Predictor Pre-Selection for Mixed-Frequency Dynamic Factor Models: A Simulation Study with an Empirical Application to GDP Nowcasting

Domenic Franjic* Karsten Schweikert†

University of Hohenheim University of Hohenheim

[Latest update: September 23, 2024]

Abstract

We investigate the performance of dynamic factor model nowcasting with pre-selected predictors in a mixed-frequency setting. The predictors are selected via the elastic net as it is common in the targeted predictor literature. A simulation study and an application to empirical data are used to evaluate different strategies for variable selection, the influence of tuning parameters, and to determine the optimal way to handle mixed-frequency data. We propose a novel cross-validation approach that connects the pre-selection and nowcasting step. In general, we find that pre-selecting provides more accurate nowcasts compared with the benchmark dynamic factor model using all variables. Our newly proposed cross-validation method outperforms the other specifications in most cases.

Keywords: Elastic net, High-dimensional, Soft-thresholding, Targeted predictors, Variable selection
JEL Classification: C32, C53, E37

*Corresponding author; Address: University of Hohenheim, Core Facility Hohenheim & Institute of Economics, Schloss Hohenheim 1 C, 70593 Stuttgart, Germany, e-mail: *franjic@uni-hohenheim.de*

†Address: University of Hohenheim, Core Facility Hohenheim & Institute of Economics, Schloss Hohenheim 1 C, 70593 Stuttgart, Germany

1 Introduction

The mixed-frequency dynamic factor model (DFM) has become a workhorse model for macroeconomic forecasting and nowcasting (Giannone et al., 2008; Bok et al., 2018). Through a combination of factor analysis and Kalman smoothing (Stock and Watson, 2002a,b; Doz et al., 2011), the model can handle big data sets[1], constructed from mixed-frequency predictors, while also exploiting the often shorter publication lags of the predictor variables. This is especially useful when forecasting a not yet released lower frequency variable, often current quarter GDP, with the help of partially available quarterly and monthly economic indicators (see, for example, Giannone et al., 2008; Bańbura et al., 2013).

By construction, the framework invites for the use of increasingly larger data sets to include predictors from all sectors of the economy. However, concerns have been raised that the resulting factors are less useful for forecasting particularly when the idiosyncratic errors are cross-correlated (Boivin and Ng, 2006). Therefore, Bai and Ng (2008) propose to employ a set of *targeted predictors* (TP) for the factor analysis. More specifically, predictors are pre-selected using the *elastic net* (EN) before the estimation of a factor model and the construction of a forecast.[2] The concept is extended to a mixed-frequency nowcasting framework by Bessec (2013) and Siliverstovs (2017).

The literature provides ambiguous results on whether the TP approach can substantially improve the forecasting accuracy (Eickmeier and Ng, 2011; Castle et al., 2013; Kim and Swanson, 2014; Bulligan et al., 2015; Kim and Swanson, 2018). Particularly, it is unknown how well the TP approach performs in mixed-frequency data structures that are common in empirical applications (Bańbura et al., 2013). In previous studies, the pre-selection step of the TP approach is mostly implemented as originally proposed by Bai and Ng (2008),

[1]The theory for quasi maximum likelihood estimation has been developed by Doz et al. (2012).

[2]In their original study, Bai and Ng (2008) introduce two different thresholding rules. Under the hard thresholding rule, a predictor is used in the final forecasting step if, independent of the other predictors, a linear regression of the target variable on said predictor results in an absolute t-statistic that exceeds a certain threshold. In the case of the soft thresholding method, the elastic net or LASSO is applied for predictor selection setting the target variable as the dependent variable. They find larger and more systematic improvements with soft thresholding.

where the number of variables aimed to be extracted is set to 30 (see, e.g., Bai and Ng, 2008; Schumacher, 2010; Eickmeier and Ng, 2011; Kopoin et al., 2013). This is in contrast to the more common approach in the literature on penalized regressions, where the tuning parameters are cross-validated using a part of the sample (Zou and Hastie, 2005; Zou et al., 2007; Hastie et al., 2015).

Consequently, the following issues in terms of model specification arise: it is unclear which algorithm should be used to solve the EN problem. Available choices are, for example, coordinate descent or LARS-EN. Although the minimiser of the EN objective function is theoretically unique, the corresponding parameter vector, and hence the placement and number of zeros, is generally not unique. Depending on this choice, it must be decided whether the tuning parameters are determined based on (i) cross-validation (CV)[3] or (ii) the maximum number of predictor variables to be selected from the full panel of predictor variables. Then, there is the additional question of how to handle mixed-frequency data before feeding it into the EN.

The current literature does not provide a conclusive statement on the performance of TP nowcasting. Only few studies report evidence for TP models to substantially improve over benchmark models when applied to different real world data sets (see, e.g., Schumacher, 2010; Kopoin et al., 2013; Girardi et al., 2017). Instead, most studies report mixed results (Eickmeier and Ng, 2011; Castle et al., 2013; Bulligan et al., 2015; Kim and Swanson, 2014, 2018). A reason for this could lie in sub-optimal specification of the pre-selection step. The current literature lacks studies that evaluate the influence of important tuning parameters and guidelines on how these parameters should be validated. Most of the research is conducted by comparing the TP performance against other models in selected empirical data sets without analyzing in detail how the tuning of the pre-selection step influences the nowcasting performance. When turning to the issue of handling mixed-frequency data in the TP framework, the literature provides even less guidance for applied researchers.

[3]In an earlier version of this paper, we also considered information criteria. However, we found that optimizing information criteria is generally dominated by the CV approach and results are thus omitted.

To the knowledge of the authors, there is no large scale simulation study that investigates the above-mentioned issues exhaustively. The paper at hand thus aims to contribute to the literature in the following way: (i) it is investigated whether any of the TP specifications can improve the nowcast accuracy of different benchmark models in a controlled mixed-frequency setting, (ii) we study the effects that the choice of the tuning parameters in the pre-selection step has on the overall model performance. Particularly, we compare the performance of data driven specifications against selecting a fixed number of predictors. (iii) We propose a new TP cross-validation strategy based on prior nowcast errors that seems to perform well compared with the existing approaches. Additionally, we assess the predictive accuracy of those methods for empirical data sets that differ in the number of predictors and spatial resolution.[4]

Our findings from the simulation study and the empirical application allow for multiple conclusions. Most notably, we find that for each DGP and real data set investigated, at least one TP specification can improve the nowcast accuracy compared with the benchmark models. Selecting a fixed number of predictors performs reasonably well, particularly if the data set contains many irrelevant variables. However, we can show that the TP models with a data driven pre-selection step frequently outperform the models with a fixed number of selected predictors. The data driven (or cross-validated) models work well in misspecified or noisy systems and achieve the best results in the empirical application. The proposed TP specification which links the pre-selecting and nowcasting steps by cross-validating based on the nowcast error yields the most consistent results, outperforming the benchmark DFM in controlled settings and in our empirical application.

The remainder of this paper is organized as follows. Section 2 discusses the necessary basics for DFM nowcasting, followed by an outline of the model specifications and their

[4]We focus on the EN as a variable selection tool in this study as originally proposed by Bai and Ng (2008). This is due to the fact that the EN is most commonly used in practice and it is also used by Bessec (2013) and Siliverstovs (2017) in the context of mixed-frequency nowcasting. Moreover, extensive comparisons between different variable selection tools in the context of factor model nowcasting are explored in the literature (see e.g. Kim and Swanson, 2014, 2018).

implementation. Section 3 describes the set-up and reports the results from our simulation study. In Section 4, the empirical data sets are introduced and the relative performance of the estimators are evaluated in a real world context. Section 5 summarizes the findings of this study.

2 Methodology

2.1 Dynamic Factor Model

In the following, we briefly describe the mixed-frequency dynamic factor models estimated with the two-step procedure proposed by Giannone et al. (2008) and Doz et al. (2011). Initially, we consider a vector of stationary monthly variables $\mathbf{z}_t = \{z_{h,t}\}_{h=1}^{H}$ and a DFM characterized by the following equations:

$$\begin{aligned} \mathbf{z}_t = \mathbf{\Lambda}\mathbf{f}_t + \boldsymbol{\xi}_t, \qquad & \mathbf{\Phi}^P(\mathbb{L})\mathbf{f}_t = \mathbf{\Upsilon}\boldsymbol{\epsilon}_t \\ \mathbb{E}(\boldsymbol{\xi}_t\boldsymbol{\xi}_t') = \mathbf{\Sigma}_{\boldsymbol{\xi}} = \operatorname{diag}(\sigma^2_{1,\boldsymbol{\xi}}, \ldots, \sigma^2_{H,\boldsymbol{\xi}}) & \\ \mathbb{E}(\boldsymbol{\xi}_t\boldsymbol{\xi}_{t-s}') = 0,\ \forall\, t, s > 0, \qquad & \mathbb{E}(\boldsymbol{\xi}_t\,(\mathbf{\Upsilon}\boldsymbol{\epsilon}_{t-w})') = 0,\ \forall\, w, t. \end{aligned} \tag{1}$$

Further, if allowing for missing data at the right end of the time index due to different publication lags, so-called ragged edges, it is assumed that

$$\sigma^2_{h,\boldsymbol{\xi}} = \begin{cases} \sigma^2_h > 0, & \text{if } z_{h,t} \text{ is available} \\ \infty, & \text{otherwise} \end{cases}. \tag{2}$$

Here, $\mathbf{f}_t' = \{f_{r,t}\}_{r=1}^{R}$ is the vector of static factors at time t, and $\mathbf{\Lambda} = \{\lambda_{h,r}\}_{h=1,r=1}^{H,R}$ is the factor loadings matrix. $\boldsymbol{\xi}_t = \{\xi_{h,t}\}_{h=1}^{H}$ is the idiosyncratic error vector at time t with diagonal covariance matrix $\mathbf{\Sigma}_{\boldsymbol{\xi}}$. The matrix $\mathbf{\Upsilon} = \{\upsilon_{r,q}\}_{r=1,q=1}^{R,Q}$ has full rank Q, and the vector of primitive shocks $\boldsymbol{\epsilon}_t = \{\epsilon_{q,t}\}_{q=1}^{Q}$ is assumed to be a white noise process with zero mean and

covariance matrix $\mathbf{I}_Q$. Furthermore, the roots of the lag polynomial $\mathbf{\Phi}^P(\mathbb{L})$, where $\mathbf{\Phi}_p = \{\phi_{r,s,p}\}_{r=1,s=1}^{R,R}$, for all $p = 1 \ldots, P$, are assumed to lie outside the unit circle.

The mixed-frequency structure of the DFM is implemented according to Mariano and Murasawa (2003). We assume that the target variable, e.g., log GDP, is integrated of order one. Further, we assume that quarterly observations are recorded in the last month of the quarter. $x_{t,q}$ denotes a quarterly variable, and $x_{t-t_m,m}$, for $t_m = 0, 1/3, 2/3$, is the corresponding latent monthly variable at the last, second, and first month of quarter t. Analogously to our empirical application, we use log GDP for our illustration. Since log GDP is a flow variable, the quarterly observations $\ln \text{gdp}_{t,q}$ are related to unobserved monthly observations $\ln \text{gdp}_{t,m}$ via

$$\ln \text{gdp}_{t,q} = \frac{1}{3}\left(\ln \text{gdp}_{t,m} + \ln \text{gdp}_{t-1/3,m} + \ln \text{gdp}_{t-2/3,m}\right), \tag{3}$$

and it holds for the quarterly growth rates, $\Delta_3 \ln \text{gdp}_{t,q} = \ln \text{gdp}_{t,q} - \ln \text{gdp}_{t-1,q}$, that

$$\begin{aligned} \Delta_3 \ln \text{gdp}_{t,q} &= \frac{1}{3}\Delta \ln \text{gdp}_{t,m} + \frac{2}{3}\Delta \ln \text{gdp}_{t-1/3,m} + \Delta \ln \text{gdp}_{t-2/3,m} \\ &+ \frac{2}{3}\Delta \ln \text{gdp}_{t-1,m} + \frac{1}{3}\Delta \ln \text{gdp}_{t-4/3,m}. \end{aligned} \tag{4}$$

Equation (4) expresses the quarterly growth rates as a function of the monthly growth rates and thereby provides us with the aggregation weights for the mixed-frequency DFM.[5]

As part of the model selection process, we determine the number of common factors by optimizing the information criterion provided in Bai and Ng (2002). To ensure some form of dimensionality reduction, the maximum value for $\hat{R}$ to be considered is $\min(\lfloor 0.5N_m^* \rfloor, 15)$ for the simulated data and $\min(\lfloor 0.5N_m^* \rfloor, 5)$ for the real world data, where N_m^* is the number of monthly targeted predictors and $\lfloor \cdot \rfloor$ is the floor function. The number of primitive shocks

[5]Quarterly GDP is often defined as the arithmetic mean of the (unobserved) monthly levels. However, the chosen representation as a geometric mean allows us to work with a linear state-space model whereas the usual link over the arithmetic mean requires a nonlinear state-space model (Proietti and Moauro, 2006). We follow the recent literature (see, e.g., Giannone et al., 2008; Bańbura and Rünstler, 2011; Andreini et al., 2021) and use the outlined approach for a linear state-space model in this paper.

Q is chosen to be estimated using the procedure in Bai and Ng (2007). For the order of the VAR processes of the common factor, we follow the convention used in Marcolino de Mattos et al. (2019), where a value for $\hat{P}$ is retrieved by computing the AIC, BIC, HQC, and final prediction error and choosing the most parsimonious model indicated by these measures.

2.2 Preparing mixed-frequency data for the pre-selection step

Since the EN is not able to handle mixed-frequency data sets automatically, it requires a preliminary step creating a balanced set of variables. The TP literature provides two ideas in this regard. The first approach proposed by Siliverstovs (2017) is a *skip-sampling* or blocking approach often used in MIDAS models. In this case, the monthly observations of a predictor variable are transformed into three quarterly variables, thereby aligning the frequencies of the quarterly variable of interest and the monthly predictors. For example, unemployment rates as a monthly predictor variable is transformed in such a way, tracking the unemployment rate at the first, second, and third month of a quarter.

Skip-sampling the monthly variables has some obvious drawbacks. The complete data matrix $\mathbf{X}_s$ is of dimension $((N_q + 3N_m) \times \tau)$, where $\tau < T$ is the point in time for which all predictors are observed. Given the number of monthly observations, N_m, skip-sampling thus increases the dimensionality of the problem considerably, which can lead to high computational costs. Moreover, since the variables are highly correlated, increasing the overall number of variables also increases the likelihood of including irrelevant variables.

Alternatively, Bessec (2013) aggregates the monthly variables into pseudo quarterly indicators. The *aggregation* approach has the advantage that the dimensionality of the problem is not further increased at the cost of losing some information from the higher frequency predictors. This study differs from Bessec (2013) in that the monthly variables are not averaged over a quarter. Instead, we follow the ideas of Mariano and Murasawa (2003) and aggregate the monthly variables using the geometric mean of the three monthly variables of

each quarter.[6] Hence, the aggregation scheme in this step is conceptually similar to the one used in the mixed-frequency DFM.

2.3 Pre-selection with the Elastic Net

The EN was first introduced in Zou and Hastie (2005) and represents a combination of the Ridge and LASSO regularization. To solve the EN problem, we consider two popular algorithms, namely LARS-EN (Zou and Hastie, 2005) and coordinate descent (Friedman et al., 2010), and put special emphasis on the tuning parameters for these algorithms.[7]

Consider the balanced vintage $\mathcal{V}_\tau$. For the variable selection step, let the target variable be the quarterly time series with index one and define $\mathbf{y} := \mathbf{x}_{1,q} = \{x_{1,t,q}\}_{t=1}^{\tau} \in \mathbf{X}_{\tau,q}$. Let $\mathbf{X} = \{x_{n,t}\}_{n=1,t=1}^{M,\tau}$, a same-frequency predictor matrix, be retrieved from the predictors in $\mathcal{V}_\tau$ by one of the methods described above, where M is either equal to $N-1$ or $N_q + 3N_m - 1$ depending on whether aggregation or skip-sampling is used. The EN problem is defined as

$$\min \left(\|\mathbf{y} - \mathbf{X}'\boldsymbol{\beta}\|_2 + l_1\|\boldsymbol{\beta}\|_1 + l_2\|\boldsymbol{\beta}\|_2 \right), \tag{5}$$

where $\boldsymbol{\beta} = \{\beta_n\}_{n=1}^{M}$ is the coefficient vector. The parameters l_1 and l_2 are, respectively, the LASSO (Tibshirani, 1996) and Ridge (Hoerl and Kennard, 1988) tuning parameter.

It is also common to define $\alpha = \frac{l_1}{(l_1+l_2)}$ and recast (5) into

$$\min \left(\|\mathbf{y} - \mathbf{X}'\boldsymbol{\beta}\|_2 + l \left(\alpha\|\boldsymbol{\beta}\|_1 + \frac{(1-\alpha)}{2}\|\boldsymbol{\beta}\|_2 \right) \right). \tag{6}$$

Here, $\alpha \in [0, 1]$ is a mixing, or weight, parameter between the two size constraints.

The first algorithm is referred to as the LARS-EN algorithm (Efron et al., 2004; Zou and Hastie, 2005) and is most commonly used in the TP literature (see, e.g., Bai and Ng, 2008;

[6] Averaging yields qualitatively similar results. Detailed results are available upon request.

[7] For the implementation of the coordinate descent algorithm, the `glmnet` function of the correspondent package by Friedman et al. (2010) is used. The LARS-EN algorithm is implemented in `C++` and integrated into `R` using the `Rcpp` package by Eddelbuettel and Francois (2011).

Bessec, 2013; Siliverstovs, 2017). Despite its popularity in the TP literature, the procedure has a potential drawback when used in a purely data driven framework. Eventually, at some point along the solution path, each variable moves into the active set. For larger data sets, this requires more steps to calculate the full solution path and the solving the problem becomes computationally costly. This is of special concern with respect to skip-sampling, where the number of variables in the model matrix is drastically increased.

An alternative algorithm, the so-called coordinate descent, has been proposed by Friedman et al. (2010). The procedure has the advantage that it is generally faster than the LARS-EN approach when applied to large cross sections. It might thus be possible to use coordinate descent in the TP framework to achieve similar forecasting error improvements with less computational costs compared to the LARS-EN algorithm.

The EN tuning parameters need to be determined by the researcher and several strategies are offered in the literature. For the coordinate descent algorithm, we follow the default settings of the `glmnet` package by Friedman et al. (2010). For the LARS-EN algorithm, the l_2 grid is chosen to be a sequence of 99 logarithmically spaced values between $\ln(1.001)$ and $\ln(10)$, where 0 is added to incorporate the pure LASSO fit. The maximum number of steps k is set to the number of variables of the given data set.[8]

For CV, an expanding window time series cross validation procedure according to Hyndman and Athanasopoulos (2018, Ch. 5.10) is implemented. Different to the classical CV approach applied to i.i.d. data, this specific approach preserves the time-series structure of the underlying observations.

For the TP specification with fixed parameters l_2 and k, we use the values reported by Bai and Ng (2008), i.e., $l_2 = 0.25$ and $k = 30$. This specification is often found in empirical studies.[9]

[8] In general, it might be necessary to include more steps if the goal is to get a complete ranking of all the variables in the data set since variables that have entered the set at any point might be forced out later down the algorithm path. In our case, using a larger number of candidate steps is found to be computationally costly but does not improve the nowcast accuracy.

[9] Note that the authors also consider different values of the tuning parameter l_2 and $k = 5$. However, they conclude that the $l_2 = 0.25$ and $k = 30$ parametrisation yields the most promising results.

2.4 Model specifications

We begin with a description of the first TP specification family that uses information contained in the data to tune the EN parameters and therefore guide the variable selection process. The main intuition behind these data driven parametrisations is that the TP models might be more effective if they were specifically trained to the underlying data instead of relying on a fixed number of targeted predictors. Furthermore, it must be assumed that in a real world scenario, the set of variables leading to the smallest nowcasting error might be time-varying. The data driven specifications thus provide a more flexible framework in this regard when compared to the fixed value approach.

To construct the nowcast at time T, we first build the balanced panel with variables in $\mathbf{X}_{\tau,q}$ and $\mathbf{X}_{\tau,m}$ by either aggregating or skip-sampling. As above, $\tau < T$ is the most recent date for which the panel $\mathcal{V}_\tau$ is balanced. Second, we use the variables contained in $\mathbf{X}_\tau$ and the target variable $\mathbf{y}_\tau$ to solve the EN problem either via LARS-EN or coordinate descent. In both cases, we validate the parameters of the EN via time-series CV using a one-step-ahead forecast. Then, $\mathcal{V}_T^* = \{\mathbf{X}_{q,T}^*, \mathbf{X}_{m,T}^*\}$ is the vintage of TPs constructed from variables for which the solution of the EN problem in the previous step has indicated non-zero coefficients. Now, we can use $\mathcal{V}_T^*$ to construct a nowcast of y_T using the mixed-frequency DFM. This is done using a repeated variable selection step with an expanding window prior to each nowcast.[10] Note that in total, there are four specifications under investigation for this TP specification family. For simplicity, we employ the following terminology. For each specification, the first part of its name is either an "A-" if the mixed-frequency data is aggregated or an "S-" if it is skip-sampled. The following part indicates whether the EN is solved via coordinate descend ("CD-") or the LARS-EN algorithm ("LE-"). This is followed by "TPN" for a target predictor nowcast.[11]

[10] We also investigated a variable selection procedure for which the variables were selected once at the beginning of the nowcasting exercise. Since the results were not significantly different, we refrain from presenting them here. However, the results are available upon request.

[11] For example, the model specification where the mixed-frequency data is aggregated and the EN is solved via coordinate descent is expressed as "A-CD-TPN".

In addition to the family of data driven models, the fixed value procedures are investigated. Restricting the model to always selecting a fixed number of predictors might lead to improvements in terms of nowcasting performance by way of a more stable variable selection step. The terminology above is also applied to the two fixed value specifications. The skip-sampling approach is denoted by "S-F-TPN" and aggregation is denoted by "A-F-TPN".

We also consider two new types of specifications that we denote by *nowcast validated TP specifications*. In these data driven specifications, the final validation of the set of targeted predictors is not achieved by the EN based on an in-sample fit but rather it is based on the nowcasting error of the previous quarter. The main intuition is that the EN eventually leads to a set of targeted predictors retrieved via a linear projection of the target variable onto the predictor space. However, factor models are used in a second step to link the predictors to the variable of interest, and a set of predictors leading to the lowest nowcasting error may not generally coincide with the set of predictors identified by the EN. Thus, a natural extension to the models above is to use the EN as a first pass filter, given a grid of values for the mixing parameter, generating 100 candidate sets, and then evaluate these sets using the previous period's nowcasting error. Moreover, if the validation is conducted as a pseudo-nowcasting exercise, i.e. the publication lags are accounted for in the nowcast step, this approach allows for a variable selection step that implicitly accounts for the release structure of all variables. The nowcast validation procedure slightly differs from the previous two model families in the second step. Here, we use $\mathbf{X}_{\tau-1}$ and $\mathbf{y}_{\tau-1}$ to solve the EN problem via coordinate descent for a grid of α values $\boldsymbol{\alpha}$, where l is cross validated. Let $\mathcal{C} := \{\mathcal{V}_{\alpha\in\boldsymbol{\alpha},\tau}\}$ be the set of vintages, where each vintage corresponds to a value in $\boldsymbol{\alpha}$, and consists of the variables for which the solution of the EN problem reports non-zero coefficients given the corresponding α value.

We use $\mathcal{C}$ to construct a nowcast of y_τ for each unique vintage in the set.[12] When doing so, we impose the release structure of the underlying data observed at T onto each $\mathcal{V}_{\alpha\in\boldsymbol{\alpha},\tau}$ to explicitly account for the importance of the timeliness of the release of each predictor. The targeted predictors set is then found as the set resulting in the smallest RMSFE. The two versions of these models are distinguished by their handling of mixed-frequency data and are denoted by "A-N-TPN" and "S-N-TPN", respectively.

The results of five benchmark models are also reported. Specifically, we analyze the forecasting errors of an AR(1), ARMA, and naive constant growth model. Furthermore, the nowcast results of a DFM using all variables, referred to as crowded DFM or CDFM, is the most important benchmark for the TP specifications. In the simulation study, we are able to compute nowcasts from a DFM that only uses the relevant variables, the so-called oracle DFM (ODFM). This is done to compare the results of using all the variables with those of using only the ones stemming from the same process as the target variable (CDFM vs. ODFM).

3 Simulation

3.1 Setup

In our simulation study, four different DGPs are investigated. Inspired by the ideas of Boivin and Ng (2006), each DGP is composed of two different models, which are structurally identical, e.g. both series admit a factor model of similar dimensions with the same dynamic structure but are parameterized differently. The models are only connected by their idiosyncratic errors, which are all drawn from the same multivariate distribution with a non-diagonal

[12] The reason for using only a single nowcast error as a validation measure is twofold. First, in an early version of this study, we found that using multiple nowcasts does not improve the overall performance of the model, but instead increases computational costs and drastically decreases the number of observations available in the validation set. Second, this validation procedure is designed to be applied to quarterly data in a real scenario. Since using more validation errors requires using older data, it is to be expected that using the most recent information is more informative and extending the set of validation errors does not improve the performance.

covariance matrix, i.e., the idiosyncratic error are correlated across models. Additionally, each DGP is considered using either a relatively high or low *signal-to-noise ratio* (SNR) in the measurement equation. For simplicity, we denote the data stemming from the same process as the target variable by *relevant* (subscript "re") while the other variables are referred to as *irrelevant* (subscripts "ir"). The main reasons for using this procedure are the following: first, cross-correlation of the error terms is the main driver of why a pre-selection step was initially considered. Therefore, we also draw from cross-correlated errors in the same process. Second, when using large data sets, it is difficult to assume that all available variables are relevant, i.e., stem from a single model. Thus, we sample from two different models to provide a more realistic scenario. Lastly, considering different SNRs is important, since the effect of an increase in the SNR on the forecasting performance is ambiguous. For example, it could be that a lower SNR increases the need for variable pre-selection, since the effects of including irrelevant predictors are stronger than in the case of a high SNR. On the other hand, when the SNR is low, the EN might not be able to correctly identify the important predictors.[13][14]

The first and second DGPs are a combination of factor models that closely resemble the theoretical underlying DFM in the nowcast step. The first DGP (DGPI) is a combination of two five-factor models, where their underlying structure is taken from Bai and Ng (2007). The second DGP (DGPII) is chosen to be a more parsimonious single factor model with each factor following a white noise process. Let $\mathbf{x}_{t,\text{re}} = \{x_{n,t,\text{re}}\}_{n=1}^{N_{\text{re}}}$ and $\mathbf{x}_{t,\text{ir}} = \{x_{n,t,\text{ir}}\}_{n=1}^{N_{\text{ir}}}$ be the

[13]Drawing data sets of equal size from two structurally identical processes does not violate the general assumptions of the Giannone et al. (2008) estimator. This is due to the fact that such a DGP can be interpreted as a single FM with a block diagonal loadings matrix consisting of $\mathbf{\Lambda}_{\text{re}}$ in the upper left corner, $\mathbf{\Lambda}_{\text{ir}}$ in the lower right corner, and zero matrices along the anti-diagonal. In the limit, the Gram matrix of the loadings matrix tends to a diagonal matrix, which is a common identification assumption in the FM literature (see Smeekes and Wijler, 2018, for a similar argument regarding sparse FMs).

[14]We refrain from using loading matrices with overlapping loadings of the relevant and irrelevant process. Such a loading matrix would result in a DGP that could be interpreted as data derived from three processes that differ not only in their "relevance" but also in the number of factors.

sets of relevant and irrelevant variables at time t. The DGPs are formally defined as

$$\begin{aligned}
\mathbf{x}_{t,\text{re}} &= \mathbf{\Lambda}_{\text{re}}\mathbf{f}_{t,\text{re}} + \boldsymbol{\xi}_{t,\text{re}} \qquad & \mathbf{f}_{t,\text{re}} &= \mathbf{\Phi}_{\text{re}}\mathbf{f}_{t-1,\text{re}} + \boldsymbol{\nu}_{t,\text{re}} \\
\mathbf{x}_{t,\text{ir}} &= \mathbf{\Lambda}_{\text{ir}}\mathbf{f}_{t,\text{ir}} + \boldsymbol{\xi}_{t,\text{ir}} \qquad & \mathbf{f}_{t,\text{ir}} &= \mathbf{\Phi}_{\text{ir}}\mathbf{f}_{t-1,\text{ir}} + \boldsymbol{\nu}_{t,\text{ir}} \\
\boldsymbol{\nu}_{t,\text{re}} &= \mathbf{\Upsilon}_{\text{re}}\boldsymbol{\epsilon}_{t,\text{re}}, \text{ and } \boldsymbol{\nu}_{t,\text{ir}} = \mathbf{\Upsilon}_{\text{ir}}\boldsymbol{\epsilon}_{t,\text{ir}} \qquad & & \text{(DGPI)} \\
\boldsymbol{\xi}_t &:= (\boldsymbol{\xi}'_{t,\text{re}}, \boldsymbol{\xi}'_{t,\text{ir}})' \sim \mathcal{N}(\mathbf{0}_N, \mathbf{\Sigma}_{\boldsymbol{\xi}}), \text{ where } N = N_{\text{re}} + N_{\text{ir}} \\
& R = 5, \; Q = 3, \text{ and } P = 1.
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{x}_{t,\text{re}} &= \boldsymbol{\lambda}_{\text{re}} f_{t,\text{re}} + \boldsymbol{\xi}_{t,\text{re}} \qquad & f_{t,\text{re}} &= \epsilon_{t,\text{re}} \\
\mathbf{x}_{t,\text{ir}} &= \boldsymbol{\lambda}_{\text{ir}} f_{t,\text{ir}} + \boldsymbol{\xi}_{t,\text{ir}} \qquad & f_{t,\text{ir}} &= \epsilon_{t,\text{ir}} \qquad \text{(DGPII)} \\
\boldsymbol{\xi}_t &:= (\boldsymbol{\xi}'_{t,\text{re}}, \boldsymbol{\xi}'_{t,\text{ir}})' \sim \mathcal{N}(\mathbf{0}_N, \mathbf{\Sigma}_{\boldsymbol{\xi}}), \text{ where } N = N_{\text{re}} + N_{\text{ir}} \\
& R = 1, \; Q = 0, \text{ and } P = 0.
\end{aligned}$$

The motivation for using the above DGPs is twofold. First, since the DGPs closely resemble the processes in the Giannone et al. (2008) framework, it can be assumed that the nowcasting step estimates the true parametrisation of the model more precisely if the variables stemming from the relevant processes have been identified correctly. However, it can be expected that the variable selection step is unstable, because the factor driven dependence structure has to be approximated linearly in the pre-selection step. Comparing DGPI and DGPII, it might be possible to infer on the role of the complexity of the underlying DGP on the variable selection and nowcasting accuracy.

The third DGP considered (DGPIII) is conceptually different from the ones above. Here, the data set contains variables that are generated from two VARMA$(1, 1)$ processes, where

the errors are cross-correlated.

$$\begin{aligned}
\mathbf{x}_{t,\text{re}} &= \boldsymbol{\Pi}_{\text{re}}\mathbf{x}_{t-1,\text{re}} + \boldsymbol{\Theta}_{\text{re}}\boldsymbol{\xi}_{t-1,\text{re}} + \boldsymbol{\xi}_{t,\text{re}} \\
\mathbf{x}_{t,\text{ir}} &= \boldsymbol{\Pi}_{\text{ir}}\mathbf{x}_{t-1,\text{ir}} + \boldsymbol{\Theta}_{\text{ir}}\boldsymbol{\xi}_{t-1,\text{ir}} + \boldsymbol{\xi}_{t,\text{ir}} \qquad \text{(DGPIII)} \\
\boldsymbol{\xi}_t := (\boldsymbol{\xi}'_{t,\text{re}}, \boldsymbol{\xi}'_{t,\text{ir}})' &\sim \mathcal{N}(\mathbf{0}_N, \boldsymbol{\Sigma}_{\boldsymbol{\xi}}), \text{ where } N = N_{\text{re}} + N_{\text{ir}}
\end{aligned}$$

Contrary to the DGPs described above, the TP approach is expected to perform comparatively well when applied to data generated by DGPIII since the model is linear in the predictors by construction. Particularly, the variable selection is expected to be more stable. The nowcasting step, however, might create more problems for the DFM since the model might not be approximated well by a small number of common factors. Therefore, including irrelevant variables and cross-correlated errors might heavily influence the nowcasting step. We expect that for the VARMA parametrisation, the TP approach should result in considerably more precise nowcasts compared to the CDFM.[15]

The last DGP (DGPIV) is a combination of two symmetric DFMs where the measurement equation includes lags of the latent factors which follow a VMA process. Again, the structure of the DFMs was taken from Bai and Ng (2007). Formally, the data are generated via

$$\begin{aligned}
\mathbf{x}_{t,\text{re}} &= (\boldsymbol{\Lambda}_{0,\text{re}} + \boldsymbol{\Lambda}_{1,\text{re}}\mathbb{L} + \boldsymbol{\Lambda}_{2,\text{re}}\mathbb{L}^2)\mathbf{f}_{t,\text{re}} + \boldsymbol{\xi}_{t,\text{re}} \\
\mathbf{f}_{t,\text{re}} &= \boldsymbol{\Psi}_{\text{re}}\boldsymbol{\epsilon}_{t-1,\text{re}} + \boldsymbol{\epsilon}_{t,\text{re}} \\
\mathbf{x}_{t,\text{ir}} &= (\boldsymbol{\Lambda}_{0,\text{ir}} + \boldsymbol{\Lambda}_{1,\text{ir}}\mathbb{L} + \boldsymbol{\Lambda}_{2,\text{ir}}\mathbb{L}^2)\mathbf{f}_{t,\text{ir}} + \boldsymbol{\xi}_{t,\text{ir}} \qquad \text{(DGPIV)} \\
\mathbf{f}_{t,\text{ir}} &= \boldsymbol{\Psi}_{\text{ir}}\boldsymbol{\epsilon}_{t-1,\text{ir}} + \boldsymbol{\epsilon}_{t,\text{ir}} \\
\boldsymbol{\xi}_t := (\boldsymbol{\xi}'_{t,\text{re}}, \boldsymbol{\xi}'_{t,\text{ir}})' &\sim \mathcal{N}(\mathbf{0}_N, \boldsymbol{\Sigma}_{\boldsymbol{\xi}}), \text{ where } N = N_{\text{re}} + N_{\text{ir}} \\
R = 6,\ Q = 2,\ &\text{and } P = \infty.
\end{aligned}$$

[15] Since the DFM is misspecified for this DGP, it is not surprising that using the information criteria provided in Bai and Ng (2002) to estimate the number of common factors is not reliable. The procedure appears to almost always select the maximum number of factors provided by the researcher. Consequently, the convention is chosen to fix the number of common factors to one and force the factor model to find a parsimonious representation of the data.

Both processes in DGPIV are highly parameterized and highly dynamic DFMs that differ from the underlying model of the Giannone et al. (2008) framework, where the factors have an autoregressive lag structure and the observed variables are not dependent on the lags of the factors explicitly. Especially since the nowcasting framework treats the lags of the factors as additional static factors, it might be expected that, in general, the nowcasting procedure is less precise. Furthermore, since the factors follow a VMA process, it is likely that P is estimated to be large thus further increasing the dimensionality of the estimated model. The variable selection step is assumed to be very unstable since a linear fit of the highly complex DGP is difficult to achieve. Summarizing, it is expected that the TP approach, if at all, only leads to small improvements of the nowcasting accuracy in this case.

For the parametrisation of the DGPs, we follow Bai and Ng (2007) and draw the factor loading matrices and vectors $\mathbf{\Lambda}_{\text{re}} = \{\lambda_{n,r,\text{re}}\}_{n=1,r=1}^{N_{\text{re}},R}$, $\mathbf{\Lambda}_{\text{ir}} = \{\lambda_{n,r,\text{ir}}\}_{n=1,r=1}^{N_{\text{ir}},R}$ and $\boldsymbol{\lambda}_{\text{re}} = \{\lambda_{n,\text{re}}\}_{n=1}^{N_{\text{re}}}$, $\boldsymbol{\lambda}_{\text{ir}} = \{\lambda_{n,\text{ir}}\}_{n=1}^{N_{\text{ir}}}$ from a standard normal multivariate distribution. This is also the case for the factor loadings corresponding to the lagged factors in DGPIV. Furthermore, for all error vectors of the factor processes it holds that $\boldsymbol{\epsilon}_{t,\text{re}} \sim \mathcal{N}(\mathbf{0}_{N_{\text{re}}}, \mathbf{I}_{N_{\text{re}}})$, $\boldsymbol{\epsilon}_{t,\text{ir}} \sim \mathcal{N}(\mathbf{0}_{N_{\text{ir}}}, \mathbf{I}_{N_{\text{ir}}})$, $\epsilon_{t,\text{re}} \sim \mathcal{N}(0,1)$, and $\epsilon_{t,\text{ir}} \sim \mathcal{N}(0,1)$. For DGPI, DGPII, and DGPIV, the covariance matrix of the idiosyncratic error $\mathbf{\Sigma}_{\boldsymbol{\xi}}$, having nonzero off-diagonal elements, is constructed by $\mathbf{\Sigma}_{\boldsymbol{\xi}} = \mathbf{I}_N \mathbf{R} \mathbf{I}_N$, where $\mathbf{R}$ is a randomly drawn correlation matrix using the `rcorrmatrix` function of the `clusterGeneration` package by Qiu and Joe (2020). For DGPIII, $\mathbf{I}_N$ is replaced by $0.005 \cdot \mathbf{I}_N$ for the construction of $\mathbf{\Sigma}_{\boldsymbol{\xi}}$, such that the simulated series have a similar magnitude compared to those drawn from the other DGPs before standardization. For the low SNR case, the variances are multiplied by a factor of 5. The starting vector for each factor, or in the case of DGPIII for the variables, is chosen to be zero, or the zero vector, and a burn-in period of 200 observations is used.

For DGPI and DGPIV, it is chosen that $\mathbf{\Phi}_{\text{re}} = \mathbf{\Phi}_{\text{ir}} = \text{diag}(0.2, 0.375, 0.55, 0.725, 0.9)$ and $\mathbf{\Psi}_{\text{re}} = \mathbf{\Psi}_{\text{ir}} = \text{diag}(0.2, 0.9)$, respectively (Bai and Ng, 2007). Further, the matrices $\mathbf{\Upsilon}_{\text{re}}$ and $\mathbf{\Upsilon}_{\text{ir}}$ are calculated via $\mathbf{\Upsilon}_{\text{re}} = \mathbf{A}_{\text{re}} \mathbf{S}_{\text{re}} \mathbf{A}_{\text{re}}$ and $\mathbf{\Upsilon}_{\text{ir}} = \mathbf{A}_{\text{ir}} \mathbf{S}_{\text{ir}} \mathbf{A}_{\text{ir}}$, where $\mathbf{A}_{\text{re}}$ and $\mathbf{A}_{\text{ir}}$ are random

orthonormal matrices computed via the `randortho` function provided by the `pracma` package by Borchers (2021), and $\mathbf{S}_{\text{re}}$ and $\mathbf{S}_{\text{ir}}$ are diagonal matrices of rank Q with elements drawn from $\mathcal{U}(0.8, 0.12)$. To ensure stationarity in DGPIII, $\mathbf{\Pi}_{\text{re}} = \mathbf{P}_{\text{re}}\mathbf{D}_{\text{re}}\mathbf{P}_{\text{re}}$ and $\mathbf{\Pi}_{\text{ir}} = \mathbf{P}_{\text{ir}}\mathbf{D}_{\text{ir}}\mathbf{P}_{\text{ir}}$, where $\mathbf{P}_{\text{re}}$ and $\mathbf{P}_{\text{ir}}$ are matrices with random entries drawn from $\mathcal{U}(-0.5, 0.5)$, and $\mathbf{D}_{\text{re}}$ and $\mathbf{D}_{\text{ir}}$ are diagonal with random entries drawn from $\mathcal{U}(0.1, 0.8)$. To guarantee invertability, $\mathbf{\Theta}_{\text{re}}$ and $\mathbf{\Theta}_{\text{ir}}$ are constructed analogously.

Since the study aims to simulate a real world nowcasting exercise, a quarter of both the relevant and irrelevant variables is aggregated via (4) to resemble quarterly observations. While this does not influence the extraction of the factors in the Giannone et al. (2008) procedure, it is expected to have an impact on the variable selection step. Considering the lag structure, only the target variable is lagged by 90 days. We refrain from imposing a random lag structure on the predictors since this would only influence the relative TP performance if the publication lags were longer than the one for the target variable, which is rarely the case in practice.[16]

The data is drawn for $(T, N_{\text{re}}, N_{\text{ir}}) = \{(200, 150, 50), (200, 100, 100), (200, 50, 150)\}$ to evaluate the TP procedure for different combinations of relevant and irrelevant variables. The expanding nowcasting window spans the last 100 observations while the first 100 are used as the initial training set.[17]

3.2 Results

To evaluate the performance of the different TP specifications, the RMSFE taken over the complete nowcasting period is reported. The results are found in Table 1 and 2 where the relative performance with respect to the CDFM is highlighted.

We arrive at four general conclusions. First, in most cases, it is found that the ODFM

[16]The lag structure of the predictors might also be of concern if one aims to forecast h-steps ahead rather then to nowcast. In this case the CV procedure would shift the focus of predicting one step ahead to h steps ahead.

[17]Note that T corresponds to the time index convention used throughout Section 2, i.e., for each DGP we train the models on 300 observations for the monthly predictors and 100 observations for the quarterly predictors and use a similar number of observations for the nowcasting window.

model outperforms the CDFM model which is itself outperformed by some of the TP specifications. This gives clear evidence that (i) TP seems to be a viable option to improve nowcast accuracy, and (ii) the optimal set of TP is not always equal to the set of relevant variables. In other words, even if the true model specification was known, targeting predictors might still improve the nowcasting performance exploiting the dependence structure in a given data set.

Second, the TP procedure achieves the worst nowcasting results for $N_{\text{re}} = 50$ and $N_{\text{ir}} = 150$. Both in terms of absolute and relative performance, a clear increase in the RMSE is found for all TP specifications. While this is to be expected, because having more irrelevant variables in the data set on average leads to a higher likelihood of selecting more irrelevant variables, it clearly shows that the TP procedure should be used with care. Simply adding (possibly irrelevant) variables to the data set does not boost the TP performance.

Third, using CV methods seems to outperform choosing a fixed number of predictors. While the A-F-TPN provides the lowest RMSFE for 5 out of 12 DGP configurations in the high SNR setting, it is also highly volatile resulting in RMSFE increases by up to 81%. The nowcast validated specifications only achieve the lowest RMSFE in 4 out of 12 cases. However, the S-N-TPN specification is very reliable beating the benchmark CDFM in 9 out of 12 DGP configurations. In the low SNR setting, the nowcast validated specifications provide the lowest RMSFE in 9 out of 12 configurations clearly outperforming the fixed specifications. This hints towards the importance of using the information contained in the data when applying the TP framework, instead of relying on ad-hoc solutions, i.e., selecting the number of predictors beforehand.

Finally, decreasing the signal-to-noise ratio considerably worsens the overall performance of the TP framework. The TP procedure often results in an RMSFE improvement of only about 1%. Note, however, that in this setting the CDFM provides results that are often only marginally better than the more parsimonious AR(1), ARMA, and naive mean benchmarks, if at all. Therefore, using TP models can still be a viable alternative to improve the nowcasting

performance of mixed-frequency DFMs in noisier settings.

[Table 1 here]

[Table 2 here]

Turning to the relative performance of each specification in more detail, we find that most TP specifications are able to provide lower RMSFEs than the most important benchmark model (CDFM). The S-N-TPN specification is found to be the one that outperforms the CDFM benchmark most often, beating the benchmark in 9 out of 12 cases. Taking the average over the DGPs, it is found that the A-LE-TPN, S-LE-TPN, and S-N-TPN provide a lower RMSFE than the CDFM. Accounting for the expectedly bad performance in the case of DGPIV (see Table 1) and only taking averages over DGPI – DGPIII, it is found that each specification results in a lower RMSFE by at least 4%. The single best performing specification is found to be the A-N-TPN for DGPIII with $N_{\text{re}} = 100$ and $N_{\text{ir}} = 100$. Here, the RMSFE is reduced by 35% compared to the CDFM.

In the case of a low SNR (Table 2), the picture changes slightly. The best performing parametrisation is again the S-N-TPN, which results in a lower RMSFE in 10 out of 12 cases. However, it must be noted that these improvements are considerably smaller than those found in the high SNR case. Averaging over all DGPs, it is now found that all model parametrisations provide a lower RMSFE by at least 1%. The single best performing specifications is, again, the A-N-TPN parametrisations for DGPIII using $N_{\text{re}} = 100$ and $N_{\text{ir}} = 100$ with a reduction of about 32%.

The simulation results do not reveal a substantial difference between using the LARS-EN and coordinate descent algorithm in the family of data driven models. However, since the former is computationally more costly than the latter, our results provide some evidence against the general notion in the TP literature on preferring the LARS-EN algorithm. The same holds true when comparing skip-sampling with aggregating the higher frequency data in the family of data driven specifications. No systematic difference can be observed between the

performance of models using skip-sampling to those using aggregation. Since skip-sampling is also computationally more costly, our findings suggest to use aggregation in empirical applications. For the nowcast validated models, the findings are somewhat more distinct. While the S-N-TPN procedure appears to be more consistent in achieving an RMSFE improvement over the CDFM, the A-N-TPN parametrisation appears to result in higher but less frequent improvements.

When TP specifications are applied to DGPs with underlying VARMA representation, the resulting nowcasts are substantially more accurate than those of the benchmark models. A possible explanation for this finding might be that estimating DFMs via the Giannone et al. (2008) framework is too noisy due to the correlation structure of the idiosyncratic error. This shows that variable pre-selection might be essential in the case where the underlying model cannot be well approximated by a DFM. Further, it suggests that simpler, more parsimonious, models like a VARMA model with pre-screened data, might be promising alternative model choices for nowcasting in similar settings.

Turning to DGPIV, a highly complex dynamic state-space-model with moving average components, we observe that while TP models are still able to achieve RMSFE reductions of about 2–4% and up to 7%, they do not yield the same RMSFE reductions that are reported for the other DGPs. In general, the TP models are more often unable to outperform the CDFM model, and some model specifications, such as those selecting a fixed number of variables in the high SNR setting, even increase the RMSFE. This might be due to the fact that the highly complex dynamics of the system cannot be captured by the EN variable selection step. This could be an indicator that, for such DGPs, a different variable selection tool might be necessary and ideally should be combined with a nowcasting step that accounts for the moving average components in the factor structure.

Besides the forecast performance, we also study the variable selection of all model candidates. As expected, the variable selection procedure is highly unstable over time. Only for DGPIII, the pre-selection step seems to be able to pick the relevant variables with sufficient

accuracy. These findings might explain why the family of fixed value models, i.e. A- and S-F-TPN, sometimes outperform the data driven and nowcast validated specifications. Since no additional systematic conclusions can be drawn from the results of the variable selection, we opt not to report them here.

4 Empirical Application

4.1 Data

To evaluate the TP specifications in a real world scenario, three different data sets are chosen. Each data set has unique features allowing us to investigate each model's performance in different settings.

The first data set contains the US GDP growth rate (log first differences) and various monthly predictor variables. We denote the data set by US. Log GDP growth is retrieved from U.S. Bureau of Economic Analysis (2024) and the set of monthly predictors is constructed based on the "FRED MD" database provided by McCracken (2024). The original data set contains 127 monthly US indicators grouped into the categories such as, for example, "Income", "Labour Market", and "Housing" (for details, see McCracken and Ng, 2016). We omit variables from the predictor set that represent disaggregates such as, for example, consumer price indices disaggregated by product category. We presume that the additional information provided by the disaggregates should be outweighed by the additional noise introduced to the estimation problem by including more highly correlated predictors. Following the results of the simulation study, they are therefore excluded. Furthermore, three additional variables are excluded due to missing observations at the beginning of the nowcasting window. This results in 87 predictors overall. We use data from 1979Q1 to 2019Q3, which corresponds to 492 monthly observations.

Second, we employ a data set used by Wollmershäuser et al. (2020) to nowcast German GDP (hereinafter referred to as GER). The main motivation behind using this data set

is its sheer size. After excluding indicators with too many missing observations, the data set includes 443 monthly predictors. Therefore, it is assumed that using TP will lead to a substantial improvement of the nowcasting accuracy by reducing the dimensionality of the estimation problem. For this data set, we use the observations from January 1991 to September 2019.

The third data set was collected by Kuck and Schweikert (2021) to nowcast regional GDP growth of the German state of Baden-Württemberg. The data set reports a mixture of 18 quarterly and 78 monthly regional, national, and international hard and soft indicators with observations from January 1996 to December 2020, from which 8 are omitted due to missing observations. We omit three additional variables so that we are able to publish the data set, since these predictors are retrieved from proprietary data sources. The main motivation behind the choice of this data set is that the TP approach is very promising in a setting with a high spatial resolution like, for example, regional GDP growth. Providing good nowcasts on a regional level is often more complicated than nowcasting GDP growth at the country level due to the limited availability of regional predictor variables. Prior studies rely on a relatively small number of regional predictors combined with many predictors that are available at an aggregated level (Eickmeier and Ng, 2011; Kopoin et al., 2013; Lehmann and Wohlrabe, 2015; Kuck and Schweikert, 2021). A priori, it is unclear whether national or international predictors have predictive power for such target variables or instead simply add noise to the factor estimation. Therefore, we assess whether TP can improve the DFM performance by eliminating these potentially irrelevant variables.

For all three data sets, we compute out-of-sample nowcasts from 2012Q1 to 2019Q3. We choose this period for several reasons. First, it is the most recent period of relative economic ease between the global financial crisis and the COVID-19 pandemic. Second, since structural breaks are likely to occur when working with real data, the length of the period used in the out-of-sample evaluation set should be chosen with care. By restricting the period to roughly eight years for the evaluation set, we hope to strike a good balance between a sufficiently

reliable result and a low probability of structural breaks.

4.2 Results

Similar to Section 3, we report the relative RMSFE in Table 3 for each model specification applied to each real world data set and additionally conduct Diebold-Mariano tests to compare the predictive performances. Again, we find that using TP can lead to substantial improvements of the RMSFE compared to the CDFM benchmark, but the results differ considerably between the selected data sets. It is found that 3 out of 8 TP specifications for the US data set, 3 out of 7 TP specifications for the GER data set, and 6 out of 8 TP specifications for the BW data set are able to beat the benchmark. These improvements can amount to 10% for the US, 16% for the GER and as high as 21% for the BW data set. It must be noted that for the case of the GER data set, it is not possible to evaluate the S-LE-TPN due to computational costs. Since we do not find that this specification consistently outperforms the other parametrisations in the simulation study, and we aim to evaluate the performance of all TP specifications using the largest available number of predictors, we refrain from omitting predictors for this specification.

Discussing the specifications in more detail, the results show that the S-CD-TPN and S-N-TPN specifications are able to beat the CDFM benchmark for all three data sets considered. These improvements can be substantial with the S-CD-TPN specification reaching the lowest RMSFE value for the US and GER data sets. The S-N-TPN specification, which performs generally well in the simulations, also reaches the lowest RMSFE value for the GER data set and narrowly beats the S-CD-TPN specification in case of the BW data set.

Comparing the performance of different TP model specification families, we find that the fixed specifications only improve the nowcasts in one instance, otherwise leading to larger RMSFE values than the CDFM benchmark. Hence, the fixed specifications that are very prominent in the nowcast literature are largely dominated by the data driven and nowcast validated specifications. This shows that the conventional strategy of cross-validating the

tuning parameters can lead to nowcasting performance gains in empirically relevant cases.

For the BW data set, it appears that the large improvements in nowcasting accuracy can be attributed to the fact that the underlying data is comprised by mostly national and international indicators with the goal of nowcasting state level GDP. Therefore, it is very likely that some, or potentially many, of the predictors are uninformative, which could lead to noisy factor estimation if these predictors are included. Thus, pre-selecting the predictors might be crucial when the task is to nowcast a variable of interest for which "relevant" predictors are only sparsely available. This is often the case for regional forecasting applications.[18] For this data set, we can achieve the largest RMSFE improvements, highlighting that eliminating some of those predictor can improve the factor extraction and consequently leads to better nowcasts. Figure 1 depicts a comparison of the model forecasts over time. It appears that the largest differences across model specifications can be observed at the turning points of the business cycle.

With respect to the GER data set, the improvements might stem from the considerable dimensionality reduction of the predictor space after the pre-selection step. Since several studies have shown that increasing the number of predictors can substantially increase the nowcasting error, the results for the GER data set might be a prime example for the potential of TP when working with a large number of predictors. However, in this case, we observe mixed results with some TP specifications improving the nowcasting results while others lead to substantially larger nowcasting errors than the CDFM. Particularly, A-CD-TPN and A-N-TPN have (significantly) larger RMSFEs compared with the CDFM. Figure 2 shows that while the performance of A-N-TPN can be attributed to an outlying prediction in 2013, A-CD-TPN generally misses its target. The very large cross-section of predictors presents a challenging problem for the EN estimator and this might explain the ambiguous results.

For the US data set, it is only possible to achieve smaller improvements. Here, the largest

[18] We repeated the nowcasting exercise for a dataset comprised only of regional indicators. Here, the performance gains by using TP are much smaller. The results are not reported but can be obtained from the authors upon request.

reduction in the RMSFE is found to be 10%. Furthermore, most of the TP specifications are not able to beat the CDFM benchmark. Figure 3 shows that the forecasts are almost identical over the full nowcasting period. A possible explanation for this might be that the data set is already well selected and does not benefit from targeted pre-selection. Nonetheless, the results still show that it is possible to slightly decrease the nowcasting error for the US data set by using TP, for example, by cross-validating the tuning parameters with previous nowcasting errors.

[Table 3 here]

[Figure 3 here]

[Figure 2 here]

[Figure 1 here]

5 Conclusion

In this paper, we investigate whether pre-selecting the mixed-frequency DFM can lead to more accurate nowcasts when the pre-selection step is implemented in a purely data driven way. Moreover, the influence of the EN tuning parameters on the model performance is explored. Our specifications also include models for which the EN parameters are set to fixed values, as is common in the TP literature, and novel procedures where the EN acts as a first pass filter and the previous period's nowcast error is evaluated to determine the final targeted predictors set.

Our results suggest that a pre-selection step can lead to more accurate nowcasts and, in some cases, TP can also outperform an oracle DFM for which only the relevant variables are selected. It is found that for every DGP and all three real world data sets, at least one TP specification minimises the RMSFE. The signal-to-noise ratio and the ratio between relevant

and irrelevant variables determine the potential improvements of the TP approach over the benchmark DFM using all predictors.

For the simulated data, it is found that the models with nowcast validated EN parameters outperform the fixed and data driven nowcast validated specifications for most DGPs. They perform particularly well if the DGP does not follow a factor structure and the EN is able to accurately select the relevant variables. Choosing a fixed number of predictors in the pre-selection step performs well if the DGP has a factor structure and variable selection via the EN tends to be difficult. For the real world data, nowcast-validated and data driven specifications generally provide lower RMSFE than the fixed models.

Our simulation results clearly show that it is not recommendable to simply expand the data set with available variables and to rely on accurate pre-selection to improve the nowcasting performance. Thus, even when pre-selecting variables can, in principle, be achieved effectively, data sets should be carefully selected based on economic intuition. Further, we do not find any considerable and systematic differences between the models using the LARS-EN algorithm and coordinate descent. The same holds true for the differences between skip-sampling and aggregation. Since using LARS-EN and skip-sampling in combination is computationally much more costly than their alternatives, we recommend to use aggregation and coordinate descent for large empirical data sets.

Regarding future research, it should be investigated whether it is possible to achieve further improvements of the pre-selection step by more heavily restricting the data driven model specifications. This might result in a variable selection that is more flexible than in the fixed specification, while at the same time reducing the variable selection instability of the completely data driven models. Additionally, the framework should be tested against more parsimonious (linear) models since the simulation study suggests that targeting predictors appears to be most effective when the underlying data is not well approximated by a factor model. For example, a VARMA model with pre-screened data could be a promising alternative to the TP DFM in such cases. Alternatively, the pre-selection step and the nowcasting

step could be combined when the DFM is estimated with the option of shrinking some factor loadings to zero (Kaufmann and Schumacher, 2017, 2019). In this way, the issue of combining a linear pre-selection method (EN) with a factor driven nowcasting model (DFM) could be avoided.

Acknowledgments

We thank Thomas Dimpfl, Robert Jung, Konstantin Kuck, and participant of the THE Workshop, the Doctoral Seminar at the University of Hohenheim, and the German Statistical Week in Münster for valuable comments and suggestions. Further, we thank Robert Lehmann for sharing the ifoCAST data with us. Access to Refinitiv Eikon, provided by the Hohenheim Datalab (DALAHO), is gratefully acknowledged.

References

Andreini, P., Hasenzagl, T., Reichlin, L., Senftleben-König, C., Strohsal, T., 2021. Now-casting German GDP. International Journal of Forecasting doi:10.1016/j.ijforecast.2021.11.009.

Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70, 191–221. doi:10.1111/1468-0262.00273.

Bai, J., Ng, S., 2007. Determining the number of primitive shocks in factor models. Journal of Business and Economic Statistics 25, 52–60. doi:10.1198/073500106000000413.

Bai, J., Ng, S., 2008. Forecasting economic time series using targeted predictors. Journal of Econometrics 146, 304–317. doi:10.1016/j.jeconom.2008.08.010.

Bańbura, M., Giannone, D., Modugno, M., Reichlin, L., 2013. Now-casting and the real-time data flow, in: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier. volume 2 of *Handbook of Economic Forecasting*, pp. 195–237. doi:10.1016/B978-0-444-53683-9.00004-9.

Bańbura, M., Rünstler, G., 2011. A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting gdp. International Journal of Forecasting 27, 333–346. doi:10.1016/j.ijforecast.2010.01.011.

Bessec, M., 2013. Short-term forecasts of french gdp: A dynamic factor model with targeted predictors. Journal of Forecasting 32, 500–511. doi:10.1002/for.2262.

Boivin, J., Ng, S., 2006. Are more data always better for factor analysis? Journal of Econometrics 132, 169–194. doi:https://doi.org/10.1016/j.jeconom.2005.01.027.

Bok, B., Caratelli, D., Giannone, D., Sbordone, A.M., Tambalotti, A., 2018. Macroeconomic Nowcasting and Forecasting with Big Data. Annual Review of Economics 10, 615–643. doi:10.1146/annurev-economics-080217-053214.

Borchers, H.W., 2021. pracma: Practical numerical math functions. URL: https://CRAN.R-project.org/package=pracma. r package version 2.3.3.

Bulligan, G., Marcellino, M., Venditti, F., 2015. Forecasting economic activity with targeted predictors. International Journal of Forecasting 31, 188–206. doi:10.1016/j.ijforecast.2014.03.004.

Castle, J.L., Clements, M.P., Hendry, D.F., 2013. Forecasting by factors, by variables, by both or neither? Journal of Econometrics 177, 305–319. doi:10.1016/j.jeconom.2013.04.015.

Doz, C., Giannone, D., Reichlin, L., 2011. A two-step estimator for large approximate dynamic factor models based on kalman filtering. Journal of Econometrics 164, 188–205. doi:10.1016/j.jeconom.2011.02.012.

Doz, C., Giannone, D., Reichlin, L., 2012. A Quasi-Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models. The Review of Economics and Statistics 94, 1014–1024. doi:10.1162/REST_a_00225.

Eddelbuettel, D., Francois, R., 2011. Rcpp: Seamless R and C++ integration. Journal of Statistical Software 40, 1–18. doi:10.18637/jss.v040.i08.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least angle regression. The Annals of Statistics 32, 407 – 499. doi:10.1214/009053604000000067.

Eickmeier, S., Ng, T., 2011. Forecasting national activity using lots of international predictors: An application to new zealand. International Journal of Forecasting 27, 496–511. doi:10.1016/j.ijforecast.2009.10.011.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. Journal of statistical software 33, 1. doi:10.18637/jss.v033.i01.

Giannone, D., Reichlin, L., Small, D., 2008. Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics 55, 665–676. doi:10.1016/j.jmoneco.2008.05.010.

Girardi, A., Golinelli, R., Pappalardo, C., 2017. The role of indicator selection in nowcasting euro-area gdp in pseudo-real time. Empirical Economics 53, 79–99. doi:10.1007/s00181-016-1151-z.

Hastie, T., Tibshirani, R., Wainwright, M., 2015. Statistical learning with sparsity: The lasso and generalizations. Chapman & Hall/CRC.

Hoerl, A.E., Kennard, R.W., 1988. Ridge regression. Wiley, New York.

Hyndman, R.J., Athanasopoulos, G., 2018. Forecasting: principles and practice. 3 ed., OTexts Melbourne.

Kaufmann, S., Schumacher, C., 2017. Identifying relevant and irrelevant variables in sparse factor models. Journal of Applied Econometrics 32, 1123–1144. doi:10.1002/jae.2566.

Kaufmann, S., Schumacher, C., 2019. Bayesian estimation of sparse dynamic factor models with order-independent and ex-post mode identification. Journal of Econometrics 210, 116–134. doi:10.1016/j.jeconom.2018.11.008.

Kim, H.H., Swanson, N.R., 2014. Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence. Journal of Econometrics 178, 352–367. doi:10.1016/j.jeconom.2013.08.033.

Kim, H.H., Swanson, N.R., 2018. Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. International Journal of Forecasting 34, 339–354. doi:10.1016/j.ijforecast.2016.02.012.

Kopoin, A., Moran, K., Pare, J.P., 2013. Forecasting regional gdp with factor models: How

useful are national and international data? Economics Letters 121, 267–270. doi:10.1016/j.econlet.2013.08.007.

Kuck, K., Schweikert, K., 2021. Forecasting Baden-Württemberg's GDP growth: MIDAS regressions versus dynamic mixed-frequency factor models. Journal of Forecasting 40, 861–882. doi:10.1002/for.2743.

Lehmann, R., Wohlrabe, K., 2015. Forecasting GDP at the regional level with many predictors. German Economic Review 16, 226–254. doi:10.1111/geer.12042.

Marcolino de Mattos, D., Costa Ferreira, P., de Valk, S., Branco Gomes, G., 2019. nowcasting: Predicting economic variables using dynamic factor models. URL: https://CRAN.R-project.org/package=nowcasting. r package version 1.1.4.

Mariano, R.S., Murasawa, Y., 2003. A new coincident index of business cycles based on monthly and quarterly series. Journal of Applied Econometrics 18, 427–443. doi:10.1002/jae.695.

McCracken, M.W., 2024. FRED-MD and FRED-QD: monthly and quarterly databases for macroeconomic research. URL: https://research.stlouisfed.org/econ/mccracken/fred-databases/. federal Reserve Bank of St. Louis.

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

Proietti, T., Moauro, F., 2006. Dynamic Factor Analysis with Non-Linear Temporal Aggregation Constraints. Journal of the Royal Statistical Society. Series C 55, 281–300. doi:10.1111/j.1467-9876.2006.00536.x.

Qiu, W., Joe, H., 2020. clusterGeneration: Random cluster generation (with specified degree of separation). URL: https://CRAN.R-project.org/package=clusterGeneration. r package version 1.3.7.

Schumacher, C., 2010. Factor forecasting using international targeted predictors: The case of german gdp. Economics Letters 107, 95–98. doi:10.1016/j.econlet.2009.12.036.

Siliverstovs, B., 2017. Short-term forecasting with mixed-frequency data: a midasso approach. Applied Economics 49, 1326–1343. doi:10.1080/00036846.2016.1217310.

Smeekes, S., Wijler, E., 2018. Macroeconomic forecasting using penalized regression methods. International journal of forecasting 34, 408–430.

Stock, J.H., Watson, M.W., 2002a. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179. doi:10.1198/016214502388618960.

Stock, J.H., Watson, M.W., 2002b. Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics 20, 147–162. doi:10.1198/073500102317351921.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x.

U.S. Bureau of Economic Analysis, 2024. Gross domestic product [gdp]. URL: https://fred.stlouisfed.org/series/GDP. retrieved from FRED, Federal Reserve Bank of St. Louis.

Wollmershäuser, T., Göttert, M., Grimme, C., Lautenbacher, S., Lehmann, R., Link, S., Menkhoff, M., Möhrle, S., Rathje, A.C., Reif, M., Sandqv, P., 2020. ifo Konjunkturprognose Winter 2020: Das Coronavirus schlägt zurück – erneuter Shutdown bremst Konjunktur ein zweites Mal aus. ifo Schnelldienst 73, 03–61.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67, 301–320. doi:10.1111/j.1467-9868.2005.00503.x.

Zou, H., Hastie, T., Tibshirani, R., 2007. On the degrees of freedom of the lasso. The Annals of Statistics 35, 2173 – 2192. doi:10.1214/009053607000000127.

Tables

Table 1: Relative root mean squared error of 100 targeted predictors nowcasts for $N = 200$ variables generated with a high signal-to-noise ratio

Model	$N_{\text{rel}} = 150,\ N_{\text{ir}} = 50$				$N_{\text{rel}} = 100,\ N_{\text{ir}} = 100$				$N_{\text{rel}} = 50,\ N_{\text{ir}} = 150$			
	DGPI	DGPII	DGPIII	DGPIV	DGPI	DGPII	DGPIII	DGPIV	DGPI	DGPII	DGPIII	DGPIV
Naive	1.22	1.06	1.37	4.25	1.19	1.06	1.10	4.18	1.22	1.06	0.99	3.86
AR	1.17	1.06	1.37	4.24	1.14	1.06	1.10	4.17	1.16	1.06	1.02	3.85
ARMA	1.17	1.06	1.37	4.24	1.14	1.06	1.10	4.17	1.16	1.06	1.02	3.85
ODFM	1.02	1.00	1.00	0.95	0.98	1.00	0.77	**0.97**	0.99	1.00	1.00	0.95
CDFM	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A-CD-TPN	**0.90**	0.95	0.97	1.77	0.94	1.02	0.94	1.47	0.99	0.99	0.95	1.63
A-LE-TPN	1.00	0.99	1.00	0.99	0.93	0.99	0.77	1.01	0.98	0.98	1.00	1.23
S-CD-TPN	0.97	1.02	0.88	1.37	1.02	1.02	0.89	1.23	0.93	1.00	0.99	1.20
S-LE-TPN	1.02	1.00	0.96	0.98	0.98	0.99	0.78	**0.97**	1.00	1.00	1.00	**0.96**
A-F-TPN	0.92	**0.81**	1.14	1.81	**0.86**	**0.79**	0.73	1.23	**0.86**	**0.95**	0.98	1.17
S-F-TPN	0.93	0.91	1.09	1.40	0.90	0.87	0.77	1.56	0.94	0.96	1.01	1.03
A-N-TPN	0.94	1.03	**0.79**	1.21	1.05	1.01	**0.65**	1.20	0.88	1.00	**0.91**	1.27
S-N-TPN	0.98	0.99	0.84	**0.97**	0.94	1.01	0.67	1.10	0.94	0.96	0.95	1.17

Note: The error measure is reported relative to the benchmark CDFM. Bold printed values indicate the lowest achieved RMSFE. The model acronyms correspond to the following specifications: AR: An AR(1) model; ARMA: An ARMA model with lag order determined recursively via the AIC; Naive: A naive unconditional mean model; ODFM: A DFM using only the relevant variables; CDFM: A DFM using all variables; A-CD-TPN (S-CD-TPN): The set of targeted predictors is retrieved via the EN, which is solved using coordinate descent, with EN parameters retrieved via CV. The data is aggregated (skip-sampled); A-LE-TPN (S-LE-TPN): The EN is solved via LARS-EN; A-N-TPN (S-N-TPN): A TP model where the set of targeted predictors is retrieved using the nowcasting error from the previous period; A-F-TPN (S-F-TPN): A TP model where the set of targeted predictors is retrieved via the EN, where the EN parameters are set so that 30 predictors are selected. The variables are newly selected for each nowcast.

Table 2: Relative root mean squared error of 100 targeted predictors nowcasts for $N = 200$ variables generated with a low signal-to-noise ratio

Model	$N_{\text{rel}} = 150,\ N_{\text{ir}} = 50$				$N_{\text{rel}} = 100,\ N_{\text{ir}} = 100$				$N_{\text{rel}} = 50,\ N_{\text{ir}} = 150$			
	DGPI	DGPII	DGPIII	DGPIV	DGPI	DGPII	DGPIII	DGPIV	DGPI	DGPII	DGPIII	DGPIV
Naive	1.01	0.99	1.33	1.22	1.00	0.99	1.15	1.10	0.99	1.00	1.00	1.00
AR	0.98	0.99	1.32	1.21	0.97	0.99	1.16	1.08	0.96	1.00	1.05	0.99
ARMA	0.98	0.99	1.32	1.21	0.97	0.99	1.16	1.08	0.96	1.00	1.05	0.99
ODFM	1.04	1.00	1.00	0.97	0.99	1.00	0.81	0.91	0.98	1.01	0.99	0.92
CDFM	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A-CD-TPN	1.02	1.03	0.96	1.00	1.01	0.98	0.91	0.98	0.99	1.00	0.99	0.94
A-LE-TPN	**0.99**	1.00	0.99	**0.98**	0.99	1.00	0.81	1.04	1.00	1.00	1.00	1.00
S-CD-TPN	1.01	1.01	0.92	1.04	1.01	1.00	0.88	1.00	1.00	1.00	0.99	1.00
S-LE-TPN	1.00	1.01	0.99	1.05	1.00	1.00	0.81	1.03	0.99	1.00	1.00	1.00
A-F-TPN	1.02	1.03	1.09	1.08	**0.98**	0.98	0.77	**0.96**	**0.97**	1.00	0.97	1.00
S-F-TPN	1.01	1.00	1.03	1.11	1.01	**0.97**	0.80	0.98	0.99	1.01	1.00	**0.93**
A-N-TPN	1.06	**0.99**	0.79	1.03	1.05	**0.97**	**0.68**	**0.96**	**0.97**	1.00	**0.93**	1.00
S-N-TPN	**0.99**	**0.99**	**0.77**	0.99	1.00	0.98	0.71	0.99	1.01	**0.98**	0.94	0.96

Note: The error measure is reported relative to the benchmark CDFM. Bold printed values indicate the lowest achieved RMSFE. The model acronyms correspond to the following specifications: AR: An AR(1) model; ARMA: An ARMA model with lag order determined recursively via the AIC; Naive: A naive unconditional mean model; ODFM: A DFM using only the relevant variables; CDFM: A DFM using all variables; A-CD-TPN (S-CD-TPN): The set of targeted predictors is retrieved via the EN, which is solved using coordinate descent, with EN parameters retrieved via CV. The data is aggregated (skip-sampled); A-LE-TPN (S-LE-TPN): The EN is solved via LARS-EN; A-N-TPN (S-N-TPN): A TP model where the set of targeted predictors is retrieved using the nowcasting error from the previous period; A-F-TPN (S-F-TPN): A TP model where the set of targeted predictors is retrieved via the EN, where the EN parameters are set so that 30 predictors are selected. The variables are newly selected for each nowcast.

Table 3: Relative root mean squared error of a quarterly GDP change targeted predictors nowcasting exercise for the US, GER, and BW data sets

Model	US	GER	BW
Naive	1.57	1.56	1.69
AR	2.76	1.78	1.99
ARMA	2.76	1.78	1.99
CDFM	1.00	1.00	1.00
A-CD-TPN	1.06**	2.16***	0.84
A-LE-TPN	1.31	0.85*	**0.79**
S-CD-TPN	**0.90****	**0.84***	0.89
S-LE-TPN	1.02		1.04
A-F-TPN	1.14***	1.24	1.01
S-F-TPN	1.06	1.04	0.87
A-N-TPN	**0.90****	1.55*	0.80*
S-N-TPN	0.98	**0.84**	0.87**

Note: The nowcasting period is 2012Q1–2019Q3 for all three data sets. The RMSFE is reported relative to the benchmark CDFM model. For all data sets it is chosen that $l_2 = 0.25$ and $k = 30$ in the case of fixed specifications. Bold printed values indicate the lowest RMSFE. The model acronyms correspond to the following specifications: AR: An AR(1) model; ARMA: An ARMA model with lag order determined recursively via the AIC; Naive: A naive mean model; CDFM: A DFM using all variables. A-CD-TPN (S-CD-TPN): The set of targeted predictors is retrieved via the EN, which is solved using coordinate descent, with EN parameters retrieved via CV. The data is aggregated (skip-sampled); A-LE-TPN (S-LE-TPN): The EN is solved via LARS-EN; A-N-TPN (S-N-TPN): A TP model where the set of targeted predictors is retrieved using the nowcasting error from the previous period; A-F-TPN (S-F-TPN): A TP model where the set of targeted predictors is retrieved via the EN, where the EN parameters are set by the researcher. The variables are newly selected for each nowcast.
We conduct Diebold-Mariano test with the CDFM as the benchmark forecasts. Statistical signficant differences are indicated by $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Figures

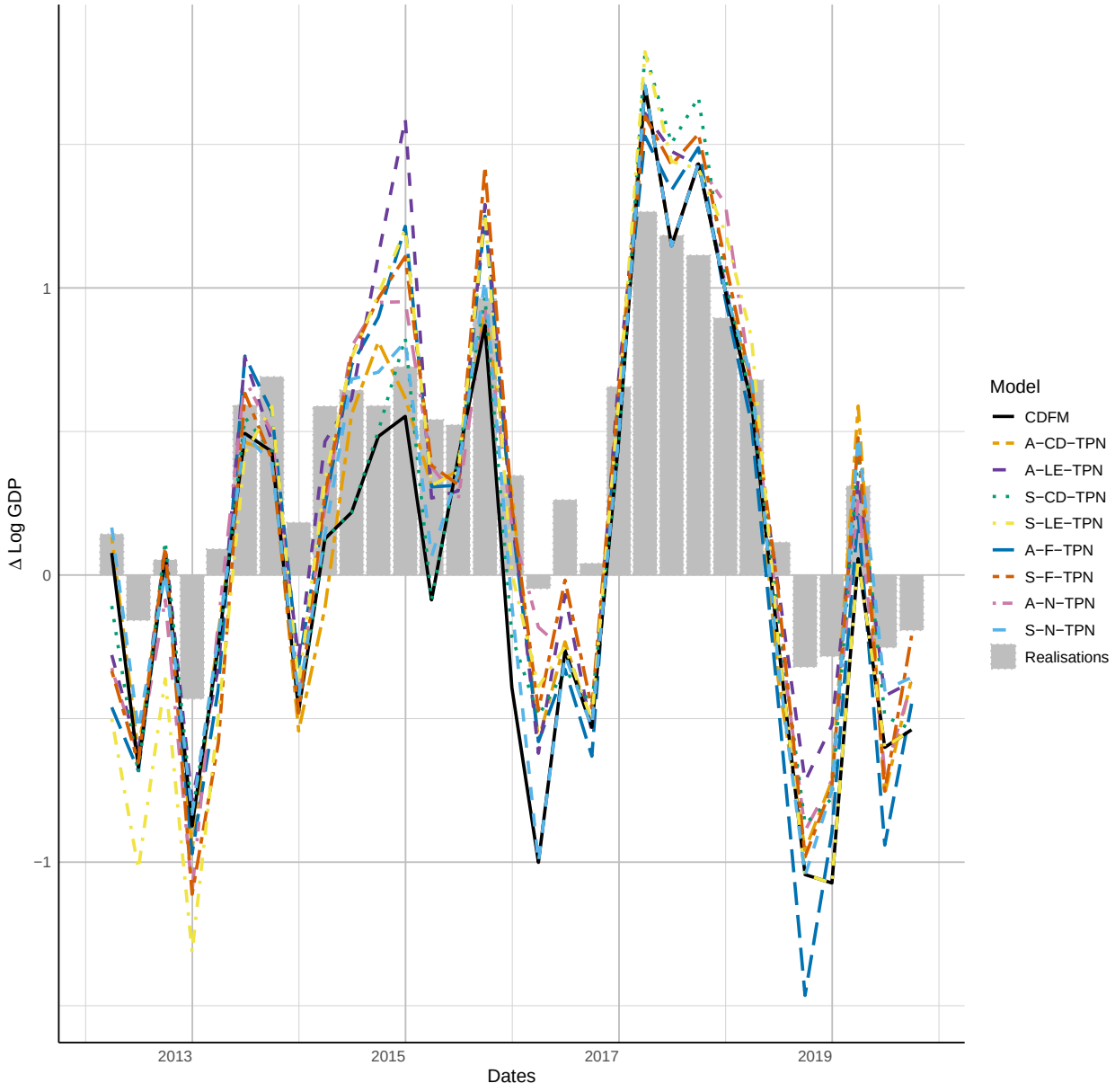

Figure 1: Plots of the model forecasts against the realizations for the BW dataset

Note: The nowcasting period is 2012Q1–2019Q3 for all three data sets. For all data sets it is chosen that $l_2 = 0.25$ and $k = 30$ in the case of fixed specifications. The model acronyms correspond to the following specifications: CDFM: A DFM using all variables. A-CD-TPN (S-CD-TPN): The set of targeted predictors is retrieved via the EN, which is solved using coordinate descent, with EN parameters retrieved via CV. The data is aggregated (skip-sampled); A-LE-TPN (S-LE-TPN): The EN is solved via LARS-EN; A-N-TPN (S-N-TPN): A TP model where the set of targeted predictors is retrieved using the nowcasting error from the previous period; A-F-TPN (S-F-TPN): A TP model where the set of targeted predictors is retrieved via the EN, where the EN parameters are set by the researcher. The variables are newly selected for each nowcast.

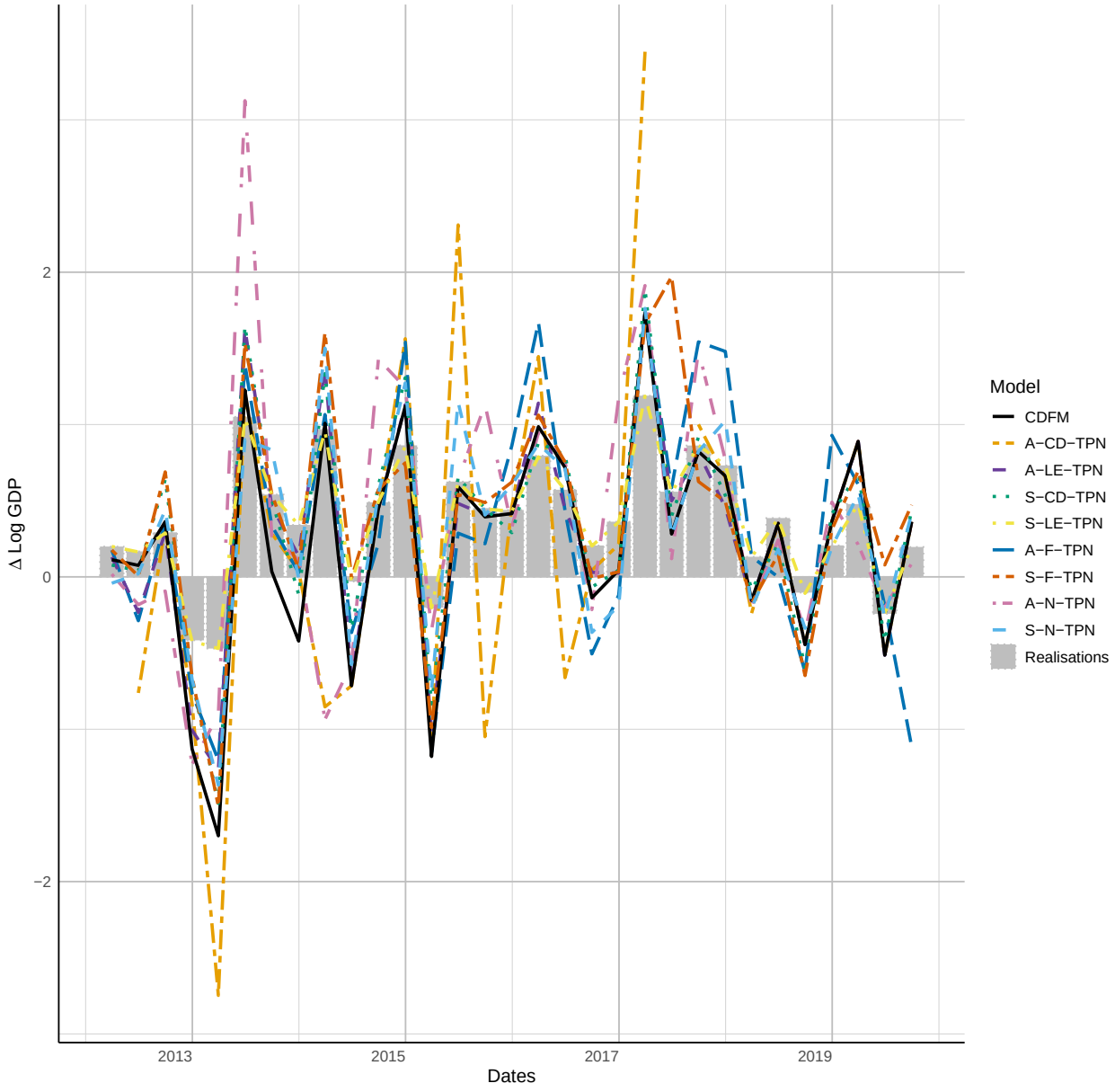

Figure 2: Plots of the model forecasts against the realizations for the GER dataset

Note: The nowcasting period is 2012Q1–2019Q3 for all three data sets. For all data sets it is chosen that $l_2 = 0.25$ and $k = 30$ in the case of fixed specifications. The model acronyms correspond to the following specifications: CDFM: A DFM using all variables. A-CD-TPN (S-CD-TPN): The set of targeted predictors is retrieved via the EN, which is solved using coordinate descent, with EN parameters retrieved via CV. The data is aggregated (skip-sampled); A-LE-TPN (S-LE-TPN): The EN is solved via LARS-EN; A-N-TPN (S-N-TPN): A TP model where the set of targeted predictors is retrieved using the nowcasting error from the previous period; A-F-TPN (S-F-TPN): A TP model where the set of targeted predictors is retrieved via the EN, where the EN parameters are set by the researcher. The variables are newly selected for each nowcast.

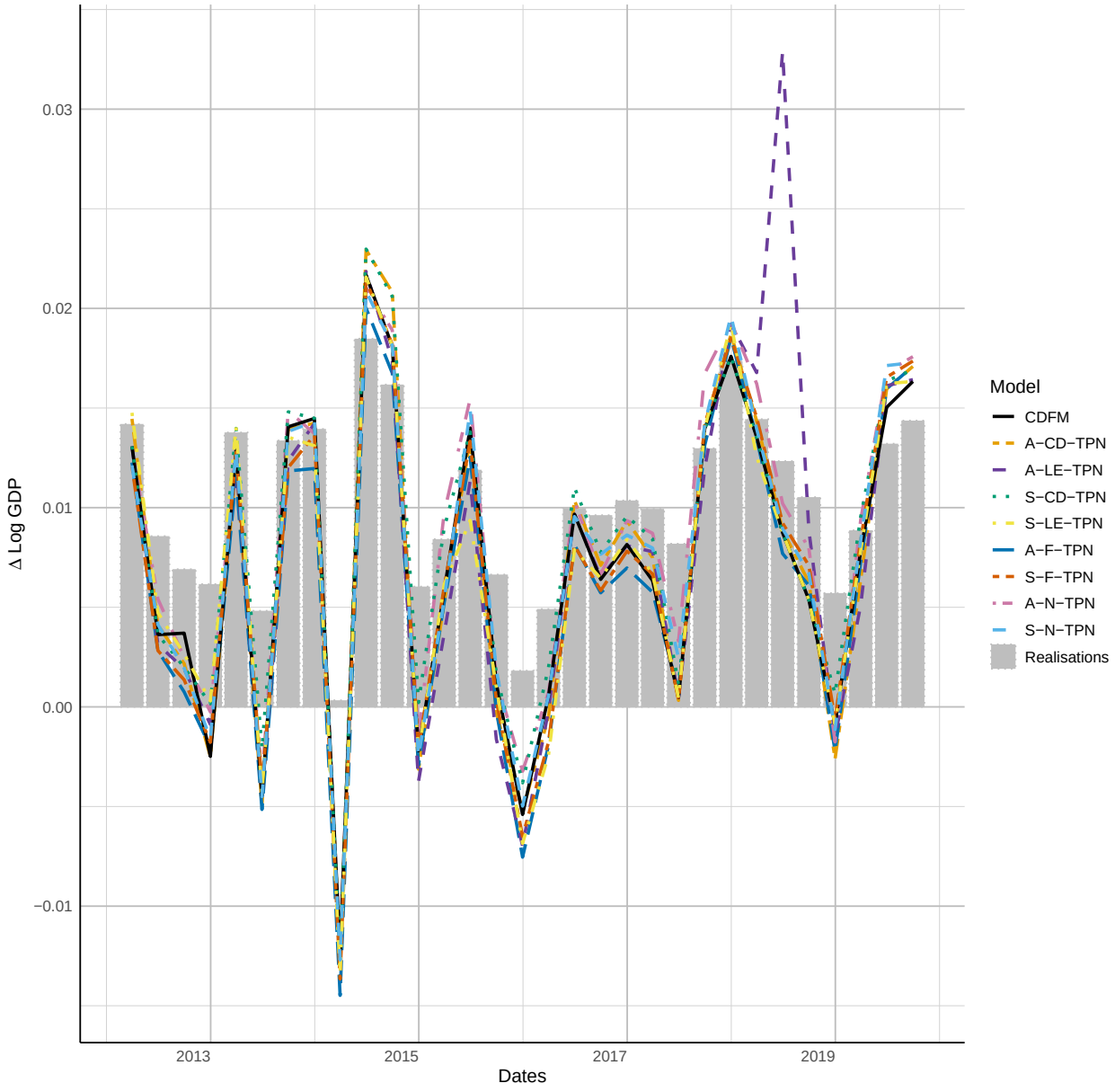

Figure 3: Plots of the model forecasts against the realizations for the US dataset

Note: The nowcasting period is 2012Q1–2019Q3 for all three data sets. For all data sets it is chosen that $l_2 = 0.25$ and $k = 30$ in the case of fixed specifications. The model acronyms correspond to the following specifications: CDFM: A DFM using all variables. A-CD-TPN (S-CD-TPN): The set of targeted predictors is retrieved via the EN, which is solved using coordinate descent, with EN parameters retrieved via CV. The data is aggregated (skip-sampled); A-LE-TPN (S-LE-TPN): The EN is solved via LARS-EN; A-N-TPN (S-N-TPN): A TP model where the set of targeted predictors is retrieved using the nowcasting error from the previous period; A-F-TPN (S-F-TPN): A TP model where the set of targeted predictors is retrieved via the EN, where the EN parameters are set by the researcher. The variables are newly selected for each nowcast.