Eigenvalue ratio estimators for the number of dynamic factors*

Marco Avarucci
University of Glasgow
Maddalena Cavicchioli†
University of Modena and Reggio Emilia
Mario Forni‡
Università di Modena e Reggio Emilia, CEPR and RECent
Paolo Zaffaroni
Imperial College London and Università di Roma La Sapienza

Abstract. We introduce three estimators for the number of common shocks of a large factor model. Two of them, the Dynamic Eigenvalue Ratio (DER) and the Dynamic Growth Ratio (DGR) are the dynamic counterparts of the eigenvalue ratio estimators (ER and GR) proposed by Ahn and Horenstein (2013). The third, the Dynamic eigenvalue Difference Ratio (DDR), is closely related to the test statistic proposed by Onatski (2009). The advantage of such estimators is that they do not require preliminary determination of discretionary parameters. In addition, they can be applied to single frequencies or frequency bands of interest. We prove consistency of such estimators and evaluate their performance under simulation. We conclude that DDR is preferable to DER and DGR and is a safer and excellent alternative to the best existing criteria. We apply DDR to the FRED-QD US data set and find that the US macro economy is driven by two main shocks. One of them has the features of a demand shock and the other of a supply shock. The demand shock explains most of the cyclical fluctuations of the main macroeconomic aggregates.

Keywords: Large dynamic factor models, dynamic principal components, dynamic factors.

JEL Classification: C01, C13, C38.

*The authors thank Marco Lippi, Alexei Onatski, Marc Hallin, Roman Liška and Alex Horenstein for sharing codes and providing useful suggestions.
†Research supported by the PRIN-MIUR Grant 2010J3LZEN and FAR 2014.
‡Research supported by the PRIN-MIUR Grant 20174ESLX7.

1 Introduction

In the last twenty years, large dynamic factor models have been applied successfully to the analysis of big panels of time series. Such data-rich environments are a common feature of macroeconomics and finance, where a few common shocks drive the comovements of many variables, so that information is scattered through a large number of interrelated series.[1]

In large factor models, the number of variables is infinite, and the observed data are thought of as an (n, T)-dimensional realization of the double-indexed process x_{it}, $i = 1, \ldots, \infty$, $t = -\infty, \ldots, \infty$.[2] Each variable x_{it} is decomposed into the sum of two unobservable components, the "common component", call it χ_{it}, and the "idiosyncratic component", say e_{it}. The idiosyncratic components are poorly correlated across sections. By contrast, the common components are driven by a small number q of unobservable shocks, say f_{jt}, $j = 1, \ldots, q$, which are the same for all cross-sectional units. These shocks, often called "dynamic factors", are loaded through one-sided linear filters, or impulse-response functions, $\lambda_{ij}(L)$, where L is the lag operator. A restriction which we shall not impose in this paper, but is often assumed in the literature, is that the common components are *contemporaneous* linear combinations of $r \geq q$ unobservable variables F_{kt}, $k = 1, \ldots, r$, often called "static factors". In such a case, we say that the model admits a static factor representation; the dynamic nature of the model comes from the fact that the static factors have a dynamic representation in the common shocks.

Estimating the number of common shocks is of course a crucial preliminary step for the estimation of large dynamic factor models. But in several situations it is also interesting in itself. In finance, the number of shocks can be interpreted as the number of sources of nondiversifiable risk. In consumer demand theory, the number of factors in budget share data provides crucial information about the demand system (Lewbel (1991)). In macroeconomics, determining the number of structural shocks can provide a useful guidance for theoretical modeling.

Our empirical application concerns macroeconomics. RBC models postulate the existence of a single supply shock that drives variables related to real economic activity (Kydland and

[1]Early theoretical contributions are Forni and Reichlin (1998), Forni, Hallin, Lippi, and Reichlin (2000, 2005), Forni and Lippi (2001), Stock and Watson (2002b), Bai and Ng (2002, 2007), Bai (2003). A partial list of early applications include forecasting (Stock and Watson (2002a,b), Marcellino, Stock, and Watson (2003), Boivin and Ng (2006), D'Agostino and Giannone (2012), structural macroeconomic analysis (Bernanke and Boivin (2003), Bernanke, Boivin, and Eliasz (2005), Favero, Marcellino, and Neglia (2005), Eickmeier (2007), Forni, Giannone, Lippi, and Reichlin (2009), Forni and Gambetti (2010)), nowcasting and business cycle indicators (Forni and Lippi (2001), Cristadoro, Forni, Reichlin, and Veronese (2005), Giannone, Reichlin, and Small (2008), Altissimo, Cristadoro, Forni, and Lippi (2010)), the analysis of financial markets (Corielli and Marcellino (2006), Ludvigson and Ng (2007, 2009), Hallin, Mathias, Pirotte Speder, and Veredas (2011)).

[2]The main feature distinguishing large approximate dynamic factor models from traditional dynamic factor models (Sargent and Sims (1977), Geweke (1977)) is the fact that the idiosyncratic components are not necessarily orthogonal to each other. This important generalization has the consequence that common and idiosyncratic components are no longer conceptually distinguishable to each other if the cross-sectional dimension is finite. This motivates the assumption of an infinite number of variables.

Prescott (1982)). In contrast, modern DSGE models feature several shocks: the one of Smets and Wouters (2007), for instance, has even seven. Recently, Angeletos, Collard, and Dellas (2020) argue that there could be just one shock, a non-inflationary demand shock, explaining the bulk of business cycle fluctuation of real macroeconomic variables. How many shocks are there in the macro economy? The answer to these important question can be given by estimating the number of dynamic factors within a factor model.

Existing criteria are not entirely satisfactory. The methods proposed by Bai and Ng (2007) and Amengual and Watson (2007) assume that the factor model can be written in the static form, a restriction that reduces their scope. Moreover, they require the preliminary determination of the number of static factors. This first step can be performed using different methods, which can provide different outcomes, conditioning the final result. Hallin and Liska (2007) propose an estimator (HL henceforth) which entails minimization of a loss function which includes a penalty term. The use of penalty functions, while common in this field, is problematic: several functional forms can be used, and each one of them can be multiplied by an arbitrary constant, which calls for calibration. The Authors propose an ingenious and effective method to calibrate their penalty function.[3] The method, however, requires evaluation of the loss function over a grid n_j, T_j, $j = 1, \dots, J$, and the outcome is sensitive to the choice of such a grid.[4] Onatski (2009) proposes a test for the null of $q = k$ against the alternative of $k < q \leq k_{\max}$. The test can be used sequentially as a device to estimate q; however, the procedure proposed in the paper (O henceforth) requires preliminary choices which are discretionary to some extent and may be a source of error.

In this paper we study three new criteria to determine the number of dynamic factors, that do not present any of these problems.

Our first two estimators are the dynamic equivalents of two criteria proposed by Ahn and Horenstein (2013) to estimate the number of static factors, i.e. the "Eigenvalue Ratio" and the "Growth Ratio". In analogy with these denominations, we call these new estimators the "Dynamic Eigenvalue Ratio" (DER) and the "Dynamic Growth Ratio" (DGR).

Let $\widehat{\mathbf{\Sigma}}_n(\omega_\ell)$ be a suitable estimate[5] of the spectral density matrix of the x's at the Fourier frequencies $\omega_\ell = 2\pi\ell/T$, $\ell = 0, 1, \dots, T-1$. Moreover, let $\hat{\mu}_{nk}(\omega_\ell)$ be the $k-$th eigenvalue of $\widehat{\mathbf{\Sigma}}_n(\omega_\ell)$ in decreasing order of magnitude. Finally, let $\hat{\mu}_{nk}$ be the average of these eigenvalues across frequencies: $\hat{\mu}_{nk} = \sum_{\ell=0}^{T-1} \hat{\mu}_{nk}(\omega_\ell)/T$. We call $\hat{\mu}_{nk}$, $k = 1, \dots, k_{\max}$, the "dynamic eigenvalues", to distinguish them from the "static eigenvalues", i.e. the eigenvalues of the variance-covariance matrix of the variables, used by Ahn and Horenstein (2013) for the number of static factors.

[3] A similar method is used also by Alessi, Barigozzi, and Capasso (2010) to determine the number of static factors.

[4] Moreover, the result is dependent on the ordering of the variables in the data set. To solve this problem, the procedure requires that the variables are reordered randomly; but this causes the result to contain a stochastic element.

[5] We consider in this paper a periodogram smoothing estimate with rectangular window.

The DER is simply the ratio of two adjacent dynamic eigenvalues, i.e

$$\mathrm{DER}_n^T(k) = \hat{\mu}_{nk}/\hat{\mu}_{n,k+1}$$

and k_{DER} is defined as the value of k maximizing $\mathrm{DER}_n^T(k)$ for k varying between 1 and a maximum value $k_{\max}$ specified by the researcher. The definition of $\mathrm{DGR}_n^T(k)$ and k_{DGR} is somewhat more involved and will be given below. In this paper we prove that k_{DER} and k_{DGR} converge in probability to q as n and T go to infinity at the same rate (provided that $k_{\max} \geq q$).

Our third estimator is closely related to the test statistic proposed by Onatski (2009), i.e.

$$\frac{\hat{\mu}_{nk}(\omega) - \hat{\mu}_{n,k+1}(\omega)}{\hat{\mu}_{n,k+1}(\omega) - \hat{\mu}_{n,k+2}(\omega)},$$

where ω is a frequency of interest. In the present paper we define the "Dynamic eigenvalue Difference Ratio" (DDR) as the ratio above, where $\hat{\mu}_{nk}(\omega)$ is replaced by $\hat{\mu}_{nk}$:

$$\mathrm{DDR}_n^T(k) = \frac{\hat{\mu}_{nk} - \hat{\mu}_{n,k+1}}{\hat{\mu}_{n,k+1} - \hat{\mu}_{n,k+2}}.$$

In words, we do not consider a specific frequency, but the average over all frequencies. Our estimator of q, k_{DDR}, is given by the value of k maximizing DDR.[6] DDR has a few advantages with respect to O. First, by averaging across frequencies, we do not have to choose a specific frequency. Second, we do not need to combine in some way the potentially different results obtained at different frequencies. Finally, not having a test, we also avoid the need to choose a significance level either. We prove that k_{DDR} converge in probability to q as n and T go to infinity at the same rate.

The only one "nuisance parameter" of DDR, DER and DGR is the size of the smoothing window used for estimation of the spectral density matrix. Hence, DDR, DER and DGR are, so to speak, "safer" than HL and O, which, besides the window size, require additional preliminary choices which might be sources of error.

To provide an intuition about our estimators, let us recall an important result in factor model theory. A dynamic factor structure is characterized by the behavior of the eigenvalues of the spectral density matrix $\Sigma_n(\omega)$ of the first n variables, as $n \to \infty$ (Forni et al. (2000); Forni and Lippi (2001)). The q largest eigenvalues diverge, whereas the others are bounded.[7] We show here that, under suitable assumptions, similar properties hold for the sample analogue of such eigenvalues, i.e. $\hat{\mu}_{nk}(\omega)$. The first q eigenvalues, properly normalized, diverge at the

[6] Indeed, in order to avoid denominators very close to 0, we 'correct' the denominator by taking, for each k, the maximum between $\hat{\mu}_{n,k+1} - \hat{\mu}_{n,k+2}$ and the smallest non-zero eigenvalue.

[7] Similarly, the existence of a well defined static factor representation is linked to the behavior of the eigenvalues of the *variance-covariance matrix* (rather than the spectral density matrix) of the first n variables, say ν_{nk} (Chamberlain and Rothschild (1983)).

same rate, whereas the others are bounded and bounded away from zero in probability. Now let us consider DER. When both eigenvalues are either large or small, the ratio should be relatively small, but, when $k = q$, the numerator is large and the denominator is small, so that the ratio is large. The intuition behind DDR is similar. When $k = q$, the numerator is large, since it is the difference between a large eigenvalue and a small one. On the other hand, the denominator is small, since it is the difference between small eigenvalues.

We evaluate the small sample performance of DER, DGR and DDR compared to each other and with HL and O, by means of a few Monte Carlo simulations.[8] We run three main experiments: (a) the one proposed by Hallin and Liska (2007); (b) the one proposed by Onatski (2009), and (c) our own experiment. Simulation results are the following: (i) DDR dominates DER and DGR in all experiments; (ii) DDR performs comparably or better than HL and O for experiments (a) and (b) and performs better than HL and O in experiment (c). We conclude that DDR is a safer and excellent alternative to existing criteria.

In principle, it is possible that a common shock has major effects on some frequency bands but not on others. For example, a demand shock may not have long-term effects on most variables related to real economic activity, like GDP growth, industrial production, employment and unemployment. An interesting feature of DDR (as indeed of DER and DGR) is that it can be applied to a single frequency or a frequency band, by averaging the eigenvalues on the relevant frequencies rather than all frequencies. To illustrate this possible usage of DDR we run a further experiment where the DGP is such that the spectral density matrix of the common components has reduced rank at specific frequencies. We show that DDR is reasonably able to detect the rank reduction.

In the empirical application we focus on a quarterly US macroeconomic data set: the FRED-QD data set by McCracken and Ng (2020). DDR provides a clear-cut result: the US macro economy is driven by two major shocks. This result holds both on specific frequency bands and on the entire $[0, \pi]$ interval. Moreover, it holds both for the whole sample and several sub-samples.

The dynamic eigenvalues, beside being the basis for our criteria, are consistent estimates of the total variance of the variables explained by the dynamic factors, decomposed by frequency. Moreover, by combining the eigenvalues with the corresponding eigenvectors, we can estimate, for each frequency, the variance explained by the common shocks for each variable Forni et al. (2000).

We find that two common shocks are enough to capture the bulk of the variance of the main macroeconomic aggregates, both at cyclical frequencies and in the long term.

The first dynamic factor explains almost nothing of the long-run variance of GDP, con-

[8] We do not consider in our simulations the methods proposed by Bai and Ng (2007) and Amengual and Watson (2007), since, as already observed, these methods assume a factor model which can be written in the static form (a restriction that we do not impose here) and require the preliminary determination of the number of static factors.

sumption, investment, unemployment rate and hours worked. Furthermore, it induces a positive covariance between GDP growth and inflation changes. It therefore has the characteristics of a demand shock. The second factor has instead the typical features of a supply shock.

The demand shock explains most of the cyclical fluctuations in GDP and other real variables. The supply shock plays a minor, but not negligible, role. Furthermore, the demand shock explains most of the variance in inflation and the federal funds rate at all frequencies. These results are incompatible with both the RBC model and the hypothesis recently suggested by Angeletos et al. (2020), that the cyclical fluctuations are explained by a single, non-inflationary demand shock.

The paper is organized as follows. Section 2 introduce notation. In Section 3 we present the model and the assumptions. In Section 4 we present our consistency results. Section 5 is devoted to our Monte Carlo exercises. The empirical study is presented in Section 6. The Appendices present the proofs and additional material related to the empirical application.

2 Notation

3 The model and the assumptions

4 Main Results

4.1 Proof Outline

4.2 Lemmata

5 Simulations

To evaluate the performance of the criteria defined in the previous sections, we run a few Monte Carlo experiments. In subsections 5.1 and 5.2 we use three different specifications of Model (**??**) to compare the performances of the criteria defined in the previous section with each other and with HL and O. The conclusion is that DDR dominates DGR and DER and performs comparably or better than HL and O.

As already observed (?), DDR (like DER and DGR) can be evaluated at a specific frequency of interest or on a frequency band. To illustrate this possible usage of our preferred criterion, in subsection 5.3 we conduct a further simulation exercise, where the DGP is such that the spectral density matrix of the common components has reduced rank at specific frequencies. We show that DDR, when evaluated at these frequencies, is able to detect the rank reduction with reasonable accuracy.

5.1 Simulation design

First experiment. The first DGP follows the one proposed by Hallin and Liska (2007), Section 5. Precisely:

i. The common shocks f_{jt}, $j = 1, \dots, q$, $t = 1, \dots, T$, $q \leq 3$, are $iid \sim \mathcal{N}(0, D_j)$, with $D_1 = 1$, $D_2 = .5$ and $D_3 = 1.5$.

ii. The idiosyncratic components are of form $e_{it} = \sum_{l=0}^{4} \sum_{k=0}^{2} g_{i,l,k}\, \varepsilon_{i+l,t-k}$, where the ε_{it}'s are $iid \sim \mathcal{N}(0, 1)$, and the $g_{i,l,k}$'s are $iid \sim \mathcal{U}_{[1,1.5]}$, where $i = 1, \dots, n$, $t = 1, \dots, T$, $l = 1, \dots, 4$, $k = 0, 1, 2$. The ε_{it}'s and the $g_{i,l,k}$'s are mutually independent and independent of the f_{jt}'s. Hence the idiosyncratic components are both autocorrelated and "locally" cross-correlated.

iii. the filters $\lambda_{ij}(L)$, $i = 1, \dots, n$, $j = 1, \dots, q$, are randomly generated (independently from the f_{jt}'s and e_{it}'s) by one of the following devices: (1) MA loadings: $\lambda_{ij}(L) = \lambda_{ij,0} + \lambda_{ij,1}L + \lambda_{ij,2}L^2$ with iid and mutually independent coefficients $(\lambda_{ij,0}, \lambda_{ij,1}, \lambda_{ij,2}) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_3)$; (2) AR loadings: $\lambda_{ij}(L) = m_{ij,0}(1 - m_{ij,1}L)^{-1}(1 - m_{ij,2}L)^{-1}$ with iid and mutually independent coefficients $m_{ij,0} \sim \mathcal{N}(0, 1), m_{ij,1} \sim \mathcal{U}_{[.8,.9]}$ and $m_{ij,2} \sim \mathcal{U}_{[.5,.6]}$.

Finally, for each i, the variance of each idiosyncratic component e_{it} and that of the corresponding common component $\chi_{it} = \sum_{j=1}^{q} \lambda_{ij}(L) f_{jt}$ are normalized to 0.5.

The artificial samples were generated with $q = 2, 3$ and (n, T)= $(60, 100)$, $(100, 100)$, $(70, 120)$, $(120, 120)$, $(150, 120)$. Notice that in this experiment we have both large and small factors (the variance of the second factor is one half of that of the first factor and one third of the variance of the third factor).

Second experiment. The second DGP is the one studied by Onatski (2009), Sections 5.1 and 5.3. The basic difference with respect to the previous DGP is that here the number of factors is kept fixed to $q = 2$, whereas the variance of the idiosyncratic components takes on different values.

Precisely:

i. The common shocks f_{jt}, $k = 1, \dots, q$, $t = 1, \dots, T$, are $iid \sim \mathcal{N}(0, I_q)$.

ii. The idiosyncratic components follow AR(1) processes both cross-sectionally and over time: $e_{it} = \rho_i \, e_{i,t-1} + v_{it}$, $v_{it} = \rho v_{i-1,t} + \epsilon_{it}$, where $\rho_i \sim iid \mathcal{U}_{[-.8,.8]}$, $\rho = .2$ and $\epsilon_{it} \sim iid \mathcal{N}(0, 1)$.

iii. The filters $\lambda_{ij}(L)$, $i = 1, \dots, n$ and $j = 1, \dots, q$, are randomly generated (independently from the f_{jt}'s and e_{it}'s) by one of the following devices: (1) MA loadings: $\lambda_{ij}(L) = m_{ij,0}(1 + m_{ij,1}L)(1 + m_{ij,2}L)$ with *iid* and mutually independent coefficients $m_{ij,0} \sim \mathcal{N}(0, 1), m_{ij,1} \sim \mathcal{U}_{[0,1]}$ and $m_{ij,2}) \sim \mathcal{U}_{[0,1]}$; (2) AR loadings: same as in the first experiment; $\lambda_{ij}(L) = m_{ij,0}(1-m_{ij,1}L)^{-1}(1-m_{ij,2}L)^{-1}$ with *iid* and mutually independent coefficients $m_{ij,0} \sim \mathcal{N}(0, 1), m_{ij,1} \sim \mathcal{U}_{[.8,.9]}$ and $m_{ij,2} \sim \mathcal{U}_{[.5,.6]}$.

For each i, the idiosyncratic component e_{it} and the common component $\chi_{it} = \sum_{j=1}^{q} \lambda_{ij}(L) f_{jt}$ are normalized so that their variances equal σ^2 and 1, respectively. Hence the idiosyncratic to common variance ratio is σ^2 for all i. Following Onatski (2009), we set $q = 2$ and (n, T, σ^2) equal to $(70, 70, 1)$, $(70, 70, 2)$, $(70, 70, 4)$, $(100, 120, 1)$, $(100, 120, 2)$, $(100, 120, 6)$, $(150, 500, 1)$, $(150, 500, 8)$, $(150, 500, 16)$.

Third experiment. The main feature of this experiment is that, unlike the previous ones, the idiosyncratic-common variance ratio differs across different cross-sectional units. The loadings are ARMA(1,2) filters and the number of factors is larger than in the previous experiments $(q = 2, 4, 6)$.

Precisely the third GDP is the following.

i. The common shocks f_{jt}, $j = 1, \dots, q$, $t = 1, \dots, T$, are $iid \sim \mathcal{N}(0, I_q)$.

ii. Same as in the second experiment. The idiosyncratic components follow AR(1) processes both cross-sectionally and over time: $e_{it} = \rho_i \, e_{it-1} + v_{it}$, $v_{it} = \rho v_{i-1t} + \epsilon_{it}$, where $\rho_i \sim iid \mathcal{U}_{[-.8,.8]}$, $\rho = 0.2$ and $\epsilon_{it} \sim iid \mathcal{N}(0, 1)$.

iii. the filters $\lambda_{ij}(L)$, $i = 1, \dots, n$ and $j = 1, \dots, q$, are randomly generated (independently from the f_{jt}'s and e_{it}'s) with ARMA loadings: $\lambda_{ij}(L) = (m_{ij,0}+m_{ij,1}L+m_{ij,2}L^2)/a_{ij,0}(1-a_{ij,1}L)$, where the coefficients are *iid* and mutually independent, $m_{ij,s} \sim \mathcal{U}_{[-1,1]}$, $s = 0, 1, 2$, and $a_{ij,r} \sim \mathcal{U}_{[-0.8,0.8]}$, $r = 0, 1$.

In this experiment we want to control for the common to idiosyncratic variance ratio without forcing all variables in the cross section to have the same ratio. To this end, for each artificial data set we compute the average sample variance of the common and the idiosyncratic components, say σ_χ^2 and σ_e^2. Then we multiply all common components by $1/\sigma_\chi$, and all idiosyncratic components by s/σ_e, with s taking on two values: (i) $s = 0.5$ (small idiosyncratic components) and (ii) $s = 1$ (large idiosyncratic components). Since variables do not have the same variance, we standardize them before estimation.

We set $q = 2, 4, 6$ and (n, T)= $(60, 80)$, $(120, 80)$, $(60, 240)$, $(120, 240)$, $(240, 480)$.

To compute DER, DGR and DDR, we use the periodogram smoothing estimator (**??**) with the bandwidth parameter $M_T = \lfloor 0.75\sqrt{T} \rceil$ and take the average of the eigenvalues evaluated in the frequency grid $\omega_\ell = 2\pi\ell/T$, $\ell = 1, \ldots, T-1$. As explained above, in order to avoid denominators very close to 0, we correct the denominator of DDR by taking, for each k, the maximum between $\hat{\mu}_{n,k+1} - \hat{\mu}_{n,k+2}$ and the smallest non-zero eigenvalue $\hat{\mu}_{n,m}$. In our simulation, such correction is active in about 5% of the ratios.

We compare our criteria with HL and O. With regard to HL, we use the log information criterion $IC_{2;n}^T$ with penalty $p_1(n, T)$ and the Bartlett lag window with parameter $M_T = \lfloor 0.75\sqrt{T} \rfloor$, which yield the best performance in the simulations shown by the authors. The method requires evaluation of the loss function over a grid n_j, T_j, $j = 1, \ldots, J$. We stick to the one proposed by the authors, i.e. $n_j = n - 10j$, $T_j = T - 10_j$, $j = 0, 1, 2, 3$.

When dealing with O, we use the procedure described in Section 5.3 of Onatski (2009). We found that the results are sensitive to the choice of the parameter m (Onatski, 2009, footnote 7). For the second experiment, we stick to Onatski's choice, which is very effective ($m = 30$ for $(n, T) = (70, 70)$, $m = 40$ for $(n, T) = (100, 120)$, $m = 65$ for $(n, T) = (150, 500)$). For the first DGP, we use $m = 15$; for the third experiment, we use $m = 15, 20, 30$ for $T = 80, 240, 480$, respectively. These values produce better results than the larger ones suggested in Onatski's paper.

For all experiments and all estimators we set $k_{\max} = 8$. For all experiments we generate 500 artificial data sets. We evaluate the results by using the percentage of correct answers.

5.2 Simulation results

Table 1 reports results for the first experiment. Boldface numbers denote the estimator(s) which perform best for each q, n, T configuration. Results for HL are very close to those reported in Hallin and Liška (2007). HL and DDR have the best performances for all q, n, T configurations. With MA loadings and AR loadings, $q = 2$, when HL has the best performance DDR ranks second and vice versa; moreover, differences are very small. With AR loadings, $q = 3$, DDR clearly outperforms HL, with the exception of the cases $(n, T) = (120, 120)$ and $(n, T) = (150, 120)$, in which results are similar. DDR uniformly outperforms DGR, which in turn does better than DER and O. In this experiment the second factor is small as compared to the first and the third ones. Hence we can conclude that DDR is reasonably able to detect small factor.

Table 2 reports results for the second experiment. Results for O are close to those reported in Onatski (2009). With MA loadings, O is the best method for all n, T, σ^2 configurations, but for the case $(n, T, \sigma^2) = (70, 70, 2)$, in which it is beaten by DDR. DDR is tied for first with O in 5 cases, performs slightly better than O in the case above and ranks second in the remaining three cases. Again, DDR uniformly dominates DGR, which in turn dominates DER. With AR loadings, DDR performs the best for all n, T, σ^2 configurations. As already noticed by Onatski (2009), HL works well for low values of σ^2, but fails dramatically for noisy data, with the exception of the AR case with $T = 500$. Onatski (2009) argues that the grid of the calibration procedure does not work with high values of σ^2.

Table 3 reports results for the third experiment. In the upper panel (small idiosyncratic components), DDR and DGR perform similarly and dominate all other estimators, with the only exception of the configuration $(q, n, T) = (4, 120, 80)$, in which HL performs slightly better. O does not perform well for $q = 4, 6$ and small T. In the lower panel (large idiosyncratic components), DDR dominates DGR, which in turn dominates DER. No method is able to correctly capture the number of factors for the (q, n, T) configurations $(4, 60, 120)$, $(4, 120, 80)$, $(6, 60, 120)$, $(6, 120, 80)$ and $(6, 60, 240)$. In the remaining rows, DDR is the method that works best in all cases except the $(6, 120, 240)$ configuration, in which HL performs slightly better. Again, O works reasonably well for $q = 2$ but not for $q = 4, 6$.

Surprisingly, HL fails for large n, T configurations, particularly in the case $q = 2$, which should be the simplest one. The fact that the performance deteriorates as n, T gets larger suggests that the calibration procedure for the penalty function does not work in these cases. Paradoxically, the procedure works much better for the same cases when the idiosyncratic components are larger (lower panel): here the problem is not that the data is noisy, but on the contrary, it is that it is not noisy enough.

To better understand what is going on here, we run a fourth experiment.

Fourth experiment. The DGP is the same of the third experiment. The parameters q, n and T are kept fixed, with values $q = 3$, $n = 100$, $T = 100$. The parameter s, which governs the average idiosyncratic variance, varies between 0.3 and 1.2.

Table 4 shows results for HL and DDR. The table reports not only the percentage of correct outcomes, but also the percentages of outcomes with $\hat{q} < q$ and $\hat{q} > q$. HL performs very well for intermediate values of s, but has problems for DGP's with both small and large idiosyncratic components. When s is very small, the number of factors is overestimated, whereas when s is very large the number of factors is underestimated. On the other hand, DDR does not loose accuracy for small idiosyncratic components; when data becomes noisier, its performance deteriorates, but less than HL. More specifically, DDR and HL have similar performances in the interval $0.50 \leq s \leq 0.80$; for all other values of s, DDR outperforms HL, and for extreme values of s the difference is very large. We conjecture that, by changing the grid n_j, T_j, $j = 1, \ldots, J$ over which to evaluate the loss function, HL would perform better. But in practice we do not know in advance how noisy data is, so that the choice of the grid is a possible source of error.

Summing up, DDR dominates DGR and DER for almost all DGPs. In Experiment 1, AR loadings, Experiment 2, AR loadings, and Experiment 3, large idiosyncratic components, DDR has the best performance for almost all parameter configurations. For Experiment 1, MA loadings, DDR and HL perform similarly and dominate the other estimators. For Experiment 3, small idiosyncratic components, DDR and DGR perform similarly and dominate the other estimators. For Experiment 2, MA loading, O has best performance, but DDR ranks either first or second in all cases. In Table 4, DDR generally outperforms HL.

We conclude that DDR is preferable to DGR and DER and has an excellent performance compared to the best existing estimators.

5.3 Using DDR on selected frequency bands

In this subsection we run a fifth Monte Carlo experiment. We use two DGPs, the Trend-cycle Model, and the Stop-band Model. In both models the spectral density matrix of the common components has reduced rank at a specific frequency. In the Trend-cycle Model, one of the common shock is loaded by all variables with impulse response functions which vanish for $L = 0$. Hence, assuming that x_t is the first difference of the I(1) vector y_t, this shock has transitory effects on all variables in y_t. In the Stop-band Model one of the common shocks is loaded by all variables with filters whose frequency response vanishes at frequency $\pi/6$, which can be interpreted as a cyclical frequency, in that it corresponds to a period of 3 years with quarterly data.

Fifth experiment. More precisely, in the fifth experiment the data are generated as follows.

i. We have two common shocks f_{jt}, $j = 1, 2$, $t = 1, \dots, T$, which are $iid \sim \mathcal{N}(0, I_2)$.

ii. The idiosyncratic components are mutually independent white noises: $e_{it} = g_i \epsilon_{it}$, where $\epsilon_{it} \sim iid\mathcal{N}(0, 1)$ and the g_i's are $iid \sim \mathcal{U}_{[-1,1]}$.

iii. Trend-cycle Model. For the first factor, the permanent shock, we have the AR loadings $\lambda_{i1}(L) = a_{i1,0}/(1 - a_{i1,1}L)$, where the coefficients are iid, mutually independent, $a_{i1,0} \sim \mathcal{U}_{[-1,1]}$, $a_{i1,1} \sim \mathcal{U}_{[-0.8,0.8]}$. For the second factor, the transitory shock, we have the ARMA loadings $\lambda_{i2}(L) = a_{i2,0}(1-L)/(1-a_{i2,1}L)$, where the coefficients are iid, mutually independent, $a_{i2,0} \sim \mathcal{U}_{[-1,1]}$, $a_{i2,1} \sim \mathcal{U}_{[0,0.7]}$.

iv. Stop-band model. For the first factor we have the same AR loadings as in the Trend-cycle model: $\lambda_{i1}(L) = a_{i1,0}/(1 - a_{i1,1}L)$, $a_{i1,0} \sim \mathcal{U}_{[-1,1]}$, $a_{i1,1} \sim \mathcal{U}_{[-0.8,0.8]}$. For the second factor we use here a rough stop-band filter whose frequency response vanishes at frequencies $\pi/6$ and $-\pi/6$. Precisely, we use the ARMA loadings $\lambda_{i2}(L) = a_{i2,0}(1 - e^{-\mathrm{i}\pi/6}L)(1 - e^{\mathrm{i}\pi/6}L)/(1 - a_{i2,1}L)$, where "i" in Roman font denotes the imaginary unit and $a_{i2,0} \sim \mathcal{U}_{[-0.5,0.5]}$, $a_{i2,1} \sim \mathcal{U}_{[0.8,0.9]}$.

In the Trend-cycle model we have $\lambda_{i2}(0) = 0$ for all i, so that there is just one factor affecting the variables at frequency 0. Hence $q = 2$ for $\omega \neq 0$ and $q = 1$ for $\omega = 0$. To get an economic interpretation of the model, assume that the variables are the growth rates of macroeconomic series related to real economic activity, like for instance GDP and its components, employment variables, industrial production indexes, hours worked and so on. Then the first factor is a long-run, permanent shock, driving the common trends of the series, expressed in log levels. For instance, it can be a technology shock, or a generic supply shock. On the other hand, the second factor has transitory effects and can be interpreted as a demand shock driving a common cycle. Both shocks may affect variables at business cycle frequencies.

In the Stop-band model, we have $\lambda_{i2}(e^{\mathrm{i}\pi/6}) = \lambda_{i2}(e^{-\mathrm{i}\pi/6}) = 0$ for all i, so that there is just one factor affecting the variables at frequency $\pi/6$, which corresponds to a period of 3 years with quarterly data. As a result, we have $q = 1$ for $\omega = \pm\pi/6$. The filters $\lambda_{i2}(L)$, $i = 1, \dots, n$, can be regarded as rough stop-band filters toning down cyclical frequencies. As an example, the amplitude response of the filter $(1 - e^{-\mathrm{i}\pi/6}L)(1 - e^{\mathrm{i}\pi/6}L)/(1 - a_{i2,1}L)$, with $a_{i2,1} = 0.85$, is plotted in Figure 1.

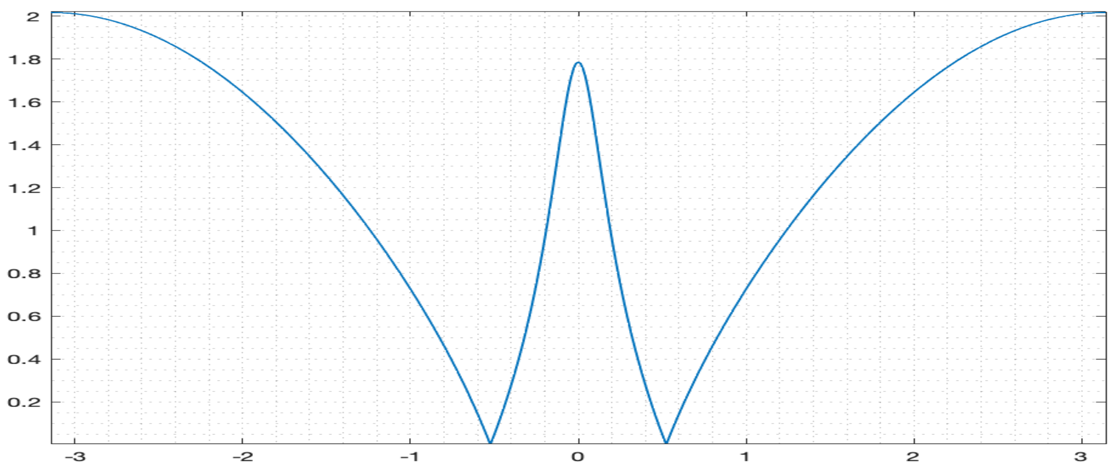

Figure 1: Amplitude frequency response of the filter $(1 - e^{-\mathrm{i}\pi/6}L)(1 - e^{\mathrm{i}\pi/6}L)/(1 - a_{i2,1}L)$ with $a_{i2,1} = 0.85$, as a function of ω, $\omega \in [-\pi, \pi]$.

As in the third experiment, for each artificial data set we compute the average sample variance of the common and the idiosyncratic components, say σ_χ^2 and σ_e^2. Then we multiply all common components by $1/\sigma_\chi$ and all idiosyncratic components by s/σ_e, with s taking on two values: (i) $s = 0.6$ (small idiosyncratic components) and (ii) $s = 1.2$ (large idiosyncratic components). Again, we standardize the variables before estimation. For this exercise, we set $n = 120$ and $T = 240$. DDR is evaluated at the points $\omega = 0$ and $\omega = \pi/6$. Moreover, it is evaluated in the long-run frequency band $[0, 2\pi/80]$, corresponding to cycles of 20 years or more with quarterly data, in the cyclical band $[2\pi/32, 2\pi/6]$, corresponding to cycles between

18 months and 8 years, and in the short-run band $[2\pi/8, \pi]$, corresponding to cycles between 6 and 18 months. Finally, we evaluate DDR in the whole interval $[0, \pi]$.

Table 5 reports the percentage of outcomes $\hat{q} = 1$, $\hat{q} = 2$ and $\hat{q} > 2$, over 500 replications. Boldface numbers denote the percentage of correct outcomes. On the long-run band $0 \leq \omega \leq 2\pi/80$ (which corresponds to periodicity greater than 20 years with quarterly data) the true number of factors is 2 for both models, but, for the Trend-cycle model, the contribution of the transitory shock to total variance is negligible, so that we consider correct the outcome $\hat{q} = 1$. On the cyclical band $2\pi/32 \leq \omega \leq 2\pi/8$ the true number of factors is 2 for both models, but, for the Stop-band model, the contribution of the non-cyclical shock to total variance is very small, so that we consider correct the outcome $\hat{q} = 1$.

In most cases, DDR is able to detect the correct number of factors in all cases. At a first sight, the worst outcome is the one of the Stop-band model, small idiosyncratic components (upper-right panel), for the cyclical band, where the result is $\hat{q} = 2$ in almost 40% of the cases. Indeed, we have already observed that the true number of factors is in fact 2. With small idiosyncratic components, in several cases the criterion is able to detect the presence of a second factor, despite the fact that its effect are very small.

6 How many shocks are there in the macro economy?

How many shocks are there driving the macroeconomic variables? How many long-term shocks are there? and how many are the ones driving the business cycle? These important issues have remained somewhat obscured in the empirical macroeconomic literature. Structural VARs, which are the most common tool in empirical macroeconomic analysis, are not able to answer these questions, since in VAR models the number of shocks is equal to the number of variables that the researcher decides to include in the model. On the contrary, factor models allow us to answer.

Existing evidence is mixed. Sargent and Sims (1977), by using a small-dimensional dynamic factor model, find that two shocks fit US macroeconomic data reasonably well. Giannone, Reichlin, and Sala (2005) have no estimator, but argue informally in favor of two shocks, basing on the explained variances of the principal component series of a large factor model. Bai and Ng (2007), using their information criteria, find 7 static factors and 4 shocks. Amengual and Watson (2007) find 7 static factors and 7 shocks. Hallin and Liska (2007) do not find a clear-cut result: four shocks but perhaps only one. Onatski (2009) finds that his proposed test cannot reject the null of 2 shocks versus the alternative of a number of shocks between 3 and 7.

The above questions are interesting in themselves, as a possible guidance for macroeconomic modeling. But they are also inextricably intertwined with others, which are central to the macroeconomic debate of the last 30-40 years: are the main cyclical shocks permanent or transitory? are they mainly demand or supply shocks? This is not the place to summarize this vast literature. We just want to mention a few contributions. The RBC model is challenged by Blanchard and Quah (1989) and Galí (1999), where it is argued that transitory shocks explain a large fraction of output and unemployment fluctuations. Beaudry and Portier (2006) argue that the bulk of the business cycle is explained by the news shock, that is, a permanent shock that anticipates future increases in productivity. The predominance of news shocks is questioned in Barsky and Sims (2011) and Forni, Gambetti, and Sala (2014). Finally, Angeletos et al. (2020) identify what they call the "main cyclical shock" and show that, on the one hand, it is disconnected from the long term, but, on the other hand, it has very modest effects on inflation.

6.1 The number of shocks driving the US macro economy

For our empirical analysis we use the US quarterly macroeconomic data set recently developed by McCracken and Ng (2020?). Of this data set we consider the $n = 216$ series starting from the first quarter of 1960. The final date of the sample is the first quarter of 2020. As for the transformations, we deviate from those suggested by the Authors for the interest rates, which are taken in levels rather than in differences; furthermore, prices and other nominal variables are taken in log-differences, rather than in double differences of the logs. The reason is that we want to avoid a possible over-differentiation, which enhances the high frequencies, of little interest for our analysis. The complete list of variables and transformations is provided in Appendix C. After the transformations, the number of observations over time is $T = 240$.

We consider the entire sample and nine sub-samples: the five 40-year sub-samples 1960Q2-2000Q1, 1965Q2-2005Q1, 1970Q2-2010Q1, 1975Q2-2015Q1 and 1980Q2-2020Q1, and the four 30-year sub-samples 1960Q2-1990Q1, 1970Q2-2000Q1 , 1980Q2-2010Q1 and 1990Q2-2020Q1. To estimate q, we compute DDR on three frequency bands: the entire $[0, \pi]$ band, the $[0, 2\pi/6]$ band, which excludes fluctuations of less than 18 months, of little interest for macroeconomic analysis, and the $[2\pi/32,\ 2\pi/6]$ cyclical band, which includes waves ranging from 18 months to 8 years. For comparison, we consider DER, DGR, HL and O (only on the whole $[0, \pi]$ band). HL is calculated as in the simulations. For O, we use $m = 20$ for the whole sample and for the 40-year sub-samples and $m = 15$ for the 30-year sub-samples. For DER, DGR and DDR we set the bandwidth parameter $M_T = \lfloor\sqrt{T}\rceil$, according to the results of the simulation exercise reported in Appendix D. For all estimators, we set $k_{\max} = 8$.

Results are reported in Table 6. Our favorite estimator DDR selects 2 factors; this finding is reasonably consistent across sub-samples. The same holds for DDRa, which excludes the short-run frequencies. DDRbc presents a mixed evidence: two cyclical factors for the whole sample and four sub-samples, just one factor for the remaining 5 sub-samples. By contrast, HL selects 5 factors, whereas O is in favor of 3 factors. As for the sub-samples, HL oscillates between 2 and 4 factors, with a prevalence of 4 factors, and O varies between 2 and 5, with a prevalence of 2 factors. Overall, DDR and DDRa are more parsimonious than HL and O and more consistent across sub-samples.

Why do DDR and HL differ so much? One possible explanation is that there are 2 large factors and 3 smaller factors, and that the latter are elusive to DDR but not to HL. This interpretation contrasts however with the simulation of Table 1, where there are factors of different variance and DDR has a performance similar to that of HL. An alternative explanation is that HL overestimates the number of factors, as in the simulation of Table 4.

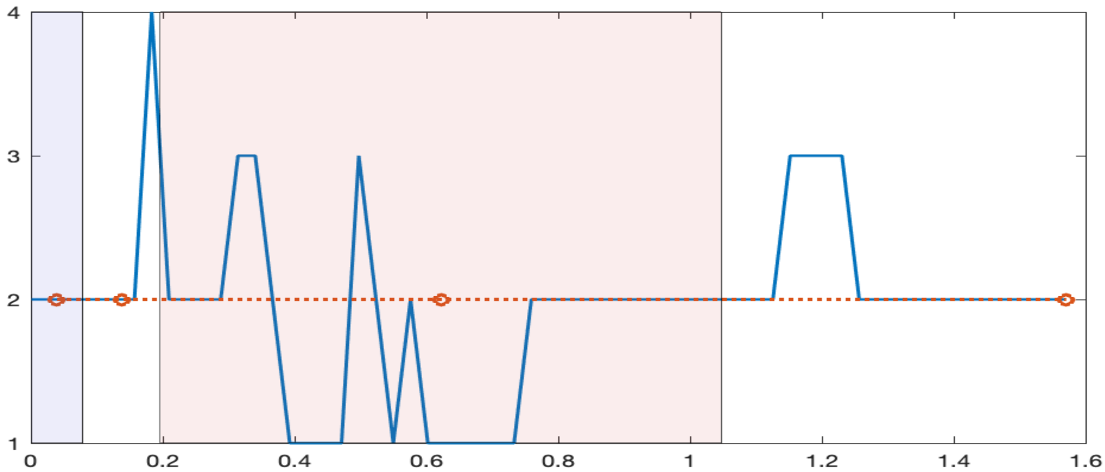

Figure 2: Number of factors estimated by DDR when evaluated at each frequency (blue solid line) and on 5 frequency bands (red dotted line with circles).

From now on we focus solely on our favorite estimator, DDR. Figure 2 shows the number of estimated factors by DDR for each single Fourier frequency in the interval $[0, \pi/4]$ and each one of the frequency bands $[0, 2\pi/80]$ (long-run), $[2\pi/80, 2\pi/32]$ (long cycles), $[2\pi/32, 2\pi/16]$ (medium cycles), $[2\pi/16, 2\pi/6]$ (short cycles) and $[2\pi/6, \pi]$ (short run). At frequency zero DDR selects 2 factors; the estimator is somewhat unstable when evaluated on single frequencies, fluctuating between 1 and 7 factors. Its value however stabilizes to 2 when averaging the eigenvalues on the five bands above. The result of 2 estimated factors on all bands is confirmed when using the bandwidths $M_T = \lfloor a\sqrt{T}\rceil$, with $a = 0.8, 0.9, 1.1, 1.2$.

Our conclusion is that there are two major shocks driving the US macro economy. The presence of other shocks cannot be ruled out, but the latter are not large and/or pervasive enough to be distinguished from the idiosyncratic components on the basis of existing data. The result of 2 shocks is in line with the findings of Sargent and Sims (1977), Giannone et al. (2005), and Onatski (2009).

6.2 Common and idiosyncratic variance

The dynamic eigenvalues represent the variance explained by the principal component series, broken down by frequency. The mean over all frequencies is the explained variance of the data set; the mean over a band is the explained variance of that band (Brillinger (1981)). Table 7 shows the explained variances of the first 6 principal component series for a few frequency bands of interest. Each variance is divided by the total variance of the variables on that band and multiplied by 100; therefore the numbers that appear in the table are the percentages of variance accounted for by the first 6 principal components. The first principal component, for instance, explains 33% of the sum of the variances of the variables, while the second one explains about 17%. On all bands, the difference between the variance explained by the second principal component and that explained by the third is relatively large, while the difference between the third and the fourth is relatively small. This is the reason why the criterion selects two factors.

The principal component series are consistent estimates of the dynamic factors (Forni et al. (2000)). Hence the explained variances above are estimates of the variance explained by the common shocks. If, as indicated by the criterion, there are only 2 common shocks, the sum of the variances of the common components is about half of the total variance and the sum of the idiosyncratic variances is the other half. At high frequencies, of little macroeconomic interest, the common variance is only 30% of the total, while on the cyclical band and in the long run the common variance is about 60%.

Of course, there are variables which have large idiosyncratic components and variables with small idiosyncratic components. The idiosyncratic components may represent sectoral, local, or foreign sources of variation, as well as measurement errors and other sources of fluctuations affecting only a small number of variables in the data set. Hence sectoral industrial production indexes, sectoral price indexes, exchange rates or variables measuring aspects of the macro economy which are poorly represented in the data set have large idiosyncratic components, whereas the main macroeconomic aggregates have small idiosyncratic components. We estimate the spectral-density matrix of the variable components driven by the k-th factors, $k = 1, 2$, by the simple formula

$$\mathbf{v}_k(\omega_\ell)\hat{\mu}_k(\omega_\ell)\mathbf{v}'_k(\omega_\ell),$$

$\omega_\ell = 2\pi\ell/T$, $\ell = 0, 1 \ldots, T-1$, where $\hat{\mu}_k(\omega_\ell)$ is the k-th eigenvalue of the spectral density matrix $\widehat{\mathbf{\Sigma}}(\omega_\ell)$, $\mathbf{v}_k(\omega_\ell)$ is the corresponding eigenvector and the prime denotes transposition and conjugation (Forni et al. (2000)).

Table 8 shows, for seven key macroeconomic variables, the variance explained by the first 2 principal components along with the variance explained by the following three, from the third to the fifth. The aim is to see how large is the explained variance for these variables when retaining only two factors and how large is the variance that we loose with respect to the choice $q = 5$ suggested by HL. We see from the table that, on the bands of macroeconomic interest, the long run, the long waves and the business cycle, 2 factors are enough to capture 80-85% of GDP growth fluctuations, about 70% of consumption, about 80% of investment, 80-90% of the unemployment rate variation, 85-90% of hours worked, about 90% of inflation and 80% of the federal funds rate. The variance that we loose by selecting $q = 2$ instead of $q = 5$ is not negligible, particularly for consumption and the interest rate; nevertheless, we can conclude that 2 factors are enough to capture the bulk of the variance of the main macroeconomic aggregates on the frequency bands of main macroeconomic interest.

6.3 A cyclical shock and a long-run shock

Now let us go on and see how much of the variance of the above variables is explained by the first factor and how much is explained by the second factor, for each frequency band. Of course the factors are not identified according to macroeconomic criteria. Indeed, imposing an economic identification scheme is beyond the scope of the present paper. Nevertheless, it turns out that the explained variances make it possible to assign a precise economic meaning to the common shocks resulting from the statistical identification and to draw some conclusions of great economic interest.

The explained variances are shown in Table 9. The first dynamic factor accounts for almost nothing of the variance of GDP and all real activity variables in the long run and the long cycles band, which is instead explained by the second factor. We therefore find ourselves, without having looked for it, in front of an identification *à la* Blanchard and Quah (1989): the first factor is a temporary shock, while the second one is a permanent shock. It is very much tempting to interpret the temporary shock as a demand shock and the long-run shock as a supply shock. To confirm this interpretation, we look at the covariances of GDP growth and inflation changes induced by the two shocks, and find that in fact such covariance is positive for the temporary shock, which therefore has the features of a demand shock, and negative for the long-run shock, which can then be regarded as a supply shock.[9] In the last column of the table are reported the explained variances at business cycle frequencies. The demand shock is the most important cyclical shock for real activity variables. It accounts for 53% of GDP growth fluctuations, 35% of consumption, 55% of investment, 62% of unemployment and 59% of hours worked. The contribution of the permanent, supply shock is not negligible, particularly for consumption, but is smaller, between 26 and 32%.

Figure 3 illustrates the same points by showing the spectral density of the seven variables above, along with the spectra of the common components driven by the two shocks. The upper-left panel refers to GDP growth. The supply shock accounts for the long run and the long- medium cycles, but explains almost nothing of short cycles and short-run frequencies,

[9] This is shown in Figure 2 below.

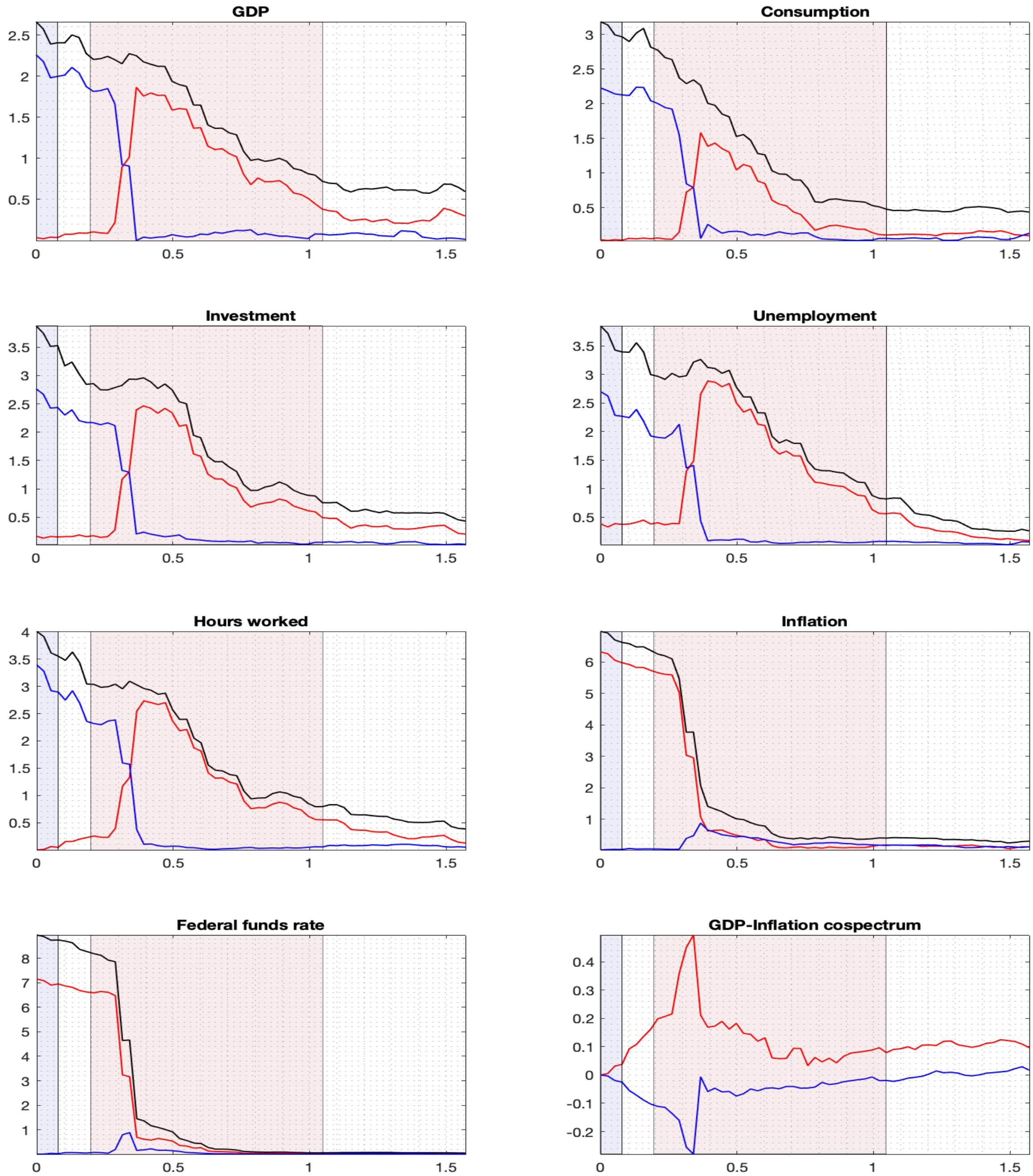

Figure 3: The estimated spectral density functions of seven variables (black line) and the components driven by the first factor (red line) and the second factor (blue line). The variables are: GDP growth, consumption growth, investment growth, unemployment rate changes, hours worked changes, GDP deflator inflation, federal funds rate. Bottom-right panel: co-spectrum of GDP growth and inflation changes produced by the first factor (red line) and the second factor (blue line). Lilac shadowed area: long-run frequencies; pink shadowed area: business cycle frequencies.

as if it were cut by a low-pass filter canceling waves of periodicity shorter than 4 years. By contrast, the demand shock explains almost all cycles of 4 years or less and almost nothing of longer cycles. A similar result holds for the other variables related to real activity: consumption, investment, unemployment and hours worked. Inflation and the interest rate are explained almost exclusively by the demand shock, both at business cycle frequencies and in the long run. The bottom-right panel of Figure 3 shows the co-spectra of GDP growth and inflation changes relative to the transitory shock (red line) and the permanent shock (blue line). As anticipated above, the transitory shock induces a a positive covariance between GDP growth and inflation changes, whereas the opposite is true for the permanent shock.[10]

The above results is in line with Blanchard and Quah (1989) and in sharp contrast with both the RBC model and the idea that news shocks explain the bulk of business cycle fluctuations. The finding by Angeletos et al. (2020), that the bulk of business cycle fluctuations is not driven by a long-run shock is confirmed. On the other hand, the hypothesis put forward by the Authors, that there could be just one non-inflationary demand shock affecting real activity variables, is clearly rejected: output fluctuations are partly explained by the supply shock, and the demand shock is closely related with inflation.

7 Summary and conclusion

[10]We consider inflation changes in place of inflation since the latter exhibits a large negative co-spectrum with GDP growth for all of the first five principal component series, owing to the 70s and the early 80s, characterized by low growth and high inflation.

References

Ahn, S. C. and A. R. Horenstein (2013). Eigenvalue ratio test for the number of factors. *Econometrica 81*, 1203–1227.

Alessi, L., M. Barigozzi, and M. Capasso (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics & Probability Letters 80*, 1806–1813.

Altissimo, F., R. Cristadoro, M. Forni, and G. Lippi, M.and Veronese (2010). New eurocoin: Tracking economic growth in real time. *The Review of Economics and Statistics 92*, 1024–1034.

Amengual, D. and M. W. Watson (2007). Consistent estimation of the number of dynamic factors in a large N and T panel. *Journal of Business & Economic Statistics 25*, 91–96.

Angeletos, G., F. Collard, and H. Dellas (2020). Business-cycle anatomy. *American Economic Review 110*, 3030–3070.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica 71*, 135–171.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*, 191–221.

Bai, J. and S. Ng (2007). Determining the number of primitive shocks in factor models. *Journal of Business & Economic Statistics 25*, 52–60.

Barsky, R. and E. Sims (2011). News shocks and business cycles. *Journal of Monetary Economics 58*, 273–289.

Beaudry, P. and F. Portier (2006). Stock prices, news, and economic fluctuations. *American Economic Review 96*, 1293–1307.

Bernanke, B. and J. Boivin (2003). Monetary policy in a data-rich environment. *Journal of Monetary Economics 50*, 525–546.

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

Blanchard, O. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review 79*, 655–673.

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

Brillinger, D. (1981). *Time Series: Data Analysis and Theory.* SIAM: Philadelphia.

Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica 51*, 1281–304.

Corielli, F. and M. Marcellino (2006). Factor based index tracking. *Journal of Banking & Finance 30*(8), 2215–2233.

Cristadoro, R., M. Forni, L. Reichlin, and G. Veronese (2005). A core inflation indicator for the euro area. *Journal of Money, Credit and Banking 37*, 539–60.

D'Agostino, A. and D. Giannone (2012). Comparing alternative predictors based on large-panel factor models. *Oxford Bulletin of Economics and Statistics 74*, 306–326.

Eickmeier, S. (2007). Business cycle transmission from the U.S. to Germany–a structural factor approach. *European Economic Review 51*, 521–551.

Favero, C. A., M. Marcellino, and F. Neglia (2005). Principal components at work: The empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics 20*, 603–620.

Forni, M. and L. Gambetti (2010). The dynamic effects of monetary policy: A structural factor model approach. *Journal of Monetary Economics 57*, 203–216.

Forni, M., L. Gambetti, and L. Sala (2014). No news in business cycles. *The Economic Journal 124*, 1168–1191.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009). Opening the black box: Structural factor models with large cross-sections. *Econometric Theory 25*, 1319–1347.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics 82*, 540–554.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005). The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association 100*, 830–840.

Forni, M. and M. Lippi (2001). The generalized dynamic factor model: Representation theory. *Econometric Theory 17*, 1113–1141.

Forni, M. and L. Reichlin (1998). Let's get real: A factor analytical approach to disaggregated business cycle dynamics. *Review of Economic Studies 65*, 453–473.

Galí, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review 89*, 249–271.

Geweke, J. (1977). *The dynamic factor analysis of economic time series models*, pp. 365–383. North-Holland Publishing Company: Amsterdam.

Giannone, D., L. Reichlin, and L. Sala (2005). Monetary policy in real time. In *NBER Macroeconomics Annual 2004, Volume 19*, pp. 161–224. National Bureau of Economic Research, Inc.

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

Hallin, M. and R. Liska (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association 102*, 603–617.

Hallin, M., C. Mathias, H. Pirotte Speder, and D. Veredas (2011). Market liquidity as dynamic factors. *Journal of Econometrics 163*, 42–50.

Kydland, F. E. and E. C. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica 50*, 1345–1370.

Lewbel, A. (1991). The rank of demand systems: Theory and nonparametric estimation. *Econometrica 59*, 711–30.

Ludvigson, S. and S. Ng (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics 83*, 171–222.

Ludvigson, S. C. and S. Ng (2009). Macro factors in bond risk premia. *The Review of Financial Studies 22*, 5027–5067.

Marcellino, M., J. Stock, and M. Watson (2003). Macroeconomic forecasting in the euro area: Country specific versus area-wide information. *European Economic Review 47*(1), 1–18.

McCracken, M. W. and S. Ng (2020). FRED-QD: A quarterly database for macroeconomic research. Working Papers 2020-005, Federal Reserve Bank of St. Louis.

Onatski, A. (2009). Testing hypotheses about the number of factors in large factor models. *Econometrica 77*(5), 1447–1479.

Sargent, T. and C. Sims (1977). Business cycle modeling without pretending to have too much a priori economic theory. Working Papers 55, Federal Reserve Bank of Minneapolis.

Smets, F. and R. Wouters (2007). Shocks and frictions in U.S. business cycles: A bayesian DSGE approach. *American Economic Review 97*, 586–606.

Stock, J. and M. Watson (2002a). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*, 147–62.

Stock, J. H. and M. W. Watson (2002b). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*, 1167–1179.

Tables

q	n	T	HL	O	DER	DGR	DDR
MA loadings							
2	60	100	**99.6**	47.6	80.8	92.4	98.4
	100	100	**99.6**	71.2	87.8	96.6	99.2
	70	120	97.8	53.2	87.2	96.2	**99.6**
	120	120	99.4	78.4	96.6	99.4	**100.0**
	150	120	99.6	83.0	97.4	99.4	**100.0**
3	60	100	62.0	15.8	27.2	52.6	**77.0**
	100	100	**91.2**	24.6	29.6	52.0	87.8
	70	120	89.8	18.8	32.8	62.2	**90.8**
	120	120	**99.0**	25.8	43.0	70.4	97.0
	150	120	**99.8**	29.6	48.0	74.2	97.4
AR loadings							
2	60	100	96.8	81.2	84.4	92.8	**98.8**
	100	100	**99.4**	93.0	90.4	94.6	**99.4**
	70	120	**99.8**	88.8	89.2	95.8	99.4
	120	120	**100.0**	96.8	94.4	97.4	**100.0**
	150	120	**100.0**	96.8	96.4	98.0	**100.0**
3	60	100	31.2	33.4	41.2	54.6	**67.6**
	100	100	62.4	54.4	49.0	66.0	**87.0**
	70	120	69.4	49.8	49.6	63.2	**81.0**
	120	120	90.8	67.6	57.8	70.0	**91.6**
	150	120	**94.6**	74.4	63.2	75.8	93.6

Table 1: First experiment described in Section 5. Percentage of correct outcomes over 500 replications. HL: Hallin and Liška (2007) estimator, O: Onatski (2009) estimator, DER: Dinamic Eigenvalue Ratio estimator, DGR: Dynamic Growth Ratio estimator; DDR: Dynamic Difference Ratio Estimator. Boldface numbers denote the estimator which performs best in each row.

n	T	σ^2	HL	O	DER	DGR	DDR
MA loadings							
70	70	1	**100.0**	**100.0**	99.6	99.8	**100.0**
70	70	2	94.6	99.8	89.2	94.0	**100.0**
70	70	4	1.6	**89.0**	49.4	58.2	77.6
100	120	1	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
100	120	2	**100.0**	**100.0**	99.6	**100.0**	**100.0**
100	120	6	3.0	**98.8**	54.0	62.2	81.4
150	500	1	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
150	500	8	**100.0**	**100.0**	99.2	99.8	**100.0**
150	500	16	40.2	**97.4**	45.6	47.0	88.8
AR loadings							
70	70	1	99.0	98.8	90.8	96.0	**99.4**
70	70	2	85.6	89.4	78.4	87.8	**98.4**
70	70	4	14.4	64.6	61.0	69.4	**84.8**
100	120	1	**100.0**	99.0	99.6	**100.0**	**100.0**
100	120	2	**100.0**	99.4	95.8	98.2	**100.0**
100	120	6	44.8	77.6	75.4	82.0	**91.4**
150	500	1	**100.0**	98.4	**100.0**	**100.0**	**100.0**
150	500	8	**100.0**	97.8	99.6	**100.0**	**100.0**
150	500	16	99.6	91.6	91.4	93.0	**99.8**

Table 2: Second experiment described in Section 5 ($q = 2$). Percentage of correct outcomes over 500 replications. HL: Hallin and Liška (2007) estimator, O: Onatski (2009) estimator, DER: Dinamic Eigenvalue Ratio estimator, DGR: Dynamic Growth Ratio estimator; DDR: Dynamic Difference Ratio Estimator. Boldface numbers denote the estimator which performs best in each row.

q	n	T	HL	O	DER	DGR	DDR
Small idiosyncratic components							
2	60	120	99.2	98.8	**100.0**	**100.0**	**100.0**
	120	80	88.6	90.8	**100.0**	**100.0**	99.8
	60	240	99.6	**100.0**	**100.0**	**100.0**	**100.0**
	120	240	35.6	**100.0**	**100.0**	**100.0**	**100.0**
	240	480	0.6	**100.0**	**100.0**	**100.0**	**100.0**
4	60	120	97.6	64.8	88.8	97.0	**98.6**
	120	80	**89.2**	36.2	67.2	86.6	88.8
	60	240	**100.0**	94.6	**100.0**	**100.0**	**100.0**
	120	240	91.6	99.2	**100.0**	**100.0**	**100.0**
	240	480	73.4	**100.0**	**100.0**	**100.0**	**100.0**
6	60	120	8.6	23.8	17.8	56.4	**57.0**
	120	80	0.8	16.4	5.0	**26.6**	21.8
	60	240	96.8	54.0	96.4	99.2	**99.8**
	120	240	98.6	87.0	**100.0**	**100.0**	**100.0**
	240	480	91.6	99.8	**100.0**	**100.0**	**100.0**
Large idiosyncratic components							
2	60	120	97.2	87.2	85.8	93.4	**99.4**
	120	80	99.0	93.4	85.4	94.2	**99.8**
	60	240	**100.0**	99.0	98.8	99.4	**100.0**
	120	240	99.8	99.8	99.8	**100.0**	**100.0**
	240	480	99.8	**100.0**	**100.0**	**100.0**	**100.0**
4	60	120	0.0	16.8	2.0	7.6	**30.0**
	120	80	0.0	**15.2**	2.0	7.4	11.8
	60	240	16.8	32.4	29.4	55.8	**80.8**
	120	240	99.8	73.4	92.0	97.8	**100.0**
	240	480	96.8	**100.0**	**100.0**	**100.0**	**100.0**
6	60	120	0.0	**8.2**	0.0	0.0	0.8
	120	80	0.0	**9.0**	0.0	0.0	0.0
	60	240	0.0	**7.8**	0.0	0.4	7.6
	120	240	**48.0**	20.8	5.4	11.2	47.0
	240	480	94.8	94.8	**100.0**	**100.0**	**100.0**

Table 3: Third experiment described in Section 5. Percentage of correct outcomes over 500 replications. HL: Hallin and Liška (2007) estimator, O: Onatski (2009) estimator, DER: Dinamic Eigenvalue Ratio estimator, DGR: Dynamic Growth Ratio estimator; DDR: Dynamic Difference Ratio Estimator. Boldface numbers denote the estimator which performs best in each row.

	HL				DDR			
s	$\hat{q} < q$	$\hat{q} = q$	$\hat{q} > q$	Total	$\hat{q} < q$	$\hat{q} = q$	$\hat{q} > q$	Total
0.30	0.0	**12.2**	87.8	100.0	0.0	**99.8**	0.2	100.0
0.35	0.0	**41.4**	58.6	100.0	0.0	**99.6**	0.4	100.0
0.40	0.0	**69.4**	30.6	100.0	0.0	**99.8**	0.2	100.0
0.45	0.0	**88.8**	11.2	100.0	0.2	**99.4**	0.4	100.0
0.50	0.0	**98.4**	1.6	100.0	0.0	**99.8**	0.2	100.0
0.55	0.0	**99.8**	0.2	100.0	0.2	**99.6**	0.2	100.0
0.60	0.0	**100.0**	0.0	100.0	0.0	**100.0**	0.0	100.0
0.65	0.0	**100.0**	0.0	100.0	0.0	**100.0**	0.0	100.0
0.70	0.0	**100.0**	0.0	100.0	0.0	**100.0**	0.0	100.0
0.75	0.6	**99.4**	0.0	100.0	0.8	**99.2**	0.0	100.0
0.80	0.4	**99.6**	0.0	100.0	0.2	**99.8**	0.0	100.0
0.85	2.6	**97.4**	0.0	100.0	1.2	**98.6**	0.2	100.0
0.90	6.6	**93.4**	0.0	100.0	3.0	**97.0**	0.0	100.0
0.95	15.0	**85.0**	0.0	100.0	3.8	**96.2**	0.0	100.0
1.00	28.2	**71.8**	0.0	100.0	5.4	**94.6**	0.0	100.0
1.05	45.0	**55.0**	0.0	100.0	11.0	**89.0**	0.0	100.0
1.10	63.2	**36.8**	0.0	100.0	15.2	**84.8**	0.0	100.0
1.15	77.4	**22.6**	0.0	100.0	18.0	**82.0**	0.0	100.0
1.20	89.4	**10.6**	0.0	100.0	26.8	**73.2**	0.0	100.0

Table 4: DGP: model of the third experiment described in Section 5, with $q = 3$, $n = 100$, $T = 100$ and $s = 0.3 + 0.05i$, $i = 0, 1, \ldots, 18$. Percentage of underestimated outcomes, correct outcomes and overestimated outcomes over 500 replications. HL: Hallin and Liška (2007) estimator, DDR: Dynamic Difference Ratio Estimator. Boldface numbers denote the percentage of correct outcomes.

	Trend-Cycle Model				Stopband Model			
Frequency band	$\hat{q}=1$	$\hat{q}=2$	$\hat{q}>2$	Total	$\hat{q}=1$	$\hat{q}=2$	$\hat{q}>2$	Total
Small idiosyncratic components								
$\omega=0$	**82.6**	13.8	3.6	100.0	1.8	**95.2**	3.0	100.0
$0\leq\omega\leq 2\pi/80$	**89.6**	10.2	0.2	100.0	1.6	**98.4**	0.0	100.0
$\omega=2\pi/12$	0.0	**100.0**	0.0	100.0	**99.8**	0.2	0.0	100.0
$2\pi/32\leq\omega\leq 2\pi/8$	0.0	**100.0**	0.0	100.0	**60.6**	39.4	0.0	100.0
$2\pi/8\leq\omega\leq\pi$	0.0	**99.4**	0.6	100.0	0.0	**96.4**	3.6	100.0
$0\leq\omega\leq\pi$	0.0	**99.8**	0.2	100.0	0.0	**98.0**	2.0	100.0
Large idiosyncratic components								
$\omega=0$	**93.2**	3.6	3.2	100.0	9.6	**82.0**	8.4	100.0
$0\leq\omega\leq 2\pi/80$	**98.4**	1.2	0.4	100.0	11.2	**87.4**	1.4	100.0
$\omega=2\pi/12$	3.8	**96.2**	0.0	100.0	**100.0**	0.0	0.0	100.0
$2\pi/32\leq\omega\leq 2\pi/8$	4.2	**95.8**	0.0	100.0	**94.6**	5.4	0.0	100.0
$2\pi/8\leq\omega\leq\pi$	0.0	**99.8**	0.2	100.0	0.6	**98.4**	1.0	100.0
$0\leq\omega\leq\pi$	0.0	**99.8**	0.2	100.0	1.4	**98.2**	0.4	100.0

Table 5: DGP: Trend-cycle Model (left panel) and Stopband Model (right panel) described in Section 5, with $q = 2$, $n = 120$, $T = 240$ and $s = 0.6$ (small idiosyncratic components), $s = 1.2$ (large idiosyncratic components). Percentage of outcomes $\hat{q} = 1$, $\hat{q} = 2$ and $\hat{q} > 2$, over 500 replications, obtained with the DDR estimator, evaluated at selected frequencies or frequency bands. Boldface numbers denote the percentage of correct outcomes. In the Trend-cycle Model one of the two common shocks has zero effect in the long-run for all variables, so that the true value of q at frequency 0 is 1. On the long-run band $0 \leq \omega \leq 2\pi/80$ (which corresponds to periodicity greater than 20 years with quarterly data) the true number of factors is 2, but the contribution of the transitory shock to total variance is negligible, so that we consider correct the outcome $\hat{q} = 1$. In the Stopband Model one of the two common shocks has zero effect at frequency $\pi/6$ for all variables, so that the true value of q at this frequency is 1. On the cyclical band $2\pi/32 \leq \omega \leq 2\pi/8$ (which corresponds to cycles between 2 and 8 years with quarterly data) the true number of factors is 2, but the contribution of the transitory shock to total variance is very small, so that we consider correct the outcome $\hat{q} = 1$.

Sample span	DDR	DDRa	DDRbc	DER	DGR	HL	O
1960Q2-2020Q1	2	2	2	2	2	5	3
1960Q2-2000Q1	2	2	1	1	1	4	2
1965Q2-2005Q1	2	2	1	1	1	4	2
1970Q2-2010Q1	2	2	2	2	2	4	2
1975Q2-2015Q1	2	2	2	2	2	4	5
1980Q2-2020Q1	2	2	2	2	2	4	2
1960Q2-1990Q1	1	2	1	1	1	3	5
1970Q2-2000Q1	1	1	1	1	1	4	4
1980Q2-2010Q1	2	3	2	2	2	2	2
1990Q2-2020Q1	2	2	1	2	2	3	3

Table 6: Number of factors detected by the competing criteria for the whole sample and nine sub-samples. DDR: Dynamic Difference Ratio Estimator. DDRa: Dynamic Difference Ratio Estimator evaluated on the $[0 \quad 2\pi/6]$ frequency band. DDRa: Dynamic Difference Ratio Estimator evaluated on the cyclical band $[2\pi/32 \quad 2\pi/6]$. DER: Dinamic Eigenvalue Ratio estimator. DGR: Dynamic Growth Ratio estimator. HL: Hallin and Liška (2007) estimator. O: Onatski (2009) estimator.

	All frequencies	Long run > 20 years	Long cycles 8-20 years	Medium cycles 4-8 years	Short cycles 1.5-4 years	Business cycle 1.5-8 years
1st PC	33.0	35.7	36.1	39.4	44.9	42.4
2nd PC	17.5	23.6	23.1	23.2	14.6	18.3
3rd PC	7.7	8.6	8.3	8.2	6.4	7.1
4th PC	6.1	6.7	6.6	6.0	4.4	5.2
5th PC	4.7	4.8	4.8	4.0	3.7	3.9
6th PC	3.9	3.3	3.5	3.3	3.1	3.3

Table 7: Percentage of overall variance explained by the first 6 principal component series by frequency band. All frequencies: $[0 \quad \pi]$; Long run: $[0 \quad 2\pi/80)$; Long cycles: $[2\pi/80 \quad 2\pi/32)$; Medium cycles: $[2\pi/32 \quad 2\pi/16)$; Short cycles: $[2\pi/16 \quad 2\pi/6)$; business cycle: $[2\pi/32 \quad 2\pi/6)$.

	PC series	All freq.	Long run > 20 years	Long cycles 8-20 years	Medium cycles 4-8 years	Short cycles 1.5-4 years	Bus. cycle 1.5-8 years
GDP	first 2	72.5	**84.2**	84.3	84.3	80.9	**82.0**
	next 3	16.2	**7.8**	7.0	5.7	9.8	**8.5**
Cons.	first 2	55.2	**69.9**	69.8	70.9	64.8	**67.3**
	next 3	15.9	**16.5**	15.7	14.9	13.0	**13.5**
Inv.	first 2	72.2	**77.9**	77.5	83.4	85.1	**84.4**
	next 3	8.7	**14.1**	13.4	5.9	3.3	**4.2**
U rate	first 2	80.2	**83.2**	81.8	82.2	90.9	**88.0**
	next 3	8.1	**11.7**	12.5	12.9	2.5	**5.8**
Hours	first 2	76.9	**87.0**	86.7	88.2	90.9	**89.8**
	next 3	8.2	**8.2**	8.4	8.0	2.0	**4.1**
Inflation	first 2	84.0	**91.5**	91.0	90.7	74.7	**86.2**
	next 3	7.4	5.2	**5.5**	6.0	12.4	**7.8**
FFR	first 2	78.2	**78.9**	78.7	81.4	73.4	**79.1**
	next 3	16.0	**17.5**	17.2	14.3	14.6	**14.8**

Table 8: Percentage of variance explained by the first 2 principal component series and the following 3 for a few selected variables, by frequency band. All frequencies: $[0 \quad \pi]$; Long run: $[0 \quad 2\pi/80)$; Long cycles: $[2\pi/80 \quad 2\pi/32)$; Medium cycles: $[2\pi/32 \quad 2\pi/16)$; Short cycles: $[2\pi/16 \quad 2\pi/6)$; business cycle: $[2\pi/32 \quad 2\pi/6)$.

	Factors	All freq.	Long run > 20 years	Long cycles 8-20 years	Medium cycles 4-8 years	Short cycles 1.5-4 years	Bus. cycle 1.5-8 years
GDP	1st	43.9	**2.2**	2.8	10.9	72.9	**52.9**
	2nd	28.6	**82.0**	81.5	73.3	8.0	**29.1**
Cons.	1st	27.2	**2.3**	2.3	6.7	53.4	**34.8**
	2nd	28.0	**67.6**	67.5	64.2	11.4	**32.5**
Inv.	1st	38.7	**5.4**	5.3	12.7	74.8	**54.7**
	2nd	33.5	**72.5**	72.2	70.7	10.3	**29.7**
U rate	1st	46.7	**10.7**	10.8	18.1	82.1	**62.1**
	2nd	33.5	**72.6**	70.9	64.1	8.8	**25.9**
Hours	1st	39.3	**0.4**	2.6	14.1	82.0	**59.3**
	2nd	37.6	**86.6**	84.2	74.2	8.9	**30.6**
Inflation	1st	**78.5**	91.4	90.1	86.2	45.6	**76.0**
	2nd	5.5	**0.1**	1.0	4.4	29.1	**10.2**
FFR	1st	74.9	**78.7**	78.1	78.4	57.5	**73.7**
	2nd	3.2	**0.2**	0.6	3.0	15.9	**5.4**

Table 9: Percentage of variance explained by the first and the second factor for a few selected variables, by frequency band. All frequencies: $[0 \quad \pi]$; Long run: $[0 \quad 2\pi/80)$; Long cycles: $[2\pi/80 \quad 2\pi/32)$; Medium cycles: $[2\pi/32 \quad 2\pi/16)$; Short cycles: $[2\pi/16 \quad 2\pi/6)$; business cycle: $[2\pi/32 \quad 2\pi/6)$.

Appendices

A Auxiliary Results

B Proofs

C Variables and transformations

The data are from FRED-QD. We retained the 216 series starting in 1960Q1. We report here the ID number and the mnemonic. For the description of each variable see McCracken and Ng (2020). Transformation codes: 1 = no transformation; 2 = first difference; 5 = log difference; 7 = first difference of the of percentage variation.

ID number	FRED-QD ID number	Transf. code	FRED-QD Mnemonic	ID number	FRED-QD ID number	Transf. code	FRED-QD Mnemonic
1	1	5	GDPC1	52	54	5	CES9091000001
2	2	5	PCECC96	53	55	5	CES9092000001
3	3	5	PCDGx	54	56	5	CES9093000001
4	4	5	PCESVx	55	57	5	CE16OV
5	5	5	PCNDx	56	58	2	CIVPART
6	6	5	GPDIC1	57	59	2	UNRATE
7	7	5	FPIx	58	60	2	UNRATESTx
8	8	5	Y033RC1Q027SBEAx	59	61	2	UNRATELTx
9	9	5	PNFIx	60	62	2	LNS14000012
10	10	5	PRFIx	61	63	2	LNS14000025
11	11	1	A014RE1Q156NBEA	62	64	2	LNS14000026
12	12	5	GCEC1	63	65	5	UEMPLT5
13	13	1	A823RL1Q225SBEA	64	66	5	UEMP5TO14
14	14	5	FGRECPTx	65	67	5	UEMP15T26
15	15	5	SLCEx	66	68	5	UEMP27OV
16	16	5	EXPGSC1	67	73	5	LNS12032194
17	17	5	IMPGSC1	68	74	5	HOABS
18	18	5	DPIC96	69	76	5	HOANBS
19	19	5	OUTNFB	70	77	1	AWHMAN
20	20	5	OUTBS	71	79	1	AWOTMAN
21	22	5	INDPRO	72	80	1	HWIx
22	23	5	IPFINAL	73	81	5	HOUST
23	24	5	IPCONGD	74	82	5	HOUST5F
24	25	5	IPMAT	75	83	5	PERMIT
25	26	5	IPDMAT	76	84	5	HOUSTMW
26	27	5	IPNMAT	77	85	5	HOUSTNE
27	28	5	IPDCONGD	78	86	5	HOUSTS
28	29	5	IPB51110SQ	79	87	5	HOUSTW
29	30	5	IPNCONGD	80	88	5	CMRMTSPLx
30	31	5	IPBUSEQ	81	89	5	RSAFSx
31	32	5	IPB51220SQ	82	90	5	AMDMNOx
32	34	1	CUMFNS	83	92	5	AMDMUOx
33	35	5	PAYEMS	84	95	5	PCECTPI
34	36	5	USPRIV	85	96	5	PCEPILFE
35	37	5	MANEMP	86	97	5	GDPCTPI
36	38	5	SRVPRD	87	98	5	GPDICTPI
37	39	5	USGOOD	88	99	5	IPDBS
38	40	5	DMANEMP	89	100	5	DGDSRG3Q086SBEA
39	41	5	NDMANEMP	90	101	5	DDURRG3Q086SBEA
40	42	5	USCONS	91	102	5	DSERRG3Q086SBEA
41	43	5	USEHS	92	103	5	DNDGRG3Q086SBEA
42	44	5	USFIRE	93	104	5	DHCERG3Q086SBEA
43	45	5	USINFO	94	105	5	DMOTRG3Q086SBEA
44	46	5	USPBS	95	106	5	DFDHRG3Q086SBEA
45	47	5	USLAH	96	107	5	DREQRG3Q086SBEA
46	48	5	USSERV	97	108	5	DODGRG3Q086SBEA
47	49	5	USMINE	98	109	5	DFXARG3Q086SBEA
48	50	5	USTPU	99	110	5	DCLORG3Q086SBEA
49	51	5	USGOVT	100	111	5	DGOERG3Q086SBEA
50	52	5	USTRADE	101	112	5	DONGRG3Q086SBEA
51	53	5	USWTRADE	102	113	5	DHUTRG3Q086SBEA

ID number	FRED-QD ID number	Transf. code	FRED-QD Mnemonic	ID number	FRED-QD ID number	Transf. code	FRED-QD Mnemonic
103	114	5	DHLCRG3Q086SBEA	160	187	5	EXCAUSx
104	115	5	DTRSRG3Q086SBEA	161	188	1	UMCSENTx
105	116	5	DRCARG3Q086SBEA	162	190	2	B020RE1Q156NBEA
106	117	5	DFSARG3Q086SBEA	163	191	2	B021RE1Q156NBEA
107	118	5	DIFSRG3Q086SBEA	164	194	5	IPMANSICS
108	119	5	DOTSRG3Q086SBEA	165	195	5	IPB51222S
109	120	5	CPIAUCSL	166	196	5	IPFUELS
110	121	5	CPILFESL	167	197	1	UEMPMEAN
111	122	5	WPSFD49207	168	198	1	CES0600000007
112	123	5	PPIACO	169	199	5	TOTRESNS
113	124	5	WPSFD49502	170	200	7	NONBORRES
114	125	5	WPSFD4111	171	201	1	GS5
115	126	5	PPIIDC	172	202	1	TB3SMFFM
116	127	5	WPSID61	173	203	1	T5YFFM
117	129	5	WPU0561	174	204	1	AAAFFM
118	130	5	OILPRICEx	175	205	5	WPSID62
119	132	5	CES2000000008x	176	206	5	PPICMM
120	133	5	CES3000000008x	177	207	5	CPIAPPSL
121	135	5	COMPRNFB	178	208	5	CPITRNSL
122	136	5	RCPHBS	179	209	5	CPIMEDSL
123	138	5	OPHNFB	180	210	5	CUSR0000SAC
124	139	5	OPHPBS	181	211	5	CUSR0000SAD
125	140	5	ULCBS	182	212	5	CUSR0000SAS
126	142	5	ULCNFB	183	213	5	CPIULFSL
127	143	5	UNLPNBS	184	214	5	CUSR0000SA0L2
128	144	1	FEDFUNDS	185	215	5	CUSR0000SA0L5
129	145	1	TB3MS	186	216	5	CES0600000008
130	146	1	TB6MS	187	217	5	DTCOLNVHFNM
131	147	1	GS1	188	218	5	DTCTHFNM
132	148	1	GS10	189	219	5	INVEST
133	150	1	AAA	190	220	1	HWIURATIOx
134	151	1	BAA	191	221	5	CLAIMSx
135	152	1	BAA10YM	192	222	5	BUSINVx
136	154	1	TB6M3Mx	193	223	1	ISRATIOx
137	155	1	GS1TB3Mx	194	224	1	CONSPIx
138	156	1	GS10TB3Mx	195	225	1	CP3M
139	157	1	CPF3MTB3Mx	196	226	1	COMPAPFF
140	158	5	BOGMBASEREALx	197	227	5	PERMITNE
141	160	5	M1REAL	198	228	5	PERMITMW
142	161	5	M2REAL	199	229	5	PERMITS
143	162	5	MZMREAL	200	230	5	PERMITW
144	163	5	BUSLOANSx	201	231	5	NIKKEI225
145	164	5	CONSUMERx	202	234	5	TLBSNNCBx
146	165	5	NONREVSLx	203	235	1	TLBSNNCBBDIx
147	166	5	REALLNx	204	236	5	TTAABSNNCBx
148	168	5	TOTALSLx	205	237	5	TNWMVBSNNCBx
149	170	5	TABSHNOx	206	238	2	TNWMVBSNNCBBDIx
150	171	5	TLBSHNOx	207	239	5	TLBSNNBx
151	172	5	LIABPIx	208	240	1	TLBSNNBBDIx
152	173	5	TNWBSHNOx	209	241	5	TABSNNBx
153	174	1	NWPIx	210	242	5	TNWBSNNBx
154	175	5	TARESAx	211	243	2	TNWBSNNBBDIx
155	176	5	HNOREMQ027Sx	212	244	5	CNCFx
156	177	5	TFAABSHNOx	213	245	5	S&P 500
157	184	5	EXSZUSx	214	246	5	S&P: indust
158	185	5	EXJPUSx	215	247	1	S&P div yield
159	186	5	EXUSUKx	216	248	5	S&P PE ratio

D Calibrating the window size

To better calibrate the window size of kDDR, kDER and kDGR for the macroeconomic data set used in the empirical application, we run a further simulation exercise. We set $n = 216$ (the number of series in the data set) and use two values for T, namely $T = 240$ (the time dimension of the whole sample) and $T = 120$ (the size of a few sub-samples used in the application). For the bandwidth parameter M_T, we use $M_T = \lfloor a\sqrt{T} \rceil$, with $a = 0.5, 0.75, 1, 1.25$. The DGP is the one used in the third experiment, Subsection 5.1, in the version with large idiosyncratic components.

Table A reports the percentage of correct outcomes over 500 replications. DDRcb denotes the Dynamic Difference Ratio Estimator evaluated on the cyclical band $[2\pi/32 \; 2\pi/6]$, which is used in the empirical application. Boldface numbers denote the window size parameter a which performs best for each value of q and each estimator.

Let us consider first the case $T = 240$, reported in the left part of the table. With $q = 2$, all bandwidths are able to capture the correct q for all replications. With $q = 4$, the bandwidth $M_T = \lfloor 0.5\sqrt{T} \rceil$ performs poorly, particularly for DDR and DDRcb. The other bandwidth perform well for all estimators. With $q = 6$, both the bandwidths with $a = 0.5$ and $a = 0.75$ have the worst performances, whereas the larger bandwidths perform well. Coming to the sample size $T = 120$ reported in the right part of the table, we see that all values of a produce good results for all estimators; however $a = 0.75$ and $a = 1$ perform somewhat better than $a = 0.5$ and $a = 1.25$. With $q = 4$, $a = 0.5$ and $a = 0.75$ perform poorly; $a = 1$ is the best choice for all but the DER estimator. Finally for $q = 6$ none of the estimators is able to capture the correct q for all values of a. Overall, we see from the table that $a = 1$ and $a = 1.25$ have the best performances. For the empirical application, we choose $a = 1$, corresponding to $M_T = \lfloor \sqrt{T} \rceil$.

		$T = 240$				$T = 120$			
q	a	DER	DGR	DDR	DDRcb	DER	DGR	DDR	DDRcb
2	0.50	100.0	100.0	100.0	100.0	97.8	99.4	99.8	97.0
	0.75	100.0	100.0	100.0	100.0	99.2	99.8	**100.0**	**99.6**
	1.00	100.0	100.0	100.0	100.0	**100.0**	**100.0**	**100.0**	97.4
	1.25	100.0	100.0	100.0	100.0	99.0	99.8	99.8	91.8
4	0.50	87.8	98.2	82.4	61.2	2.8	12.2	0.0	2.0
	0.75	**100.0**	**100.0**	**100.0**	99.6	39.8	63.2	50.4	70.0
	1.00	99.8	99.8	**100.0**	**100.0**	57.4	**77.8**	**94.6**	**73.0**
	1.25	99.8	**100.0**	**100.0**	99.8	**60.8**	77.0	92.2	60.2
6	0.50	0.2	1.6	0.0	0.0	0.0	0.0	0.0	0.0
	0.75	48.6	66.8	31.8	50.0	0.2	0.2	0.0	1.2
	1.00	84.4	91.2	95.2	**93.6**	0.8	0.8	0.4	**20.8**
	1.25	**91.0**	**95.2**	**99.6**	92.4	**1.2**	**1.4**	**15.0**	19.4

Table A: Calibration of the bandwidth parameter for the empirical application. The DGP is the model of the third experiment described in Section 5, large idiosyncratic components, with $q = 2, 4, 6$, $n = 216$, $T = 240, 120$. The bandwidth parameter is $M_T = \lfloor a\sqrt{T} \rceil$ with $a = 0.5, 0.75, 1, 1.25$. The Table reports the percentage of correct outcomes over 500 replications. DER: Dinamic Eigenvalue Ratio estimator, DGR: Dynamic Growth Ratio estimator; DDR: Dynamic Difference Ratio Estimator; DDRbc: Dynamic Difference Ratio Estimator evaluated on the cyclical band $[2\pi/32\ 2\pi/6]$. Boldface numbers denote the window size parameter a which performs best for each value of q and each estimator.