A wavelet-based multivariate multiscale approach for forecasting

Working Papers 2016

António Rua
A wavelet-based multivariate multiscale approach for forecasting

Working Papers 2016

António Rua
A wavelet-based multivariate multiscale approach
for forecasting

António Rua
Banco de Portugal
NOVA School of Business and
Economics

July 2016

Abstract
In an increasingly data rich environment, factor models have become the workhorse
approach for modelling and forecasting purposes. However, factors are non-observable and
have to be estimated. In particular, the space spanned by the unknown factors is typically
estimated via principal components. Herein, it is proposed a novel procedure to estimate
the factor space resorting to a wavelet based multiscale principal component analysis.
Through a Monte Carlo simulation study, it is shown that such an approach allows
to improve both factor model estimation and forecasting performance. In the empirical
application, one illustrates its usefulness for forecasting GDP growth and inflation in the
United States.

JEL: C22, C40, C53

Keywords: Wavelets, Multiscale Principal Components, Factor models, Forecasting.

Acknowledgements: I would like to thank all seminar participants at the 36th International
Symposium on Forecasting for helpful comments and suggestions on an earlier draft of the
paper.
E-mail: antonio.rua@bportugal.pt
1. Introduction

In an increasingly data-rich environment, factor models have become one of the most popular forecasting tools in the literature and among practitioners in central banks and international institutions. In fact, there is by now a huge strand of literature using factor models to forecast macroeconomic variables, namely GDP growth and inflation. See Stock and Watson (1999, 2002a), Banerjee and Marcellino (2006), Giannone, Reichlin and Small (2008) for the United States, Marcellino, Stock and Watson (2003), Camba-Méndez and Kapetanios (2005), Angelini et al. (2011) for the euro area, Artis, Banerjee and Marcellino (2005) for the United Kingdom, Barhoumi, Darné and Ferram (2010, 2013) for France, Schumacher (2007, 2010) and Schumacher and Breitung (2008) for Germany, Rünstler et al. (2009) for several European countries, among many others.

Typically, the factors in such forecasting models are estimated via principal components analysis. Such an approach draws on the work by Stock and Watson (1998, 2002b) and Bai and Ng (2002) who have shown that the principal components are consistent estimators of the true latent factors when the cross-section dimension and the number of observations tend to infinity. Moreover, feasible forecasts, constructed using the estimated factors and estimated parameters, are shown to be asymptotically efficient.

However, the finite sample performance of the principal components estimator worsens when the relative explanatory power of the factors decreases vis-à-vis the idiosyncratic components. Intuitively, when the relative importance of the idiosyncratic term increases, it becomes more difficult to distinguish the common from the idiosyncratic component. Based on Monte Carlo evidence, Boivin and Ng (2006) and Bai and Ng (2008) show that the estimated factors and forecasts are negatively affected by decreasing the relative importance of the common component. Such findings have been reinforced by the theoretical results concerning the finite sample properties of the principal components estimator, within a weakly influential factor asymptotics framework, as provided by Johnstone and Lu (2009) and Onatski (2012).

To cope with the above mentioned issue, we suggest a wavelet-based approach. Although wavelet analysis has been developed in other fields, such an approach has already proved to be useful in economics and finance. See, for instance, the pioneer work of Ramsey and Zhang (1996, 1997) and Ramsey and Lampart (1998a,b)). Recent applications of wavelets in the literature can

In particular, we resort to a wavelet-based multiscale principal components analysis to improve factor model estimation and forecasting performance. Multiscale principal components analysis has been initially proposed by Bakshi (1998) for multivariate statistical process control. The multiscale principal components analysis is a generalisation of principal components analysis and involves decomposing each variable on a selected family of wavelets. A principal components analysis is conducted independently at each scale and combined in an efficient scale-recursive way to yield the multiscale model. Hence, multiscale principal components analysis harvest the benefits of both principal components and wavelet analysis. On the one hand, the relationship between the variables is decorrelated by principal components while, on the other hand, each variable is decorrelated by the wavelet decomposition.

Herein, we show how one can take advantage of the multiscale principal components analysis to enhance the estimation of the space spanned by the true latent factors and to improve factor model forecasting behaviour. We provide a Monte Carlo simulation study to assess the suggested approach and we find that it delivers noteworthy gains vis-à-vis the principal components estimator. Furthermore, to illustrate the empirical usefulness of the suggested wavelet-based multivariate multiscale principal components approach, we assess its performance for forecasting GDP growth and inflation in the United States. Drawing on the large dataset compiled by Stock and Watson (2012), we find that the proposed approach outperforms significantly the factor model advocated therein which relies on principal components.

The paper is organized as follows. In section 2, the wavelet multiscale decomposition is overviewed and the multiscale principal components analysis is discussed. In section 3, the design of the Monte Carlo exercise is described and the corresponding simulation results are reported. In section 4, the empirical application is conducted by forecasting GDP growth and inflation in the United States and the results of the out-of-sample forecasting exercise are presented. Finally, section 5 concludes.
2. A wavelet-based multiscale approach

2.1. From wavelets to multiscale decomposition

The term wavelet denotes a small wave. The wave refers to the condition that this function $\psi(\cdot)$ is oscillatory. The smallness refers to the condition that it is of finite length, that is, compactly supported. Thus, a wavelet should satisfy two basic properties namely the integral of $\psi(\cdot)$ is zero,

$$\int \psi(t)dt = 0 \quad (1)$$

that is, the average value of the wavelet in the time domain must be zero; and the square of $\psi(\cdot)$ integrates to unity,

$$\int \psi^2(t)dt = 1 \quad (2)$$

which means that $\psi(\cdot)$ is limited to an interval of time. Although it departs from zero for a limited interval of time, the excursions above zero must cancel out with the excursions below zero. Thus, a wavelet is any function that integrates to zero and is square integrable.\(^1\)

The wavelet transform decomposes a time series in terms of some basis functions, the wavelets, analogous to the use of sines and cosines in Fourier analysis. Wavelets are a family of basis functions $\psi_{\tau,s}(t)$ that are localized in both time and frequency and are obtained by translation and dilation of the mother wavelet $\psi(t)$ as

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}}\psi \left( \frac{t - \tau}{s} \right) \quad (3)$$

where $\tau$ determines the time position (translation parameter), $s$ is the scale (dilation parameter) and $\frac{1}{\sqrt{s}}$ is for energy normalization across the different scales ($\|\psi_{\tau,s}\|^2 = \|\psi\|^2$).

---

\(^1\) In addition, it should also satisfy the so-called admissibility condition,

$$0 < \int_0^{+\infty} \left| \frac{\hat{\psi}(\omega)}{\omega} \right|^2 d\omega < +\infty$$

where $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$, that is, $\hat{\psi}(\omega) = \int_{-\infty}^{+\infty} \psi(t)e^{-i\omega t}dt$, so as to allow for the reconstruction of the time series without loss of information.
The Discrete Wavelet Transform (DWT) involves the discretization of the translation and dilation parameters (see, for example, Percival and Walden (2000)). In particular, these parameters are discretized dyadically as $s = 2^{-j}$ and $\tau = 2^{-j}k$. Any time series $x(t)$ can be decomposed as a weighted sum of dyadically discretized orthonormal basis functions as

$$x(t) = \sum_k d_{1,k} \psi_{1,k}(t) + \cdots + \sum_k d_{J,k} \psi_{J,k}(t) + \sum_k a_{J,k} \varphi_{J,k}(t)$$

(4)

where $J$ is the number scales, $d_{j,k}$ is the detail coefficient at scale $j$ and location $k$ whereas $a_{J,k}$ is the scaling function coefficient at the coarsest scale $J$ and location $k$. The terms $\sum_k d_{j,k} \psi_{j,k}(t)$ for $j = 1, 2, ..., J$ and $\sum_k a_{J,k} \varphi_{J,k}(t)$ represent the detail and smooth components, respectively. The wavelet transform coefficients are given by

$$d_{j,k} = \int x(t) \psi_{j,k}(t) dt$$

$$a_{J,k} = \int x(t) \varphi_{J,k}(t) dt$$

(5)  (6)

where

$$\psi_{j,k}(t) = 2^{-j/2} \psi \left(2^{-j} t - k\right)$$

$$\varphi_{J,k}(t) = 2^{-J/2} \varphi \left(2^{-J} t - k\right)$$

(7)  (8)

The scaling function or father wavelet $\varphi(.)$ captures the low frequency content of the series that is not captured by wavelets at the corresponding or finer scales.

For a given family of wavelets, efficient methods to compute the wavelet decomposition are based on the convolution of the time series with the corresponding wavelet filter $H$ and scaling filter $G$ (see, for example, Percival and Walden (2000)). Thus the coefficients can be obtained by

$$d_j = H_j x$$

$$a_j = G_j x$$

(9)  (10)

---

2. When the number of observations, $T$, is divisible by $2^J$ there are $T/2^J$ $d_{j,k}$ coefficients at scale $j = 1, ..., J - 1$, while at scale $J$ there are $T/2^J$ $d_{J,k}$ coefficients and $T/2^J$ $a_{J,k}$ coefficients. In total, there are $T$ wavelet coefficients, that is, $T = T/2^1 + T/2^2 + ... + T/2^{J-1} + T/2^J + T/2^J$. 

---
where \( x \) denotes the data vector, \( d_j \) is the vector of detail coefficients at scale \( j \), \( H_j \) is obtained by applying the \( G \) filter \( j - 1 \) times and the filter \( H \) once, \( a_j \) is the vector of scaling coefficients at scale \( j \) and \( G_j \) is obtained by applying the \( G \) filter \( j \) times. Hence, the DWT is implemented via the Mallat pyramid algorithm where in the first stage it consists in transforming \( x \) into first level wavelet coefficients \( d_1 \) and first level scaling coefficients \( a_1 \). In the second stage, it transforms the vector \( a_1 \) into the second level details \( d_2 \) and second level scaling coefficients \( a_2 \) and so on. This means that at the \( j^{th} \) stage (for \( j = 2, \ldots, J \)) the vector \( a_{j-1} \) is transformed, likewise \( x \) in the first stage, by applying the filters \( H \) and \( G \) to obtain \( d_j \) and \( a_j \). At the end of the \( J^{th} \) stage one can form the DWT coefficient vector as

\[
W = \begin{bmatrix}
    d_1 \\
    \vdots \\
    d_j \\
    a_J
\end{bmatrix}
\]  

(11)

One can also write

\[
W = Wx
\]

(12)

where \( W \) is \( T \times T \) real valued orthonormal matrix defining the DWT and satisfying \( W'W = I \) (\( T \times T \) identity matrix). Hence, from (9)-(12),

\[
W = \begin{bmatrix}
    H_1 \\
    \vdots \\
    H_J \\
    G_J
\end{bmatrix}
\]  

(13)

Based on (12) one can write

\[
x = W'W
\]

(14)

and drawing on the orthonormality of \( W \) one can show that
\[ \|x\|^2 = x'x \]
\[ = (W'W)'(W'W) \]
\[ = (H_1'd_1 + ... + H_J'd_J + G_J'a_J)'(H_1'd_1 + ... + H_J'd_J + G_J'a_J) \]
\[ = (d_1'1 + ... + d_J'H_J + a_J'G_J)'(H_1'd_1 + ... + H_J'd_J + G_J'a_J) \]
\[ = d_1'd_1 + ... + d_J'H_J + a_J'G_J \]
\[ = \|d_1\|^2 + ... + \|d_J\|^2 + \|a_J\|^2 \] (15)

Equation (15) implies that the DWT is an energy-preserving transformation allowing to decompose the variance of \( x \) on a scale-by-scale basis. This is a key feature which has been exploited by Fan and Gençay (2010) to distinguish between a white noise and a unit root process. As discussed by Fan and Gençay (2010), a white noise process has more energy at the lowest scale detail coefficients \( d_1 \) while declining towards the highest scale. In contrast, in the case of a unit root process the energy is basically concentrated in the scaling coefficients \( a_J \). Hence, a more (less) persistent series has more (less) energy at higher scales and less (more) at lower scales. We will also take advantage of this feature here and show how it can be used to improve factor model estimation.

2.2. Multiscale principal components analysis

In the previous section, we considered the case of a single series \( x \). Now, let \( X \) denote a data matrix \( T \times N \). Suppose that one applies the DWT to \( X \), that is, \( WX \). Firstly, one should note that the variance-covariance matrix of \( WX \) is the same as that of \( X \) since

\[ \|WX\|^2 = (WX)'WX = X'W'WX = X'X = \|X\|^2 \] (16)

An important implication of this result is that for principal component analysis, focusing on \( WX \) or \( X \) does not change the analysis. In fact, (16) implies that the loadings obtained with principal components of \( X \) and \( WX \) are identical. Moreover, the principal components of \( WX \) are the wavelet transform of the principal components of \( X \). To see this, let \( Z \) be the principal components matrix of \( X \) and \( A \) the corresponding loadings matrix, then

\[ X = ZA' \Leftrightarrow WX = (WZ)A' \] (17)
By extending the proof of (15) to the multivariate case and making use of (9) and (10) it can also be shown that

$$\|WX\|_2^2 = \|H_1X\|_2^2 + \ldots + \|H_JX\|_2^2 + \|G_JX\|_2^2$$

(18)

The above results are key to the subsequent analysis. Not only the variance-covariance of the wavelet transform is the same of the original data matrix but one can also decompose it terms of the contribution at multiple scales. Such a result allows one to conduct a principal component analysis at each scale independently of the other scales.

This leads to the so-called multiscale principal component analysis (as initially proposed by Bakshi (1998)). It consists in the following steps. First, the DWT is performed for each column of $X$ which produces the matrices $D_1, \ldots, D_J$ containing the detail coefficients and $A_J$ containing the scaling coefficients. For each scale, one selects the appropriate number of principal components or suppress the detail $j$. Then, reconstruct a new matrix $\tilde{X}$ containing the main features of the original matrix $X$. Finally, perform the principal component analysis of $\tilde{X}$. Note that, since, in general, $\text{rank}(\tilde{X}) = \text{rank}(X)$, the last step is needed to reduce dimensionality. Moreover, if no dimension reduction is performed at any scale, that is, all principal components are retained at each scale, then $\tilde{X}$ will be identical to $X$.

To show how can multiscale principal component analysis enhance factor model estimation let us consider the following factor structure in matrix form as

$$X = FA' + e$$

(19)

where $F$ is the $T \times r$ matrix of non-observable factors, $A$ is the $N \times r$ matrix of (unknown) factor loadings and $e$ is the $T \times N$ matrix of idiosyncratic errors. When both $N \to \infty$ and $T \to \infty$, Stock and Watson (1998, 2002b), Bai and Ng (2002), Bai (2003) and Amengual and Watson (2007) have shown that, under slightly different sets of assumptions regarding the data generating processes of

---

3. Without loss of generality, we consider the factor model within the static framework. A dynamic factor model with $q$ factors can be written as a static factor model with $r$ factors, where $r$ is finite. Key results regarding dynamic factor models with large datasets can be found in Forni et al. (2000, 2004, 2005).
the factors and the idiosyncratic components\textsuperscript{4}, the space spanned by the true factors can be consistently estimated via the first principal components.

However, the finite sample performance of the principal components estimator deteriorates when the explanatory power of the factors decreases vis-à-vis the explanatory power of the idiosyncratic errors. Based on a Monte Carlo analysis, Boivin and Ng (2006) and Bai and Ng (2008), show that the factor estimates and forecasts are adversely affected by decreasing the relative importance of the common component. Intuitively, when the importance of the idiosyncratic error is magnified, it becomes harder to disentangle the common from the idiosyncratic component in the data. This issue has been further investigated within a weakly influential factor asymptotics framework to assess the finite sample properties of the principal components estimator in a context of a relatively weak explanatory power of the factors. Johnstone and Lu (2009) show the inconsistency of the principal components estimator for the one-factor model with \textit{i.i.d} Gaussian factor and \textit{i.i.d} Gaussian idiosyncratic terms while Onatski (2012) extends the findings to more general dynamics.

The above discussion suggests that a possible way to improve the performance of the principal components estimator consists in downweighting the noise. As mentioned earlier, when discussing the multiscale decomposition of a time series, the noise tends to be reflected, to a larger extent, at the lowest scale. On the other hand, factors are typically smooth and are therefore better captured by the scaling function. Hence, to improve factor model estimation, we suggest to suppress the lowest detail components of $X$ while retaining the first principal components extracted at higher scales. In other words, since at the lowest scale the signal-to-noise ratio is expected to be very low one can suppress it without losing too much information while avoiding the problem of estimating via principal components in a context of weak factors. The data at the lowest scale should be basically uninformative for the factor structure. At higher scales, by retaining the first principal components on scale-by-scale basis it allows to be more sensitive to scale-varying signal features and can potentially enhance factor model estimation and forecasting.

\textsuperscript{4} The typical assumptions allow for some heteroskedasticity and limited dependence of the idiosyncratic components in both the time and cross-section dimensions, as well as for moderate correlation between the latter and the factors.
3. A Monte Carlo study

In this section, we conduct a Monte Carlo simulation study to support the above suggested approach.

3.1. The setup

Let us define a factor model with a data generating process given by

\[ x_{it} = \sum_{j=1}^{r} \Lambda_{ij} f_{jt} + e_{it} \]  \hspace{1cm} (20)

with \( i = 1, \ldots, N \) and \( t = 1, \ldots, T \) or simply in vector notation

\[ X_t = \Lambda F_t + e_t \]  \hspace{1cm} (21)

for \( t = 1, \ldots, T \),

\[ A(L)F_t = u_t \]  \hspace{1cm} (22)

with \( u_t \) i.i.d. \( \mathcal{N}(0, I_r) \),

\[ D(L)e_t = v_t \]  \hspace{1cm} (23)

with \( v_t \) i.i.d. \( \mathcal{N}(0, T) \),

\[ A_{ij}(L) = \begin{cases} 
1 - a_1 L - a_2 L^2 & \text{if } i = j \\
0 & \text{if } i \neq j 
\end{cases} \text{ for } i, j = 1, \ldots, r \]  \hspace{1cm} (24)

where \( a_1 = 2b \cos(\varphi) \) with \( \varphi \in (0, \pi) \) and \( a_2 = -b^2 \),

\[ D_{ij} = \begin{cases} 
1 - dL & \text{if } i = j \\
0 & \text{if } i \neq j 
\end{cases} \text{ for } i, j = 1, \ldots, N \]  \hspace{1cm} (25)

\[ A_{ij} \text{ i.i.d. } \mathcal{N}(0, 1) \]  \hspace{1cm} (26)

for \( i = 1, \ldots, N, j = 1, \ldots, r \),

\[ T_{ij} = \tau |i-j| (1 - d^2) \sqrt{\alpha_i \alpha_j} \]  \hspace{1cm} (27)

for \( i, j = 1, \ldots, N \) and
\[ \alpha_i = \frac{\beta_i}{1 - \beta_i (1 + a_2)} \frac{1 - a_2}{(1 - a_2)^2 - a_1^2} \sum_{j=1}^{r} A_{ij}^2 \]  

\((28)\)

A similar model has been used, for example, in Stock and Watson (2002b), Doz et al. (2012) and Pinheiro et al. (2013). Note that, herein we generate the factors as AR(2) processes in such a way that the resulting factors may display a cyclical behaviour (see, for example, Bierens (2001) and Castro et al. (2013)). This is motivated by the body of empirical literature where factors extracted from macroeconomic datasets are used as business cycle indicators (see the seminal work of Stock and Watson (1989, 1991, 1993) as well as Forni et al. (2001), Altissimo et al. (2001), Rua and Nunes (2005), Valle e Azevedo et al. (2006), Altissimo et al. (2010) among others).

In particular, \((24)\) generates factors with a cyclical pattern of \(2\pi/\varphi\) time periods with parameter \(b\) controlling for the persistence. In the simulations, we set \(\varphi\) equal to \(2\pi/23\) so as to generate a cyclical pattern with the average duration of post-war US business cycles according to NBER business cycle dating committee, that is, 23 quarters. As in Doz et al. (2012) simulation study, we considered a smooth process for the factors and set \(b\) equal to 0.9.

As regards the idiosyncratic components, the model allows a limited cross-correlation as \(T\) is a Toeplitz matrix with the parameter \(\tau\) controlling for the degree of cross-correlation. The dynamics of the idiosyncratic components is governed by an AR(1) process with autoregressive parameter \(d\). For the time being, we set \(\tau = 0\) and \(d = 0\) (see also Doz et al. (2012)).

The parameter \(\beta_i\) corresponds to the ratio between the variance of the idiosyncratic component \(e_{it}\) and the variance of the corresponding variable \(x_{it}\). This parameter controls for the relative importance between the common and the idiosyncratic components. As found for the US dataset considered by Stock and Watson (2012), we set this ratio equal to 0.6 with the number of factors \(r\) equal to 5.

In the case of the wavelet-based multiscale approach, we consider a Daubechies wavelet filter of length 4 and \(J = 1\) as in Fan and Gençay (2010). Hence, we obtain two components namely a detail and a smooth component, and corresponding coefficients, \(D_1\) and \(A_1\). As mentioned earlier, the suggested approach involves suppressing the lowest detail while computing the first \(r\) principal components for the highest scale and finally retaining \(r\) factors from the corresponding reconstructed data matrix.
Several measures are computed for the comparison between the standard principal components estimator and the multiscale approach and the results are based on 1000 sample draws. Following Stock and Watson (2002b), Doz et al. (2012) and Pinheiro et al. (2013) we compute the trace $R^2$ as

$$R^2_{\hat{F}, \hat{F}} = \frac{\mathbb{E} \left[ \text{tr} \left( \hat{F}' \hat{F} \left( \hat{F}' \hat{F} \right)^{-1} \hat{F}' \hat{F} \right) \right]}{\mathbb{E} \left[ \text{tr} \left( \hat{F}' \hat{F} \right) \right]}$$

(29)

where $\mathbb{E} [\cdot]$ denotes the expectation estimated by averaging the relevant statistic over the 1000 draws and $\hat{F}$ are the estimated factors. This statistic is a measure of fit of the multivariate regression of the true factors on the estimated factors, and is commonly used because the factors are identified only up to a rotation. A value higher and closer to one denotes a better estimation of the space spanned by the true factors.

Besides assessing the improvement in terms of the estimation of the factor space, we also evaluate the potential forecasting gains that might result from the suggested wavelet-based multiscale approach. Hence, similarly to Stock and Watson (2002b), we consider the scalar variable to be forecasted as generated by

$$y_{t+1} = \beta' F_t + \varepsilon_{t+1}$$

(30)

where $\beta$ is a $r \times 1$ vector of 1s and $\varepsilon_{t+1} \text{i.i.d. } \mathcal{N}(0,1)$ and independent of the other errors in the above specified factor model. The out-of-sample forecast is given by $\hat{y}_{t+1|T} = \sum_{j=1}^{r} \hat{\beta}_j \hat{f}_j$, where $\hat{\beta}$ are the OLS coefficients in the regression of $y_{t+1}$ onto $\hat{f}_j$, $j = 1, ..., r$, $t = 1, ..., T - 1$ and $\hat{f}_j$ denote the estimated factors either via principal components or via the wavelet-based multiscale approach.

The mean squared forecast error (MSFE) is computed for both approaches.

### 3.2. Simulation results

Let us consider different sizes for the cross-section and sample length $(N, T = 25, 50, 75, 100, 150, 200)$. The simulation results are presented in Table 1. In particular, we report the trace $R^2$ for the wavelet-based multiscale principal components ($R^2_{\text{WMSPC}}$), the relative trace $R^2$ ($R^2_{\text{WMSPC}} / R^2_{\text{PC}}$) and the relative MSFE ($\text{MSFE}_{\text{WMSPC}} / \text{MSFE}_{\text{PC}}$) between the two approaches.
A wavelet-based multivariate multiscale approach for forecasting

Table 1. Monte Carlo results

<table>
<thead>
<tr>
<th></th>
<th>$N = 25$</th>
<th>$N = 50$</th>
<th>$N = 75$</th>
<th>$N = 100$</th>
<th>$N = 150$</th>
<th>$N = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2_{WMSPC}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T = 25$</td>
<td>0.72</td>
<td>0.75</td>
<td>0.77</td>
<td>0.77</td>
<td>0.78</td>
<td>0.77</td>
</tr>
<tr>
<td>$T = 50$</td>
<td>0.79</td>
<td>0.84</td>
<td>0.86</td>
<td>0.87</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>$T = 75$</td>
<td>0.81</td>
<td>0.87</td>
<td>0.89</td>
<td>0.90</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>$T = 100$</td>
<td>0.82</td>
<td>0.88</td>
<td>0.91</td>
<td>0.92</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>$T = 150$</td>
<td>0.83</td>
<td>0.90</td>
<td>0.92</td>
<td>0.93</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>$T = 200$</td>
<td>0.84</td>
<td>0.90</td>
<td>0.93</td>
<td>0.94</td>
<td>0.95</td>
<td>0.96</td>
</tr>
</tbody>
</table>

| $R^2_{WMSPC} / R^2_{PC}$ |          |          |          |            |          |          |
| $T = 25$ | 1.09     | 1.05     | 1.04     | 1.03       | 1.02      | 1.01      |
| $T = 50$ | 1.12     | 1.07     | 1.05     | 1.03       | 1.02      | 1.02      |
| $T = 75$ | 1.13     | 1.07     | 1.05     | 1.03       | 1.02      | 1.02      |
| $T = 100$ | 1.14    | 1.07     | 1.05     | 1.03       | 1.02      | 1.02      |
| $T = 150$ | 1.14    | 1.07     | 1.05     | 1.03       | 1.02      | 1.02      |
| $T = 200$ | 1.14    | 1.07     | 1.05     | 1.03       | 1.02      | 1.02      |

| MSFE$_{WMSPC} / MSFE_{PC}$ |          |          |          |            |          |          |
| $T = 25$ | 0.69     | 0.82     | 0.84     | 0.86       | 0.86      | 0.90      |
| $T = 50$ | 0.76     | 0.78     | 0.84     | 0.87       | 0.91      | 0.92      |
| $T = 75$ | 0.72     | 0.78     | 0.82     | 0.84       | 0.90      | 0.94      |
| $T = 100$ | 0.77    | 0.82     | 0.86     | 0.91       | 0.94      | 0.95      |
| $T = 150$ | 0.79    | 0.83     | 0.84     | 0.90       | 0.93      | 0.91      |
| $T = 200$ | 0.77    | 0.83     | 0.87     | 0.89       | 0.95      | 0.96      |

Firstly, as the size of the cross-section and sample length increase, the trace $R^2$ for the wavelet-based multiscale principal components increases while approaching one when both $N$ and $T$ are large. Even for small $N$ and $T$, the trace $R^2$ is relatively high. More importantly, the trace $R^2$ of wavelet-based multiscale approach is always higher than that of principal components for all $N$ and $T$ considered, as $R^2_{WMSPC} / R^2_{PC}$ is always above one. That is, the wavelet-based multiscale approach delivers a better estimation of the space spanned by the true factors whatever the size of the cross-section and sample length. One should note that the gains vary depending on $N$ and $T$. In particular, the relative improvement over the principal components estimator, in terms of the trace $R^2$, is higher for small $N$ attaining up to 14 per cent. As one increases $N$ the gains decrease but even for very large $N$ and $T$ one still records a gain of around 2 per cent.

On top of the previous findings, the wavelet-based multiscale approach also leads to noteworthy forecasting gains vis-à-vis principal components with the ratio $MSFE_{WMSPC} / MSFE_{PC}$ being below one for all $N$ and $T$. In particular,
the forecasting improvement is higher for smaller datasets with the gains reaching more than 30 per cent. Note that, even for relatively large datasets the gains are substantial ranging from more than 20 per cent up to 10 per cent in most cases. Overall, the simulation results highlight the usefulness of the suggested approach for the estimation of the factor space and for factor-augmented forecasting.

<table>
<thead>
<tr>
<th>( N = 25 )</th>
<th>( N = 50 )</th>
<th>( N = 75 )</th>
<th>( N = 100 )</th>
<th>( N = 150 )</th>
<th>( N = 200 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R^2_{WMSPC} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T = 25 )</td>
<td>0.75</td>
<td>0.76</td>
<td>0.77</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td>( T = 50 )</td>
<td>0.81</td>
<td>0.85</td>
<td>0.87</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>( T = 75 )</td>
<td>0.83</td>
<td>0.89</td>
<td>0.90</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>( T = 100 )</td>
<td>0.84</td>
<td>0.90</td>
<td>0.92</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>( T = 150 )</td>
<td>0.85</td>
<td>0.92</td>
<td>0.94</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>( T = 200 )</td>
<td>0.86</td>
<td>0.92</td>
<td>0.94</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>( R^2_{WMSPC}/R^2_{PC} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T = 25 )</td>
<td>1.07</td>
<td>1.04</td>
<td>1.03</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>( T = 50 )</td>
<td>1.10</td>
<td>1.05</td>
<td>1.03</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>( T = 75 )</td>
<td>1.12</td>
<td>1.06</td>
<td>1.03</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>( T = 100 )</td>
<td>1.12</td>
<td>1.05</td>
<td>1.03</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>( T = 150 )</td>
<td>1.13</td>
<td>1.05</td>
<td>1.03</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>( T = 200 )</td>
<td>1.13</td>
<td>1.05</td>
<td>1.03</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>( MSFE_{WMSPC}/MSFE_{PC} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T = 25 )</td>
<td>0.75</td>
<td>0.83</td>
<td>0.88</td>
<td>0.88</td>
<td>0.89</td>
</tr>
<tr>
<td>( T = 50 )</td>
<td>0.79</td>
<td>0.84</td>
<td>0.87</td>
<td>0.87</td>
<td>0.91</td>
</tr>
<tr>
<td>( T = 75 )</td>
<td>0.72</td>
<td>0.75</td>
<td>0.87</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>( T = 100 )</td>
<td>0.78</td>
<td>0.85</td>
<td>0.89</td>
<td>0.90</td>
<td>0.96</td>
</tr>
<tr>
<td>( T = 150 )</td>
<td>0.78</td>
<td>0.85</td>
<td>0.89</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>( T = 200 )</td>
<td>0.77</td>
<td>0.86</td>
<td>0.89</td>
<td>0.92</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table 2. Monte Carlo results with heteroskedastic idiosyncratic components

As a sensitivity analysis, we relax the assumption of homoskedastic idiosyncratic components. As in Doz et al. (2012) and Pinheiro et al. (2013), we consider heteroskedastic idiosyncratic components by setting \( \beta_i \ i.i.d. U([0.1,0.9]) \). From the results presented in Table 2, one can conclude that the above findings are basically unchanged. In Table 3, we present the simulation results allowing for moderate serial correlation and cross-correlation among idiosyncratic components by setting \( \tau = 0.5 \) and \( d = 0.5 \) (see Doz et al. (2012)) departing from the exact factor model. Although, in general, the trace \( R^2 \) is slightly lower and both the estimation and forecasting gains are smaller,
the wavelet-based approach continues to outperform the principal components approach.

<table>
<thead>
<tr>
<th></th>
<th>$N = 25$</th>
<th>$N = 50$</th>
<th>$N = 75$</th>
<th>$N = 100$</th>
<th>$N = 150$</th>
<th>$N = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T = 25$</td>
<td>0.73</td>
<td>0.76</td>
<td>0.77</td>
<td>0.77</td>
<td>0.78</td>
<td>0.77</td>
</tr>
<tr>
<td>$T = 75$</td>
<td>0.75</td>
<td>0.82</td>
<td>0.84</td>
<td>0.85</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>$T = 100$</td>
<td>0.73</td>
<td>0.83</td>
<td>0.87</td>
<td>0.89</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>$T = 150$</td>
<td>0.72</td>
<td>0.84</td>
<td>0.89</td>
<td>0.91</td>
<td>0.93</td>
<td>0.94</td>
</tr>
<tr>
<td>$T = 200$</td>
<td>0.72</td>
<td>0.85</td>
<td>0.89</td>
<td>0.92</td>
<td>0.94</td>
<td>0.94</td>
</tr>
</tbody>
</table>

\[ R^2_{WMSPC} \]

<table>
<thead>
<tr>
<th></th>
<th>$T = 25$</th>
<th>$T = 50$</th>
<th>$T = 75$</th>
<th>$T = 100$</th>
<th>$T = 150$</th>
<th>$T = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T = 25$</td>
<td>1.04</td>
<td>1.02</td>
<td>1.01</td>
<td>1.01</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$T = 50$</td>
<td>1.05</td>
<td>1.03</td>
<td>1.02</td>
<td>1.01</td>
<td>1.01</td>
<td>1.00</td>
</tr>
<tr>
<td>$T = 75$</td>
<td>1.06</td>
<td>1.03</td>
<td>1.02</td>
<td>1.01</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>$T = 100$</td>
<td>1.07</td>
<td>1.04</td>
<td>1.02</td>
<td>1.01</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>$T = 150$</td>
<td>1.08</td>
<td>1.04</td>
<td>1.02</td>
<td>1.01</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>$T = 200$</td>
<td>1.09</td>
<td>1.04</td>
<td>1.02</td>
<td>1.01</td>
<td>1.01</td>
<td>1.01</td>
</tr>
</tbody>
</table>

\[ MSFE_{WMSPC}/MSFE_{PC} \]

<table>
<thead>
<tr>
<th></th>
<th>$T = 25$</th>
<th>$T = 50$</th>
<th>$T = 75$</th>
<th>$T = 100$</th>
<th>$T = 150$</th>
<th>$T = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T = 25$</td>
<td>0.81</td>
<td>0.95</td>
<td>0.98</td>
<td>0.97</td>
<td>0.99</td>
<td>1.01</td>
</tr>
<tr>
<td>$T = 50$</td>
<td>0.93</td>
<td>0.91</td>
<td>0.95</td>
<td>0.96</td>
<td>0.99</td>
<td>0.97</td>
</tr>
<tr>
<td>$T = 75$</td>
<td>0.88</td>
<td>0.89</td>
<td>0.94</td>
<td>0.93</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>$T = 100$</td>
<td>0.89</td>
<td>0.93</td>
<td>0.94</td>
<td>0.97</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>$T = 150$</td>
<td>0.89</td>
<td>0.93</td>
<td>0.92</td>
<td>0.95</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>$T = 200$</td>
<td>0.86</td>
<td>0.91</td>
<td>0.92</td>
<td>0.94</td>
<td>0.97</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Table 3. Monte Carlo results with serial and cross-correlation in the idiosyncratic components

4. Forecasting US GDP growth and inflation

To illustrate the empirical usefulness of the suggested wavelet-based multivariate multiscale principal components approach, we evaluate its performance for forecasting GDP growth and inflation in the United States.

4.1. Data

In particular, we resort to the large dataset of Stock and Watson (2012) which comprises 143 quarterly series for the United States spanning 49 years, from 1960 up to 2008. This macro dataset includes GDP and its
components, industrial production and capacity utilization, employment and hours, unemployment, housing starts, inventories and new orders, consumer prices and commodity prices, hourly earnings and unit labor costs, interest rates and spreads, monetary aggregates, exchange rates, stock prices and consumer expectations. As described in Stock and Watson (2012), the series are transformed by taking logarithms and/or differencing. In particular, first differences of logarithms are used for real variables, first differences are used for nominal interest rates and second differences of logarithms for prices. See Stock and Watson (2012) for a detailed list and further details on the series.

4.2. Design of the exercise

Based on the above mentioned dataset, Stock and Watson (2012) conduct an empirical comparison of various forecasting methods designed for a large number of predictors. They find that factor model forecasts using the first five principal components as predictors outperform all the alternative methods. Hence, this evidence suggests setting such a forecasting model as the benchmark to beat.

The factor model to be considered to forecast $y$ is based on the least squares estimation of equation

$$y_{t+h} = \beta_0 + \sum_{j=1}^{r} \beta_j \hat{f}_j + \sum_{i=1}^{p} \gamma_i y_{t+1-i} + \varepsilon_{t+h} \quad (t = p, \cdots, T - h)$$

(31)

where $y_{t+h}$ denote the variable to be forecasted with a forecasting horizon of $h$ periods. In the case of GDP, $y_{t+h}$ is the $h$-period growth whereas for inflation, $y_{t+h}$ is the $h$-period change in inflation. Regarding the factors, we set $r$ equal to five as in Stock and Watson (2012) with $\hat{f}_j$ denoting the estimated factors either via principal components or via the wavelet-based multiscale approach. As usual, the number $p$ of autoregressive terms is determined by the standard BIC criterion.

We consider several forecasting horizons namely $h = 1, \ldots, 8$ and the out-of-sample period runs from 1985Q1 up to 2008Q4 as in Stock and Watson (2012)
which corresponds to about half of the sample size. The \( h \)-step ahead forecasts for the variable \( y \) are obtained via a recursive forecasting exercise with recursive factor estimation, parameter estimation, model selection, and so forth.

The forecasting performance is evaluated by comparing the MSFE. In particular, we report the relative MSFE \( (MSFE_{W_{MSPC}}/MSFE_{PC}) \) between the two approaches. To reinforce the findings, we assess the statistical significance of the forecasting gains by computing the Harvey et al. (1997) modified version of the Diebold and Mariano (1995) test.

### 4.3. Empirical results

In Table 4, we present the results for GDP growth and for both inflation and core inflation, based on the CPI all items and CPI excluding food and energy, respectively. The entries in the table correspond to the relative MSFE between the wavelet multiscale principal components approach and the standard principal components. Hence, a ratio lower than one means that the former outperforms the latter. In the case of the wavelet-based approach, as in the Monte Carlo study, we consider \( J = 1 \) and suppress the lowest detail while computing the first \( r \) principal components for the highest scale and retaining \( r \) factors from the corresponding reconstructed data matrix.\(^6\)

<table>
<thead>
<tr>
<th>Horizon ( h )</th>
<th>GDP growth</th>
<th>Inflation</th>
<th>Core inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h = 1 )</td>
<td>0.90</td>
<td>1.05</td>
<td>0.76***</td>
</tr>
<tr>
<td>( h = 2 )</td>
<td>0.78</td>
<td>0.98</td>
<td>0.65***</td>
</tr>
<tr>
<td>( h = 3 )</td>
<td>0.81</td>
<td>0.91*</td>
<td>0.64***</td>
</tr>
<tr>
<td>( h = 4 )</td>
<td>0.80*</td>
<td>0.98</td>
<td>0.66***</td>
</tr>
<tr>
<td>( h = 5 )</td>
<td>0.73**</td>
<td>0.94</td>
<td>0.66***</td>
</tr>
<tr>
<td>( h = 6 )</td>
<td>0.71**</td>
<td>0.88</td>
<td>0.64***</td>
</tr>
<tr>
<td>( h = 7 )</td>
<td>0.73**</td>
<td>0.92</td>
<td>0.68***</td>
</tr>
<tr>
<td>( h = 8 )</td>
<td>0.72**</td>
<td>0.78</td>
<td>0.70**</td>
</tr>
</tbody>
</table>

**Table 4. Out-of-sample forecasting evaluation**

Note: The entries in the table denote the relative MSFE between the wavelet-based and principal components approach. The *** denote statistical significance of the forecasting gains at the 10, 5 and 1 per cent significance levels respectively.

\(^6\) In addition, we considered the case of \( J > 1 \) with the number of factors determined on a scale-by-scale basis but the forecasting performance, in overall terms, does not improve over the case where \( J = 1 \).
Overall, one can see that for almost all variables and forecasting horizons, the wavelet based approach improves on the factor model based on standard principal components. Although the relative gains seem to be a bit lower for the one-quarter ahead horizon, in most cases one obtains noteworthy forecasting improvements. In particular, for GDP growth the gains attained are, on average, more than 20 per cent and near 30 per cent in some horizons. Even for the one-quarter ahead forecasts, where the improvement is the lowest, one obtains a 10 per cent gain. In the case of inflation, the ratio is, in general, lower than one but the magnitude of the gains are smaller and only around 10 per cent, on average. However, when one focuses on core inflation, the forecasting improvements are striking reaching more than 30 per cent, on average. The gains range from 24 per cent for one-quarter ahead forecasts up to 36 per cent. These findings clearly support empirically the wavelet-based multiscale approach.

In the above forecasting exercise, we discarded the lowest detail by selecting no principal components at the lowest scale and set the number of factors to be retained informed by the work of Stock and Watson (2012). However, in practice, the number of factors to be considered is usually not known. To assess how one can cope with such a practical issue, we resort to the recently proposed criterion by Onatski (2010) which determines the number of factors based on the empirical distribution of the eigenvalues. Several reasons support the use of this criterion. On the one hand, Onatski (2010) shows that this criterion performs the best among a set of alternative criteria suggested in the literature, including the well-known Bai and Ng (2002) criteria. On the other hand, this criterion is particularly suitable in a context of weakly influential factors as shown by Onatski (2012). This is quite relevant in our case given that, as discussed earlier, one expects the explanatory power of the factors to be relatively small namely at the lowest scale.

Drawing on the above criterion, we find that no factors are selected at the lowest scale while the number of factors for the highest scale and for the reconstructed data is two. Note that the first finding provides additional empirical cross-validation for the above suggested procedure of supressing the lowest detail whereas the fact that only two factors are retained is in line with the results of Onatski (2012) using US data. In Table 5, we present the corresponding relative forecasting behaviour.
A wavelet-based multivariate multiscale approach for forecasting

<table>
<thead>
<tr>
<th>Horizon</th>
<th>GDP growth</th>
<th>Inflation</th>
<th>Core inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = 1$</td>
<td>1.00</td>
<td>1.01</td>
<td>0.70***</td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.88</td>
<td>0.99</td>
<td>0.63**</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>0.84</td>
<td>0.96</td>
<td>0.63**</td>
</tr>
<tr>
<td>$h = 4$</td>
<td>0.83</td>
<td>1.05</td>
<td>0.63**</td>
</tr>
<tr>
<td>$h = 5$</td>
<td>0.77</td>
<td>0.95</td>
<td>0.49**</td>
</tr>
<tr>
<td>$h = 6$</td>
<td>0.81</td>
<td>0.84</td>
<td>0.40**</td>
</tr>
<tr>
<td>$h = 7$</td>
<td>0.86</td>
<td>0.88</td>
<td>0.44*</td>
</tr>
<tr>
<td>$h = 8$</td>
<td>0.86</td>
<td>0.79</td>
<td>0.45*</td>
</tr>
</tbody>
</table>

Table 5. Out-of-sample forecasting evaluation with criterion determined number of factors

Note: The entries in the table denote the relative MSFE between the wavelet-based and principal components approach. The *, **, *** denote statistical significance of the forecasting gains at the 10, 5 and 1 per cent significance levels respectively.

The results in Tables 4 and 5 are broadly similar from a qualitatively point of view. As in the previously discussed exercise, we also find that for almost all variables and forecasting horizons, the wavelet based approach improves the forecasting performance. In particular, in the case of GDP growth, the forecasting gains are, on average, 15 per cent which is slightly lower than reported in Table 4. For inflation, the results are relatively close in overall terms. Finally, for core inflation, the improvement is even larger attaining, on average, 45 per cent, and being particularly visible for longer horizons. Hence, even if the number of factors selected at each scale is determined resorting to a criterion, the wavelet-based approach continues to deliver noteworthy forecasting gains.

5. Conclusions

In a context of growing data availability, the use of factor models has become widespread as it allows to take on board large datasets in a intuitive and parsimonious way. The estimation of the latent factors, which are subsequently used to obtain factor-augmented forecasts, is usually done via the principal components technique. Although it consistently estimates the space spanned by the true factors when the cross-section dimension and the number of observations tend to infinity, the finite sample performance of the principal components estimator deteriorates substantially when the explanatory power of the factors decreases vis-à-vis the explanatory power of the idiosyncratic
errors. Naturally, this impacts negatively on the forecasting performance of factor models.

Herein, we suggest a wavelet-based approach to cope with the above mentioned shortcoming. In particular, we propose estimating the factor model through a wavelet-based multiscale principal components analysis. Such an approach merges the benefits of principal components, which captures the relationship among the variables, and wavelet analysis, which enhances the decomposition of each variable dynamics. Based on a Monte Carlo simulation study, we show that it improves factor model estimation and forecasting performance.

Furthermore, we apply the suggested procedure to forecast GDP growth and inflation in the United States which are key variables for policymaking. We find that the wavelet-based approach delivers noteworthy forecasting gains over a wide range of forecasting horizons. On average, the forecasting improvement is more than 20 per cent in the case of GDP growth, around 10 per cent for inflation and more than 30 per cent for core inflation. These findings seem promising and reinforce the usefulness of wavelets to enhance forecasting performance.
References


WORKING PAPERS

2013

01|13  Macroeconomic forecasting using low-frequency filters  
   João Valle e Azevedo | Ana Pereira

02|13  Everything you always wanted to know about sex discrimination  
   Ana Rute Cardoso | Paulo Guimarães | Pedro Portugal

03|13  Is there a role for domestic demand pressure on export performance?  
   Paulo Soares Esteves | António Rua

04|13  Ageing and fiscal sustainability in a small euro area economy  
   Gabriela Castro | José R. Maria | Ricardo Mourinho Félix | Cláudia Rodrigues Braz

05|13  Mind the gap! The relative wages of immigrants in the Portuguese labour market  
   Sónia Cabral | Cláudia Duarte

06|13  Foreign direct investment and institutional reform: Evidence and an application to Portugal  
   Paulo Júlio | Ricardo Pinheiro-Alves | José Tavares

07|13  Monetary policy shocks: We got news!  
   Sandra Gomes | Nikolay Iskrev | Caterina Mendicino

08|13  Competition in the Portuguese Economy: Estimated price-cost margins under imperfect labour markets  
   João Amador | Ana Cristina Soares

09|13  The sources of wage variation: a three-way high-dimensional fixed effects regression model  
   Sonia Torres | Pedro Portugal | John T. Addison | Paulo Guimarães

10|13  The output effects of (non-separable) government consumption at the zero lower bound  
   Valerio Ercolani | João Valle e Azevedo

11|13  Fiscal multipliers in a small euro area economy: How big can they get in crisis times?  
   Gabriela Castro | Ricardo M. Félix | Paulo Julio | Jose R. Maria

12|13  Survey evidence on price and wage rigidities in Portugal  
   Fernando Martins

13|13  Characterizing economic growth paths based on new structural change tests  
   Nuno Sobreira | Luis C. Nunes | Paulo M. M. Rodrigues

14|13  Catastrophic job destruction  
   Anabela Carneiro | Pedro Portugal | José Varejão

15|13  Output effects of a measure of tax shocks based on changes in legislation for Portugal  
   Manuel Coutinho Pereira | Lara Wemans

16|13  Inside PESSOA - A detailed description of the model  
   Vanda Almeida | Gabriela Castro | Ricardo M. Félix | Paulo Júlio | José R. Maria

17|13  Macropuodential regulation and macroeconomic activity  
   Sudipto Karmakar

18|13  Bank capital and lending: An analysis of commercial banks in the United States  
   Sudipto Karmakar | Junghwan Mok
2014

1|14 Autoregressive augmentation of MIDAS regressions
   Cláudia Duarte

2|14 The risk-taking channel of monetary policy – exploring all avenues
   Diana Bonfim | Carla Soares

3|14 Global value chains: Surveying drivers, measures and impacts
   João Amador | Sónia Cabral

4|14 Has US household deleveraging ended? a model-based estimate of equilibrium debt
   Bruno Albuquerque | Ursel Baumann | Georgi Krustev

5|14 The weather effect: estimating the effect of voter turnout on electoral outcomes in Italy
   Alessandro Sforza

6|14 Persistence in the banking industry: fractional integration and breaks in memory
   Uwe Hassler | Paulo M.M. Rodrigues | Antonio Rubia

7|14 Financial integration and the great leveraging
   Daniel Carvalho

8|14 Euro area structural reforms in times of a global crisis
   Sandra Gomes

9|14 Labour demand research: towards a better match between better theory and better data
   John T. Addison | Pedro Portugal | José Varejão

10|14 Capital inflows and euro area long-term interest rates
   Daniel Carvalho | Michael Fidora

11|14 Misallocation and productivity in the lead up to the Eurozone crisis
   Daniel A. Dias | Carlos Robalo Marquesz | Christine Richmond

12|14 Global value chains: a view from the euro area
   João Amador | Rita Cappariello | Robert Stehrer

13|14 A dynamic quantitative macroeconomic model of bank runs
   Elena Mattana | Ettore Panetti

14|14 Fiscal devaluation in the euro area: a model-based analysis
   S. Gomes | P. Jacquinot | M. Pisani

15|14 Exports and domestic demand pressure: a dynamic panel data model for the euro area countries
   Elena Bobeica | Paulo Soares Esteves | António Rua | Karsten Staehr

16|14 Real-time nowcasting the US output gap: singular spectrum analysis at work
   Miguel de Carvalho | António Rua
2015

1|15  Unpleasant debt dynamics: can fiscal consolidations raise debt ratios?
   Gabriela Castro | Ricardo M. Félix | Paulo Júlio | José R. Maria

2|15  Macroeconomic forecasting starting from survey nowcasts
   João Valle e Azevedo | Inês Gonçalves

3|15  Capital regulation in a macroeconomic model with three layers of default
   Laurent Clerc | Alexis Derviz | Caterina Mendicino | Stephane Moyen | Kalin Nikolov | Livio Stracca | Javier Suarez | Alexandros P. Vardoulakis

4|15  Expectation-driven cycles: time-varying effects
   Antonello D’Agostino | Caterina Mendicino

5|15  Seriously strengthening the tax-benefit link
   Pedro Portugal | Pedro S. Raposo

6|15  Unions and collective bargaining in the wake of the great recession
   John T. Addison | Pedro Portugal | Hugo Vilares

7|15  Covariate-augmented unit root tests with mixed-frequency data
   Cláudia Duarte

8|15  Financial fragmentation shocks
   Gabriela Castro | José R. María | Paulo Júlio | Ricardo M. Félix

9|15  Central bank interventions, demand for collateral, and sovereign borrowing cost
   Luís Fonseca | Matteo Crosignani | Miguel Faria-e-Castro

10|15 Income smoothing mechanisms after labor market transitions
   Nuno Alves | Carlos Martins

11|15 Decomposing the wage losses of displaced workers: the role of the reallocation of workers into firms and job titles
   Anabela Carneiro | Pedro Raposo | Pedro Portugal

12|15 Sources of the union wage gap: results from high-dimensional fixed effects regression models
   John T. Addison | Pedro Portugal | Hugo Vilares

13|15 Assessing european firms’ exports and productivity distributions: the compnet trade module
   Antoine Berthou | Emmanuel Dhyne | Matteo Buganelli | Ana-Maria Cazacu | Calin-Vlad Demian | Peter Harasztosi | Tibor Lalinsky | Jaanika Meriküll | Filippo Oropallo | Ana Cristina Soares

14|15 A new regression-based tail index estimator: an application to exchange rates
   João Nicolau | Paulo M. M. Rodrigues

15|15 The effect of bank shocks on firm-level and aggregate investment
   João Amador | Arne J. Nagengast

16|15 Networks of value added trade
   João Amador | Sónia Cabral

17|15 House prices: bubbles, exuberance or something else? Evidence from euro area countries
   Rita Fradique Lourenço | Paulo M. M. Rodrigues
2016

1|16  A mixed frequency approach to forecast private consumption with ATM/POS data
Cláudia Duarte | Paulo M. M. Rodrigues | António Rua

2|16  Monetary developments and expansionary fiscal consolidations: evidence from the EMU
António Afonso | Luís Martins

3|16  Output and unemployment, Portugal, 2008–2012
José R. Maria

4|16  Productivity and organization in Portuguese firms
Lorenzo Caliendo | Luca David Opromolla | Giordano Mion | Esteban Rossi-Hansberg

5|16  Residual-augmented IVX predictive regression
Matei Demetrescu | Paulo M. M. Rodrigues

6|16  Understanding the public sector pay gap
Maria M. Campos | Evangelia Papapetrou | Domenico Depalo | Javier J. Pérez | Roberto Ramos

7|16  Sorry, we’re closed: loan conditions when bank branches close and firms transfer to another bank
Diana Bonfim | Gil Nogueira | Steven Ongena

8|16  The effect of quantitative easing on lending conditions
Laura Blattner | Luísa Farinha | Gil Nogueira

9|16  Market integration and the persistence of electricity prices
João Pedro Pereira | Vasco Pesquita | Paulo M. M. Rodrigues | António Rua

10|16  EAGLE-FLI | A macroeconomic model of banking and financial interdependence in the euro area
N. Bokan | A. Gerali | S. Gomes | P. Jacquinot | M. Pisani

11|16  Temporary contracts’ transitions: the role of training and institutions
Sara Serra

12|16  A wavelet-based multivariate multiscale approach for forecasting
António Rua
A mixed frequency approach to forecast private consumption with ATM/POS data

Working Papers 2016

Cláudia Duarte | Paulo M. M. Rodrigues | António Rua