Interpreting Big Data in the Macro Economy: A Bayesian Mixed Frequency Estimator[1]

David Kohns[2] Arnab Bhattacharjee[3]

October 14, 2019

Abstract

More and more are Big Data sources, such as Google Trends, being used to augment nowcast models. An often neglected issue within the previous literature, which is especially pertinent to policy environments, is the interpretability of the Big Data source included in the model. We provide a Bayesian modeling framework which is able to handle all usual econometric issues involved in combining Big Data with traditional macroeconomic time series such as mixed frequency and ragged edges, while remaining computationally simple and allowing for a high degree of interpretability. In our model, we explicitly account for the possibility that the Big Data and macroeconomic data set included have different degrees of sparsity. We test our methodology by investigating whether Google Trends in real time increase nowcast fit of US real GDP growth compared to traditional macroeconomic time series. We find that search terms improve performance of both point forecast accuracy as well as forecast density calibration not only before official information is released but also later into GDP reference quarters. Our transparent methodology shows that the increased fit stems from search terms acting as early warning signals to large turning points in GDP.

Keywords- Big Data, Machine Learning, Interpretability, Illusion of Sparsity, Density Nowcast, Google Search Terms.

[1]We would like to thank Laurent Ferrara, Alessandro Giovonnelli, Nicolas Bonino-Gayoso, Markus Kaiser, Mark Schaffer, Gary Koop, Aubrey Poon and the conference participants of the Scottish Economic Society as well as the International Symposium on Forecasting for their invaluable comments and suggestions.

[2]Scottish Graduate Program in Economics and Spatial Economics & Econometrics Centre

[3]Spatial Economics & Econometrics Centre

1 Introduction

The objective of nowcast models is to produce 'early' forecasts of the target variable which exploits the real time data publication schedule of the explanatory data set. Nowcasting is particularly relevant to central banks and other policy environments who conduct forward looking policies on the basis of key economic indices such as GDP or inflation which are published with a lag of up to 7 weeks with respect to their reference period.[4] It is now common to combine, next to traditional macroeconomic data, ever more information from Big Data sources such as internet search terms, satellite data, scanner data, etc. (Buono et al., 2018) which are available in near real time.

A data source which has garnered much attention in the recent nowcast literature is Google Trends (GT), Google's search term indices. GT provides on a scale of 1 to 100, for a given time frame and location, the popularity of certain search terms entered into the Google search engine. Due to their timeliness as compared to conventional macro data and ability to function as an index of sentiment of demand and supply (Scott and Varian (2014)), they have celebrated wide spread use in nowcasting applications in many disparate fields of economics (see Choi and Varian (2012) and Li (2016) for surveys). They have proven especially useful in applications where searches are directly related to the variable of interest, such as unemployment data where internet search engines provide the dominant funnel through which job seekers find jobs (Smith (2016)). However, the verdict is still out whether Google Trends can be useful in nowcasting aggregate economic variables such as GDP, to which intuitive and theoretical links to search behaviour is less clear. Given that - compared to traditional macro data - Google Trends have a much shorter publication lag, potential improvements in fit could be substantial, especially before official data is released.

Next to the well known challenges of using Big Data sources for nowcasting such as different sampling frequencies and asynchronous data publications (ragged edges), a recent question which has been posed by Giannone et al. (2018) to the for- and nowcasting literature, is whether dense or sparse modeling techniques are appropriate in high dimensional settings. Since macro data tend to co-move strongly, and therefore dense modelling techniques might be preferred, Big Data sources such as Google Trends likely provide only a small number of search terms pertinent to GDP growth nowcasting, such that sparse models might be appropriate. Hence, there is uncertainty over the degree of sparsity one should impose (Giannone et al. (2018)).

Our proposed nowcasting method has in part been motivated by 4 key findings in the broader mixed frequency forecasting literature. Firstly, Giannone et al. (2018) highlighted that traditional approaches to deal with high dimensionality in macro forecasting mostly rely on dimension reduction through variants of factor estimation or LASSO style regularisation. These, however, impose the assumption that the underlying DGPs are respectively dense or sparse which they show can lead to sub-par forecasting performance when the model space provides support for models with dense *and* sparse components. This finding is reinforced by Cross et al. (2019) and Boivin and Ng (2006). Instead, Giannone et al. (2018) propose to use a prior which stays agnostic about the underlying degree of sparsity which is shown to improve forecasting accuracy. Secondly, forecast accuracy can be improved by opening the estimation procedure up to model uncertainty. As found in many previous forecast applications, combining forecasts

[4]The exact lag in publications of GDP and inflation depends as well on which vintage of data the econometrician wishes to forecast. Since early vintages of aggregate quantities such as GDP can display substantial variation between vintages, this is not a trivial issue.

over competing models through Bayesian Model Averaging (BMA) outperforms forecast accuracy compared to single model forecasts (see Steel (2017) for a survey) and, further, provides a formal probabilistic way of obtaining predictive inference. Third, leaving high frequency observations unrestricted can increase nowcasting fit if the frequency mismatch is not too wide, such as in nowcasting GDP with typical monthly macroeconomic data (Foroni and Marcellino (2014)). Lastly, to allow generalisability to other nowcast applications, it is important to be able to account for time series attributes such as non-stationarity or seasonality in a single modelling framework to which Scott and Varian (2014) provide a convenient state space framework.

Our modelling approach unifies these findings, while also addressing the econometric issues of mixed frequency and ragged edges in a computationally simple way by augmenting the Bayesian Structural Time series model (BSTS) by Scott and Varian (2014) with unrestricted mixed frequency data sampling, U-MIDAS Foroni and Marcellino (2014). One of the main advantages of our approach is that the tractability of the estimation procedure allows for a high degree of interpretability of our Big Data relative to competing nowcasting methodologies. We test our methodology by investigating whether U.S. real GDP growth nowcast accuracy in real time can be increased when including Google Trends alongside an updated version of the Giannone et al. (2016) data set which serves as our representative macro data set used for nowcasting GDP. We find that Google Trends increase nowcast point as well as density fit not only before official information is released but also later into GDP reference quarters. We show that the increased fit stems from search terms acting as early warning signals to large turning points in GDP.

To place our proposed econometric methodology into the existing mixed frequency nowcasting literature, we use the taxonomy provided by Bańbura et al. (2013) and focus on related mixed frequency methods designed for Big Data contexts (a more exhaustive list of references can be found in Bańbura et al. (2013) and Foroni and Marcellino (2014)).

Bańbura et al. (2013) classify mixed frequency nowcasting methods into full system and partial system estimators. While full system estimators treat all included variables as endogenous and therefore model them simultaneously, reduced system estimators only model the target variable, such that all other explanatory variables are treated as independent regressors. In the full system model space, solutions to deal with large data sets generally fall into mixed frequency factor or dynamic factor models (DFM) (e.g.: Giannone et al. (2008), Bańbura et al. (2013), Bok et al. (2018)) and mixed frequency VARs (e.g.: Koop (2009), Schorfheide and Song (2015), Brave et al. (2019)). In large K small T contexts, these methods provide different answers of how to conduct dimension reduction of the regressor space. In DFM models, this is handled through lower dimensional factors on the cross-section of the explanatory variables, while a commonly chosen way to address the curse of dimensionality in MF VARs, is to rely on shrinkage methods, most prominently Bayesian shrinkage (for a survey of such methods, see Koop (2009)).

Albeit widespread adoption of using DFM models for nowcasting in policy environments such as at the New Work Fed (Bok et al. (2018)) and Atlanta Fed (Higgins (2014)), the focus of these models is a continuous real time monitoring purpose. Drawbacks of these methods are that, empirically, it can be hard to interpret the effect of a single regressor in the system and, theoretically, as shown by Boivin and Ng (2006), using factor methods based on large data bases can lead to sub-par forecast fit if idiosyncratic errors are cross-correlated or when

factors that drive the forecast in smaller data sets are dominated in larger data sets. While interpretability is less of an issue in mixed frequency VARs, their computational complexity can become prohibitive in high dimensional settings with large frequency mismatches, as is the case when using Big Data sources on aggregate economic time series such as GDP.

Within partial system nowcasting, the most commonly used models are Bridge equations and mixed data sampling (MIDAS) regressions. In bridge models, high frequency observations are averaged within the reference period of the lower frequency target variable. Such models are considered in studies such as Bencivelli et al. (2012), Diron (2008) or Baffĩgi et al. (2004). The MIDAS approach, as originally popularised by Ghysels et al. (2004), utilises additive distributed lag (ADL) functions as kernels such as the Almon or Beta function, to allow for a parsimonious way in linking high frequency data to low frequency data. Only recently have these methods been adapted to high dimensional settings. Ferrara and Simoni (2019), use ridge regression based on bridged data to find a subset of regressors to optimise forecast accuracy, while in MIDAS settings, Marcellino et al. (2018) first retrieve principal components of the regressors and then apply MIDAS regression. Foroni and Marcellino (2014) on the other hand show in simulations as well as a nowcasting exercise on U.S. GDP growth that if the frequency mismatch between the lower frequency and higher frequency variable is not too large, such as in a quarterly to monthly mismatch, leaving the intra-period coefficients unconstrained can increase forecasting fit. With large data sets, however, this approach either requires shrinkage methods (Carriero et al. (2015)) or model averaging (Chikamatsu et al. (2018)). The verdict is still out whether in Big Data settings, U-MIDAS sampling can increase nowcasting fit compared to simpler Bridge models or MIDAS sampling.

Most related to our empirical methodology is the paper by Carriero et al. (2015) who use U-MIDAS sampled data from the FRED-MD data base to nowcast U.S. real GDP growth, where parsimony is enforced by a Minnesota style prior. In their Bayesian setup, they allow for stochastic volatility and time varying parameters which they show, increases point fit and density calibration. The Minnesota prior is adapted from large Baysian VAR applications such as in Litterman (1986) and Bańbura et al. (2010), and assumes that the modelled variables follow a unit root process and that lags receive increasingly stronger shrinkage which is different from shrinkage on exogenous variables. More generally, Minnesota style priors assume that the underlying DGP is sparse. Although sensible to specific time series, these assumptions do not necessarily apply to general nowcasting settings and necessitate specification of hyperparameters for the various degrees of shrinkage.

Relative to the previous literature on partial system mixed frequency estimators, our contribution is that first, we use a George and McCulloch (1997) style SSVS prior with a Dirac point mass to conduct variable selection on U-MIDAS sampled data, which compared to the Minnesota prior of Carriero et al. (2015), is a more automatic way of sparsifying the model as it stays agnostic about underlying degree of sparsity. Second, we allow for greater amount of flexibility to model general time series aspects such as unit roots and seasonality by using the BSTS framework of Scott and Varian (2014) and extend it to mixed frequency. Lastly, through unrestricted data sampling our approach allows for tractability of the Big Data included. Importantly, we believe that the last aspect may give more economic insights into the information inherent in Big Data sources which will help bridge the gap to a more structural approach to nowcasting.

In the Google Trends nowcasting domain, studies that address the problem of mixed fre-

quency and high dimensionality remain relatively sparse. Notable exceptions are Li (2016) who estimates the mixed frequency dynamic factor model of Bańbura et al. (2013) with Google Trends on U.S. initial jobless claims, Koop and Onorante (2019) who estimate dynamic model averaging models with transition probabilities based on Google Trends to nowcast various financial as well as aggregate economic time series, Scott and Varian (2014) who address high dimensionality, but not mixed frequency and lastly Ferrara and Simoni (2019), mentioned above, compare the gains of using Google Trends along side hard and soft economic data to nowcast real Euro area GDP growth. We contribute to the Google Trends literature by being the first paper to nowcast U.S. real GDP growth with Google Trends and macro data. Further, we are the first paper within this domain that addresses nowcast density calibration as well as potentially differing density structures of our data sets.

The rest of the paper will continue as follows. In section 2, we will describe our model for nowcasting. In the third section, we describe our method of pre-selecting search terms, as well as the macroeconomic data set for nowcasting. Section 4 presents and discusses our results and section 5 concludes.

2 Methodology

Model Setup

The original Bayesian Structural Time Series Model, as proposed by Scott and Varian (2014) consists of three distinct parts. First, a "structural time series model" (STM) after Harvey (2006), which delineates stochastic trends, seasonality and "irregular" regression components within a modular normal state space system. Second, a Spike-and-Slab (SNS) prior on the regression parameters to conduct variable selection and lastly, BMA to aggregate the information of the forecasts during the sampling process to form a unified forecast density. We add a fourth component to this model, namely, U-MIDAS sampling applied to the higher frequency regression component. We will discuss each part in turn.

State Space Model

The STM model of Harvey (2006) in particular, is a local linear trend model (LLT) which adds a regression component in a linearly additive fashion to the observation equation. In contrast to the original formulation of the model, we found that the data prefer a local trend model (LT) which we report below[5]:

$$y_t = \mu_t + z_t + \epsilon_t, \epsilon_t \sim N(0, \sigma^2) \tag{1}$$

$$\mu_t = \mu_{t-1} + w_t, w_t \sim N(0, W) \tag{2}$$

$$z_t = \beta' X_{t,M} \tag{3}$$

The subscript $t = 1, \cdots, T$ refers to quarters. Our response variable in the observation equation $y = (y_1, \cdots, y_T)$, is related to an unobserved trend component, $\mu = (\mu_1, \cdots, \mu_T)$ and a regression component, $z_t = \beta' X_{t,M}$, where $X_{t,M} = [x_{1,t,M}, \cdots, x_{K,t,M}]$ and $\beta = [\beta_1, \cdots, \beta_K]$ is the $K \times 1$ regression coefficient vector. The subscript after t indicates the observation within a given quarter, which in our case relates to monthly observations, i.e., $M = (1, 2, 3)$. We denote

[5]Note also, that we omit a seasonal state space equation, as our estimations performed better with pre-deseasonalised data.

the entire $T \times K$ regressor Matrix as $\boldsymbol{X}$. Note that the regressors have already been vertically realigned with quarterly growth according to the U-MIDAS methodology. This vertical realignment is visualised in a stylised fashion for a single regressor below:

$$\left(\begin{array}{c|ccc} y_{1stquarter} & x_{Mar} & x_{Feb} & x_{Jan} \\ y_{2ndquarter} & x_{Jun} & x_{May} & x_{Apr} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \end{array}\right) \tag{4}$$

Hence, the original dimensionality of the regressor matrix, prior to U-MIDAS sampling, is $TM \times K/M$. Adding the regression component through an identity in equation (3) in the observation equation has the benefit of only increasing the state dimensionality by 1. Equation (2) describes the stochastic evolution of the unobserved state equation which follows a unit root process. We assume that both the observation equation as well as the state equation have Gaussian errors with constant scalar variances, (σ^2, W) respectively. This, however, can straightforwardly be relaxed by allowing for stochastic volatility which we leave for future research.

To ease notation, we collect all parameters to be estimated in $\theta : \beta, \sigma^2, W$ and states in $\boldsymbol{\alpha} : \{\mu\}_{1:T}$. Bold symbols refer to vectors containing observations from 1 to T. Parameters and states are estimated by the Kalman filter which iterates between updating and prediction equations (see Petris & Petrone (2008)). During this process, the Kalman filter conveniently produces at every point in time a filtered estimate of the state, along with a predictive density (Harvey (2006)). This will be used in the empirical part to evaluate one-step-ahead forecasts.

Spike-and-Slab Priors

One of the main advantages of using a linear state space system is that each component is modular such that the states and regression parameters can be modelled independently once conditioned on either Scott and Varian (2014). Thus, conditional on sampling the states, the observation equation collapses into a normal linear regression model. I.e., $\tilde{y} = \boldsymbol{y} - \boldsymbol{\mu} \sim N(\boldsymbol{X}\beta, \sigma^2)$ to which we apply SNS priors to select a subset of regressors from our big data source.

The SNS prior we use in our application is based on the conjugate point mass zero SSVS prior of George & McCulloch (1997). It has a long tradition in economics (see Ishwaran and Rao (2005)) and is a fairly automatic way of conducting variable selection where regressors are sorted on the basis of relative marginal inclusion probabilities into a spike component, which shrinks the coefficient to exactly zero, or a slab component, which lets the parameters be estimated relatively freely. It is often formulated as a hierarchical prior as follows:

$$p(\beta, \sigma^{-2}, \gamma) = p(\beta|\gamma, \sigma^{-2}) p(\sigma^{-2}|\gamma) p(\gamma) \tag{5}$$

$\gamma = (\gamma_1, \cdots, \gamma_K)$ is an unknown indicator variable which takes the value of 1 if $\beta_j \neq 0$ and 0 if $\beta_j = 0$. β_γ is the subset β where $\beta_j \neq 0$. The prior further assumes that regression effects allocated to the spike component are independent of each other and independent of the subset β_γ, where elements may be dependent. The joint prior distribution is:

$$p(\beta|\gamma) = p_{slab}(\beta_\gamma) \prod_{j:\gamma_j=0} p_{spike}(\beta_j) \tag{6}$$

We define the spike distributions formally as Dirac point mass functions, $p(\beta_j) = p(\beta_j|\gamma = 0) = \Delta_0(\beta_j)$. For the slab distributions, we choose the widely used Zellner's g-prior which specifies the prior mean, b_γ, as zero and the covariance as a scalar multiple of the Fisher information, $\Omega_\gamma = g(X'_\gamma X_\gamma)^{-1}$:

$$\beta_\gamma|\gamma, \sigma^2 \sim N(b_\gamma, \sigma^2\Omega_\gamma) \tag{7}$$

This prior formulation allows us to control the strength of the prior information respective data information through the hyper-parameter g. The larger the chosen g, the more weight is put on the data. As recommended by George and McCulloch (1997), we set g=100. For the error variance, we use the usual uninformative prior which is commonly used as it makes $log(\sigma^2)$ uniform (Smith and Kohn, 1996):

$$p(\sigma) \propto 1/\sigma \tag{8}$$

Finally, we model γ through a Bernoulli prior which has the intuitive appeal of needing only one hyperparameter, $\pi_j = \pi \forall j$, which can be interpreted as the expected model size:

$$\gamma \sim \prod_j \pi_j^{\gamma_j}(1 - \pi_j)^{1-\gamma_j} \tag{9}$$

In summary, the researcher need only specify hyperparameters for the mean prior of the coefficients (b_γ), the relative strength of the prior vs data information (g) and the expected model size (π)[6].

In relation to comparable priors for variable selection and shrinkage, the SNS prior as defined above seems a natural fit for applications involving Big Data. It requires little input, and therefore little prior knowledge of the underlying relationship between the large data set and the response variable. Further, it remains agnostic about whether the underlying data representation is a-priori dense, sparse or both. As Giannone et al. (2018) have shown, SNS priors are able to separate variable selection through inclusion probabilities and shrinkage of the regression coefficients through the prior covariance structure which makes them an efficient prior when the posterior model space includes representations with modes on models which are relatively large and relatively small. This seems a natural assumption when combining macro-data sets with strong co-variation and large data bases such as Google Trends with potentially many unimportant search terms, i.e., sparse data. This flexibility contrasts starkly with LASSO style shrinkage methods which assume a-priori sparsity, or factor methods, which assume a-prior density.

The only other prior input needed from the researcher to conduct posterior simulation is a prior for the state variance and starting values for the Kalman filter. These are standard and given in the appendix.

The Posterior

The full posterior can be written as:

$$p(\theta, \boldsymbol{\alpha}|\mathbf{y}) = p(\theta)p(\alpha_0)\prod_{t=1}^{T} p(y_t|\alpha_t, \theta)p(\alpha_t|\alpha_{t-1}, \theta) \tag{10}$$

[6]Additionally, as in Scott and Varian (2014), we use an independent Gamma prior on the state variance W. Exact specification can e found in the appendix

α_0 refers to the starting value of the state and θ refers to the collection of parameters unrelated to the states, α. Since the joint posterior of states and parameters does not pertain to a known form, we draw in Gibbs steps from the conditional posteriors $p(\theta|\alpha, y)$ and $p(\alpha|\theta, y)$. These draws produce a Markovian chain $p(\theta, \alpha)_0, p(\theta, \alpha)_1, \cdots$ which after a certain burn-in period, result in a stationary distribution $p(\theta, \boldsymbol{\alpha}|\mathbf{y})$ (Scott, 2015). The entire sampling algorithm can be found in the appendix.

Posterior simulation of the Parameters

Making use of the independence between the state component and the regression component, the conditional posterior for θ can be factored into $p(\theta|\alpha, y) = p(\beta, \gamma, \sigma, W|\boldsymbol{\alpha}, \mathbf{y}) = p(W|\boldsymbol{\alpha}, \mathbf{y})p(\beta, \gamma, \sigma|\boldsymbol{\alpha}, \mathbf{y})$. Note that conditional on the states, the posteriors are readily available from standard conjugate posterior distributions which can be evaluated analytically and therefore offer computationally an advantage over non-conjugate prior setups (the exact posterior distributions are listed in the appendix). Analytical distributions for the posterior of γ are additionally necessary so as to circumvent reducibility of the Markov chain of parameter draws, stemming from the non-proper Dirac point-mass prior in (6) (George and McCulloch, 1997). In fact, marginalising over all the regression parameters, all information necessary to conduct variable selection is in the posterior of γ, which is interpreted as the posterior inclusion probability of a given regressor (George and McCulloch, 1997). Suppressing the condition on the states and state parameters for readability, the marginal posterior inclusion probability can be computed using Bayes' Law $p(\gamma|y) \propto p(y|\gamma)p(\gamma)$:

$$\begin{aligned} p(y|\gamma) &\propto \int_\sigma \left\{ \int_\beta p(y|\beta_\gamma, \sigma^2) p(\beta_\gamma|\sigma^2) d\beta_\gamma \right\} p(\sigma^2) d\sigma^2 \\ &\propto (1+g)^{-q_\gamma/2} S(\gamma)^{-T/2} \end{aligned} \tag{11}$$

q_γ is defined as the length of a given β_γ and $S(\gamma) = y'y - \frac{c}{1+c} y' X_\gamma (X_\gamma' X_\gamma)^{-1} X_\gamma' y$, where X_γ is the design matrix for which $\beta \neq 0$ (George & McCulloch, 1997). Plugging in our prior for γ, we evaluate:

$$p(\gamma|y) \propto (1+g)^{-q_\gamma/2} S(\gamma)^{-T/2} \prod_j^K \pi_j^{\gamma_j} (1-\pi_j)^{1-\gamma_j} \tag{12}$$

After sampling γ, we explore the conditional posteriors of the remaining parameters in θ.

Posterior simulation of the states

The posterior of the states, $p(\alpha|\theta, y)$ are simulated using the classic Kohn & Carter (1994) Gibbs sampling algorithm. It exploits the Markov property of the state space whose conditional posterior can be written out as (suppressing the condition on θ for readability):

$$p(\alpha|y) = p(\tilde{\alpha}_T|\tilde{y}_T) \prod_{t=1}^{T-1} p(\alpha_t|\alpha_{t-1}, \tilde{y}_t) \tag{13}$$

Where, $\tilde{y}_t$ and $\tilde{\alpha}_t$ refer to $[y_1, \cdots, y_t]$ and $[\alpha_1, \cdots, \alpha_t]$ respectively.

BMA

Each draw in the sampling process above produces a model which can be used to obtain forecasts at any point in time for y_{t+1}. The formal Bayesian way to combine these forecasts into one

coherent density is to acknowledge model uncertainty by simulating forecasts over the entire model space. Evaluating the entire model space, however, would entail estimating 2^K possible models which with current computational methods is not feasible. As the SSVS algorithm explores models with low residual variance and relatively few predictors, the algorithm concentrates on regions in the model space with high likelihood. Let y^* be the latent forecast based on $\boldsymbol{y}$ and ϕ be a vector of all parameters and states, then the forecast distribution is:

$$p(y^*|\boldsymbol{y}) = \int p(\boldsymbol{y^*}|\phi, \boldsymbol{y})p(\phi|\boldsymbol{y})d\phi \tag{14}$$

Hence, each prediction is simulated with draws from the posterior coefficient space which by construction accounts for model likelihood and sparsity (Scott & Varian, 2014). For the empirical application, we compute the expectation of (15), $E(y^*|\boldsymbol{y})$, as our frequentist equivalent to a point forecast. These are evaluated on a 1-step-ahead basis to the actual realisations. Since our method provides a whole forecast density, we additionally check for correct calibration by visually inspecting the probability integral transforms (PIT). This method was first emphasised by Diebold et al. (1998) and has the advantage over other indicators of density fit, such as coverage rates, of providing a more general indicator of goodness of fit. This generalisibility stems from PITs providing an indication of density fit irrespective of the econometricians' loss function. PIT histograms, under optimal density nowcast calibration, should form an uniform(0,1) distribution.

3 Data

As early data vintages of U.S. GDP can exhibit substantial variation compared to final vintages (Croushore (2006); Romer (2012); and Sims (2002)), it is not trivial which data to use in evaluating nowcast models on historical data. Further complications can arise through changing definitions or methods of measurements of data (Carriero et al. (2015)). However, as in our application, only a few explanatory variables have recorded real time vintages (see Giannone et al. (2016)) and our training window is restricted to begin with 2004 only - since this is the earliest data point for the Google Trends data base - we decided to use final vintages of our data. We therefore consider a pseudo-real time data set: we use the latest vintage of data, but, at each point of the forecast horizon, we use only the data published up to that point in time.

The target variable for this application is deseasonalised U.S. real GDP growth (GDP growth) of as downloaded from the FRED website. The deseasonalisation pertains here to the X-13-ARIMA method and was performed prior to download from the FRED-MD website. We found that this improved forecast accuracy compared to modeling it in our state space system. This might be due to the small sample size. As Google Trends are only available from 01/01/2004-01/06/2019, at the time of download, the period under investigation pertains to the same period in quarters (61 quarters). We split the data set into a training sample of 45 quarters (2004q2-2015q2) and a forecast sample of 15 quarters (2015q3-2019q1).

The macro data set pertains to an updated version of the data base of the seminal paper by Giannone et al. (2016) (henceforth, 'macro data'). It contains 13 time series which are closely watched by professional and institutional forecasters (Giannone et al. (2016)) such as real indicators (industrial production, house starts, total construction expenditure etc.), price data (CPI, PPI, PCE inflation), financial market data (BAA-AAA spread) and credit, labour and economic uncertainty measures (volume of commercial loans, civilian unemployment, economic

uncertainty index etc.)(Giannone et al. (2016)). Table 2 gives an overview over all data along with FRED codes.

Our sample of search terms comprises 27 Google Trends which have been chosen based on the root term methodology as in Koop and Onorante (2019) and Bock (2018). In general, there is no consensus on how to optimally pre-screen search terms for the final estimation. Methods which have been proposed by the previous literature fall into: (i) pre-screening through correlation with the target variable (e.g.: Scott and Varian (2014); Choi and Varian (2012) and references therein) or through cross-validation (Ferrara and Simoni (2019)), (ii) use of prior economic intuition where search terms are selected through backward induction (e.g.: Smith (2016); Ettredge et al. (2005); Askitas and Zimmermann (2009)), and (iii) root terms, which similarly specify a list of search terms through backward induction, but additionally download "suggested" search terms from the Google interface. This serves to broaden the semantic variety of search terms in a semi-automatic way. From the authors' perspective, the root term methodology currently provides the best guarantee of finding economically relevant Google Trends, as methodologies based on pure correlation do not preclude spurious relationships (Scott and Varian (2014)).

Since search terms can display seasonality, we deseasonalise all Google Trends by the Loess filter which is implemented with the "stl" command in R [7]. Finally, our pseudo-real time calendar can be found in table 1 and has been constructed after the data's real publication schedule. It comprises in total 30 vintages which are estimated until the final GDP observation is released.

4 Results

In order to gauge the added benefit of search terms, we compare in- as well as out-of-sample fit between three variants of our model: a full model which contains the Google Trends data base as well as monthly macro variables, an only macro data model, and lastly a completely univariate model which estimates (1)-(2) without the inclusion of regressors, z_t. Without z_t, the model collapses into a local trend model which can be shown is equivalent to estimating an ARIMA(0,1,1) model on U.S. real GDP growth (Petris et al., 2008). This will serve as our benchmark. We abbreviate our models as F-MFBSTS, M-MFBSTS & U-BSTS respectively. All prior specifications and starting values are kept the same across models and sample sizes so as to maintain comparability. In order to robustify our results, we tested expected model sizes between 5 and 20 so as to incorporate beliefs of relatively small and large models. All subsequent results are based on the best performing prior model size hyperparameter of 12, unless stated otherwise.

As described in section 3, the state space system in (1)-(3), at each point in time, iterates through forecast and updating equations which produce one-step-ahead forecast distributions for the observation equation and likelihoods which we use to describe the fit of our models.

To summarise pure in-sample performance, we found it most instructive, as recommended by Scott and Varian (2014), to look at the mean absolute cumulative one-step-ahead forecast errors of our models. These are shown in the upper panel of Figure 1. The lower panel shows

[7] To mitigate any inaccuracy stemming from sampling error, we downloaded the set of Google Trends seven times between 01/07-08/07/2019 and took the cross-sectional average. Since we used the same IP address and googlemail account, there might still be some unaccounted measurement error which could be further mitigated by using web-crawling.

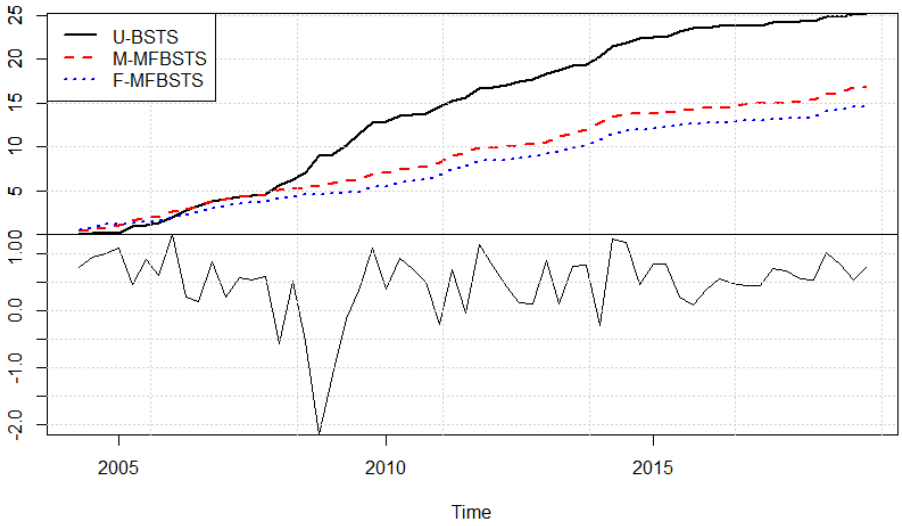


Figure (1) Cumulative absolute one-step-ahead forecast error comparison

the realisation of U.S. real GDP growth. 4 points emerge from this figure: the first is that the full model performs best, which can be seen from the blue dotted line being consistently below the cumulative absolute errors of the macro and univariate model. Secondly, the monthly information in both the full and macro model significantly improve fit, as indicated by the near linear increase of cumulative absolute errors compared to the univariate model. Thirdly, this is especially true around the financial crisis in 2007/2008, where the monthly information efficiently captures the large trough. The last salient point is that starting around the time of the financial crisis, the search terms' augmented model provides the largest improvement in fit compared to the macro and univariate model.

Inclusion Probabilities

The economic intuition behind these results can further be unpacked by consulting the marginal posterior inclusion probabilities which are computed as in (13) and are shown in figure 2, along with their respective sign when drawn. To keep the discussion and figures tractable, we report in figure 2 only the top ten most drawn regressors for the full model [8]. The length of the bars indicate posterior inclusion probability and the color of the bars indicate the relative probability of positive vis-a-vis negative signs of the variables on a black-white spectrum [0,1]. Definitions along with descriptive stats for all variables can be found in the appendix. The posterior inclusion probabilities as seen in the graph, firstly, confirm that Google Trends do play a role in GDP estimation, as 4 out of the top 10 most drawn variables are search indices, with the search term "gdp growth" receiving the highest posterior inclusion probability among them. These trends, secondly, cluster around the topic of recession fears or financial distress which, as economic intuition might suggest, are negatively associated with GDP growth. Hence, these trends can be interpreted as expectations of downside risks to GDP growth, which, as

[8] Results for the macro model can be found in the appendix.

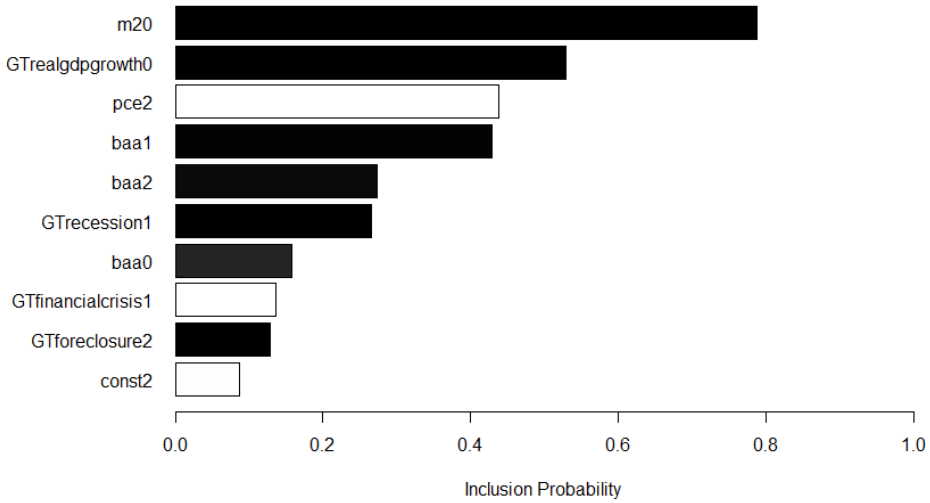


Figure (2) Marginal posterior inclusion probabilities for F-MFBSTS. The prefix GT signifies variables which are Google Trends. The number [0,1,2] appended to a variable indicates the position in a given quarter, with 0 being the latest month. The color of the bar is determined by the probability of the sign of a given variable being positive, where 1 is white and 0 black.

shown in figure 1, provide large improvements of fit around large turning points. This insight in particular corroborates findings from previous studies with Google Trends such as Choi and Varian (2012) who find that Google Trends perform well in predicting large troughs and peaks. Thirdly, the signs of the macro variables in figure 2, too, mostly conform to economic intuition. For example, credit spread variables (BAA-AAA) which have a strong negative correlation to GDP growth are often used in macro models to gauge financial distress, as increased spread points toward higher perceived future risk (Giannone et al. (2016)). Similarly, PCE inflation has a positive sign, which coheres with general theoretical as well as empirical insights from real business cycle analysis (see Romer (2012)). M2 variables, however, which are the monthly growth in the M2 money base, have counter intuitive negative signs. Economic theory predicts that an increase in money growth increases GDP under sticky prices. However, if money growth increases subsequent to declining GDP, then the contemporaneous relationship may be negative, which is what is likely picked up here. Fourthly, the marginal inclusion probabilities also highlight the benefit of leaving the intra-quarter monthly coefficients unconstrained, as only certain months feature frequently in model draws (e.g., M2-0 and PCE-2). To determine the added benefit of unconstrained mixed frequency estimation, we compare the in-sample results of our full model with our full model based on bridged data (as in Ferrara and Simoni (2019)). Figure 3 shows that fit is markedly improved through unrestricted data sampling. Lastly, comparing inclusion probabilities between the full model and monthly model in figure 9 of the appendix, the strongest macro variables are retained. This points toward the fact that only a hand-full of macro variables consistently drive fit.

Consistent with the inclusion probabilities, we find that the top 10 models by likelihood agree with a core selection of macro variables driving fit. We report posterior means and stan-

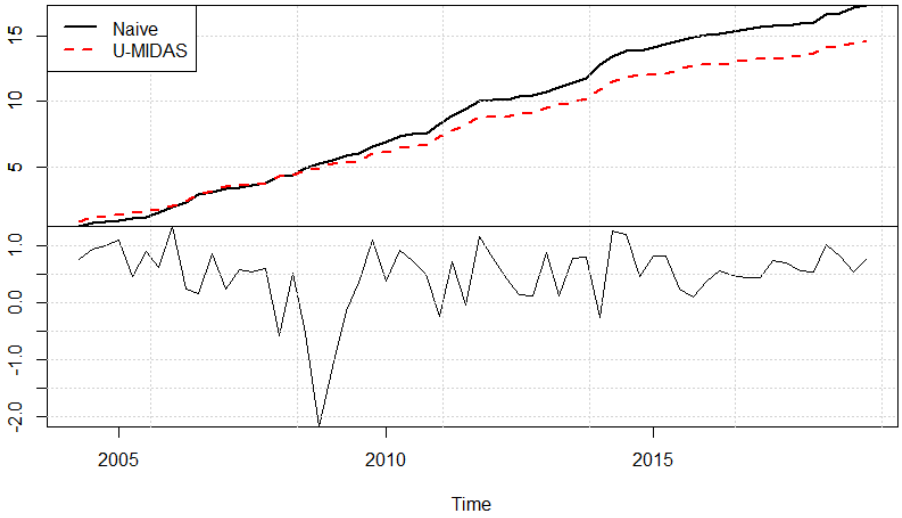


Figure (3) Upper Panel: Both naive & U-MIDAS models have been estimated on the full data set and same prior specifications as in Figure 1. The U-MIDAS model pertains to the full model in Figure 1.

dard deviations for the top 10 models in table 4 in the appendix. However, table 4, also makes clear that contrary to the macro data, there is less support for a consistent presence of particular Google Trends search terms (with the expection of the term "gdp growth"). Those included, however, cluster around topics of financial distress. Similarly to Giannone et al. (2018), we find that the posterior size of the models are well identified, as can be seen from figure 4 below.

To investigate the stability of our results across prior expected model sizes, we tested hyperparameter values for π between 5 and 20. Up to $\pi = 15$, we find that model sizes, regressors and estimation precision are very similar. Interestingly, we find that for larger expected model sizes, $\pi > 15$, posterior model sizes centre on two different modes. This is shown exemplary for model size of 18 in figure 5. It shows a mode around 3-5, and another on larger models with 13-15 regressors. These larger models select in more macro variables, but also many financial distress search terms (see figure 10 in the appendix). The forecast performance of these models, however, is severely impaired by estimation uncertainty such that the out of sample performance fairs worse than the univariate local trend model. For this reason, the out-of-sample results below are based on the best performing π of 12. The bi-modality in the model space does however motivate future research to include prior specifications which allow for varying degrees of shrinkage alongside variable selection such that estimation error can be better controlled for in larger models. The finding that the model size is well identified, but certain variable groups included less so, underscores the need for Bayesian model averaging so as to incorporate information from various model variants.

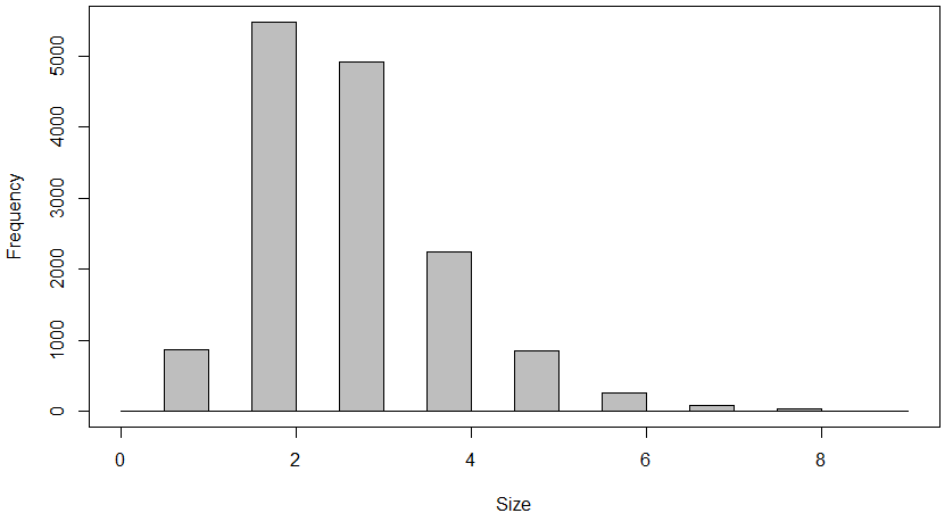


Figure (4) *Posterior Distribution of model sizes of the MF-MFBSTS for* $\pi = 12$.

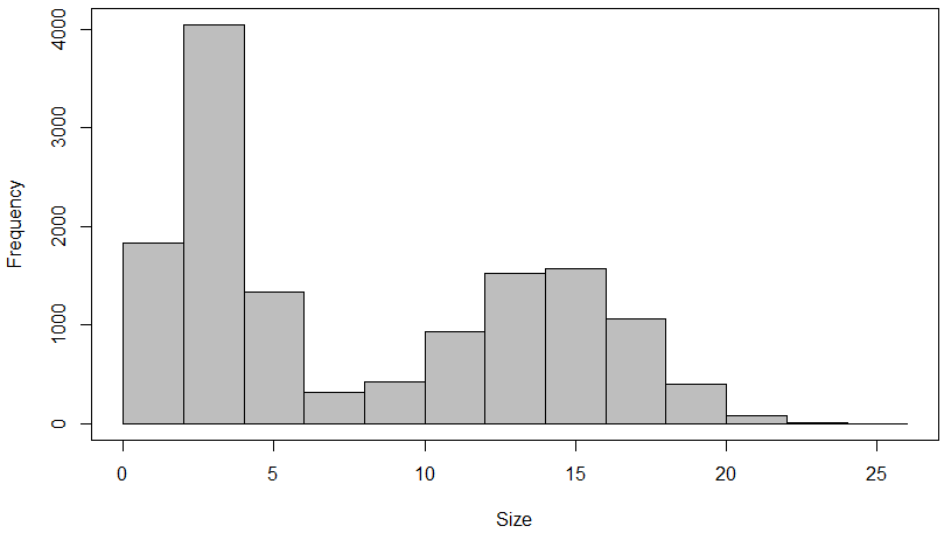


Figure (5) *Posterior distribution of model sizes of the F-MFBSTS model for* $\pi = 18$

Real-Time Nowcasting

To evaluate out-of-sample fit in real time, we firstly estimate the state and parameters of our models on our training sample of 45 periods, and then simulate forecast distributions using the vintages of the pseudo-real time data calendar in table 2. We treat every variable prior to its release, as indicated by the 'Release' column in table 2, as unobserved and set its value to zero. E.g., the first time the model is updated (for a given reference period) is on the last day of the first month, which we signify by vintage 1. All but the variables fedfunds-2 rate and the BAA-2 spread are set to zero. Importantly, we do not re-estimate the model states nor the regression parameters for each vintage, but take the historical states and the parameters as given. Further, since we want the Google Trends to reflect the most recent information upon publication, we set their release date to the last day of the given month. [9]

Figure 6 plots the real time mean squared forecast errors (RTMSFE). The RTMSFE show, as expected, that with the arrival of more information throughout a given quarter, nowcast accuracy increases. Similar to our in-sample comparisons, we find that the full model outperforms the monthly model throughout all vintages and the univariate model in later vintages. Up until vintage 23, however, the univariate model outperforms both model variants with regressors. This rather surprising finding can be attributed the fact that the local trend has been fitted in the training sample assuming that all regressorts at all time point are available. In early vintages, therefore, the local trend displays substantially more variation than in latter ones. This feature could potentially be remedied by choosing a tighter prior on the state variances, which is under active investigation by the authors.

While the symmetry in the RTMSFE decline for the full and macro model underpins results from in-sample estimations that both models select the same driving macro variables, a decisive difference appears in vintage 3 which is when Google Trends are released. Forecast accuracy for the full over the macro model increases at this point by nearly 19%. Subsequent Google Trends releases, however, do not increase fit further over the released macro variables. This finding echoes nowcast results of previous research using survey indicators for nowcasting which indicates that the information inherent in the chosen search terms provide valuable information about GDP expectations early on in the quarter, however, with fundamental macro information being released, these add little extra value (Giannone et al. (2016); Aastveit et al. (2018)).

As mentioned above, optimally calibrated nowcast densities should result in their PIT series forming uniform(0,1) distributions. In figure 8 we report PITs for all models concentrating on the most illustrative vintages. The PITs show, consistent with point-nowcasts, that the full model also provides the best density calibration, having the the most uniform looking PIT. Surprisingly, the macro model density provides PITs throughout the first three vintages which pile onto the lowest pentile, indicating that the nowcast distribution centres on values consistently above the realisations. The tent-like shape of the PITs for both the full and univariate model, indicate that the forecast distributions are too dispersed. This result, however, might be an artifact of the short nowcast time-frame.

To Shed some further light on the improved coverage of including search terms in the estimation algorithm, we plot in figure 7 the for the last vintage in our data publication schedule

[9] An alternative approach would be to download Google Trends at higher frequencies and average to obtain monthly Trends in real time. This, however, would refute our approach of determining intra quarter weights in a data driven fashion. Previous research on nowcasting GDP showed further that information at higher frequencies only improves nowcasts marginally if at all (Bańbura et al. (2013))

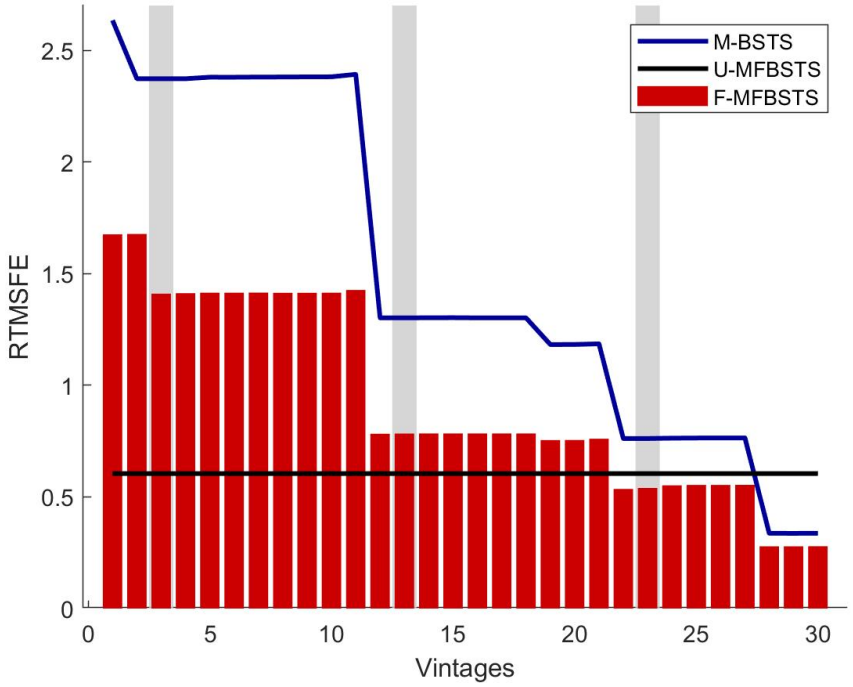
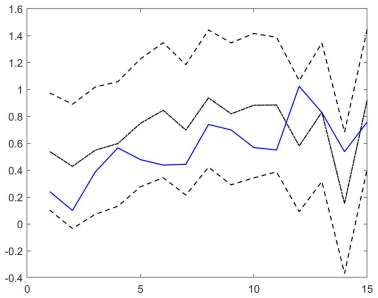
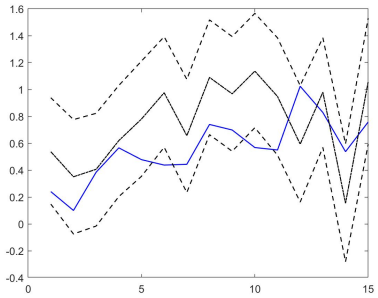


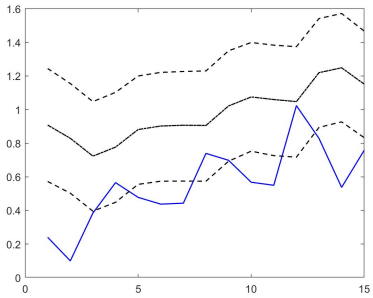
Figure (6) RTMSFE of the U-BSTS, M-MFBSTS and F-MFBSTS models over 30 data vintages. Publication vintages of Google Trends are highlighted with grey bars.



(a) F-MFBSTS



(b) M-MFBSTS



(c) U-BSTS

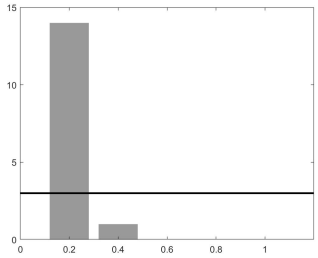
Figure (7) Black dotted lines are the mean of the forecast distributions. Dashed lines represent the 15th and 85th intervals of the forecast distributions. U.S. real growth realisations are in blue.

mean forecasts of our model variants along with 15th and 85th percentile of their respective nowcast distributions. The full model, provides clear improvements on the coverage of the forecast distribution, albeit being more widely dispersed.

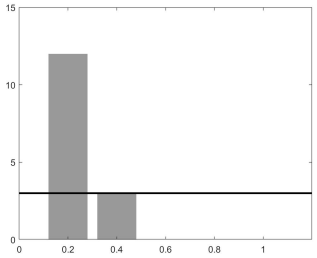
5 Conclusion

Our results indicate that Google Trends improve nowcast point fit, especially early into the quarter when official information is scant. Albeit search terms improving density nowcasts, the PITs show that correct calibration needs further improvements, as for the chosen forecast horizon, the densities remain too dispersed. A possible remedy - which we leave for future research - has been offered by Carriero et al. (2015) who show that in their sample, stochastic volatility is able to provide improvements on calibration of U.S. real GDP growth density nowcasts.

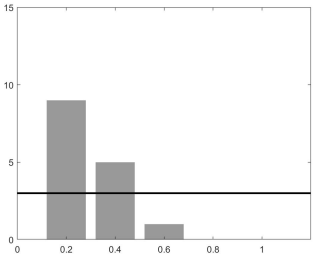
Next to improvements in nowcast fit, the tractability of our methodology is able to yield economic explanations to the improved fit. We have shown that the model selects search terms which cluster around the topic of financial distress and therefore carry expectations of downturns in GDP which can act as early warning indicators. This insight in particular provides promising input for research using search term data for structural modeling. Further, our methodology has shown that model uncertainty is pervasive in GDP nowcasting with Big Data and that there is need to be mindful of differing density structures when combining data sets.



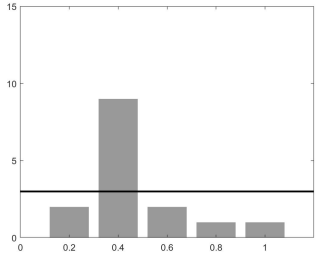
(a) PIT F-MFBSTS - Vintage 1



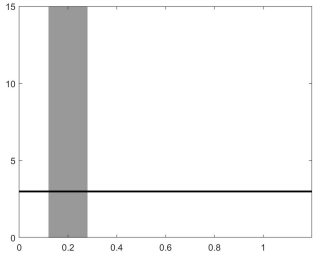
(b) PIT F-MFBSTS - Vintage 2



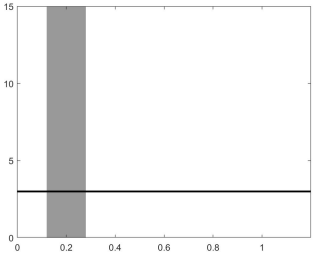
(c) PIT F-MFBSTS - Vintage 21



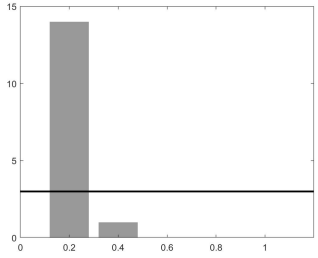
(d) PIT F-MFBSTS - Vintage 30



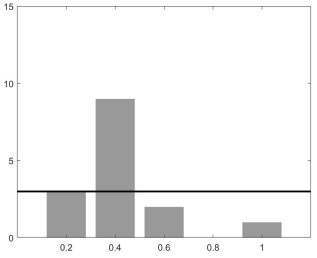
(e) PIT M-MFBSTS - Vintage 1



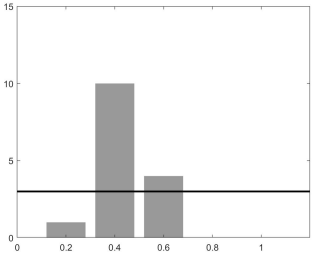
(f) PIT M-MFBSTS - Vintage 2



(g) PIT M-MFBSTS - Vintage 20



(h) PIT M-MFBSTS - Vintage 30



(i) PIT U-BSTS

Figure (8) PIT historgrams for all models

References

Aastveit, K. A., F. Ravazzolo, and H. K. van Dijk (2018): "Combined Density Nowcasting in an Uncertain Economic Environment," *Journal of Business and Economic Statistics*, 36, 131–145.

Askitas, N. and K. F. Zimmermann (2009): "Google Econometrics and Unemployment Forecasting," *Applied Economics Quarterly*, 55, 107–120.

Baffĭgi, A., R. Golinelli, and G. Parigi (2004): "Bridge models to forecast the euro area GDP," *International Journal of Forecasting*, 20, 447–460.

Bańbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013): "Now-casting and the real-time data flow," in *Handbook of Economic Forecasting*, vol. 2, 195–237.

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

Bencivelli, L., M. G. Marcellino, and G. Moretti (2012): "Selecting Predictors by Using Bayesian Model Averaging in Bridge Models," *SSRN Electronic Journal*.

Bock, J. (2018): "Quantifying Macroeconomic Expectations in Stock Markets Using Google Trends," *SSRN Electronic Journal*.

Boivin, J. and S. Ng (2006): "Are more data always better for factor analysis?" *Journal of Econometrics*, 132, 169–194.

Bok, B., D. Caratelli, D. Giannone, A. M. Sbordone, and A. Tambalotti (2018): "Macroeconomic Nowcasting and Forecasting with Big Data," *Annual Review of Economics*, 10, 615–643.

Carriero, A., T. E. Clark, and M. Marcellino (2015): "Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility," *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 178, 837–862.

Choi, H. and H. Varian (2012): "Predicting the Present with Google Trends," *Economic Record*, 88, 2–9.

Croushore, D. (2006): "Chapter 17 Forecasting with Real-Time Macroeconomic Data," .

Diron, M. (2008): "Short-term forecasts of euro area real GDP growth: An assessment of real-time performance based on vintage data," *Journal of Forecasting*, 27, 371–390.

Ettredge, M., J. Gerdes, and G. Karuga (2005): "Using web-based search data to predict macroeconomic statistics," .

Ferrara, L. and A. Simoni (2019): "When Are Google Data Useful to Nowcast Gdp? An Approach Via Pre-Selection and Shrinkage," *SSRN Electronic Journal*.

Foroni, C. and M. Marcellino (2014): "A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates," *International Journal of Forecasting*, 30, 554–568.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004): "The MIDAS Touch: Mixed Data Sampling Regression Mod," *Finance*.

Giannone, D., M. Lenza, and G. E. Primiceri (2018): "Economic Predictions with Big Data: The Illusion of Sparsity," *SSRN Electronic Journal*.

Giannone, D., F. Monti, and L. Reichlin (2016): "Exploiting the monthly data flow in structural forecasting," *Journal of Monetary Economics*, 84, 201–215.

Giannone, D., L. Reichlin, and D. Small (2008): "Nowcasting: The real-time informational content of macroeconomic data," *Journal of Monetary Economics*, 55, 665–676.

Harvey, A. (2006): "Chapter 7 Forecasting with Unobserved Components Time Series Models," .

Koop, G. (2009): "Bayesian Econometric Methods for Empirical Macroeconomics," *Methods*, 1–14.

Koop, G. and L. Onorante (2019): " Macroeconomic Nowcasting Using Google Probabilities ," 17–40.

Litterman, R. B. (1986): "Forecasting with bayesian vector autoregressions—five years of experience," *Journal of Business and Economic Statistics*, 4, 25–38.

Marcellino, M., F. Papailias, G. L. Mazzi, G. Kapetanios, and D. Buono (2018): "Big Data Econometrics: Now Casting and Early Estimates," *SSRN Electronic Journal*.

Scott, S. L. and H. R. Varian (2014): "Predicting the present with Bayesian structural time series," *International Journal of Mathematical Modelling and Numerical Optimisation*, 5, 4–23.

Sims, C. A. (2002): "The role of models and probabilities in the monetary policy process," in *Brookings Papers on Economic Activity*, 2, 1–62.

Smith, P. (2016): "Google's MIDAS Touch: Predicting UK Unemployment with Internet Search Data," *Journal of Forecasting*, 35, 263–284.

Zhu, Y. (2010): *Preface to the third issue of Journal of Cardiovascular Disease Research*, vol. 1.

6 Appendix

6.1 Posterior Distributions

Regression Posteriors

$$\beta_\gamma|\sigma^2,\gamma,W,\boldsymbol{\alpha},\boldsymbol{y} \sim N(\tilde{\beta}_\gamma,\sigma^2 V_\gamma)$$
$$V_\gamma = (\Omega_\gamma^{-1} + X_\gamma' X_\gamma)^{-1} = \frac{g}{1+g}(X_\gamma' X_\gamma)^{-1}$$
$$\tilde{\beta}_\gamma = V_\gamma(\Omega_\gamma^{-1} b_\gamma + X_\gamma' y) = V_\gamma X_\gamma' y$$
$$\sigma^2|\gamma,W,\boldsymbol{\alpha},\boldsymbol{y} \sim \Gamma^{-1}(s_T,S_T)$$
$$s_T = \frac{T-1}{2}$$
$$S_T = \frac{1}{2}(y'y + b_\gamma'\Omega^{-1}b_\gamma - \tilde{\beta}_\gamma' V_\gamma \beta_\gamma)$$
$$= (y'y - y'X_\gamma V_\gamma X_\gamma' y)$$

State Variance Posteriors

$$W|\beta,\gamma\sigma^2,\boldsymbol{\alpha},\boldsymbol{y} \sim \Gamma^{-1}(s_T^s,S_T^s)$$
$$s_T^s = \frac{df_0 + n - 1}{2} = \frac{n-1}{2}$$
$$S_T^s = \frac{ss_0 + \sum_{t=2}^{T}(\mu_t - \mu_{t-1})^2}{2}$$

6.2 Sampling Algorithm

The sampling algorithm involves iterating through the following steps[10]:

1. Choose starting value α_0 and compute $\ddot{y} = \boldsymbol{y} - \boldsymbol{X}\beta_\gamma$ and sample α from $p(\boldsymbol{\alpha}|\boldsymbol{y},\gamma,\beta,\sigma^2,W)$. Simulate state draws from the Kohn & Carter (1994) algorithm.
2. Compute $\bar{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{\alpha}$ and sample $(\gamma,\beta\,\sigma^2)$ from the posterior $p(\beta|\sigma^2,\gamma,\boldsymbol{y},\boldsymbol{\alpha},W)$ $p(\sigma^2|\gamma,\boldsymbol{y},\boldsymbol{\alpha},W)p(\gamma|\boldsymbol{y},\boldsymbol{\alpha},\boldsymbol{W})$:
   (a) Sample each element in γ_j from γ from $p(\gamma_j = 1|\gamma_{\gamma/j})$:
   $$p(\gamma_j = 1|\gamma_{/j},\boldsymbol{y}) = \frac{1}{1+\frac{1-\pi}{\pi}R_j}, R_j = \frac{p(\boldsymbol{y}|\gamma_j = 0,\gamma_{/j})}{p(\boldsymbol{y}|\gamma_j = 1,\gamma_{/j})}$$
   (b) Sample the observation error variance σ^2 from $\Gamma^{-1}(s_T,S_T)$.
   (c) Draw γ and set $\beta_j = 0$ if $\gamma_j = 0$. Sample non-zero elements from $N(\tilde{\beta},\sigma^2 V)$
   (d) Sample W from $\Gamma^{-1}(s_T^s,S_T^s)$
3. Repeat 1.-2. until convergence.

The relative posterior inclusion probabilities in R_j are computed using the marginal likelihood formula in (12). We chose starting value $\alpha_0 = 0.5$ and an uninformative state variance prior.

[10]Note, that we suppress the state conditionals in step 2. for readibility

6.3 Tables

	Top 10 Models by Likelihood									
Variables	1	2	3	4	5	6	7	8	9	10
hours-0	-0.2713	-0.2749	0	0	0	0	-0.4921	0	0	-0.6062
	(0.03832)	(0.03832)	(0.03832)	(0.03832)	(0.03832)	(0.03832)	(0.03832)	(0.03832)	(0.03832)	(0.03832)
hours-1	0	0	0	0	0	0	0	0.9860	0	0
	(0.07958)	(0.07958)	(0.07958)	(0.07958)	(0.07958)	(0.07958)	(0.07958)	(0.07958)	(0.07958)	(0.07958)
hours-2	0	0	0	0	-0.3407	-0.3867	0	0	0	0
	(0.09564)	(0.09564)	(0.09564)	(0.09564)	(0.09564)	(0.09564)	(0.09564)	(0.09564)	(0.09564)	(0.09564)
BAA-0	0	0	-0.3407	-0.3867	0	0	0	0	0	0
	(0.14396)	(0.14396)	(0.14396)	(0.14396)	(0.14396)	(0.14396)	(0.14396)	(0.14396)	(0.14396)	(0.14396)
BAA-1	0	0	-0.4643	-0.5216	-0.2380	-0.1420	-0.2065	0	0	-0.2045
	(0.18712)	(0.18712)	(0.18712)	(0.18712)	(0.18712)	(0.18712)	(0.18712)	(0.18712)	(0.18712)	(0.18712)
BAA-2	-0.2534	-0.2341	0	0	0	0	0	-0.2385	-0.1892	0
	(0.13392)	(0.13392)	(0.13392)	(0.13392)	(0.13392)	(0.13392)	(0.13392)	(0.13392)	(0.13392)	(0.13392)
M2-0	-0.0075	-0.0084	-0.0068	-0.0067	-0.0066	-0.0058	-0.0073	0	-0.0085	-0.0066
	(0.00499)	(0.00499)	(0.00499)	(0.00499)	(0.00499)	(0.00499)	(0.00499)	(0.00499)	(0.00499))	(0.00499)
PCE-2	0.5618	0.3872	0.5425	0.4305	0.5790	0.4856	0.6564	0	0.5103	0.4124
	(0.14029)	(0.14029)	(0.14029)	(0.14029)	(0.14029)	(0.14029)	(0.14029)	(0.14029)	(0.14029)	(0.14029)
Unrate-2	0	0	0	0	0	0	0	-0.0326	0	0
	(0.01483)	(0.01483)	(0.01483)	(0.01483)	(0.01483)	(0.01483)	(0.01483)	(0.01483)	(0.01483)	(0.01483)
GT - brentoil - 0	0	0	0	0	0	0	0	-0.0326	0	0
	(0.00080)	(0.00080)	(0.00080)	(0.00080)	(0.00080)	(0.00080)	(0.00080)	(0.00080)	(0.00080)	(0.00080)
GT - industrialproduction - 1	0	0	0	0	0	0	0	0.0298	0	0
	(0.00554)	(0.00554)	(0.00554)	(0.00554)	(0.00554)	(0.00554)	(0.00554)	(0.00554)	(0.00554)	(0.00554)
GT - dowjones - 0	0	0	-0.0005	0	0	0	0	0	0	0
	(0.00185)	(0.00185)	(0.00185)	(0.00185)	(0.00185)	(0.00185)	(0.00185)	(0.00185)	(0.00185)	(0.00185)
GT - financialcrisis - 1	0.0378	0.0395	0.0388	0.0343	0.0307	0.0350	0.0356	0	0.0503	0.0498
	(0.00913)	(0.00913)	(0.00913)	(0.00913)	(0.00913)	(0.00913)	(0.00913)	(0.00913)	(0.00913)	(0.00913)
GT - recession - 1	-0.0234	-0.0213	-0.0280	-0.0264	-0.0254	-0.0253	-0.0318	0	-0.0260	-0.0204
	(0.00480)	(0.00480)	(0.00480)	(0.00480)	(0.00480)	(0.00480)	(0.00480)	(0.00480)	(0.00480)	(0.00480)
GT - merger - 0	0	0	0	0	-0.0184	0	0	0	0	0
	(0.00067)	(0.00067)	(0.00067)	(0.00067)	(0.00067)	(0.00067)	(0.00067)	(0.00067)	(0.00067)	(0.00067)
GT - merger - 1	0	0	0	-2.33E-05	0	0	0	0	0	0
	(0.00110)	(0.00110)	(0.00110)	(0.00110)	(0.00110)	(0.00110)	(0.00110)	(0.00110)	(0.00110)	(0.00110)
GT - real gdp growth - 0	-0.0477	-0.0467	-0.0478	-0.0444	-0.0473	-0.0605	-0.0499	0	-0.0556	-0.0703
	(0.02819)	(0.02819)	(0.02819)	(0.02819)	(0.02819)	(0.02819)	(0.02819)	(0.02819)	(0.02819)	(0.02819)
GT - insolvency - 0	0	0	0	0	0	0	0	-0.0521	0	0
	(0.01080)	(0.01080)	(0.01080)	(0.01080)	(0.01080)	(0.01080)	(0.01080)	(0.01080)	(0.01080)	(0.01080)
GT - insolvency - 2	0	0	0	0	-0.0005	0	0	0	0	0
	(0.00055)	(0.00055)	(0.00055)	(0.00055)	(0.00055)	(0.00055)	(0.00055)	(0.00055)	(0.00055)	(0.00055)

Table (1) Posterior coefficients draws of the top 10 models by marginal likelihood. Posterior standard deviations are reported in parentheses below.

Vintage	Timing	Release	Variable Name	Pub. lag	Transformation	FRED Code
1	Last day of month 1	Fed. funds rate & credit spread	fedfunds & baa	m	3	FEDFUNDS & BAAY10
2	Last day of month 1	Google Trends		m	4	-
3	1st bus. day of month 2	Economic Policy Uncertainty Index	uncertainty	m-1	1	USEPUINDXM
4	1st Friday of month 2	Employment situation	hours & unrate	m-1	2	AWHNONAG & UNRATE
5	Middle of month 2	CPI	cpi	m-1	2	CPI
6	15th-17th of month 2	Industrial Production	indpro	m-1	2	INDPRO
7	3rd week of month 2	Credit & M2	loans & m2	m-1	2	LOANS & M2
8	Later part of month 2	Housing starts	housst	m-1	1	HOUST
9	Last week of month 2	PCE & PCEPI	pce & pce2	m-1	2	PCE & PCEPI
10	Last day of month 2	Fed. funds rate & credit spread	fedfunds & baa	m	3	FEDFUNDS & BAAY10
11	Last day of month 2	Google Trends		m	4	-
12	1st bus. day of month 3	Economic Policy Uncertainty Index	uncertainty	m-1	1	USEPUINDXM
13	1st bus. day of month 3	Construction starts	construction	m-2	1	TTLCONS
14	1st Friday of month 3	Employment situation	hours & unrate	m-1	2	AWHNONAG & UNRATE
15	Middle of month 3	CPI	cpi	m-1	2	CPI
16	15th-17th of month 3	Industrial Production	indpro	m-1	2	INDPRO
17	3rd week of month 3	Credit & M2	loans & m2	m-1	2	LOANS & M2
18	Later part of month 3	Housing starts	housst	m-1	1	HOUST
19	Last week of month 3	PCE & PCEPI	pce & pce2	m-1	2	PCE & PCEPI
20	Last day of month 3	Fed. funds rate & credit spread	fedfunds & baa	m	3	FEDFUNDS & BAAY10
21	Last day of month 3	Google Trends		m	4	-
22	1st bus. day of month 4	Economic Policy Uncertainty Index	uncertainty	m-1	1	USEPUINDXM
23	1st bus. day of month 4	Construction starts	construction	m-2	1	TTLCONS
24	1st Friday of month 4	Employment situation	hours & unrate	m-1	2	AWHNONAG & UNRATE
25	Middle of month 4	CPI	cpi	m-1	2	CPI
26	15th-17th of month 4	Industrial Production	indpro	m-1	2	INDPRO
27	3rd week of month 4	Credit & M2	loans & m2	m-1	2	LOANS & M2
28	Later part of month 4	Housing starts	housst	m-1	1	HOUST
29	Last week of month 4	PCE & PCEPI	pce & pce2	m-1	2	PCE & PCEPI
30	Later part of month 5	Housing starts	housst	m-2	1	HOUST

Table (2) Pseudo real time calendar based on actual publication dates. Transformation: 1 = monthly change, 2 = monthly growth rate, 3 = no change, 4 = LOESS decomposition. Pub. lag: m = refers to data for the given month within the reference period, m-1 = refers to data with a months' lag to publication in the reference period, m-2 = refers to data with 2 months' lag to publication in the reference period.

Varname	Mean	Min	Max	Std.
hours	-0.01	-1.16	1.47	0.29
baa	6.66	4.2	9.3	1.36
loans	0.43	-2.85	3.63	0.89
cpi	0.18	-1.76	1.37	0.26
fedfunds	2.61	0.07	6.54	2.25
housst	0.41	-343	279	94.18
indpro	0.15	-4.33	2.05	0.64
m2	34.90	-37.70	213.50	30.40
pce	0.38	-1.32	2.76	0.39
pce2	0.15	-1.13	0.96	0.18
const	2801	-35597	29267	10711.9
unrate	-0.17	-8.51	8	2.65
uncertainty	0.07	-81.11	103.77	19.98

Table (3) Monthly Macro Summary Stats

Google Trend	Mean	Min	Max	Std.
bankdefault	0.03	-21.50	38.61	8.94
bankruptcy	0.03	-15.23	27.60	5.10
default	0.07	-36.96	40.32	12.44
derivatives	-0.01	-13.64	22.48	4.55
dow jones	-0.16	-17.03	45.58	7.13
economic condition	-0.01	-39.17	46.98	11.82
fed loans	0.02	-13.63	23.25	7.52
fed	0.03	-21.50	38.61	8.94
federal reserve system	0.01	-8.13	19.99	4.12
financial crisis	0.03	-20.66	33.55	6.32
foreclosure	0.00	-20.69	60.62	6.73
gdp growth	6.87	-20.69	60.62	7.89
house prices	0.01	-11.82	44.03	5.29
industrial production	0.02	-8.04	15.60	3.41
insolvency	0.15	-6.5	22.78	6.34
jobless	0.29	-7.9	13.01	5.34
jobless claims	0.59	-8.99	5.97	2.74
jobs	0.81	-4.99	14.46	6.91
loans	0.01	-14.59	20.97	5.92
real gdp growth	-0.03	-17.31	31.40	8.22
stockmarket	-0.05	-15.66	8.41	7.69
us economy	-0.09	-36.90	26.58	8.74
us default	-38.39	0.00	19.17	5.22
recession	0.00	-12.29	24.01	3.10
stock market	0.01	-6.48	27.50	6.14
unemployment	0.01	-34.60	29.98	7.38
US default	-0.01	-20.09	35.51	8.02
US growth	-0.06	-23.86	61.20	7.90

Table (4) Google Trends Summary Statistics - GDP

6.4 Figures

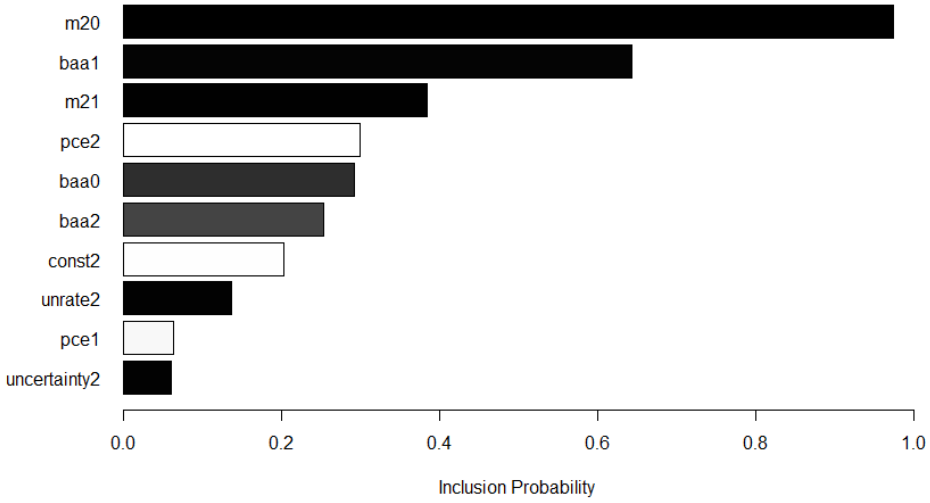


Figure (9) Marginal posterior inclusion probabilities for F-MFBSTS. The prefix GT signifies variables which are Google Trends. The number [0,1,2] appended to a variable indicates the position in a given quarter, with 0 being the latest month.

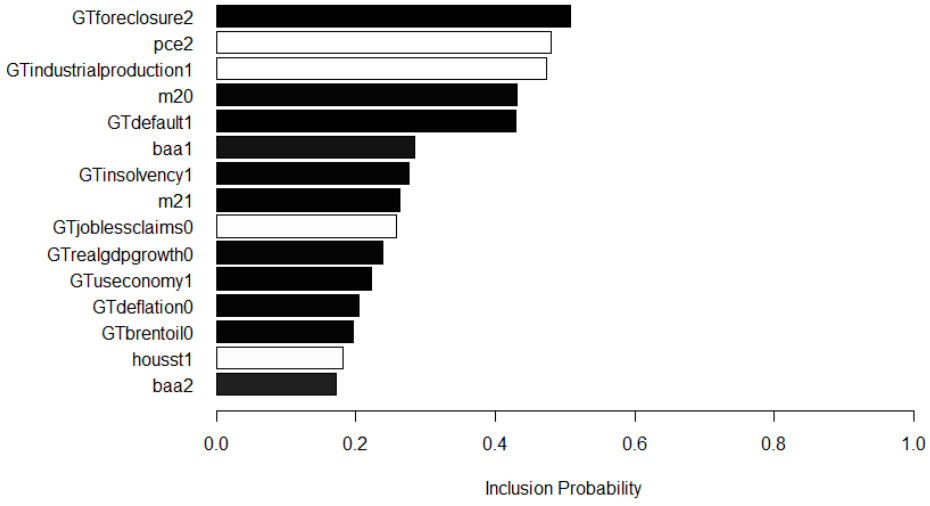


Figure (10) Marginal posterior inclusion probabilities for F-MFBSTS based on $\pi = 18$. The prefix GT signifies variables which are Google Trends. The number [0,1,2] appended to a variable indicates the position in a given quarter, with 0 being the latest month.