Pooling-based data interpolation and backdating*

Massimiliano Marcellino
IEP - Bocconi University, IGIER and CEPR

Revised Version - May 2006

Abstract

Pooling forecasts obtained from different procedures typically reduces the mean square forecast error and more generally improves the quality of the forecast. In this paper we evaluate whether pooling interpolated or backdated time series obtained from different procedures can also improve the quality of the generated data. Both simulation results and empirical analyses with macroeconomic time series indicate that pooling plays a positive and important role also in this context.

Key words: Pooling, Interpolation, Factor Model, Kalman Filter, Spline

JEL Classification: C32, C43, C82.

*I am grateful to an anonymous Referee for helpful comments and to Andrea Ghiringhelli for excellent research assistance. This research is partly supported by the Italian MIUR.

1 Introduction

Bates and Granger (1969) showed that a combination of forecasts may perform better than each of the single constituent forecasts, and there now exists a vast amount of empirical evidence to support their claim, see e.g. Timmermann (2005) for a recent overview. As discussed by Hendry and Clements (2004), possible reasons for the good performance of forecast pooling may be model misspecification and parameter non-constancy that are attenuated by weighting.

Forecasting can be considered as a problem of estimation of missing observations at the end of a time series. But estimation of missing observations at the beginning of a time series (i.e. backdating) or of periodically missing observations (i.e. data interpolation, e.g. obtaining monthly values starting from quarterly data) is also quite relevant for empirical analysis and has attracted considerable attention in the literature, see e.g. Marcellino (1998) or Proietti (2004) for recent overviews.

In this paper we propose to apply pooling techniques developed in the forecasting literature to improve the quality of backdated and/or interpolated data. To the best of our knowledge, this is the first attempt to bring together pooling and interpolation.

The analysis is complicated by the fact that the forecasts can be compared with realized values after some time, while typically missing observations remain missing. This limits somewhat the range of feasible pooling techniques. Yet, simple combination methods such as averaging, possibly after trimming extreme values, work quite well in practice when compared with more sophisticated techniques, see e.g. Stock and Watson (1999).[1]

Our starting point is the paper by Angelini, Henry and Marcellino (2003). They introduce a new method for data interpolation or backdating, based on the idea of summarizing large information sets with a factor model, and then using the estimated factors to recover the missing observations. Using simulation experiments and empirical applications, they compare the factor

[1] In a few cases more sophisticated methods might be adopted. For example, in the case of missing observations at the beginning of the sample, if the available observable sample is long enough, it could be split into two subsamples, and one of them used to test alternative backdating procedures. The procedure would mimick what is known as predictive least squares in the forecasting literature. However, the latter is often dominated by simpler pooling techniques, see e.g. Stock and Watson (1999).

based method with the traditional Chow and Lin (1971) multivariate approach, and with univariate methods based either on parametric AR models combined with the Kalman filter, or on the use of (nonparametric) spline functions. They find that the multivariate methods typically outperform the univariate approaches, but the ranking of factor-based interpolation and Chow-Lin is not clear cut.

We consider the same set of interpolation/backdating methods as Angelini et al. (2003), but also a set of pooling procedures. In particular, we look at the mean, trimmed mean and median either of all methods or of only univariate or multivariate procedures. To benchmark the analysis, we use the same set up for the simulation experiments as Angelini et al. (2003). In particular, the series to be interpolated is generated either as a linear combination of factors, or as a linear combination of a limited set of regressors, or as an AR(1) process.

As expected, the best method is the one suited for the specific data generating mechanism, e.g. factor based interpolation when the series of interest is generated as a combination of factors. But the interesting and clear-cut result we find is that pooling the multivariate methods is a close second best for *any* generating mechanism. Moreover, using the pooled interpolated / backdated data usually reduces the bias in subsequent econometric analyses using the generated variables.

A set of empirical applications with data for four European countries, the euro area and the US also support the usefulness of the pooling procedures.

Therefore, we conclude that the good performance of pooling for forecasting is confirmed also for backdating and data interpolation, and this represents a new, interesting and important result for empirical analysis.

The structure of the paper is as follows. In section 2 we describe in more details the interpolation/backdating procedures under evaluation. In section 3 we discuss the design of the simulation experiments and the related results. In section 4 we present the empirical applications. Finally, in section 5 we summarize the main findings of the paper and conclude.

2 Interpolation and back-dating procedures

Let us consider a weakly stationary time series y_t^o whose realizations can be only partly observed. In particular, the observed values of y_t^o can be thought of as realizations of the process $y = \{y_\tau\}_{\tau=1}^{\infty} = \{\omega(L)y_{kt}^o\}_{t=1}^{\infty}$, where τ indicates the aggregate temporal frequency (e.g. quarters), k the frequency of aggregation (e.g. 3 if t is measured in months), L is the lag operator, and $\omega(L) = \omega_0 + \omega_1 L + ... + \omega_{k-1} L^{k-1}$ characterizes the aggregation scheme. For example, $\omega(L) = 1 + L + ... + L^{k-1}$ in the case of flow variables and $\omega(L) = 1$ for stock variables.

The $n \times 1$ vector of weakly stationary time series X_t can instead be observed in each time period t.

We stack the observations on X_t, y_t^o, and y_τ into $\underset{nT\times 1}{\mathbf{X}}$, $\underset{T\times 1}{\mathbf{Y}^o}$ and $\underset{s\times 1}{\mathbf{Y}}$, where T and s are, respectively, the number of disaggregate and aggregate observations. If we construct the aggregator matrix $\overline{\mathbf{W}}$ as

$$\underset{(s+nT)\times(n+1)T}{\overline{\mathbf{W}}} = \begin{pmatrix} \underset{s\times T}{\mathbf{W}} & \underset{s\times nT}{\mathbf{0}} \\ \underset{nT\times T}{\mathbf{0}} & \underset{nT\times nT}{\mathbf{I}} \end{pmatrix},$$

$$\underset{s\times T}{\mathbf{W}} = \begin{pmatrix} \omega_0, \omega_1, ..., \omega_{k-1} & 0, 0, ..., 0 & ... & 0, 0, ..., 0 \\ 0, 0, ..., 0 & \omega_0, \omega_1, ..., \omega_{k-1} & ... & 0, 0, ..., 0 \\ ... & & & \\ 0, 0, ..., 0 & 0, 0, ..., 0 & & \omega_0, \omega_1, ..., \omega_{k-1} \end{pmatrix},$$

then $\mathbf{Z} = \overline{\mathbf{W}}\mathbf{Z}^o$, where $\mathbf{Z}^o = (\mathbf{Y}^{o\prime} : \mathbf{X}')'$, $\mathbf{Z} = (\mathbf{Y}' : \mathbf{X}')'$, and : denotes stacking.

The structure of the matrix $\mathbf{W}$ can be easily modified to deal with missing observations at the beginning of the sample, or elsewhere. In particular, since the processes under analysis are weakly stationary, the results for backdating are also applicable to the case of forecasting, i.e. missing observations on the y variable at the end of the sample, assuming that the corresponding values of the X variables are known.

Several procedures are available to estimate the values of $\mathbf{Y}^o$ given those of $\mathbf{Z}$. In this section we describe first the single competing interpolation and backdating procedures that we consider in the simulation experiments, and then the methods to combine the results of the single procedures into a pooled estimate of the missing observations.

To construct the parametric univariate estimator (that relies only on information on $\mathbf{Y}$), we assume that y_t^o is generated by an AR(3) model. We compute the optimal estimator of the systematically missing observations using the Kalman filter, and the smoother, according to the formulae in

Harvey and Pierse (1984), see also Kohn and Ansley (1986), Nijman and Palm (1986), and Gomez and Maravall (1994). For backdating only the smoother is used.

The nonparametric univariate estimator is based on spline functions, see e.g. Micula and Micula (1998). The tension factor, which indicates the curviness of the resulting function is set equal to one. Values close to zero would imply that the curve is approximately the tensor product of cubic splines, while if the tension factor is large the resulting curve is approximately bi-linear. A major advantage of this estimator is that it does not require assumptions on the disaggregate generating mechanism, which is typically unknown and cannot be uniquely recovered from the generating mechanism at the aggregate level because of aliasing problems, see e.g. Marcellino (1998, 1999) for additional details. On the other hand, this method cannot be used for backdating.

The first multivariate estimator (that relies on information both on $\mathbf{Y}$ and on $\mathbf{X}$) implements the Chow and Lin (1971) procedure, allowing for an AR(1) structure in the errors of the regression. Basically, the y variable is regressed on (some of) the variables in X at the aggregate level, and the estimated coefficients are combined with the observable values of the regressors at the disaggregate level to obtain an estimate for the missing values in y^0.

Five variables out of the n in $\mathbf{X}$ are included as regressors, and they are selected among the set of available variables on the basis of their correlation at the aggregate level with the variable to be disaggregated. This reflects common empirical practices, where, typically, variables involved in Chow-Lin procedures are pre-selected on the basis of their economic relevance with respect to the estimated variable.

The second multivariate estimator exploits the factors extracted from a dynamic factor model estimated with a large data set. Specifically, we assume that X_t admits the factor representation

$$\underset{n\times 1}{X_t} = \underset{n\times p}{\Lambda}\underset{p\times 1}{F_t} + \underset{n\times 1}{e_t}, \tag{1}$$

where p, the number of factors, is substantially smaller than n, namely, a few common forces drive the joint evolution of all variables. Precise conditions on the factors, F_t, and the idiosyncratic errors, e_t, can be found in Stock and Watson (2002a, 2002b).

In the first step of this factor based interpolation procedure, the factors are estimated as the first principal components of the X variables, i.e., without using information on the y variable to be interpolated (see Stock and Watson (2002a, 2002b) for a theoretical justification of the principal component based estimator for the factors and a detailed discussion of its properties). The Chow and Lin (1971) method is then applied to recover the missing values of y_t^o using (three) estimated factors as regressors rather than some selected variables in X.

In the second step of the interpolation procedure, the interpolated y^0 variable obtained in the first step is added to the balanced panel of the X variables, the factors are re-estimated, the Chow and Lin (1971) procedure is applied with the new set of factors, a new set of interpolated values for y^0 are obtained, and they are used to construct another balanced panel, another set of factors, etc. The procedure is repeated until the estimates of the factors do not change substantially in successive iterations. If the fit of the Chow and Lin (1971) regression in the second step is lower than that in the first step, the procedure is stopped and the balanced factor based estimator is used.

This iterative interpolation procedure, introduced by Angelini et al. (2003), refines the EM algorithm that Stock and Watson (2002a, 2002b) proposed to extract factors from unbalanced panels. Moreover, following the same line of reasoning as in Stock and Watson (2002a, 2002b) in a forecasting context, the fact that the estimated rather than the true factors are used in the interpolation/backdating procedure does not affect the quality of the fit of the Chow-Lin regression, at least asymptotically, see also Bai (2003) and Bai and Ng (2004).

Both multivariate procedures can be applied to interpolation and backdating. Therefore, so far we have five methods for interpolation: Kalman filter (K-filter), Kalman smoother (K-smoother), spline, Chow-Lin, factor-based (DFM); and three methods for backdating: Kalman smoother, Chow-Lin, and factor-based.

For interpolation we then consider five pooling procedures: the average of the five single methods (P-Mean), of the three univariate methods (P-Uni), of the two multivariate methods (P-Multi), the median of the five single methods (P-Median), and a trimmed mean where the highest and lowest values are discarded (P-TrimMean). For backdating we evaluate four pooling methods: the average of the three single methods (P-Mean), of the two multivariate methods (P-Multi), the median of the three single methods (P-Median), and a trimmed mean where the highest of the three values is discarded (P-TrimMean).

Finally, it is worth mentioning that changes in the specification of the single estimators under analysis (e.g. increasing the order of the AR model, or using more factors or more variables, or changing the trimming procedure) in general do not affect the results of the simulation experiments.

3 Simulation experiments

In this section we evaluate the relative performance of the alternative interpolation / backdating methods by means of simulation experiments. In the first subsection we describe the design of the experiments. In the second subsection we discuss the results. In the third subsection we comment upon the outcome of a set of sensitivity analyses. In the final subsection we run additional simulations to evaluate the biases that can emerge when using the interpolated / backdated data in subsequent econometric analysis.

3.1 The experimental design

Following Angelini et al. (2003), we consider three different generating mechanisms (DGM) for the variables:

$$\begin{aligned} X_t &= \Lambda F_t + e_t, \\ y_t^o &= \beta' F_t + \varepsilon_t, \end{aligned} \tag{2}$$

$$\begin{aligned} X_t &= \Lambda F_t + e_t, \\ y_t^o &= \beta' Z_t + \varepsilon_t, \end{aligned} \tag{3}$$

and

$$\begin{aligned} X_t &= Q X_{t-1} + u_t, \\ y_t^o &= \gamma y_{t-1}^o + v_t, \end{aligned} \qquad \begin{pmatrix} u_t \\ v_t \end{pmatrix} = P \begin{pmatrix} e_t \\ \varepsilon_t \end{pmatrix}. \tag{4}$$

In the first specification (2) both the X and the y variables are generated by a factor model. The number of factors is set equal to 3, the factors are independent AR(1) processes with root equal to 0.8, and the elements of Λ

and β are independent draws from a uniform distribution over the interval $[0, 1]$. In the second specification (3) $Z_t = (x_{1t}, x_{2t}, x_{3t})'$, so that y depends on some of the variables in X rather than on the factors. In the third specification (4) y and each of the variables in X are AR(1) processes, each with root equal to 0.8 (Q is a diagonal matrix).

In the first two DGMs e_t and ε_t are i.i.d. $N(0, 1)$ errors, uncorrelated across themselves, while in the third DGM the errors are correlated across variables (the elements of the matrix P are independent draws from a uniform distribution), but not over time. In all cases X_t contains 50 variables while y_t^o is univariate, and the sample size is set equal to 100.

When the DGM is (2, DFM) we expect the factor estimator to be the best, but the Chow and Lin (1971) method should also perform well since the number of regressors (five) is larger than the number of factors, so that the former can provide a good approximation for the latter. When the DGM is (3, Chow-Lin) the Chow-Lin method is expected to generate the lowest loss function, but the factor based interpolation approach could also perform well when the factors have a high explanatory power for the Z variables, since the model for y_t^o in (3) can be written as:

$$y_t^o = \beta' S \Lambda F_t + \beta' S e_t, + \varepsilon_t,$$

where S is the selection matrix

$$S = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \end{pmatrix}.$$

When data are generated according to (4, AR-SUR) the univariate estimators should be ranked first, but the multivariate methods could also perform well due to the correlation across variables generated by the variance covariance matrix of the error terms. Since in practice the DGM is not known, we are particularly interested in whether any pooling methods provides a robust interpolation procedure, in the sense of working reasonably well for *all* DGMs. Hence, though more complicated DGMs could be used, those in (2)-(4) already provide a good framework to evaluate the relative merits of the alternative interpolation methods.

We set the disaggregation frequency at 4, so that only 25 values of y_t^o can be observed out of 100. This mimics disaggregation of annual data into quarterly data. We analyze both stock and flow variables. Next we also consider the case of missing observations at the beginning of the series, assuming that 20 starting values of y_t^o are unobservable. For each case we run 1000 replications, and rank the estimators on the basis of the average absolute and mean square interpolation/backdating error (MAE and MSE, respectively). We also compute percentiles of the distribution of the absolute and mean square disaggregation error, which provides information on the robustness of the performance of the estimators.

3.2 Results

The Monte Carlo results are summarized in Table 1 for the three DGMs: DFM, Chow-Lin and AR-SUR. We only report the average MSE over replications, since the MAE leads to similar rankings of the various interpolation methods, and the mean and the median of the distribution of the disaggregation errors over replications are usually very close. More generally, the ranking of the methods is rather uniform when evaluated at different quantiles, these results are available upon request.

When the data are generated by the DFM in (2), the figures in panel A of Table 1 clearly show that the factor method performs best, as expected and in line with the findings in Angelini et al. (2003). Yet, the (very close) second best is P-Multi, the simple average of the factor and Chow-Lin interpolated values. The three univariate methods perform rather poorly and also the other four pooling procedures, where at least one univariate method appears in the combined interpolated series, are beaten by Chow-Lin. However, it is worth noting that P-Uni works better than each of its three constituent single univariate components for both stock and flow variables.

When the data on the y variable are generated by the Chow-Lin specification in (3), the P-Multi yields the lowest MSE for both stock and flow variables, see panel B of Table 1, with the Chow-Lin interpolation method as the second best. The ranking is reverted for backdating, but the differences are minor. The univariate methods are again substantially worse, but also in this case pooling them into P-Uni systematically lowers the MSE.

When the data are generated by the AR-SUR DGM in (4), in turn, the relative behavior of the univariate methods improves substantially, see panel C of Table 1. In the case of stock variables, the average of the five single methods, P-Mean, yields the lowest MSE, and the trimmed average, P-TrimMean is the second best. P-Multi is in this context the worst pooling procedure, but it is still better than any of the single interpolation methods. For flow variables the performance of the multivariate methods deteriorates substantially, likely because the averaging underlying this aggregation scheme decreases the correlation across y and the X variables. As a consequence, the performance of P-Multi also deteriorates, but P-TrimMean and P-Mean remain the best interpolation methods. The ranking of P-Multi improves again for the case of backdating, where it is the first best followed by Chow-Lin. However, in this case all methods give satisfactory results, with minor differences in the resulting MSEs.

In summary, pooling appears to work quite well also for interpolation and backdating. Typically, the pooled interpolated series perform better than each of their single components. In particular, P-Multi is ranked first or second best in most experiments, which suggests it as a candidate for a robust interpolation procedure. We now evaluate whether these results are robust to several changes in the experimental design.

3.3 Sensitivity analysis

The first two additional experiments we consider introduce misspecification into the models that underlie the single interpolation methods. In this context, the relative performance of the pooling procedures can further improve, since the method most suited for each DGM now relies on a misspecified model. First, we consider a misspecification of the number of factors. In particular, there are ten factors in the DGM in (2) but only five of them are used in the factor based interpolation procedure. The figures in Table 2, which again refer to the average MSE, indicate that P-Multi becomes the first-best for all types of interpolation and backdating.

In the second additional experiment we use the Chow-Lin DGM in (3) but now y depends on all the variables in X, i.e. $Z_t = X_t$ in (3). This is the case for which the factor based method should be the most appropriate, since the y variable depends on 50 regressors and there are not enough observations to run a Chow-Lin procedure with all the potentially relevant variables included. Therefore, we can either summarize the 50 variables by means of (three) estimated factors or select those (five variables) that present the highest correlation with y. The values in Table 2 indicate that now for all types of interpolation or backdating the losses from the Chow-Lin method are ten times larger than those from the factor approach, which becomes the first best. Yet, P-Multi is the second best, and it is by far better than Chow-Lin (or a univariate method).

Next, we increase the variability of the idiosyncratic error of the factor model in (2) and (3), specifically the idiosyncratic error is $2e_t$ rather than e_t. This complicates the estimation of the factors, so that the performance of factor based interpolation could be expected to deteriorate. On the other hand, since the variability of the error term in the y equation, ε_t in (2) and (3), remains the same, while the variance of each of the variables in X increases, the Chow-Lin method could perform even better than in the standard case with the DGM in (3). The consequences on the pooling procedures can be hardly determined a priori in this case. The results are reported in Tables 2 for the DGM (2) and (3). When the DGM is (2), the factor based method deteriorates by about 20% with respect to the values reported in Table 1, but it remains the first best, closely followed by P-Multi. The losses of the factor based method increase to 30-40% when the DGM is (3) with respect to Table 1, while the Chow-Lin method improves substantially, as expected, but the differences with P-Multi remain minor.

In Table 3 we evaluate the role of the number of variables and observations, In particular, we compare the benchmark case of $N = 50, T = 100$ with $N = 100, T = 100$, $N = 50, T = 200$, and $N = 100, T = 200$. We have chosen to focus for simplicity on the case where there are ten factors in the DGM in (2) but only five of them are used in the factor based interpolation procedure. The figures indicates that there are only minor changes in the results, with P-Multi providing the lowest losses in most of the cases. This indicates that the factors and the loadings are already quite accurately estimated when $N = 50, T = 100$.

In Table 3 we also consider what happens if a fewer number of observations are missing (specifically, the disaggregation frequency lowers from 4 to 2), and when the persistence in the factors decreases (the root of the AR processes lowers from 0.8 to 0.4). The results are interesting and in line with intuition. A lower number of missing observations decreases the loss, since the sample becomes more informative, while lower persistence of the factors increases the loss, since the quality of the estimated factors decreases.

Finally, we consider backdating 20% of missing observations when the variables are generated using the three DGMs in (2)-(4) but both y and the variables in X are subject to larger errors at the beginning of the sample. Specifically, an error term $5u$ is added to the first 20% of observations on y and each of the X variables, where u is i.i.d. standard normal. This in an interesting case since larger measurement errors at the beginning of the sample can be expected for some macroeconomic time series. For these DGMs we would expect an improved performance of the univariate methods, since the use of regressors is complicated by the measurement error, with uncertain results for the pooling procedures. Table 4 confirms the better behavior of univariate methods and, as a consequence, the range of the losses across interpolation methods shrinks substantially with respect to the standard case in Table 1. However, even in this case, P-Multi is the first best for two DGMs, and the second best for the third one.

In summary, these additional experiments provide additional evidence in favour of the pooling techniques, and in particular of P-Multi.

3.4 Using the interpolated values

As noted by Angelini et al. (2003), it may be worth assessing the extent to which using the interpolated series instead of the actual ones would impact on possible subsequent econometric exercises. For example, we can expect the dynamic properties of the interpolated series and its relationships with other variables to be somewhat affected, with the extent of the bias depending on the goodness of the disaggregation method but also on the specific econometric characteristic under analysis. In particular, in this subsection we study the autocorrelation properties of the interpolated data as well as regression results.

We generate the data according to the three DGMs in (2)-(4). Then we compute the difference (ρ) between the first order autocorrelation coefficients for the actual and interpolated series, and the absolute value of the difference (β) of the estimated coefficient of y_t in the regression $w_t = y_t + u_t$, with u_t i.i.d. N(0,1), using actual and interpolated data for both w_t and y_t.

The results are reported in three panels of Table 5 for the three types of DGMs, again for stock and flow variables, and for 20% of missing observations at the beginning of the sample. As in the previous experiments, we focus on the average of the empirical distribution of ρ and β over 1000 replications for the competing interpolation methods, but details on the quantiles of the distribution are available upon request.

Four main comments can be made. First, as in the case of Angelini et al. (2003), the ranking of the disaggregation methods in terms of bias reflects that of Table 1, which suggests that minimizing the mean square disaggregation error is a good criterion to minimize also the bias in subsequent econometric analyses with the interpolated series. Second, the size of ρ and β is much smaller in the case of missing observations at the beginning of the sample than for interpolation of stock and flow variables, which is again in line with the results in Table 1 and is mainly due to the lower fraction of missing data, i.e. 20% versus 75% in the case of stock and flow variables. Third, when the DGM is multivariate, univariate interpolation procedures lead to major biases in the estimated dynamics, i.e. large values of ρ, and viceversa when the DGM is univariate. Fourth, even when ρ is large, the corresponding value of β is small, indicating that the estimation of dynamic relationships can be more affected by interpolation than contemporaneous relationships, which is also a sensible result. Finally, and most important for us, the pooling procedures yield gains also in this context, and P-Multi performs particularly well, except in the case of stock variables with DGM (4), due to the bad results for both the factor based method and Chow-Lin in this case.

4 Empirical applications

In this section we compare the relative merits of the interpolation methods in practice using three datasets for, respectively, four European countries, the euro area, and the US. In the first subsection we describe the data, in the second one the results.

4.1 The data

To start with, we consider quarterly series for GDP growth and inflation (measured as the quarter on quarter change in the private consumption deflator) for France, Germany, Italy and Spain, i.e. the four largest countries in the euro area, over the period 1977:3-1999:2. For these countries, we also collect two datasets with either price or real variables. The price dataset contains variables such as CPI, GDP deflator, export and import deflators, etc., overall 29 series (in growth rates). The variables in the real dataset include, among others, GDP components, capacity utilization, industrial production, employment and the unemployment rate, etc., a total of 39 series (transformed to achieve stationarity when necessary). The two datasets are extracted from the one used in Angelini et al. (2003), and a Data Appendix available upon request contains a detailed list of all the series employed in the current analysis.

Since the four European countries are still rather heterogenous, to evaluate the relative merits of the interpolation procedures with a more homogeneous dataset we then consider 64 quarterly variables (plus inflation and GDP growth) for the euro area, over the period 1980:1-2003:4. The series include both nominal and real variables and are extracted from an updated version of the dataset in Fagan, Henry and Mestre (2001), details are provided in the Data Appendix.

Finally, we consider a monthly dataset for the US that includes CPI inflation, Industrial Production and other 123 varied macroeconomic variables, over the period 1961:1:2003:12, for a total of 516 observations. The dataset is an updated version of the balanced panel in Stock and Watson (2002a, 2002b), and it let us evaluate the interpolation methods for a different frequency and a substantially longer time period. Again, details on the variables are provided in the Data Appendix.

4.2 Empirical results

In the first application, we carry out two kinds of interpolation exercises that mimic what we did in the simulation experiments. First, we drop all the observations but those corresponding to the last quarter of each year (or to the last month of each quarter for the US). Second, we drop the initial 20% of the observations. In both cases, we interpolate the missing observations so as to recreate them, and then compare the interpolated with the actual values. The price deflator is treated as a stock variable and GDP (or IP) growth as a flow.

As in the simulation experiments, we extract three factors from each dataset to implement the factor based interpolation procedure. Previous work by Stock and Watson (1998) for the US, and Marcellino, Stock and Watson (2003) and Angelini et al. (2001) for European countries and the euro area have shown that a limited number of factors are sufficient to explain a substantial proportion of the variability of all the series. We use the same setup as in the simulations also for the Chow-Lin method, namely, five regressors are selected from the datasets used for factor extraction, following the procedure outlined in the previous section. For the univariate methods, we consider the filtered and smoothed estimate from the Kalman filter assuming an AR(3) model for both inflation and GDP (or IP) growth, and the spline. Finally, for the pooling procedures, we evaluate the five alternative combination methods for interpolation describe in the previous section, and the four ones for backdating.

The comparison of the interpolation/backdating methods is based on the mean square and the results are summarized in Table 6.[2]

For the interpolation of missing infra-year data, in the case of the inflation rates, the P-Multi method delivers the best results for each of the four European countries and for the euro area. It is also better than most single methods for the US. For the latter country, the lowest MSE and MAE are achieved by P-Mean, the average of all single methods, since for the US the univariate interpolation procedures work reasonably well while they lead to major losses for the other countries. For GDP growth, the spline method is the best choice for the four European countries and Chow-Lin for the US (IP), but a pooled method is the second best in all the five cases. P-Multi is the first best for the euro area.

When estimating missing observations concentrated at the beginning of the sample, the differences across the methods shrink, and the best procedures appear to be P-Multi and Chow-Lin (first or second best in, respectively, 8 cases out of 12 and 7 cases out of 12, based on the MSE).

Overall, these results are in line with the outcome of the simulation experiments. In particular, the better performance of the univariate methods for GDP growth can be related to the fact that measurement error is a more serious issue for GDP than for inflation.

In the second application, again as in the simulation experiments, we compute the difference (ρ) between the first order autocorrelation coefficients

[2] The ranking based on the mean absolute disaggregation errors is very similar, detailed results are available upon request.

for the actual and interpolated series for each variable and country. For France, Italy and Spain we then compute the difference of the estimated coefficients in a regression of inflation or GDP growth for each country on the same variable for Germany (β), using actual and interpolated series. For the euro area and the US, β is computed as the difference of the estimated coefficient in a regression of inflation on GDP growth, and viceversa. The results are presented in Table 7.

The biases in dynamics (ρ) are smaller for inflation than for GDP growth, due to the higher persistence of the former. For interpolation of inflation, the differences across methods in the values of ρ are in general minor, and the lowest values are achieved either by a multivariate procedure or by P-Multi. For GDP growth, the differences widen, the performance of Chow-Lin further improves, and the second best is either DFM or P-Multi. In the case of backdating, the ranking of the methods shrinks for both inflation and GDP growth, and the pooling procedures confirm their good performance, in particular, P-Mean and P-Median for inflation and P-TrimMean for GDP growth.

The biases in regression coefficients (β) are harder to be interpreted. Actually, an interpolation method can introduce large errors but if these errors are similar for the two series that are interpolated and then used in the regression, the bias in β can turn out to be small. This could explain the good performance of the spline for inflation, it generates the lowest biases in four cases out of five, while spline based interpolation did not perform well in Table 6. Using interpolated GDP growth rates, Chow-Lin introduces the smallest distortions in all cases, followed by P-Multi. In the case of backdating, the results are more varied, but P-Multi and P-TrimMean work quite well for, respectively, inflation and GDP growth.

Finally, to evaluate the robustness of the results we have (a) increased the number of factors to five, as the number of regressors in the Chow-Lin method; (b) decreased the number of regressors in the Chow-Lin method to three, as the number of factors in the base case; (c) used the consumer price index instead of the consumption deflator for the European countries. Although there were some changes in the resulting figures, the ranking of the interpolation methods was virtually unaltered in all cases.

5 Conclusions

In this paper we propose a pooling approach to interpolation and estimation of missing observations and compare it with a number of more standard single techniques, using both artificially generated and actual time series. Since in general realizations of the missing observations do not become available, contrary to the forecasting case where the target values can be observed with some delay, only pooling based on simple averaging methods is considered. However, the forecasting literature suggests that even these simple pooling methods can be effective.

In the Monte Carlo experiments, deleted data from artificial series are re-estimated using the whole range of considered methods (Kalman filter and smoother, Spline, Chow-Lin, factor models, and different averaging schemes). Using a sample of 25 years of quarterly data for 50 series, we examine the cases where stock and flow variables are only available at the annual frequency, and that where there are 20% of missing starting data. The artificial series are generated using AR(1) models for each variable with correlated errors, or factor models where the variable of interest can depend either on the factors or on a small subset of other variables. As a sensitivity analysis, the factor based interpolation can be based on a number of factors largely inferior to that of the DGM, or there can be larger idiosyncratic components, or the variables can be subject to measurement error. The relative ranking of the procedures is evaluated in two ways. First, by means of the Mean Absolute (interpolation / backdating) Error, Mean Squared Error and the quantiles of the absolute or squared difference between the interpolated series and the original "true" one. Second, on the basis of the biases that emerge in subsequent econometric analysis using the interpolated or backdated values.

The main conclusion of the simulation experiments is that each single interpolation/backdating method is optimal when the data are generated according to the hypotheses underlying that method (e.g. factor based interpolation for the factor DGM). Yet, the pooling procedures, and in particular P-Multi that combines the Chow-Lin and the factor based approaches, provide a close second best across *all* experiments, and this is an important result since in practice the DGM is not known and the first best cannot be achieved.

The stable, robust and on average satisfactory performance of the pooling methods emerges also from the empirical analysis, where we consider the same issues as in the simulation experiments but using actual macroeconomic time series for the largest European countries in the euro area, for the euro area as a whole, and for the US.

An interesting issue for further research in this area is the evaluation of pooling using more sophisticated single interpolation/backdating procedures, such as the dynamic extension of Chow-Lin in Santos Silva and Cardoso (2001) or the structural model based interpolation in De Jong (2005), or possibly of more complicated pooling schemes for those cases where the missing observations become available with some delay and can be compared with the interpolated values. Yet, while there can be further improvements in the quality of the interpolated data, we believe that the main message of this paper, namely that pooling different procedures is beneficial, will remain valid or even further strengthened.

References

[1] Angelini, E., Henry, J. and M. Marcellino (2003), "Interpolation with a large information set", *Journal of Economic Dynamics and Control*, forthcoming.

[2] Angelini, E., Henry, J. and R. Mestre (2001), "Diffusion index based inflation forecasts for the Euro area", European Central Bank WP 61.

[3] Bai, J. (2003), "Inferential theory for factor models of large dimension", *Econometrica*, 71, 135-171.

[4] Bai, J. and S. Ng (2004), "Confidence intervals for diffusion index forecasts with a large number of predictors", mimeo.

[5] Bates, J.M. and C.W.J. Granger (1969), "The Combination of Forecasts", *Operations Research Quarterly*, 20, 415-468.

[6] Chow, G.C. and Lin, A. (1971), "Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time-Series by Related Series." *Review of Economics and Statistics*, 53, 372-375.

[7] De Jong. P. (2005), "Smoothing and interpolation with the state space model", in Harvey, A. and Proietti, T. (eds.), *Readings in Unobserved Component Models*, Oxford University Press, forthcoming.

[8] Fagan, G., Henry, J. and Mestre, R. (2001), "An area-wide model for the Euro area", ECB Working Paper No. 42, European Central Bank.

[9] Gomez, V. and A. Maravall (1994), "Estimation, Prediction and Interpolation for Non Stationary Time Series with the Kalman Filter." *Journal of the American Statistical Association*, 89, 611-624.

[10] Harvey, A.C. and R.G. Pierse (1984), "Estimating Missing Observations in Economic Time Series." *Journal of the American Statistical Association*, 79, 125-131.

[11] Hendry, D.F., and M.P. Clements (2004), "Pooling of Forecasts", *Econometrics Journal*, 7, 1-31.

[12] Kohn, R. and C.F. Ansley (1986), "Estimation, Prediction and Interpolation for ARIMA models with Missing Data." *Journal of the American Statistical Association*, 81, 751-761.

[13] Marcellino, M. (1998), "Temporal disaggregation, missing observations, outliers, and forecasting: a unifying non-model based procedure", *Advances in Econometrics*, 13, 181-202.

[14] Marcellino, M. (1999), "Some consequences of temporal aggregation for empirical analysis", *Journal of Business and Economic Statistics*, 17, 129-136.

[15] Marcellino, M., Stock, J.H. and M. W. Watson (2003), "Macroeconomic forecasting in the Euro area: country specific versus euro wide information", *European Economic Review*, 47, 1-18.

[16] Micula, G. and S. Micula (1998), *Handbook of Splines*, Kluwer Academic Publishers.

[17] Nijman, T.E. and F. C. Palm (1986), "The Construction and Use of Approximations for Missing Quarterly Observations: A Model-Based Approach." *Journal of Business and Economic Statistics*, 4, 47-58.

[18] Proietti, T. (2004), "Temporal disaggregation by state-space methods: Dynamic regression models revisited", mimeo.

[19] Santos Silva, J.M. and Cardoso, F.N. (2001), "The Chow-Lin method using dynamic models", *Economic Modeling*, 18, 269-280.

[20] Stock, J.H. and M.W. Watson. (1999) "A comparison of linear and non-linear univariate models for forecasting macroeconomic time series", in Engle, R. and White, H.. (eds), *Cointegration, causality, and forecasting: A festschrift in honor of Clive W.J. Granger*, Oxford: Oxford University Press, 1-44.

[21] Stock, J.H. and M.W. Watson (2002a), "Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economic Statistics*, 20, 147-162.

[22] 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.

[23] Timmermann, A.G. (2005) "Forecast Combination", in G. Elliott, C.W.J. Granger and A. Timmermann (eds.) *Handbook of Economic Forecasting*, forthcoming.

Table 1. Disaggregation error

	A			B			C		
	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%
DFM:	**0,305**	**0,294**	**0,073**	0,284	0,273	0,07	0,621	0,991	0,144
Chow-Lin:	0,384	0,368	0,083	*0,253*	*0,252*	**0,052**	0,557	1,216	*0,137*
Spline:	1,172	0,627	.	1,172	0,622	.	0,551	0,237	.
K-filter:	0,869	0,649	.	0,872	0,644	.	0,717	0,375	.
K-smoother:	0,874	0,648	0,239	0,888	0,642	0,237	0,673	0,359	0,226
P-Mean:	0,481	0,405	0,093	0,45	0,37	0,077	**0,381**	*0,234*	0,142
P-Median:	0,603	0,536	0,077	0,598	0,524	0,062	0,455	0,24	0,139
P-TrimMean:	0,49	0,436	0,099	0,465	0,409	0,085	*0,427*	**0,218**	0,149
P-Multi:	*0,317*	*0,307*	*0,075*	**0,239**	**0,236**	*0,055*	0,518	0,868	**0,133**
P-Uni:	0,824	0,618	.	0,835	0,614	.	0,512	0,283	.

Note: The table reports the average MSE, computed over 1000 replications.
A The DGM is DFM as in (2)
B The DGM is Chow Lin as in (3)
C The DGP is AR as in (4)
Bold indicates the lowest loss, italics the second lowest.

Table 2. Disaggregation error, other experiments

	A			B		
	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%
DFM:	*0,214*	0,234	*0,062*	**0,004**	**0,004**	**0,001**
Chow-Lin:	0,225	*0,23*	0,067	0,056	0,053	0,011
Spline:	1,417	0,762	.	1,028	0,539	.
K-filter:	0,951	0,756	.	0,815	0,579	.
K-smoother:	0,967	0,757	0,238	0,841	0,574	0,227
P-Mean:	0,46	0,382	0,081	0,294	0,212	0,03
P-Median:	0,687	0,628	0,064	0,49	0,429	0,007
P-TrimMean:	0,476	0,437	0,089	0,31	0,255	0,036
P-Multi:	**0,193**	**0,203**	**0,061**	*0,017*	*0,016*	*0,003*
P-Uni:	0,967	0,737	.	0,765	0,543	.

	C			D		
	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%
DFM:	**0,339**	**0,339**	**0,084**	0,421	0,403	0,104
Chow-Lin:	0,467	0,463	0,106	**0,214**	**0,222**	**0,037**
Spline:	1,161	0,621	.	1,265	0,673	.
K-filter:	0,869	0,643	.	0,906	0,686	.
K-smoother:	0,886	0,641	0,24	0,929	0,686	0,235
P-Mean:	0,513	0,426	0,106	0,492	0,414	0,081
P-Median:	0,632	0,54	0,095	0,632	0,567	0,08
P-TrimMean:	0,531	0,461	0,114	0,524	0,479	0,094
P-Multi:	*0,358*	*0,356*	*0,088*	*0,25*	*0,251*	*0,053*
P-Uni:	0,832	0,612	.	0,877	0,658	.

Note: The table reports the average MSE, computed over 1000 replications.
A The DGM is DFM Mis-specified as in (2) but with 10 factors in the DGM and 5 used in the factor based procedures
B The DGM is Chow Lin as in (3), but Zt=Xt.
C The DGP is as in (2), but the error term in the factor model is 2*et
D The DGP is as in (3), but the error term in the factor model is 2*et
Bold indicates the lowest loss, italics the second lowest.

Table 3: DGP is DFM mis-specified

	A			B			C		
	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%
DFM:	*0,214*	0,234	*0,062*	0,293	0,305	*0,053*	*0,215*	*0,219*	**0,025**
Chow-Lin:	0,225	*0,23*	0,067	*0,277*	*0,276*	0,06	0,248	0,248	0,03
Spline:	1,417	0,762	.	1,299	0,684	.	1,273	0,676	.
K-filter:	0,951	0,756	.	0,944	0,69	.	0,852	0,675	.
K-smoother:	0,967	0,757	0,238	1,144	0,689	0,248	0,853	0,675	0,128
P-Mean:	0,46	0,382	0,081	0,474	0,376	0,073	0,439	0,366	0,037
P-Median:	0,687	0,628	0,064	0,691	0,584	0,056	0,644	0,573	0,027
P-TrimMean:	0,476	0,437	0,089	0,498	0,432	0,08	0,456	0,413	0,04
P-Multi:	**0,193**	**0,203**	**0,061**	**0,241**	**0,244**	**0,052**	**0,207**	**0,21**	*0,026*
P-Uni:	0,967	0,737	.	0,956	0,67	.	0,877	0,659	.

	D			E			F		
	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%
DFM:	*0,208*	*0,214*	**0,023**	*0,15*	*0,156*	*0,055*	0,515	0,513	*0,092*
Chow-Lin:	0,242	0,245	0,028	0,17	0,17	0,063	*0,445*	*0,431*	0,099
Spline:	1,271	0,677	.	1,009	0,54	.	1,4	0,75	.
K-filter:	0,859	0,702	.	0,607	0,525	.	0,941	0,831	.
K-smoother:	0,86	0,68	0,103	0,607	0,526	0,258	0,988	0,794	0,22
P-Mean:	0,44	0,366	0,033	0,316	0,267	0,077	0,558	0,489	0,103
P-Median:	0,646	0,572	0,026	0,454	0,389	0,058	0,705	0,647	0,094
P-TrimMean:	0,457	0,412	0,036	0,322	0,292	0,084	0,577	0,535	0,11
P-Multi:	**0,201**	**0,207**	*0,024*	**0,144**	**0,147**	**0,054**	**0,414**	**0,407**	**0,09**
P-Uni:	0,881	0,664	.	0,635	0,496	.	0,937	0,754	.

Note: The table reports the average MSE, computed over 1000 replications.
Bold indicates the lowest loss, italics the second lowest.
The DGM is DFM Mis-specified as in (2) but with 10 factors in the DGM and 5 used in the factor based procedures
A N=50, T=100 B N=100, T=100 C N=50, T=200
D N=100, T=200 E Aggregation Frequency=2 instead of 4 F Persistent parameter of factors is 0.4 instead of 0.8

Table 4. Backdating with measurement error, 20% missing obs

	DGM AR				DGM DFM				DGM CHOW LIN			
	Average	0,25	0,5	0,75	Average	0,25	0,5	0,75	Average	0,25	0,5	0,75
DFM:	37,104	27,316	34,71	44,363	2,105	1,359	1,901	2,688	*4,809*	2,189	3,908	*6,372*
Chow-Lin	**36,912**	*27,283*	**34,486**	*44,333*	2,033	1,304	*1,77*	2,613	4,882	2,3	3,991	6,441
K-smooth	37,738	28,018	35,346	45,257	2,17	1,363	1,93	2,794	5,14	2,27	4,177	6,867
P-Mean:	37,208	27,463	34,787	44,629	*2,02*	**1,286**	1,779	2,613	4,816	*2,14*	*3,889*	6,4
P-Median	37,15	27,47	34,736	44,481	2,026	1,296	1,772	*2,608*	4,828	2,146	3,904	6,42
P-TrimMe	37,198	27,44	34,617	44,444	2,047	1,323	1,817	2,617	4,858	2,191	3,95	6,517
P-Multi:	*36,996*	**27,281**	*34,579*	**44,28**	**2,003**	*1,29*	**1,759**	**2,556**	**4,747**	**2,129**	**3,821**	**6,306**

Note: The table reports the mean and percentiles of the empirical distribution of the MSE, computed over 1000 replications,
when the the first 20% of the observations on y and each of the X are equal to their values plus $5*u_t$ and u_t is iid N(0,1)
Bold indicates the lowest loss, italics the second lowest.

Table 5. Properties of interpolated data, DGM is DFM

	A			B			C		
	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%	STOCK	FLOW	MISS OBS 20%
DFM:	**0,095**	*0,146*	**0,031**	0,103	0,154	0,031	0,6	0,195	0,025
Chow-Lin:	0,107	**0,143**	0,034	**0,088**	**0,124**	**0,028**	0,452	0,373	0,03
Spline:	0,671	0,776		0,677	0,778		**0,09**	0,162	
K-filter:	0,344	0,574		0,346	0,582		0,301	0,095	
K-smoother:	0,383	0,594	0,065	0,391	0,603	0,065	0,25	0,097	0,027
P-Mean:	0,357	0,542	0,041	0,366	0,549	0,039	0,188	**0,069**	**0,025**
P-Median:	0,406	0,679	0,033	0,422	0,687	0,03	0,229	0,085	**0,025**
P-TrimMean:	0,317	0,579	0,042	0,33	0,588	0,041	0,234	*0,084*	**0,025**
P-Multi:	*0,098*	0,15	**0,031**	*0,093*	*0,145*	**0,028**	0,483	0,19	**0,025**
P-Uni:	0,506	0,682		0,52	0,687		*0,124*	0,096	

Note: The table reports the difference (ρ) between the first order autocorrelation coefficient for the actual and interpolated series in the left-hand side, and the absolute value of the difference (β) of the estimated coefficient of y_t in the regression $w_t = y_t + u_t$, with u_t iid N(0,1), using actual and interpolated data for both y and w, in the right-hand side.

A The DGM is DFM as in (2) B The DGM is Chow Lin as in (3) C The DGP is AR as in (4)

Bold indicates the lowest loss, italics the second lowest.

Table 6. Estimation of quarterly data, empirical example

INFLATION

							MISSING OBSERVATIONS 20%					
	DE	ES	FR	IT	US	AW	DE	ES	FR	IT	US	AW
DFM:	0,23	*0,114*	0,053	0,064	0,531	0,191	0,094	0,114	0,051	0,062	0,165	0,066
Chow-Lin:	0,235	0,13	*0,04*	*0,03*	0,478	0,2	0,1	*0,084*	*0,026*	**0,019**	0,132	**0,053**
Spline:	0,571	0,285	0,106	0,1	0,486	0,236	.	.	.	.	.	.
K-filter:	0,535	0,487	0,135	0,153	0,439	0,241	.	.	.	.	.	.
K-smoother:	0,485	0,487	0,127	0,138	0,44	0,24	0,278	0,1	0,596	5,858	0,104	0,933
P-Mean:	*0,222*	0,195	0,054	0,06	**0,34**	*0,156*	**0,068**	0,089	0,091	0,774	*0,094*	0,193
P-Median:	0,337	0,246	0,065	0,074	0,403	0,205	0,08	**0,083**	0,043	0,054	0,101	0,071
P-TrimMean:	0,261	0,207	0,059	0,069	*0,36*	0,182	0,119	0,089	0,217	*0,031*	**0,083**	0,352
P-Multi:	**0,183**	**0,1**	**0,031**	**0,029**	0,451	**0,15**	*0,084*	0,095	**0,022**	*0,031*	0,133	*0,054*
P-Uni:	0,462	0,393	0,101	0,115	0,427	0,233	.	.	.	.	.	.

REAL GDP GROWTH

							MISSING OBSERVATIONS 20%					
	DE	ES	FR	IT	US	AW	DE	ES	FR	IT	US	AW
DFM:	0,129	0,121	0,317	0,133	0,216	0,197	0,058	0,263	0,094	0,092	0,077	0,1
Chow-Lin:	0,071	1,364	0,146	0,104	**0,004**	*0,173*	0,038	0,285	0,059	**0,072**	**0,004**	**0,014**
Spline:	**0,058**	**0,069**	**0,097**	**0,078**	0,436	0,558	.	.	.	.	.	.
K-filter:	0,244	0,252	0,253	0,133	0,511	0,465	.	.	.	.	.	.
K-smoother:	0,129	0,122	0,162	0,102	0,493	0,463	0,112	**0,103**	0,058	0,136	0,282	0,23
P-Mean:	*0,067*	0,112	0,111	*0,078*	0,218	0,26	0,048	*0,138*	**0,038**	0,086	0,064	0,043
P-Median:	0,068	*0,081*	*0,106*	0,086	0,351	0,359	0,049	0,191	*0,041*	0,086	0,066	0,036
P-TrimMean:	*0,067*	0,087	*0,106*	0,083	0,273	0,306	**0,029**	0,265	0,068	0,092	0,106	0,049
P-Multi:	0,074	0,405	0,173	0,09	*0,056*	**0,12**	*0,034*	0,263	0,056	*0,075*	*0,019*	*0,027*
P-Uni:	0,108	0,107	0,125	0,097	0,459	0,464	.	.	.	.	.	.

Note: Inflation is treated as a stock variable, GDP growth as a flow variable. The Table reports the MSE.
DE: Germany; ES: Spain; FR: France; IT: Italy; US: United States; AW: euro area
Bold indicates the lowest loss, italics the second lowest.

Table 7. Properties of interpolated data, empirical example

INFLATION

	DE	ES	FR	IT	US	AW	ES	FR	IT	US	AW
DFM:	**0,012**	**0,016**	0,011	*0,012*	0,103	*0,009*	0,008	0,041	**0,002**	0,009	0.045
Chow-Lin:	0,043	0,023	0,012	**0,011**	**0,012**	**0,004**	0,023	0,119	0,053	0,063	0.004
Spline:	0,093	0,047	0,021	0,021	0,096	0,021	**0,003**	**0,012**	0,024	**0,001**	**0.000**
K-filter:	0,076	0,049	*0,01*	0,014	0,111	0,01	0,047	0,044	0,031	0,024	0.022
K-smoother:	0,093	0,05	*0,01*	0,015	0,111	0,01	0,041	*0,04*	0,026	0,024	0.014
P-Mean:	0,087	0,049	0,016	0,019	0,11	0.017	0,014	0,068	*0,022*	0,013	0.006
P-Median:	0,091	0,05	0,014	0,019	0,114	0.015	0,007	0,063	0,026	0,03	*0.002*
P-TrimMean:	0,091	0,05	0,016	0,019	0,114	0.016	0,015	0,064	0,023	0,021	0.002
P-Multi:	*0,019*	*0,025*	**0,008**	0,015	*0,082*	0.012	*0,005*	0,092	0,024	*0,008*	0.026
P-Uni:	0,095	0,052	0,016	0,018	0,115	0.016	0,031	0,041	0,025	0,019	0.010

MISSING OBSERVATIONS 20%

	DE	ES	FR	IT	US	AW	ES	FR	IT	US	AW
DFM:	0.036	0.060	0.004	*0.008*	**0,01**	0,026	0,055	0,187	**0,013**	0,015	0,411
Chow-Lin:	0.035	**0.055**	0.010	0.011	*0,011*	**0,008**	0,054	0,044	0,047	0,074	**0,071**
K-smoother:	0.021	0.057	0.006	0.015	0,015	0,045	0,223	0,433	1,413	**0,005**	0,411
P-Mean:	*0.009*	0.058	**0.002**	**0.001**	0,013	0,021	0,122	0,194	0,617	0,03	0,306
P-Median:	**0.003**	**0.055**	**0.002**	0.009	0,013	0,026	*0,021*	*0,14*	0,171	0,021	0,364
P-TrimMean:	0.014	0.056	0.006	0.011	*0,011*	0,035	0,128	0,28	0,237	*0,006*	0,344
P-Multi:	0.018	0.058	**0.002**	0.011	*0,011*	*0,013*	**0,014**	**0,09**	*0,015*	0,042	*0,245*

REAL GDP GROWTH

	DE	ES	FR	IT	US	AW	ES	FR	IT	US	AW
DFM:	0,748	**0,015**	*0,18*	*0,205*	0,425	0,297	0,524	0,375	0,471	0,061	0,204
Chow-Lin:	**0,228**	*0,019*	**0,062**	**0,122**	**0,005**	**0,025**	**0,156**	**0,052**	**0,069**	**0,024**	**0,019**
Spline:	0,939	0,078	0,312	0,416	0,459	0,346	0,515	0,367	0,442	0,05	0,211
K-filter:	0,844	0,065	0,263	0,313	0,35	0,296	0,518	0,372	0,436	0,06	0,235
K-smoother:	0,847	0,076	0,288	0,315	0,35	0,297	0,524	0,373	0,435	0,06	0,235
P-Mean:	0,828	0,069	0,279	0,338	0,405	0,299	0,507	0,359	0,424	0,057	0,223
P-Median:	0,867	0,075	0,287	0,347	0,415	0,311	0,537	0,382	0,45	0,071	0,246
P-TrimMean:	0,861	0,075	0,286	0,339	0,42	0,311	0,531	0,381	0,448	0,065	0,236
P-Multi:	*0,593*	0,026	0,196	0,219	*0,282*	*0,201*	*0,396*	*0,26*	*0,332*	*0,032*	*0,127*
P-Uni:	0,895	0,077	0,297	0,367	0,415	0,323	0,53	0,378	0,442	0,061	0,238

MISSING OBSERVATIONS 20%

DFM:	0,062	**0,025**	0,049	0,055	0,025	0,054	0,093	**0,004**	0,022	*0,028*	0,022
Chow-Lin:	0,067	0,036	**0,008**	0,056	**0,001**	0,067	0,059	0,025	0,025	0,034	0,025
K-smoother:	**0,022**	0,041	0,07	0,054	0,008	**0,022**	**0,028**	0,062	0,067	0,029	0,067
P-Mean:	0,062	0,038	0,051	**0,05**	0,019	0,058	0,078	0,012	*0,009*	0,035	*0,009*
P-Median:	0,08	*0,034*	0,054	*0,053*	0,019	0,074	0,072	0,014	0,013	0,035	0,013
P-TrimMean:	*0,029*	0,036	*0,036*	0,071	*0,004*	*0,027*	*0,049*	0,027	**0,002**	**0,023**	**0,002**
P-Multi:	0,071	*0,034*	0,038	0,054	0,024	0,066	0,086	*0,005*	0,029	0,034	0,029

Note: The table reports the difference (ρ) between the first order autocorrelation coefficient for the actual and interpolated series in the left-hand side, and the absolute value of the difference (β) of the estimated coefficient of y_t in the regression $w_t = y_t + u_t$ in the right-hand side,
where for each European country y is the same variable as w but for Germany, while for the euro area and the US
w is inflation and y output growth, and viceversa
Bold indicates the lowest loss, italics the second lowest.