Nowcasting US GDP with artificial neural networks

Loermann, Julius and Maas, Benedikt

May 2019

Nowcasting US GDP with artificial neural networks

Julius Loermann*   Benedikt Maas†

May 2019

Abstract

We use a machine learning approach to forecast the US GDP value of the current quarter and several quarters ahead. Within each quarter, the contemporaneous value of GDP growth is unavailable but can be estimated using higher-frequency variables that are published in a more timely manner. Using the monthly FRED-MD database, we compare the feedforward artificial neural network forecasts of GDP growth to forecasts of state of the art dynamic factor models and the Survey of Professional Forecasters, and we evaluate the relative performance. The results indicate that the neural network outperforms the dynamic factor model in terms of now- and forecasting, while it generates at least as good now- and forecasts as the Survey of Professional Forecasters.

JEL classification: C32, C53, C55, E32

*Hamburg University, Department of Economics, Von-Melle-Park 5, 20146 Hamburg, Germany; E-mail: *Julius.Ferdinand.Loermann@uni-hamburg.de*

†Hamburg University, Department of Economics, Von-Melle-Park 5, 20146 Hamburg, Germany; E-mail: *Benedikt.Maas@uni-hamburg.de*

1. Introduction

Policy makers regularly request information on the current state of the economy. This also applies to central bankers before they adjust their monetary policy stance, as well as to economists who work in major finance departments and who have to make a decision about a short term budget plan. However, because GDP is measured on a quarterly frequency, often with considerable time lags, the GDP growth of the current period needs to be estimated as accurately as possible. This current quarter forecast is often referred to as nowcasting, as defined for instance by Bańbura et al. (2013). When the nowcast of GDP growth is conducted for the current quarter, more timely and higher frequency information are available. The combination of several monthly indicators might help to extract signals about the current state of the economy. Unfortunately this approach does face some difficulties because monthly indicators are also only available with certain publication lags, leading to the so called "ragged edge" problem that was first described by Wallis (1986).

To overcome these challenges, central banks commonly apply the dynamic factor model (DFM) approach of Giannone et al. (2008). This modeling framework shrinks information from a large dataset into a few underlying factors, while at the same time applying Kalman-Filtering techniques to fill up the missing data of the ragged edge within the dataframe. Although this approach provides a unified framework incorporating dynamic imputation and nowcasting, the central predictor equation remains linear, making the model potentially unfit to generalize to non-linear patterns.

This paper addresses several of the above described issues by applying a machine learning framework to now- and forecasting the GDP growth of the United States between 1999 and 2018. Specifically, an artificial neural network (ANN) approach is chosen within the combined filter and wrapper approach of Crone and Kourentzes (2010) and Kourentzes et al. (2014). The resulting multilayer perceptron (MLP) is a highly non-linear, flexible and dynamic framework that enables us to automatically fit the most appropriate neural network architecture, while simultaneously allowing for the selection of the most relevant monthly indicators for now- and forecasting the current and future values of GDP growth. The MLP approach to now- and forecasting is applied to the real-time monthly vintages of the FRED-MD database, enabling us to conduct quasi real-time forecasts of US GDP growth between 1999 and 2018.

While no other study of nowcasting has applied ANNs, or machine learning algorithms in general, several other studies have applied neural networks to forecast macroeconomic and financial variables. For example, Tkacz (2001) uses neural networks to forecast the Canadian GDP growth rate between 1989 and 1992 by applying lagged GDP growth and several other financial variables, such as yield spreads and monetary aggregates. While Tkacz (2001) uses quarterly data to forecast Canadian GDP by using ANNs, Heravi et al. (2004) apply ANNs to forecast monthly industrial production between 1978 and 1995 for the three largest European economies, with data provided by Eurostat. In contrast to Tkacz (2001), Heravi et al. (2004) only use the information contained in the lagged values of monthly y-o-y growth of industrial production. Most recently, Jung et al. (2018) use recurrent neural networks, elastic nets and super learners to forecast GDP growth of seven major advanced and developing economies. They combine the World Economic Outlook Database from the IMF with data from the International Country Risk Guide. These combined databases contain quarterly and annual data of records of national accounts, monetary, trade, labor market variables as well as business and consumer confidence indices and several risk metrics. The time span ranges from 1970 to 2010, and forecasts are conducted up until 2010. Besides these studies, ANNs have also been used in financial forecasting. For instance, Kuan and Liu (1995) investigate the forecasting ability of ANNs and recurrent ANNs for daily exchange rates of five major currencies between 1980 and 1985. More recently Torres and Qiu (2018) apply recurrent neural networks to daily data of several crypto-currencies, exchange rates, commodities and stocks between 2013 and 2017.

Our contribution is to demonstrate that feedforward artificial neural networks can be applied to nowcasting, as well as to forecasting. The reason why we focus on feed forward ANNs, in contrast to more sophisticated recurrent ANNs, is that there are several efficient procedures to input variable and network architecture selection. By applying the combined filter and wrapper approach of Crone and Kourentzes (2010) and Kourentzes et al. (2014), we demonstrate how the implementation of ANNs can be automatized in the process of now- and forecasting. To demonstrate the usefulness of the ANN approach, we conduct a now- and forecasting competition between the standard DFM methodology of Giannone et al. (2008) and Survey of Professional Forecasters (SPF). Our main findings can be summarized as follows: the applied ANN approach beats the DFM significantly in now- and forecasting, when the evaluation metrics are the root mean squared forecast error (RMSFE) and mean absolute forecast error (MAFE). Compared to the SPF, the ANN produces smaller RMSFEs and MAFEs in now- and short term forecasting; however, the results are not significantly different from zero. Therefore, the ANN performs better than the DFM and is at least as good as the SPF.

The remainder of the paper is organized as follows. Section 2 describes the DFM methodology of Giannone et al. (2008) and the combined filter and wrapper approach of Crone and Kourentzes (2010) and Kourentzes et al. (2014). Section 3 describes the applied data, especially the real-time vintages of the FRED-MD database and the SPF data, and the real-time forecasting setup. Section 4 presents the empirical results, while Section 5 concludes.

2. Econometric framework

2.1. Dynamic factor model

To exploit information of many monthly potential predictor variables and obtain an early estimate of quarterly GDP growth, Giannone et al. (2008) combine a DFM and the Kalman smoother. Their two-step approach solves three problems of nowcasting: it deals with the mixed frequency issue of combining monthly predictor variables and quarterly GDP, it can handle a large number of potential predictor variables, and it can cope with the ragged edge problem of the underlying data. Additionally, it has the potential to capture the essential dynamics of the time series of the panel.

The rest of this section summarises the approach of Giannone et al. (2008). The theory behind the two-step estimator is derived in Doz et al. (2011).

Let $x_{t|v_j}$ be an $n \times 1$ vector of stationary monthly indicator variables available for the vintage v_j, which is transformed so as to correspond to a quarterly quantity when observed at the end of the quarter.[1] Giannone et al. (2008) assume the following factor structure of the transformed monthly indicators:

$$x_{t|v_j} = \mu + \lambda F_t + \varepsilon_{t|v_j}, \tag{2.1}$$

where μ is a constant, F_t is an $r \times 1$ vector of common factors, λ is an $n \times r$ matrix of factor loadings, and $\varepsilon_{t|v_j}$ is an $r \times 1$ vector of idiosyncratic components. It is assumed that the common components given by $\chi_t = \lambda F_t$ are linear functions of a few $r < n$ unobserved common factors that capture most of the variation of the underlying dataset, while the idiosyncratic components are driven by variable-specific shocks. The dynamics of the factors are modeled as follows:

$$F_t = AF_{t-1} + Bu_t, \tag{2.2}$$

where B is a $r \times q$ matrix of full rank q, A is a $r \times r$ matrix with all roots of $det\,(I_r - Az)$ lying outside the unit circle, and u_t is a $q \times 1$ vector of white noise shocks to the common factors. The idiosyncratic error term vector $\varepsilon_{t|v_j}$ is assumed to be white noise, cross-sectionally orthogonal, as well as orthogonal to the common shock vector u_t. In terms of the parametrisation, Giannone et al. (2008) choose two static factors and two common shocks, hence $r = q = 2$, which we are going to follow.

Equations (2.1) and (2.2) set up a state space framework which allows standard Kalman filter techniques to estimate the common factors. The estimation is conducted as follows: first, by ignoring observations that are not available for all the variables of the dataset, a balanced panel is

[1] See the Appendix of Giannone et al. (2008) for the data transformations.

created from the original ragged edge dataset. Then, principal components are derived from this balanced panel and the parameters of (2.2) are estimated by ordinary least squares (OLS) regression. In the second step, the common factors are estimated by running the Kalman smoother using the entire ragged edge dataset, where true parameters in the state space specification are replaced by parameter estimates. Hence, when no observation is available, the filter produces a forecast of the common factors. Having obtained the estimated factors from the unbalanced panel, the nowcasts of GDP growth finally appear as the fitted values of an OLS regression of the quarterly GDP series on the quarterly estimated factors:

$$\widehat{y}_{t|v_j} = \alpha + \beta^{'}\widehat{F}_{t|v_j}, \tag{2.3}$$

where $\widehat{y}_{t|v_j}$ is the estimated quarterly GDP growth, and $\widehat{F}_{t|v_j}$ are the estimated common factors.

2.2. Artificial neural network

This section describes the set-up of our main nowcasting machine, the ANN. ANNs can be modeled as flexible frameworks, which belong to the class of machine learning algorithms. As with most machine learning techniques, the ANN can be used for classification—which is a predictive exercise in which the dependent variable is qualitative—and for forecasting continuous variables. The description that follows is based on Crone and Kourentzes (2010), Kourentzes et al. (2014) and Ord et al. (2017).

A simple example of a one-layer feed-forward ANN with I input- and H hidden nodes, also called neurons, within a time series context is given by

$$y_{t+1} = \sum_{h=1}^{H} \beta_h g\left(\sum_{i=1}^{I} \gamma_{hi} p_i - \gamma_{0i}\right) - \beta_0, \tag{2.4}$$

where $\mathbf{p} = [y_t, ..., y_{t-n}, \mathbf{x'_{t+1}}, ..., \mathbf{x'_{t-k}}]$ is the vector of inputs containing lags of the dependent variable y_{t+1} and contemporaneous, as well as lagged values of further explanatory variables $\mathbf{x'_{t+1}}$. The one layer perceptron, as it is frequently called, can be generalised to a MLP. The coefficient vectors $\beta = [\beta_1, ..., \beta_H]$ and $\gamma = [\gamma_1, ..., \gamma_I]$ are the so called output layer and hidden layer weights. The two coefficients β_0 and γ_{0i} are called biases. A bias decides whether or not a neuron gets activated. A neuron is said to be activated whenever for a given h, the weighted sum of the previous layer neurons—in this case, when $\sum_{i=1}^{I} \gamma_{hi} p_i$ exceeds its bias. The function $g()$ is referred to as the squashing function and it maps the weighted sum of previous layer neurons minus bias to some interval (a,b), typically $(-1,1)$ or $(0,1)$. Hence, common choices for $g()$

are the hyperbolic tangent, $g(x) = \frac{e^{2x}-1}{e^{2x}+1}$, or the sigmoid, $g(x) = \frac{e^x}{e^x+1}$. The reason behind using a squashing function is purely practical because networks tend to train better when neurons can only take values from limited data ranges (see Ord et al. (2017)). When a variable is predicted by the ANN, the final output is retransformed to its original scale by applying the squashing functions inverse $g^{-1}(x)$ within the output layer. The output and input variables can be of any scale, meaning that (as within a linear regression) the networks weights would account for scaling. However, Ord et al. (2017) argue that, similar to the application of the squashing function, networks tend to train better when the in- and outputs are bijectively transformed into a common scale. When all of the variables are on a common scale, the networks weights' only task is to capture the non-linear relationship between the inputs and the output. An additional advantage is that neurons are less likely to be saturated; that is, for example for the logistic squashing function taking values close to 1 (0), regardless of the input being for example 100000 and 10000000 (-100000 and -10000000). The most common scaling practice is to use the generalised min-max transformer

$$z(x) = (b-a)\frac{x - min(x)}{max(x) - min(x)} + a, \tag{2.5}$$

which maps a metric variable x into the interval $[a,b]$. The question of the right values for a and b is an empirical one and depends upon the forecasting performance of the resulting model. Given an ANN specification, the interval that maximises forecast accuracy on some validation data can be chosen. Typical choices to start with are $a = 0$ and $b = 1$.

The described ANN can be interpreted as a flexible and highly parametrised non-linear autoregressive distributed lag model (ARDL). The weights and biases are estimated or, in the language of machine learning, trained. The most commonly used training method is the backpropagation algorithm of Rumelhart et al. (1986). This is a variant of gradient descent in which, after a random initialisation, the sum of squared errors between predicted and actual outcome are minimised by adjusting the weights and biases of the ANN. The detected minimum is a local one. Note, however, that in the context of machine learning finding, the global minimum is not desirable because the trained model would over-fit the training data, leading to poor out-of-sample performance.

Before the ANN is used for the nowcasting exercise, the input variables—that is, the number of lags of y_{t+1} and the additional explanatory variables and their lag structure $\mathbf{x'_{t+1}}, \ldots, \mathbf{x'_{t-k}}$—and the network architecture—that is, the number of hidden layers, the number of neurons per layer and the type of the squashing function—have to be selected. Crone and Kourentzes

(2010) provide a combined filter and wrapper approach to select these features of ANNs for the purpose of time series forecasting. The filter selects the input variables before the network architecture is selected and the wrapper selects the network architecture given the input selection afterwards. Lachtermacher and Fuller (1995) demonstrate that an efficient way of input variable selection for ANNs is by applying stepwise regressions; that is, linear ARDL models. Within this procedure, all of the variables and lags that are not significant on a 5 % level are deleted step-by-step. However, stepwise selection is only valid in a statistical sense when the applied time series is stationary (see Sims et al. (1990)), which is the case here as the FRED-MD series are transformed to stationarity (see Appendix A.1 for details).[2]

Given the selected input variables, the network architecture is selected by applying the following wrapper: the data are split into a training set, containing 80 % of the sample, and a validation set, containing the remaining 20 %. Afterwards, different network architectures with varying numbers of hidden layers and neurons per layer are trained on the training set by applying the backpropagation algorithm. Then, the predictive performance of the different networks is evaluated on the validation set. The network architecture resulting in the smallest mean squared error (MSE) on the validation set is selected. Because the ANN is able to approximate any type of function, there is the danger of overfitting the test data when the network architecture becomes more complex. The minimisation of the MSE on the validation set takes this concern into account and therefore reduces the negative predictive effects of overfitting. However, the danger of overfitting on the validation set remains. To minimise that potential effect, we apply a 5-fold-cross validation scheme and minimise the MSE on the different folds.[3]

As discussed earlier, the final network architecture and the inputs are selected by the combined filter and wrapper approach of Crone and Kourentzes (2010). When the final network is trained via the backpropagation algorithm, the resulting coefficients correspond to a local minimum of the loss function. Because there are multiple local minima, the random initialisation of the backpropagation algorithm results in different weights and therefore different forecasts. Kourentzes et al. (2014) suggest to retrain the network for multiple random initialisa-

[2] It should be noted that the full filter procedure of Crone and Kourentzes (2010) includes the construction of further input variables that account for deterministic seasonality and deterministic trends. This variable construction is refereed to as an interative neural filter (INF) because it applies an ANN on the target variable using trigonometric functions containing the seasonal length of the inputs. The seasonal length is detected by minimising the euclidean distance between vectors of the target variable, which have been constructed by splitting the target vector into subvectors of different (equal) length. The INF accounts in a flexible way for deterministic seasonality of y_{t+1} and $\mathbf{x'_{t+1}}$. However, the FRED-MD data are deseasonalised and transformed to stationarity. Therefore, the INF procedure is not applied in this paper.

[3] Besides the training of the weights and biases within each network, the selection of the number of neurons and layers, as well as the selection of the squashing function and the transformation interval of the inputs and the target is referred to as supervised learning in the machine learning terminology.

tions, which results in a distribution of forecasts. The mode of a kernel density estimate of the different forecasts shows superiority in terms of forecast accuracy when compared to a single forecast applying only one trained network. This so called ensemble operator approach is also applied here for 100 different random initialisations.

Finally, to deal with the ragged edge problem, we fill up missing values by applying univariate $ARMA(p,q)$ forecasts of each single time series. The lag-lengths of the $ARMA(p,q)$ models are selected via Akaike information criterion (AIC).

3. Data and forecasting design

3.1. Data

The data behind the MLP and the DFM comes from FRED-MD, the monthly database for Macroeconomic Research of the Federal Reserve Bank of St. Louis, which is described extensively in McCracken and Ng (2016).

FRED-MD is a large macroeconomic database that i designed for the empirical analysis of "big data". The database is publicly available and updated in real-time on a monthly basis.[4] It consists of 134 monthly time series and is classified into eight categories: (1) output and income, (2) labor market, (3) housing, (4) consumption, orders and inventories, (5) money and credit, (6) interest and exchange rates, (7) prices and (8) stock market. A full list of the data and its transformation is given in Appendix A.1. The time series start in January 1959 and vintages of the whole database are available since August 1999.

3.2. Real-time forecasting setup

For the training of the MLP and the estimation of the DFM, we use the information available at the end of the second month of the quarter, hence we use the initial releases of the FRED-MD database at the end of February, May, August, and November. We estimate the models recursively using only information available at each point where the nowcasts are computed. The first nowcast is conducted for 1999Q3.

[4]The FRED-MD database is available for download under the following link: `https://research.stlouisfed.org/econ/mccracken/fred-databases/`.

3.3. Survey of Professional Forecasters

The US GDP growth forecasts of the MLP and DFM are compared with the median forecasts of the SPF.

The SPF is the oldest survey of macroeconomic forecasts in the United States, and it is conducted and published by the Federal Reserve Bank of Philadelphia.[5]

The SPF is a quarterly forecast and is released around the 15th of the second month in the middle of every quarter; that is, mid-February, mid-May, mid-August and mid-November. The forecasters publish a nowcast of the current quarter, and one-quarter- up to four-quarters-ahead forecasts.

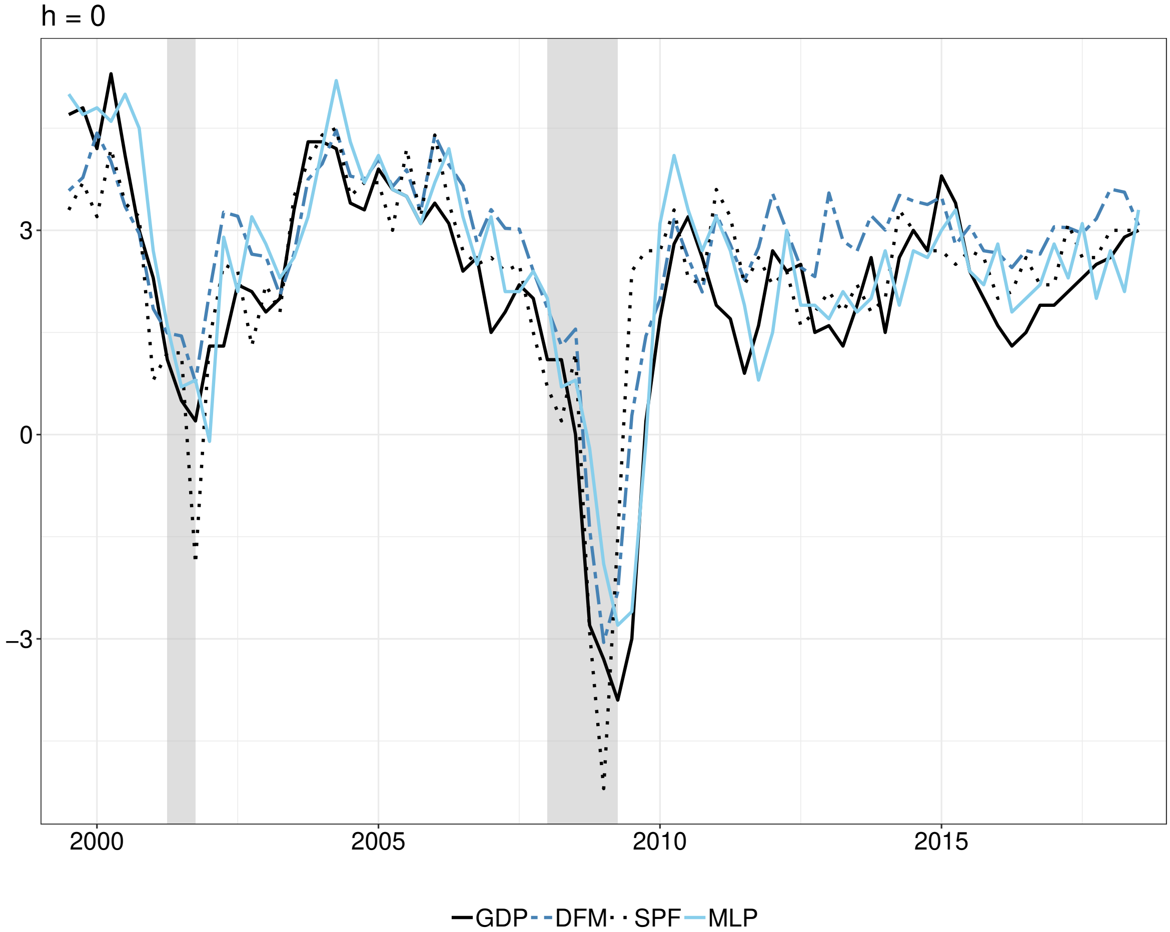
4. Empirical results

Figure 1 plots the actual realisation of GDP growth together with the nowcasts (horizon $h = 0$) of the SPF, the DFM and the MLP. It can be seen that the SPF is sometimes over-pessimistic at the end of recessions, while being fairly accurate at the beginning. For instance, during the 2008 recession, the SPF nowcast tracks the actual GDP growth tightly. Between 2007:Q3 and 2008:Q3—with the exception of 2008:Q2 where it is with a nowcast of 1.2 % GDP growth above the actual GDP growth too over-optimistic—, it is fairly accurate with absolute errors between 0.1 % and 0.5 %. In contrast, in 2009:Q1, the SPF overestimates the actual GDP decline of -3.3 % with -5.2 % by a large extent.

The DFM nowcasts are most of the time above the GDP growth and hence too optimistic. This is especially true during the 2008 recession and the recovery period that followed. For example, between 2010 and 2015, and between 2015 and 2018 the DFM nowcasts are almost always above the actual realisations and compared to the SPF and the MLP the DFM is the furthest away from actual GDP growth. The MLP is fairly accurate during the 2008 recession because its nowcast error is most often smaller compared to the SPF. Exceptions are 2008:Q1 and 2008:Q3, where the SPF produces nowcast errors of -0.4 and 0.1 percentage points, while the MLP produces nowcast errors of -0.9 and -2.6 percentage points. In the following expansionary period between 2009 and 2018, the MLP nowcast is close to the SPF nowcast, while it produces superior nowcasts approximately half of the time.

[5] More information about the SPF can be found in Croushore (1993). All of the SPF releases can be found under the following link: `https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/`.

Figure 1: Realised GDP growth versus nowcasts.

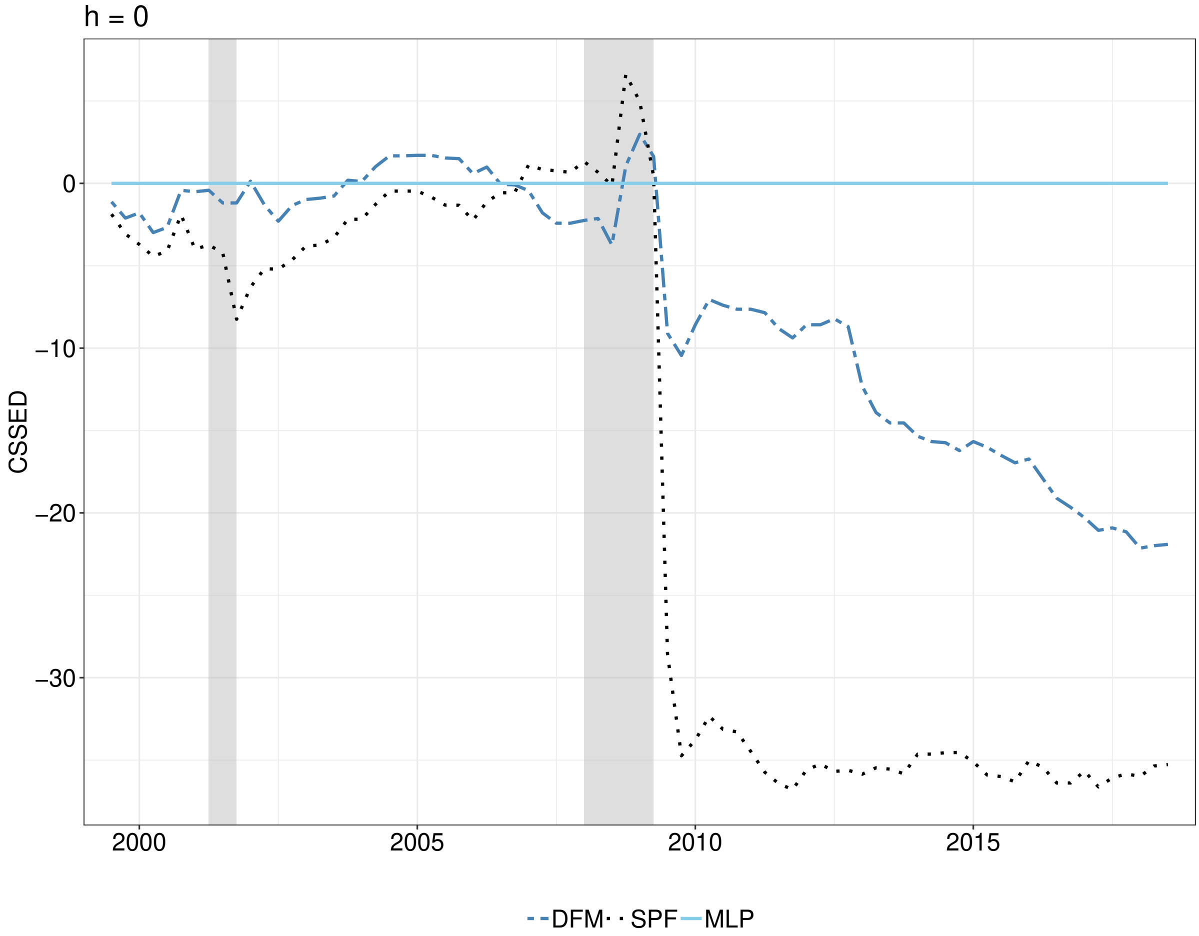


Notes: Realised GDP growth versus nowcasts of the DFM, the SPF, and the MLP. NBER recessions are highlighted by gray shading.

These findings mirror themselves in Figure 2, where the cumulative sum of squared forecast error differences (CSSED) between the DFM, the SPF and the MLP are plotted. The baseline model here is the MLP such that the squared error difference between the SPF versus MLP and the DFM versus the MLP are added up and plotted. Whenever the dotted black line (dashed-blue line) is below the zero line, the MLP outperforms the SPF (DFM) in terms of nowcasting.

It is apparent that, especially before and after the 2008 recession, the MLP is superior to the SPF. Shortly before and up to the middle of the 2008 recession, the SPF and the MLP are equally accurate, while the described precision of the SPF shows up in CSSED values larger than zero from the mid of the recession until its end. The major fall of the CSSED value can be attributed to the over-pessimistic nowcast of the SPF at the end of the crisis, while the stable negative values reflect the equal nowcasting accuracy in the following recovery period.

Figure 2: CSSED plots for nowcasts.



Notes: This graph shows the CSSED for the nowcast. The CSSED is computed as $CSSED_{m,\tau} = \sum_{\tau=R}^{T} \left(\hat{e}^2_{bm,\tau} - \hat{e}^2_{m,\tau}\right)$, where $\hat{e}^2_{bm,\tau}$ denotes the squared forecast error of the MLP. Values above zero indicate that the alternative model outperforms the MLP, while values below zero mean that the MLP outperforms the competing model. NBER recessions are highlighted by gray shading.

The CSSED for the DFM versus the MLP indicates that before the 2008 recession both are equal in terms of nowcasting GDP growth, while especially after the crisis the MLP gets better and better, which is indicated by the downward sloping dashed blue line.

To summarise the visual analysis of the nowcast horse-race between the SPF, the DFM and the MLP, one can follow that the MLP produces at least as accurate nowcasts as the SPF, while the absolute errors that it makes are less severe. The DFM, however, tends to be too optimistic most of the time. An explanation of the visually detected superiority of the MLP might be stated as follows. The flexibility of the MLP in terms of input variable selection and the possibility of non-linear functional fitting when necessary give it a clear advantage over the DFM. The DFM can only re-weight different variables through the factor extraction procedure while keeping the model linear in its predictor equation, making it unable to fit accurately to potential non-linear periods, such as recessions and subsequent recovery phases. An explanation of the MLPs superiority towards the SPF is more difficult to give because it is not clear what type of model the SPF uses for its nowcast.

The next step is to analyse whether the MLP significantly beats the SPF and the DFM in nowcasting ($h = 0$) and out of sample forecasting ($h > 0$). To start the statistical analysis, the relative forecast performance of the three competitor models is evaluated versus a naive constant growth model. Table 1 reports the relative RMSFE of the naive benchmark model versus the three competitor approaches. Values smaller than one indicate that the competitor model has a smaller RMSFE than the naive constant growth model. In addition, the significance of the relative forecast performance is tested by the Diebold and Mariano (1995) test. The nowcasting/forecasting period ranges from 1999:Q3 to 2018:Q3.

Focusing at first on nowcasting, it is apparent that all three competitors beat the naive benchmark model significantly on a 1 % level at a horizon of $h = 0$. Moreover, one can see that the MLP performs best in terms of relative RMSFE versus the benchmark, while the SPF is ordered second and the DFM last. The same pattern occurs for the one- ($h = 1$) and two-steps-ahead ($h = 2$) forecasting horizons. On a three-steps-ahead horizon, the MLP is the only model of the three competitors, which beats the naive benchmark model on a 5 % significance level. Finally, on a four step ahead horizon, the MLP is again the only model that can significantly beats the benchmark at least on a 10 % level, while the SPF and the DFM are not able to generate superior forecasting performance.

Table 1: Nowcasts and forecasts of GDP: out-of-sample evaluation for DFM, SPF and MLP vs. naive benchmark.

model	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
DFM	0.477^{***}	0.624^{***}	0.812^{***}	0.950	1.035
SPF	0.518^{***}	0.547^{***}	0.724^{***}	0.809	0.878
MLP	0.398^{***}	0.537^{***}	0.694^{***}	0.839^{**}	0.942^{*}

Notes: This table reports the relative RMSFE of GDP growth for the DFM, the SPF, and the MLP relative to a naive constant growth model for GDP. Evaluation sample: 1999Q3 − 2018Q3. A value below one indicates that the competitor model beats the naive benchmark model. The stars denote statistical significance at 10 %(*), 5 %(**) and 1 %(***) level of the Diebold and Mariano (1995) test.

As Table 1 shows, the MLP generates the smallest relative RMSFE towards the naive benchmark model for the nowcasting and for all forecasting horizons up to $h = 4$. The next step is to test whether the MLP is also able to beat the DFM and the SPF directly. Table 2 reports the relative RMSFE of the MLP versus the two competitor models. A value of the relative RMSFE that is smaller than one means that MLP generates a smaller RMSFE than the DFM or the SPF. The evaluation period is again 1999:Q3 to 2018:Q3.

The DFM is significantly outperformed on all horizons. In terms of nowcasting, the relative RMSFE is 0.834 and the difference is significant on a 10 % significance level. In terms of forecasting, the MLP beats the DFM significantly on a 1 % level from a one- up to a four-steps-ahead forecasting horizon. Comparing the nowcasting performance of the SPF with the MLP in terms of relative RMSFE, it is apparent that with a value of 0.768 the MLP outperforms the SPF; however, the result is not statistically significant. Turning towards forecasting, Table 2 shows that the MLP beats the SPF up to a two-steps-ahead forecasting horizon, which is again insignificant. For for three- and four steps ahead forecasting horizons, the SPF generates smaller RMSFEs, which are also insignificant. This indicates that when the RMSFE is used as a forecasting evaluation method, the MLP is at least as good as the SPF in terms of nowcasting and short-term forecasting, while it significantly outperforms the DFM on every horizon.

Table 2: Nowcasts and forecasts of GDP: out-of-sample evaluation for MLP, RMSFE.

model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
DFM	0.834^{*}	0.860^{**}	0.854^{**}	0.884^{***}	0.910^{***}
SPF	0.768	0.982	0.958	1.037	1.073

Notes: This table reports the relative RMSFE of GDP growth for the MLP relative to the DFM and the SPF. Evaluation sample: 1999Q3 − 2018Q3. A value below one indicates that the MLP beats the competitor model. The stars denote statistical significance at 10%(*), 5%(**) and 1%(***) level of the Diebold and Mariano (1995) test.

A similar picture occurs when the MAFE is used as a nowcasting/forecasting evaluation method. This scenario, everything else kept equal, is presented in Table 3. The MLP still beats the DFM in terms of nowcasting performance; however, the relative MAFE is 0.791 and the performance advantage of the MLP is now significant on a 5 % level. When it comes to forecasting, the MLP significantly outperforms the DFM up to a four-quarters-ahead horizon. Again, similar to the RMSFE evaluation presented in Table 2, the MLP generates a lower MAFE compared to the SPF on horizons from $h = 0$ to $h = 2$, while still being insignificant. The conclusion remains the same; the MLP outperforms the DFM on every horizon, while the SPF is insignificantly outperformed when it comes to nowcasting and short-term forecasting.

Table 3: Nowcasts and forecasts of GDP: out-of-sample evaluation for MLP, MAFE.

model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
DFM	0.791^{**}	0.800^{***}	0.845^{***}	0.844^{***}	0.849^{***}
SPF	0.883	1.034	1.089	1.163	1.154

Notes: This table reports the relative MAFE of GDP growth for the MLP relative to the DFM and the SPF. Evaluation sample: 1999Q3 – 2018Q3. A value below one indicates that the MLP beats the competitor model. The stars denote statistical significance at $10\,\%(^{*})$, $5\,\%(^{**})$ and $1\,\%(^{***})$ level of the Diebold and Mariano (1995) test.

To get more granular insights and to test for robustness, Table 4 reports the RMSFE and MAFE of the MLP and, additionally, the relative values of these compared to the respective evaluation metrics of the SPF and the DFM, while the following variable groups have been excluded from the training and evaluation samples: Output and income variables (G1), labor market variables (G2), housing variables (G3), consumption, orders and inventory variables (G4), money and credit variables (G5), interest rates and exchange rates (G6) and stock market variables (G8). In addition, Figure 3 depicts a heatmap as a visual inspection of the specific variables used within the respective quarter by the MLP.[6] All of the variables that appear within the heatmap have been used at least once, while non-depicted variables are not used at all. The baseline results where none of the groups have been removed are presented in the first row of the upper and lower part of the table, respectively.

The nowcasting evaluation presented in Table 4 indicates robustness as the relative RMSFEs and MAFEs do not change by a large extent. However, some differences should be noticed. When the RMSFE is used as the evaluation metric, the nowcasting superiority of the MLP towards the DFM becomes significant on a 5 % level, when labor market variables (G2) or interest rates and exchange rates (G6) are excluded. These two groups of variables are most often used, as can be seen in Figure 3. Because labor market variables tend to be lagging behind production or financial market variables, the performance gain due to exclusion might be explained by efficiency gains due to increased parsimony. In contrast, exchange rates and interest rates are forward looking financial variables and hence should increase the information gain of the MLP. However, these fast moving financial variables tend to be noisy and therefore may reduce the signal extraction by the MLP, ultimately resulting in a lower forecasting performance.

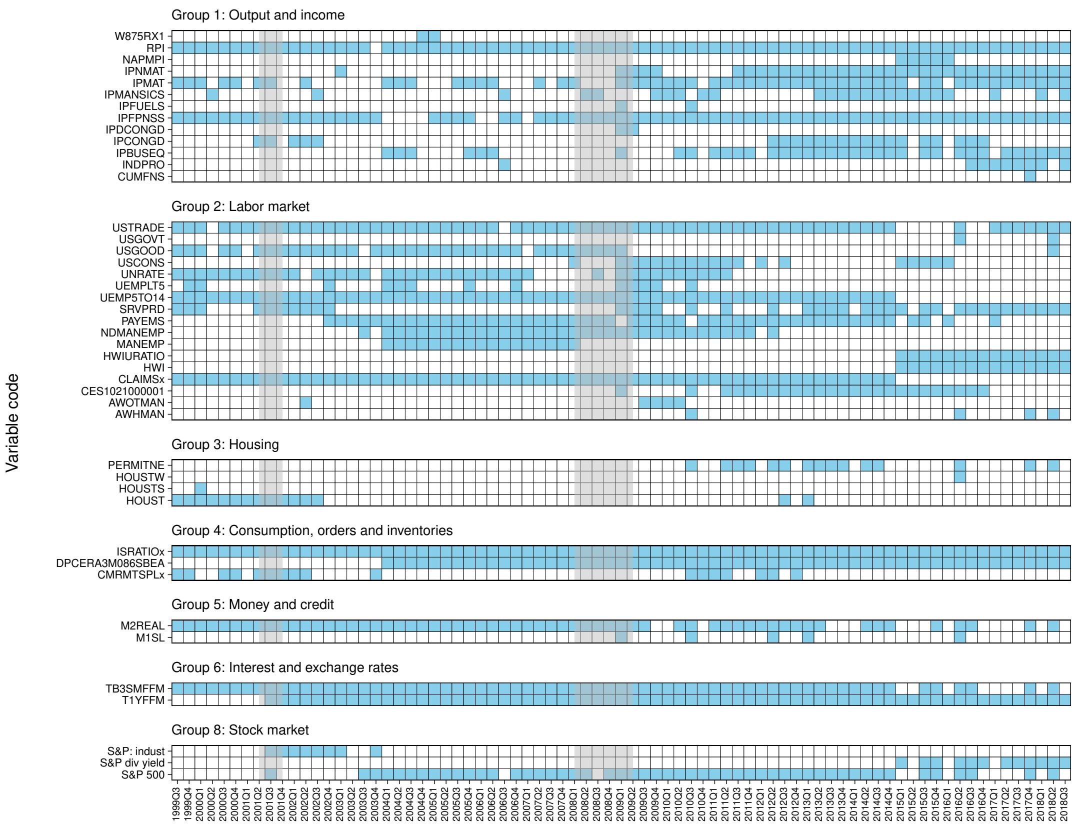
[6] Please note that the structure of the FRED-MD database has slightly changed over time. For details see the historical vintages of FRED-MD 1999-08 to 2014-12.

Table 4: Nowcasts of GDP: out-of-sample RMSFE and MAFE evaluation with different data groups for MLP.

	$h = 0$					
	$RMSFE_{MLP}$	$\frac{RMSFE_{MLP}}{RMSFE_{DFM}}$	$\frac{RMSFE_{MLP}}{RMSFE_{SPF}}$	$MAFE_{MLP}$	$\frac{MAFE_{MLP}}{MAFE_{DFM}}$	$\frac{MAFE_{MLP}}{MAFE_{SPF}}$
All Groups	0.812	0.834*	0.768	0.640	0.791**	0.883
less G1	0.813	0.835*	0.769	0.649	0.802**	0.895
less G2	0.743	0.763**	0.703*	0.597	0.738***	0.823*
less G3	0.848	0.871	0.802	0.661	0.817**	0.912
less G4	0.930	0.955	0.880	0.694	0.858*	0.957
less G5	0.810	0.832*	0.766	0.642	0.794**	0.886
less G6	0.780	0.801**	0.738	0.651	0.805**	0.898
less G7	0.865	0.888	0.818	0.673	0.832**	0.928
less G8	0.832	0.854*	0.787	0.665	0.822**	0.917

Notes: This table reports the RMSFE and MAFE of GDP growth for the MLP. Evaluation sample: 1999Q3 − 2018Q3. The stars denote statistical significance at 10 %(*), 5 %(**) and 1 %(***) level of the Diebold and Mariano (1995) test.

Figure 3: Heatmap of the variables used within the MLP.

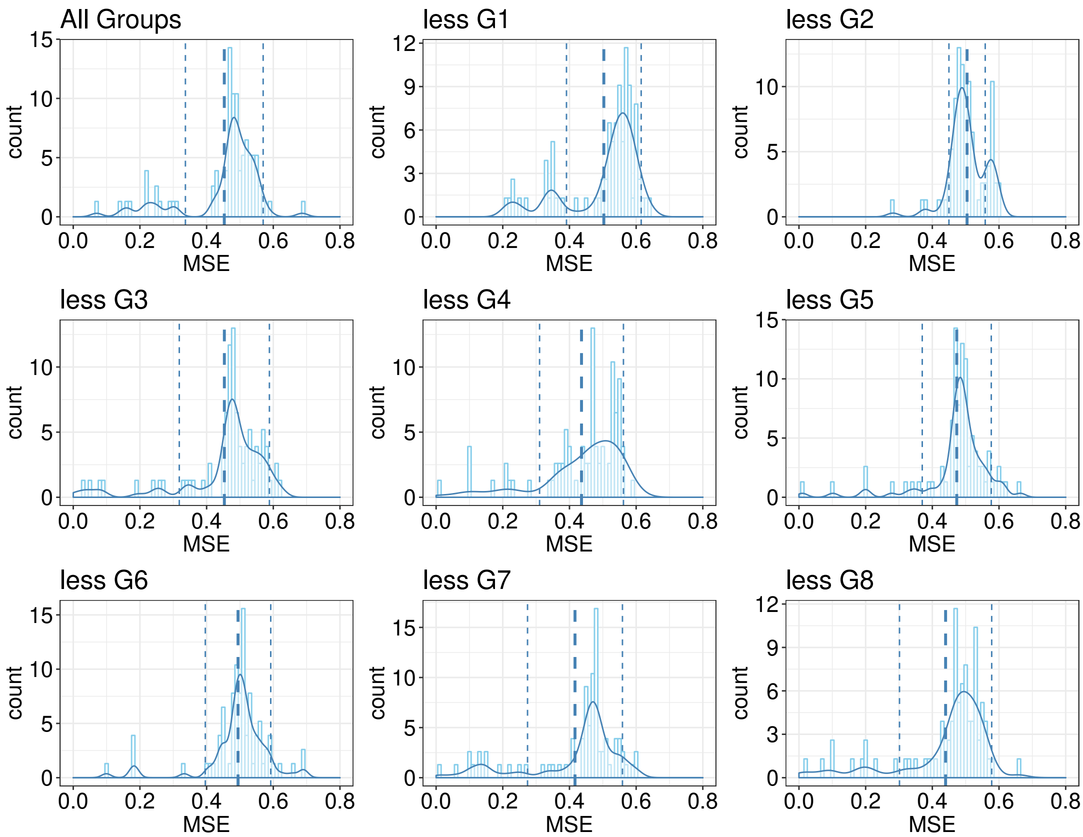


Notes: This graph shows a heatmap of all variables that have been used at least once by the MLP within the respective quarter. The blue squares indicate usage while the while squares indicate non-usage. NBER recessions are highlighted by gray shading.

Figure 4 plots the distribution of MSEs, which are created during the cross validation when the MLP is trained. The histogram and kernel density estimates of the MSEs are plotted for the baseline MLP, where all groups are included and for all MLPs where one of the above described groups are excluded.

When labor market variables (G2) are excluded, one can see that the distribution of MSEs spreads out much less when compared to the baseline and that there are less outliers towards the left tail of the distribution. Since these are MSEs, which are generated during the training of the MLP, very small values indicate overfitting on the respective fold. A larger number of folds where the MLP overfits leads to poorer out-of-sample performance on unseen data. It seems to be the case that the backward looking labour market variables produce the higher degree of overfitting during the training. This may happen because the additional benefit and associated information gain of an inclusion is relatively small given that these variables simply aggregate all previously known movements of production indicators with a time lag. The applied filter approach for input variable selection described in Section 2.2 comes to its limits in this situation because it selects mechanically based on significance within ARDL models. A similar reduction in MSEs can be seen when the fast moving financial variables (G6) are excluded. Again, there are fewer outliers towards the left tail of the distribution. There also seems to be less overfitting during the training when this group is excluded. In this example, a possible explanation for increased nowcast performance might be a higher signal to noise ratio, resulting from the exclusion of noisy variables like exchange- and interest rates.

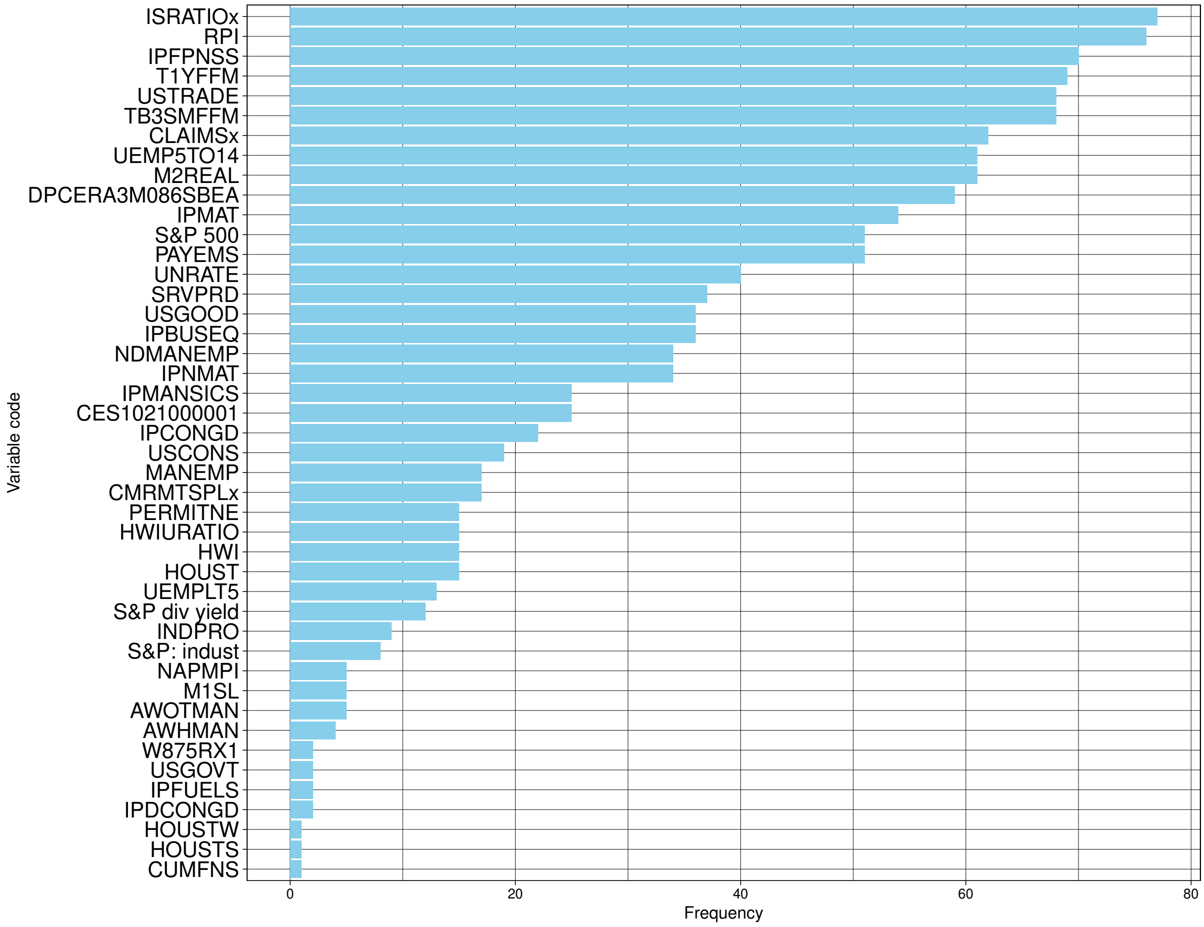
Figure 4: Histogram and densities of MSE.



Notes: This graph shows histograms and kernel density estimates of the MSEs, which result from the cross validation during the training of the MLP. The dashed vertical line represents the mean, while the two dotted lines represent the standard deviation around the mean.

Figure 5 demonstrates the frequency of variables used within the MLP during the time period the nowcasting exercise is conducted. To avoid being too descriptive at this point, the focus will be on the 10 most frequent variables applied. Together with Figure 3, an attempt can be made to open up the black box of the MLP approach. The 10 most frequently used variables from top to bottom are: inventory sales (ISRATIOx), real personal income (RPI), industrial production of final products and non-industrial supplies (IPFPNSS), 1-year treasury minus fed funds rate (T1YFFM), all employees in retail trade (USTRADE), three months treasury rate minus fed funds rate (TB3SMFFM), initial claims (CLAIMSx), civilian unemployment rate with a duration of 5 to 14 weeks (UNEMP5TO14), real M2 money stock (M2REAL) and real personal consumption expenditures (DPCERA3M086SBEA). They are used between 77 and 58 times during the nowcasting and forecasting exercise. Hence, out of the 10 most frequently selected variables, the MLP picks mostly real activity variables (4/10), some financial variables (3/10) and some labor market variables (3/10). Half of the real variables consist of production variables (inventory to sales and industrial production of non-industrial supplies) and the other half of income variables (real personal income and real personal consumption expenditures). The financial indicators mostly consist of variables describing the yield curve (1-year treasury minus fed funds rate and 3-month treasury minus fed funds rate). Because the yield curve is one of the most used variables in the prediction of business cycle turning points, it positively confirms that the artificial neural network also selects these variables most frequently. Interestingly, labor market variables are selected frequently because they are supposed to be lagging behind real and financial movements. However, the most depicted labor market variable is initial claims, which is an indicator of the labor market as a whole and therefore serves as a forecast of unemployment itself. Which variables are used shortly before and during crisis periods is especially interesting. For example, in Figure 3 one can see that in the years before the great financial crisis of 2007 to 2008, the housing variables are not used at all. This may reflect the ability of the MLP to recognise when a variable is moving apart, for example through an inflating bubble process as it was clearly the case for the housing market. Yield curves seem to be indicative before, after and during recessions as they are used by the MLP around the Dotcom and housing bubbles of 2001 and 2007. The same is true for the most frequently used production, income and consumption variables that were discussed earlier. Labor market variables make up the largest block of applied variables. However, finding a consistent pattern among the labor market indicators, especially with regard of when and which of those variables are used, is difficult and shows the limits of the presented attempt to opening up the black box of an artificial neural network in economic forecasting and nowcasting.

Figure 5: Bar chart demonstrating the frequency of variables used by the MLP.



5. Conclusion

This paper applies a machine learning framework to economic now- and forecasting of US GDP growth. Artificial neural networks are applied to the monthly vintages of the FRED-MD database. These monthly indicators are used within the flexible MLP framework of Crone and Kourentzes (2010), Kourentzes et al. (2014) and Ord et al. (2017), which is able to automatically select the timely variables that are most informative for economic now- and forecasting of the quarterly GDP growth. The variables are selected by applying a linear ARDL filter approach of Kourentzes et al. (2014), where the most significant variables and corresponding lags are selected previously to the network training. The network architecture is selected by the wrapper approach of Kourentzes et al. (2014), which applies a cross-validation scheme to select the number of hidden layers and associated nodes. The MLP is trained with data from the 1950s up until the late-1990s. Afterwards, the first now- and forecasts are conducted, while the MLP is continuously retrained within an expanding window scheme. This approach enables us to get a highly flexible regression framework, which is able to fit to potentially any non-linearity that might occur throughout time. While the described functional flexibility might serve as an advantage over traditional linear frameworks, the resulting black box is a clear disadvantage.

The applied MLP framework is used in a now- and forecasting competition against the DFM approach of Giannone et al. (2008) and the SPF. These two approaches have been chosen because the DFM approach is the most commonly applied framework for nowcasting applied by central banks, while the SPF is one of the most hard-to-beat competitors in macroeconomic forecasting. All three frameworks are tested against a naive constant growth model framework and outperform it in terms of the RMSFE and the MAFE. When the models are tested against each other in terms of now- and forecasting accuracy, it is found that the MLP significantly outperforms the DFM on horizons of $h = 0, 1, ..., 4$ in terms of RMSFE and MAFE. Against the SPF, the MLP generates smaller RMSFEs and MAFEs on a horizon of $h = 0$. However, this result is not significant. The results are robust to the omission of subgroups of variables. To summarise the results, one can conclude that the flexible MLP framework generates significantly better results compared to the DFM approach, while the results are at least as good as the SPF nowcasts.

A potential field of future research could be the application of a recurrent neural network (RNN) to nowcasting. When MLPs are a nonlinear generalisation of ARDL models, RNNs can be seen as non-linear generalisations of autoregressive moving average distributed lag (ARMADL) models. However, a major disadvantage is that, to the best of our knowledge, RNNs do not have efficient filters and wrappers, making an automation of the network structure difficult.

References

Bańbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2013). Now-casting and the real-time data flow. In *Handbook of economic forecasting*, volume 2, pages 195–237. Elsevier.

Crone, S. F. and Kourentzes, N. (2010). Feature selection for time series prediction–a combined filter and wrapper approach for neural networks. *Neurocomputing*, 73(10-12):1923–1936.

Croushore, D. D. (1993). Introducing: the Survey of Professional Forecasters. *Business Review-Federal Reserve Bank of Philadelphia*, 6:3–15.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3).

Doz, C., Giannone, D., and Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164(1):188–205.

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676.

Heravi, S., Osborn, D. R., and Birchenhall, C. (2004). Linear versus neural network forecasts for european industrial production series. *International Journal of Forecasting*, 20(3):435–446.

Jung, J.-K., Patnam, M., and Ter-Martirosyan, A. (2018). An algorithmic crystal ball: Forecasts-based on machine learning. IMF Working Paper 230.

Kourentzes, N., Barrow, D. K., and Crone, S. F. (2014). Neural network ensemble operators for time series forecasting. *Expert Systems with Applications*, 41(9):4235–4244.

Kuan, C.-M. and Liu, T. (1995). Forecasting exchange rates using feedforward and recurrent neural networks. *Journal of applied econometrics*, 10(4):347–364.

Lachtermacher, G. and Fuller, J. D. (1995). Back propagation in time-series forecasting. *Journal of Forecasting*, 14(4):381–393.

McCracken, M. W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

Ord, K., Fildes, R., and Kourentzes, N. (2017). *Principles of Business Forecasting–2nd Ed.* wessex, Incorporated.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088):533.

Sims, C. A., Stock, J. H., and Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica: Journal of the Econometric Society*, pages 113–144.

Tkacz, G. (2001). Neural network forecasting of Canadian GDP growth. *International Journal of Forecasting*, 17(1):57–69.

Torres, D. G. and Qiu, H. (2018). Applying recurrent neural networks for multivariate time series forecasting of volatile financial data.

Wallis, K. F. (1986). Forecasting with an econometric model: The "ragged edge" problem. *Journal of Forecasting*, 5(1):1–13.

A. Appendix

A.1. FRED-MD database

The TCODE column denotes the following data transformation for a series x: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1.0)$. The FRED column gives mnemonics in FRED followed by a short description.

Some series require adjustments to the raw data available in FRED. These variables are tagged by an asterisk to indicate that they have been adjusted and thus differ from the series from the source. For a detailed summary of the adjustments see McCracken and Ng (2016).

Group 1. Output and income

	ID	*tcode*	*FRED*	*Description*
1	1	5	RPI	Real Personal Income
2	2	5	W875RX1	Real personal income ex transfer receipts
3	6	5	INDPRO	IP Index
4	7	5	IPFPNSS	IP: Financial Products and Nonindustrial Supplies
5	8	5	IPFINAL	IP: Final Products (Market Group)
6	9	5	IPCONGD	IP: Consumer Goods
7	10	5	IPDCONGD	IP: Durable Consumer Goods
8	11	5	IPNCONGD	IP: Nondurable Consumer Goods
9	12	5	IPBUSEQ	IP: Business Equipment
10	13	5	IPMAT	IP: Materials
11	14	5	IPDMAT	IP: Durable Materials
12	15	5	IPNMAT	IP: Nondurable Materials
13	16	5	IPMANSICS	IP: Manufacturing (SIC)
14	17	5	IPB51222s	IP: Residential Utilities
15	18	5	IPFUELS	IP: Fuels
16	19	1	NAPMPI	ISM Manufacturing: Production Index
17	20	2	CUMFNS	Capacity Utilization: Manufacturing

Group 2: Labor market

	ID	*tcode*	*FRED*	*Description*
1	21*	2	HWI	Help-Wanted Index for United States
2	22*	2	HWIURATIO	Ratio of Help Wanted/No. Unemployed
3	23	5	CLF160OV	Civilian Labor Force
4	24	5	CE160V	Civilian Employment
5	25	2	UNRATE	Civilian Unemployment Rate
6	26	2	UEMPMEAN	Average Duration of Unemployment (Weeks)
7	27	5	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks
8	28	5	UEMP5TO14	Civilians Unemployed for 5-14 Weeks
9	29	5	UEMP15OV	Civilians Unemployed - 15 Weeks and Over
10	30	5	UEMP15T26	Civilians Unemployed for 15-26 Weeks
11	31	5	UEMP27OV	Civilians Unemployed for 27 Weeks and Over
12	32*	5	CLAIMSx	Initial Claims
13	33	5	PAYEMS	All Employees: Total nonfarm
14	34	5	USGOOD	All Employees: Goods-Producing Industries
15	35	5	CES1021000001	All Employees: Mining and Logging: Industries
16	36	5	USCONS	All Employees: Construction
17	37	5	MANEMP	All Employees: Manufacturing
18	38	5	DMANEMP	All Employees: Durable Goods
19	39	5	NDMANEMP	All Employees: Nondurable Goods
20	40	5	SRVPRD	All Employees: Service-Providing Industries
21	41	5	USTPU	All Employees: Trade, Transportation and Utilities
22	42	5	USWTRADE	All Employees: Wholesale Trade
23	43	5	USTRADE	All Employees: Retail Trade
24	44	5	USFIRE	All Employees: Financial Activities
25	45	5	USGOVT	All Employees: Government
26	46	1	CES0600000007	Avg Weekly Hours: Goods-Producing
27	47	2	AWOTMAN	Avg Weekly Overtime Hours: Manufacturing
28	48	1	AWHMAN	Avg Weekly Hours: Manufacturing
29	49	1	NAPMEI	ISM Manufacturing: Employment Index
30	127	6	CES0600000008	Avg Hourly Earnings: Goods-Producing
31	128	6	CES2000000008	Avg Hourly Earnings: Construction
32	129	6	CES3000000008	Avg Hourly Earnings: Manufacturing

Group 3: Housing

	ID	tcode	FRED	Description
1	50	4	HOUST	Housing Starts: Total New Privately Owned
2	51	4	HOUSTNE	Housing Starts: Northeast
3	52	4	HOUSTMW	Housing Starts: Midwest
4	53	4	HOUSTS	Housing Starts: South
5	54	4	HOUSTW	Housing Starts: West
6	55	4	PERMIT	New Private Housing Permits (SAAR)
7	56	4	PERMITNE	New Private Housing Permits: Northeast (SAAR)
8	57	4	PERMITMW	New Private Housing Permits: Midwest (SAAR)
9	58	4	PERMITS	New Private Housing Permits: South (SAAR)
10	59	4	PERMITW	New Private Housing Permits: West (SAAR)

Group 4: Consumption, orders and inventories

	ID	tcode	FRED	Description
1	3	5	DPCERA3M086SBEA	Real personal consumption expenditures
2	4*	5	CMRMTSPLx	Real Manu. and Trade Industries Sales
3	5*	5	RETAILx	Retail and Food Services Sales
4	60	1	NAPM	ISM: PMI Composite Index
5	61	1	NAPMNOI	ISM: New Orders Index
6	62	1	NAPMSDI	ISM: Supplier Deliveries Index
7	63	1	NAPMII	ISM: Inventories Index
8	64	5	ACOGNO	New Orders for Consumer Goods
9	65*	5	AMDMNOx	New Orders for Durable Goods
10	66*	5	ANDENOx	New Orders for Nondefense Capital Goods
11	67*	5	AMDMUOx	Unfilled Orders for Durable Goods
12	68*	5	BUSINVx	Total Business Inventories
13	69*	2	ISRATIOx	Total Business: Inventories to Sales Ratio
14	130*	2	UMSCENTx	Consumer Sentiment Index

Group 5: Money and credit

	ID	*tcode*	*FRED*	*Description*
1	70	6	M1SL	M1 Money Stock
2	71	6	M2SL	M2 Money Stock
3	72	5	M2REAL	Real M2 Money Stock
4	73	6	AMBSL	St. Louis Adjusted Monetary Base
5	74	6	TOTRESNS	Total Reserves of Depository Institutions
6	75	7	NONBORRES	Reserves of Depository Institutions
7	76	6	BUSLOANS	Commercial and Industrial Loans
8	77	6	REALLN	Real Estate Loans at All Commercial Banks
9	78	6	NONREVSL	Total Nonrevolving Credit
10	79*	2	CONSPI	Nonrevolving consumer credit to Personal Income
11	131	6	MZMSL	MZM Money Stock
12	132	6	DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding
13	133	6	DTCTHFNM	Total Consumer Loans and Leases Outstanding
14	134	6	INVEST	Securities in Bank Credit at All Commercial Banks

Group 6: Interest and exchange rates

	ID	*tcode*	*FRED*	*Description*
1	84	2	FEDFUNDS	Effective Federal Funds Rate
2	85*	2	CP3Mx	3-Month AA Financial Commercial Paper Rate
3	86	2	TB3MS	3-Month Treasury Bill
4	87	2	TB6MS	6-Month Treasury Bill
5	88	2	GS1	1-Year Treasury Rate
6	89	2	GS5	5-Year Treasury Rate
7	90	2	GS10	10-Year Treasury Rate
8	91	2	AAA	Moody's Seasoned Aaa Corporate Bond Yield
9	92	2	BAA	Moody's Seasoned Baa Corporate Bond Yield
10	93*	1	COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS
11	94	1	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS
12	95	1	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS
13	96	1	T1YFFM	1-Year Treasury C Minus FEDFUNDS
14	97	1	T5YFFM	5-Year Treasury C Minus FEDFUNDS
15	98	1	T10YFFM	10-Year Treasury C Minus FEDFUNDS
16	99	1	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS
17	100	1	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS
18	101	5	TWEXMMTH	Trade Weighted U.S. Dollar Index: Major Currencies
19	102*	5	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate
20	103*	5	EXJPUSx	Japan / U.S. Foreign Exchange Rate
21	104*	5	EXUSUKx	U.S. / U.K. Foreign Exchange Rate
22	105*	5	EXCAUSx	Canada / U.S. Foreign Exchange Rate

Group 7: Prices

	ID	*tcode*	*FRED*	*Description*
1	106	6	WPSFD49207	PPI: Finished Goods
2	107	6	WPSFD49502	PPI: Finished Consumer Goods
3	108	6	WPSID61	PPI: Intermediate Materials
4	109	6	WPSID62	PPI: Crude Materials
5	110*	6	OILPRICEx	Crude Oil, spliced WTI and Cushing
6	111	6	PPICMM	PPI: Metals and metal products
7	112	1	NAPMPRI	ISM Manufacturing: Prices Index
8	113	6	CPIAUCSL	CPI: All Items
9	114	6	CPIAPPSL	CPI: Apparel
10	115	6	CPITRNSL	CPI: Transportation
11	116	6	CPIMEDSL	CPI: Medical Care
12	117	6	CUSR0000SAC	CPI: Commodities
13	118	6	CUSR0000SAD	CPI: Durables
14	119	6	CUSR0000SAS	CPI: Service
15	120	6	CPIULFSL	CPI: All Items less Food
16	121	6	CUSR0000SA0L2	CPI: All Items less Shelter
17	122	6	CUSR0000SA0L5	CPI: All Items less Medical Care
18	123	6	PCEPI	Personal Cons. Expend.: Chain Index
19	124	6	DDURRG3M086SBEA	Personal Cons. Expend.: Durable Goods
20	125	6	DNDGRG3M086SBEA	Personal Cons. Expend.: Nondurable Goods
21	126	6	DSERRG3M086SBEA	Personal Cons. Expend.: Services

Group 8: Stock market

	ID	*tcode*	*FRED*	*Description*
1	80*	5	S&P 500	S&P's Common Stock Price Index: Composite
2	81*	5	S&P: indust	S&P's Common Stock Price Index: Industrials
3	82*	2	S&P div yield	S&P's Composite Common Stock: Dividend Yield
4	83*	5	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio