Multivariate Singular Spectrum Analysis for Forecasting Revisions to Real-time Data

KERRY PATTERSON[*], HOSSEIN HASSANI[**], SAEED HERAVI[†] & ANATOLY ZHIGLJAVSKY[‡] [a]

[*]School of Economics, University of Reading, UK
[**]School of Mathematics, Cardiff University, UK
[†]Cardiff Business School, Cardiff University, UK
[‡]School of Mathematics, Cardiff University, UK

ABSTRACT: Real-time data on national accounts statistics typically undergoes an extensive revision process leading to multiple vintages on the same generic variable. The time between the publication of the initial and final data is a lengthy one and raises the question of how to model and forecast the final vintage of data – an issue that dates from seminal articles by Mankiw et. al., (1984), Mankiw and Shapiro (1986) and Nordhaus (1987). To solve this problem, we develop the nonparametric method of multivariate Singular Spectrum Analysis, MSSA, for multi-vintage data. MSSA is much more flexible than standard methods of modelling that involve at least one of the restrictive assumptions of linearity, normality and stationarity. The benefits are illustrated with data on the UK Index of Industrial Production: neither the preliminary vintages nor the competing models are as accurate as forecasts using MSSA.

KEY WORDS: nonparametric methods, data revisions, trajectory matrix, reconstruction, Hankelisation, recurrence formula, forecasting.

Correspondence Address: Kerry Patterson, Department of Economics, University of Reading, RG6 6AA, UK. Email: k.d.patterson@reading.ac.uk. The authors acknowledge the helpful comments of three referees and the editor.

[a] The authors are grateful to an anonymous referee and the editor for their constructive comments on an earlier version of this paper.

Introduction

Much economic data, particularly for the national income and product accounts, is subject to a revision process, resulting in what is known as real-time data[1]. The stylised measurement process is that data is available on a variable, generically denoted y, at time t for an observation period (1, …, t – s), s ≥ 0, where s is the publication lag; then at t + 1 (say), data on ostensibly the same generic variable becomes available for (1, …, t – s + 1) and, typically, at least some data on the overlapping sample period differs, see Patterson (1995a). This process continues until, eventually, there is a sub-sample of data at time t + m that is (generally) unchanging, which concludes the revision process for that part of the data sample. The initial and revised data at different stages are referred to as different vintages of observations on y. We adopt the notation that there are $v = 1, \ldots, m$ vintages, with m being the last in a conditional sense; that is there is no presumption that this is the 'true' data measured without error, although there may be a temptation for users to make that assumption, at least implicitly. Rather, the 'final' vintage is the result of the conclusion of the measurement process. Typically one finds that the sequence of vintages become less variable as m is approached (which implies that the preliminary vintages are not efficient forecasts).

The preliminary vintages naturally predate the final vintage and can thus be regarded as forecasts of the final vintage and, indeed, forecasts of any subsequent vintage. This was the focus of the seminal early articles by Mankiw et. al., (1984), Mankiw and Shapiro (1986), Mork (1987) and Nordhaus (1987), which shaped much subsequent research. That research established the view that preliminary vintages may be rational (or efficient) forecasts in the sense that subsequent revisions are 'news' relative to the information set available at the time the preliminary vintages were formed; alternatively, the preliminary vintages may be less than efficient forecasts of later vintages. The two limiting cases are where the revisions are orthogonal to the preliminary vintage (pure 'news') and where revisions are orthogonal to a later vintage (pure 'noise'). Much research has considered this question and for further references, the interested reader is referred to the website maintained by Dean Croushore (2009).

The problems faced by policy makers, who need to take a timely view on the development of the economy and, at the same time, base their decisions on uncertain data, has led to a continuing interest in the properties of data revisions, see, for example, research at the Bank of England, Castle and Ellis (2002), Ashley et. al., (2005) and Cunningham et. al., (2007), at the Bureau of Economic Analysis, see Fixler and Nalewaik (2006) and at the Japanese Cabinet Office, see Kawagoe (2007). Research on methods has also continued, for example, an interesting parametric method, which develops an earlier state space approach to modeling data revisions (see, for example, Harvey, 1983, and Howrey, 1984) is due to Jacobs and van Norden (2010); this reflects the current research interest in modeling data revisions in a systems approach (for further references, see endnote 1). A recent contribution by Clements and Galvão (2009) illustrates the continuing interest in the impact of data revisions on forecasting models.

In a recent study Aruoba (2008) notes: "It is of great interest for practitioners and policy makers to find ways of exploiting the potential forecastability in real time …". It is this forecastability that is the primary concern of this study. We construct forecasts of the final vintage by different methods and evaluate them against the baseline of the preliminary vintages. In a sense this is a development of the 'news versus noise' debate, since if we can improve upon a preliminary vintage as a forecast of a later vintage using the same information set, then that preliminary vintage cannot be an efficient forecast of the later vintage.

We consider two approaches to constructing forecasting models. Typically, the time series models used for forecasting are parametric and based on at least one of the restrictive assumptions of normality, stationarity and linearity; however, quite often financial and economic time series data are non-Gaussian and may be generated by processes that are nonstationary and/or nonlinear. The system based parametric approaches to modelling the structure of data revisions, such as a state space formulation, have also been applied within an essentially linear framework. Parametric models also require a specification that should match the underlying data generation process, typical choices being ARMA and VAR models. In contrast methods that do not depend on these assumptions are likely to be useful for modelling and forecasting economic data and, therefore, in the second approach, we construct models based on Singular Spectrum Analysis (SSA), which is a nonparametric method that does not embody the standard assumptions and exploits an analysis of the data in phase space rather than parameter space.

The plan of this paper is as follows. The next section gives an overview to enable the key issues to be outlined as a guide to this study. The subsequent section outlines the framework for the interpretation of multi-vintage data; the MSSA methodology is presented in the fourth section; the fifth section considers the framework for forecast assessment; and the sixth section illustrates the application of SSA/MSSA and alternative parametric methods to the *IIP*, including an assessment of nonlinear dependence and non-normality; some concluding remarks are reserved for the final section. (An appendix gives details of the specification of the VAR models used in this study).

Overview

The SSA^2 method consists of two complementary stages: decomposition and reconstruction (or estimation), with each stage comprising two separate steps. At the first stage, the series is decomposed into mutually orthogonal components and in the second stage the original series is reconstructed selecting those components that reduce the noise in the series. The reconstructed series is then used for forecasting new data points. The methods of SSA can be applied to a single series or jointly to several series and in the latter case it is referred to as multivariate SSA or MSSA; as in the case of parametric modelling, two or more series may be related, which in the context of MSSA has a correspondence in terms of matched components of the several series.

A particular concern of this study is to show how the flexibility of the MSSA can be used in forecasting data subject to multiple revisions, comparing that method with forecasts obtained by more conventional parametric models. In so doing, we make a distinction between forecasting 'within-sample' and forecasting 'out-of-sample'. By the nature of the data measurement process, data for different vintages, $v = 1, \ldots, m$, is published sequentially; that is letting the data for the v-th vintage of the generic variable y for time period t be denoted $y_t^{(v)}$, it is necessarily the case that publication of this vintage precedes publication of vintages $y_t^{(v+1)}$, $y_t^{(v+2)}$ and so on until the final vintage $y_t^{(m)}$. Thus, one forecasting problem is to forecast $y_t^{(m)}$ having available data on some of the preceding vintages relating to time t and preceding time periods. This is a within-sample forecasting[3] problem, provided that the sample is interpreted as including the latest data on the earliest vintage.

In contrast, the out-of-sample problem occurs when the data period to be forecast starts later than the period for which the most recent data on the earliest vintage is available. In this case, by definition, the first vintage of the data is not available and hence neither that nor any other vintage can act as the default forecast. This is the kind of problem in which it is necessary to 'fix' a sample period or more precisely in our case an information set, on which estimation of the phase space or parameter space is based; given that information, the underlying algorithm projects the data forward in time.

We have found the MSSA approach to be beneficial for a number of time series that are subject to revision. To illustrate the benefits we report results for the Index of Industrial Production (*IIP*) for the U.K, as in Patterson (2002). This variable is taken as a key indicator of the state of the U.K's industrial base and is regarded as a leading (monthly) indicator of the general state of the U.K economy. We are able to benchmark the improvement in forecasting using MSSA relative to standard linear models, such as ARMA and VAR models.

In the case of within-sample forecasting, the comparison also includes the use of the preliminary vintages as the default forecasts; additionally in this context, the comparison is extended to recursive updating of the modelling space. Forecasting performance is evaluated in a number of ways. In addition to a standard comparison using the root mean squared error, we consider whether differences in forecasts are statistically significant, see Diebold and Marino (1995) and Harvey et. al., (1997), and whether the methods are able to forecast the direction of change. Overall, we find that there are significant gains in using MSSA.

The General Structure of Real-time Data

To establish some notation, let the variable of interest (the 'generic' variable) be denoted y and let $y_t^{(v)}$ be the v-th vintage ($v = 1, \cdots, m$) of y for the period t, where $v = 1$ indicates the initially published data and $v = m$ the finally published data. The structure of the published data at time t, assuming a one period publication lag, can be represented[4] in a data matrix Y_t of dimension $(t-1)\times m$, with each column

designating a vintage and each row a time period, as follows, where n.a indicates data not available at time = *t*:

$$Y_t = \begin{bmatrix} y_1^{(1)} & y_1^{(2)} & y_1^{(3)} & \cdots & y_1^{(m)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{t-m}^{(1)} & y_{t-m}^{(2)} & y_{t-m}^{(3)} & \vdots & y_{t-m}^{(m)} \\ \vdots & \vdots & \vdots & \vdots & n.a \\ y_{t-3}^{(1)} & y_{t-3}^{(2)} & y_{t-3}^{(3)} & n.a & n.a \\ y_{t-2}^{(1)} & y_{t-2}^{(2)} & n.a & n.a & n.a \\ y_{t-1}^{(1)} & n.a & n.a & n.a & n.a \end{bmatrix} \quad (1)$$

Note that publication from a particular issue of the data source at time *t* traces back a diagonal of the observation matrix Y_t, which is a composite of data of different vintages; for example, at time *t* the published data is $y_{t-1}^{(1)}$, $y_{t-2}^{(2)}$, …, $y_{t-m}^{(m)}$. This is sometimes referred to as 'real-time' data. The first complete set of vintages available at time *t* is $y_{t-m}^{(v)}$, $v = 1, \ldots, m$. The time series of the *v*-th vintage of data is denoted $Y^{(v)} = (y_1^{(v)}, \ldots, y_{T_v}^{(v)})'$ where T_v is the length of $Y^{(v)}$ and $v = 1, \ldots, m$.

Multivariate Singular Spectrum Analysis (MSSA) for Multi-vintage Data

The data vintages are part of a measurement system characterised by $m > 1$ vintages of the generic variable *y* and this section contains a brief description of the method of MSSA developed for the case of multiple data vintages.

Each time series component of the multivariate system is viewed as the sum of unobservable components: the signal comprises components such as the trend, oscillations or periodic movements, and noise. This decomposition excludes the possibility of chaotic dynamics and we, therefore, also conducted a test, based on Lyapunov exponents, for the absence of this feature from our data[4]; for the general issues see, for example, Rivero et. al., (2005) and for a practical test see Saïda (2007).

The aim of MSSA is to extract the signal leaving the residual; more generally, the algorithm can also extract groups corresponding to components of the signal. The two stages to the process are decomposition and reconstruction, each of which comprises two steps. Finally, the MSSA algorithm provides forecasts via a linear recurrence formula.

Stage 1: *Decomposition: Embedding and Singular Value Decomposition (SVD)*

Step 1: *Embedding*

Embedding is a mapping that translates a one-dimensional time series into a multi-dimensional series through the use of subsets of the original series. The key output in this stage is the trajectory matrix, generically referred to as *X*. This is a matrix that is formed by taking a window of observations of length *L* and moving this throughout the sample. Such a procedure will be familiar from time series analysis that focuses on calculating moving averages or recursive estimation with a moving window; here moving vectors of observations are created.

To see how this works we take a window comprising the first L_v observations of $Y^{(v)}$, then drop the first observation and add the (L_v + 1)-th observation to create another same length window (vector). This process is continued, with the data organised into a matrix $X^{(v)}$, of dimension $L_v \times K_v$, where $K_v = T_v - L_v + 1$. The resulting trajectory matrix $X^{(v)}$ is:

$$X^{(v)} = \begin{bmatrix} y_1^{(v)} & y_2^{(v)} & \cdots & y_{T_v-L_v+2}^{(v)} & y_{T_v-L_v+1}^{(v)} \\ y_2^{(v)} & y_3^{(v)} & \cdots & y_{T_v-L_v+1}^{(v)} & y_{T_v-L_v+2}^{(v)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{L_v-1}^{(v)} & y_{L_v}^{(v)} & \cdots & y_{T_v-2}^{(v)} & y_{T_v-1}^{(v)} \\ y_{L_v}^{(v)} & y_{L_v+1}^{(v)} & \cdots & y_{T_v-1}^{(v)} & y_{T_v}^{(v)} \end{bmatrix} \quad (2)$$

In this set-up T_v and L_v are allowed to differ depending on *v*; thus, in general $X^{(v)}$ is of dimension $L_v \times K_v$ where $K_v = T_v - L_v + 1$. In practice, a common L_v is chosen in the case that the trajectory matrices are stacked horizontally to obtain the system trajectory matrix. For simplicity of exposition, we also assume $T = T_1 = \ldots = T_m$ implying $K_1 = \ldots = K_m = K$; this assumption is relaxed below.

The trajectory matrix for the system as a whole comprising $Y^{(v)}$, $v = 1, \ldots, m$, is obtained by stacking the trajectory matrices horizontally to form the trajectory matrix of the multidimensional series. The resulting trajectory matrix X of dimension $L \times mK$, is given by:

$$\begin{aligned} \mathrm{X} &= \left(X^{(1)}; \ \ldots; \ X^{(v)}; \ \ldots; \ X^{(m)}\right) \quad (3) \\ &= \left(X_1^{(1)} \ \ldots \ X_K^{(1)}; \ \ldots; \ X_1^{(v)} \ \ldots \ X_K^{(v)}; \ \ldots; \ X_1^{(m)} \ \ldots \ X_K^{(m)}\right) \end{aligned}$$

Notice that each of the *m* blocks of *K* columns corresponds to the trajectory matrix for a particular vintage. *X* is the trajectory matrix for the system of data vintages given by $Y = (Y^{(1)}, \ldots, Y^{(v)}, \ldots, Y^{(m)})$ which, in this simplified case, is a vector of dimension $mT \times 1$, where $Y^{(v)} = (y_1^{(v)}, \ldots, y_T^{(v)})'$. The case where T_v and so K_v are not equal is easily accommodated; in that case because the individual trajectory matrices are stacked horizontally, they can be of different column dimensions. Thus, *X* is of dimension $L \times \sum_{v=1}^{m} K_v$ and *Y* is of dimension $\sum_{v=1}^{m} T_v \times 1$.

The trajectory matrix is an example of a Hankel matrix in which the diagonal elements are equal for all combinations where the sum of the row (*i*) and column (*j*) indices are equal to a constant; that is, $X_{ij} = X_{ji}$ for $ij = c$. Visually, the diagonals are those on a line from the South-West to the North-East of the matrix, which are referred to as the Hankel diagonals. This is a characteristic that is used in the second

stage in which the original series are reconstructed using the principal components obtained in the next step.

Step 2: *obtain the singular value decomposition, SVD, of the system trajectory matrix*
The second step in stage 1 is to construct the SVD of the trajectory matrix X and represent it as a sum of $d \leq L$ rank-one, mutually orthogonal elementary matrices. First define the matrix $C = XX'$ and denote by $\lambda_1, \ldots, \lambda_d$ the ordered non-negative eigenvalues of C, such that $\lambda_1 \geq \ldots \geq \lambda_d \geq 0$; where $d \leq L$, with $d = L$ if all $\lambda_i \geq 0$. The corresponding eigenvectors are $\{U_i\}_{i=1}^{d}$; the factor vectors are $\{V_i\}_{i=1}^{d}$ where $V_i = X'U_i / \sqrt{\lambda_i}$ are of dimension $mK \times 1$. The principle component vectors are $\sqrt{\lambda_i} V_i$ and the eigentriple that forms the basis of the SVD is $(\sqrt{\lambda_i}, U_i, V_i)$.

The trajectory matrix X is decomposed into the sum of d elementary matrices $X_i = \sqrt{\lambda_i} U_i V_i'$, such that:

$$X = \sum_{i=1}^{d} X_i \qquad (4)$$

The matrices $\mathrm{X_i}$ are referred to as elementary matrices, which have rank 1 and are, by construction, mutually orthogonal. The SVD given by (4) is optimal in the sense that among all the matrices of rank $r < d$, the matrix $X^{(r)} = \sum_{i=1}^{r} X_i$ provides the best approximation to the trajectory matrix X in the norm sense, such that $\| X - X^{(r)} \|$ is a minimum.

The contribution of the component X_i to the expansion (3) is given by its eigenvalue λ_i as a share in the sum of the eigenvalues, that is $\lambda_i / \sum_{j=1}^{d} \lambda_j$. The singular spectrum (hence the description singular spectrum analysis) refers to a graph of the ordered eigenvalues, $\lambda_1 \geq \ldots \geq \lambda_d \geq 0$ and is useful in deciding which principal components to include in the reconstruction step of the SSA method. If none of the eigenvalues are negative, then the singular spectrum is a graph of the L ordered eigenvalues.

The aim of the first two steps is to achieve separability of the components in the decomposition of the series, for example, in a simple case into the trend, any periodic elements and the noise. In this respect the selection of L is critical, which must be large enough to allow a separation into the components. If L is too small, not all of the components will be captured and it must, therefore, be large enough so that each column of $X^{(v)}$ captures an essential part of the (time series) behaviour of $Y^{(v)}$. On the other hand, if L is too large there will be too few 'windows' to follow the evolving behaviour of $Y^{(v)}$. The degree of separability can be assesed empirically by, for example, means of the weighted correlation coefficients – see equation (14) below. The use of the SVD has an anology with principal components analysis (PCA), in which orthogonal components of the orginal series are extracted and a reduced

dimensionality can be achieved. In SVD, the directions determined by the d eigenvectors, $\{U_i\}_{i=1}^{d}$, are orthogonal and maximise the variation in that direction, just as in PCA.

Stage 2: *Reconstruction, Hankelisation and Gouping*
Step 3: *diagonal averaging (block Hankelisation)*
In the first stage of this step, the elementary matrices X_i, corresponding to the i-th principal component in the SVD, are used to (re)construct series of the same length as the original series. Note that X is in m blocks, one for each vintage of data; thus each block is Hankelised and then the m resulting $T_v \times 1$ vectors are stacked vertically into a $\sum_{v=1}^{m} T_v \times 1$ vector; if T_v = T then the resulting vector is $mT \times 1$. We briefly describe the Hankelisation procedure for one of these blocks.

The Hankelisation procedure may be represented by first rearranging X_i so that the sum of each Hankel diagonal is one element in a $T \times 1$ vector, say $X'_{H,i}$ and then premultiplying by a diagonal matrix $H = \text{diag}(h_i)$, with diagonal elements that are the inverse of the number of elements in the corresponding row of $X_{H,i}$. That is:

$$\tilde{X}_i = HX'_{H,i}$$

$$= \begin{pmatrix} x_{1,1}^{(i)} \\ (x_{2,1}^{(i)} + x_{1,2}^{(i)})/2 \\ (x_{3,1}^{(i)} + x_{2,2}^{(i)} + x_{1,3}^{(i)})/3 \\ \vdots \\ (x_{L,K-2}^{(i)} + x_{L-1,K-1}^{(i)} + x_{L-2,K}^{(i)})/3 \\ (x_{L,K-1}^{(i)} + x_{L-1,K}^{(i)})/2 \\ x_{l,K}^{(i)} \end{pmatrix} \qquad (5)$$

where $x_{j,k}^{(i)}$ is the (i,k)-th element of X_i. Note that the sum of the subscripts in each row is the same; thus, the general element sums and then averages the Hankel diagonal elements for each row element. The result is a $T \times 1$ vector of the time series components corresponding to the i-th principal component of the trajectory matrix. The $T \times r_v$ matrix $\tilde{X}$ of all reconstructed components is defined with typical column vector $\tilde{X}_i$, thus $\tilde{X} = [\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{r_v}]$.

Step 4: *grouping*
Let r denote the trajectory dimension resulting from grouping in the multivariate case and let r_v denote the dimension of the univariate trajectory space for the v-th vintage. Golyandina and Stepanov (2005) show that $r_{min} \le r \le r_{max}$, where $r_{min} = \max\{r_v : v = 1, \ldots, m\}$ and $r_{max} = \sum_{v=1}^{m} r_v$. Thus, the ceiling to the multivariate trajectory dimension is simply that obtained when there are, in a sense, no common or matched

components amongst the m vintages; the presence of matching components reduces the dimension of the multivariate system and indicates that the system is interrelated, which should result in gains when forecasting the series. The selection of $r < r_{min}$ leads to a loss of precision as parts of the signals in all series will be lost. From the other side, if $r > r_{max}$ then noise is included the reconstructed series. The selection of $r \cong r_{min}$ (keeping $r > r_{min}$) is a good choice for highly interrelated series sharing several common components. The selection of $r \cong r_{max}$ is necessary when the series analysed have very little relation to each other.

The result of the grouping step is the reconstructed series, or estimated signal in this case, for all m vintages; that is:

$$\tilde{Y} = \sum_{i \in I} \tilde{X}_i \qquad (6)$$

So that $\tilde{Y}$ is the estimation counterpart of the actual series Y and of dimension $mT \times 1$ or, more generally, $\sum_{v=1}^{m} T_v \times 1$.

Forecasting multiple series

The SSA forecasting method can be applied to time series that satisfy the linear recurrence formula, which for a single series, is given by:

$$y_t = \sum_{j=1}^{L-1} a_j y_{t-j} \qquad (7)$$

where $\{a_j\}_{j=1}^{L-1}$ is a sequence of constant coefficients. The class of time series governed by the linear recurrence formula is wide, including harmonics, polynomials and exponential time series, see Golyandina and Stepanov (2005). Within the estimation sample, fitted values of y_t are just given by the reconstructed series; otherwise, forecasts are dynamic and given by application of the recurrence formula.

Letting a ^ over a variable denote an estimated or forecasted value, then for a single series, forecasts are obtained as follows.

Estimation sample, for $t = 1, \ldots, T$:

$$\hat{y}_t^{(v)} = \tilde{y}_t^{(v)} \qquad (8a)$$

where $\tilde{y}_t^{(v)}$ is the t-th element of $\tilde{Y}^{(v)}$.

Dynamic forecasts, $t = T + 1, \ldots, T + h$:

$$\hat{y}_{T+h}^{(v)} = \sum_{j=h}^{L-1} a_j y_{T+h-j}^{(v)} + \sum_{j=1}^{h-1} a_j \hat{y}_{T+h-j}^{(v)} \quad \text{for } h \le L-1 \qquad (8b)$$

$$\hat{y}_{T+h}^{(v)} = \sum_{j=1}^{L-1} a_j \hat{y}_{T+h-j}^{(v)} \quad \text{for } h > L-1 \qquad (8c)$$

Note that (8b) uses the actual values of y_{T+h-j} where they are available, that is where $T + h - j \leq T$.

For simplicity of exposition assume that K_v = K, then in the case of multiple series, there is one $(K-1)\times 1$ vector of recurrence coefficients for each of the v = 1, …, m time series. Let $\hat{Z}$ denote the $m(K-1)\times 1$ overall vector of vertically stacked vectors of the last $K-1$ forecast values, that is:

$$\hat{Z} \quad = \left(\hat{y}^{(1)}_{T+h-K+1} \quad \dots \quad \hat{y}^{(1)}_{T+h-1}; \quad \dots; \quad \hat{y}^{(m)}_{T+h-K+1} \quad \dots \quad \hat{y}^{(m)}_{T+h-1}\right)' \tag{9}$$

Then the forecasts are obtained by application of an extension of the linear recurrence rule to the multivariate case. The general representation is:

$$\hat{F} \quad = \Psi\hat{Z} \tag{10}$$

where $\hat{F} = (\hat{y}^{(1)}_{T+h}\ \hat{y}^{(2)}_{T+h}\ \dots\ \hat{y}^{(m)}_{T+h})'$. Where actual values are available they replace forecasted values.

The multivariate recurrence coefficients are given in the Ψ matrix of dimension $m\times m(K-1)$, where:

$$\Psi \quad = (I_m - \Pi\Pi')^{-1}\Pi Q' \tag{11}$$

The components of Ψ are defined as follows. To obtain Q, first note that an output of the SVD is the set of factor vectors $V_j = X'U_i / \sqrt{\lambda_j}$, j = 1, …, r, which are each of dimension $(mK\times 1)$; next, as in the univariate case, a new vector is created, but with the m elements at K, $2K$, …, mK excluded. The resulting vector is denoted $V_j^{\nabla m}$, and these vectors, j = 1, …, m, are collected into the $m(K-1)\times r$ matrix Q, such that:

$$Q \quad = \left(V_1^{\nabla s}; \dots; \quad V_j^{\nabla s}; \quad \dots; \quad V_r^{\nabla s}\right) \tag{12}$$

The set of excluded elements from V_j becomes the j-th column vector of the $m\times r$ matrix $\Pi = [\Pi_1 \quad \dots \quad \Pi_j \quad \dots \quad \Pi_r]$, that is $\Pi_j = (\pi_j^{(1)} \quad \dots \quad \pi_j^{(v)} \quad \dots \quad \pi_j^{(m)})'$, such that:

$$\Pi \quad = \begin{bmatrix} \pi_1^{(1)} & \dots & \pi_j^{(1)} & \dots & \pi_r^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \pi_1^{(v)} & \dots & \pi_j^{(v)} & \dots & \pi_1^{(v)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \pi_1^{(m)} & \dots & \pi_j^{(m)} & \dots & \pi_r^{(m)} \end{bmatrix} \tag{13}$$

Selection of parameters: L

In principle more than two groups corresponding to the signal and the system noise or residual, can be identified in MSSA; however, in our case, interest centres on this two-fold split, the signal providing the basis on which to forecast the series. MSSA requires the selection of two parameters: the window length L and the number of elementary matrices r from which to form the index set I. Their selection depends on the structure of the data and the purpose of the analysis. There are some particular considerations of the structure of the multi-vintage data that affect the choice of L.

In the case of multi-vintage data subject to the kind of data measurement process outlined in second section, the data vector for each vintage is of a different length. For example, consider a bivariate system with vintages v and m, (which we find to work well for IIP). The data vectors are then of dimension T_v and T_m, respectively, where $T_v - T_m = m - v > 0$; this is evident from the data matrix Y_t, see (1) above. In effect within-sample forecasting relates to 'squaring' off the matrix Y_t, replacing the not available (n.a) observations by forecasts using the SSA recurrence relationships; the use of $Y^{(v)}$ and $Y^{(m)}$ implies that L should be chosen such that $L \geq m - v + 1$ so that at least one observation of the final vintage is used in the linear recurrent formula. In the case of a system with more than two vintages, L should be chosen such that $L \geq m - \min(v) + 1$, where $\min(v)$ is the minimum of the v values.

Otherwise, L should be selected to enable separability of the components, here into the signal and the noise. For example, if L is small, the trajectory matrix would be influenced by the system noise, so L should be sufficiently large to capture the local trend without being too influenced by the noise. The maximum number of principal components in the SVD is L, so the singular spectrum can be helpful in deciding on the selection of L. The first elementary matrix X_1 with the norm $\sqrt{\lambda}_1$ has the largest contribution to the norm of $\mathbf{X}$ in (3) and the last elementary matrix X_d, with the norm $\sqrt{\lambda}_d$, has the smallest contribution to the norm of $\mathbf{X}$.

Selection of parameters: r

The aim in choosing the value of r is to separate the noise and signal; the SSA decomposition is successful if the resulting additive components of the series are separated from each other. The weighted correlation or *w*-correlation, see Golyandina et. al., (2001) can be used in determining separability. It is a geometrically motivated measure of dependence between two series, say, $Z_T^{(i)}$ and $Z_T^{(j)}$, defined as:

$$\rho_{i,j}^{w} = \frac{\langle Z_T^{(i)}, Z_T^{(j)} \rangle_w}{\| Z_T^{(i)} \|_w \| Z_T^{(j)} \|_w} \qquad (14)$$

where $\langle Z_T^{(i)}, Z_T^{(j)} \rangle_w = \sum_{t=1}^{T} w_t z_t^{(i)} z_t^{(j)}$, $\| Z_T^{(i)} \|_w = \sqrt{\langle Z_T^{(i)}, Z_T^{(i)} \rangle_w}$ and $z_t^{(i)}$ is a typical element of $Z_T^{(i)}$. The series are *w*-orthogonal if $\rho_{i,j}^{w} = 0$, which is necessary for separability. The weights are:

$w_t = t$ for $1 \le t \le L^*$

$w_t = L^*$ for $L^* < t \le K^*$

$w_t = T - t$ for $K^* < t \le T$

where $L^* = \min(L, K)$, $K^* = \max(L, K)$.

The w-correlation, $\rho_{i,j}^w$, is calculated for pairs of reconstructed components corresponding to pairs of the L eigenvalues in the SVD. Large values of $\rho_{i,j}^w$ between the reconstructed components indicate that these are candidates to be put into one group, corresponding to the same element in the MSSA decomposition

The next stage is to obtain the sequence of cumulative *w*-correlations. The original time series is first reconstructed using just the first component $Z_T^{(1)} = \tilde{X}_1$, considering the remaining components as the noise element, say $Z_T^{(2)} = \sum_{j=2}^{m} \tilde{X}_j$ corresponding here to eigentriples 2, …, m. Then the *w*-correlation between $Z_T^{(1)}$ and $Z_T^{(2)}$ is calculated, denote it C(1).

Next, take the first two eigenvalues, obtain the reconstructed series and consider the rest, that is using eigentriples 3, …, m, as noise, and again calculate the *w*-correlation, denote it C(2). Continue, so that a sequence of cumulative *w*-correlations is constructed as C(k), $k = 1, \ldots, L - 1$, between the reconstructed series from the first k components and the remaining $L - k$ components. The existence of structure in the series is indicated by local minima and maxima in the graph of C(k) against k. For example, a typical pattern is a decline in the cumulative *w*-correlations corresponding to a separation of the components as k increases; the first local minimum indicates the first separation and subsequent local maxima suggest possible secondary structure.

Forecasting assessment framework

In this section, the forecasting problem is considered in greater detail; this is a problem to be considered in 'real time', that is as it would present itself to a practitioner who has available at time t the data summarised in the data matrix Y_t of equation (1). As noted in the introduction, two aspects of the forecasting problem are distinguished depending on the information that is available at the time of forecasting.

Forecasting within-sample

The first of the real-time forecasting problems at time t is to complete the final column of the data matrix Y_t, with forecasts of $y_{t-s}^{(m)}$, $s = m - 1, \ldots, 1$; (forecasts are indicated by ^ above). This is referred to as a within-sample problem in the sense that the time subscript, $t - 1$, of the last forecast observation, $\hat{y}_{t-1}^{(m)}$, matches that of the most recent observation, $y_{t-1}^{(1)}$, in the data matrix at time t. In contrast, the out-of-sample problem refers to forecasts that extend the data matrix.

The data matrix Y_t of (1) is reproduced below, but with the final column completed by forecasts:

$$\hat{Y}_t = \begin{bmatrix} y_1^{(1)} & y_1^{(2)} & y_1^{(3)} & \cdots & y_1^{(m)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{t-m}^{(1)} & y_{t-m}^{(2)} & y_{t-m}^{(3)} & \cdots & y_{t-m}^{(m)} \\ y_{t-m+1}^{(1)} & y_{t-m+1}^{(2)} & y_{t-m+1}^{(3)} & \vdots & \hat{y}_{t-m+1}^{(m)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{t-3}^{(1)} & y_{t-3}^{(2)} & y_{t-3}^{(3)} & & \hat{y}_{t-3}^{(m)} \\ y_{t-2}^{(1)} & y_{t-2}^{(2)} & & & \hat{y}_{t-2}^{(m)} \\ y_{t-1}^{(1)} & & & & \hat{y}_{t-1}^{(m)} \end{bmatrix} \qquad (15)$$

One way of obtaining the forecasts is to project forward using only data in the final column, as in the case of a univariate model using either standard ARMA modelling or univariate SSA. At time t, the information set on which to base the univariate forecast is fixed at $\Omega_{t-m}^{(m)}$, that is the final column of the data matrix:

$$\Omega_{t-m}^{(m)} = (y_1^{(m)}, \ldots, y_{t-m}^{(m)})' \qquad (16)$$

The *h*-step ahead forecasts are formed conditional on $\Omega_{t-m}^{(m)}$, and the conditional within-sample *h*-step ahead forecast of $y_{t-v}^{(m)}$ is denoted:

$$\hat{y}_{t-v,h}^{(m)} \mid \Omega_{t-m}^{(m)} \qquad v = m - h;\ h = 1, 2, \ldots, m-1 \qquad (17)$$

For *h* varying, this generates the sequence $\{\hat{y}_{t-v,h}^{(m)} \mid \Omega_{t-m}^{(m)}\}_{h=1}^{m-1}$. Alternatively, fixing *h* (for example, for the 1-step ahead forecast, $h = 1$) and allowing *t* to vary, generates the sequence $\{\hat{y}_{t-v,h}^{(m)} \mid \Omega_{t-m}^{(m)}\}_{t=T}^{T+n}$; it is these sequences for $h = 1, 2, \ldots, m-1$, which are used in constructing measures of forecasting comparison.

In fact, this univariate procedure makes no use of information in Y_t other than that of the final column with terminal observation, $y_{t-m}^{(m)}$. In this case, the estimation sample for the univariate model of the m-th vintage is $(1, \ldots, t-m)$, so that forecasting first takes up the within-sample difference $(t-m+1, \ldots, t-1)$ and then moves to the out-of-sample period $(t, t+1, \ldots, t+n)$, $n \geq 1$.

Alternative forecasts are provided by the preliminary vintages. Generally, $y_{t-v}^{(v)}$, available at $t - v + 1$, is an *h*-step ahead forecast of $y_{t-v}^{(m)}$, where $h = m - v$, $v = 1, \ldots, m-1$ implying $h = m-1, \ldots, 1$. In more detail, at time *t*, $y_{t-1}^{(1)}$ is published, this is the first forecast internal to the measurement system at *t* of the final vintage $y_{t-1}^{(m)}$, where the latter is not available until a further $m-1$ time periods have elapsed. Also at time *t*, $y_{t-2}^{(2)}$ is published, which can be regarded as the $(m-2)$-period ahead forecast of $y_{t-2}^{(m)}$; this is the second available forecast of $y_{t-2}^{(m)}$, the first being the $(m-1)$-period

ahead forecast at time t, $y^{(1)}_{t-2}$; this process continues through to the publication at time t of $y^{(m-1)}_{t-(m-1)}$, which is the one-period ahead internal forecast of $y^{(m)}_{t-(m-1)}$, being the eleventh in the series of internal forecasts of $y^{(m)}_{t-(m-1)}$.

As an alternative to the univariate approach, the different vintages can be modelled jointly in a multivariate system. This could take the form of a VAR model or, in the present context, the MSSA algorithm can be applied to the system. For example, consider a bivariate system comprising vintages v and m, then the data at time t comprises the vectors $Y^{(v)}_{t-v} = (y^{(v)}_1, \ldots, y^{(v)}_{t-v})'$ and $Y^{(m)}_{t-m} = (y^{(m)}_1, \ldots, y^{(m)}_{t-m})'$. The trajectory matrix that follows from this system is an example of equation (10) and the MSSA algorithm is then applied to the bivariate trajectory matrix.

This procedure can be repeated for vintages v, $v < m$, and m. In principle, a sequence of $(v + 1)$-th order MSSA models could be constructed starting with $Y^{(1)}_{t-1}$ and $Y^{(m)}_{t-m}$, continuing with $Y^{(1)}_{t-1}$, $Y^{(2)}_{t-2}$ and $Y^{(m)}_{t-m}$ and so on through to the m-th order system comprising $Y^{(1)}_{t-1}$, $Y^{(2)}_{t-2}$, …, $Y^{(m-1)}_{t-(m-1)}$ and $Y^{(m)}_{t-m}$. This can be reduced to a sequence of bivariate models if vintage v data in a sense encompasses data of previous vintages (for example, in adding nothing significant in a mean squared error sense to the accuracy of the forecasts). We found that the simpler bivariate approach was as effective as the corresponding higher order model and, therefore, the empirical results reported below relate to this case.

The information set for the within-sample bivariate forecasts is:

$$\Omega^{(v,m)}_{t-v} = (y^{(v)}_1, \ldots, y^{(v)}_{t-v})'; (y^{(m)}_1, \ldots, y^{(m)}_{t-m})' \qquad (18)$$

That is $\Omega^{(m)}_{t-m}$ is appended by the v-th column of Y_t. The conditional within-sample h-step ahead bivariate forecast of $y^{(m)}_{t-v}$ is denoted:

$$\hat{y}^{(m)}_{t-v,h} \mid \Omega^{(v,m)}_{t-v} \qquad v = m - h;\ h = 1, 2, \ldots, m - 1 \qquad (19)$$

At time t, this forecast uses information on $y^{(v)}$ to time $t - v$ and on $y^{(m)}$ to time $t - m$ to produce h-step ahead forecasts of $y^{(m)}$ in order to complete the final column of $\hat{Y}_t$.

These forecasts, whether based on a univariate or bivariate model, are recursive in the sense that as t increments by one at each step in the evaluation sample, the underlying model is updated by the new sample observation.

Forecasting out-of-sample

The second of the forecasting problems arises when the forecast period is beyond the within-sample period. The forecast is then necessarily ahead of the last period for which vintage v data is available in Y_t; thus, neither the final vintage data nor the v-th

vintage data are available. This occurs at time $t-(v-1)$ in forecasting $y_{t+s}^{(m)}$, $s=-(v-1), \ldots, 0, 1, \ldots, n$, using $Y_{t-v}^{(v)}$ where $v < m$. In the case of univariate forecasts, the information set available is identical to the within-sample case and so the distinction applies to the VAR and the bivariate version of MSSA and not to the ARMA model and univariate SSA.

In the case of a bivariate SSA model, the data on $Y_{t-v}^{(v)}$ and $Y_{t-m}^{(m)}$ are used jointly to provide forecasts of $y_{t+s}^{(v)}$ and $y_{t+s}^{(m)}$. The bivariate forecasts at time $t-(v-1)$ are conditioned on the information set denoted $\hat{\Omega}_{t-v}^{(v,m)}$, defined as follows:

$$\hat{\Omega}_{t-v}^{(v,m)} = (y_1^{(v)}, \ldots, y_{t-v}^{(v)})'; \{y_1^{(m)}, \ldots, y_{t-m}^{(m)}; (\hat{y}_{t-m+1}^{(m)}, \ldots, \hat{y}_{t-v}^{(m)} \mid \Omega_{t-v}^{(v,m)})\}' \qquad (20)$$

The difference between $\hat{\Omega}_{t-v}^{(v,m)}$ of (24) and $\Omega_{t-v}^{(v,m)}$ of (22), is that the missing values in the final column of Y_t are replaced by their forecasts from the bivariate system.

We assume that the data is published monthly (as in the case of the *IIP*) and that the problem is to construct an out-of-sample forecast of an annual cycle of data, that is month 1 through to month 12; for a particular estimation sample this generates a sequence of out-of-sample forecasts with index i = 1, .., 12. A sequence of annual forecast cycles is then generated over the forecast period, $t+1, \ldots, t+n$. The sequence of *h*-step ahead out-of-sample forecasts comprises those forecasts with the same value of *h*. In each case, for both the bivariate models considered, the final vintage is combined with a different preliminary vintage, $v = 1, \ldots, m-1$. This gives rise to two dimensions of evaluation: the index of the horizon length, *h*, and the index of the preliminary vintage, *v*, which is combined with the final vintage.

For a particular annual cycle, the (conditional on $\hat{\Omega}_{t-v}^{(v,m)}$) out-of-sample forecast data vector for the final vintage, where the underlying model comprises the *v*-th and final vintage, is:

$$\hat{Y}_t^{(v,m)} = (\hat{y}_{t,1}^{(m)}, \hat{y}_{t+1,2}^{(m)}, \ldots, \hat{y}_{t+(12-1),12}^{(m)} \mid \hat{\Omega}_{t-v}^{(v,m)})' \qquad v = 1, \ldots, m-1 \qquad (21)$$

The sequence of *h*-step ahead forecasts is $\{\hat{y}_{t+j,h}^{(m)} \mid \hat{\Omega}_{t+j-v}^{(v,m)}\}_{j=0}^{n}$, for $h = 1, \ldots, 12$.

Forecast evaluation criteria

The forecast performance is evaluated in terms of two criteria: the relative root mean squared error, rrmse, and the accuracy in determining changes in direction. Throughout, *T* is the start of the evaluation period, *m* is the final vintage and *n* is the forecast horizon.

Within-sample

Univariate *h*-step ahead: in the univariate case, the rmse relative to an alternative forecast (indicated by $\tilde{y}$), is given by:

$$\text{rrmse}(h) = \left(\frac{\sum_{j=0}^{n} \left((\hat{y}_{T+j-v,h}^{(m)} \mid \Omega_{T+j-m}^{(m)}) - y_{T+j-v}^{(m)} \right)^2}{\sum_{j=0}^{n} \left((\hat{y}_{T+j-v,h}^{(m)} \mid \Omega_{T+j-m}^{(m)}) - y_{T+j-v}^{(m)} \right)^2} \right)^{1/2} \qquad (22a)$$

This measure is calculated for each $h = 1, 2, \ldots, 11$ and $v = m - h$.

The idea is to evaluate the h-step ahead forecast by the univariate SSA method for $h = 1, 2, \ldots, 11$, and to scale it relative to an alternative method. In the case of the univariate forecasts, the scaling is relative to the ARMA forecast of the same horizon length.

Bivariate h-step ahead: in this case use is made of vintages other than the m-th in forming the forecast. The MSSA and VAR models each use vintage v and m data over the forecast period; alternatively, the preliminary vintage v data implies an $h = m - v$ step head forecast. Scaled by the rmse for the corresponding preliminary vintage, the rrmse is:

$$\text{rrmse}(h) = \left(\frac{\sum_{j=0}^{n} \left((\hat{y}_{T+j-v,h}^{(m)} \mid \Omega_{T+j-v}^{(v,\,m)}) - y_{T+j-v}^{(m)} \right)^2}{\sum_{j=0}^{n} \left(y_{T+j-v}^{(v)} - y_{T+j-v}^{(m)} \right)^2} \right)^{1/2} \qquad (22b)$$

This measure is also evaluated for each $h = 1, 2, \ldots, 11$ and v is obtained from $v = m - h$. In this case it is sensible to scale the numerator rmse by the rmse of the preliminary vintage data of the same horizon length as these are competing candidates for forecasting the final vintage.

Bivariate out-of-sample forecasts

In this case there are h-step forecasts, $h = 1, \ldots, 12$, available for each value of v; conditional on v, the rrmse is given by,

$$\text{rrmse}(h|\, v) = \left(\frac{\sum_{j=0}^{n} \left((\hat{y}_{T+j,h}^{(m)} \mid \hat{\Omega}_{T+j-v}^{(v,\,m)}) - y_{T+j-v}^{(m)} \right)^2}{\sum_{j=0}^{n} \left((\tilde{y}_{T+j,h}^{(m)} \mid \tilde{\Omega}_{T+j-v}^{(m)}) - y_{T+j-v}^{(m)} \right)^2} \right)^{1/2} \qquad (22c)$$

If $\text{rrmse}(h) < 1$, then the SSA forecasts outperform the denominator forecast in a root mean squared error sense; alternatively, if $\text{rrmse}(h) > 1$, then the SSA forecasts are worse than the denominator forecast. The Diebold-Mariano (1995) test, modified by Harvey et. al. (1997)[6], denoted S_1^H, is calculated to check if the difference between the two forecasting procedures is statistically significant.

Predicting turning points

For some purposes it may be more harmful to make a smaller prediction error yet wrongly forecast the direction of change than to make a larger directionally correct error, see Ash et. al., (1997) and Heravi et. al., (2004). Clements and Smith (1999) suggest that model forecasts may be better evaluated by assessing their accuracy in

forecasting the direction of change. Thus, we also compute the percentage of forecasts that correctly predict the direction of the annual change in y. That is we compare the sign of $(y_t^{(m)} - y_{t-12}^{(m)})$ with the sign of the predicted annual change from the preliminary vintage and, as the case may be, the univariate or bivariate SSA model and the ARMA or VAR model (see the appendix for detail of the VAR specification)..

The results in the tables that follow in section 5, present the proportion of times that the forecast is of the right sign together with the significance of the test statistic constructed as follows. Let Z_t , $t = T + 1, \ldots, T + n$, take the value 1 if the forecast correctly predicts the direction of change and 0 otherwise. The Moivre-Laplace limit theorem implies that for large samples, the test statistic $DC = 2(\overline{Z} - 0.5)n^{1/2}$, where $\overline{Z}$ is the mean of Z_t over the forecast period, is aysmptotically distributed as standard normal. When $\overline{Z}$ is significantly larger than 0.5, the forecast is said to have the ability to predict the direction of change; alternatively if $\overline{Z}$ is significantly smaller than 0.5, the forecasts tend to give the wrong direction of change.

Empirical Analysis

The data on IIP: structure and summary statistics

The IIP is published in Table 7 of the monthly Digest of statistics (MDS) as "output of the total production industries". The overall data period for this study comprises 423 monthly observations for 1972:1 to 2007:3 on 12 vintages of seasonally adjusted IIP. The first vintage, which is published one month after the latest month of published data, refers to the first publication in the MDS. The second vintage refers to the next published figure and so on. For this study we take the 12^{th} vintage as the final vintage (m), so that there are 12 vintages of data on the same generic variable.

Descriptive statistics

The revisions to the data are defined as $r_t^v \equiv y_t^{(v)} - y_t^{(m)}$, with % revisions given by $pr_t^{(v)} \equiv 100(y_t^{(v)} - y_t^{(m)})/y_t^{(v)}$. To give an idea of the structure of the revisions, some summary statistics for $pr_t^{(v)}$ are presented in Table 1.

'Table 1 here'

The mean revision is negative for all vintages, showing that the estimates systematically under-predict the final vintage. The standard deviations also indicate that the early vintages are relatively noisy predictions of the final vintage; for example, the standard deviation of the first revision is twice the standard deviation of the 8^{th} revision. This decline in the standard deviations as v increases is an implication of the 'noise' characterisation of revisions, which is considered further below.

Time series characteristics of the data

To illustrate the time series characteristics of the data on the *IIP*, we briefly consider some background material. Figure 1 graphs the first and final vintages of *IIP* and Figure 2 shows their first differences. It is evident from Figure 1 that there is a trend in the data, the critical question in this context being whether the trend is better characterised as a stochastic or deterministic trend. There is, however, no evident

trend in the first differenced series, see Figure 2. (Craft, Leybourne and Mills, 1989, found that a measure of industrial production during the industrial revolution that imposed two unit roots worked well, but the data characteristics of that period of growth do not translate to the 20th/21st century data period considered here; moreover, Newbold and Agiakloglou (1991) found that the application of some standard unit root tests did not suggest two unit roots for that historical period.)

'Figures 1 and 2 about here'

To investigate further, Table 2 reports the psuedo-t test statistics due to Dickey and Fuller, DF, (Fuller 1996), Elliott, Rothenberg and Stock, ERS, (1996) and Shin and Fuller, SF, (1998). The ADF and ERS statistics use an autoregressive model, whereas the SF procedure uses exact ML estimation of the unconditional likelihood function with an ARMA model, or ARIMA model. (Note that the ERS-type tests are not sensible when the alternative is not local-to-unity as in the case of the test for a second unit root.)

The sequential testing procedure follows Dickey and Pantula (1987), with lag lengths selected alternately by AIC and BIC. Note that the test statistics for two unit roots indicate unambiguous rejection of the null hypothesis; the sample values being well in excess, in absolute value, of the respective 1% critical values. In the case of the tests for two unit roots, the alternative is a single unit root with an intercept, that is a drifted random walk; neither an explosive alternative nor nonstationarity about a linear trend are suggested by the data and, indeed, the qualitative results are as strong for rejection if this alternative is entertained. In the case of the test statistics for a single unit root, the null hypothesis is not rejected, the alternative in this case being stationarity around a linear trend. Thus, we conclude that we cannot reject the null hypothesis of a single unit root for any of the vintages of *IIP*. The parametric forecasting models for *IIP*, therefore, imposed a unit root.

'Table 2 about here'

News or noise?

A longstanding interest in assessing the characteristics of data revisions is whether they can be viewed as 'news' or 'noise'. To ensure stationarity the revisions are analysed in first differences, thus first define $z_t^{(v)} \equiv \Delta y_t^{(v)}$, where Δ is the first difference operator, so that the revisions are defined as $r_{z,t}^{v} \equiv z_t^{(v)} - z_t^{(m)}$. If the revisions are news, then the *v*-th revision $r_{z,t}^{v}$ is not forecastable given $z_t^{(v)}$, since all information has been efficiently incoprorated into $z_t^{(v)}$; as a result, $r_{z,t}^{v}$ should be orthogonal to $z_t^{(v)}$. If the revisions are noise, then $r_{z,t}^{v}$ is orthogonal to $z_t^{(m)}$, simply being a measurement error relative to $z_t^{(m)}$. These hypotheses are referred to respectively as the efficient forecasts hypothesis (EFH) and the measurement error hypothesis (MEH). The standard framework testing these hypotheses is:

'News': EFH

$$z_t^{(m)} = \beta_1^* + \beta_2^* z_t^{(v)} + \xi_t^* \qquad v = 1, \ldots, m-1$$

$$\Rightarrow$$

$$r_{z,t}^{v} = \beta_1 + \beta_2 z_t^{(v)} + \xi_t$$

where $\beta_1 = -\beta_1^*$, $\beta_1 = 1 - \beta_2^*$, $\xi_t = -\xi_t^*$ and H_0: $\beta_1 = \beta_2 = 0$.

'Noise': MEH

$$z_t^{(v)} = \alpha_1 + \alpha_2^* z_t^{(m)} + \zeta_t \qquad v = 1, \ldots, m-1$$

$$\Rightarrow$$

$$r_{z,t}^{v} = \alpha_1 + \alpha_2 z_t^{(m)} + \zeta_t$$

where $\alpha_2 = \alpha_2^* - 1$ and H_0: $\alpha_1 = \alpha_2 = 0$.

In each case the alternative hypothesis is the negation of the null hypothesis. Estimation of these regressions should allow for serial correlation in the regression disturbances, which is induced by the implict forecasting horizon. As Brown and Maital (1981) and Mork (1987) note, the serial correlation is in the form of a moving average error process of order *j*, where *j* is the elapse time between compilation of the the preliminary and final vintages; see Patterson and Heravi (1992) for details in a revisions context. The results for testing the EFH and MEH for *IIP* are reported in Table 3, where the table entries are the test statistic (an F test on the appropriate H_0) with the corresponding p-value shown beneath the test statistic.

'Table 3 about here'

The results are unambiguous. First, the EFH is uniformly rejected for all vintages except the last, indicating that the revisions are in principle forecastable; it is only with the very last vintage that the results indicate, at least in a linear context, that the final revision not forecastable and is 'news'. Second, the MEH is not rejected for any of the vintages. Taken together these results are promising in the sense of offering improvments on taking the preliminary vinatges alone as implicit forecast of the final vintage.

Assessing nonlinear dependence

If a time series is the output of a non-deterministic, linear dynamic system, then measures of linear association such as the standard correlation coefficient, ρ, can be used for measuring dependencies, for example, between two time series. However, if the data are outputs from a nonlinear process, a measure should be used that has the ability to capture the nonlinearities of series. Granger and Lin (1994) defined a standardised measure based on the mutual information between the series (see also Darbellay, 1998, and Granger et. al., 2004), referred to as the global correlation coefficient, defined as:

$$\lambda(X, Y) = [1 - exp\{-2I(X, Y)\}]^{1/2} \qquad (23)$$

$I(X,Y)$ is the mutual information of two continuous random variables X and Y given by:

$$I(X, Y) = \int_x \int_y f(x,y) \ln \frac{f(x,y)}{f(x)f(y)} dxdy \qquad (24)$$

where $f(x, y)$ is the joint probability density function (pdf) of X and Y, and $f(x)$ and $f(y)$ are the marginal pdfs of X and Y, respectively.

The coefficient $\lambda(X,Y)$ varies between 0 and 1 and measures the dependence, whether linear or nonlinear, between X and Y. As to the former, an important property of $\lambda(X, Y)$ follows if X and Y are bivariate normal with correlation coefficient ρ; in that case $\lambda(X,Y) = |\rho|$, see Granger and Lin (1994). The null hypothesis of no dependence corresponds to $\lambda(X,Y) = 0$ and the alternative hypothesis to $\lambda(X,Y) > 0$; the critical values, obtained by simulation, are provided in Granger et. al., (2004).

The main difficulty in estimating the mutual information and global information measures from empirical data lies in the fact that the relevant pdfs are unknown. One solution to this problem is to approximate the densities by means of histograms, but a histogram with an arbitrary bin size would not be the best way, because it can cause underestimation or overestimation of the empirical mutual information. To overcome this problem we use the method of marginal equiquantization, which establishes a homogeneous partition of the space into equiprobable cells. The partition is stopped when local independence is found between cells using the chi-square test for that purpose, see Darbellay (1998).

Table 4 shows the results of measuring global correlation and linear correlation between the revisions $r_t^{(v)}$ and the preliminary vintage, $y_t^{(v)}$, and between $r_t^{(v)}$ and the final vintage, $y_t^{(m)}$. All table entries are significant at the 5% level as indicated by *. These results strongly suggest that there is a significant degree of nonlinear dependence among the series; note that, uniformly across the vintages, global correlation exceeds linear correlation, suggesting that nonlinearity is a feature of the data. This observation

'Table 4 here'

Multivariate non-normality

The SSA method does not assume linearity or normality of the data either in finite samples or asymptotically. To assess the normality aspect of our data set, we used the Doornik-Hansen Omnibus, DHO, multivariate normality test[7], see Doornik and Hansen (1994). This is a multivariate version of Shenton and Bowman's (1977) univariate omnibus test for normality, based on transformed skewness and kurtosis coefficients.

The DHO(p) test statistic is approximately distributed as $\chi^2(2p)$, where p is the number of series being considered jointly. Two cases of interest reported here. In the

first case, the test is applied pairwise between $r_t^{(v)}$ and $y_t^{(v)}$, and also pairwise between $r_t^{(v)}$ and $y_t^{(m)}$ (as in Table 2), with the results reported in Table 3. All table entries exceed the 99% quantile, indicating that there is strong evidence of non-normality.

The second case is when all 12 vintages are considered together, so that the null hypothesis corresponds to multivariate normality with $p = 12$. In this case DHO(12) is distributed as $\chi^2(24)$, with a 95% quantile of $\chi^2_{0.95}(24) = 36.42$; the sample value of the test statistic is 3,850, leading to an unequivocal rejection of the null hypothesis of normality. Thus, taken together, the measures of nonlinear dependence and tests for (non)normality show that it would be unwise to adopt a linear forecasting framework, whether univariate or multivariate, for IIP.

'Table 5 here'

Empirical selection of the parameters

As in the identification stage case of parameteric models, such as ARMA and VAR models, it is necessary to select some parameters to determine a practical SSA or MSSA model for forecasting. Considerable guidance is provided by the examples in Golyandina et. al., (2001) and, in addition, Hossian and Thomakos (2010) provide some examples using economic and financial time series. Two essential tools in this part of the process are a graph of the singular spectrum and the w-correlations $\rho_{i,j}^w$. To illustrate the choice of L, the singular spectrum for the first vintage is shown in Figure 3. The window length is $L = 13$ and it is clear that this value is sufficient to enable the separation required in the second stage.

'Figure 3 here'

The next step is to calculate $\rho_{i,j}^w$ for pairs of reconstructed components corresponding to pairs of the $L = 13$ eigenvalues in the SVD and Figure 4 shows $\rho_{i,j}^w$ for the $L = 13$ reconstructed components of $Y^{(1)}$ in a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1.

'Figure 4 here'

It is apparent that separating the eigentriples into three groups comprising 1-5, 6 and 7, and 8-13, leads to a decomposition of the trajectory matrix into three almost orthogonal blocks. Therefore, in this case the system noise can be associated either with components 8-13 or with components 6-13. The two choices are assessed by the cumulative *w*-correlations.

As noted in section 3, the existence of structure in the series is indicated by local minima and maxima in the w-correlations. This is shown in Figure 5 for the first vintage series. The first minimum occurs at point $k = 5$, thus we choose the

reconstructed series from the first 5 eigentriples as signal and the rest, components 6–13, as noise.

'Figure 5 here'

Forecast performance
Throughout this section, T = December, 2000; n = 63 (governed by data availability) and m = 12, determined by the data measurement process.

Within-sample performance
The results for the within-sample, univariate forecasts are presented in Table 4. (Throughout the tables $^{+}$, * and ** indicate significance at the 10%, 5% and 1% levels, respectively.) It is evident that the SSA forecasts are superior to the ARMA[8] forecasts at all but the shortest forecast horizons, and the differences are significant at the 1% level by the S_1^H test statistic. As expected the rmse of both methods increases as h increases. There is a uniform rmse advantage to SSA over ARMA, which increases with h; for example, the rmse ratio is 83% for h = 11, to 86% for h = 6 and 97% for h = 1.

'Table 6 here'

The direction of change results are presented in Table 7. The SSA forecasts are better than the ARMA forecasts at all but the shortest horizon (h = 1), with 62% accuracy for h = 11 increasing to 74% for h = 2 compared to 40% to 60%, respectively, for the ARMA model. The results of the *DC* test indicate significance at at the 10% level for h = 11 and at the 1% level for h = 10, …, 1.

'Table 7 here'

Continuing the within-sample comparison, the MSSA forecasts are compared to the VAR forecasts and the appropriate preliminary vintage in Table 8 for rmse and in Table 9 for the direction of change. (The appendix details the nature of the VAR and the lag selection methods.) It is clear from Table 8 that whilst the VAR cannot improve upon the preliminary vintage, the MSSA is significantly better than both alternatives at all horizons, even very short horizons. Note, in particular, that MSSA is able to improve upon the preliminary vintage (and, comparing Tables 6 and 8, there is a substantial gain over univariate SSA). For example, the rmse gain relative to the preliminary vintage varies from 11% for h = 11 through to 4% for h = 1; in turn these represent gains of 39% and 24% relative to the VAR. The gains are significant at the 1% level using the S_1^H test except for h = 1, 2, when they are significant at the 5% level.

'Table 8 here'

The second half of the table shows that bias reductions for MSSA are between 25% and 29%, relative to the preliminary vintage, whereas whilst there are some reductions for the VAR these are not as great as for MSSA and there are also cases, h = 7 – 11,

when the bias increases. (The bias ratio is the ratio of the absolute bias for VAR or MSSA relative to the absolute bias of the preliminary vintage.)

The force of these results is emphasised in predicting the direction of change, as shown in Table 9; it is clear that MSSA uniformly dominates the VAR model and only fails to dominate the preliminary vintage in two cases when the forecasting horizon is short. Thus, in general, for the 'backward' looking task of completing the vector of observations for the final vintage, MSSA is to be preferred both to the VAR and to the preliminary vintage.

'Table 9 here'

Out-of-sample performance

Turning to the out-of-sample forecasts, the results are reported in Tables 10 and 11. First note that, by definition, the preliminary vintages are now not available out of sample. Table 10, therefore, compares the MSSA forecasts directly with the VAR forecasts. The results are presented by the length of the forecast horizon, h, and the vintage that is combined with the final vintage in the multivariate method (either VAR or MSSA), see Table 10. As expected, ceteris paribus, the rmse decreases as the forecast horizon decreases, and decreases as the preliminary vintage (which is used in the bivariate models) approaches the final vintage. In all cases, the MSSA method outperforms the VAR model.

'Tables 10 and 11 here'

For example, when vintage 1 data is combined with the final vintage data, (see the first three rows of the table), MSSA gives a reduction of 19% for $h = 1$ and 16% for $h = 11$. This scale of reduction is typical throughout as v increases; for example, there are substantial gains even when the penultimate and final vintages are combined in the MSSA method, the exception being when the forecast is at its shortest horizon, $h = 1$. Throughout the reductions are generally significant at the 1% level; for example, all comparisons using $v = 1$ and $m = 12$ are significant and 9 are significant using $v = 11$ and $m = 12$.

The dominance of MSSA is confirmed by the direction of change test results, see Table 11. Overall, the VAR departs little from the random result of a 50% rate of detecting the direction of change, only improving slightly as the forecasting horizon shortens. On the other hand, the MSSA method often achieves a significant difference (at levels varying from 10% to 1%) from the random case, typically with 60-70% of the changes in direction correctly identified, increasing to 80% or more as the forecasting horizon shortens and the combining data vintage approaches the final vintage.

Concluding remarks

The need for timely data to enable an assessment of current economic conditions is ever present and in recent years has been heightened by the delegation of some policy tasks to central banks, such as the Bank of England in the UK. In principle it is

possible to wait until data is finalised and so provides, in the view of the data agency, the best possible guide to the underlying economic concept, in practice there is a continuing demand for the publication of key economic time series at the earliest practicable date. This gives rise to what are referred to as vintages of data on the same generic variable, the vintage number referring to the stage in the revisions process. The fact that there are different stages in the collection and subsequent publication of different vintages of the same generic variable has been recognised in the concept of the data measurement process. This stylised process applies to a large number of key data series.

Because there is a time lapse between publication of the preliminary vintages and final vintage of data for revised real-time data, it seems natural to use the most recently available vintage of data as the best available estimate or forecast of the final vintage. Indeed, such a practice applies widely to components of the National Income and Product Accounts. However, in effect, the preliminary vintage is being taken as a forecast of the final vintage and from this perspective, it is not a forgone conclusion that it is the best forecast.

The issues raised in early contributions by Mankiw et. al., (1984), Mankiw and Shapiro (1986) and Mork (1987), in the news versus noise debate, continue to be relevant see Aruoba (2008). In a policy context, the key question is whether it is possible to improve upon the preliminary vintages in order to provide more accurate forecasts of subsequent vintages? If the preliminary vintages are (fully) efficient forecasts as elucidated by Mankiw and Shapiro (op. cit.,), then the answer is no; however, there is evidence in the case of the *IIP*, both using the standard linear testing framework and the nonlinear framework of SSA, that the preliminary vintages can be improved upon. Typically, attempts to 'beat' the preliminary vintages involve parametric models, which involve a number of assumptions, for example they necessarily impose a structure and usually involve a dsitributional assumption. In a more general approach, this study has developed the method of Multivariate Singular Spectrum Analysis as a nonparametric alternative to typical parametric models such as ARMA and VAR models.

SSA is flexible enough to accommodate data generating processes that are nonlinear, non-Gaussian and nonstationary, in contrast to standard models that embody one or more of these assumptions. Our empirical results for the UK *IIP* strongly indicate that the forecasts using a bivariate version of MSSA are better than using a preliminary vintage as an estimate of the final vintage or basing forecasts on ARMA or VAR models. A consequence of this finding is that the preliminary vintages of the *IIP* are not fully efficient forecasts of the final vintage.

One of the merits of the nonparametric approach of MSSA is its flexibility. We anticipate, therefore, that it would perform well across the range of series for which data revisions are standard practice and that the gains shown here are likely to occur in other applications. Of course, such a statement has to be supported by empirical results and, as part of further research on this topic, we are considering the applicability of MSSA to a number of series other than the *IIP* and we hope to report

our results at a later stage. We are also undertaking an exercise comparing developments of the state space approach, see, for example, Patterson (1995b, 1995c,) and Jacobs and van Norden (2010), with MSSA. An alternative modelling framework is that of a cointegrated set of revisions as suggested by Patterson and Heravi (2004a) and further developed by Hecq and Mazzi (2008). The most ambitious application in this context would be to link multiple revisions and multiple variables, thus exploiting possible data revision relations between variables, for example, consumption and income or GDP and industrial production. Such an extension necessarily leads to an increase in the dimension of the system being considered, but offers gains by exploiting the correlations between revisions and between variables; the advantage of MSSA over both the state space approach and (standard) cointegration modelling, is that it is able to exploit nonlinear correlations automatically if they are present in the data and its extension to high-dimensional systems is straightforward both theoretically and practically.

Endnotes

[1]For an excellent reference source for articles and working papers on data revisions see, for example, Croushore (2008). Recent work includes Knetsch and Reimers (2009) and Ghosh (2009) and earlier references of interest include Conrad and Corrado (1979), Howrey (1984), Harvey et. al., (1983) and Patterson (1995b), Patterson (2002) and Garratt and Vahey (2006); on forecasting with data subject to revision see, for example, Mariano and Tanizaki (1995), Patterson (1995c), Ghosh and Lien (1997), Stark and Croushore (2002), Busetti (2006) and Croushore (2006).

[2]The SSA method, which is widely applied in other sciences, incorporates elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamic systems and signal processing. The genesis of SSA is usually associated with Broomhead and King (1986), and a thorough description of the theoretical and practical foundations of the SSA method, with many examples, can be found in Danilov and Zhigljavsky (1997) and Golyandina et. al. (2001). For a more elementary treatment see Elsner and Tsonis (1996); and for a comparison between SSA and other techniques for forecasting time series, see Hassani (2007), Hassani and Zhigljavsky (2008). For an example of the flexibility of SSA in other areas see Aldrich and Barkhuizen (2003) and Salgado and Alonso (2006). For extensions to the multivariate case, several forecasting procedures and change-point detection problems see Golyandina et. al., (2001) and Moskvina and Zhigljavsky (2003). Also, Hassani et. al., (2009) show that the results obtained by the SSA method are more accurate than those obtained by ARIMA and GARCH models.

[3]The terms 'within-sample' and 'out-of-sample' are intended to distinguish two quite distinct forecasting and practical problems best seen in terms of the data matrix that is constructed from a typical DMP involving revised data, see equation (1): within-sample means 'filling in' a data matrix; out-of-sample means extending the data matrix. There is another distinction that is relevant for univariate models, which involves the sample used for estimation. This is necessarily within the sample as defined here, but will not coincide with it except for the first vintage of data.

[4]The test for chaotic dynamics was implemented in the form suggested by Saïda (2007), who also provides a MATLAB routine for implementation. The test is based on the Lyapunov exponent of a series, with a positive value indicating divergence and chaos and a negative value indicating stability. For example, in the case of the final vintage we obtained a Lyapunov exponent of –0.044, with rejection of chaotic dynamics; similar results were obtained for the other vintages; further details of the test and results can be provided on request.

[5]As noted by Patterson and Heravi (2004b), there are other ways of organising data subject to revision. They introduce the distinction between definitions of vintage based on revision-stage and, alternatively, real-time. The analysis here follows the literature on the empirical analysis of data revisions that defines vintage as relating to the stage of revision of the DMP, see, for example, Howrey (1978), Mork (1987) and, Akritidis (2003) and Richardson (2002, 2003). The revision-stage definition of a vintage relates to the (approximate) consistency of the successive and progressive stages of the measurement process. For example, in the context of GDP, there are flash estimates, 75-day estimates and so on through to the (conditionally) final estimate; each represents a different stage in the DMP, largely as a result of incremental information that becomes available to the data collection agency. Another use of the word vintage in this general context is to capture the idea of a real-time vintage, see, for example, Croushore and Stark (2001), rather than a revision-stage vintage. A real-time vintage is simply the run of data available, in a particular publication, at a particular point in time. Overall, the data is the same, but it is organised differently.

[6]This test takes the following form: denote the forecast errors for two forecasting methods, i = 1, 2, for T + h = j as $e_{i,j}$. Typically, forecast evaluation takes a function of these errors, such as the squared error or root squared error, over T or h, to evaluate their relative performance; this function is generically denoted $f(e_{i,j})$ and the difference $d_j = f(e_{1,j}) - f(e_{2,j})$ is considered. Diebold and Mariano (1995) suggested the test statistic $S_1 = \hat{V}(\bar{d})^{-1/2}\bar{d}$, where $\bar{d} = \sum_{j=1}^{n} d_j$ and $\hat{V}(\bar{d})$ is an estimator of the variance of $\bar{d}$ based on the autocovariances of d_j. Harvey et. al., (1997, equation 9) modify this statistic using an approximately unbiased estimator of the variance of $\bar{d}$. The resulting test statistic is $S_1^H = C \times S_1$, where $C = [\{T + 1 - 2h + T^{-1}h(h+1)\}/T]^{1/2}$; additionally they suggest using critical values from the 't' distribution, with (T – 1) degrees of freedom. Our forceast evaluation is, as in Harvey et. al., (ibid), based on $f(e_{i,j})$ defined as the squared error.

[7]The DHO statistic is *DHO(p)* = $Z_1^{'}Z_1 + Z_2^{'}Z_2$ where p is the number of series, $Z_1^{'} = (z_{11}, \ldots, z_{1p})$ and $Z_2^{'} = (z_{21}, \ldots, z_{2p})$; z_{1i} is a transformation of the standard univariate skewness coefficient $\sqrt{b_1}$, applied to the *i*-th series, due to D'Agostino (1970), and z_{2i} is a transformation of the standard kurtosis coefficient $\sqrt{b_2}$, from a gamma distribution to χ^2 and then to standard normal, applied to the *i*-th series; see Doornik and Hansen (ibid, Appendix) for details.

[8]As noted, preliminary tests strongly suggested a unit root in each vintage of the *IIP*, confirming the results of Patterson (2002); a unit root was, therefore, imposed in the ARMA and VAR modelling, with the remaining orders selected alternately by AIC/BIC with BIC producing the better forecasts; a constant was included to capture the drift in the data series. See aslo the appendix

Appendix: parametric models, selection procedure

In order that the parametric models were considered equally with the SSA and MSSA models, careful consideration was given to the selection of the lag lengths, since as noted by Lütkepohl (2005, section 4.3) this is likely to be a critical issue in forecasting performance. Lütkepohl (op. cit.,) suggested that the lag selection criterion should be one that reflects the purpose of the model, which, in this case, is to provide a good model for forecasting, and he suggests AIC (which differs from the final prediction error, FPE only by a term of order $O(T^{-2})$). In addition, Raffalovich et. al., (2008) have suggested that BIC performs well (and better than AIC) in bivariate VARs, which are used here as the parameteric specification. We selected the lag length alternately by AIC and BIC and found that selection by BIC resulted in better performance and, therefore, it is these that are reported in the text.

The comparison VAR is specified to put it on an equal footing, in terms of information dating, with MSSA. For example, the within-sample forecasts are based on the information set $\Omega_{t-v}^{(v,m)} = (y_1^{(v)}, \ldots, y_{t-v}^{(v)})'; (y_1^{(m)}, \ldots, y_{t-m}^{(m)})'$, see equation (18) in the text. To illustrate the VAR specification, consider the simplest case with lag length = 1, then the bivariate VAR is given by:

$$\begin{pmatrix} z_{t-v}^{(v)} \\ z_{t-m}^{(m)} \end{pmatrix} = \begin{pmatrix} \mu_1^v \\ \mu_2^v \end{pmatrix} + \begin{bmatrix} \alpha_{11}^v & \alpha_{12}^v \\ \alpha_{21}^v & \alpha_{22}^v \end{bmatrix} \begin{pmatrix} z_{t-v-1}^{(v)} \\ z_{t-m-1}^{(m)} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t-v}^v \\ \varepsilon_{t-m}^m \end{pmatrix} \quad v = 1, \ldots, m-1 \qquad \text{(A2.1)}$$

Where $z_{t-v}^{(v)} \equiv \Delta y_{t-v}^{(v)}$, Δ is the first difference operator and the inclusion of a constant allows for drift. It is more informative to write this specification making the lead transparent, so that the time index is shifted by m periods, which gives:

$$\begin{pmatrix} z_{t+s}^{(v)} \\ z_t^{(m)} \end{pmatrix} = \begin{pmatrix} \mu_1^v \\ \mu_2^v \end{pmatrix} + \begin{bmatrix} \alpha_{11}^v & \alpha_{12}^v \\ \alpha_{21}^v & \alpha_{22}^v \end{bmatrix} \begin{pmatrix} z_{t+s-1}^{(v)} \\ z_{t-1}^{(m)} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t+s}^v \\ \varepsilon_t^m \end{pmatrix}$$

where s = m – v. For example, for s = 1 and hence v = 11, then $z_t^{(m)}$ and $z_{t+s-1}^{(v)}$ are contemporaneous; that is for the one-step ahead within-sample forecast, the preliminary vintage in the same row is (linearly) combined with the preceding value of the final vintage, that is $z_{t-1}^{(m)}$. In contrast an ARMA model just uses the latter.

REFERENCES

Akritidis, L. (2003) Revisions to quarterly GDP growth and expenditure components, *Economic Trends*, pp. 69–85.

Aldrich, C. & Barkhuizen, M. (2003) Process system identification strategies based on the use of singular spectrum analysis, *Minerals Engineering*, 16, pp. 815–826.

Aruoba, S. B. (2008) Data Revisions are not well-behaved, *Journal of Money, Credit and Banking*, 40, pp. 319–340

Ash, J. C. K. Smyth, D. J. & Heravi, S. (1997) The accuracy of OECD forecasts for Japan, *Pacific Economic Review*, 2, pp. 25–44.

Ashley, J. Driver, R. Hayes, S. & Jeffery, C. (2005), Dealing with Data Uncertainty. *Bank of England Quarterly Bulletin (Spring)*, pp. 23–29.

Broomhead, D. S. & King, G. P. (1986) Extracting qualitative dynamics from experimental data, *Physica D*, 20, pp. 217–236.

Brown, B. W. & and Maital, S. (1981) What do economists know? An empirical study of exeperts' exepectations, *Econometrica*, 49, pp. 491-504.

Busetti, F. (2006) Preliminary data and econometric forecasting: an application with the Bank of Italy quarterly model, *Journal of Forecasting*, 25, pp. 1–23.

Castle, J. & Ellis, C. (2002) Building a real-time database for GDP(E). *Bank of England Quarterly Bulletin* (Spring), pp. 42–49.

Clements, M. P. & Smith, J. (1999) A Monte Carlo investigation of forecasting performance of empirical SETAR model, *Journal of Applied Econometrics*, 14, pp. 123–141.

Clements, M. K. & Galvão, A. B. (2009) Real-time forecasting with autoregressive models in the presence of data revisions, University of Warwick, ESG, working paper.

Conrad, W. & Corrado, C. (1979) Application of the kalman filter to revisions in monthly retail sales estimates, *Journal of Economic Dynamics and Control*, pp. 177–98.

Crafts, N. F. R. Leybourne, S. J. & Mills, T. C. (1989) Trends and cycles in British industrial production, 1700-1913, *Journal of the Royal Statistical Society*, Series A, 152, 43-60.

Croushore, D. & Stark, T. (2001) A real-time data set for macroeconomists, *Journal of Econometrics*, 105, pp. 111–130.

Croushore, D. (2006) Forecasting with real-time macroeconomic data, in: Graham Elliott, Clive W.J. Granger, and Allan Timmermann, (eds.), *Handbook of Economic Forecasting* (Amsterdam, North-Holland).

Croushore, D. (2008) Frontiers of real-time data analysis. *Federal Reserve Bank of Philadelphia, Working Paper* No. 08-04.

Croushore, D. (2009) http://oncampus.richmond.edu/~dcrousho/docs/realtime_lit.pdf.

Cunningham, A. Eklund, J. Jeffery, C. Kapetanios, G. & Labhard. V. (2007) A state space approach to extracting the signal from uncertain data, Bank of England Working Paper No. 336.

D'Agostino, R. B. (1970) Transformation to normality of the null distribution of g1, *Biometrika*, 57, pp. 679–681.

Danilov, D. & Zhigljavsky, A. (1997) Principal components of time series: the 'caterpillar' method, *University of St. Petersburg, St. Petersburg*, (in Russian).

Darbellay, G. (1998) An adaptive histogram estimator for mutual information, *UTIA Research Report* 1889, Academy of Sciences, Prague.

Dickey, D. A. & S. G. Pantula. (1987) Determining the order of differencing in autoregressive processes, *Journal of Business & Economics Statistics*, 5, pp. 455-461

Diebold, F. X. & Mariano, R. S. (1995) Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, pp. 253–63.

Doornik, J. A. & Hansen, H. (1994) An omnibus test for univariate and multivariate normality, *Discussion Paper*, Nuffield College, Oxford, UK.

Elliott, G. Rothenberg, T. J. & Stock, J. H. (1996) Efficient tests for an autoregressive root, *Econometrica*, 64, pp. 813-836.

Elsner, J. B. & Tsonis, A. A. (1996) *Singular spectral analysis, a new tool in time series analysis*, (Plenum Press, New York and London)

Fixler, D. J. & Nalewaik, J. J. (2006) News, noise, and estimates of the 'True' unobserved state of the economy. US. Bureau of Economic Analysis Working paper, January.

Fuller, W. (1996) *An Introduction to Statistical Time Series*, 2nd edition, New York: John Wiley.

Garratt, A. & Vahey, S. P. (2006) UK real-time macro data characteristics, *Economic Journal*, 116, pp. F119-F135.

Ghosh, S. & Lien, D. (1997) Forecasting with preliminary data, *Journal of Forecasting*, 16, pp. 463–473.

Ghosh, S. (2009) United States trade balance announcements: the nature of its data revisions, *Oxford Bulletin of Economics and Statistics*, 71, pp. 371 – 383.

Golyandina, N. & Stepanov, D. (2005) SSA-based approaches to analysis and forecast of multidimensional time series, *Proceedings of the Fifth Workshop on Simulation*, pp. 293–296.

Golyandina, N. Nekrutkin, V. & Zhigljavsky, A. (2001) *Analysis of Time Series Structure: SSA and Related Techniques*, (Chapman & Hall, London).

Granger, C. & Lin, J. L. (1994) Using the mutual information coefficient to identify lags in nonlinear models, *Journal of Time Series Analysis*, 15, pp. 371–384.

Granger, G. Maasoumi, E. & Racine, J. (2004) A dependence metric for possibly nonlinear processes, *Journal of Time Series Analysis*, 25, pp. 649–669.

Harvey, A. C. McKenzie, C. R. Blake, D. P. C. & Desai, M. J. (1983) Irregular data revisions, in Arnold Zellner, ed., *Applied Time Series Analysis of Economic Data*, pp. 329–347. Washington DC. U.S Department of Commerce, Economic Research Report ER-5.

Harvey, D. Leybourne, S. & Newbold, P. (1997) Testing the equality of prediction mean squared errors, *International Journal of Forecasting*, 13, pp. 281-291.

Hassani, H. (2007) Singular spectrum analysis: methodology and comparison, *Journal of Data Science*, 5, pp. 239–257.

Hassani, H. & Zhigljavsky, A. (2008) Singular spectrum analysis: methodology and application to economics data, *Journal of System Science and Complexity*, 22, 372–394.

Hassani, H. Dionisio, A. & Ghodsi, M. (2009) The effect of noise reduction in measuring the linear and nonlinear dependency of financial markets, *Journal of Nonlinear Analysis: Real World Application*, 11, 392-402.

Hassani, H. Heravi, S. & Zhigljavsky, A. (2009) Forecasting European industrial production with singular spectrum analysis, *International Journal of Forecasting*, 25, pp. 103–118.

Hassani, H. & D. Thomakos (2010) A review on singular spectrum analysis for economic and financial time series, *Statistics and Its Interface*, 3, pp. 377-397.

Hecq, A. & Mazzi, G. J. (2008) Advanced revision analysis for economic time series and their role in improving forecasting accuracy, paper presented to the 5th Eurostat Colloqium on Modern Tools for Business Cycle Analysis.

Heravi, S. M. Osborn, D. R. & Birchenhall, C. R. (2004) Linear versus neural network forecasts for European industrial production series, *International Journal of Forecasting*, 20, pp. 435-446.

Howrey, E. P. (1978) The use of preliminary data in economic forecasting, *The Review of Economics and Statistics*, 60, pp. 193–200.

Howrey, E. P. (1984) Data revision, reconstruction and prediction: an application to

inventory investment, *The Review of Economics and Statistics*, 66, pp. 386–393.

Jacobs, J. P. A. M. & van Norden, S. (2010) Modelling data revisions: measurement error and dynamics of "true" values, *Journal of Econometrics* (forthcoming).

Kawagoe, M. (2007) How is economic growth rate revised afterwards? Real-time data analysis of real GDP growth rate, *Quarterly Journal of National Accounts of Japan*, No.134, Economic and Social Research Institute, Cabinet Office, Japanese Government.

Knetsch, T. A. & Reimers, H. E. (2009) Dealing with benchmark revisions in real-time data: the case of German production and orders statistics, *Oxford Bulletin of Economics and Statistics*, 71, pp. 209-235.

Lütkepohl. H. (2005), *New Introduction To Time Series Analysis*, Springer, Berlin.

Mankiw, N. G. & Shapiro, M. D. (1986) News or Noise? an Analysis of GNP Revisions, *Survey of Current Business*, pp. 20-25.

Mankiw, N. G. Runkle, M. & Shapiro, M. D. (1984), Are preliminary announcements of the money stock rational forecasts?, *Journal of Monetary Economics*, 14 pp. 15 - 27.

Mariano, R. S. & Tanizaki, H. (1995) Prediction of final data with use of preliminary and/or revised data, *Journal of Forecasting*, 14, pp. 351–380.

Mork, R. A. (1987) Ain't behavin: forecast errors and measurement errors in early GNP estimates, *Journal of Business and Economic Statistics*, 5, pp. 165–175.

Moskvina, V. G. & Zhigljavsky, A. (2003) An algorithm based on singular spectrum analysis for change-point detection. *Communication in Statistics - Simulation and Computation*, 32, pp. 319–352.

Newbold, P., & Agiakloglou, C. (1991), Looking for evolving growth rates and cycles in British industrial production, 1700-1913, *Journal of the Royal Statistical Society*, Series A, 154, pp. 341-348.

Nordhaus, W. D. (1987) Forecasting efficiency: concepts and applications, *Review of Economics and Statistics*, pp. 667-674.

Patterson, K. D. (1995a) An integrated model of the data measurement and data generation processes with an application to consumers' expenditure, *Economic Journal* 105, pp. 54–76.

Patterson, K .D. (1995b) A state space approach to forecasting the final vintage of revised data with an application to the index of industrial production, *Journal of Forecasting*, 14, pp. 337–350.

Patterson, K. D. (1995c) Forecasting the final vintage of real personal disposable income: A state space approach, *International Journal of Forecasting*, 11, pp. 395–405.

Patterson, K. D. (2002) Modelling the data measurement process for the index of production, *Journal of the Royal Statistical Society* Series A, 165, pp. 279–296.

Patterson, K. D. & Heravi, S. M. (1992), Efficient forecasts or measurement errors? Some evidence for revisions to United Kingdom GDP growth rates, *The Manchester School*, LX, pp. 249-263.

Patterson, K.D. & Heravi, S.M. (2004a), Revisions to Official Data on U.S. GDP: A Multivariate Assessment of Different Vintages, *Journal of Official Statistics*, 20, 573-602.

Patterson, K. D. & Heravi, S. M. (2004b) Rebasing, common cycles, and some practical implications of modelling data revisions, *Journal of Official Statistics*, 20, 4, pp. 1–15.

Raffalovich, L. E. Deane, G. D. Armstrong. D. & Tsao, H-S. (2008) Model selection procedures in social research: Monte-Carlo simulation results, *Journal of Applied Statistics*, 35, pp. 1093-1114.

Richardson, C. (2002) Revisions to GDP: a time profile, *Economic Trends*, July, pp. 1–8.

Richardson, C. (2003) Revisions analysis: a time series approach, *Economic Trends*, December, pp. 86–89.

Rivero, S. S. Rodriguez, F. F. & Felix, A. (2005) Testing chaotic dynamics via Lyapunov exponents, *Journal of Applied Econometrics*, 20, pp. 911–930.

Saïda, A. B. (2007) Using the Lyapunov exponent as a practical test for noisy chaos, available at SSRN: http://ssrn.com/abstract=970074.

Salgado, D. R. & Alonso, F. J. (2006) Tool wear detection in turning operations using singular spectrum analysis, *Journal of Materials Processing Technology*, 171, pp. 451–458

Shenton, L. R. & Bowman, K. O. (1977) A bivariate model for the distribution of sqrt(b_1) and b_2, *Journal of the American Statistical Association*, 72, pp. 206–211.

Shin, D. W. & Fuller, W. A. (1998) Unit root tests based on unconditional maximum likelihood estimation for the for the autoregressive moving average, *Journal of Time Series Analysis* 19, pp. 591–599.

Stark, T. & Croushore, D. (2002) Forecasting with a real-time data set for macroeconomists, *Journal of Macroeconomics*, 24, pp. 507−31.

Figure captions

Figure 1. First and final vintages of IIP

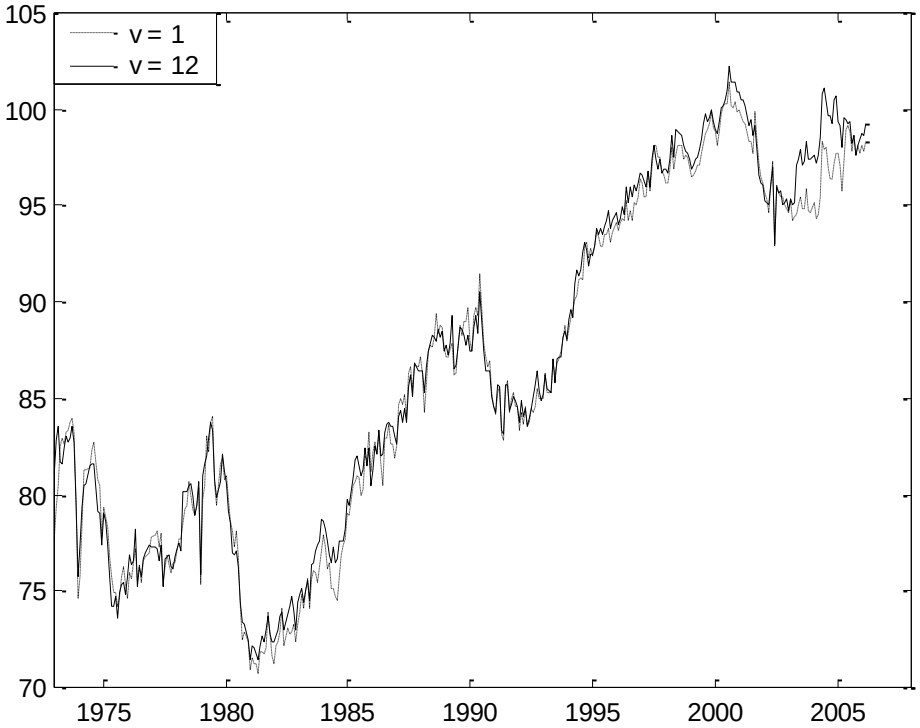

Figure 2. First differences of first and final vintages of IIP

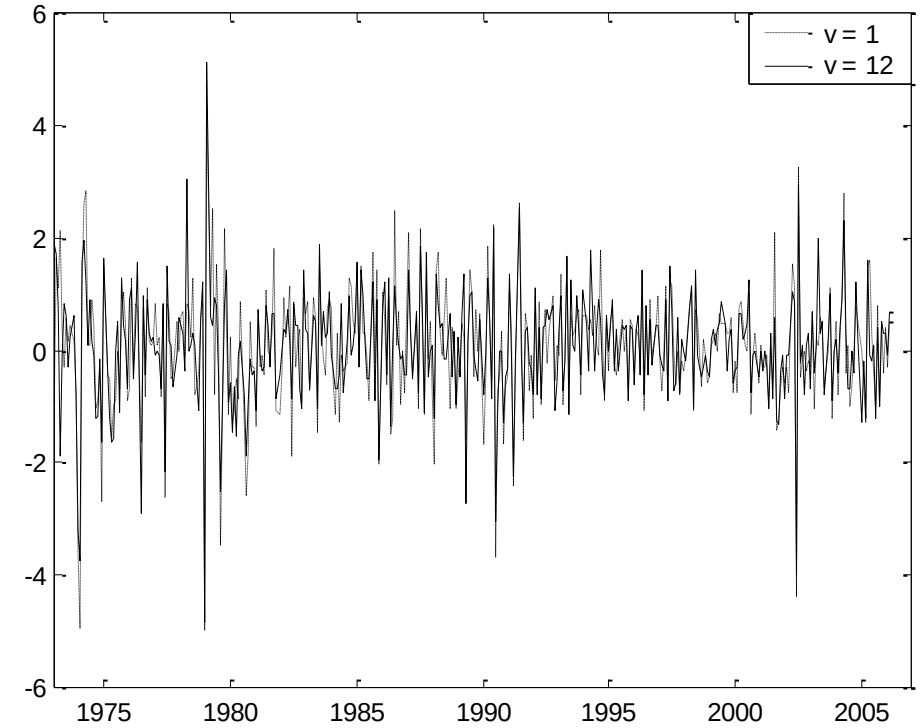

Figure 3. Logarithm of the 13 eigenvalues for $Y^{(1)}$

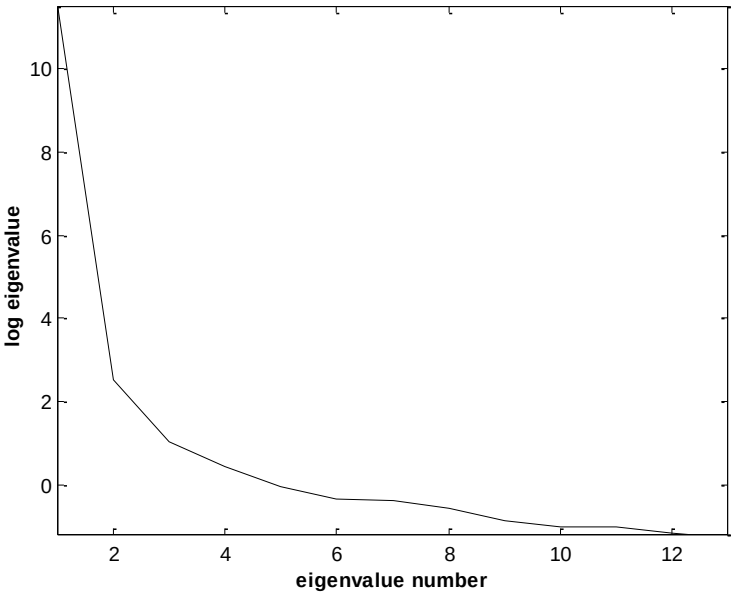

Figure 4. Matrix of w-correlations for the 13 reconstructed components of $Y^{(1)}$

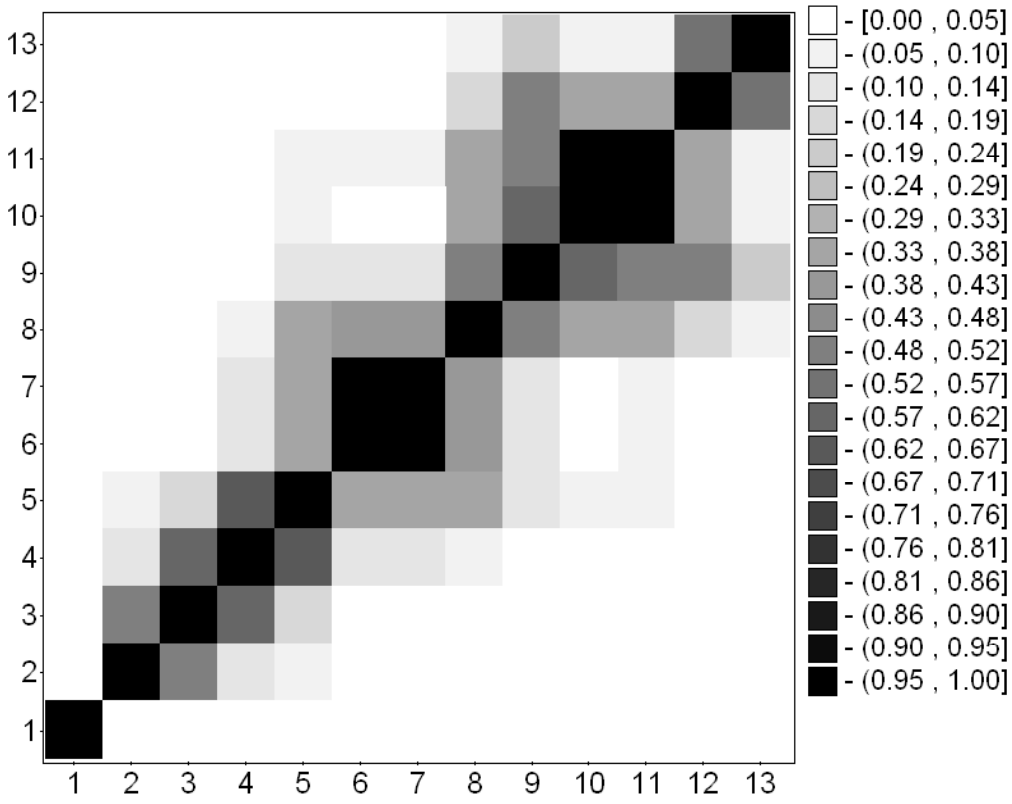

Figure 5. Cumulative w-correlations for the first vintage series

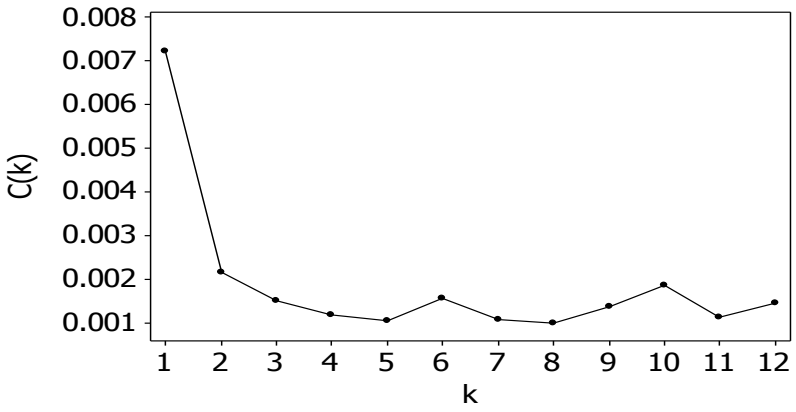

Table 1. Revisions to IIP: summary statistics, $pr_t^{(v)}$

v =	1	2	3	4	5	6	7	8	9	10	11
Mean	-0.46	-0.42	-0.37	-0.34	-0.28	-0.23	-0.17	-0.14	-0.11	-0.08	-0.06
Mean absolute	0.83	0.73	0.65	0.59	0.52	0.45	0.38	0.33	0.27	0.22	0.17
Standard deviation	1.04	0.94	0.85	0.78	0.71	0.65	0.58	0.52	0.45	0.40	0.35

Table 2. Unit root test statistics for *IIP*

Testing for two unit roots												
v =	1	2	3	4	5	6	7	8	9	10	11	final
ADF(AIC)	-16.23	-16.78	-16.23	-23.67	-24.23	-16.57	-16.62	-16.59	-16.77	-16.54	-16.86	-16.35
ADF(BIC)	-22.28	-16.78	-22.61	-23.67	-24.23	-24.00	-23.47	-23.85	-23.85	-23.89	-23.73	-23.32
SF(AIC)	-17.83	-18.02	-16.42	-16.77	-16.66	-16.21	-16.05	-16.09	-16.07	-16.28	-16.03	-16.31
SF(BIC)	-16.03	-15.55	-16.42	-16.77	-16.66	-16.21	-16.05	-16.09	-16.07	-16.28	-16.03	-16.31
Testing for a single root												
ADF(AIC)	-2.38	-2.37	-2.52	-2.37	-2.30	-2.31	-2.12	-2.27	-2.09	-2.28	-2.14	-2.33
ADF(BIC)	-2.32	-2.11	-2.29	-2.22	-2.16	-2.15	-2.21	-2.18	-2.17	-2.15	-2.20	-2.20
ERS(AIC)	-2.17	-2.16	-2.34	-2.22	-2.15	-2.18	-2.00	-2.14	-1.99	-2.17	-2.02	-2.19
ERS(BIC)	-1.89	-1.90	-1.90	-1.97	-1.90	-1.84	-1.86	-1.84	-1.82	-1.80	-1.83	-1.87
SF(AIC)	-2.49	-2.54	-2.49	-2.34	-2.29	-2.26	-2.29	-2.24	-2.24	-2.21	-2.29	-2.31
SF(BIC)	-2.33	-2.37	-2.31	-2.22	-2.13	-2.12	-2.14	-2.11	-2.10	-2.11	-2.14	-2.19

Notes: the table entries are for the appropriate pseudo-t test, with lag lengths selected alternately by AIC and BIC, each case being indicated in parentheses. ADF is the standard augmented DF test, and ERS and SF are the Elliott, Rothenberg and Stock (1996) and maximum likelihood versions of this test, respectively. The tests for two unit roots allow for an intercept whereas the test statistics for a single unit root, allow for a linear trend. The 1%, 5% and 10% critical values for each of the tests are given below:

Critical values	trend			intercept		
	1%	5%	10%	1%	5%	10%
ADF	–3.96	–3.41	–3.13	–3.43	–2.86	–2.57
SF	–3.88	–3.31	–3.02	–3.36	–2.70	–2.40
ERS	–3.48	–2.89	–2.57			

Table 3. News or noise?

v =	1	2	3	4	5	6	7	8	9	10	11
EFH	33.62	48.04	26.82	23.86	19.87	18.58	15.03	9.49	5.87	3.44	1.53
p-val.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.22
MEH	1.09	1.49	0.83	0.69	0.89	1.92	1.14	1.57	1.73	1.70	1.82
p-val.	0.34	0.23	0.43	0.50	0.42	0.15	0.32	0.21	0.18	0.18	0.16

Note: table entries are values of the F statistics for testing H_0, with p-value in the second row.

Table 4. Correlations between the revisions $r_t^{(v)}$ and the preliminary vintage, $y_t^{(v)}$, and between $r_t^{(v)}$ and the final vintage, $y_t^{(m)}$ of *IIP*

Vintage: $r_t^{(v)}$		1	2	3	4	5	6	7	8	9	10	11
λ	Preliminary Vintage	0.434*	0.450*	0.423*	0.422*	0.402*	0.398*	0.405*	0.359*	0.313*	0.297*	0.296*
	Final vintage	0.423*	0.435*	0.403*	0.403*	0.404*	0.368*	0.344*	0.317*	0.303*	0.296*	0.292*
Linear Correlation	Preliminary vintage	0.157*	0.137*	0.128*	0.123*	0.143*	0.145*	0.142*	0.132*	0.132*	0.100*	0.081*
	Final vintage	0.253*	0.225*	0.209*	0.198*	0.211*	0.207*	0.198*	0.183*	0.176*	0.139*	0.114*

Note: * indicates significance at the 5% level.

Table 5. Results of the *DHO* test for pairwise normality: $r_t^{(v)}$ and the preliminary vintage, $y_t^{(v)}$; and $r_t^{(v)}$ and the final vintage, $y_t^{(m)}$ of *IIP*

Vintage: $r_t^{(v)}$	1	2	3	4	5	6	7	8	9	10	11
Preliminary Vintage	102.9**	116.1**	122.3**	135.7**	202.3**	239.3**	332.7**	384.9**	306.4**	349.0**	289.7**
Final Vintage	107.1**	117.8**	122.2**	134.4**	199.9**	234.5**	324.6**	377.3**	301.4**	344.0**	287.3**

Note: ** indicates table entries significant at the 1% level (or better): $\chi^2_{0.99}(4)$ = 13.28.

Table 6. RRMSE: ARMA and SSA (within-sample)

Method	*h*										
	11	10	9	8	7	6	5	4	3	2	1
ARMA	2.44	2.37	2.26	2.10	1.94	1.82	1.64	1.51	1.40	1.22	1.01
SSA	2.02	1.95	1.78	1.77	1.73	1.56	1.53	1.41	1.38	1.17	0.98
RRMSE	0.83**	0.82**	0.79**	0.84**	0.89**	0.86**	0.93**	0.93**	0.99	0.96	0.97

Table 7. Direction of Change: ARMA vs SSA (within-sample)

Method	*h*										
	11	10	9	8	7	6	5	4	3	2	1
ARMA	0.40	0.35	0.44	0.48	0.51	0.59	0.59	0.55	0.57	0.60	0.75**
SSA	0.62 $^{+}$	0.63**	0.63**	0.66**	0.65*	0.69**	0.71**	0.72**	0.71**	0.74**	0.78**

Table 8. RRMSE: preliminary vintage, VAR and MSSA (within-sample)

Horizon	*h*										
	11	10	9	8	7	6	5	4	3	2	1
Vintage	*v*										
	1	2	3	4	5	6	7	8	9	10	11
Method											
PV	1.71	1.60	1.48	1.37	1.28	1.14	1.02	0.92	0.79	0.66	0.54
VAR	2.50	2.44	2.24	2.04	1.91	1.71	1.59	1.42	1.20	0.98	0.68
RRMSE	1.46	1.53	1.51	1.49	1.49	1.50	1.56	1.54	1.52	1.48	1.26
MSSA	1.53	1.43	1.33	1.23	1.16	1.06	0.97	0.86	0.76	0.65	0.52
RRMSE	0.89**	0.89**	0.90**	0.90**	0.91**	0.93**	0.95**	0.93**	0.96**	0.98*	0.96*
L: r	13: 5	12: 5	11: 5	10: 5	9: 5	8: 4	7: 4	6: 3	5: 3	4: 3	3: 2
	Bias										
	1	2	3	4	5	6	7	8	9	10	11
PV	-0.75	-0.71	-0.63	-0.56	-0.49	-0.40	-0.33	-0.27	-0.21	-0.16	-0.11
VAR	0.57	0.56	0.45	0.47	0.43	0.39	0.34	0.29	0.23	0.18	0.13
Bias ratio	0.76	0.79	0.71	0.84	0.88	0.98	1.03	1.07	1.10	1.13	1.18
MSSA	-0.55	-0.52	-0.46	-0.42	-0.35	-0.29	-0.25	-0.20	-0.16	-0.12	-0.08
Bias ratio	0.73	0.73	0	0.75	0.71	0.72	0.76	0.74	0.76	0.75	0.73

Table 9. Direction of change: preliminary vintage, VAR and MSSA (within-sample)

Horizon	*h*										
	11	10	9	8	7	6	5	4	3	2	1
Vintage	*v*										
	1	2	3	4	5	6	7	8	9	10	11
Method											
PV	0.75**	0.76**	0.76**	0.73**	0.81**	0.86**	0.89**	0.92**	0.92**	0.94**	0.97**
VAR	0.57	0.49	0.53	0.54	0.61*	0.64*	0.53	0.63*	0.64*	0.70**	0.92**
MSSA	0.79**	0.84**	0.84**	0.79**	0.87**	0.87**	0.92**	0.89**	0.90**	0.94**	0.97**

Table 10. RRMSE: VAR and MSSA (out-of-sample)

Method	*h*										
	11	10	9	8	7	6	5	4	3	2	1
Vintage: 1											
VAR	3.46	3.44	3.37	3.26	3.25	3.13	3.01	2.88	2.74	2.64	2.54
MSSA	2.89	2.79	2.71	2.59	2.59	2.51	2.42	2.31	2.22	2.10	2.07
RRMSE	0.84**	0.81**	0.80**	0.79**	0.80**	0.80**	0.80**	0.80**	0.81**	0.80**	0.81**
Vintage: 2											
VAR	3.50	3.47	3.36	3.33	3.24	3.10	2.97	2.82	2.71	2.61	2.51
MSSA	2.85	2.71	2.67	2.55	2.51	2.44	2.34	2.25	2.15	2.06	20.1
RRMSE	0.81**	0.78**	0.79**	0.77**	0.77**	0.79**	0.79**	0.80**	0.79**	0.79**	0.80**
Vintage: 3											
VAR	3.47	3.42	3.31	3.28	3.18	3.04	2.91	2.75	2.65	2.54	2.44
MSSA	2.80	2.67	2.64	2.43	2.42	2.37	2.29	2.18	2.08	2.01	1.95
RRMSE	0.81**	0.78**	0.80**	0.74**	0.76**	0.78**	0.79**	0.79**	0.78**	0.79**	0.80**
Vintage: 4											
VAR	3.44	3.40	3.33	3.19	3.05	2.89	2.75	2.63	2.51	2.41	2.27
MSSA	2.72	2.62	2.60	2.37	2.35	2.30	2.20	2.09	2.00	1.95	1.87
RRMSE	0.79**	0.77**	0.78**	0.74**	0.77**	0.80**	0.80**	0.79**	0.80**	0.81**	0.82**
Vintage: 5											
VAR	3.41	3.35	3.23	3.08	2.95	2.78	2.65	2.54	2.41	2.28	2.08
MSSA	2.67	2.56	2.53	2.30	2.28	2.20	2.10	2.01	1.97	1.85	1.79
RRMSE	0.78**	0.76**	0.78**	0.75**	0.77**	0.79**	0.79**	0.79**	0.82**	0.81**	0.86**
Vintage: 6											
VAR	3.39	3.30	3.17	3.01	2.86	2.72	2.60	2.48	2.31	2.11	1.95
MSSA	2.60	2.49	2.47	2.44	2.25	2.12	2.06	1.95	1.89	1.79	1.73
RRMSE	0.77**	0.75**	0.78**	0.81**	0.79**	0.78**	0.79**	0.79**	0.82**	0.85**	0.89**
Vintage: 7											
VAR	3.28	3.18	3.03	2.88	2.74	2.61	2.49	2.32	2.12	1.95	1.75
MSSA	2.55	2.42	2.40	2.38	2.20	2.05	1.99	1.88	1.81	1.69	1.65
RRMSE	0.78**	0.76**	0.79**	0.83**	0.80**	0.79**	0.80**	0.81**	0.85**	0.87**	0.94*
Vintage: 8											
VAR	3.15	3.03	2.90	2.75	2.63	2.49	2.33	2.12	1.95	1.75	1.60
MSSA	2.48	2.35	2.33	2.24	2.13	1.94	1.89	1.77	1.72	1.62	1.53
RRMSE	0.79**	0.78**	0.80**	0.81**	0.81**	0.78**	0.81**	0.83**	0.88**	0.93*	0.96
Vintage: 9											
VAR	2.99	2.88	2.76	2.63	2.50	2.33	2.13	1.97	1.83	1.61	1.44
MSSA	2.43	2.27	2.24	2.10	2.03	1.84	1.81	1.69	1.64	1.53	1.45
RRMSE	0.81**	0.79**	0.81**	0.80**	0.81**	0.79**	0.85**	0.86**	0.90**	0.95*	0.99
Vintage: 10											
VAR	2.87	2.77	2.66	2.51	2.36	2.16	2.01	1.80	1.62	1.46	1.37
MSSA	2.36	2.19	2.16	1.95	1.83	1.77	1.69	1.56	1.53	1.39	1.35
RRMSE	0.82**	0.79**	0.81**	0.78**	0.78**	0.82**	0.84**	0.87**	0.94*	0.95*	0.99
Vintage: 11											
VAR	2.68	2.60	2.51	2.33	2.20	2.02	1.84	1.68	1.53	1.36	1.22
MSSA	2.28	2.11	2.11	1.85	1.77	1.67	1.60	1.45	1.42	1.25	1.21
RRMSE	0.85**	0.81**	0.84**	0.79**	0.80**	0.83**	0.87**	0.86*	0.93*	0.95	1.00

Table 11. Direction of change: VAR and MSSA (out-of-sample)

Method	*h*										
	11	10	9	8	7	6	5	4	3	2	1
Vintage: 1											
VAR	0.47	0.48	0.47	0.50	0.47	0.53	0.58	0.47	0.48	0.53	0.60
MSSA	0.55	0.56	0.58	0.59	0.59	0.60	0.60	0.59	0.60	0.58	0.59
Vintage: 2											
VAR	0.53	0.52	0.49	0.54	0.47	0.55	0.58	0.52	0.51	0.59	0.60
MSSA	0.56	0.55	0.56	0.59	0.59	0.61 $^{+}$	0.60	0.59	0.62 $^{+}$	0.60	0.61 $^{+}$
Vintage: 3											
VAR	0.49	0.48	0.47	0.52	0.47	0.55	0.58	0.48	0.49	0.54	0.60
MSSA	0.57	0.57	0.56	0.60	0.60	0.62 $^{+}$	0.59	0.60	0.62 $^{+}$	0.62 $^{+}$	0.63*
Vintage: 4											
VAR	0.49	0.48	0.45	0.50	0.46	0.52	0.53	0.47	0.49	0.51	0.59
MSSA	0.59	0.58	0.57	0.60	0.62 $^{+}$	0.61 $^{+}$	0.61 $^{+}$	0.61 $^{+}$	0.63*	0.65**	0.64*
Vintage: 5											
VAR	0.44	0.46	0.45	0.48	0.44	0.52	0.54	0.52	0.52	0.51	0.59
MSSA	0.59	0.58	0.58	0.61 $^{+}$	0.63*	0.64*	0.63*	0.62 $^{+}$	0.64*	0.67**	0.68**
Vintage: 6											
VAR	0.47	0.44	0.45	0.46	0.47	0.52	0.51	0.50	0.49	0.51	0.59
MSSA	0.61	0.59	0.60	0.62 $^{+}$	0.65*	0.65*	0.64*	0.63*	0.66**	0.69**	0.70**
Vintage: 7											
VAR	0.48	0.49	0.49	0.51	0.50	0.49	0.50	0.52	0.54	0.53	0.55
MSSA	0.61	0.60	0.61 $^{+}$	0.64*	0.67**	0.67**	0.67**	0.64**	0.67**	0.72**	0.73**
Vintage: 8											
VAR	0.49	0.50	0.51	0.50	0.51	0.53	0.52	0.53	0.53	0.57	0.56
MSSA	0.63 $^{+}$	0.62 $^{+}$	0.64*	0.66*	0.69**	0.69**	0.68**	0.66**	0.69**	0.74**	0.75**
Vintage: 9											
VAR	0.52	0.50	0.52	0.51	0.51	0.53	0.54	0.55	0.54	0.57	0.57
MSSA	0.65*	0.64*	0.66*	0.68**	0.71**	0.71**	0.69**	0.67**	0.71**	0.77**	0.78**
Vintage: 10											
VAR	0.51	0.53	0.53	0.50	0.53	0.57	0.56	0.55	0.57	0.59	0.59
MSSA	0.68**	0.66*	0.67**	0.70**	0.72**	0.72**	0.71**	0.68**	0.73**	0.80**	0.81**
Vintage: 11											
VAR	0.51	0.50	0.51	0.54	0.54	0.60	0.59	0.55	059	0.60	0.75**
MSSA	0.70**	0.69**	0.71**	0.73**	0.75**	0.74**	0.73**	0.72**	0.75**	0.84**	0.86*