

Meta-Residuals in Dynamic Factor Models

Sorin T. Pascu*

September 18, 2017

Abstract

This paper extends the factor extraction in a dynamic factor model (DFM) by obtaining factors from forecasted data. The emerging residuals are shown to be useful in tracing the impulse transmission in the system. These residuals are used in empirical applications in forecasting and impulse responses. A possible solution for the price puzzle, the counterintuitive increase in inflation after a hike in the federal funds rate, is proposed.

1 Introduction

The work of Sims (1980) introduced impulse responses in vector autoregressive (VAR) systems as a new part of structural analysis. Since then, the models have evolved and are used in many different areas. In macroeconomics, many variables can be included in the analysis. This leads to a degree of freedom problem, since the number of parameters grows rapidly with the number of variables. Yet reducing the number of variables leads to a missing information problem.

Dynamic factor models (DFM) use the information available in large datasets by extracting factors that contain much of this information. This common component is combined with an idiosyncratic part for each individual series. DFM are used for forecasting and structural analysis. For the latter, impulses are set in factors and the responses in the series are obtained using an impact matrix (Stock and Watson 2016). Kilian and Lütkepohl (2017) point out that the impulse transmission between the series is only partly captured by this procedure and that an important shock can enter through the idiosyncratic part of a series. In addition, the representation with the impact matrix is not the Wold moving average representation of the factor model, as assumed by Stock and Watson (2016).

This paper extends the factor extraction in a DFM context by obtaining factors from forecasted data. The emerging residuals are denominated meta-residuals, and this method is used to partially improve forecasts as well as to analyze impulse responses. The inner life of the DFM without exogenous impulses is examined. Impulses are set and responses are obtained in individual series. Meta-residuals are shown to be useful in tracing the impulse transmission through factors within the system.

*I am grateful to authors who have published their data and program code, and I particularly thank James H. Stock and Mark W. Watson for this.

The rest of this article is arranged as follows: Section 2 introduces the DFM and its estimation. Section 3 proposes a new way of forecasting with DFM, introduces meta-residuals, and applies the theory on the so-called Stock-Watson dataset FRED-MD¹. Section 4 turns to structural analysis by examining impulse responses with meta-residuals in a DFM empirically and theoretically. A solution for the price puzzle, the counterintuitive increase in inflation after a hike in the federal funds rate, is proposed. Section 5 concludes.

2 Model Setup

A dynamic factor model is a time series process for a K -dimensional vector x_t that depends on an r -dimensional vector of unobserved factors f_t and a K -dimensional process v_t

$$x_t = \lambda_0 f_t + \lambda_1 f_{t-1} + \cdots + \lambda_{q^*} f_{t-q^*} + v_t. \quad (2.1)$$

The first part of the equation involving the factors is the common component, while the v_t term is the idiosyncratic part. It is introduced as VAR to explain the residual correlation and is defined as

$$v_t = A_1 v_{t-1} + \cdots + A_p v_{t-p} + u_t, \quad (2.2)$$

with $A_i, i = 1, \dots, p$ being diagonal matrices, u_t white noise with diagonal covariance matrix Σ_u and the factor evolution being assumed as

$$f_t = \Gamma_1 f_{t-1} + \cdots + \Gamma_s f_{t-s} + \eta_t. \quad (2.3)$$

The model is called an exact dynamic factor model as opposed to an approximate factor model, in which the idiosyncratic part allows for some correlation between the variables and thus has a non-diagonal Σ_u .

A shorter form is provided here:

$$x_t = \lambda(L) f_t + v_t, \quad v_t = A(L) v_{t-1} + u_t, \quad f_t = \Gamma(L) f_{t-1} + \eta_t \quad (2.4)$$

where

$$\lambda(L) = \lambda_0 + \lambda_1 L + \cdots + \lambda_{q^*} L^{q^*} \quad (2.5)$$

$$A(L) = A_1 L + \cdots + A_p L^p = \text{diag}[a_1(L), \dots, a_k(L)] \quad (2.6)$$

$$\Gamma(L) = \Gamma_1 L + \cdots + \Gamma_s L^s. \quad (2.7)$$

The notation is similar to Kilian and Lütkepohl (2017)², who also include an historical overview. An intuitive exemplifying representation of this process with $q^* = 2, p = 2$ and $s = 1$ is provided

here:

$$\begin{aligned}
 & \leftarrow f_t = \Gamma_1 f_{t-1} + \eta_t \\
 & \leftarrow f_{t-1} = \Gamma_1 f_{t-2} + \eta_{t-1} \\
 & \leftarrow f_{t-2} = \Gamma_1 f_{t-3} + \eta_{t-2} \\
 & \leftarrow v_{t-1} = A_1 v_{t-2} + A_2 v_{t-3} + u_{t-1} \\
 & \leftarrow v_{t-2} = A_1 v_{t-3} + A_2 v_{t-4} + u_{t-2} \\
 & \leftarrow u_t \text{ being white noise}
 \end{aligned}$$

$$x_t = \lambda_0 f_t + \lambda_1 f_{t-1} + \lambda_2 f_{t-2} + A_1 v_{t-1} + A_2 v_{t-2} + u_t$$

The model representation can be stacked with R static factors merged in a matrix F containing the r dynamic factors of the matrix f (see also Kilian and Lütkepohl (2017) or section 6.3 of the appendix). This is the so-called static form.

Its estimation is described in Stock and Watson (2005b); see also Kilian and Lütkepohl (2017):

Step 1 Regress the individual time series on their own lags to obtain an initial estimate for the filtered data $(I - \hat{A}(L))X$.

Step 2 Compute an initial $\hat{F}$ as the R principal components of $(I - \hat{A}(L))X$ corresponding to the R biggest eigenvectors.

Step 3 Estimate Λ by K individual regressions $(I - \hat{A}(L))X = \hat{\Lambda}\hat{F} + \hat{u}$. Filter the data with the new estimate of $(I - \hat{A}(L))$. Compute F as the R principal components of $(I - \hat{A}(L))X$ corresponding to the R biggest eigenvectors.

Step 4 Iterate step 3 until convergence.

Stock and Watson (2005b) compute the eigenvectors of $((I - \hat{A}(L))X)'(I - \hat{A}(L))X$ for $\hat{\Lambda}$ or $(I - \hat{A}(L))X((I - \hat{A}(L))X)'$ for $\hat{F}$, depending on which computation is faster ($K > T$ or $K < T$). The remaining part is computed through OLS³ as $\hat{F} = \hat{\Lambda}'(I - \hat{A}(L))X$ (or $\hat{\Lambda} = ((I - \hat{A}(L))X)\hat{F}'$). However, Stock and Watson (2005b) regress the data x on the factors F and the lags of the data x . In contrast, this article regresses on the factors F and the lags of the idiosyncratic part v , according to the model setup (2.1).

The obtained static factors are only a linear transformation of the true dynamic factors. Following Stock and Watson (2005b), the dynamic factors f_t are estimated by first regressing $\hat{F}_t$ on $\hat{F}_{t-1}$ to obtain the residuals $\hat{U}_t$. The first r principal components of the covariance matrix $\hat{\Sigma}_U = T^{-1}\sum_t \hat{U}_t \hat{U}'$ of these residuals are denoted as $\hat{W}$. The primitive dynamic factors are estimated as

$$\hat{f}_t = \hat{W}'\hat{F}_t, \quad (2.8)$$

while $\hat{\eta}_t = \hat{W}'\hat{U}_t$ (Kilian and Lütkepohl (2017)).

Finally, since

$$(I - A(L))x_t = (I - A(L))\lambda(L)f_t + u_t$$

$$\begin{aligned}(I - A(L))x_t - u_t &= (I - A(L))\lambda(L)f_t \\ (I - A(L))^{-1}((I - A(L))x_t - u_t) &= \lambda(L)f_t\end{aligned}\tag{2.9}$$

the coefficients of $\lambda(L)$ are obtained by regressing the left side of (2.9) on the newly obtained dynamic factors f ⁴.

3 Forecasting with Meta-Residuals

3.1 Theoretical Aspects

Forecasting DFM can be done by iterating forward

$$E[X_{t+1}|X_t, f_t, X_{t-1}, f_{t-1}\dots] = \alpha^f(L)f_t + A(L)X_t\tag{3.1}$$

where

$$\alpha^f(L) = \lambda_0\Gamma(L) - A(L)\lambda(L) + L^{-1}(\lambda(L) - \lambda_0).$$

See Stock and Watson (2016) for its derivation (or section 6.1 of the appendix).

In this article, factors are also extracted from forecasted data with the restriction of given parameters $\Lambda, W, \lambda(L), \Gamma(L)$ and $A(L)$. In addition, the already extracted factors need to remain unchanged. This is achieved by the following procedure:

Step 1 Forecast X_{t+1} with (3.1).

Step 2 Extract new static factors $\hat{F}_{t+1}$ through OLS from $(I - \hat{A}(L))x_{t+1} = \Lambda\hat{F}_{t+1}$ and obtain estimates of unchanged past factors and new future factors⁵. Those estimates allow for the computation of the common component $\hat{\Lambda}\hat{F}$ including the future $\hat{\Lambda}\hat{F}_{t+1}$.

Step 3 Use the previously estimated $\hat{W}$ to obtain the dynamic factors $\hat{f}$ including $\hat{f}_{t+1}$ with $\hat{f} = \hat{W}'\hat{F}$.

Step 4 Compute the other forecast periods $t + 2, \dots, t + h$ as above in steps 1 - 3.

In order to obtain unchanged dynamic factors $\hat{f}$, the matrix $\hat{W}$ from (2.8) is assumed constant and used, in a sense, as invariant loadings for $\hat{F}$.

The data consist of 123 series from the data set FRED-MD for macroeconomic research. All data are adjusted for outliers, transformed to be stationary according to the codes provided by FRED and, finally, demeaned and standardized (see section 6.4 for details).

The number of static factors R is 7, as determined by the information criterion $ICp1$ introduced by Bai and Ng (2002). Criteria from Bai and Ng (2007) determine the number of dynamic factors r to be 3. The lag numbers for the idiosyncratic part p is set to 4, as well as the number of past lags for the common component q^* (thus, including the contemporaneous lag, the common component has 5 lags; see section 6.2 for further details).

Figure 1 shows the end of the first dynamic factor extracted through the initial procedure (solid green line) with the corresponding first dynamic factor extracted through the second procedure with constant coefficients (dotted blue line). The latter also has future values. Note that the past is almost unchanged⁶.

In all figures, the last available data point is marked with a dashed, vertical line and the NBER recession dates are shaded in gray.

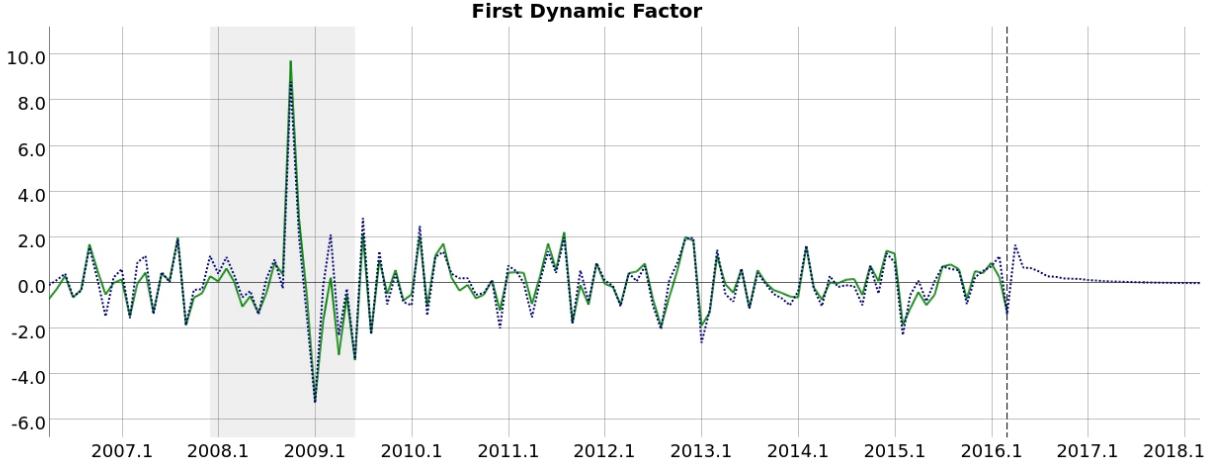


Figure 1: The last 120 values of the first dynamic factor extracted through the initial procedure (solid green line) and through the second procedure with constant coefficients (dotted blue line).

The differences between the extracted factors and the forecasted factors are denominated meta-residuals and written as η^{meta} . In the following, the factors extracted from forecasts are referred to as $\hat{f}^{extract}$ and the forecasted dynamic factors, based on $\hat{f}^{extract}$, as $\hat{f}^{forecast}$. Meta-residuals u^{meta} for the idiosyncratic part arise through the factor extraction.

Here are the formulas for the first forecast period:

$$\hat{f}_{t+1}^{forecast} = \Gamma_1 \hat{f}_t + \dots + \Gamma_s \hat{f}_{t-s+1} \quad (3.2)$$

$$\hat{f}_{t+1}^{extract} = \Gamma_1 \hat{f}_t + \dots + \Gamma_s \hat{f}_{t-s+1} + \eta_{t+1}^{meta} \quad (3.3)$$

$$\hat{f}_{t+1}^{extract} - \hat{f}_{t+1}^{forecast} = \eta_{t+1}^{meta}. \quad (3.4)$$

The extraction is done similarly in every forecast period, using previously extracted factors (possibly combined with past factors due to (2.3))⁸.

3.2 Empirical Results for Forecasting

The extracted factors were examined in the previous section with the focus on the unchanged past. Figure 2 shows $\hat{f}^{forecast}$ for 24 periods ahead, compared to $\hat{f}^{extract}$. Because the factors live longer, the factor structure can be better exploited for forecasting. The formula here is:

$$E[X_{t+1}|X_t, f_t, X_{t-1}, f_{t-1}...] = \alpha^f(L)f_t^{forecast} + A(L)X_t. \quad (3.5)$$

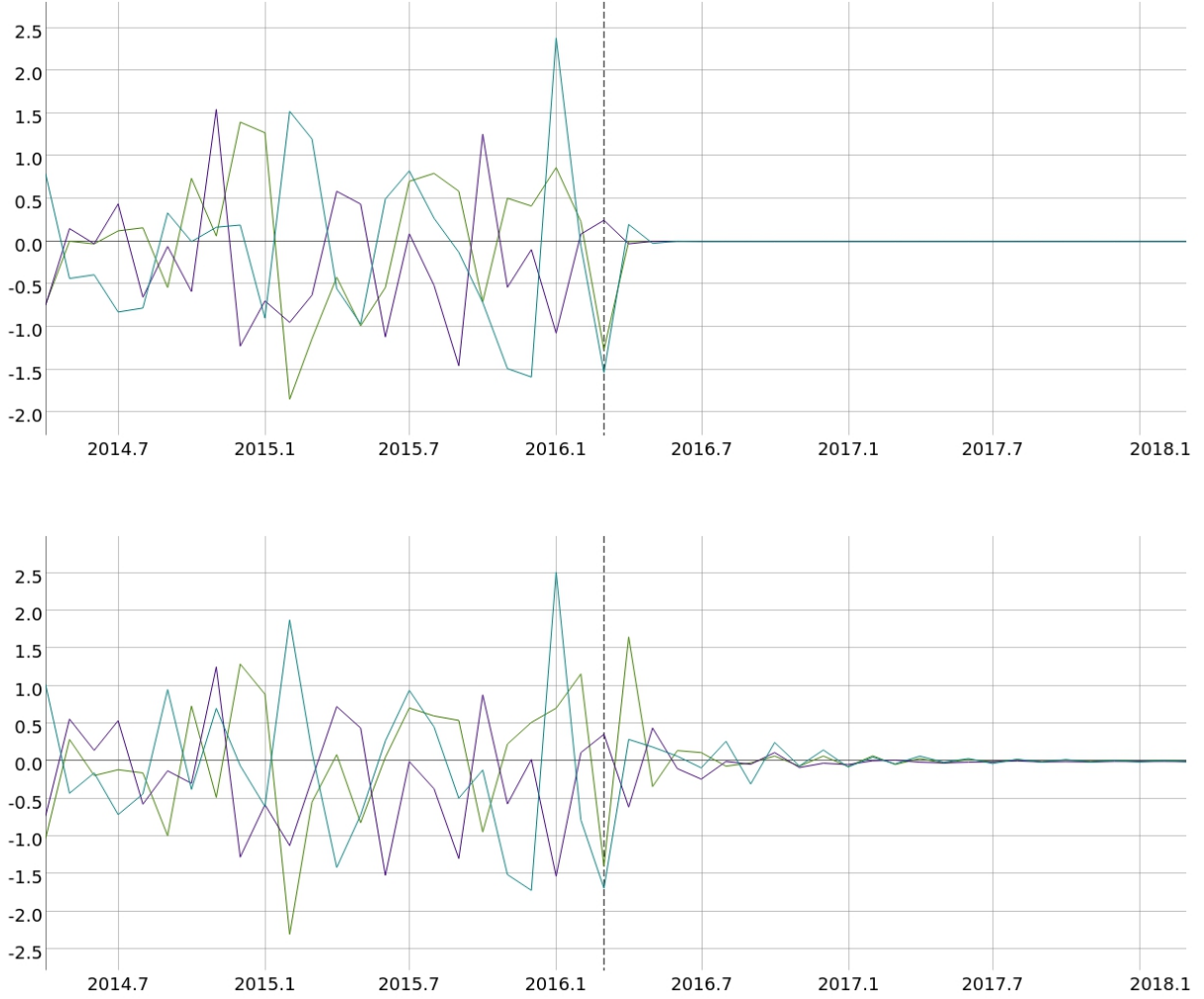


Figure 2: $\hat{f}^{forecast}$ (upper part) and $\hat{f}^{extract}$

The forecasts from (3.1) and (3.5) are compared using the relative mean squared error (MSE) based on the univariate AR(4) forecast. Thus, values smaller than 1 are better than the univariate AR(4) forecast. By definition, the forecast value of the first period is the same for both methods. The transformation, demeaning and standardization of data, as well as the estimation of factors and loadings, occurs recursively. All regressions start in 1984.01, where Stock and Watson (2009) found a break in the loadings. The periods for forecast comparison start in 1985.1, when the first 12-months-ahead forecast is available.

Table 1: Relative MSE for INDPRO for the periods 1985.01/end (relative to univariate AR(4)) for the 1-12 periods ahead forecasts.

Method	1	2	3	4	5	6	7	8	9	10	11	12
DFM	1,05	1,03	1,16	1,08	1,15	1,11	1,09	1,07	1,06	1,05	1,04	1,03
DFM meta	1,05	0,22	0,29	0,12	0,12	0,14	0,11	0,26	0,12	0,99	0,44	8,46

Table 1 compares the two forecast methods from (3.1) and (3.5). A first glance suggests that the forecasts have improved considerably; however, some issues remain. Figure 3 shows the plots of the pseudo out-of-sample forecasts for 6 and 11 months. The second graphic shows a massive drop in the forecast with meta-residuals that occurs after the great recession, within one period, likely related to a structural break. The drop improves the relative MSE, but this improvement is purely accidental. The forecasts with a higher horizon, e.g. the 12-months-ahead forecast, show an evident deterioration in the relative MSE value⁹. Table 2 shows the relative MSE for the period up to the great recession (1985.01 - 2007.11), when the pure DFM forecasts are partially better than the DFM forecasts with meta-residuals.

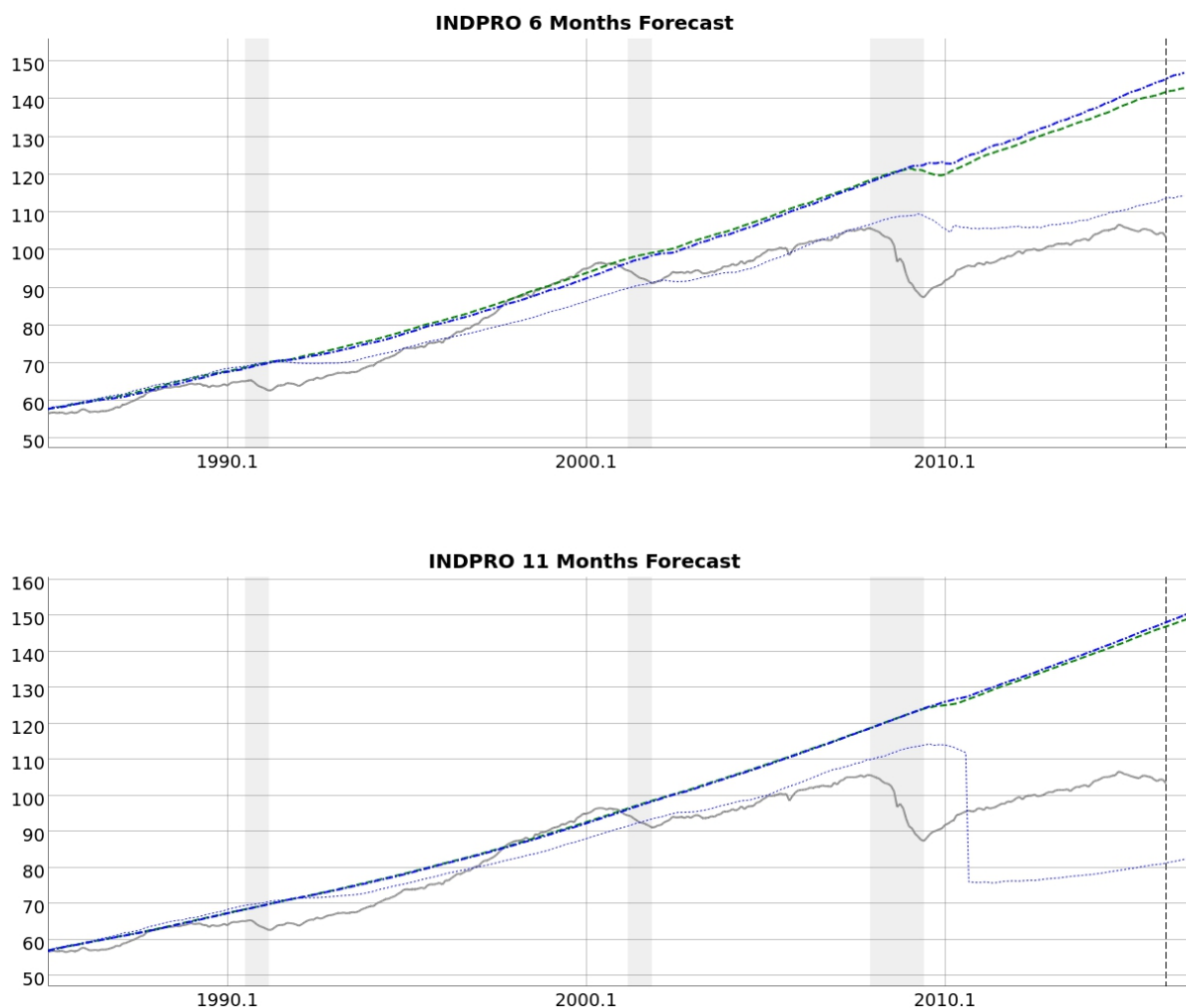


Figure 3: Data (solid gray) and forecasts: AR(4) (dashed green), DFM (dotted dashed blue), and DFM with meta-residuals (dotted blue).

Another problem is that forecasts for other series show significantly poorer results, at least with these parameter values. As seen in tables 3 and 4 for the series industrial production of consumer goods (IPCONGD), values in which the DFM forecasts have improved, or one of the methods is better than the univariate AR(4) benchmark, are barely found.

Table 2: Relative MSE for INDPRO for the periods 1985.01/2007.11 (relative to univariate AR(4)) for the 1-12 periods ahead forecasts.

Method	1	2	3	4	5	6	7	8	9	10	11	12
DFM	0,97	0,41	0,58	0,58	0,81	0,85	0,89	0,92	0,94	0,95	0,97	0,98
DFM meta	0,97	0,96	1,42	0,63	0,51	0,39	0,36	0,35	0,36	0,37	0,37	0,39

Table 3: Relative MSE for IPCONGD for the periods 1985.01/end (relative to univariate AR(4)) for the 1-12 periods ahead forecasts.

Method	1	2	3	4	5	6	7	8	9	10	11	12
DFM	1,41	1,48	1,18	1,20	1,06	1,04	1,04	1,01	1,01	1,01	1,00	1,00
DFM meta	1,41	1,62	1,40	1,18	1,10	1,02	1,03	0,95	1,06	0,89	1,22	0,72

3.3 A Closer Look at Meta-Residuals

The procedure of acquiring meta-residuals only shifts data from one process to another within one period, without changing the forecasted data for that period. Therefore, the forecasts based on meta-residuals do not differ from the Stock and Watson forecasts in $t + 1$.

$$E[X_{t+1}|X_t, f_t, X_{t-1}, f_{t-1}\dots] = \alpha^f(L)f_t + \lambda_0\eta_{t+1}^{meta} + A(L)X_t + u_{t+1}^{meta} = \alpha^f(L)f_t + A(L)X_t$$

and in general it holds:

$$\alpha^f(L)f_{t+h}^{forecast} + \lambda_0\eta_{t+h+1}^{meta} + A(L)X_{t+h} + u_{t+h+1}^{meta} = \alpha^f(L)f_{t+h}^{forecast} + A(L)X_{t+h}$$

where the forecast $f_{t+h}^{forecast}$ is the one period ahead forecast based on $f_{t+h-1}^{extract}$.

Therefore, it holds:

$$\begin{aligned}\lambda_0\eta_{t+h}^{meta} + u_{t+h}^{meta} &= 0 \\ \lambda_0\eta_{t+h}^{meta} &= -u_{t+h}^{meta}\end{aligned}\tag{3.6}$$

The forecasts of the data series are not directly improved by using meta-residuals. Rather, the forecasts of the factors are improved through extracting better estimates in previous periods from forecasted data.

Equation (3.6) shows one of the main results of this article: the meta-residuals arise due to the transmission between the common component and the idiosyncratic part, and can be used to trace this transmission. Since the individual part is supposed as a collection of univariate processes

Table 4: Relative MSE for IPCONGD for the periods 1985.01/2007.11 (relative to univariate AR(4)) for the 1-12 periods ahead forecasts.

Method	1	2	3	4	5	6	7	8	9	10	11	12
DFM	1,77	1,78	1,16	1,30	1,06	1,02	1,03	1,01	1,00	1,01	1,00	1,00
DFM meta	1,77	2,19	1,80	1,63	1,41	1,22	1,13	1,12	1,11	1,10	1,08	1,07

($A_i, i = 1, \dots, p$ being diagonal matrices, u_t white noise with diagonal covariance matrix Σ_u), the whole interaction within the system can be traced through meta-residuals. If the assumptions for the idiosyncratic part seem too restrictive, the model could be built as a hierarchical model (Moench, Ng, and Potter (2013), Hallin and Liška (2011)) with groups and subgroups such as the regional housing market, local housing market, etc. One could then compute (sub-) group meta-residuals.

Figure 4 shows the data corresponding to (3.6) for four series: industrial production, industrial production of consumer goods, federal funds rate and core inflation (series INDPRO, IPCONGD, FEDFUNDS and CPIULFSL¹⁰; the data are still transformed¹¹). The data support the conclusions in (3.6) only for the last three series. For industrial production, however, $-u_i^{meta}$ seems to be mirrored, suggesting that negation is not necessary here.

There is a technical difficulty that has not yet been mentioned: the factors are not sign-identified. This might not be relevant if the corresponding loadings and the idiosyncratic part can be adapted¹². McCracken and Ng (2016) suggest simply flipping the entire data series if the sign is not correct. However, it is not clear how to proceed with the sign of meta-residuals. One could flip the signs according to (3.6), such that the meta-residuals for the series INDPRO look like those of the series IPCONGD. However, because the formula (3.6) is new, no changes are made to the data. Instead, the impulse response analysis is continued with the series IPCONGD, FEDFUNDS and CPIULFSL. The evolution of the meta-residuals of these series corresponds to (3.6) after period $t + 3$ or $t + 4$, but with a slight shift.

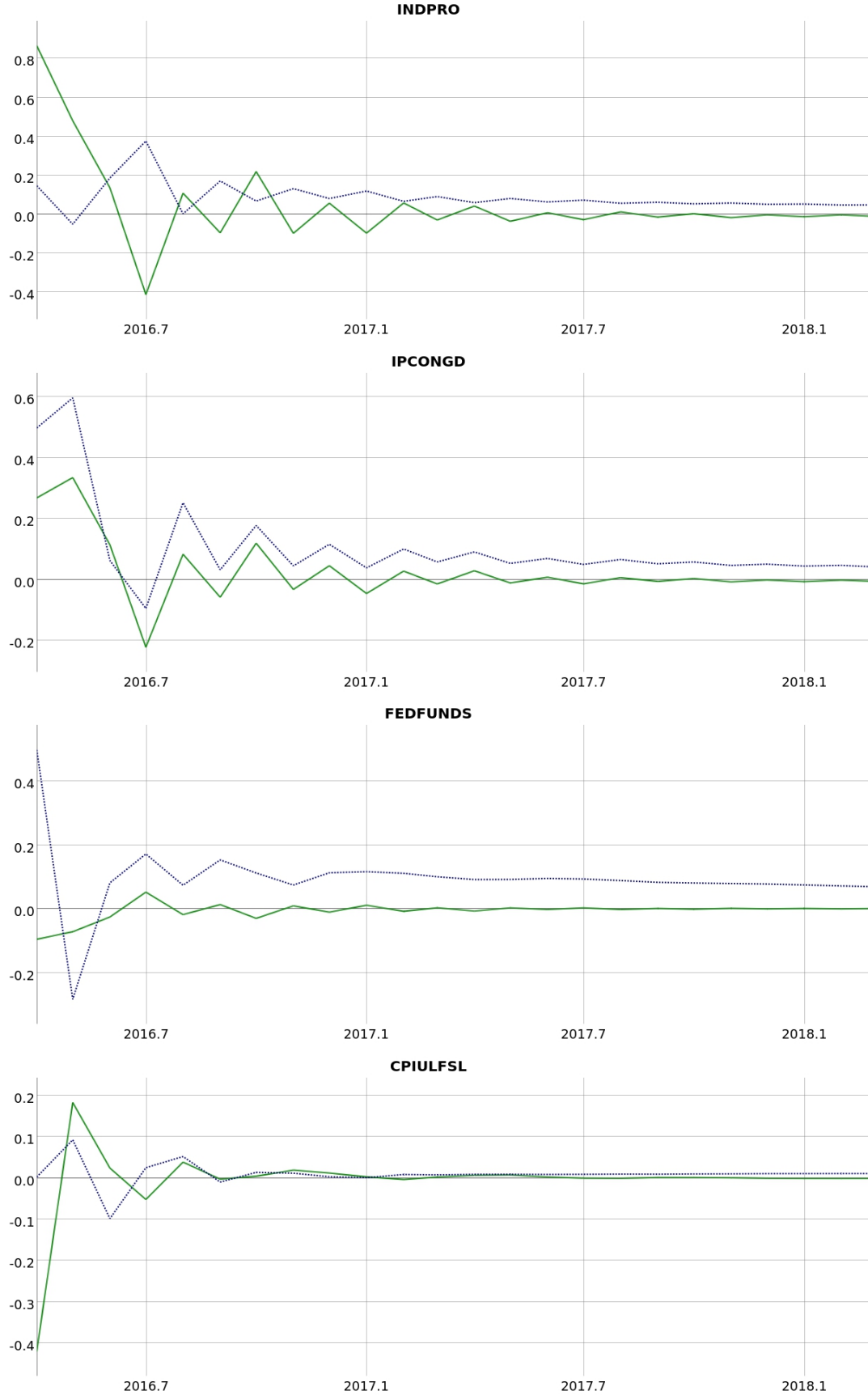


Figure 4: Meta-residuals for industrial production, industrial production of consumer goods, federal funds rate, and inflation corresponding to (3.6). The solid green line represents $\lambda_{0i}\eta^{meta}$ and the dotted blue line $-u_i^{meta}$ of the respective series i . All data are future data, such that the y-axis corresponds to $(t + 1)$.

4 Impulses and their Transmission

Process (2.1) is stationary and therefore has a Wold moving average (MA) representation. Thus, both subprocesses $\lambda(L)f_t$ and v_t , have a Wold MA representation as well (see (Lütkepohl 2005), p. 26). In sections 4.1, 4.2, and 4.3, the two subprocesses are considered independent, and no transmission of impulses occurs between them. This assumption will be dropped in section 4.4 and in the empirical part in section 4.5.

In the beginning, the process is simplified to $q^* = 0$, $p = 1$, and $s = 1$:

$$x_t = \lambda_0 f_t + v_t$$

$$\begin{aligned} \leftarrow f_t &= \Gamma_1 f_{t-1} + \eta_t \\ \leftarrow v_{t-1} &= A_1 v_{t-2} + u_{t-1} \\ x_t &= \lambda_0 f_t + A_1 v_{t-1} + u_t \end{aligned}$$

In the following, the indexes in λ_0 , Γ_1 , and A_1 remain, although the subprocesses contain one matrix each and thus need no index.

4.1 Impulse in the Idiosyncratic Part Residuals u_t

The relevant equations are the same as in Lütkepohl (2005), p. 51, and are repeated here. In order to isolate the impulse, the variables are assumed to have a value of zero prior to time $t = 0$, corresponding to their mean. The factors are set to zero.

The impulse is set in the idiosyncratic part of the first variable: the corresponding residual is set to one in period 0, that is, $u_{1,0} = 1$.

From the general equation $x_t = \lambda f_t + v_t = \lambda f_t + A_1 v_{t-1} + u_t$, it follows:

$$\begin{aligned} x_0 &= \lambda * 0 + v_0 = \lambda * 0 + A_1 * 0 + u_0 = u_0 \\ x_1 &= 0 + A_1 v_0 = A_1 u_0 \text{ since all future residuals } u_{1,2,3,\dots} = 0 \\ x_2 &= 0 + A_1 v_1 = A_1^2 u_0 \\ &\vdots \\ x_i &= A_1 v_{i-1} = A_1^i u_0 \end{aligned}$$

A process with several individual series and an impulse in the idiosyncratic part of the first series would have a response in $(t + i)$ of

$$x_i = A_1^i \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (4.1)$$

Because A_1 is diagonal and the factors are assumed to be zero, the variables have no effect on each other.

4.2 Impulse in the Factor Evolution Residuals η_t

In order to isolate the impulse, the factors are assumed to have a value of zero prior to time $t = 0$, corresponding to their mean, and the residual of the last factor is set to one, that is, $\eta_{r,0} = 1$. The idiosyncratic part is set to zero.

From the general equation $x_t = \lambda f_t + v_t = \lambda_0(\Gamma_1 f_{t-1} + \eta_t) + A_1 v_{t-1} + u_t$, it follows:

$$x_0 = \lambda_0 f_0 + 0 = \lambda_0(\Gamma_1 f_{-1} + \eta_0) = \lambda_0(\Gamma_1 * 0 + \eta_0) = \lambda_0 \eta_0$$

$$x_1 = \lambda_0 f_1 + 0 = \lambda_0(\Gamma_1 f_0 + \eta_1) = \lambda_0(\Gamma_1 f_0 + 0) = \lambda_0 \Gamma_1 \eta_0 \text{ since all future residuals } \eta_{1,2,3,\dots} = 0$$

$$x_2 = \lambda_0 f_2 + 0 = \lambda_0(\Gamma_1 f_1 + \eta_2) = \lambda_0(\Gamma_1 f_1 + 0) = \lambda_0 \Gamma_1^2 \eta_0$$

$\vdots$

$$x_i = \lambda_0 f_i + 0 = \lambda_0(\Gamma_1 f_{i-1} + \eta_i) = \lambda_0(\Gamma_1 f_{i-1} + 0) = \lambda_0 \Gamma_1^i \eta_0$$

A process with two dynamic factors and an impulse in the last factor looks like the following:

$$x_i = \lambda_0 \Gamma_1^i \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (4.2)$$

The factor impulse will affect all variables and will be considered later.

4.3 Impulses in Both Residuals u_t and η_t without Transmission

The common component and the idiosyncratic part are regarded here as two independent, stationary processes, each of them having a Wold MA representation. The derivations from the previous sections are combined and given here in more detail. The terms containing the factors and series prior to time $t = 0$ (assumed to be zero, corresponding to their mean) and the terms containing future residuals (also assumed to be zero) are shown here in boldface.

Starting from the general equation $x_t = \lambda f_t + v_t = \lambda f_t + A_1 v_{t-1} + u_t$, it follows:

$$\begin{aligned} x_0 &= \lambda_0 f_0 & + & v_0 \\ &= \lambda_0(\Gamma_1 \mathbf{f}_{-1} + \eta_0) & + & \mathbf{A}_1 \mathbf{v}_{-1} & + & u_0 \\ &= \lambda_0 \Gamma_1 \mathbf{f}_{-1} & + & \lambda_0 \eta_0 & + & \mathbf{A}_1 \mathbf{v}_{-1} & + & u_0 \\ \\ x_1 &= \lambda_0 f_1 & + & v_1 \\ &= \lambda_0(\Gamma_1 f_0 + \boldsymbol{\eta}_1) & + & A_1 v_0 & + & \mathbf{u}_1 \\ &= \lambda_0 \Gamma_1 f_0 & + & \boldsymbol{\lambda}_0 \boldsymbol{\eta}_1 & + & A_1(\mathbf{A}_1 \mathbf{v}_{-1} + u_0) & + & \mathbf{u}_1 \\ &= \lambda_0 \Gamma_1(\Gamma_1 \mathbf{f}_{-1} + \eta_0) & + & \boldsymbol{\lambda}_0 \boldsymbol{\eta}_1 & + & \mathbf{A}_1^2 \mathbf{v}_{-1} & + & A_1 u_0 & + & \mathbf{u}_1 \\ &= \lambda_0 \Gamma_1^2 \mathbf{f}_{-1} + \lambda_0 \Gamma_1 \eta_0 & + & \boldsymbol{\lambda}_0 \boldsymbol{\eta}_1 & + & \mathbf{A}_1^2 \mathbf{v}_{-1} & + & A_1 u_0 & + & \mathbf{u}_1 \\ \\ x_2 &= \lambda_0 f_2 & + & v_2 \\ &= \lambda_0(\Gamma_1 f_1 + \boldsymbol{\eta}_2) & + & A_1 v_1 & + & \mathbf{u}_2 \\ &= \lambda_0 \Gamma_1 f_1 & + & \boldsymbol{\lambda}_0 \boldsymbol{\eta}_2 & + & A_1(A_1 v_0 + \mathbf{u}_1) & + & \mathbf{u}_2 \\ &= \lambda_0 \Gamma_1(\Gamma_1 f_0 + \boldsymbol{\eta}_1) & + & \boldsymbol{\lambda}_0 \boldsymbol{\eta}_2 & + & A_1^2 v_0 & + & \mathbf{A}_1 \mathbf{u}_1 & + & \mathbf{u}_2 \\ &= \vdots \\ &= \lambda_0 \Gamma_1^3 \mathbf{f}_{-1} & + & \lambda_0 \Gamma_1^2 \eta_0 & + & \lambda_0 \Gamma_1 \boldsymbol{\eta}_1 & + & \boldsymbol{\lambda}_0 \boldsymbol{\eta}_2 & + & \\ & & + & \mathbf{A}_1^2 \mathbf{v}_{-1} & + & A_1^2 u_0 & + & \mathbf{A}_1 \mathbf{u}_1 & + & \mathbf{u}_2 \\ \\ &= \vdots \end{aligned}$$

$$\begin{aligned}
x_i &= \lambda_0 f_i + v_i \\
&= \vdots \\
&= \lambda_0 \Gamma_1^{i+1} f_{-1} + \lambda_0 \Gamma_1^i \eta_0 + \lambda_0 \Gamma_1^{i-1} \eta_1 + \lambda_0 \Gamma_1^{i-2} \eta_2 + \cdots + \lambda_0 \Gamma_1 \eta_{i-1} + \lambda_0 \eta_i + \\
&\quad A_1^i v_{-1} + A_1^i u_0 + A_1^{i-1} u_1 + A_1^{i-2} u_2 + \cdots + A_1 u_{i-1} + u_i
\end{aligned}$$

A process with an impulse in the last dynamic factor and another impulse in the first series would thus be as follows:

$$x_i = \lambda_0 \Gamma_1^i \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} + A_1^i \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (4.3)$$

The factor impulse affects all variables and will be considered later.

4.4 Impulses in Both Residuals u_t and η_t with Transmission

Out of the previous section evolves (4.4), the impulse response with transmission function for time point $(t + i)$. The future residuals are assumed to be non-zero. The generalization to more lags has the same form, and it is derived in section 6.3 of the appendix.

$$\begin{aligned}
x_i &= \lambda_0 f_i + v_i \\
&= \vdots \\
&= \lambda_0 \Gamma_1^{i+1} f_{-1} + \lambda_0 \Gamma_1^i \eta_0 + \lambda_0 \Gamma_1^{i-1} \eta_1 + \lambda_0 \Gamma_1^{i-2} \eta_2 + \cdots + \lambda_0 \Gamma_1 \eta_{i-1} + \lambda_0 \eta_i + \\
&\quad A_1^i v_{-1} + A_1^i u_0 + A_1^{i-1} u_1 + A_1^{i-2} u_2 + \cdots + A_1 u_{i-1} + u_i
\end{aligned} \quad (4.4)$$

The two sum members $\lambda_0 \Gamma_1^{i+1} f_{-1}$ and $A_1^i v_{-1}$ belong to the past and are supposed as constant over all impulses. They cannot be ignored, the reason for which will be shown later.

In the following, the meta-residuals are used as future residuals in (4.4). A new perspective thus emerges: the meta-residuals are now considered impulses that have arisen through changes inside the system and not exogenously. Some important considerations are mentioned below.

First, there are responses without exogenous impulses, where the meta-residuals are assumed to be parts of impulses. In other words, the system evolves with edges and those edges have consequences. The difference between a version of the system with an exogenous impulse and a version without can be treated as an impulse response. This way of computing an impulse response may not be appealing, but the numerical computation of meta-residuals requires it.

Second, the responses are dependent on the state of the system. Therefore, the values of the past factors and idiosyncratic parts cannot be ignored, and an increase in the federal funds rate of 0.25 would be, for example, different if it occurs from 4.0 to 4.25 or from 0.75 to 1.0.

Further, it is not possible to set an impulse and equally divide it for every part (common component and idiosyncratic part), as assumed in the previous section. It is also not possible to set an impulse directly in a factor; rather, it must be set in the series¹³. An impulse in a series generates

contemporaneous responses in factors that instantaneously become impulses for all other series, further generating contemporaneous responses in those series. Finally, from the model assumptions, the matrices A_i are diagonal, and because of (3.6), the complete interaction can be tracked through meta-residuals.

A contemporaneous response of series j to an impulse in series i is outlined in residuals as follows:

$$\text{impulse}_i \begin{cases} \rightarrow \lambda_{0i}\eta \rightarrow \eta \rightarrow \lambda_{0j}\eta \rightarrow u_j \\ \rightarrow u_i \end{cases}$$

The arrows do not represent a sequential succession. An impulse is distributed instantaneously to the system. The distribution of the impulse occurs through extraction of the principal components of the series containing the impulse, and not through direct separation of the impulse itself.

Rather than setting an impulse to an individual series and generating responses from other series, an impulse is set to the system and generates a response from the entire system.

4.5 Empirical Results for Structural Analysis

In order to avoid combining residuals and meta-residuals in the last available period, the impulse is set in $(t + 1)$. Thus, the past values of the factors and idiosyncratic part of the terms $\lambda_0\Gamma_1^{i+1}f_{-1}$ and $A_1^iv_{-1}$ in (4.4) are the last available values. Setting the impulse in $(t + 1)$ also ensures that past data is not changed. However, setting the impulse in the last available period should lead to similar results.

An impulse of 0.25 is set in the federal funds rate in the following way. The value should be added to the forecasted series in $(t + 1)$. The demeaning, scaling, and transformation of the forecast value must first be inverted. The impulse is then added and the series is transformed, demeaned, and standardized again (according to the original transformation code and the original center and scale, such that the past is not changed). Thus, the new value corresponding to the series is computed, not the impulse itself. The latter is calculated as the corresponding difference. To simplify the notation, the corresponding difference is denominated delta-impulse. The difference between a response corresponding to (4.4) based only on meta-residuals and one also containing a delta-impulse is denominated delta-response¹⁴.

Table 5 shows the federal funds rate meta-residuals corresponding to (3.6), before and after setting the impulse, for the period in which the impulse is set [2016.04 as $(t + 1)$]. Equation (3.6) is not fulfilled in this period. The sum of the differences comes close to the delta-impulse, which, after the aforementioned transformations, is 0.6884857.

Table 5: Contemporaneous split of impulse between $\lambda_0\eta$ and u for federal funds rate.

Type	$\lambda_0\eta$	u	Sum
series with delta-impulse	-0.1120421	0.1874181	
series without delta-impulse	-0.1212405	-0.4503821	
difference	0.009198454	0.6378002	0.6469987

First, impulse responses based only on meta-residuals, without the exogenous impulse, are shown. The left part of Figure 5 depicts the common component and the idiosyncratic part of

the response (upper and lower parts of (4.4), respectively). The values are still transformed, and it is not clear how partial responses could be transformed back to levels. Because the data are transformed, demeaned and standardized, intuitive interpretation is difficult. For example, the inversion of demeaning and standardizing turns all values of the industrial production of consumer goods except $(t + 2)$ (2016.05, in the upper right subfigure) positive, such that the response transformed back to levels is an increase, as seen in Figure 6. The left figure for the industrial production of consumer goods corresponds to (3.6)¹⁵. The federal funds rate seems also to correspond to (3.6). However, the inflation responses do not correspond, and their sum is much shorter than the other two, ending after about 12 months (lower right subfigure).

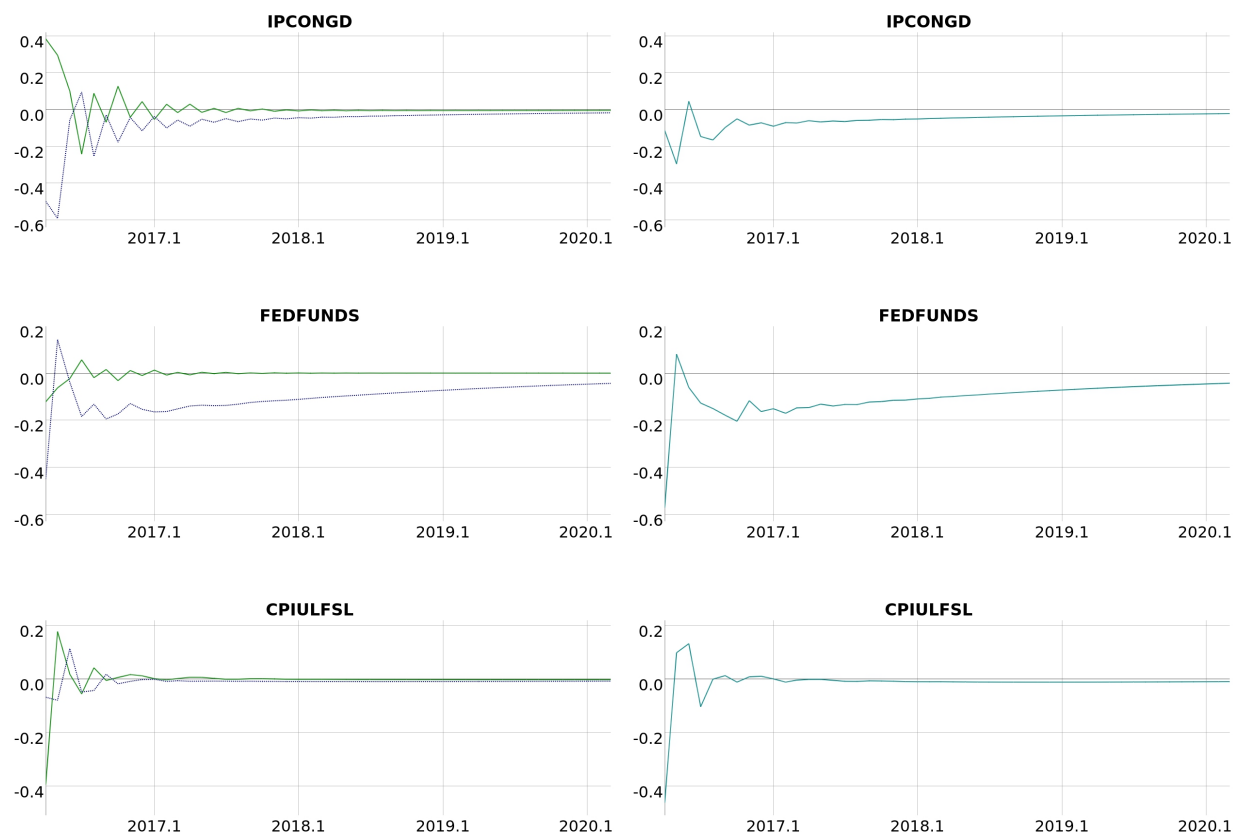


Figure 5: Responses based only on meta-residuals, without delta-impulses. On the left side, the common component (solid green) and the idiosyncratic parts (dotted blue) are shown, and the total response is depicted on the right as their sum. The y-axes on the left and right are the same, and the total impulse response can be seen transformed to levels in Figure 6. All data are future data, such that the y-axis corresponds to $(t + 1)$.

Next, Figure 6 shows the impulse responses based only on meta-residuals and based on meta-residuals with delta-impulses, as well as the delta-responses. The impulse has almost no effect on the industrial production of consumer goods, and the delta-response is correspondingly small. In the conventional impulse responses, based on VAR processes, the federal funds rate tends to go back to zero. Here, however, the federal funds rate remains on the new level.

The delta-response of the inflation series could contribute to a solution of the price puzzle found by Sims (1986). After a hike in the interest rate, conventional impulse responses show an

immediate increase in inflation, only subsequently followed by a decrease in inflation. Standard theory, intuition, and empirical evidence would predict an immediate decrease in inflation, with this contradiction being referred to as the price puzzle. The central bank cannot set the federal funds rate discretionary to any amount, thus zeroizing the meta-residuals. Because the meta-residuals retain influence, it is mandatory to take them into account by computing the delta-responses. The increase in inflation corresponding to the price puzzle cannot be seen in the delta-response of inflation¹⁶.

An impulse response is, in a sense, a forecast for an impulse, revealing the future effects of a current impulse. It concentrates on the contemporaneous lags, as can also be seen in the formulas. It is also interesting to examine the forecasts for inflation, computed based on meta-residuals in two variants: with and without delta-impulse. Figure 7 shows the two forecasts on the left, and their difference is shown in the right sub-figure. Both forecasts show an increase in the first periods, similar to the form of the price puzzle impulse response. A pseudo out-of-sample analysis of those forecasts combined with the analysis of the corresponding meta-residuals and delta-responses could clarify the price puzzle.

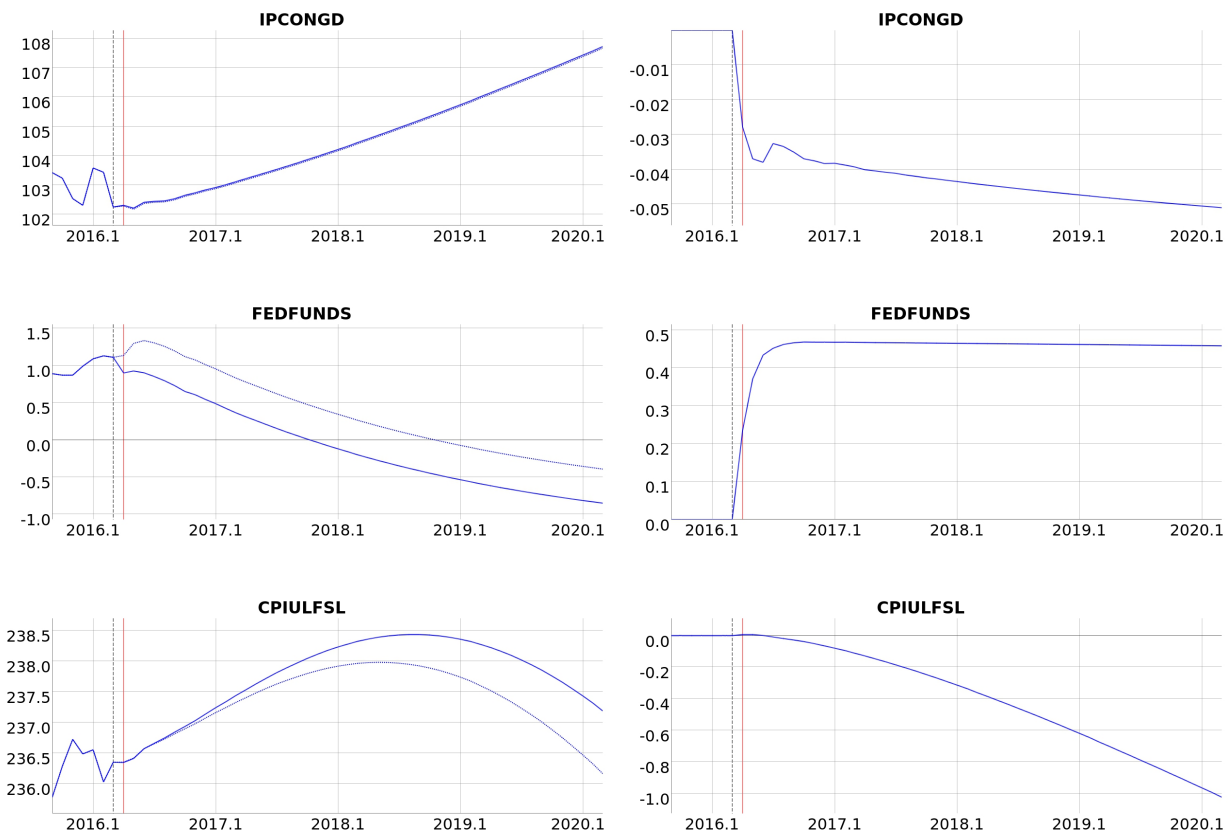


Figure 6: Left: impulse responses based only on meta-residuals (solid blue) and on meta-residuals with delta-impulses (dotted blue). Right: delta-responses. The y-axis differs in the left and right pictures, and all values are transformed back to levels. The impulse period is marked with a solid red line.

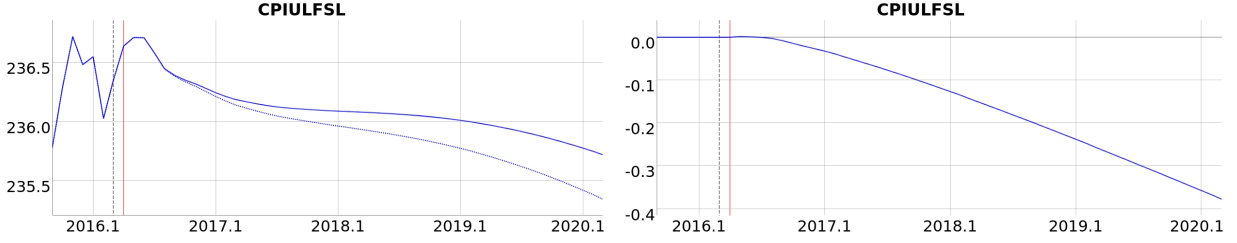


Figure 7: Inflation forecasts with (dotted) and without (solid) exogenous impulse in the federal funds rate (left) and their difference (right). The y-axes differ in the left and right pictures, and values are transformed back to levels. The impulse period is marked with a solid red line.

Figure 8 shows the split of the delta-responses in their common components and idiosyncratic parts. Such a clear delta-response in industrial production is a result of changes in the federal funds rate. The impulse is transmitted through tiny changes in the common component of the federal funds rate delta-response. This common component delta-response is changed and amplified twice through loadings λ_{0i} and λ_{0j} . The common component of series j receives the impulse from series i through $\lambda_{0i}\eta \rightarrow \eta \rightarrow \lambda_{0j}\eta$. The main changes generated by the interest rate are not $\lambda_{0i}\eta$ and $\lambda_{0j}\eta$ as shown in Figure 8, but rather η . The effects are shown as delta-responses of the factors in Figure 10. Furthermore, inverting the demeaning and standardization of the values shown in Figure 8 would create a more pronounced shape.

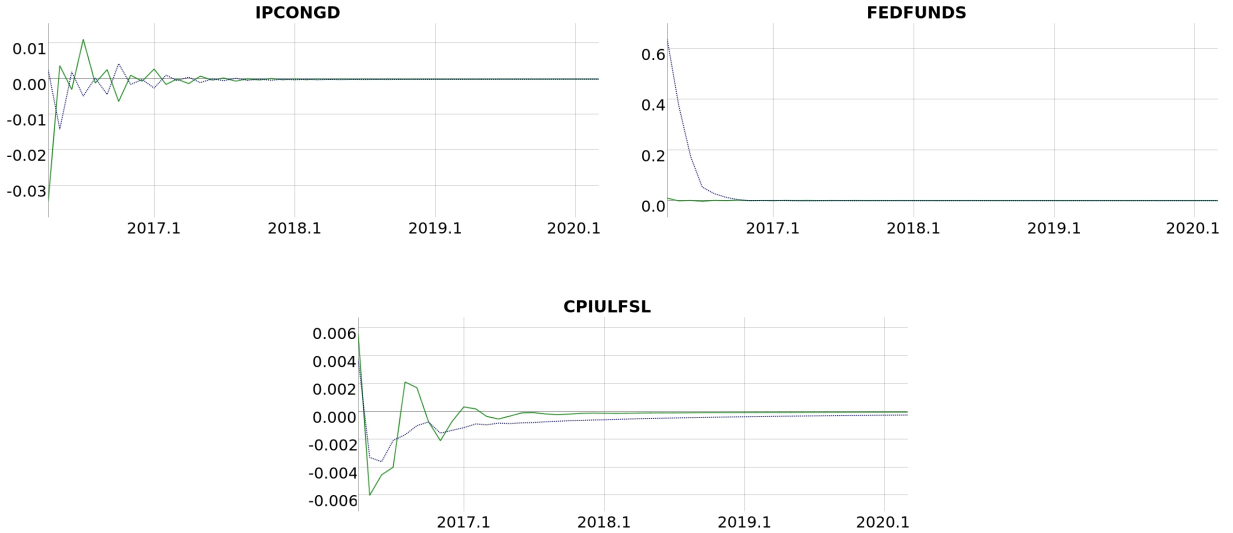


Figure 8: Common component (solid green) and idiosyncratic part (dotted blue) of delta-responses. Values are still transformed. All data are future data, such that the y-axis corresponds to $(t + 1)$.

Figure 9 shows the main object in the impulse transmission context, based only on meta-residuals, without an exogenous impulse. It is computed as

$$\Gamma_1^{i+1} \mathbf{f}_{-1} + \Gamma_1^i \eta_0 + \Gamma_1^{i-1} \eta_1 + \Gamma_1^{i-2} \eta_2 + \cdots + \Gamma_1 \eta_{i-1} + \eta_i. \quad (4.5)$$

As mentioned before, the meta-residuals are not responsible for the transmission, but rather arise due to the transmission, being somewhat like its shadow. This leads to the delta-responses of the

dynamic factors, depicted in Figure 10. They are computed as differences from (4.5) (version with delta-impulse minus version without delta-impulse). As aforementioned, the contemporaneous factor responses can be seen as impulses for the system that generate contemporaneous responses in all series other than federal funds rate. It is also worth noting that, because it adapts to the others, the third and weakest factor lives longest.

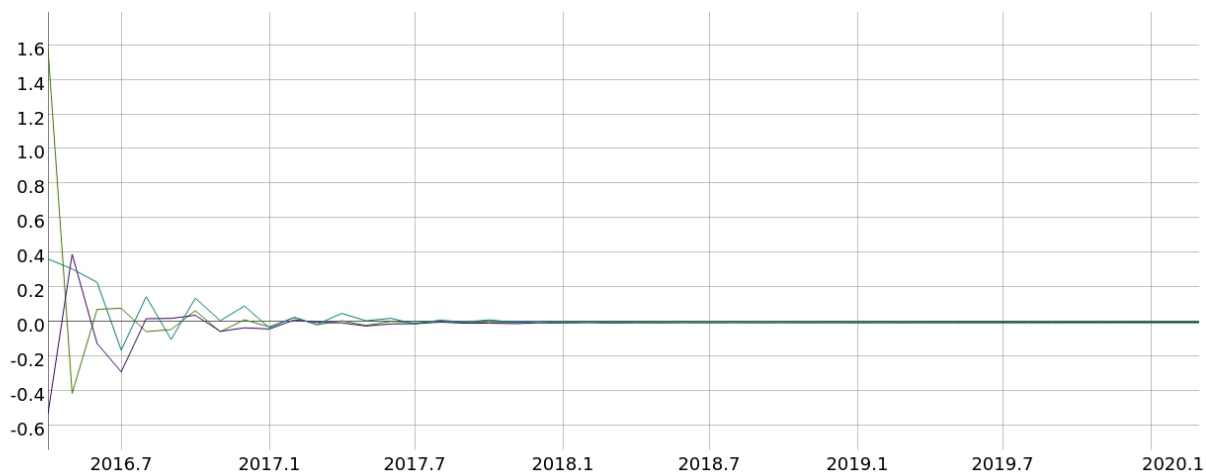


Figure 9: Recursive part of the impulse response of the common component (part without λ_0). All data are future data, such that the y-axis corresponds to $(t + 1)$. The object is based only on meta-residuals, without exogenous impulse.

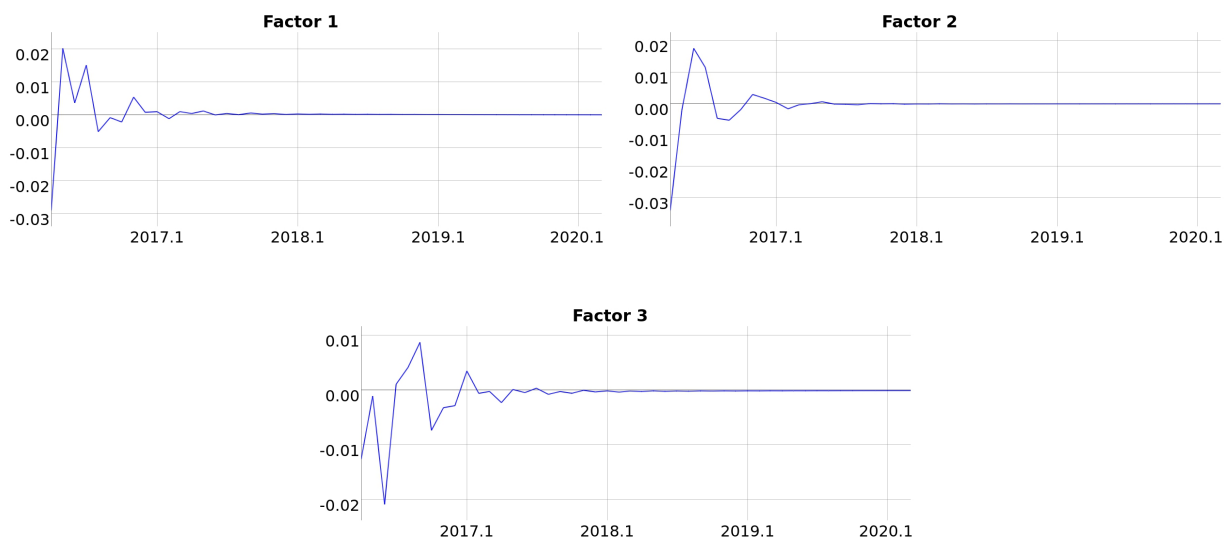


Figure 10: Delta-responses of the three dynamic factors. The values are still transformed. All data are future data, such that the y-axis corresponds to $(t + 1)$.

5 Conclusions

Extracting factors from forecasted data uncovers meta-residuals. These residuals can be used to improve forecasts and trace the impulse transmission in the system, and they can be interpreted as impulses that arise from within the system. The corresponding impulse responses need to be subtracted from responses created by exogenous impulses, in order to separate the effect of the latter. This separation requires the introduction of delta-responses. These delta-responses prove to be a possible solution for the price puzzle.

Future work needs to address the technical difficulties regarding the sign of meta-residuals. In such work, attention should focus on the system as a whole, such that all consequences of changing a sign are known and controlled.

Second, it would be interesting to explore whether the meta-residuals described here have relatives. In the much larger class of state-space or hidden Markov model that includes DFM (Stock and Watson 2016), what are the corresponding meta-residuals? Moreover, is it possible to extract factors from a second DFM (say, from similar data series from the European Union) and hyperfactors as "factors from factors" of those systems? Can hypermeta-residuals be useful for tracing the impulse transmission between the two economic spaces? Is it possible to connect DFM from different domains - for example, some market prices and climate? A long road remains between the methods described here and the impulse transmission between different systems.

Besides the aforementioned hierarchical analysis, changes could also be traced as responses of meta-residuals, as opposed to responses of factors or individual series, as shown in this article.

Finally, using different systems in which impulses can easily be controlled, possibly in bioinformatics, might bring to light changes in the structure of the DFM. If the common component goes toward zero, a regular VAR arises.

6 Appendix

6.1 Derivation of Forecast Formulas

The notation in (2.4) was slightly changed compared with Kilian and Lütkepohl (2017) since they are easier to use for forecasting formulas¹⁷.

$$X_{t+1} = \lambda(L)f_{t+1} + v_{t+1}$$

The common component is split into future lag 0, $\lambda_0 f_{t+1} = \lambda_0(\Gamma(L)f_t + \eta_{t+1})$, and present and past lags, $\sum_j \lambda_j f_{t-j+1}$.

$$X_{t+1} = \lambda_0(\Gamma(L)f_t + \eta_{t+1}) + \sum_j \lambda_j f_{t-j+1} + A(L)v_t + u_{t+1}$$

With

$$\sum_j \lambda_j f_{t-j+1} = L^{-1}(\lambda(L) - \lambda_0)f_t \text{ and } A(L)v_t = A(L)(X_t - \lambda(L)f_t)$$

we get

$$\begin{aligned} X_{t+1} &= \lambda_0(\Gamma(L)f_t + \eta_{t+1}) + L^{-1}(\lambda(L) - \lambda_0)f_t + A(L)(X_t - \lambda(L)f_t) + u_{t+1} \\ X_{t+1} &= \lambda_0\Gamma(L)f_t + \lambda_0\eta_{t+1} + L^{-1}(\lambda(L) - \lambda_0)f_t + A(L)X_t - A(L)\lambda(L)f_t + u_{t+1} \\ X_{t+1} &= \lambda_0\Gamma(L)f_t - A(L)\lambda(L)f_t + L^{-1}(\lambda(L) - \lambda_0)f_t + \lambda_0\eta_{t+1} + A(L)X_t + u_{t+1} \end{aligned}$$

and with

$$\alpha^f(L) = \lambda_0 \Gamma(L) - A(L) \lambda(L) + L^{-1}(\lambda(L) - \lambda_0)$$

we get

$$X_{t+1} = \alpha^f(L) f_t + \lambda_0 \eta_{t+1} + A(L) X_t + u_{t+1}.$$

Taking expectations delivers (3.1) :

$$E[X_{t+1} | X_t, f_t, X_{t-1}, f_{t-1} \dots] = \alpha^f(L) f_t + A(L) X_t. \quad (3.1)$$

After period $t + 1$, $\hat{f}^{extract}$ will be available and can be used for computing $\hat{f}^{forecast}$, and thus for forecasting,

$$\begin{aligned} E[X_{t+h+1} | X_{t+h}, f_{t+h}, X_{t+h-1}, f_{t+h-1} \dots] &= \alpha^f(L) f_{t+h}^{forecast} + A(L) X_{t+h} \\ &= \alpha^f(L) f_{t+h}^{forecast} + \lambda_0 \eta_{t+h+1}^{meta} + A(L) X_{t+h} + u_{t+h+1}^{meta} \end{aligned} \quad (3.5)$$

At the border between present and future, extracted factors from the past can be used for the lag polynomial $\alpha^f(L) f_{t+h}^{extract}$ when it is necessary for $\Gamma(L)$.

Repeating the same derivation without replacing $\sum_j \lambda_j f_{t-j+1} = L^{-1}(\lambda(L) - \lambda_0) f_t$ delivers

$$\begin{aligned} X_{t+1} &= \lambda_0 \Gamma(L) f_t - A(L) \lambda(L) f_t + \sum_j \lambda_j f_{t-j+1} + \lambda_0 \eta_{t+1} + A(L) X_t + u_{t+1} \\ E[X_{t+1} | X_t, f_t, X_{t-1}, f_{t-1} \dots] &= \beta^f(L) f_t + \sum_j \lambda_j f_{t-j+1} + A(L) X_t, \end{aligned}$$

where

$$\beta^f(L) = \lambda_0 \Gamma(L) - A(L) \lambda(L)$$

and

$$\beta^f(L) f_t + \sum_j \lambda_j f_{t-j+1} = (\lambda_0 \Gamma(L) - A(L) \lambda(L)) f_t + \sum_j \lambda_j f_{t-j+1} = \alpha^f(L) f_t$$

The $\beta^f(L)$ version seems easier to program, whereas the $\alpha^f(L)$ formula is more elegant. The code of the paper computes the Stock and Watson version with $\alpha^f(L)$.

6.2 Parameter Specification

In order to determine the correct parameters, it might be useful to keep in mind the order of the computation and hence the time order of setting the parameters. First, after choosing the series involved in the DFM - for example, by excluding some series like the oil price - the number of lags of the idiosyncratic part p and the number of static factors R are set and used within the same function: the factor extraction. Second, the number of dynamic factors r is set. Then, the lags of the common component, q^* , is chosen. An implicit parameter is the time period: the series are monthly data (and not, for example, quarterly data).

It helps also to take a look at (6.4), which shows that $R = r(q + 1)$ and since $q \leq (p + q^*)$ this leads to $R \leq r(p + q^* + 1)$. The inequality $q \leq (p + q^*)$ was deducted by multiplying out the matrix polynomial $\Lambda(L) = (I - A(L)) \lambda(L)$, which also leads, by assuming $p, q^* > 0$ to $q \geq (\max(p, q^*) + 1)$.

The number of lags of the idiosyncratic part, p , is set to 4 as in Stock and Watson (2005a), and it makes a big difference to $p = 6$.

The number of static factors R can be chosen by using information criteria. ICp_1 , ICp_2 , and ICp_3 from Bai and Ng (2002) lead to 8, 7, and 11, respectively, and the criterion from Otter, Jacobs, and Reijer (2014) leads to 26. Regarding the first set of criteria, McCracken and Ng (2016) mention that it is not important which criterion is used as long as the same criterion is kept throughout the analysis. $R = 7$ is chosen corresponding to ICp_1 .

Moreover, r is found to be 3 using different ways of computing it; for example, when using information criteria from Bai and Ng (2007) with all values for R from 4 to 30, $r = 3$ is always one of the results. It is also found to be 3 in a similar data set by Stock and Watson (2016) while using the Amenguel-Watson criterion.

Some final remarks follow. The ICp_1 from Bai and Ng (2002) coincides often in a recursive, pseudo out-of-sample computation with the local minimum value of the criterion by Otter, Jacobs, and Reijer (2014), which is smaller than 10. Examining the values of the information criteria in the recursive, pseudo out-of-sample computation (e.g., visually) will possibly lead to observation of jumps - those are changes in the number of dynamic factors r being shown as changes in the numbers of static factors R . The magnitude of those jumps will approximate q : since $R = r(q + 1)$, an increase in r by one leads to an increase in R by $(q + 1)$ (and thus being around 4 or 5 and not 1 or 12). Good results are obtained with $q = 4$, and the inequalities above are also fulfilled. Note that the main model involves a vector regression on the factor evolution (not a vector autoregression).

Computing the model with different lags for the factor evolution s leads to an almost identical result, such that $s = 1$ is chosen. Starting the regressions right after the structural break found by Stock and Watson (2009) in 1984.01 makes a big difference. Starting, for example, seven months earlier changes the results considerably (the first estimation periods affect all following periods).

Finally, if parameters are searched, for example, with a new data set, it is reasonable to assume that the journey will take long. Therefore, saving the extracted static factors "Fhat" (the code provides the ability to do so) helps to decrease the computation time for subsequent trials. In an pseudo out-of-sample forecast, this possibly leads to folders having data in the size of some gigabyte. A prior analysis for structural breaks might be necessary.

6.3 Generalization of Impulse Response Formula for q^* , s and p Lags

The idiosyncratic part of the impulse response in (4.4) is

$$A_1^i v_{-1} + A_1^i u_0 + A_1^{i-1} u_1 + A_1^{i-2} u_2 + \dots + A_1 u_{i-1} + u_i \quad (6.1)$$

Similar to Lütkepohl (2005), define

$$\mathbf{A} = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_r & 0 & \dots & 0 & 0 \\ 0 & I_r & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \dots & I_r & 0 \end{bmatrix}, \quad \mathbf{U}_t = \begin{bmatrix} u_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The regular impulse responses are set in u_0 , and the responses are given by

$$X_i = \mathbf{A}^i X_0$$

It can be seen that the regular impulse responses are the elements of the upper left-hand (K x K) block of A^i .

The impulse responses with meta-residuals for the idiosyncratic part are built in the very same way: the upper left matrix being raised to the power $i - k$ for each meta-residual as many times k as they already exist - namely $i - 0$ for u_0 , $i - 1$ for u_1 , $i - 2$ for u_2 , etc., leading to (6.1).

The common component part of the impulse response in (4.4) is

$$\lambda_0 \Gamma_1^{i+1} f_{-1} + \lambda_0 \Gamma_1^i \eta_0 + \lambda_0 \Gamma_1^{i-1} \eta_1 + \lambda_0 \Gamma_1^{i-2} \eta_2 + \cdots + \lambda_0 \Gamma_1 \eta_{i-1} + \lambda_0 \eta_i \quad (6.2)$$

For its derivation for general lags q^* consider the following representation (see also Kilian and Lütkepohl (2017)). Left multiplying model (2.1) with $(I - A(L))$ leads to $(I - A(L))x_t = \Lambda(L)f_t + u_t$, where $\Lambda(L) = (I - A(L))\lambda(L)$ is a matrix polynomial of order $q \leq p + q^*$.¹⁸ Assuming without loss of generality $q \geq s$, the model can be represented in static form as

$$(I - A(L))x_t = \Lambda F_t + u_t, \quad F_t = \Gamma F_{t-1} + G\eta_t \quad (6.3)$$

with similar notations as before $\Lambda = [\lambda_0, \lambda_1, \dots, \lambda_q]$, $F_t = (f'_t, \dots, f'_{t-q})'$ and

$$\Gamma = \begin{bmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_q & \Gamma_{q+1} \\ I_r & 0 & \cdots & 0 & 0 \\ 0 & I_r & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & I_r & 0 \end{bmatrix}, \quad G = \begin{bmatrix} I_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (6.4)$$

F_t are referred to as static factors, while f_t are designated as (primitive) dynamic factors. (6.2) can be derived similar to (6.1) from (6.3) and (6.4) and by noticing the selection matrix G and the first (r x r) element of Λ as being λ_0 . The sum of (6.2) and (6.1) results in (4.4).

6.4 Data Description

Monthly data in the Stock and Watson format are used, provided by FRED online at <https://research.stlouisfed.org/econ/mccracken/fred-databases/monthly/current.csv>.

The series "Help-Wanted Index for United States" (HWI) and "Ratio of Help Wanted/No. Unemployed" (HWIURATIO) are provided up to 2016.03, such that the sample ends there. It is not clear if this series will be continued with internet data, which better suits a modern economy. The series "ANDENOx," "ACOGNO," "TWEXMMTH," and "UMCSENTx" were not used since they have too many missing values, and "OILPRICEx" was not used due to lack of variation in the early part of the sample, similar to the proposal of McCracken and Ng (2016).

Data are grouped as: (1) output and income; (2) labor market; (3) housing; (4) consumption, orders, and inventories; (5) money and credit; (6) interest and exchange rates; (7) prices; and (8) stock market. The resulting data set has no missing values; it starts in 1962.01 and ends in 2016.03.

The data are preprocessed in three steps:

1. Outliers adjustment as in Stock and Watson (2005b) is done to all series, once at the beginning of the analysis (and not, e.g., recursively at every step during the pseudo out-of-sample forecast). Observations with absolute median deviations larger than six times the inter-quartile range are replaced with the median value of the preceding five observations.

2. The series are transformed according to the following codes, mentioned in the last column: (1) no transformation, (2) $(1 - L)x_t$, (3) $(1 - L)^2x_t$, (4) $\log(x_t)$, (5) $(1 - L)\log(x_t)$, (6) $(1 - L)^2\log(x_t)$, and (7) $(1 - L)(x_t/Lx_t - 1)$.

3. Data are demeaned and standardized.

Steps 3 and 2 are inversely processed at the end, after the forecast and impulse response calculation.

Series tagged with an asterisk have been adjusted by FRED and thus differ from the series from the source; see McCracken and Ng (2016) for details.

Group ID	FRED ID	FRED mnemonics	Description	Transf.
1	1	RPI	Real Personal Income	5
1	2	W875RX1	Real personal income ex transfer receipts	5
1	6	INDPRO	IP Index	5
1	7	IPFPNSS	IP: Final Products and Nonindustrial Supplies	5
1	8	IPFINAL	IP: Final Products (Market Group)	5
1	9	IPCONGD	IP: Consumer Goods	5
1	10	IPDCONGD	IP: Durable Consumer Goods	5
1	11	IPNCONGD	IP: Nondurable Consumer Goods	5
1	12	IPBUSEQ	IP: Business Equipment	5
1	13	IPMAT	IP: Materials	5
1	14	IPDMAT	IP: Durable Materials	5
1	15	IPNMAT	IP: Nondurable Materials	5
1	16	IPMANSICS	IP: Manufacturing (SIC)	5
1	17	IPB51222S	IP: Residential Utilities	5
1	18	IPFUELS	IP: Fuels	5
1	19	NAPMPI	ISM Manufacturing: Production Index	1
1	20	CUMFNS	Capacity Utilization: Manufacturing	2
2	21*	HWI	Help-Wanted Index for United States	2
2	22*	HWIURATIO	Ratio of Help Wanted/No. Unemployed	2
2	23	CLF16OV	Civilian Labor Force	5
2	24	CE16OV	Civilian Employment	5
2	25	UNRATE	Civilian Unemployment Rate	2
2	26	UEMPMEAN	Average Duration of Unemployment (Weeks)	2
2	27	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	5
2	28	UEMP5TO14	Civilians Unemployed for 5-14 Weeks	5
2	29	UEMP15OV	Civilians Unemployed - 15 Weeks & Over	5
2	30	UEMP15T26	Civilians Unemployed for 15-26 Weeks	5
2	31	UEMP27OV	Civilians Unemployed for 27 Weeks and Over	5
2	32*	CLAIMSx	Initial Claims	5
2	33	PAYEMS	All Employees: Total nonfarm	5
2	34	USGOOD	All Employees: Goods-Producing Industries	5
2	35	CES1021000001	All Employees: Mining and Logging: Mining	5
2	36	USCONS	All Employees: Construction	5
2	37	MANEMP	All Employees: Manufacturing	5
2	38	DMANEMP	All Employees: Durable goods	5
2	39	NDMANEMP	All Employees: Nondurable goods	5
2	40	SRVPRD	All Employees: Service-Providing Industries	5
2	41	USTPU	All Employees: Trade, Transportation & Utilities	5
2	42	USWTRADE	All Employees: Wholesale Trade	5
2	43	USTRAD	All Employees: Retail Trade	5
2	44	USFIRE	All Employees: Financial Activities	5
2	45	USGOVT	All Employees: Government	5
2	46	CES0600000007	Avg Weekly Hours : Goods-Producing	1
2	47	AWOTMAN	Avg Weekly Overtime Hours : Manufacturing	2
2	48	AWHMAN	Avg Weekly Hours : Manufacturing	1
2	49	NAPMEI	ISM Manufacturing: Employment Index	1
2	127	CES0600000008	Avg Hourly Earnings : Goods-Producing	6
2	128	CES2000000008	Avg Hourly Earnings : Construction	6
2	129	CES3000000008	Avg Hourly Earnings : Manufacturing	6
3	50	HOUST	Housing Starts: Total New Privately Owned	4
3	51	HOUSTNE	Housing Starts, Northeast	4
3	52	HOUSTMW	Housing Starts, Midwest	4

3	53	HOUSTS	Housing Starts, South	4
3	54	HOUSTW	Housing Starts, West	4
3	55	PERMIT	New Private Housing Permits (SAAR)	4
3	56	PERMITNE	New Private Housing Permits, Northeast (SAAR)	4
3	57	PERMITMW	New Private Housing Permits, Midwest (SAAR)	4
3	58	PERMITS	New Private Housing Permits, South (SAAR)	4
3	59	PERMITW	New Private Housing Permits, West (SAAR)	4
4	3	DPCERA3M086SBEA	Real personal consumption expenditures	5
4	4*	CMRMTSPLx	Real Manu. and Trade Industries Sales	5
4	5*	RETAILx	Retail and Food Services Sales	5
4	60	NAPM	ISM : PMI Composite Index	1
4	61	NAPMNOI	ISM : New Orders Index	1
4	62	NAPMSDI	ISM : Supplier Deliveries Index	1
4	63	NAPMII	ISM : Inventories Index	1
4	64	ACOGNO	New Orders for Consumer Goods	5
4	65*	AMDMNOx	New Orders for Durable Goods	5
4	66*	ANDENOx	New Orders for Nondefense Capital Goods	5
4	67*	AMDMUOx	Unfilled Orders for Durable Goods	5
4	68*	BUSINVx	Total Business Inventories	5
4	69*	ISRATIOx	Total Business: Inventories to Sales Ratio	2
4	130*	UMCSENTx	Consumer Sentiment Index	2
5	70	M1SL	M1 Money Stock	6
5	71	M2SL	M2 Money Stock	6
5	72	M2REAL	Real M2 Money Stock	5
5	73	AMBSL	St. Louis Adjusted Monetary Base	6
5	74	TOTRESNS	Total Reserves of Depository Institutions	6
5	75	NONBORRES	Reserves Of Depository Institutions	7
5	76	BUSLOANS	Commercial and Industrial Loans	6
5	77	REALLN	Real Estate Loans at All Commercial Banks	6
5	78	NONREVSL	Total Nonrevolving Credit	6
5	79*	CONSPI	Nonrevolving consumer credit to Personal Income	2
5	131	MZMSL	MZM Money Stock	6
5	132	DTCOLNVHFM	Consumer Motor Vehicle Loans Outstanding	6
5	133	DTCTHFM	Total Consumer Loans and Leases Outstanding	6
5	134	INVEST	Securities in Bank Credit at All Commercial Banks	6
6	84	FEDFUNDS	Effective Federal Funds Rate	2
6	85*	CP3Mx	3-Month AA Financial Commercial Paper Rate	2
6	86	TB3MS	3-Month Treasury Bill:	2
6	87	TB6MS	6-Month Treasury Bill:	2
6	88	GS1	1-Year Treasury Rate	2
6	89	GS5	5-Year Treasury Rate	2
6	90	GS10	10-Year Treasury Rate	2
6	91	AAA	Moody's Seasoned Aaa Corporate Bond Yield	2
6	92	BAA	Moody's Seasoned Baa Corporate Bond Yield	2
6	93*	COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS	1
6	94	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS	1
6	95	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS	1
6	96	T1YFFM	1-Year Treasury C Minus FEDFUNDS	1
6	97	T5YFFM	5-Year Treasury C Minus FEDFUNDS	1
6	98	T10YFFM	10-Year Treasury C Minus FEDFUNDS	1
6	99	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS	1
6	100	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS	1
6	101	TWEXMMTH	Trade Weighted U.S. Dollar Index: Major Currencies	5
6	102*	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate	5
6	103*	EXJPUSx	Japan / U.S. Foreign Exchange Rate	5
6	104*	EXUSUKx	U.S. / U.K. Foreign Exchange Rate	5
6	105*	EXCAUSx	Canada / U.S. Foreign Exchange Rate	5
7	106	WPSFD49207	PPI: Finished Goods	6
7	107	WPSFD49502	PPI: Finished Consumer Goods	6
7	108	WPSID61	PPI: Intermediate Materials	6
7	109	WPSID62	PPI: Crude Materials	6
7	110*	OILPRICEx	Crude Oil, spliced WTI and Cushing	6
7	111	PPICMM	PPI: Metals and metal products:	6
7	112	NAPMPRI	ISM Manufacturing: Prices Index	1
7	113	CPIAUCSL	CPI : All Items	6
7	114	CPIAPPSL	CPI : Apparel	6
7	115	CPITRNSL	CPI : Transportation	6
7	116	CPIMEDSL	CPI : Medical Care	6

7	117	CUSR0000SAC	CPI : Commodities	6
7	118	CUUR0000SAD	CPI : Durables	6
7	119	CUSR0000SAS	CPI : Services	6
7	120	CPIULFSL	CPI : All Items Less Food	6
7	121	CUUR0000SA0L2	CPI : All items less shelter	6
7	122	CUSR0000SA0L5	CPI : All items less medical care	6
7	123	PCEPI	Personal Cons. Expend.: Chain Index	6
7	124	DDURRG3M086SBEA	Personal Cons. Exp: Durable goods	6
7	125	DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods	6
7	126	DSERRG3M086SBEA	Personal Cons. Exp: Services	6
8	80*	S&P 500	S&P's Common Stock Price Index: Composite	5
8	81*	S&P: indust	S&P's Common Stock Price Index: Industrials	5
8	82*	S&P div yield	S&P's Composite Common Stock: Dividend Yield	2
8	83*	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio	5
8	135*	VXOCLSx	VXO	1

Acknowledgment

This article would not have been possible without the FRED dataset, and I would like to thank data specialists at FRED for taking care of revisions and data changes in the macroeconomic US dataset FRED-MD, and the Federal Reserve Bank of St. Louis for providing data open to the public.

I would also like to thank the R community for creating and maintaining the software at no charge.

References

- Bai, Jushan and Serena Ng (2002). “Determining the Number of Factors in Approximate Factor Models”. In: *Econometrica* 70.1, pp. 191–221. ISSN: 1468-0262. DOI: [10.1111/1468-0262.00273](https://doi.org/10.1111/1468-0262.00273).
- (2007). “Determining the number of primitive shocks in factor models”. In: *Journal of Business & Economic Statistics* 25.1. DOI: [10.1198/073500106000000413](https://doi.org/10.1198/073500106000000413).
- Hallin, Marc and Roman Liška (2011). “Dynamic factors in the presence of blocks”. In: *Journal of Econometrics* 163.1, pp. 29–41. DOI: [10.1016/j.jeconom.2010.11.004](https://doi.org/10.1016/j.jeconom.2010.11.004).
- Kilian, Lutz and Helmut Lütkepohl (2017). *Structural Vector Autoregressive Analysis*. Cambridge University Press, forthcoming.
- Lütkepohl, Helmut (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.
- McCracken, Michael W. and Serena Ng (2016). “FRED-MD: A monthly database for macroeconomic research”. In: *Journal of Business & Economic Statistics* 34.4, pp. 574–589. DOI: [10.1080/07350015.2015.1086655](https://doi.org/10.1080/07350015.2015.1086655).
- Moench, Emanuel, Serena Ng, and Simon Potter (2013). “Dynamic hierarchical factor models”. In: *Review of Economics and Statistics* 95.5, pp. 1811–1817. DOI: [10.1162/REST_a_00359](https://doi.org/10.1162/REST_a_00359).
- Otter, Pieter W., Jan P.A.M. Jacobs, and Ard H.J. Reijer (2014). *A criterion for the number of factors in a data-rich environment*. University of Groningen, Faculty of Economics and Business.
- Sims, Christopher A. (1980). “Macroeconomics and reality”. In: *Econometrica: Journal of the Econometric Society*, pp. 1–48. DOI: [10.2307/1912017](https://doi.org/10.2307/1912017).
- (1986). “Are forecasting models usable for policy analysis?” In: *Quarterly Review* Win, pp. 2–16.
- Stock, James H. and Mark W. Watson (2005a). “An empirical comparison of methods for forecasting using many predictors”. In: *Manuscript, Princeton University*.
- (2005b). *Implications of dynamic factor models for VAR analysis*. Tech. rep. National Bureau of Economic Research.
- (2009). “Forecasting in dynamic factor models subject to structural instability”. In: *The Methodology and Practice of Econometrics. A Festschrift in Honour of David F. Hendry* 173, p. 205. DOI: [10.1093/acprof:oso/9780199237197.001.0001](https://doi.org/10.1093/acprof:oso/9780199237197.001.0001).
- (2016). “Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics”. In: *Handbook of Macroeconomics*. Ed. by JB Taylor and H Uhlig. Vol. 2. Elsevier, pp. 415–525. DOI: [10.1016/bs.hesmac.2016.04.002](https://doi.org/10.1016/bs.hesmac.2016.04.002).

References: R Packages

- Bates, Douglas and Dirk Eddelbuettel (2013). “Fast and Elegant Numerical Linear Algebra Using the RcppEigen Package”. In: *Journal of Statistical Software* 52.5, pp. 1–24. URL: <http://www.jstatsoft.org/v52/i05/>.
- Hyndman, Rob J. and Han Lin Shang (2016). *ftsa: Functional Time Series Analysis*. R package version 4.7. URL: <http://cran.r-project.org/package=ftsa>.
- Novomestky, Frederick (2012). *matrixcalc: Collection of functions for matrix calculations*. R package version 1.0-3. URL: <http://CRAN.R-project.org/package=matrixcalc>.
- Pascu, Sorin T. (2017). *dfm.SVAR: Dynamic Factor Models and SVAR Analysis*. version 0.1.0. Appstam Consulting GmbH. Berlin, European Union. Unpublished package.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <http://www.R-project.org/>.
- Ryan, Jeffrey A. and Joshua M. Ulrich (2014). *xts: eXtensible Time Series*. R package version 0.9-7. URL: <http://CRAN.R-project.org/package=xts>.
- Shang, Han Lin (2013). “ftsa: An R Package for Analyzing Functional Time Series”. In: *The R Journal* 5.1, pp. 64–72. URL: <http://journal.r-project.org/archive/2013-1/shang.pdf>.
- Temple Lang, Duncan and the CRAN team (2016). *RCurl: General Network (HTTP/FTP/...) Client Interface for R*. R package version 1.95-4.8. URL: <https://CRAN.R-project.org/package=RCurl>.
- Tsay, Ruey S. (2015). *MTS: All-Purpose Toolkit for Analyzing Multivariate Time Series (MTS) and Estimating Multivariate Volatility Models*. R package version 0.33. URL: <http://CRAN.R-project.org/package=MTS>.
- Vanderkam, Dan et al. (2017). *dygraphs: Interface to ‘Dygraphs’ Interactive Time Series Charting Library*. R package version 1.1.1.4. URL: <http://CRAN.R-project.org/package=dygraphs>.

Notes

¹Federal Reserve Economic Data Monthly Database for Macroeconomic Research provided by the Federal Reserve Bank of St. Louis

²Kilian and Lütkepohl (2017) define the filters in left form $A(L) = I - A_1L - \dots$ and $\Gamma(L) = I - \Gamma_1L - \dots$, which is changed in the notation here since it is easier to use for the forecasts.

³If the true coefficients were known, one could obtain the factors by OLS, without using principal component extraction.

⁴Regressing the common component $\hat{\Lambda}\hat{F}$ on the newly obtained dynamic factors $\hat{f}$ would seem more appealing, but delivers poorer impulse transmission results later in the analysis.

⁵Note that the real process does not change the past factors but rather the loading estimates. Adding data from January 1990 to the factor estimation from December 1989 leads to better estimates of the loadings and thus changes the factor estimates.

⁶Step 3 of the initial factor extraction starts with OLS for $\hat{\Lambda}$ and ends with PCA extraction for $\hat{F}$. Using OLS for $\hat{F}$ (and PCA for $\hat{\Lambda}$) would make the two versions of the first dynamic factor identical for past values.

⁷Greek for transmission: " $\mu\epsilon\tau\acute{\alpha}\delta\omicron\sigma\eta$ "

⁸Note that η_{t+1}^{meta} is not $E[\eta_{t+1}]$.

⁹These kinds of accidental drops can be valuable if they occur in the correct period and predict large recessions.

¹⁰The series "Consumer Price Index for All Urban Consumers: All Items Less Food and Energy" (CPILFESL) is not included in the data set, and "Consumer Price Index for All Urban Consumers: All Items Less Food" (CPIULFSL) is taken instead.

¹¹A transformation back to levels for factors is not possible since different transformation codes are involved.

¹²There are two principal component extractions: in the estimation of the static factors $\hat{F}$ and in the extraction of the dynamic factors $\hat{f}$ out of $\hat{F}$.

¹³To set an impulse directly in a factor is a technical exercise that might reveal more about system behavior, but is not done here.

¹⁴The code used in this paper allows the setting of multiple impulses in different periods; for example, an oil price shock in $(t + 4)$ as a replacement impulse (set to 100 USD) and a federal funds rate in $(t + 6)$ as a delta-impulse (subtract 0.25).

¹⁵Note that the minus from $-u_{t+h}^{meta}$ is missing here.

¹⁶The responses are reported in the differences of levels (per period), such that a tiny increase in the delta-response of core inflation seems negligible.

¹⁷ The lag polynomials $\lambda(L)$, $\Psi(L)$, and $\delta(L)$ in the Stock and Watson (2016) notation correspond here to $\lambda(L)$, $\Gamma(L)$, and $A(L)$, and the residuals η and v correspond to η and u . The idiosyncratic dynamic is denoted by e , which corresponds here to v . j denotes the lags of $\lambda(L)$. Thus, λ_j is the matrix of lag j , and λ_0 is the first and contemporaneous matrix.

¹⁸ Expanding the matrix polynomial leads to the sum of the two orders p and q^* .