Bank of Finland

BoF Economics Review

4 • 2020

# Nowcasting Finnish GDP growth using financial variables: a MIDAS approach

**Olli-Matti Laine**, Economist
**Annika Lindblad**, Economist

## Abstract

We analyse the performance of financial market variables in nowcasting Finnish quarterly GDP growth. Especially, we assess if prediction accuracy is affected by the sampling frequency of the financial variables. Therefore, we apply MIDAS models that allow us to forecast quarterly GDP growth using monthly or daily data without temporal aggregation in a parsimonious way. Our results show that financial market data nowcasts Finnish GDP growth relatively well. When it comes to individual variables, ratios like average price-to-earnings, average price-to-book or average dividend yield track GDP growth well. Our results suggest that the sampling frequency of financial market variables is not crucial: the forecasting accuracy of daily, monthly and quarterly data is similar.

**Keywords: MIDAS, Nowcasting, Financial markets, GDP**

**JEL codes: E44, G00, E37**

---

The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Bank of Finland.

We thank the referee Eleonora Granziera (Norges Bank) for useful and insightful comments on this research.

# 1. Introduction

Financial markets provide information about investors' expectations on a daily basis. This information may be useful in predicting or nowcasting GDP growth, because asset prices are based on expected future cash flows, which in turn are linked to macroeconomic conditions.

There are many previous studies that have shown that including financial market variables in short-term forecasting or nowcasting models is useful (e.g., Friedman and Kuttner, 1993; Estrella and Mishkin, 1998; Henry, Olekalns and Thong, 2004; Chionis, Gogas and Pragidis, 2010; Nyberg, 2010). Junttila and Korhonen (2011) show that dividend yields and short-term interest rates are relevant for forecasting output growth in multiple countries, including Finland. According to their results, financial variables are especially useful in turbulent times. Kuosmanen and Vataja (2014) find similar results. Although many studies have shown the usefulness of financial market information in forecasting, Alessi, Ghysels, Onorante, Peach and Potter (2014) argue that central banks have not utilised this information in the best possible way.

Even though there are many studies analysing the predictive power of financial market data, the optimal sampling frequency we should use when including financial market data in a model has not been widely studied. Should we care about the daily or monthly movements, or is it better to look at more temporally aggregated data? Typically, if the variables in a model are measured at different frequencies, the high frequency variables are aggregated to the level of the lowest frequency variable.

In this paper we study the usefulness of including financial variables in a nowcasting model for the Finnish quarter-on-quarter GDP growth rate and whether the choice of frequency matters for nowcasting. The analysis is conducted in a mixed data sampling (MIDAS) regression framework (see, for example, Ghysels, Santa-Clara and Valkanov, 2004; Ghysels, Sinko, and Valkanov, 2007, for details on the MIDAS framework). MIDAS models allow us to study how time aggregation of variables affects prediction accuracy. At least, Andreou, Ghysels and Kourtellos (2013) and Ferrara, Marsilli and Ortega (2014) have shown that including daily financial variables in a MIDAS model improves forecasts, but they do not assess if it matters whether one uses daily or monthly data. We also assess different ways to utilise financial variables in the MIDAS framework. The earlier literature has considered, for example, individual variables, forecast combinations and principal components (Andreou et al., 2013; Ferrara et al., 2014).

Our results suggest that it does not matter significantly and systematically for nowcast accuracy whether one uses financial market data at the daily, monthly or quarterly frequency when nowcasting Finnish GDP growth. However, there may be some practical reasons to prefer higher frequency data, such as the ability to update the nowcast on a monthly, weekly or even daily frequency. On the other hand, increasing the frequency also brings some challenges, in particular by increasing the noisiness of the data, which might complicate the analysis. Ultimately, our results suggest that the choice of frequency can be made based on data availability and the needs of the forecaster, without significantly compromising nowcast accuracy.

To gauge the importance of financial data for nowcasting GDP we compare their forecasting ability to that of industrial production growth, which is a traditional and good predictor of GDP growth. Our results imply that financial variables predict GDP growth as accurately as industrial production. Because industrial production is observed with more than a one month lag, the results suggest that we can, without loss of accuracy, nowcast GDP earlier using financial variables. Different kinds of financial ratios – like the dividend yield – turn out to be especially good at nowcasting Finnish GDP growth, which is in line with the results in Junttila and Korhonen (2011). The results also provide some evidence that forecast accuracy can be improved by combining a financial market based forecast to forecasts based on macroeconomic variables. However, these improvements are not statistically significant.

The rest of the article is organised as follows. Section 2 introduces the MIDAS regression framework. Section 3 summarises the data. Section 4 analyses the nowcasting performance of the different financial variables and analyses the effects of increasing the frequency of the financial data. Section 5 considers the best way of utilising financial data to nowcast Finnish GDP growth. Section 6 concludes.

## 2. MIDAS

MIDAS models were introduced by Ghysels et al. (2004), Ghysels et al. (2005), Ghysels et al. (2006) and Ghysels et al. (2007). The central idea of the MIDAS approach is to explain a low frequency variable by variables measured at a higher frequency, without aggregation and in a parsimonious way.

The standard MIDAS model with one explanatory variable can be written as follows:

(1) $y_t = \beta_0 + \beta_1 \sum_{h=0}^{d} \theta_h \, x_{tm-h} + u_t,$

where $y_t$ is a low frequency variable (GDP growth in our models), $x_{tm-h}$ is a high frequency variable (financial market variable or industrial production in our models) and $d$ is the number of lags of the explanatory variable included in the model.[1] We get $m$ observations from the high frequency variable as we get one observation from the low frequency variable. For example, if we explained a quarterly variable by a monthly variable, $m$ would be 3. If the number of lags, $d$, was 2, then we would explain the quarterly variable by all the monthly observations of the high frequency variable from the given quarter. If $\beta_1 \neq 0$, there is a connection between the low and the high frequency variables. Function $\theta_h$ is a polynomial that weights the contemporaneous observation of the high frequency variable and its lags in a parsimonious way. In this paper we use the (normalised) exponential Almon lag polynomial:

(2) $\theta_h = \frac{e^{\lambda_1 (h+1) + \lambda_2 (h+1)^2}}{\sum_{s=0}^{d} e^{\lambda_1 (s+1) + \lambda_2 (s+1)^2}}.$

The normalisation ensures the weights sum up to one, which allows the separate identification of $\beta_1$. The exponential Almon lag polynomial allows flexible, such as decaying or hump-shaped, weighting schemes. Parameters $\lambda_1$ and $\lambda_2$ are estimated simultaneously with the other model parameters, and together with the number of lags govern the shape of the weighting scheme. If $\lambda_1 = \lambda_2 = 0$ the lags have equal weights $(1/d)$ and we essentially include a moving average of the past $d$ lags in the MIDAS model. The MIDAS model is estimated using non-linear least squares (NLS).

Due to the lag polynomial MIDAS models are especially useful when the number of lags is large, as it allows including a large number of, for example, daily lags without increasing the parameter space. However, when only a few lags are included a so called unrestricted MIDAS (U-MIDAS) model, which does not include a weighting scheme and thus estimates a separate regression parameter for each lag, can be used (Foroni, Marcellino and Schumacher, 2015). The U-MIDAS regression model can be written as:

(3) $y_t = \beta_0 + \sum_{h=0}^{d} \beta_h \, x_{tm-h} + u_t.$

In this case the model can be estimated using OLS.

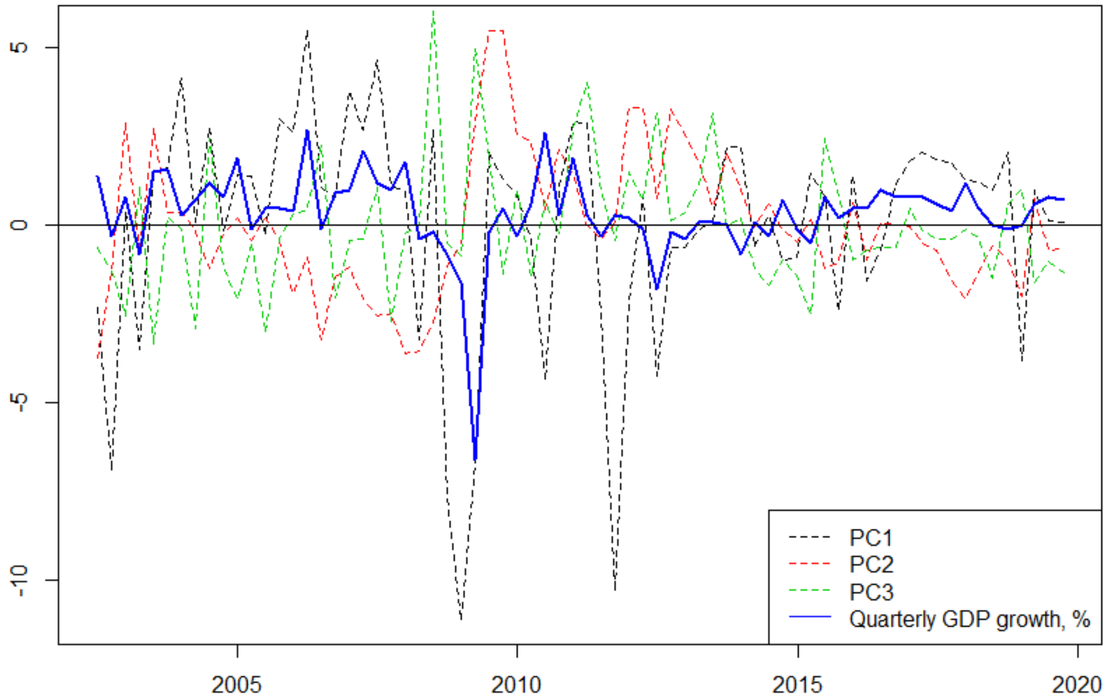
## 3. Data

We use daily, monthly and quarterly data from Q2/2002 to Q3/2019. The Finnish GDP data is from Statistics Finland. We use the latest vintage of the quarterly real GDP growth rate.[2] When it comes to financial variables, we assess the predictive power of stock indices and interest rates from Finland, Germany and the USA. In addition, we use the average

*[1] The AR-MIDAS model by Andreou et al. (2013), which includes lagged values of the dependent variable, could be used for determining whether financial variables include useful information for nowcasting GDP in addition to lagged GDP. However, in this paper we concentrate on evaluating the usefulness of different financial variables and the impact of varying the frequency at which the variables are sampled in a MIDAS model for GDP. Thus, exploring the AR-MIDAS model in this context is left for future work.*

*[2] GDP statistics are typically revised substantially. We use the latest available vintage at the time of writing (data collected on December 31st, 2019) as we consider it the best available estimate for the actual, final growth rate. However, we recognize that there are also arguments for using an earlier vintage, such as first release data.*

**Figure 1. This figure shows the development of the first three principal components together with quarterly GDP growth**




*The principal components are calculated from the data summarised in Appendix A (excluding industrial production). The data are from Q2/2002 to Q3/2019 and the frequency is quarterly.*

price-to-earnings ratios, price-to-book ratios and dividend yields from the same countries. The predictive power of the oil price, expected stock market volatility implied by Eurostoxx 50 index options and the EUR/USD exchange rate are also considered. All the data are obtained via Bloomberg. Stock indices and the oil price are in log-differences. Interest rates and the exchange rate are in differences. We also calculate the spread between the German 10-year yield and the 12-month yield. Financial ratios, stock market volatility and the interest rate spread are in levels. For comparison, we nowcast GDP growth also using the monthly growth rate of industrial production, for which the latest vintage is used. The transformations are done to achieve stationarity and are based on previous studies (for example, Bernanke, Boivin and Eliasz, 2005; Marcellino and Schumacher, 2010). All the employed time series are listed (together with their sources and transformations) and plotted in Appendix A.

To summarise the information in the financial data we use principal component analysis to extract common factors from the explanatory data. Figure 1 plots the quarterly GDP growth rate together with the first three principal components (PCs) based on standardised data (excluding industrial production). The principal components are here calculated from quarterly series for visual reasons. The principal components capture how the financial markets as a whole have developed over time. The figure shows that there is a connection between financial markets and real GDP growth. However, the relationship seems to be time-varying and at its strongest during turbulent times.[3] Especially the first PC seems to capture well some of the (negative) spikes in GDP growth. Corresponding figures for the individual financial variables can be found in Appendix A.

[3] *The first PC has a correlation of 0.48 with GDP growth, the second PC a correlation of -0.24, and the third PC a correlation of -0.35.*

## 4. Nowcasting GDP growth using financial data

Next, we assess how well different variables sampled at different frequencies nowcast GDP growth. We conduct a rolling window analysis, where the first estimation sample is from Q2/2002 to Q4/2011 and the first out-of-sample period is Q1/2012. We have chosen 2011 to be the last year of the first estimation sample, in order to see how well the models predict the sharp decline in Finnish quarterly GDP in Q2 2012.[4] Altogether we produce 31 pseudo out-of-sample forecasts, which is a relatively small number of out-of-sample observations.

In the models using daily data we include 62 daily lags. In our sample, every quarter includes at least 63 working days. By choosing 62 daily lags we avoid technical issues relating to “missing data” in the code.[5] In the models with monthly data the number of lags is two and in quarterly models zero, which means that only the contemporaneous quarter is included. The number of lags in the monthly and quarterly models is chosen such that they are comparable to the model using daily data. All the models therefore utilise data from one quarter only. We consider including more lags in Section 5.

The models using quarterly data are naturally unrestricted as there are no lags. For the monthly models we consider both the restricted and the unrestricted versions as the number of lags is small. The daily models are restricted, as it is infeasible to estimate 62 individual regression parameters. The parameters are always re-estimated in the rolling window analysis. Appendix B shows the estimation results that are obtained using the whole sample from Q2/2002 to Q3/2019.

Table 1 reports the average root-mean-square errors (RMSEs) from the rolling window analysis. The predictive variables are ordered based on the RMSEs of the quarterly models. The average price-to-earnings ratio in the United States has, perhaps surprisingly, been the best predictor for Finnish GDP growth, regardless of the frequency used. The price-to-earnings and price-to-book ratios especially for the Finnish and the US stock market have also performed well. The dividend yields for the Finnish and German stock markets produce accurate forecasts, while implied volatility, the EUR/USD exchange rate, and the stock indices tend to perform worse. However, some of the results are sensitive to the frequency used for the explanatory variables. For example, industrial production performs relatively better on the monthly frequency, while the SPX, HEX and DAX indices seem to improve as predictors as the sampling frequency increases. The relative performance of the various interest rates also seems to vary depending on the frequency used.

In addition to the individual financial variables we also include the first principal component based on the financial data, which also produces accurate nowcasts, indicating that combining information from a large set of financial indicators provides useful information for nowcasting Finnish GDP growth. If industrial production growth is taken as a benchmark predictor, we can see that on the monthly frequency at least the average price-to-earnings ratio in the United States, the German dividend yield and the PC produce more accurate nowcasts. Regardless of the frequency some financial variables thus seem useful for nowcasting Finnish GDP growth, in particular considering that they are available much earlier than industrial production.

The results are overall in line with previous results. For example, the dividend yields perform well, as concluded by Junttila and Korhonen (2011). As a small open economy Finland is significantly affected by global fluctuations, which might explain why foreign variables, such as the average price-to-earnings ratio of the S&P 500, nowcast well.

All in all, using a higher frequency does not seem to clearly and consistently improve forecasts. The differences in the RMSEs also tend to be small, and whether performance improves or deteriorates as the sampling frequency increases depends on the variable. The average RMSE of the quarterly models is 0.68, of the restricted monthly models 0.70, of the unrestricted monthly models 0.76 and of the daily models 0.68. Interestingly, despite the small number of monthly lags the MIDAS model utilising a weighting scheme performs generally better than the unrestricted model.

---

[4] *Appendix E reports the results excluding the out-of-sample forecast for the period Q2/2012.*

[5] *We use the R package midasr by Virmantas Kvedaras and Vaidotas Zemlys-Balevicius to estimate the models (see also Ghysels, Kvedaras and Zemlys, 2016).*

**Table 1: Out-of-sample RMSEs of MIDAS regression models**

| Explanatory variable | Quarterly | Monthly | Monthly (unrestricted) | Daily |
|---|---|---|---|---|
| SPX_pe | 0.52 | 0.55 | 0.56 | 0.55 |
| DAX Dividend yield | 0.58 | 0.58 | 0.60 | 0.58 |
| HEX Dividend yield | 0.60 | 0.66 | 0.66 | 0.61 |
| HEX_pb | 0.60 | 0.63 | 0.72 | 0.64 |
| PC1 | 0.63 | 0.61 | 0.61 | 0.64 |
| SPX_pb | 0.64 | 0.66 | 0.76 | 0.67 |
| HEX_pe | 0.65 | 0.69 | 0.64 | 0.68 |
| FI_10y | 0.66 | 0.85 | 0.95 | 0.67 |
| OIL | 0.66 | 0.66 | 0.70 | 0.67 |
| DE 10y-1y | 0.66 | 0.66 | 0.65 | 0.66 |
| DE_10y | 0.67 | 0.77 | 0.91 | 0.66 |
| DE_5y | 0.67 | 0.72 | 0.72 | 0.68 |
| DE_7y | 0.67 | 0.77 | 0.83 | 0.68 |
| EURUSD | 0.67 | 0.77 | 0.78 | 0.67 |
| DAX_pe | 0.68 | 0.67 | 0.85 | 0.68 |
| Industrial production | 0.68 | 0.62 | 0.66 | - |
| OMX Hels Telec | 0.69 | 0.73 | 0.77 | 0.83 |
| DE_1y | 0.70 | 0.64 | 0.71 | 0.65 |
| FI_5y | 0.70 | 0.76 | 0.79 | 0.69 |
| OMX Hels Industrials | 0.71 | 0.73 | 0.75 | 0.70 |
| OMX Hels Utilities | 0.71 | 0.66 | 0.71 | 0.77 |
| SPX Dividend yield | 0.71 | 0.70 | 0.73 | 0.70 |
| DAX_pb | 0.72 | 0.67 | 0.80 | 0.78 |
| OMX Hels Hlth Care | 0.72 | 0.81 | 0.99 | 0.67 |
| OMX Hels Technology | 0.72 | 0.70 | 0.85 | 0.75 |
| SPX | 0.74 | 0.77 | 0.72 | 0.64 |
| HEX | 0.75 | 0.71 | 0.79 | 0.67 |
| Eurostoxx 50 volatility | 0.77 | 0.74 | 0.78 | 0.77 |
| DAX | 0.84 | 0.78 | 0.84 | 0.65 |
| OMX Hels Basic Metal | 0.84 | 0.82 | 0.87 | 0.72 |

*The models are ordered based on the RMSEs of the quarterly models. The abbreviations for the variables are explained in Appendix A. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q2/2002 to Q4/2011 and the first out-of-sample forecast is Q1/2012. Thus, the results are based on 31 out-of-sample periods.*

## 5. Do financial variables improve GDP predictions?

The results in Section 4 showed that financial variables can be useful for predicting Finnish GDP growth. In this section we aim at finding the best nowcasting specification when using financial market variables in a MIDAS framework. Thus, we consider how increasing the lag length in the MIDAS models impacts forecasting performance as well as consider combining the information in financial variables by simple forecast averaging, i.e., we take an average of the MIDAS model nowcasts from Section 4. However, as some of these individual models produced consistently inferior nowcasts, we also combine only the nowcasts produced by MIDAS models driven by financial ratios (price-to-earnings, price-to-book, dividend yield), which turned out to be the most accurate class of predictors in Section 4. As benchmark models we consider the MIDAS model based on monthly growth in industrial production as well as a simple AR(1) model for GDP growth.

We use monthly data as a compromise between daily and quarterly data because our previous results did not suggest a preference for any specific frequency. Monthly data provides us with relatively timely information, allowing the nowcast to be updated within the quarter, while being less noisy than daily data. In addition, technical issues limited our use of daily data to within-quarter data. We now relax this assumption using monthly data and consider models that have between 2 and 11 monthly lags.

Table 2 reports the RMSEs of these models. According to the results, the average forecast based on only financial ratios produces the most accurate nowcasts during the out-of-sample period from Q1/2012 to Q3/2019 regardless of the number of lags included.[6] Summarising the information by principal component analysis lead to slightly lower RMSEs than combining all the MIDAS model forecasts, and the PC based forecast thus provides a viable alternative if one does not wish to preselect the models to be combined. However, none of the three purely financial market based forecasts in Table 2 outperforms the forecast based on industrial production in a statistically significant way. In this sample, all the models nowcast GDP roughly as accurately as an AR(1) model (RMSE 0.58).[7]

Regarding the choice of lag length, for industrial production growth including only two lags leads to inferior nowcasting performance, but for the other models forecasting performance is similar regardless of the lag length chosen. A longer lag length can be generally preferred based on the idea that the exponential Almon polynomial should be able put the weight of any excess lags to zero, and therefore it can be argued that including too many lags is less detrimental than including too few lags. By comparing the weighting schemes plotted in Appendices B, C and D we can clearly see that two monthly lags tends to be too few for the weights to decay to zero, whereas for many variables, such as the first PC, eleven lags seems to be the most suitable lag length. On the other hand, it could be argued that the importance of financial data lies in it being able to quickly reflect the current economic situation and changes in it, which would indicate that putting at least most of the weight on near-term data is desirable. This would advocate including only two lags, as forecasting performance is similar regardless of the number of included lags.

Financial market data and real variables might include different types of information, useful for forecasting GDP during different time periods, depending, for example, on the origin of the downturn or upturn. Therefore, it could be useful to combine the forecasts produced by the MIDAS model based on industrial production growth and the MIDAS models based on financial ratios. The resulting nowcast results in the lowest RMSEs when using at least 5 lags, while the difference is weakly statistically significant when using two lags.

Finally, in Figure 2 we compare the MIDAS model nowcasts to realised GDP growth, in order to assess whether there are clear differences in model performance over time. It is clear from the figure that none of the models capture the sharp decline in GDP in 2012 very

---

[6] *All the weight functions are plotted in Appendices B, C and D.*

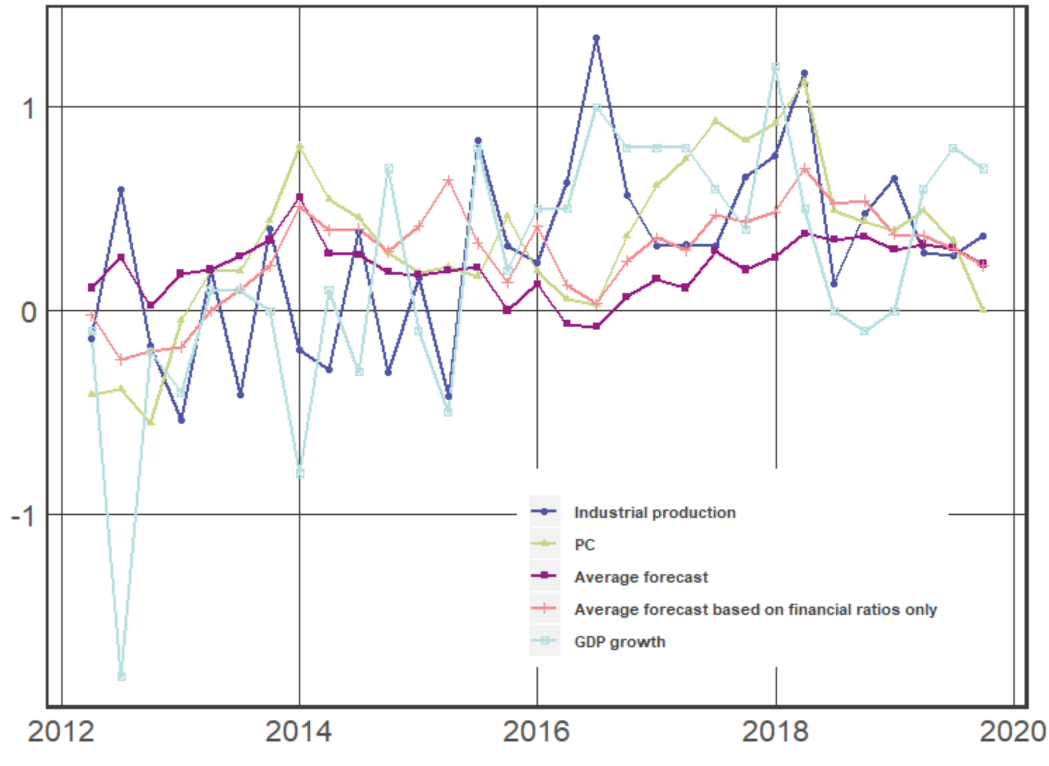
[7] *It should be noted that, first of all, the latest vintage of the GDP data is used, i.e., the data includes revisions. Thus, the performance of the AR(1) model might be better than it would have been in real time. Secondly, GDP is released with a roughly two month lag in Finland, meaning it is unavailable in real-time.*

**Table 2: RMSEs of MIDAS regression models**

| | 2 lags | 5 lags | 11 lags |
|---|---|---|---|
| Industrial production | 0.71 | 0.57 | 0.60 |
| PC | 0.60 | 0.62 | 0.59 |
| Average forecast | 0.63 | 0.66 | 0.65 |
| Average forecast: financial ratios only | 0.55 | 0.56 | 0.58 |
| Average forecast: financial ratios only and industrial production | 0.60* | 0.51 | 0.54 |

*In the model using industrial production the only explanatory variable is the MoM growth rate of the volume of industrial production. In the PC model, the only explanatory variable is the first principal component of the financial market variables (see Appendix A). 'Average forecast' is the simple average of all the financial variable based forecasts (forecasts produced using the financial market variables listed in Appendix A one at a time). 'Average forecast based on financial ratios only' is the average of the models in which the explanatory variable is the price-to-earnings ratio, the price-to-book ratio or the dividend yield. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q1/2003 to Q4/2011 and the first out-of-sample forecast is Q1/2012. Thus, the results are based on 31 out-of-sample periods. To test the statistical significance of the RMSE differences between the industrial production based forecast and the other forecasts, we use the Diebold-Mariano test assuming no heteroscedasticity or autocorrelation because the forecast horizon is zero. Asterisks *,** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.*

**Figure 2: Out-of-sample nowcasts of the different models (number of lags 11) and GDP growth**




*The nowcasts are calculated from a rolling window analysis, in which the first estimation sample is from Q2/2002 to Q4/2011 (GDP data from Q1/2003).*

well, or the recent slowdown in growth in 2018. Especially in the second quarter of 2012 all the models produce very large forecast errors. The model driven by industrial production growth performs particularly badly while financial variables were relatively more useful for nowcasting GDP. However, especially due to the limited number of observations available for evaluation we also consider the root mean square forecast errors excluding the forecast error for Q2/2012 in Appendix F. The results confirm that in particular the RMSE of the industrial production driven model improves by the exclusion of Q2/2012, but the average forecast combining financial ratios and industrial production growth still outperforms this forecast in a weakly statistically significant way when two lags are used. Industrial production growth nowcasts GDP growth particularly well around 2015-2016, and also seems more volatile than the nowcasts based on financial variables. As expected, especially the two combination forecasts are relatively stable, and thus fail to account for the strong variation in quarterly GDP growth.

Overall, financial market variables and more traditionally used predictors, here represented by industrial production, seem to nowcast Finnish GDP approximately equally well. Therefore, the usefulness of financial variables for nowcasting GDP arises from the fact that they are available earlier than real variables. This issue is highlighted by the fact that the analysis above is conducted using the final, revised vintage of industrial production data, which could give an advantage to industrial production as a predictor of GDP. This favours using financial variables, which are available immediately and are unrevised, when nowcasting quarterly GDP growth.

# 6. Conclusion

We conclude that financial market variables are useful for nowcasting Finnish quarterly GDP growth. This is especially true for the price-to-earnings ratios, price-to-book ratios and dividend yields. Forecast combinations of these variables nowcast GDP relatively accurately. The main advantage of financial market data is its immediate availability. Unlike many other variables that have been traditionally used for forecasting or nowcasting, financial market based forecasts may be updated on a daily basis. This is very useful, for example, at the time of crisis, when financial markets might react immediately, but many sentiment indicators, let alone real variables, are published with a delay. Using financial variables in combination with macroeconomic variables when nowcasting GDP thus seems beneficial.

On the other hand, our results show that one cannot improve nowcast accuracy by increasing the frequency from quarterly to monthly or daily. This indicates that short-term fluctuations of financial markets do not contain additional useful information for nowcasting GDP, which is not already included in corresponding lower frequency data.

## References

Alessi, L., Ghysels, E., Onorante, L., Peach, R., & Potter, S. (2014). Central bank macroeconomic forecasting during the global financial crisis: The European Central Bank and Federal Reserve Bank of New York experiences. *Journal of Business & Economic Statistics 32(4)*, 483-500.

Andreou, E., Ghysels, E., & Kourtellos, A. (2013). Should macroeconomic forecasters use daily financial data and how? Journal of Business & Economic Statistics, 31(2), 240-251.

Bernanke, B. S., Boivin, J., & Eliasz, P. (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics 120(1)*, 387-422.

Chionis, D., Gogas, P., & Pragidis, I. (2010). Predicting European Union recessions in the euro era: The yield curve as a forecasting tool of economic activity. *International Advances in Economic Research 16(1)*, 1-10.

Estrella, A., & Mishkin, F. S. (1998). Predicting US recessions: Financial variables as leading indicators. *Review of Economics and Statistics 80(1)*, 45-61.

Ferrara, L., Marsilli, C., & Ortega, J. P. (2014). Forecasting growth during the Great Recession: is financial volatility the missing ingredient? Economic Modelling, 36, 44-50.

Foroni, C., Marcellino, M., & Schumacher, C. (2015). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 178(1)*, 57-82.

Friedman, B. M., & Kuttner, K. (1993). Why does the paper-bill spread predict real economic activity? In Business cycles, indicators and forecasting (pp. 213-254). University of Chicago Press.

Ghysels, E., Kvedaras, V., & Zemlys, V. (2016). Mixed Frequency Data Sampling Regression Models: The R Package midasr. *Journal of Statistical Software*, *72(4)*, 1 - 35. doi:http://dx.doi.org/10.18637/jss.v072.i04

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). The MIDAS touch: Mixed data sampling regression models. Discussion paper, UCLA.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics (76)*, 509-548.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics 131(1-2)*, 59-95.

Ghysels, E., Sinko, A., & Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews, 26(1)*, 53-90.

Henry, O. T., Olekalns, N., & Thong, J. (2004). Do stock market returns predict changes to output? Evidence from a nonlinear panel data model. Empirical Economics, 29(3), 527-540.

Junttila, J., & Korhonen, M. (2011). Utilizing financial market information in forecasting real growth, inflation and real exchange rate. *International Review of Economics & Finance 20(2)*, 281-301.

Kuosmanen, P., & Vataja, J. (2014). Forecasting GDP growth with financial market data in Finland: Revisiting stylized facts in a small open economy during the financial crisis. *Review of Financial Economics 23(2)* 90-97.

Marcellino, M., & Schumacher, C. (2010). Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP. *Oxford Bulletin of Economics and Statistics 72(4)*, 518-550.

Nyberg, H. (2010). Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting 29(1-2)*, 215-230.

# Appendix

## A

The following table lists all the used variables, their abbreviations, sources and transformations. All the data are obtained via Bloomberg.

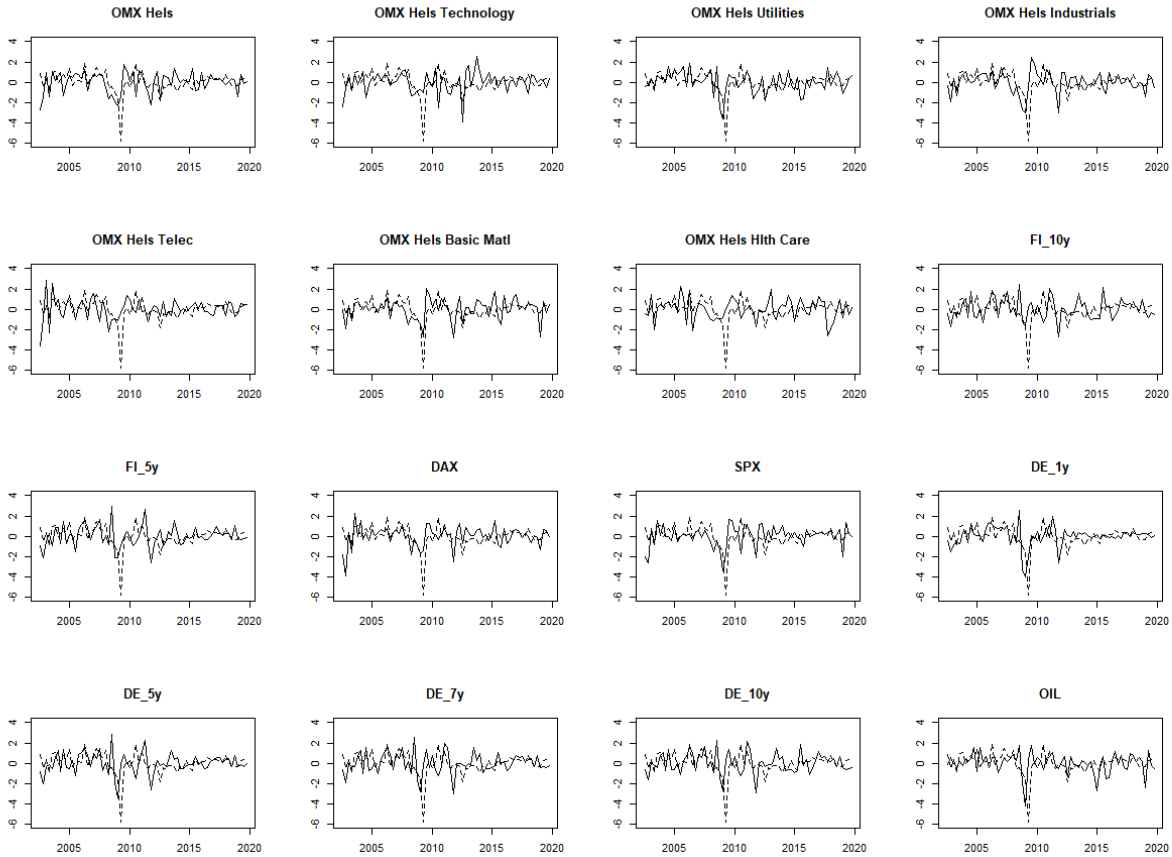
**Table A 1**

| Variable | Category | Abbreviation | Source | Transformation |
|---|---|---|---|---|
| Average ratio of the price of a stock and the company's earnings per share in OMX Helsinki | Financial ratio | HEX_pe | NASDAQ OMX Helsinki | No transformation |
| Average ratio of the stock price to the book value per share in OMX Helsinki. | Financial ratio | HEX_pb | NASDAQ OMX Helsinki | No transformation |
| Average dividend yield in OMX Helsinki | Financial ratio | HEX Dividend yield | NASDAQ OMX Helsinki | No transformation |
| Average ratio of the stock price to the book value per share in DAX | Financial ratio | DAX_pe | Deutsche Börse | No transformation |
| Average ratio of the stock price to the book value per share in DAX. | Financial ratio | DAX_pb | Deutsche Börse | No transformation |
| Average dividend yield in DAX | Financial ratio | DAX Dividend yield | Deutsche Börse | No transformation |
| Average ratio of the stock price to the book value per share in S&P 500 | Financial ratio | SPX_pe | Standard and Poor's | No transformation |
| Average ratio of the stock price to the book value per share in S&P 500 | Financial ratio | SPX_pb | Standard and Poor's | No transformation |
| Average dividend yield in S&P 500 | Financial ratio | SPX Dividend yield | Standard and Poor's | No transformation |
| The yield of Finland government bond with maturity of 10 years | Interest rate | FI_10y | Bloomberg | Difference |
| The yield of Finland government bond with maturity of 5 years | Interest rate | FI_5y | Bloomberg | Difference |
| The yield of Germany government bond with maturity of 12 months | Interest rate | DE_1y | Bloomberg | Difference |
| The yield of Germany government bond with maturity of 5 years | Interest rate | DE_5y | Bloomberg | Difference |
| The yield of Germany government bond with maturity of 7 years | Interest rate | DE_7y | Bloomberg | Difference |
| The yield of Germany government bond with maturity of 10 years | Interest rate | DE_10y | Bloomberg | Difference |
| The spread between German 10 year yield and 12 month yield | Interest rate | DE 10y-1y | Bloomberg | No transformation |
| Implied volatility on Eurostoxx 50 index options with a rolling 30 day expiry | Other | Eurostoxx 50 volatility | Deutsche Börse | No transformation |
| The price of oil (brent) | Other | OIL | Deutsche Börse | Log-difference |
| The price of euro in dollars | Other | EURUSD | Deutsche Börse | Difference |

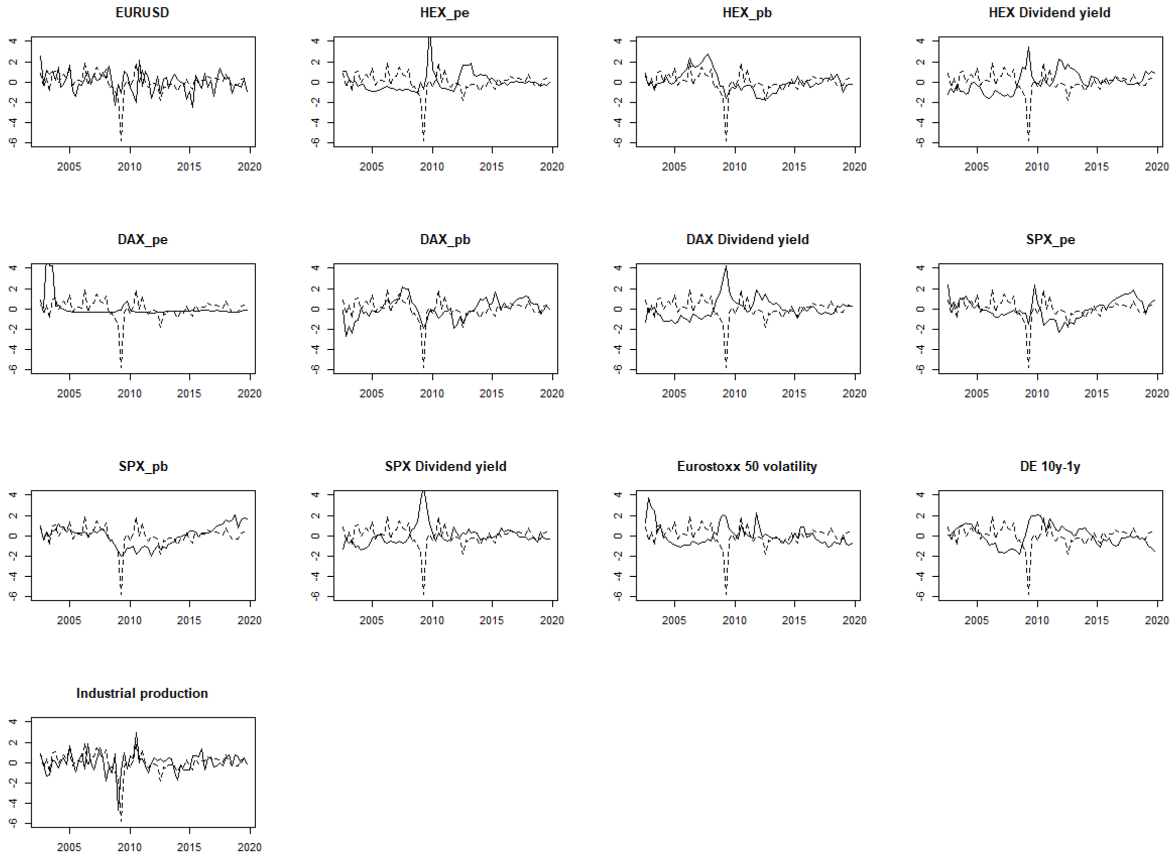
| | | | | |
|---|---|---|---|---|
| Finland Industrial Production Volume, MoM growth rate, SA | Real economy | Industrial production | Statistics Finland | No transformation |
| OMX Helsinki, price index | Stock index | OMX Hels | NASDAQ OMX Helsinki | Log-difference |
| OMX Helsinki Technology, price index | Stock index | OMX Hels Technology | NASDAQ OMX Helsinki | Log-difference |
| OMX Helsinki Utilities, price index | Stock index | OMX Hels Utilities | NASDAQ OMX Helsinki | Log-difference |
| OMX Hels Industrials, price index | Stock index | OMX Hels Industrials | NASDAQ OMX Helsinki | Log-difference |
| OMX Hels Telecommunication, price index | Stock index | OMX Hels Telec | NASDAQ OMX Helsinki | Log-difference |
| OMX Helsinki Basic Materials, preice index | Stock index | OMX Hels Basic Matl | NASDAQ OMX Helsinki | Log-difference |
| OMX Helsinki Health Care, price index | Stock index | OMX Hels Hlth Care | NASDAQ OMX Helsinki | Log-difference |
| S&P 500, price index | Stock index | SPX | Standard and Poor's | Log-difference |
| DAX, price index | Stock index | DAX | Deutsche Börse | Log-difference |

The following figures plot all the (transformed) variables (solid lines) in quarterly frequency together with quarterly GDP growth (dashed lines). All the variables have been demeaned and divided by their standard deviations to make interpretation easier.

### Figure A 1



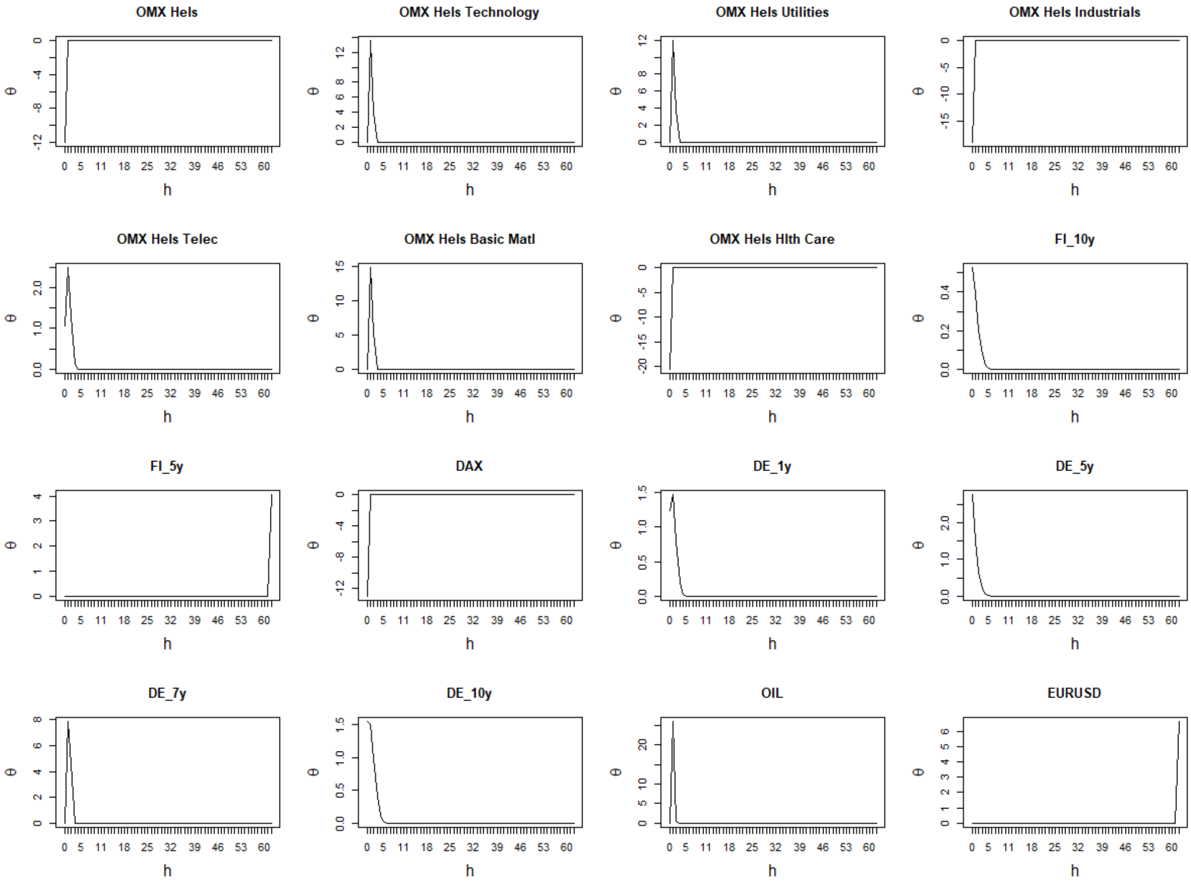
## Figure A 2


EURUSD
HEX_pe
HEX_pb
HEX Dividend yield
DAX_pe
DAX_pb
DAX Dividend yield
SPX_pe
SPX_pb
SPX Dividend yield
Eurostoxx 50 volatility
DE 10y-1y
Industrial production
-6 -4 -2 0 2 4
2005 2010 2015 2020

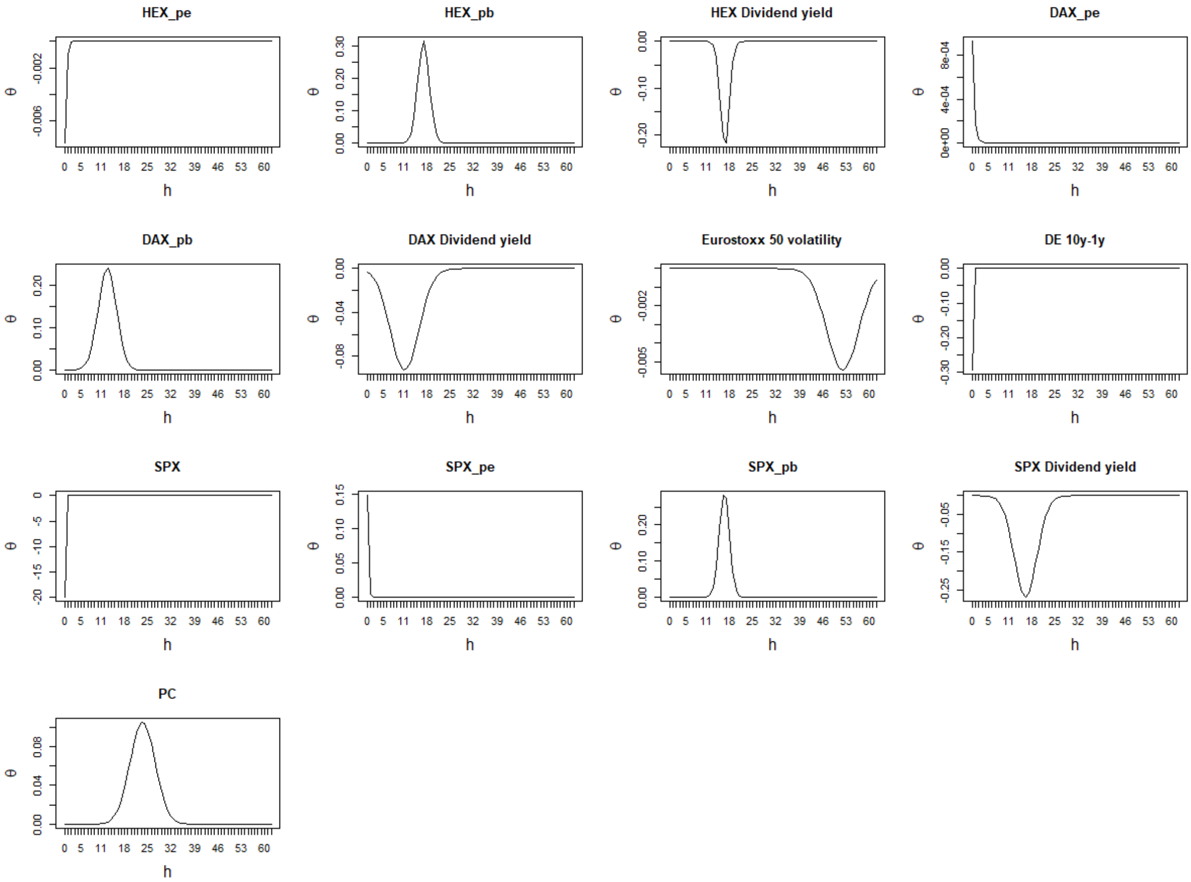
## B

The following figures show the weighting schemes that are estimated using daily data from Q2/2002 to Q3/2019.

**Figure B 1**

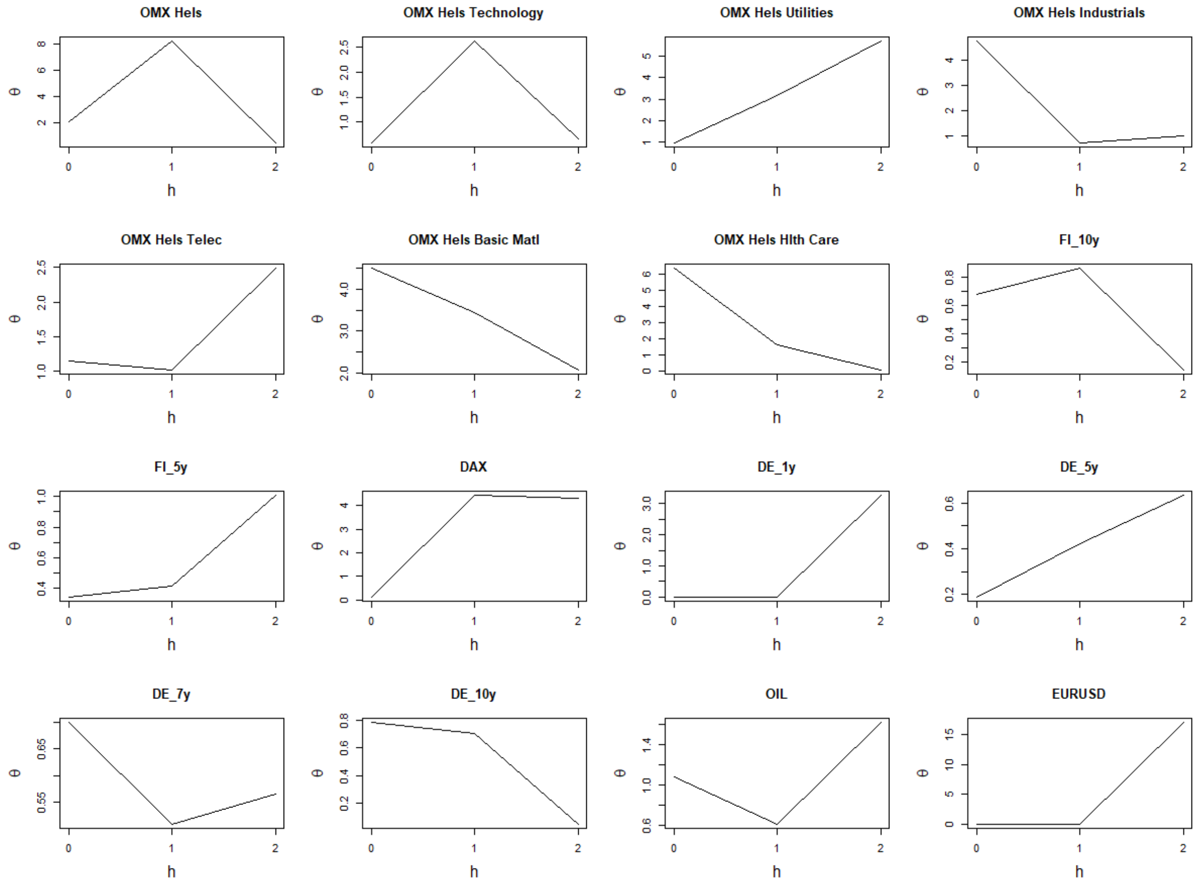


## Figure B 2

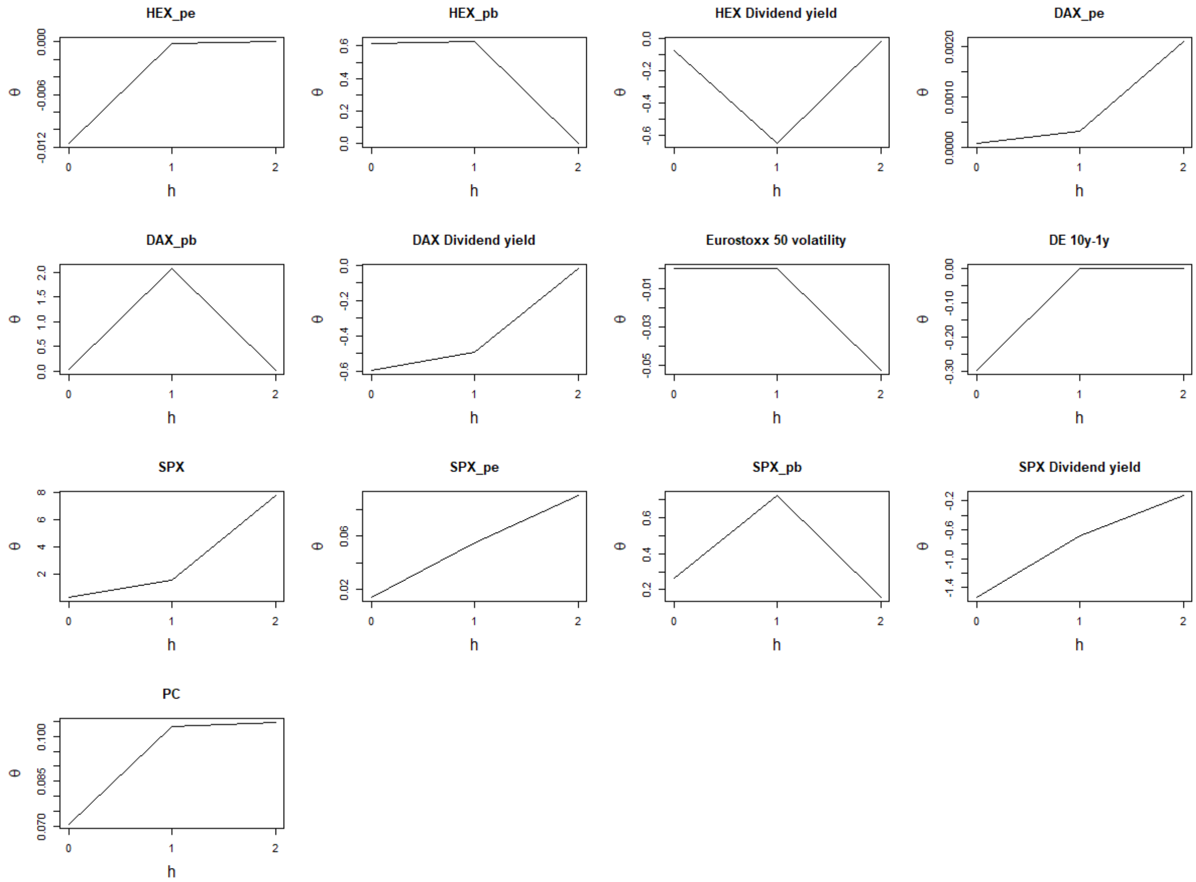



The following figures show the weighting schemes that are estimated using monthly data from Q2/2002 to Q3/2019.

## Figure B 3

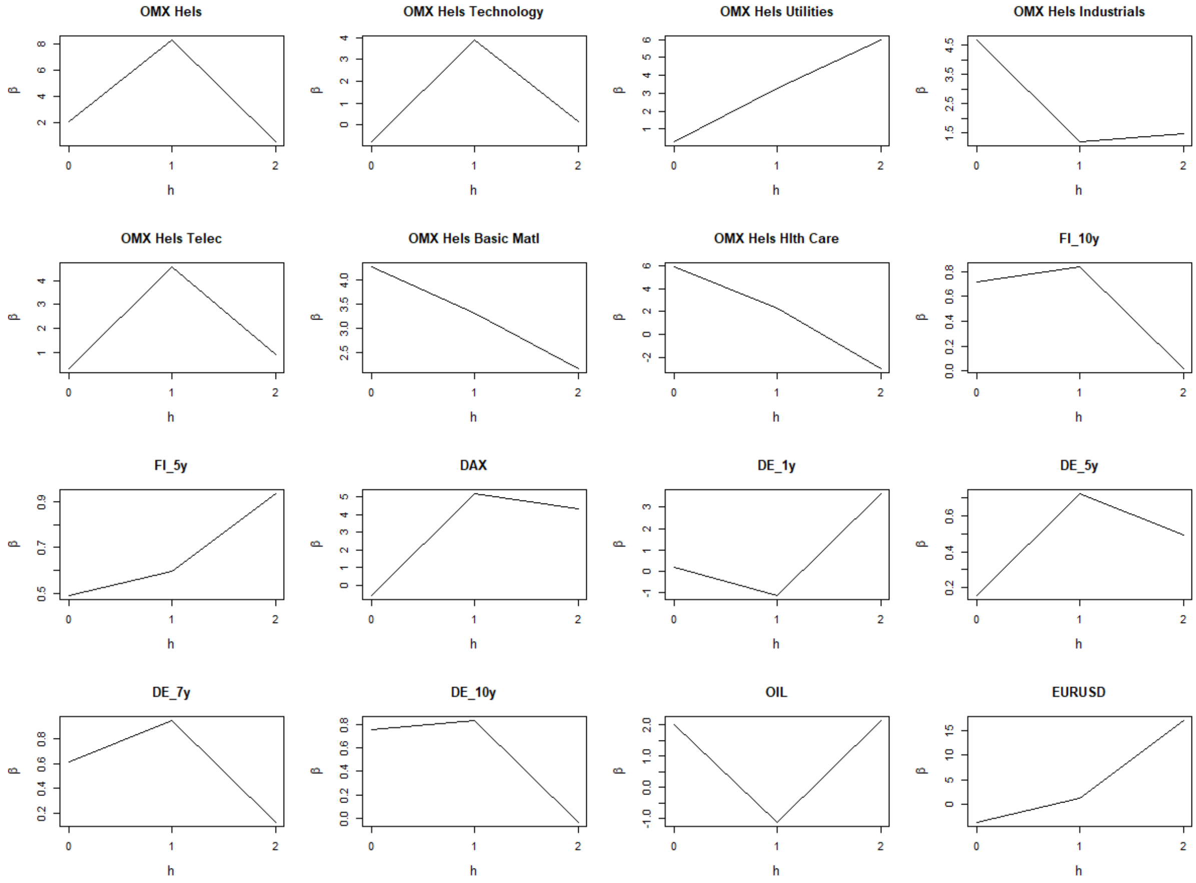



## Figure B 4

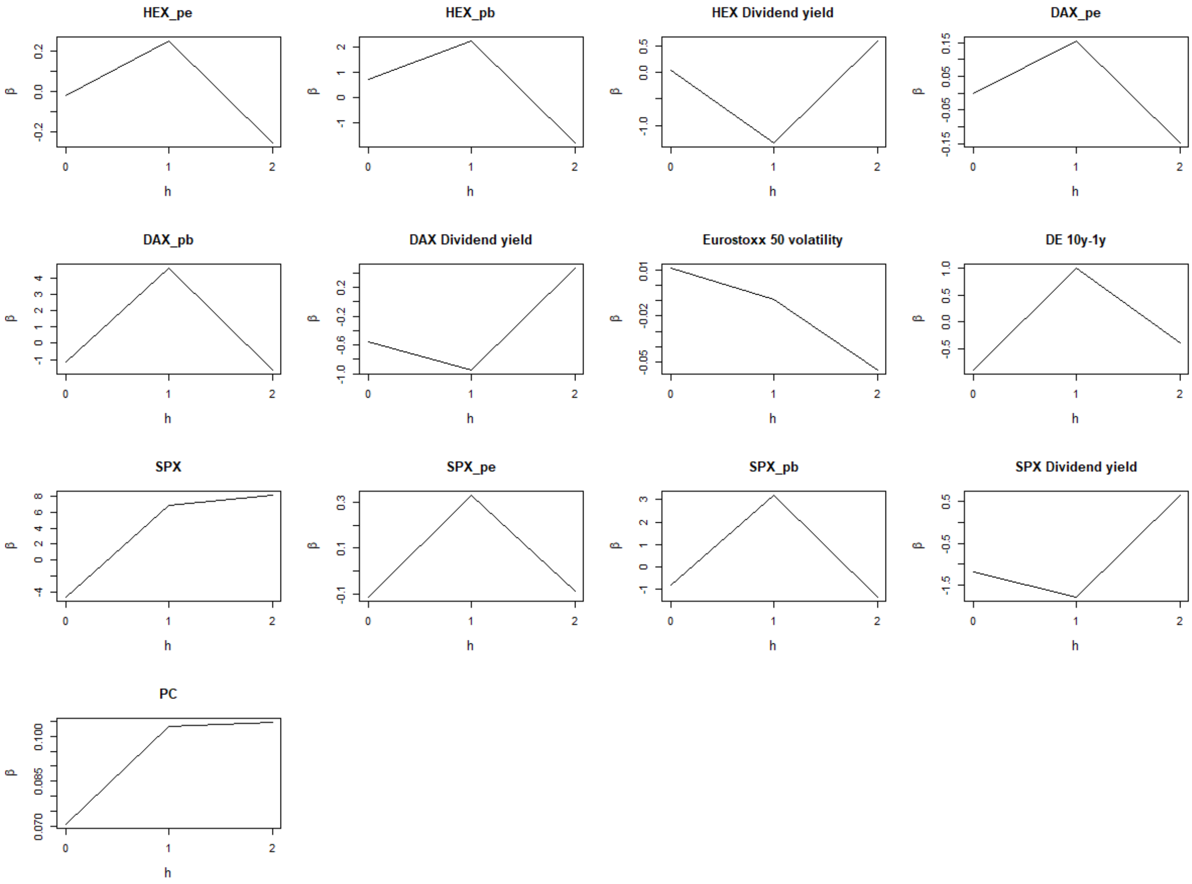


The following figures show the coefficient estimates of unrestricted model that is estimated using monthly data from Q2/2002 to Q3/2019.

**Figure B 5**

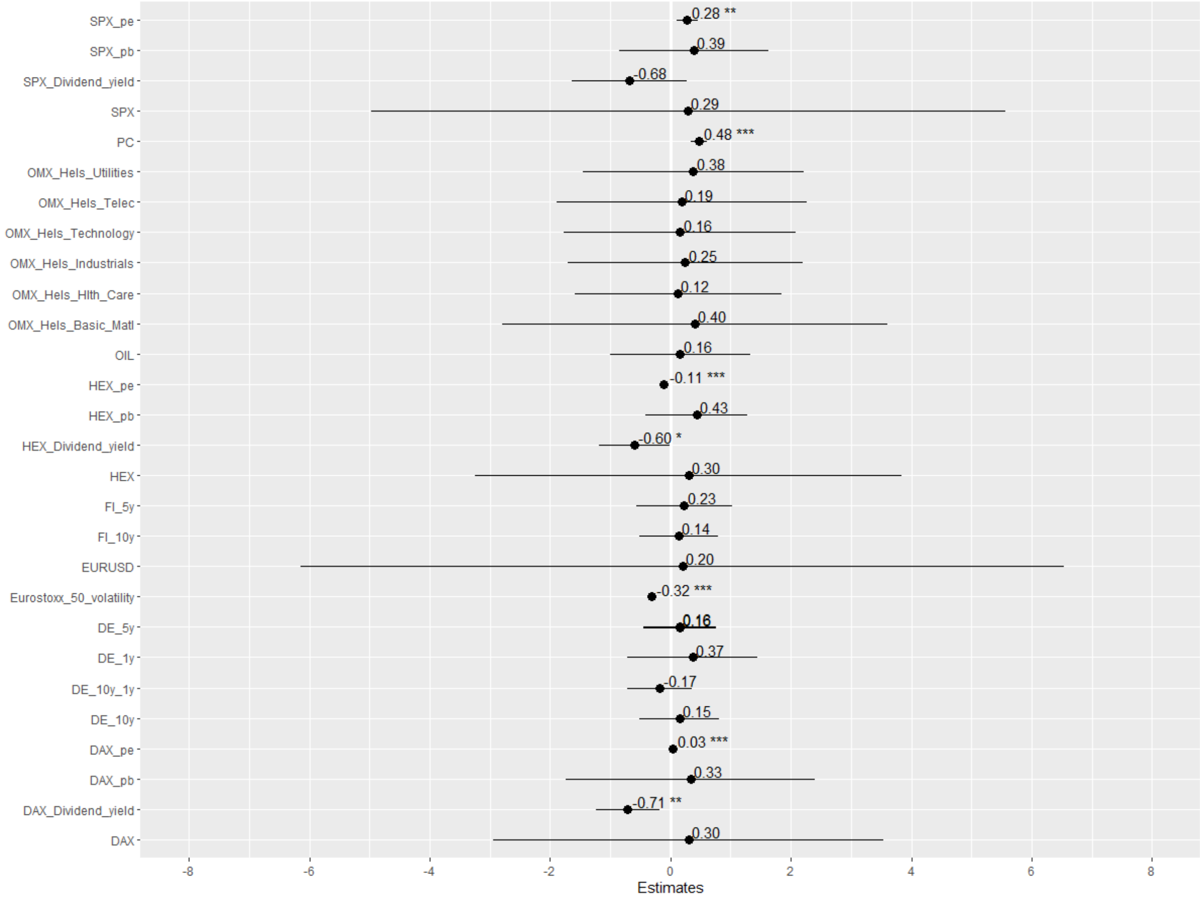


## Figure B 6




The following figure shows the standardised coefficient estimates for quarterly models. The lines represent 95 per cent confidence intervals based on heteroscedasticity and autocorrelation robust standard errors.

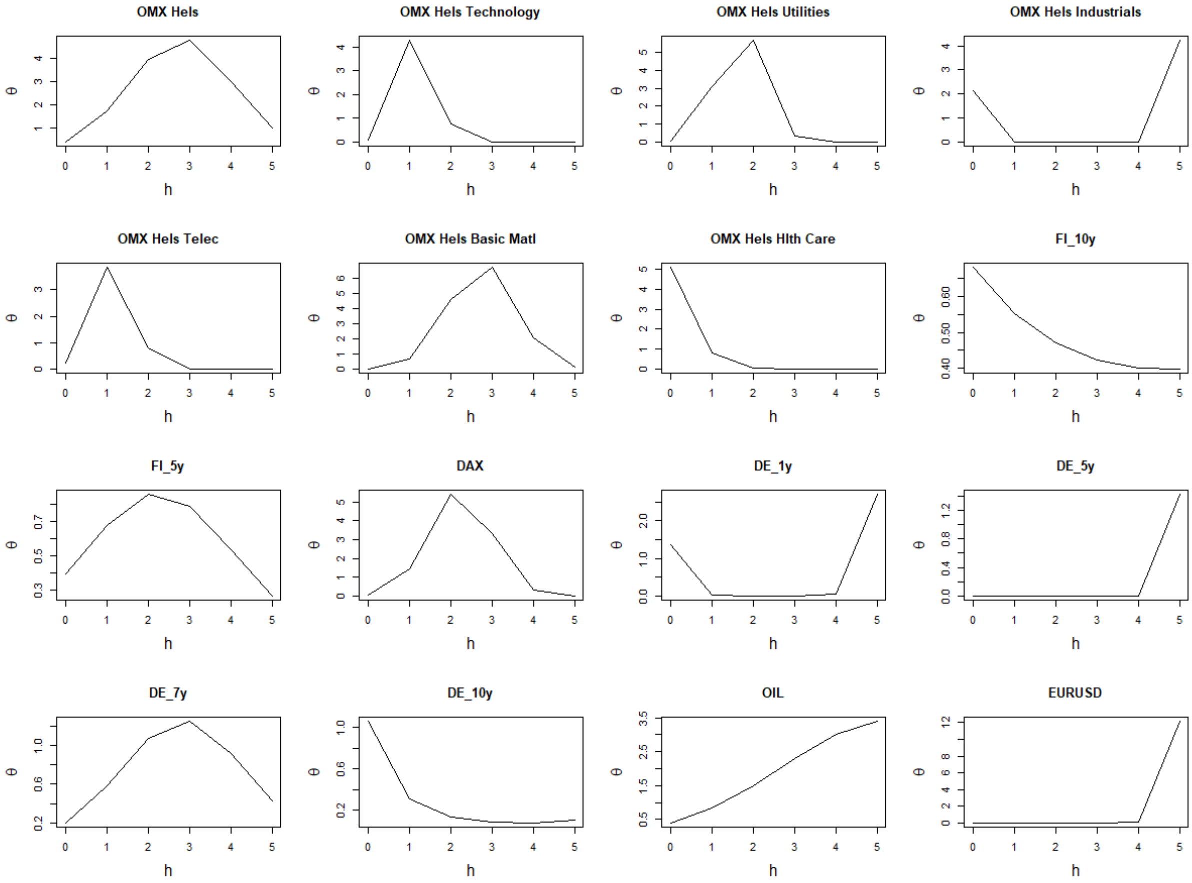
## Figure B 7


SPX_pe
SPX_pb
SPX_Dividend_yield
SPX
PC
OMX_Hels_Utilities
OMX_Hels_Telec
OMX_Hels_Technology
OMX_Hels_Industrials
OMX_Hels_Hlth_Care
OMX_Hels_Basic_Matl
OIL
HEX_pe
HEX_pb
HEX_Dividend_yield
HEX
FI_5y
FI_10y
EURUSD
Eurostoxx_50_volatility
DE_5y
DE_1y
DE_10y_1y
DE_10y
DAX_pe
DAX_pb
DAX_Dividend_yield
DAX
0.28 **
0.39
-0.68
0.29
0.48 ***
0.38
0.19
0.16
0.25
0.12
0.40
0.16
-0.11 ***
0.43
-0.60 *
0.30
0.23
0.14
0.20
-0.32 ***
0.16
0.37
-0.17
0.15
0.03 ***
0.33
-0.71 **
0.30
-8
-6
-4
-2
0
2
4
6
8
Estimates

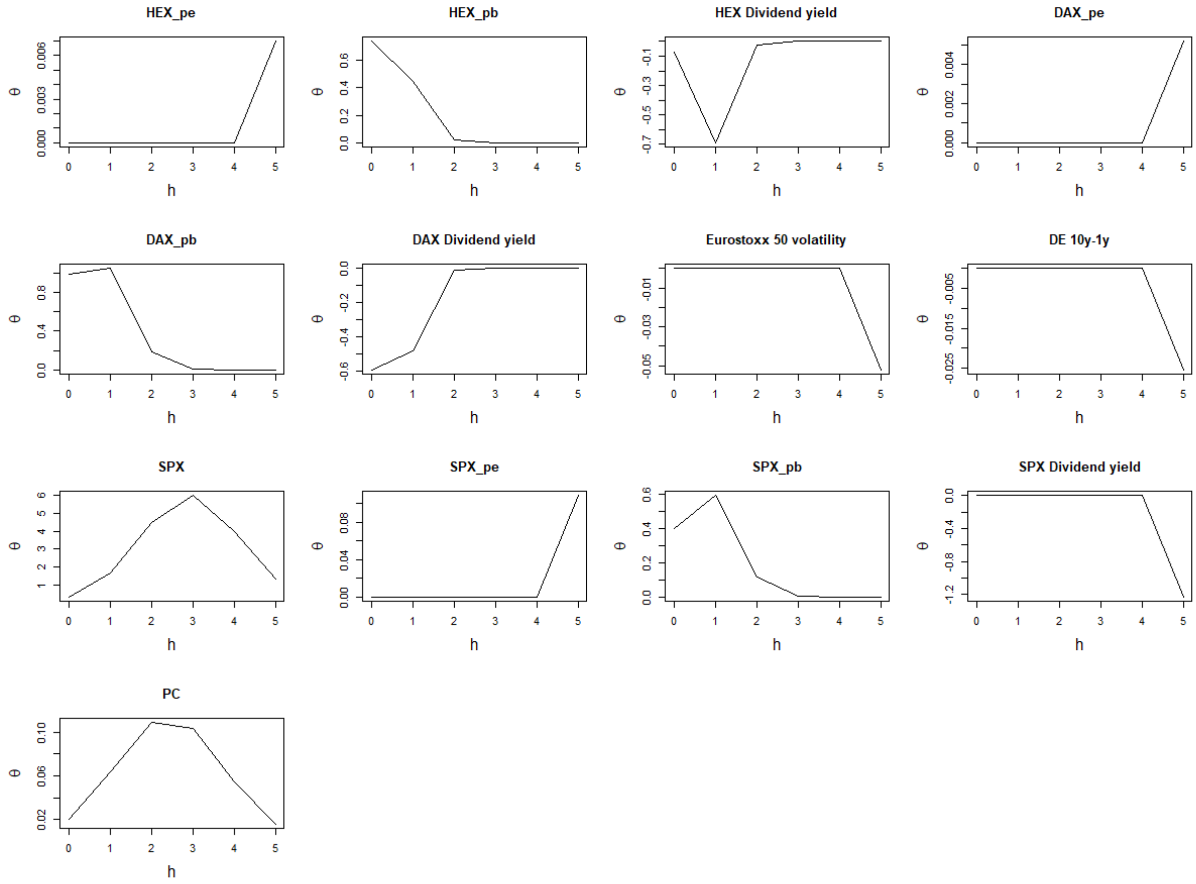
# C

The following figures show the weighting schemes that are estimated using monthly data from Q3/2002 to Q3/2019 and assuming 5 lags.

**Figure C 1**



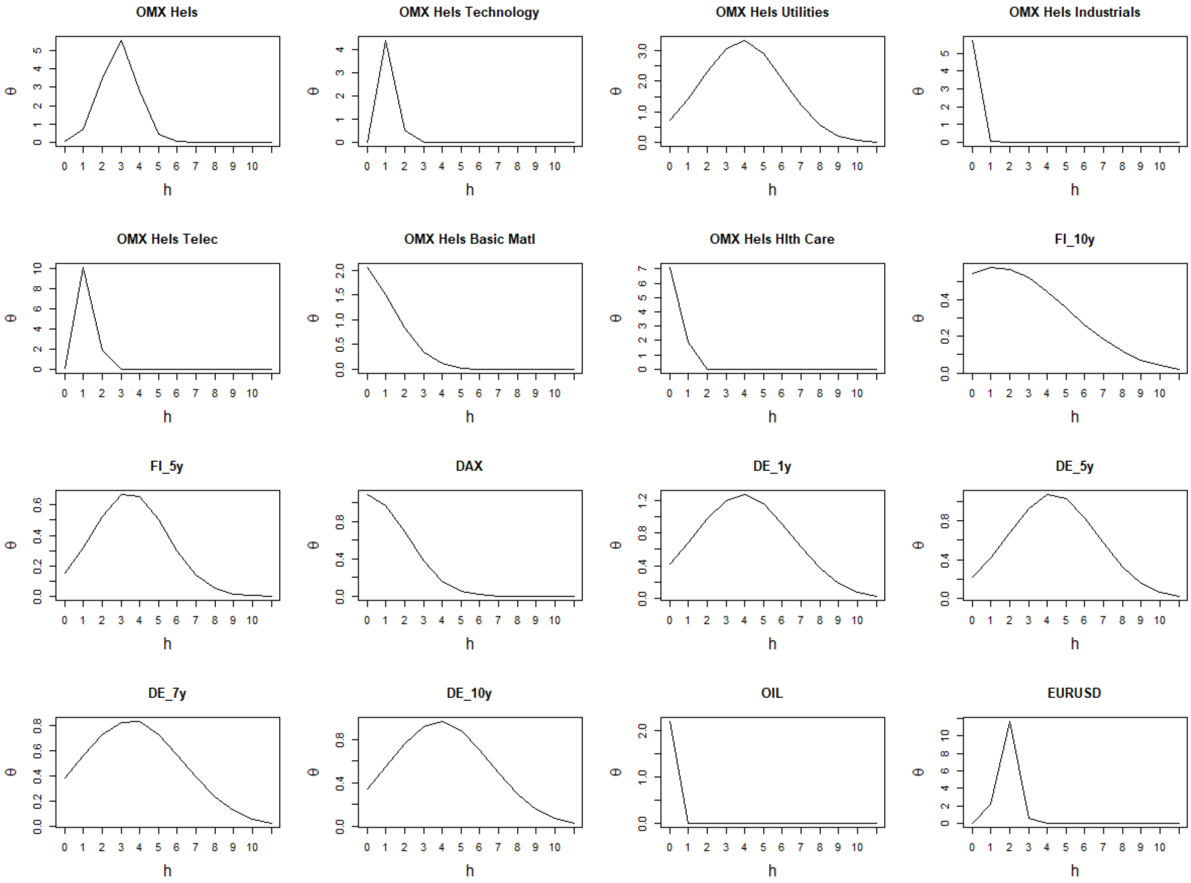
## Figure C 2


HEX_pe
HEX_pb
HEX Dividend yield
DAX_pe
DAX_pb
DAX Dividend yield
Eurostoxx 50 volatility
DE 10y-1y
SPX
SPX_pe
SPX_pb
SPX Dividend yield
PC
θ
h

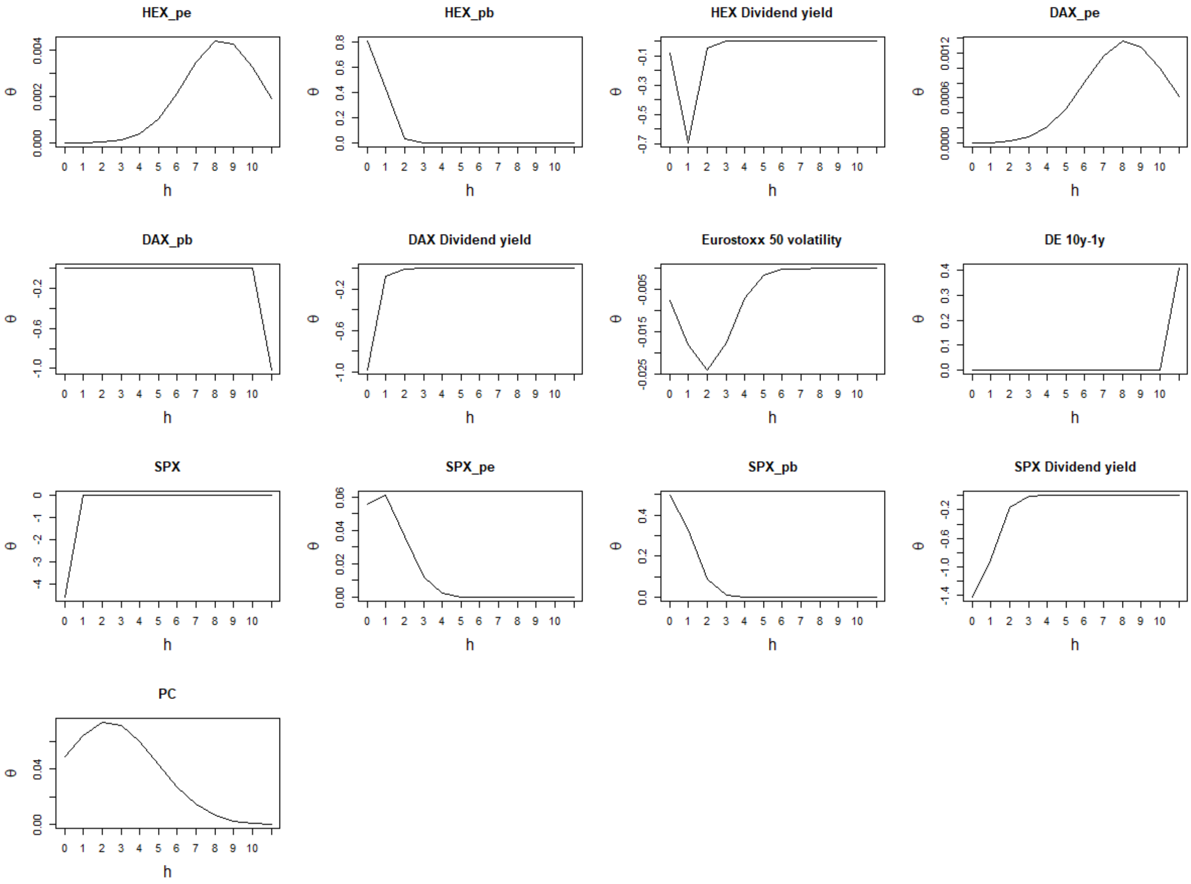
# D

The following figures show the weighting schemes that are estimated using monthly data from Q1/2003 to Q3/2019 and assuming 11 lags.

**Figure D 1**



## Figure D 2


HEX_pe
HEX_pb
HEX Dividend yield
DAX_pe
DAX_pb
DAX Dividend yield
Eurostoxx 50 volatility
DE 10y-1y
SPX
SPX_pe
SPX_pb
SPX Dividend yield
PC

# E

The following table replicates Table 1 excluding the out-of-sample forecast for the period Q2/2012. The forecast error in Q2/2012 is clearly larger than in other periods, and is excluded here in order to control for the effect of this potential outlier.

**Table E 1: Out-of-sample RMSEs of MIDAS regression models**

| Explanatory variable | Quarterly | Monthly | Monthly (unrestricted) | Daily |
|---|---|---|---|---|
| SPX_pe | 0.42 | 0.47 | 0.49 | 0.45 |
| HEX_pe | 0.53 | 0.56 | 0.64 | 0.55 |
| DE 10y-1y | 0.54 | 0.54 | 0.57 | 0.54 |
| Industrial production | 0.54 | 0.46 | 0.50 | - |
| DE_10y | 0.55 | 0.59 | 0.77 | 0.54 |
| FI_10y | 0.55 | 0.69 | 0.83 | 0.54 |
| FI_5y | 0.55 | 0.57 | 0.59 | 0.53 |
| DAX Dividend yield | 0.56 | 0.57 | 0.60 | 0.57 |
| DAX_pe | 0.56 | 0.56 | 0.78 | 0.56 |
| DE_5y | 0.56 | 0.60 | 0.60 | 0.55 |
| DE_7y | 0.56 | 0.62 | 0.71 | 0.55 |
| OIL | 0.56 | 0.56 | 0.56 | 0.62 |
| HEX_pb | 0.57 | 0.61 | 0.72 | 0.63 |
| DE_1y | 0.58 | 0.54 | 0.63 | 0.52 |
| EURUSD | 0.58 | 0.69 | 0.69 | 0.54 |
| PC | 0.59 | 0.58 | 0.58 | 0.65 |
| SPX_pb | 0.59 | 0.63 | 0.74 | 0.63 |
| HEX Dividend yield | 0.60 | 0.66 | 0.65 | 0.61 |
| OMX Hels Telec | 0.61 | 0.68 | 0.72 | 0.59 |
| OMX Hels Industrials | 0.62 | 0.60 | 0.61 | 0.58 |
| OMX Hels Hlth Care | 0.62 | 0.72 | 0.91 | 0.62 |
| SPX Dividend yield | 0.62 | 0.61 | 0.66 | 0.61 |
| SPX | 0.66 | 0.68 | 0.69 | 0.57 |
| DAX_pb | 0.67 | 0.62 | 0.79 | 0.73 |
| Eurostoxx 50 volatility | 0.67 | 0.65 | 0.67 | 0.67 |
| OMX Hels Utilities | 0.69 | 0.64 | 0.71 | 0.62 |
| HEX | 0.72 | 0.71 | 0.80 | 0.58 |
| OMX Hels Technology | 0.72 | 0.71 | 0.86 | 0.61 |
| DAX | 0.79 | 0.75 | 0.82 | 0.59 |
| OMX Hels Basic Matl | 0.82 | 0.79 | 0.84 | 0.64 |

*The models are ordered based on the RMSEs of the quarterly models. The abbreviations for the variables are explained in Appendix A. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q2/2002 to Q4/2011 and the first out-of-sample forecast is Q1/2012. The out-of-sample forecast for the period Q2/2012 is excluded. Thus, the results are based on 30 out-of-sample periods.*

F

The following table replicates Table 2 excluding the out-of-sample forecast for the period Q2/2012. The forecast error in Q2/2012 is clearly larger than in other periods, and is excluded here in order to control for the effect of this potential outlier. Comparing Table 2 to Table F 1 reveals that especially the RMSE of the industrial production driven model improves by the exclusion of Q2/2012. However, the average forecast using the financial ratios and industrial production still outperforms the industrial production driven forecast in a weakly statistically significant way when using two lags, and performs equally well using more lags. Using only the financial ratios for forecasting still produces competitive forecasts as well.

**Table F 1: RMSEs of MIDAS regression models**

| | 2 lags | 5 lags | 11 lags |
|---|---|---|---|
| Industrial production | 0.56 | 0.39 | 0.42 |
| PC | 0.57 | 0.56 | 0.54 |
| Average forecast | 0.54 | 0.55* | 0.55 |
| Average forecast: financial ratios only | 0.50 | 0.50 | 0.52 |
| Average forecast: financial ratios only and industrial production | 0.50* | 0.39 | 0.42 |

*In the model using industrial production the only explanatory variable is the MoM growth rate of the volume of industrial production. In the PC model, the only explanatory variable is the first principal component of the financial market variables (see Appendix A). 'Average forecast' is the simple average of all the financial variable based forecasts (forecasts produced using the financial market variables listed in Appendix A one at a time). 'Average forecast based on financial ratios only' is the average of the models in which the explanatory variable is the price-to-earnings ratio, the price-to-book ratio or the dividend yield. The RMSEs are calculated from a rolling window analysis, in which the first estimation sample is from Q1/2003 to Q4/2011 and the first out-of-sample forecast is Q1/2012. The out-of-sample forecast for the period Q2/2012 is excluded. Thus, the results are based on 30 out-of-sample periods. To test the statistical significance of the RMSE differences between the industrial production based forecast and the other forecasts, we use the Diebold-Mariano test assuming no heteroscedasticity or autocorrelation because the forecast horizon is zero. Asterisks *,** and*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.*