinearities of the data, we use r latent common factors:

$$x_t = \mu + \Lambda F_t + \xi_t, \tag{1}$$

where x_t is the tth column of the $N \times t$ data matrix; μ is an $N \times 1$ vector of constants; Λ is an $N \times r$ vector of factor loadings; the $r \times 1$ vector $F_t = (F_{1t}, \ldots, F_{rt})'$ is the set of factors that are orthogonal to the $N \times 1$ vector of idiosyncratic error terms ξ_t. Furthermore, we assume that ξ_ts are cross-sectionally orthogonal Gaussian white noises with $E(\xi_t \xi'_{t-s}) = 0 \; \forall \, t, s > 0$. Let $E(\xi_t \xi'_t) = \Sigma^{\xi} = diag(\sigma^2_{\xi_1}, \ldots, \sigma^2_{\xi_N})$. Following GRS, we set $r = 2$ and assume that F_t can be specified as a first order vector autoregression:

$$F_t = AF_{t-1} + \zeta_t, \tag{2}$$

where ζ_t is an $r \times 1$ vector of common shocks and A is an $r \times r$ coefficient matrix with all the roots of $det(I_r - Az)$ lying outside the unit circle. We assume that the common shocks are Gaussian white noises independent from ξ_t. Let $E(\zeta_t \zeta'_t) = \Sigma^{\zeta}$. To complete the model, we use a bridge equation to forecast y using the common factors:

$$y_t = \alpha + \beta' F_t + \varepsilon_t \tag{3}$$

where α is a scalar; β is an $r \times 1$ vector of coefficients; and ε_t is a Gaussian white noise with variance σ^2_{ε}. For nowcasting, this equation is dated at t, and y_t is not known at this point due to its publication lag.

We estimate the model using a Gibbs sampler, which consists of nine blocks for the following groups of model variables: Σ^{ξ}, $\{\mu, \Lambda\}$, A, Σ^{ζ}, F_t, x^{miss}, $\{\alpha, \beta'\}$, σ^2_{ε}, y^f_t, where x^{miss} is the set of missing values of x and y_t^f is the nowcast of y_t.[4] Since our y is quarterly and the xs are monthly, we estimate equation (3) using the data from the third month of each quarter. When estimating equations (1) and (2), we use the data from all the months. Following Carter and Kohn (1994), we generate all the factors in one multi-move Gibbs sampling block, which converges faster and is computationally more efficient compared to single-move alternatives. Following Tanner and Wong (1987), we generate x^{miss} in a data augmentation step. Finally, for each of the blocks, we condition on the simulated samples from the other subsets of the model variables and all available data on x and y. Similar to the one discussed in GRS, our estimation procedure can also produce forecasts on any day of the month.

3. The Google Trends Index and the Baseline Data Set

Google Trends provide anonymized, categorized, and aggregated data on samples of actual search requests made to Google.[5] For a given search term and geographic area, an index is available since 2004. This index is calculated as the number of searches of the term divided by the total number of searches originated from the geographic area, then scaled to the range of 0 to 100.[6] We construct an index we call the Google Recession Index (GRI) using the publicly available Google Trends data on over 50 recession-related search terms and geographic areas (listed in Table 1).[7] Specifically, we take the median of these Google Trends indexes and then scale the median to the range of 0 to 100. The resulting index is shown in Figure 1, which clearly displays the increased search volume on recession related terms during the 2008 recession.[8]

---

[4] We make standard distributional assumptions. The full algorithm is described in detail in the extended working paper version of this study available on the author's website.
[5] https://trends.google.com/
[6] Google Trends data on topics and categories were criticized for issues related to their real-time characteristics and replicability (see Lazer et al. (2014) and Li (2016)). But the criticism does not apply to data on individual search terms.
[7] We compiled the list based on Google Trends' "top related queries," Koop and Onorante (2013), and our own judgment.
[8] Other forecasting applications using similar information include Askitas and Zimmermann (2009), D'Amuri and Marcucci (2010), Koop and Onorante (2013), and Scott and Varian (2014a,b).

Our baseline data set contains monthly observations on close to 200 macroeconomic variables dating back to 1982. The variables include monetary aggregates, prices, employment statistics, survey data, housing, banking balance sheet figures, etc. The real-time data vintages are available on a weekly basis between 2005Q2 and 2014Q2.[9] We use only the vintages closest to the first business day of each month. All of the variables are transformed to induce stationarity using the same procedure in GRS. When evaluating the nowcasts, we consider two sets of actual values: the first releases and the most recent values.

In the exercises below, we report nowcasts of annualized quarterly real GDP growth rates made at the beginning of each of the three months of the quarter and the first month of the next quarter. For any quarter, the first release of the actual value becomes available at the end of the first month of the next quarter. That is, for each quarter's actual value, we report four nowcasts; all made using the appropriate real-time data vintages.

4. Nowcasting Exercises and Results

4.1 Using the GRI in the standard GRS model

Adding the GRI as just another variable is arguably the most straightforward way of augmenting the baseline data. Using the standard GRS model (not estimated using the Bayesian approach), we produce three sets of nowcasts from April 2006 to June 2014: The first and the second set of nowcasts are made using the baseline data without the GRI. The third is made using the augmented data set. The training sample used to produce the first set of nowcasts starts from 1982. For the second and the third set of nowcasts, the training sample starts from 2004, when the

[9] The data set is also used in Giannone et al. (2010). For more information, refer also to GRS. Croushore and Stark (2001) and Orphanides (2001) discuss the need for real-time data.

GRI became available. For the entire evaluation sample as well as three subsamples, Table 2 reports the RMSEs of these forecasts (#1 to #3 in the table) calculated using the first release and the most recent actual values.[10] Figure 2 compares the three sets of nowcasts. With and without the GRI, the two sets of nowcasts made using the short training sample are very similar. The accuracy improves with the use of the longer training sample (without the GRI). All three sets of nowcasts stayed above zero until December 2008. These observations suggest that merely adding the GRI to the data set does not help detecting the onset of the recession, nor does it help to improve the accuracy of the nowcasts overall.

4.2 Using the GRI and the Bayesian approach

When estimating the dynamic factor model described in Section 2 using the Bayesian approach, we can use the GRI to guide the selection of the priors, instead of simply adding it to the baseline data set.[11] The resulting "popularity priors" can be set up in various ways. We propose to set the mean of the prior of α, which now becomes α_t due to its time-varying priors, to -3.5%, the weighted average of the actual growth rate over all post-WWII recession episodes, and set the variance as:

$$Variance(\alpha_t):\begin{cases} 1 & \text{when the GRI is between 0 and 9,} \\ 0.9 & \text{when the GRI is between 10 and 19,} \\ \vdots & \vdots \\ 0.1 & \text{when the GRI is between 90 and 99,} \\ 0.05 & \text{when the GRI is 100.} \end{cases} \qquad (4)$$

[10] We performed the Diebold-Mariano tests for all the nowcasts, using #1 as the baseline. The tests failed to reject the null of equal accuracy in all cases. Due to the small number of observations, we suggest interpreting the test results with caution.

[11] We also estimated the model using the Bayesian approach with uninformed priors. As expected, simply changing the estimation method did not lead to any significant change in the nowcasts (compared to standard GRS nowcasts made without the GRI). These results, as well as details on the priors and convergence analysis for the Gibbs sampler are provided in the extended working paper version of this study.

This setup associates higher values of the GRI with decreasing levels of uncertainty regarding whether the economy is in a recession. Subsequently, we use the label BGRS-Var for this model. Alternatively, we can use the GRI to guide the prior mean, while holding the variance fixed at 1:

$$Mean(\alpha_t): \begin{cases} 4 & \text{when the GRI is between 0 and 9,} \\ 3.25 & \text{when the GRI is between 10 and 19,} \\ \vdots & \vdots \\ -2.75 & \text{when the GRI is between 90 and 99,} \\ -3.5 & \text{when the GRI is 100,} \end{cases} \tag{5}$$

where the value 4% is the average of the actual values over non-recession periods. Subsequently, we use the label BGRS-Mean for this model.

Figure 3 compares the first and the most recent releases of the actual values, the nowcasts made using the two Bayesian models, and the nowcasts made using the standard GRS model without the GRI. The corresponding RMSEs are reported in Table 2 (#4, #5, and #1, respectively). Figure 3 clearly shows the advantage of the BGRS-Var in early detection of the 2008Q3 turning point: The nowcast of the BGRS-Var model turns negative in October 2008, earlier than that of the standard GRS model. The BGRS-Mean nowcasts are also lower than the standard GRS nowcasts, although the differences are smaller than those between the nowcasts of the BGRS-Var model and the standard GRS model. Since the GRI tracks the volume of searches and not levels of output or economic activity, associating the GRI with the level of real GDP growth may be less effective, especially when the association is based on a simple linear relationship. We thus argue that the better alternative is to use the GRI to guide the prior variance.

Figure 3 also shows that, the cost of false positives is small compared to the gain from using the GRI and the BGRS-Var model to predict the turning points. Consider simple "rule-of-thumb"-type forecasters who predict a recession whenever the real-time nowcast turns negative. They would have been able to identify the turning point in 2008Q3 earlier using the BGRS-Var model than the standard GRS model without the GRI, at the cost of only one false positive in 2011Q3.[12] Note that the nowcasts of the BGRS-Var model were below the thresholds for only one month. That is, for those who predict a turning point only after two consecutive quarters of negative growth, there would have been no false positive.

Furthermore, the advantage of using the BGRS models in real time comes without sacrificing the overall forecast accuracy. As shown in Table 2, over the entire evaluation sample period from 2005Q2 to 2014Q3, the BGRS-Var nowcasts made in the third month of a quarter and the first month of the following quarter are more accurate than the nowcasts made without the GRI. When evaluated against the most recent actual values, the BGRS-Var model outperforms the standard GRS model in all four months. Using the BGRS-Var model, a forecaster would have been closer to the "truth" more often during the three years from 2007 to 2009: In 56.3% of the time, the BGRS nowcast is closer to the advance GDP estimates than the standard GRS nowcast. The number increases to 63% if the forecaster targets the revised GDP values. As for the BGRS-Mean model, while it does not always outperform the baseline nowcasts, the differences are not statistically significant.

Figure 4 plots the posterior means of α_t for the two Bayesian models. It clearly shows the role of the "popularity priors" as well as how best to utilize the information contained in the GRI. In the BGRS-Mean model, when the GRI is used to guide the selection of the mean of the priors while keeping the variance at a high level, the posterior is mostly driven by the data. In contrast, when the GRI is used to guide the prior variance, the posterior more clearly reflects the dynamics of the GRI.

[12] If the forecasters were to use 1% instead of 0% as the threshold, they would have been able to precisely identify the onset of the 2008 recession in real time, at the cost of two false positives – one in 2010Q4 and the other in 2011Q3.

The above evidence about the most recent recession suggests that the GRI may contain useful information for nowcasting US real GDP growth and its turning points. In addition, to use this information effectively, one should consider a Bayesian estimation procedure where the GRI is used to guide the prior variance. These results are robust to the choice of the search terms we used to construct the GRI. We repeated our exercises using the Google Trends index for only the term "recession" and came to the same conclusions. However, using the index of an individual search term, such as "recession," is likely to cause more false positives in the predictions of turning points, because the indexes for individual search terms are often more volatile than the GRI.

5. Concluding Remarks

We proposed a Google Recession Index using Google Trends data on search terms related to "recession" and used the index and a Bayesian dynamic factor model to nowcast US real GDP growth rate. We argued that, rather than including the GRI as "just another variable" in the information set, forecasters should adopt the Bayesian estimation strategy we proposed in the paper, where higher values of the GRI are associated with decreased uncertainty on the economy being in a recession. We compared the nowcasts made using our proposed procedure to those made using an alternative prior specification and those made using the standard dynamic factor model of GRS without the GRI. The results suggested that the GRI contains information useful for nowcasting real GDP growth. Using this information in our preferred Bayesian model, forecasters could have identified the turning point of the 2008 Great Recession at an earlier point in time than forecasters who do not use the GRI, without sacrificing the accuracy of the nowcasts during other parts of the sample periods, including the non-recession periods. Our work highlights the potential of using internet search data, such as the Google Trends, as an early gauge of important and widespread changes in the economy.

References

Askitas, N., and K. Zimmermann (2009): “Google Econometrics and Unemployment Forecasting,” *IZA Discussion Paper* No. 4201.

Banbura, M., D. Giannone, and L. Reichlin (2011): “Nowcasting,” in Michael P. Clements and David F. Hendry, editors, *Oxford Handbook on Economic Forecasting*, forthcoming.

Camacho, M. and G. Perez-Quiros (2010): “Introducing the EURO-STING: Short Term Indicator of Euro Area Growth,” *Journal of Applied Econometrics* (forthcoming).

Carter, C.K., and P. Kohn (1994): “On Gibbs Sampling for State Space Models,” *Biometrika* 81, 541-553.

Croushore, D. and T. Stark (2001): “A Real-Time Data Set for Macroeconomists,” *Journal of Econometrics* 105(1), 111-130.

D' Agostino, A., D. Giannone, M. Lenza and M. Modugno (2015): “Nowcasting Business Cycles: a Bayesian Approach to Dynamic Heterogeneous Factor Models,” *Finance and Economics Discussion Series* 2015-066. Washington: Board of Governors of the Federal Reserve System.

D'Amuri, F. and J. Marcucci (2010): “Google it! Forecasting the US Unemployment Rate with a Google Job Search Index,” *Nota di Lavoro della Fondazione Eni Enrico Mattei* 31.2010.

Doz, C., D. Giannone, D. and L. Reichlin (2011): “A Two-step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering,” *Journal of Econometrics* (164), 188-205.

Giannone, D., M. Modugno, L. Reichlin, and D. Small (2010): “Nowcasting in Real-Time,” *mimeo*, presented at the *16th International Conference on Computing in Economics and Finance, Society for Computational Economics*, London, UK. July 13.

Giannone, D., L. Reichlin, and L. Sala (2004): “Monetary Policy in Real Time,” in: Gertler, Mark, Rogoff, Kenneth (Eds.), *NBER Macroeconomics Annual*. MIT Press, 161–200.

Giannone, D., L. Reichlin, and D. Small (2008): “Nowcasting: The Real-Time Informational Content of Macroeconomic Data,” *Journal of Monetary Economics* (55), 665-676.

Kim, Chang-Jin, and Charles R Nelson (1998): “Business Cycle Turning Points, a New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching,” *The Review of Economics and Statistics* 80, 188-201.

Koop, G., and L. Onorante (2013): “Macroeconomic Nowcasting Using Google Probabilities,” *mimeo*.

Kose, M.A, C. Otrok, and C.H. Whiteman (2003): “International Business Cycles: World, Region, and Country-Specific Factors,” *American Economic Review*, 93: 1216–1239.

Lahiri, K., and J.G. Wang (2013): "Evaluating Probability Forecasts for GDP Declines Using Alternative Methodologies," *International Journal of Forecasting* 29(1), 175-190.

Lazer, D., R. Kennedy, G. King and A. Vespignani (2014): "The Parable of Google Flu: Traps in Big Data Analysis" *Science*, Vol. 343, Issue 6176, pp. 1203-1205.

Li, Xinyuan (2016): "Nowcasting with Big Data: is Google useful in the Presence of other Information?" *mimeo, London Business School.*

Loungani, P. (2001): "How Accurate are Private Sector Forecasts? Cross-country Evidence from Consensus Forecasts of Output Growth," *International Journal of Forecasting* 17(3), 419-432.

McCulloch, R., and P.E. Rossi. (1994): "An Exact Likelihood Analysis of the Multinomial Probit Model." *Journal of Econometrics* 64: 207-240.

O'Brien, M. (2014): "How the Fed Let the World Blow up in 2008," *The Atlantic*, February 26, 2014.

Orphanides, A. (2001): "Monetary Policy Rules Based on Real-Time Data," *American Economic Review* 91(4), 964-985.

Scott, S.L. and H.R. Varian (2014a): "Predicting the Present with Bayesian Structural Time Series," *Int. J. Mathematical Modeling and Numerical Optimisation,* Vol. 5, Nos. 1/2.

Scott, S.L. and H.R. Varian (2014b): "Bayesian Variable Selection for Nowcasting Economic Time Series," in *Economics of Digitization,* Avi Goldfarb, Shane Greenstein, and Catharine Tucker (Eds), NBER Press.

Stekler, H. (1972): "An Analysis of Turning Point Forecasts," *American Economic Review* 62 (4), pp. 724-729.

Tanner, M. A., and W. H. Wong (1987): "The Calculation of Posterior Distributions by Data Augmentation." *Journal of the American Statistical Association*, 82, 528-549.

Wright, J. (2013): "Evaluating Real-Time VAR Forecasts with an Informative Democratic Prior," *Journal of Applied Econometrics*, 28, pp.762-776.

Zarnowitz, V. (1986): "The Record and Improvability of Economic Forecasting," *NBER Working Paper 2099*.

Table 1. Search terms used to construct the Google Recession Index

This table lists all the search terms whose corresponding Google Trends data are used to construct the Google Recession Index.

Search term (geography)	Search term (geography)
Recession (Worldwide)	Depression (Worldwide)
Recession (United States)	Depression Economy (Worldwide)
Financial Crisis (Worldwide)	Depression Economy (United States)
Financial Crisis (United States)	Economic Depression (United States)
Great Depression (Worldwide)	Economic Depression(Worldwide)
Great Depression (United States)	Financial Collapse (Worldwide)
Depression (United States)	Financial Collapse (United States)
Crisis (United States)	Meaning of Recession (Worldwide)
Crisis Economy (United States)	What is Recession (Worldwide)
Crisis Economy (Worldwide)	What is Recession (United States)
Recession Inflation (United States)	Define Recession (United States)
Recession Inflation (Worldwide)	Define Recession (Worldwide)
Recession Deflation (Worldwide)	Define Economic Depression (United States)
Meaning of Recession (United States)	Define Economic Depression (Worldwide)
Economy (United States)	Economy Recession (Worldwide)
Economy (Worldwide)	Depression Recession (United States)
Deflation (Worldwide)	Depression Recession (Worldwide)
Deflation (United States)	2008 Recession (United States)
US GDP Growth (United States)	2008 Recession (Worldwide)
US GDP Growth (Worldwide)	Recession Definition (United States)
Crisis (Worldwide)	Recession Definition (Worldwide)
Systemic Crisis (Worldwide)	The Recession (United States)
Systemic Risk (Worldwide)	The Recession (Worldwide)
Systemic Risk (United States)	Economic Recession (United States)
Financial Market Collapse (Worldwide)	Economic Recession (Worldwide)
Economic Slowdown (Worldwide)	US Recession (United States)
Economic Slowdown (United States)	US Recession (Worldwide)
Slowdown Economy (Worldwide)	Economy Recession (United States)

Table 2. Evaluating competing sets of nowcasts

For each set of nowcasts and sample/subsample period, this table reports the RMSEs calculating using the first and the most recent releases of the actual values. Among the five sets of nowcasts evaluated here, #1, #4, and #5 are made using a training sample starting from 1982. #2 and #3 are made using a training sample starting from 2004.

Evaluation sample and forecasts	Timing of forecasts (when forecasts are made) and vintage of actual values							
	First releases of actual values				Most recent actual values			
	First month, current quarter	Second month, current quarter	Third month, current quarter	First month, next quarter	First month, current quarter	Second month, current quarter	Third month, current quarter	First month, next quarter
Full sample: 2005Q2 to 2014Q2 (37 quarters)								
1. GRS_noGRI_fullsample	2.15	1.99	1.64	1.62	2.94	2.75	2.35	2.16
2. GRS_noGRI_shortsample*	3.06	2.20	1.80	1.56	3.61	2.93	2.65	2.20
3. GRS_withGRI_shortsample*	2.92	2.19	1.81	1.58	3.44	2.92	2.65	2.21
4. BGRS_Mean	2.53	2.10	1.70	1.64	3.24	2.89	2.38	2.12
5. BGRS_Var	2.36	2.07	1.52	1.61	2.91	2.72	2.16	1.88
Before the 2008 recession: 2005Q2 to 2007Q4 (11 quarters)								
1. GRS_noGRI_fullsample	1.41	1.35	1.29	1.59	1.65	1.68	1.58	1.84
2. GRS_noGRI_shortsample*	1.80	1.44	1.50	1.93	1.96	1.25	1.61	1.72
3. GRS_withGRI_shortsample*	1.86	1.48	1.61	1.94	1.71	1.28	1.79	1.74
4. BGRS_Mean	1.49	1.44	1.34	1.58	1.79	1.66	1.62	1.79
5. BGRS_Var	1.28	1.39	1.24	1.46	1.58	1.52	1.48	1.76
During the 2008 recession: 2008Q1 to 2009Q2 (6 quarters)								
1. GRS_noGRI_fullsample	3.35	2.76	1.91	1.58	5.19	4.58	3.68	2.56
2. GRS_noGRI_shortsample*	5.16	3.25	2.62	1.15	6.66	5.12	4.67	3.22
3. GRS_withGRI_shortsample*	4.73	3.17	2.57	1.18	6.28	5.06	4.64	3.22
4. BGRS_Mean	3.54	2.96	1.82	1.53	5.53	5.03	3.78	2.19
5. BGRS_Var	3.22	2.65	1.64	2.10	4.63	4.74	3.36	1.59
After the 2008 recession: 2009Q3 to 2014Q2 (20 quarters)								
1. GRS_noGRI_fullsample	2.02	2.00	1.72	1.65	2.53	2.47	2.19	2.20
2. GRS_noGRI_shortsample*	2.51	2.02	1.58	1.52	2.62	2.40	2.02	1.95
3. GRS_withGRI_shortsample*	2.48	2.04	1.58	1.54	2.58	2.40	2.01	1.96
4. BGRS_Mean	2.62	2.10	1.83	1.71	2.91	2.52	2.17	2.26
5. BGRS_Var	2.50	2.17	1.62	1.51	2.80	2.38	2.01	2.03

* The evaluation sample starts from 2006Q2 instead of 2005Q1 due to the shortened training sample used to produce the forecasts.

Figure 1. The Google Recession Index vs real GDP growth, 2004Q1-2014Q2

This figure compares the Google Recession Index (monthly, right axis) with US real GDP growth (quarterly, left axis). The actual values are the first releases and the shaded area represents the 2008 recession according to NBER chronology. The figure clearly shows the spike in the GRI at the start of the recession, long before the NBER announcement.

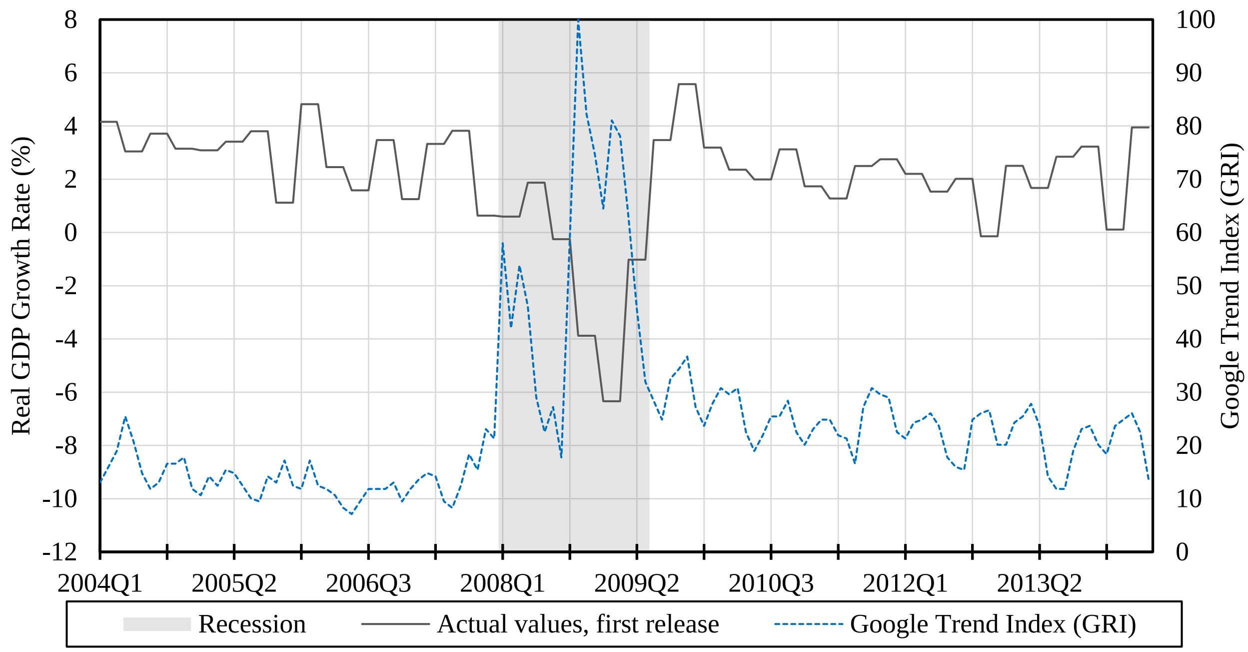
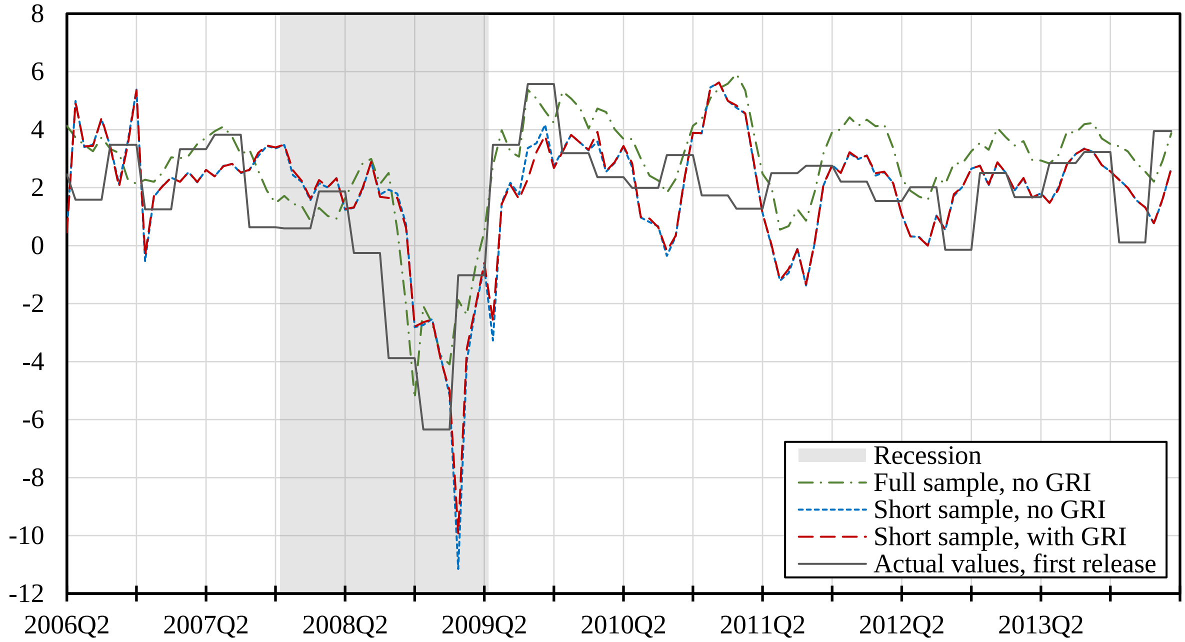


Figure 2: Nowcasts of the GRS model with and without GRI in the data set, 2006Q2-2014Q2

This figure compares the nowcasts produced by the GRS model with and without the GRI in the data set. “Full sample” means the training sample used to estimate the model starts from 1982. “Short sample” means the training sample starts from 2004. GRI is unavailable before 2004. The actual values are the first releases and the shaded area represents the 2008 recession according to NBER chronology. The top plot shows the entire evaluation sample. The bottom plot shows the same data but only the period from 2007Q1 to 2009Q4, so that the differences between the competing series of forecasts are clearly visible. The figure shows that adding the GRI does not improve the forecasts in any discernable way when using the standard GRS model.



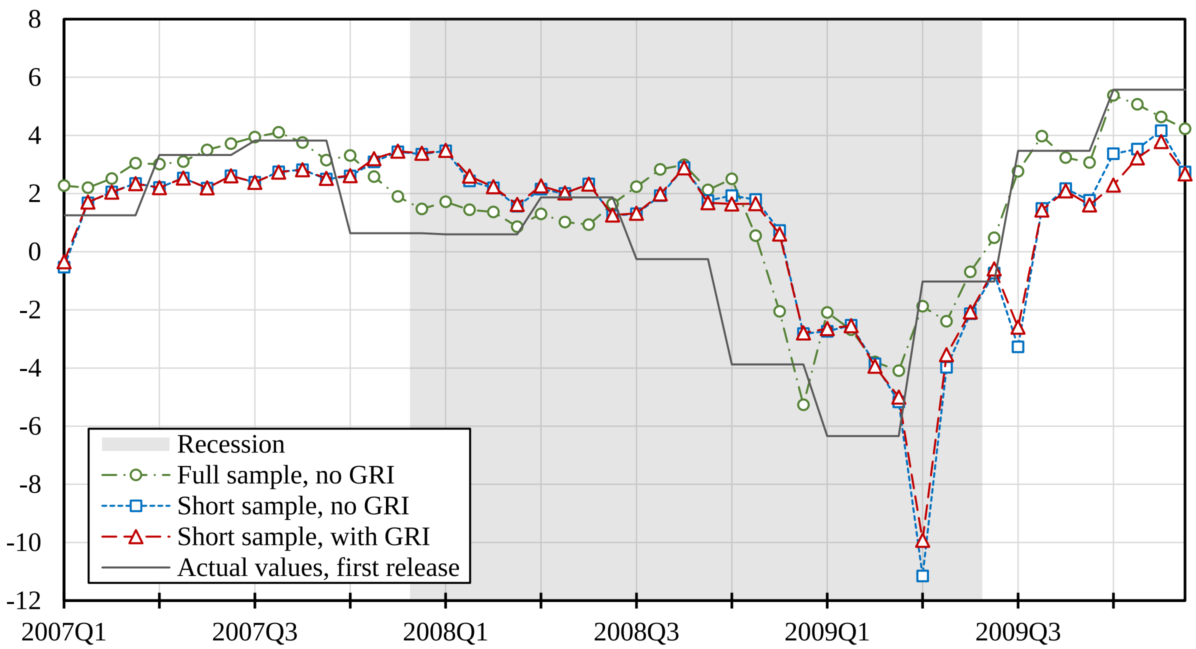
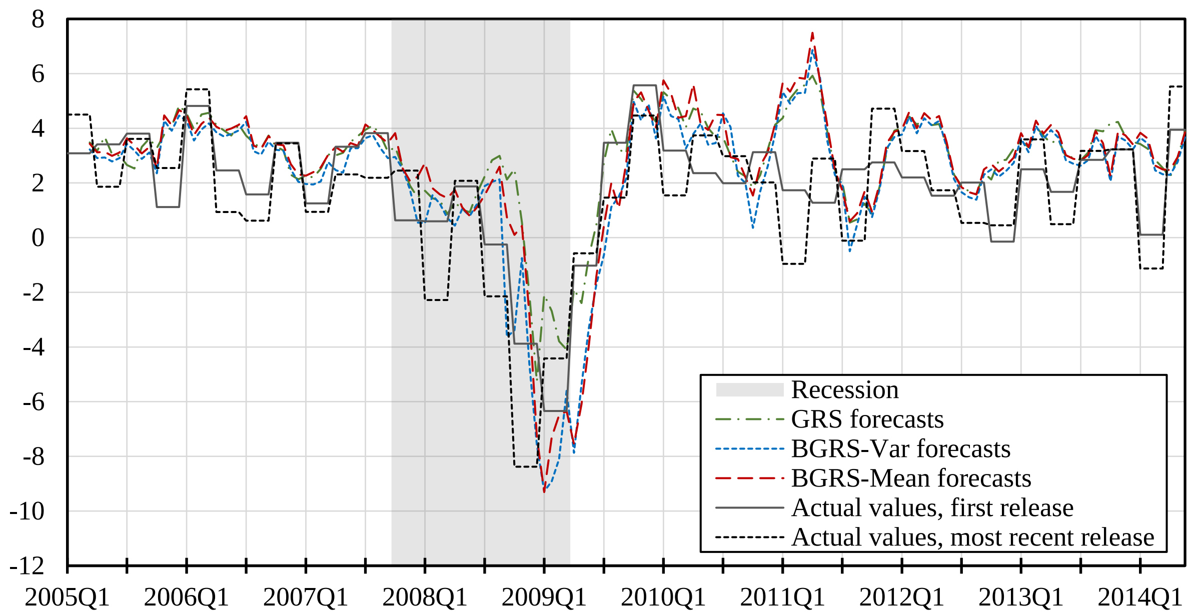


Figure 3. Nowcasts of the GRS model and the Bayesian GRS models, 2005Q1-2014Q2

This figure compares the following sets of nowcasts: (1) those produced by the GRS model without the GRI in the data set (GRS); (2) those produced by the Bayesian GRS model where the GRI is used to guide the selection of the variance of the prior (BGRS-Var); and (3) those produced by the Bayesian GRS model where the GRI is used to guide the selection of the mean of the prior (BGRS-Mean). The training sample used to estimate all three models starts from 1982. The shaded area represents the 2008 recession according to NBER chronology. The top plot shows the entire evaluation sample. The bottom plot shows the same data but only the period from 2007Q1 to 2009Q4, so that the differences between the competing series of forecasts are clearly visible. The figure shows that the information in the GRI, when extracted using the BGRS-Var model, best help predicting the onset of the 2008 recession.



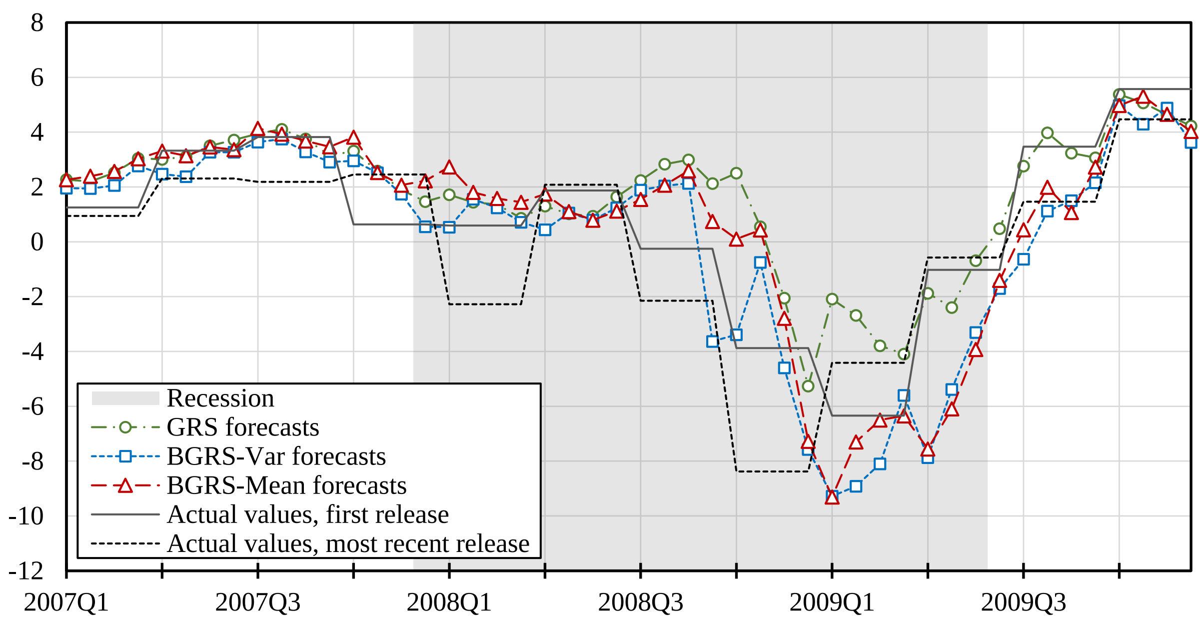


Figure 4. Posterior medians of alpha in the two Bayesian GRS models, 2005Q2-2014Q2

This figure shows the posterior medians along with the 90% equal-tailed interval of the alphas in the BGRS-Mean (top plot) and BGRS-Var (bottom plot) model. The shaded area represents the 2008 recession according to NBER chronology.