Journal of Official Statistics, Vol. 20, No. 4, 2004, pp. 573–602

Revisions to Official Data on U.S. GNP: A Multivariate Assessment of Different Vintages

K. D. Patterson[1] *and S. M. Heravi*[2]

Although there is a substantial literature on revisions to data published by official agencies, relatively little work has been undertaken on multivariate aspects of the data measurement process (DMP) producing different vintages of the GNP variable. This is particularly so for nonstationary time series. With a focus on U.S. real Gross National Product (GNP), we show that a number of interesting questions can be answered within a multivariate framework. Defining a "well-behaved" DMP as one generating a single stochastic trend in a multiple vintage data set, we can then assess whether this is the case for GNP. We also consider whether the short-run properties of the different vintages share the same dynamic structure. Further, given multiple vintages on the same "generic" variable, is it the case that one vintage, for example the final vintage, in a well defined sense, dominates the others, and that alone can be used? We show that the idea of single (final) vintage representation is related to the idea that data revisions arise through measurement errors, and contrast this with the interpretation of revisions as forecast errors. Also, the existence of multiple vintages of GNP enables a different approach to the much-researched question of whether GNP has a unit root. This can be formulated as the null hypothesis of trend stationarity in the multivariate Johansen framework. Inter alia we show the importance of the concept of weak exogeneity, and how tests for stationarity of revisions and homogeneity of vintages can be formulated and tested.

Key words: Data measurement process; cointegration; common trends; common features; weak exogeneity.

1. Introduction

1.1. Preliminary discussion of issues

Much of the data used to analyse aggregate relationships in the economy is first published in preliminary form and then revised, sometimes quite extensively, before it becomes what can be regarded as the "final" data. Revisions to data, and particularly to GNP, the focus of this study, arise for several reasons. For example, given constant definitions and methodology, the Bureau of Economic Analysis (BEA) has to balance timeliness and incomplete information and therefore in due course, as more source data is incorporated,

[1] University of Reading, Department of Economics, Reading, Berkshire RG6 6AA, UK. Email: k.d.patterson@reading.ac.uk
[2] Cardiff Business School, Colum Drive, Cardiff CF10 3EU, Wales, UK. Email: HeraviS@cardiff.ac.uk
Acknowledgments: Detailed and constructive comments from three referees and the Associate Editor, were much appreciated. This research was supported by grants from the British Academy and the Nuffield Foundation. We are grateful to staff at the U.S. Bureau of Economic Analysis and the U.S. Bureau of Labor Statistics for responding promptly and informatively to our various inquiries. Professor Lars-Erik Öller kindly made available to us a copy of his report, with Karl-Gustav Hansson, on an international comparison of data revisions. Comments from Jesus Gonzalez and Tommaso Proietti on the permanent-transitory decomposition were invaluable.

this gives rise to revisions. (See, for example, Seskin and Sullivan (2000) and Grimm and Parker (1998).) Thus, a figure for GNP is published as quickly as possible after the period has ended, with the knowledge that some components of GNP have been based on less than complete information and, as a result, will be revised as the missing information becomes available. This leads to the view that revisions arise from measurement errors; another view is that the preliminary data is constructed as an efficient forecast of later data. These topics are discussed in greater detail in Section 3.4.

In addition, especially in recent years, there have been a number of important developments in the structure of the macroeconomy that have had implications for the measurement of GNP – see Landefeld and Fraumeni (2001) on measuring the "new" economy. Thus, together with the regular process of revision given constant definitions, there are changes arising out of the development of the "new" economy; for example, computer software has been reclassified from intermediate production to investment, thus increasing GNP. In a related development, the practice of keeping weights fixed for five years or so in constructing price and quantity indices when new products, for example, cellular phones, computers and computer games, are being introduced is likely to distort the decomposition of value into price and quantity. A recent important methodological change was the move to chained price and quantity indices; a move that has been echoed in developments in many European countries.

What is clear from this process is that the data for GNP is not constant; what is the data at one point in time is not necessarily the data at a later point. This creates a need for an awareness of the properties of the data. For example, in terms of topical comment and policy-making a particular downturn in GNP, as shown in the latest published figures, would have different implications if, in due course, the GNP figures were expected to be revised up to show that the downturn was not as severe as at first indicated. Economic research directed at explaining trends and cycles in aggregate output may well be sensitive to which version of the data is being used in the analysis; and where several variables are used, there may be differences in revision patterns amongst the variables. As a further example, European harmonisation of systems of national accounts for different countries, because budget contributions are based on measures of real (constant price) GNP/GDP, should take notice of the effects of data revisions on the indicator variables.

In general terms, we are interested in two characteristics of different measurements of data on a particular variable. (The terminology we adopt is that revisions give rise to different "vintages" of data on a "generic" variable, for example GNP.) These relate to the central tendency or trend of the series of data measurements and to the short-run or cyclical movements about the trend. A user might reasonably anticipate that at whichever point in the production of the data the generic variable is measured, the trend will be common to different measurements (vintages). Perhaps, but with less certainty, the same anticipation may be held for short-run movements about the trend. The trend/cycle distinction is a critical one in econometrics related to the distinction between nonstationary and stationary variables, and this distinction is a key part of our analysis.

The study of data revisions on National Accounts data is well established – see, for example Holden and Peel (1982a, b), Mork (1987), Patterson (1995), Patterson and Heravi (1991a, b, c), Siklos (1996), Young (1987) and Zellner (1958). For an extensive survey and assessment of revisions in an international context, – see Öller and Hansson (2002).

Steel and McLaren (2000) consider the revision of trend estimates when the X11 and X11ARIMA procedures are used. There are also regular articles by BEA staff on the revisions process in the Survey of Current Business (SCB) – see, for example Parker (1997), Moulton, Seskin, and Sullivan (2001), Moulton, Parker, and Seskin (1999), Seskin and Sullivan (2000), Seskin (1998, 1999). And there is recent interest in the use of "real time" data in econometrics – see, for example, Croushore and Stark (2001).

The present study differs from previous ones in developing an explicitly multivariate approach for nonstationary time series, to solve a number of related problems. Central to this is the concept of the data measurement process, denoted the DMP, which we briefly describe and is used in Patterson (1995, 2002a) to model the widely scrutinised and revised UK Index of Production. The number of revisions to a particular variable, regarded as the generic variable of the process, is denoted $m - 1$, arising from the m different vintages of data on that variable. Each of these m vintages is a random variable, viewed as part of a multivariate system, resulting in a realisation or outcome, which is published in the case of GNP in the Survey of Current Business (SCB).

The DMP defines a multivariate distribution for the m random variables corresponding to the m vintages of data on the generic variable. Because there are multiple vintages of data, the DMP is multivariate even though a single variable, GNP in this case, is being considered. As an example of the process of revision, and the realisations it generates, consider GNP for 1996q1. This was first published in the Survey of Current Business in June 1996 as 6,811.6 ($U.S. 1992, seasonally adjusted at annual rates). This is a realisation from the first vintage of the DMP for GNP. The first vintage was then revised to 6,818.6 in the July 1996 Survey of Current Business, which is a realisation from the second vintage of the DMP; and then to 6,814.9 in the August 1996 Survey of Current Business. (See Grimm and Parker (1998) for further details on the timing of "estimates" of GDP. The first publication of GNP lags behind GDP by one publication month in the SCB.) This figure was then subject to a number of further revisions and a change in the base year to 1996. The download for 1996q1 from the BEA website in November 2001 was 7,703.1, which is taken as the "final" value, although in principle further retrospective revisions are still possible.

It is clear from previous work on the time series properties of macroeconomic aggregates, and GNP in particular, that a multivariate framework must allow for nonstationary time series. We use the standard notation that a series that becomes stationary after differencing once is denoted $I(1)$; a series stationary without differencing, and nonstationary when summed, is denoted $I(0)$. An $I(1)$ series has a stochastic trend driven by the cumulation of past shocks rather than a deterministic trend. A related key concept for $I(1)$ variables is that of cointegration, the idea that whilst short-run deviations are possible, in the long run the series are tied together. Thus, we might observe different vintages not drawing the same picture in the short run. For example, if we take a thin "slice" out of a graph of the different vintages of GNP, we may observe differences between the series; but taking two points distant in time, the series have mapped out the same general picture.

1.2. Questions of interest

The existence of multiple vintages of data on the same generic variable, here GNP, raises some interesting questions about the relationships between the vintages. It may be helpful if at the outset we summarise the key areas that are considered in detail in the remainder of the article.

1.2.1. Do different vintages share the same long-run movements?

Is the DMP for multi-vintage GNP "well-behaved" in the sense of providing a single stochastic trend (sometimes also referred to as a common trend), and hence a single permanent – that is, $I(1)$ long-memory – component for GNP? It is probably the implicit presumption of most users that this is the case; however, this is a testable assumption and a recent analysis of multiple vintages of UK GNP (Patterson 2002b) revealed two stochastic trends, with obvious difficulties of interpretation. The possibility of multiple stochastic trends may well increase with the changes in methodology and definition arising from recent developments in the way GNP is calculated. Hence, we separate those vintages that are part of the regular revisions process from the (conditionally) final vintage, which incorporates major changes. The final vintage is conditional in the sense that later revisions may be undertaken that will result in further changes to the data. We are then interested in whether the cointegration "bond" between and amongst vintages is maintained in the light of such changes.

1.2.2. Do different vintages share the same short-run movements?

Whereas the previous question is concerned with the existence of a common trend (or trends), involving a focus on the long run, it is also of interest to know if there are common short-run properties in the data. In particular, are there common features of serial correlation? That is, can the $I(0)$ cycles be removed analogously to the way that cointegration removes the $I(1)$ property in the levels of the data? Key articles in this area are those of Engle and Kozicki (1993) and Vahid and Engle (1993, 1997) and recent applications include Mills and Holmes (1999) and Issler and Vahid (2001). The question of the existence of a single common trend relates to whether different vintages are giving a consistent picture of sustained movements in GNP, whereas the question of the existence of a common cycle relates to concerns around the particular timing of short-run movements.

1.2.3. Can we use just one vintage of the data?

Can a single vintage be taken to represent, in a well-defined sense, the properties of the complete m-vintage data set? The implicit assumption of most users is probably that the final vintage should be used, with others discarded. The permanent-transitory (P-T) decomposition due to Gonzalo and Granger (1995) (see also Stock and Watson (1988), Granger and Haldrup (1997) and Proietti (1997)) is developed to provide an answer to this question. We show that, in general, the long-memory component(s) of GNP is a combination of all vintages of data, but that it is possible to formulate a set of testable restrictions to assess whether a single vintage alone is responsible for the long-memory component of a time series. Also related to this line of analysis is whether subsystems of the variables can be separated so that the permanent components of each subsystem depend only upon variables in the individual subsystem – see Granger and Haldrup (op. cit.).

1.2.4. Do data revisions arise from measurement error or efficient forecasts?

We show that the P-T decomposition can also be used to develop the measurement error hypothesis (MEH) and the efficient forecast hypothesis (EFH) extended to nonstationary variables as an explanation of the existence of data revisions. The MEH captures the idea that different vintages of data arise because initial and preliminary vintages contain measurement errors that are removed later in the process. This description matches a key part of the official explanation for data revisions – see, for example, Seskin and Sullivan (2000). On the other hand, as a large part of the literature on data revisions indicates, the MEH is contrasted with the efficient forecast view that the preliminary vintage is an optimal forecast of a later vintage. In this view the revisions are news relative to the information set available at the time of the preliminary vintage. See for example Mankiw, Runkle, and Shapiro (1984) and Mankiw and Shapiro (1986) for the U.S. and Patterson and Heravi (1992) for the UK.

1.2.5. Can we say more about the "Is there a unit root in U.S. GNP?" debate?

Can any light be thrown on the question of whether there is a unit root in U.S. GNP? The seminal study was that of Nelson and Plosser (1982), with continuing interest represented by Murray and Nelson (2000); other studies are too numerous to list, but would include, for example, Mocan (1994), Perron and Phillips (1987), Rudebsuch (1992, 1993), Stock and Watson (1986), and Walton (1986). Despite continuing research, none (to our knowledge) has considered the implications of the multivariate process generating the data for the tests that have been undertaken. There are two particularly interesting consequences of modelling published data as realisations from a multivariate DMP for GNP, as follows.

Panel unit root tests

Given that a multivintage data set generates T observations on m vintages, what is effectively available is a panel data set on which to base unit root tests, with likely gains in efficiency. For example, in our sample there are $130 \times 4 = 520$ observations rather than the 130 available from a single vintage. Panel unit root tests have been used in a number of other areas to improve the power of standard univariate tests – see, for example, Abauf and Jorion (1990) and Taylor and Sarno (1998) on pooling country data on real exchange rates and Balz (1998) on pooling different term interest rates. In the panel context, we show not just that GNP may have one unit root but also that a set of vintages for GNP may have more than one common trend, a possibility that could not be considered with univariate unit root tests.

The null hypothesis of trend stationarity

We can ask whether as the null rather than alternative hypothesis, trend stationarity of each vintage is in the cointegration space. The framework to answer this question is the cointegrating VAR or vector equilibrium correction model, VEqCM (see for example Johansen (1995a) and Hendry (1995)).

Overall, Question 1.2.1 is at the centre of the analysis. It proposes a plausible behavioural motive of the data production agency. That is, even though several vintages of data are provided to users at different stages in the publication process, the aim, whether by refinements to reduce measurement errors, or forecasts that are efficient, is to provide vintages that are linked by having a single trend and, perhaps more hopefully, common short-run (cyclical) movements.

This article is organized as follows. Section 2 defines the notation and data for this study. Section 3 describes the analytical framework and Section 4 reports the empirical results. A summary and some concluding remarks are provided in Section 5.

2. Data and Notation

2.1. Data

An initial analysis of the data suggested that the first three vintages of data on GNP are part of a process of regular revision. These revisions are, for example, primarily due to the incorporation of information from more source data, and tend to be in the nature of refinements to the originally published data and, generally, do not incorporate definitional changes, or if so these are relatively minor. Where definitional changes are present they result in what we have referred to elsewhere (see especially Patterson and Heravi (1991b)) as a "traceback" effect due to the retrospective application of a revised definition. In this respect, a redefinition has the same practical effect as a rebasing; that is, it changes the constant price reference period, which implies changes to the expenditure weights. Under a fixed base period scheme, the opportunity afforded by regular rebasing, usually every five years, is often taken to include redefinitions. Where redefinitions have been coincidental with rebasing for the first three vintages, they will have been treated as a rebasing effect. (The rebasing method we use is described in Patterson and Heravi (1991b).)

The data for the first three vintages of GNP in constant prices was obtained from successive issues of the Survey of Current Business, giving a usable sample of 130 observations from 1965q4 to 1998q1, and expressed on a common base period of 1992 $U.S. (The change from "headlining" GDP rather than GNP led to a minor change in the publication schedule starting with the October 1991 issue of the Survey of Current Business (SCB).)

The "final" series for GNP was downloaded from the BEA website in November 2001, and is expressed in 1996 $U.S. This download represents a "real time" data set, typical of what researchers and economic commentators would refer to for comment and analysis. There will be some heterogeneity in the vintages included in this and similar downloads, since data for earlier quarters has been revised more than that for later quarters. This data incorporates conceptual and definitional changes that have been described in a series of articles in the SCB – see, for example, Moulton (2000), Moulton and Sullivan (1999) and Seskin and Sullivan (2000). It was apparent from the data that the 2001 download incorporated changes that could not be handled by standard rebasing methods (to enable all series to be expressed in the same reference period in terms of constant price units). The final series was, therefore, left on its original base and is an example of the incorporation of changes made outside the regular revisions process. The question of comparability across different base years turns out to be one that can be resolved within the analytical framework proposed in Section 3, with empirical results as shown in Section 4.4.

2.2. Notation

For convenience we use the following notation: the generic variable of interest is denoted y; in Section 3 this will be the natural logarithm of constant price GNP for the United States. Use of upper case Y refers to the level of a series. A particular vintage is indicated by the addition of a superscript denoted v, where $v = 1, \ldots, m$; thus m is the (conditionally) final vintage. The subscript indicates the period to which the data refers. In the data analysis the first three vintages for a particular time period are, therefore, y_t^1, y_t^2, and y_t^3. As noted, we take the November 2001 BEA website download as the final for the present study and, for simplicity and conformity with the vintage notation, this is denoted y_t^4.

The notation y_t *without* a superscript is the m by 1 column vector of different vintages of data, that is $y_t = \left(y_t^1, y_t^2, \ldots, y_t^m\right)'$; we use m to denote the final vintage in an analysis, which may be y_t^3 or y_t^4 depending on whether three or four vintages are used. $f(y_t)$ is the joint probability density function of y_t. A *sample* of observations and vintages, that is realizations or outcomes of the process $f\left(y_t^v\right)$ for $t = 1, \ldots, T$, and $v = 1, \ldots, m$, can be arranged into a matrix of dimensions $T \times m$ corresponding to the number of observations and the number of vintages. A total, or cumulative, revision is defined as $y_t^m - y_t^v$ and a sequential revision is defined as $y_t^{v+1} - y_t^v$; relative percentage measures are defined by dividing the revisions by y_t^v and multiplying by 100. In the tables reporting empirical results, a number without decimal places indicates a normalisation or imposed restriction.

2.3. Scale of revisions relative to the 3rd vintage

To give an indication of the scale of the initial revisions, Table 1 reports the %mean and the %mean absolute revision, %mar, relative to the 3rd vintage, respectively. The positive mean revision reflects the tendency for GNP to be revised up over time, that is on average $Y_t^v < Y_t^3$ for $v = 1, 2$. However, there are oppositely signed revisions, as indicated by the difference between the mean revision and the mean absolute revision, with the latter just below 1/3%. The % cumulative revisions decline; the larger % sequential revision is between the second and third vintage rather than the first and second vintage.

Table 1. %Mean and %mean absolute revision (%mar) to real GNP relative to the 3rd vintage

	$v = 1$	$v = 2$
%mean revision	0.138	0.106
%mar	0.315	0.253

Notes: %mar $= 100\left|\left(Y_t^3 - Y_t^v\right)\right|/Y_t^v$ for $v = 1, 2$.

3. Analytical Framework

3.1. Cointegration and common trends

This section sets out the analytical framework for a multivariate DMP generating m potentially nonstationary variables, with a view to addressing the five areas of interest highlighted in the introduction. By way of preliminaries, we first consider the central idea that different vintages should be cointegrated.

As is well known, a linear combination of $I(1)$ variables is in general also $I(1)$, the exception being where the variables share a stochastic trend resulting in a reduction in the order of integration from $I(1)$ to $I(0)$ for a particular linear combination (more precisely for a space of transformations that are observationally equivalent). Cointegration of $I(1)$ vintages makes sense from the perspectives of a data production agency and users for different vintages of data pertaining to a *single* generic variable, in this case GNP. It says that whichever vintage of data is used, it should be tied through cointegration to all other vintages.

We expect not only that y_t^v and $y_t^{v+\delta}$ should be cointegrated but also that if they are on the same constant price basis, then $y_t^{v+\delta} - y_t^v$, that is the revision, should be stationary. If cointegration exists in the vector $y_t = \left(y_t^1, y_t^2, \ldots, y_t^m\right)'$, then by the Granger Representation Theorem (see Engle and Granger 1987, and Johansen 1995a), y_t has vector autoregressive (VAR), vector equilibrium correction model (VEqCM) and vector moving average (VMA) representations. Each of these can be useful in providing information about the process generating the data vintages, and we consider the VAR, VEqCM and VMA in turn.

An appropriate analytical framework is the m-variate, pth order cointegrating VAR due to Johansen (1988, 1995a); that is:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \Psi D_t + \varepsilon_t \tag{1}$$

where $\varepsilon_t \sim \text{iid}(0, \Omega)$ and Ω is positive definite. The role of the vector D_t is to allow specification of the deterministic terms. Cointegration is associated with a reduced rank for $\Pi = \alpha\beta'$, where α and β are each of dimension $m \times r$ and of rank r, with $1 \leq r \leq (m-1)$; the cointegrating rank is r. The r vectors given by $\beta' y_t$ are the cointegrating – or equilibrium correction – vectors. Hence, with cointegration, (1) is referred to as the vector equilibrium correction model or VEqCM.

Dual to the r-dimensional cointegration space is the s-dimensional space of common trends, where $r + s = m$. The common trends – see Stock and Watson (1988) – arise from the vector moving average, VMA, representation, given by:

$$y_t = C(1)\mu t + C(1)\sum_{i=1}^{t} \varepsilon_t + C^*(L)(\varepsilon_t + \mu) \tag{2}$$

where $C^*(L)$ is a matrix polynomial; the VAR is assumed to have a constant so that $\Psi D_t = \mu$, and $C(1) = \beta_\perp(\alpha_\perp{}'\left(I - \sum_{i=1}^{p-1}\Gamma_i\beta_\perp\right)^{-1}\alpha_\perp{}'$. The symbol $\perp$ indicates the orthogonal complement of a matrix. Whilst β defines the cointegration space, $\alpha_\perp$ defines the space of the stochastic trends.

If $r = m$, then there are no unit roots, hence y_t is $I(0)$ not $I(1)$; for y_t to be $I(1)$ there must be at least one stochastic trend and $r < m$. If $D_t = (1, t)'$, then the testing procedure allows for the possibility that y_t is $I(0)$ about a deterministic trend. With an appropriate specification of the cointegrating vector when $1 \leq r \leq (m-1)$, we can also ask whether individual series in y_t are $I(0)$. In the case $s = m$, none of the stochastic trends are common, and there is no cointegration given that $m = r + s$.

3.2. Common factors and a permanent-transitory decomposition

In a development due to Gonzalo and Granger, hereafter GG, (1995) (see also Granger and Haldrup (1997) and Johansen (1998)) the s common trends are given a representation in terms of the m observable variables in the system. The GG representation is very useful from the perspective of multiple vintages of data since it may be possible to associate particular vintages with the common trend – or common factor as it is referred to in terms of observables (see also Patterson (2000b)). Below we show that when $s = 1$, the condition of weak exogeneity of the vth vintage allows a representation in terms of that vintage alone.

The GG decomposition is a permanent-transitory decomposition (a *P-T* decomposition into separate $I(1)$ and $I(0)$ components) of the m vector y_t in terms of the observable variables. In particular the *P-T* decomposition is $y_t = y_t^P + y_t^T$, where the permanent, or long memory, component is $y_t^P = \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp y_t$ and the transitory, or short memory, component is $y_t^T = \alpha(\beta'\alpha)^{-1}\beta' y_t$. For reference define $\Gamma_1 \equiv \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}$ and $\Phi_1 \equiv \alpha(\beta'\alpha)^{-1}$, and note that since the rank of a matrix product cannot exceed the smallest rank in the product, Γ_1 is of rank s and Φ_1 is of rank r. The common factors are $F_t \equiv \alpha'_\perp y_t$ and Γ_1 is the loading matrix on these. It is evident that y_t^P is $I(1)$ because $\alpha_\perp$ defines the space of common trends, and that y_t^T is $I(0)$ because $\beta' y_t$ are the cointegrating vectors and, hence, $I(0)$. The factor model exists provided $(\beta, \alpha_\perp)$ has full rank, so that $\Pi \equiv \alpha\beta'$ has no more than $m - r$ eigenvalues equal to 0 – see GG (op. cit., Proposition 3) and Johansen (1998).

The essence of the *P-T* definition is that the transitory component does not Granger-cause the permanent component and the transitory component is covariance stationary whereas the permanent component is difference stationary. Under specific conditions, it is these aspects that enable us to relate the revisions to the measurement error hypothesis. This is discussed further below in Section 3.4.

What do we expect to find in the case of multi-vintage data? In what we might term a "well-behaved" DMP, $r = m - 1$ and there is, therefore, one common trend, that is $s = 1$. We also term this full inter-vintage cointegration. This single common trend can be interpreted as "driving" the process that generates $m > 1$ vintages of data. In terms of observables, the corresponding single common factor is $f_{1t} = \sum_{v=1}^{m} b_{1v} y_t^v$, where b_{1v} is the vth element of the $1 \times m$ weighting vector $\alpha'_\perp$. The permanent component of the vth vintage is, therefore, $\gamma_{v1}\sum_{v=1}^{m} b_{1v} y_t^v$, where $\Gamma_1 = (\gamma_{11}, \ldots, \gamma_{m1})'$, which is the common factor scaled by γ_{v1}.

It is as well not to take the well-behaved property of the DMP for granted. There is some evidence that controverts this presumption. For the UK, Hendry (1983, 1994) notes problems with different vintages of data used in modelling the consumption function. Caplan and Daniel (1992), Cook (1994) and Jenkinson and Brand (2000) report some of the initiatives taken by the UK's main data agency following considerable difficulties of interpretation with the scale of changes to different vintages of National Accounts data. This was confirmed by Patterson (2002b), who found two stochastic trends in some measures of UK GNP. For the U.S., the results in Siklos (1996) cast doubt on uniform acceptance of well-behaved DMPs in the National Income and Product Accounts; and the recent set of changes in the comprehensive revision to GNP – see Moulton and Sullivan (1999) – must leave this an open question a priori.

3.3. *Identification of the cointegrating vectors and weak exogeneity*

3.3.1. Identification

We next consider identification of the cointegrating vectors (Johansen and Juselius (1992), Johansen (1995b)). A sufficient condition for just identification of the cointegrating vectors is that there are $r-1$ independent restrictions on each vector. Restrictions in excess of this number are overidentifying, and can be tested using a likelihood ratio test statistic asymptotically distributed as $\chi^2(q)$, where q is the number of overidentifying restrictions.

Just-identifying restrictions follow from the view that all pairs of vintages should cointegrate. One could, for convenience, regard all vintages as pairwise cointegrating with the final vintage. Equivalently, the sequential vintages could be regarded as cointegrating. This follows from the transitivity of the cointegrating relations, or we can simply show that the implied cointegrating vectors from either parameterisation give rise to the same cointegration space. That is Π can be written as $\Pi = \alpha\beta'$ or $\Pi = \alpha\kappa^{-1}\kappa\beta'$, where κ is a full rank matrix. Whichever, this then implies that vintage i pairwise cointegrates with vintage j.

We illustrate this argument for the case $m = 4$, $r = 3$ and $r - 1 = 2$. Thus if, for example, we take β parameterised to reflect pairwise cointegration of vintage v with vintage $v + 1$ and a normalisation of the ith column of β on the ith vintage, $i = 1, 2, 3$, then κ, and the connection between parameterisations, is given as follows.

$$\kappa\beta' = \begin{bmatrix} 1 & -\beta_{21} & \beta_{21}\beta_{32} \\ 0 & 1 & -\beta_{32} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \beta_{21} & 0 & 0 \\ 0 & 1 & \beta_{32} & 0 \\ 0 & 0 & 1 & \beta_{43} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \beta_{41} \\ 0 & 1 & 0 & \beta_{42} \\ 0 & 0 & 1 & \beta_{43} \end{bmatrix} \tag{3}$$

It is evident that pairwise cointegration of vintages provides a sufficient set of 2 (independent) identifying restrictions for each cointegrating vector.

Note that although (3) implies that the different vintages are cointegrated, it does not imply that the revisions are stationary, or equivalently that the levels are homogenous of degree 1. This requires, in addition, $\beta_{21} = \beta_{32} = \beta_{43} = -1$. (The transformation matrix κ above will now be a matrix with 1's on the diagonals, nonzero upper triangular elements and 0's elsewhere.) With $m = 4$, stationarity of the revisions provides 3 testable restrictions and, in general, $m - 1$ restrictions that can be tested with a likelihood ratio test statistic distributed as $\chi^2(q)$ where $q = m - 1$.

3.3.2. Weak exogeneity

The permanent component (and the common factor) is *not* generally a single vintage but a linear combination of the m vintages. However, the situation is simplified in the case of full inter-vintage cointegration if one of the vintages – let us take as an example the final vintage – is weakly exogenous (for the parameters of interest α and β); for brevity we refer to this as weak exogeneity, see, for example, Johansen (1992).

For example, if $r = m - 1$ then weak exogeneity of the final vintage corresponds to the restriction that the last row of α is a $1 \times (m-1)$ zero vector, which can be tested with a likelihood ratio test statistic asymptotically distributed as $\chi^2(m-1)$. The matrix $\alpha_\perp$ can then be constructed with a 0 in each of the first $m - 1$ rows and a nonzero element in the mth row, normalised at 1. This construction satisfies $\alpha'_\perp \alpha = 0$ and $(\alpha, \alpha_\perp)$ of full rank and the common factor is $F_t = \alpha'_\perp y_t = y_t^m$. Simple calculations then show that the permanent component is a just a function of the final vintage.

We illustrate this result with $m = 4$ and $r = m - 1 = 2$, then the *P-T* decomposition, assuming the just-identification parameterisation in (3) and weak exogeneity of the final vintage, is:

$$\begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & -\beta_{41} \\ 0 & 0 & 0 & -\beta_{42} \\ 0 & 0 & 0 & -\beta_{43} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} + \begin{bmatrix} 1 & 0 & 0 & \beta_{41} \\ 0 & 1 & 0 & \beta_{42} \\ 0 & 0 & 1 & \beta_{43} \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} \tag{4}$$

Stationarity of the revisions implies $\beta_{41} = \beta_{42} = \beta_{43} = -1$, which combined with weak exogeneity of the final vintage implies that the final vintage is *the* common factor, *and* that the permanent component and the transitory component are the cumulative revisions.

That is, $y_t = y_t^P + y_t^T$ with $y_t^p = \left(y_t^m, \ldots, y_t^m\right)'$ and $y_t^T = (y_t^1 - y_t^m, \ldots, y_t^{m-1} - y_t^m, 0)'$. Thus (4) becomes

$$\begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} + \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} \tag{5}$$

This analysis shows that the implicit presumption that the final vintage is the permanent component rests upon a number of testable conditions. The decompositions in (4) and (5), and their generalisations, are not unique because it is only the dimension of the cointegration space, and hence the common trends space, that is unique. The choice of a particular rotation is reflected in κ, with that choice guided by potentially interesting interpretations of the resulting cointegrating vectors and adjustment coefficients, as is the case with (4) and (5). Further, we can relate these conditions to the measurement error hypothesis, which is a leading explanation for data revisions.

3.4. The Measurement Error Hypothesis (MEH) and the Efficient Forecast Hypothesis (EFH)

The idea that data revisions are due to measurement errors is central both to research and official explanations for multiple vintages – see, for example, Mork (1987) for the U.S. and Patterson and Heravi (1992) for the UK. Preliminary vintages have to satisfy the need for timely data but they contain errors due to inaccurate measurement, which are reduced by subsequent revisions. Seskin and Sullivan (2000) note of the revisions process that later "estimates incorporate source data that are more complete, more detailed, and otherwise more appropriate than those that were previously incorporated." According to the measurement error hypothesis, which has hitherto been stated in the literature for $I(0)$ variables, the preliminary vintage(s) is(are) equal to the final vintage plus an orthogonal measurement error.

In contrast to the MEH, the efficient forecast hypothesis, EFH, views the preliminary data as an efficient forecast of later vintages. For example, on this view a preliminary vintage is optimal in a mean squared error sense in predicting any succeeding vintage given information available at the time that the preliminary vintage is formed. As a result, revisions relative to later vintages should be orthogonal to the preliminary vintage.

For stationary data it is possible to draw out a number of contrasting implications of the MEH and EFH that allow them to be distinguished – see, for example, Patterson and Heravi (1992). The MEH implies that y_t^v is inherently noisier than $y_t^{v+\delta}$ for $\delta > 0$, as indicated, for example, by their respective variances, whereas the reverse is true for the EFH. The MEH implies that the regression of the revision $y_t^{v+\delta} - y_t^v$ on $y_t^{v+\delta}$ should yield a statistically insignificant coefficient on $y_t^{v+\delta}$ (because the measurement error is orthogonal to the conditionally final vintage); on the other hand, $y_t^{v+\delta} - y_t^v$ is correlated with y_t^v. In contrast, the EFH implies that the regression of $y_t^{v+\delta} - y_t^v$ on y_t^v should yield an insignificant coefficient on y_t^v (because the revision is "news" as far as y_t^v is concerned); on the other hand $y_t^{v+\delta} - y_t^v$ is correlated with $y_t^{v+\delta}$.

However, when the time series are generated by nonstationary processes, this framework is invalid. Both the MEH and the EFH imply that $y_t^{v+\delta} - y_t^v$ is $I(0)$ given that $y_t^{v+\delta}$ and y_t^v are each $I(1)$, thus the proposed regressions involve regressing an $I(0)$ variable on an $I(1)$ variable, which will result in a regression coefficient of zero asymptotically. Further, typically variances are estimated by the ensemble counterpart. For example, suppose we wanted to compare the variance of y_t^v with the variance of $y_t^{v+\delta}$. For *stationary* variables, it is sensible to estimate these variances by the corresponding sample variances for $t = 1, \ldots, T$, on the assumption that this is an ensemble of realisations from the same process for each value of t; however, when the variables are (second-order) nonstationary the concept of a single variance no longer makes sense, the variances are indexed by t and the sample variance does not estimate any of these variances.

When the data are nonstationary, we can exploit the GG decomposition to distinguish between the MEH and the EFH. A defining feature of the GG decomposition is that shocks to the transitory components are orthogonal to the permanent components in the long run, in the sense that they have no *long-run* effect on the permanent components, and hence no effect on the level of y_t. Thus, they may have a short-run effect but the total multiplier of Δy_t^P with respect to y_t^T is zero. In the previous section we derived the conditions that enable the transitory components to be interpreted as the cumulative revisions, with no transitory component – or in this context no measurement error – on the final vintage. A straightforward implication of these conditions is that the revisions are measurement errors, which are in the long run orthogonal to the final vintage. The conditions for this orthogonality, and hence the MEH interpretation, to hold are, therefore, stationarity of the revisions plus weak exogeneity of the final vintage.

On the other hand an implication of the EFH is that transitory components should be orthogonal in the long run to the initial vintage not to the final vintage. That is, $y_t = y_t^P + y_t^T$ but now with $y_t^P = (y_t^1, \ldots, y_t^1)'$ and $y_t^T = (0, y_t^2 - y_t^1, \ldots, y_t^m - y_t^1)'$. In the EFH interpretation, y_t^1 is the permanent $I(1)$ component and the revisions relative to y_t^1 are the transitory components and hence $I(0)$. Thus the revisions, which are the transitory components, can be interpreted as "news" relative to y_t^1. More complex situations, where some vintages fit the MEH and some fit the EFH, can arise and are dealt with below; it helps to understand these if we first consider the case where y_t^1 is the permanent component in greater detail.

The conditions for y_t^1 to be the permanent $I(1)$ component and revisions to be "news" are obtained as follows. First, return to (3) and choose (a rotation) $\tilde{\kappa}$ such that $\tilde{\kappa}\beta'$ is in the observationally equivalent form given by:

$$\tilde{\kappa}\beta' = \begin{bmatrix} \beta_{11} & 1 & 0 & 0 \\ \beta_{12} & 0 & 1 & 0 \\ \beta_{13} & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

If y_t^1 is weakly exogenous and $\beta_{11} = \beta_{12} = \beta_{13} = -1$, then (5) becomes, as required:

$$\begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} \tag{7}$$

That is y_t^1 is the permanent $I(1)$ component and the revisions are the $I(0)$ components.

Of note is that the (homogeneity) restriction $\beta_{1j} = -1,\ i = 1, 2, 3$, is observationally equivalent to the MEH restriction $\beta_{4j} = -1,\ i = 1, 2, 3$, and just reflects the fact that according to *both* the MEH and EFH the linear combinations given by any pair of vintages are stationary. Thus, the essence of what distinguishes the MEH and the EFH is whether, given that homogeneity is not rejected, the final vintage or the first vintage is weakly exogenous, respectively. This *is* a testable distinction.

Now consider the situation in which y_t^2 is the permanent $I(1)$ component so that $y_t^P = (y_t^2, \ldots, y_t^2)'$ and $y_t^T = (y_t^1 - y_t^2, 0, \ldots, y_t^4 - y_t^2)'$. Then:

$$\begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} + \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} \tag{8}$$

How is this structure to be interpreted? Now y_t^1 is no longer an efficient forecast of later vintages since the revision $y_t^2 - y_t^1$ is orthogonal to y_t^2 rather than y_t^1, so that the revision is a measurement error. However, y_t^2 is an efficient forecast of y_t^3 and y_t^4 in the sense that the corresponding revisions $y_t^3 - y_t^2$ and $y_t^4 - y_t^2$, respectively, are orthogonal to y_t^2. Tests to distinguish the MEH and EFH are reported in Section 4.

3.5. Serial correlation common features and common cycles

In Section 3.1, we considered how to assess whether there were common long-run features in the multi-vintage data set, the central question of that section being whether there was one common (stochastic) trend in m vintages of data. In this section we move the focus to short-run movements in the data. In this case the central idea is that the short-run correlations amongst the vintages may (also) be common. These common aspects, whether referring to the long run or the short run, are referred to as *common features*.

A serial correlation common feature is said to exist if there exists a linear combination of Δy_t which is an innovation with respect to information prior to t, (Vahid and Engle, hereafter VE, 1993). Intuitively, a common feature occurs when the Δy_t move together; it refers to short-run movements rather than the long-run tendencies in y_t that are captured by cointegration. There can be $k < m$ independent linear combinations, individually called cofeature vectors, with this property. Together they form an $m \times k$ matrix, say Φ, of rank k, such that $\Phi' \Delta y_t$ is a $k \times 1$ vector of innovations. VE (1993, Propositions 1 and 2) show that these linear combinations remove the same number, k, of cycles in the VMA given by (2), that is $\Phi' C^*(L)\varepsilon_t = 0$. Removing k cycles means that $m - k$ cycles rather than m are common. Given this equivalence, a test for a serial correlation common feature of at least order k in Δy_t is a test for at most $m - k$ common cycles. Just as cointegration in an m-vector of $I(1)$ series reduces the number of common trends, so the presence of common features reduces the number of $I(0)$ cycles.

In the present context we note the following. If $r = m - 1$, that is, the number of stochastic trends has been reduced from m to 1, a description we characterised as a well-behaved DMP, then k can only be 0 or 1. If $k = 0$ there is no reduction in the maximum number m of common cycles, whereas if $k = 1$ there are $m - 1$ common cycles, which is the same as the cointegrating rank. (This is an application of Vahid and Engle's Theorem 1.) The cofeature vector, if it exists, will be independent of the cointegrating vectors.

Vahid and Engle (op. cit.) suggest the following test statistic:

$$C(p,k) = -(T - p - 1)\sum_{i=1}^{k} \ln(1 - \rho_i^2)$$

The ρ_i^2 are the squared canonical correlations, in increasing order, between $X_t = (\Delta y_{1t}, \ \Delta y_{2t}, \ldots, \Delta y_{mt})'$ and $W_t = (\Delta y_{1t-1}, \ \Delta y_{2t-1}, \ldots, \Delta y_{mt-1}, \ldots, \Delta y_{1t-p}, \Delta y_{2t-p}, \ldots, \Delta y_{mt-p}, \ \hat{\beta}_1' y_{t-1}, \ldots \hat{\beta}_r' y_{t-1})'$ where $\wedge$ denotes a consistent estimator, here the Johansen estimator. The test statistic is asymptotically distributed as $\chi^2(df)$ with degrees of freedom given by $df = k(k + mp + r - m)$. The sequence of test statistics $C(p, 1 \le k \le k^{\max})$, is used to determine the number of serial correlation common features. The ρ_i^2 can be calculated by exploiting the dual eigenvalue problem (see Johansen 1995a, Lemma A9).

Alternatively, and of use if it is found that $k > 0$, VE show that a likelihood ratio test can be constructed by first carrying out FIML estimation of the system given by:

$$\begin{bmatrix} I_k & \phi^{*\prime} \\ 0_{(m-k)\mathrm{x}k} & I_{m-k} \end{bmatrix} X_t = \begin{bmatrix} 0_{k\mathrm{x}(mp+r)} \\ \tilde{\Gamma}_1, \ldots, \tilde{\Gamma}_p, \tilde{\alpha} \end{bmatrix} W_t + \tilde{\mu} + \nu_t \tag{9}$$

and then comparing it with the unrestricted reduced form VEqCM, which differs only in replacing the first k equations by equations analogous to the last $m - k$ equations. A test of the restrictions has df degrees of freedom and is equivalent to $C(p, k)$.

4. Empirical Analysis

This section provides the results of a number of hypothesis tests related to the previous analysis. It may be helpful at the outset to give a brief summary of the issues of interest.

In Section 4.1, we address the central question of how many stochastic trends there are in the data given by the first three vintages in the first instance, and then with inclusion of the final vintage, taken as a download from the BEA website for the purposes of this study. For example, in the first case a cointegrating rank of two implies one stochastic trend, whereas a cointegrating rank of one implies two stochastic trends.

In Section 4.2, a multivariate perspective is used to consider the question of whether real GNP has a unit root and is, therefore, difference stationary or trend stationary. In Section 4.3, we assess whether any variables can be excluded from the analysis; that is, do any of them play no part in the cointegrated system? The analysis in these two sections is related by consideration of whether a deterministic trend can be excluded from the system; we conclude that it can be excluded.

In Section 4.4, we assess two component issues related to the distinction between the measurement error hypothesis and the efficient forecast hypothesis; a necessary condition for both is that the (log) revisions should be stationary. Then the distinction hinges on which vintage, if any, is weakly exogenous.

In Section 4.5, the empirical analysis relates back to the question of how to compare the relatively homogenous first three vintages with the last. The latter is on a different constant price basis from the former, owing to the methodological change from a constant base year to a chained index. We show how this can be achieved using the results of the previous empirical analysis.

Finally, in Section 4.6, we consider whether the short-run movements in the different vintages share a serial correlation feature (analogous to the idea that in the long run the different vintages share a single stochastic trend). This hypothesis is firmly rejected, so that whilst the different vintages trace out the same long-run movement, there are departures in short-term behaviour of the vintages.

When a number of hypothesis tests are undertaken, as in this section, the cumulative, or overall, size of the testing procedure will, in general, exceed the individual size (that is, type one error) for each test. For example, two separate tests that are independent, each carried out at the α significance level, will result in a cumulative type one error of $1 - (1 - \alpha)^2$, for example 9.75% for $\alpha = 5\%$. As, in general, we do not know the extent of dependence between tests, one strategy to reduce the cumulative size is to reduce the size of each of the tests in the sequence, for example carry out each of n tests at a size of α/n (using the Bonferroni equality). Alternatively a single joint test, where available, could be used. (We are grateful to a referee for drawing attention to these points.) Although we are carrying out a number of tests, the results are unambiguous and robust in this respect; for example, often the p-values associated with the test statistics are virtually zero.

4.1. Cointegrating rank and number of common trends

The starting point is the VEqCM given by (1), with a lag order of 2 determined by standard techniques – both AIC and SIC lead to this choice (see, for example, Johansen (1995a) and Patterson (2000a)). The VEqCM allows a trend in the data space, which is not necessary in the cointegration space as suggested by application of the Pantula (1989) principle – see Hansen and Juselius (1995, especially Section 6.2). We also return to the question of the inclusion of the trend in Sections 4.2 and 4.3, in assessing the hypothesis of trend stationarity for individual data vintages.

We first briefly consider the situation with the first three vintages, with the test statistics for cointegrating rank reported in Table 2a. Consideration is also given to small sample corrections to the trace test statistic, which have been suggested by Ahn and Reinsel (1988) (see also Reimers (1992)), Hansen and Rahbeck (2000) and Johansen (2001). Ahn and Reinsel suggest multiplying the trace statistic by $T^{-1}(T - kp) \equiv CF_{AR}$, Hansen and Rahbeck suggest $T^{-1}(T - (k - 1)p) \equiv CF_{HR}$, where p is the lag length, k is the number of variables in the analysis, ($k = m$ for the cases considered here), and T is the effective sample size. For $m = 3$, $CF_{AR} = 0.95$ and $CF_{HR} = 0.97$; for $m = 4$ these are 0.94 and 0.95, respectively. The conclusions drawn from the test statistics in Tables 2 and 3 are robust to these small sample adjustments.

Table 2a. Test statistics for cointegrating rank (first three vintages)

Rank	$\hat{\lambda}$	Trace	Quantiles	
null			95%	95%(simulated)
$r = 0$	0.35	84.98	29.7	33.0
$r \leq 1$	0.21	30.39	15.4	16.7
$r \leq 2$	0.00	0.12	3.76	4.09

Note: $\hat{\lambda}$ are the estimated eigenvalues and, throughout, asymptotic 95% quantiles are from Osterwald-Lenum (1992); simulated quantiles are indicated by 95%(simulated).

Table 2b. Univariate diagnostic tests

	Equation 1	Equation 2	Equation 3
Normality			
Jarque-Bera	18.92 [0.00]	18.13 [0.00]	30.88 [0.00]
kurtosis	1.76 [0.00]	1.81 [0.00]	2.44 [0.00]
skewness	−0.37 [0.10]	−0.23 [0.29]	−0.10 [0.65]
ARCH	4.14 [0.13]	4.69 [0.10]	3.14 [0.21]

Note: Kurtosis and skewness statistics are from Kendall and Stuart (1958), programmed in RATS. Marginal significance level (p-value) of test statistic shown in [.].

Johansen (2001) shows that small sample adjustments of this kind are one part of an adjustment based on the Bartlett (1937) correction principle; in particular the correction factor is also a function of the number of common trends. Following Johansen (2001), we simulated the finite sample distributions using the empirical model as a base. In the first instance we generated simulations using the multivariate normal distribution to draw the "errors." We refer to this as the *mnid* (multivariate normal, independently distributed) case. For a variation on this bearing in mind the diagnostic tests reported below, which show excess kurtosis in the residual distributions relative to the normal distribution, we drew the ε_t from the empirical distribution function (that is, the distribution of the residuals from the estimated equations), a procedure known as "bootstrapping."

The simulated quantiles in the *mnid* case, based on 25,000 replications, were about 10% larger than the asymptotic quantiles (an example in Johansen (2002) based on fewer observations than in our case, $T = 53$, resulted in an increase of about 16% in the quantiles). For the variation with bootstrapped errors, we found the 95% quantiles were between 8% and 12% smaller than in the *mnid* case. Thus, the bootstrapped variation had the effect of returning the 95% quantiles to close to where they were before the small sample adjustment was made. In view of this we report the 95% quantiles from the simulated distributions in the *mnid* case in Tables 2 and 3, along with the standard asymptotic quantiles. It is clear that use of the small sample adjusted quantiles rather than the quantiles from the asymptotic distribution does not affect the force of the results reported here. Some diagnostic test statistics for nonnormality and ARCH effects are reported in Table 2b.

Diagnostic tests for residual autocorrelation suggest no problems at conventional significance levels; for example, the Lagrange Multiplier, LM, test for first order residual correlation in the four-equation system resulted in $LM(1) = 7.79[0.56]$ and the test for fourth order resulted in $LM(4) = 5.85[0.76]$. (Throughout, the p-value of the test statistic is shown in [.] following the test statistic.)

There are also no indications of ARCH effects. The Jarque-Bera tests indicate nonnormality in the residual distributions and separate test statistics for (excess) kurtosis and skewness suggest that this nonnormality is due to leptokurtic residual distributions with symmetry being maintained; in turn, this leptokurtic property is inherited from the revisions. Johansen (1995a) notes that the asymptotic properties of the maximum likelihood estimation and testing methods depend on an assumption of independent and identically distributed errors rather than Gaussian errors. And, as reported above, our simulations to obtain small sample adjusted quantiles indicate that our results are robust to the excess kurtosis.

The test statistics in Table 2a unambiguously suggest a cointegrating rank of 2. For example, the null of $r \leq 2$ against the alternative of $r = 3$ is not rejected, with a test statistic virtually 0 compared to the 95%(simulated) quantile of 4.09; however, the trace test statistic for $r \leq 1$ is 30.39 compared to the 95%(simulated) quantile of 16.7. There is a very clear distinction between the first two eigenvalues and the last eigenvalue.

In summary, these results suggest that:

(i) the hypothesis that all the series are stationary can be rejected: there is at least one stochastic trend driving the first three vintages; obversely, there is at least one linear combination of the three nonstationary series that is stationary.

(ii) the nonrejection of a cointegrating rank of two implies that just one stochastic trend is responsible for the $I(1)$ component in each of the three vintages. This is reassuring from a data user's point of view, since the chance selection of one vintage rather than another will not lead to a series with different long-run characteristics. In a sense the revisions are benign in not producing series that will wander apart given sufficient time.

In starting with $m = 3$, we can consider whether adding the final vintage, so that m is increased by one, increases the cointegrating rank. The hypothesis here relates to the idea that because of the substantial nature of some recent revisions to GNP, particularly the reclassification of computer software and the move to current rather than base weighted indices, the final vintage whilst still nonstationary is separated from the first three vintages. (On the concept of separation in general, see Konishi and Granger (1992) and Granger and Haldrup (1997).)

The results for $m = 4$ are reported in Table 3. These results are again unambiguous: the null hypothesis $r \leq 2$ is rejected with a trace statistic of 25.67 compared to the 95% quantile (simulated) of 16.7; however, the null hypothesis of $r \leq 3$ is not rejected. Thus, there is a firm indication that adding the final vintage increases the rank of the three variable system from 2 to 3, so separation is rejected. Thus, despite the nature of recent definitional and methodological changes to GNP, the final vintage does still have a relationship – that is it is cointegrated with – the earlier vintages. The substantial changes did not break the cointegration "bond" amongst vintages.

Table 3a. Cointegrating rank: first three vintages and final vintage (2001 download)

Rank	$\hat{\lambda}$	Trace	Quantiles	
null			95%	95%(simulated)
$r = 0$	0.36	116.0	47.2	52.1
$r \leq 1$	0.22	58.01	29.7	33.0
$r \leq 2$	0.18	25.67	15.4	16.7
$r \leq 3$	0.00	0.06	3.76	4.09

Note: $\hat{\lambda}$ are the estimated eigenvalues and, throughout, asymptotic 95% quantiles are from Osterwald-Lenum (1992); simulated quantiles are indicated by 95%(simulated).

Table 3b. Univariate diagnostic tests

	Equation 1	Equation 2	Equation 3	Equation 4
Normality				
Jarque-Bera	12.20 [0.00]	11.69 [0.01]	24.27 [0.00]	12.72 [0.00]
kurtosis	1.44 [0.00]	1.41 [0.00]	2.05 [0.00]	1.44 [0.00]
skewness	−0.27 [0.22]	−0.27 [0.23]	−0.239 [0.29]	−0.31 [0.16]
ARCH	3.31 [0.19]	2.97 [0.23]	2.28 [0.32]	0.025 [0.99]

Note: Kurtosis and skewness statistics are from Kendall and Stuart (1958), programmed in RATS. Marginal significance level (p-value) of test statistic shown in [.].

The diagnostic test results for residual autocorrelation, reported in Table 3a, again suggest no problems at conventional significance levels with the four-equation system; for example, LM(1) = 14.235[0.58] and LM(4) = 13.693[0.62]. There are also no indications of ARCH effects. The nonnormality arises from leptokurtic residual distributions with

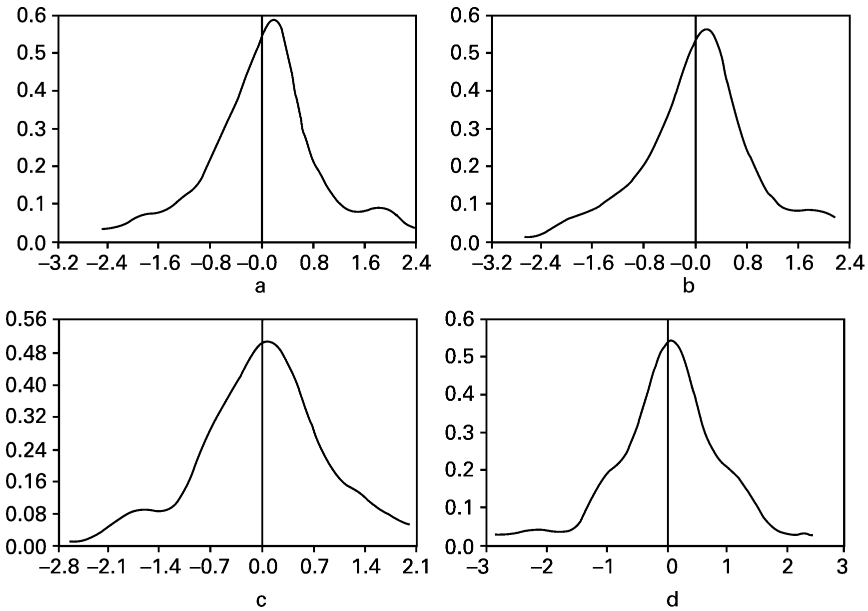

Fig. 1. Standardised equation residuals (estimated density functions)

symmetry being maintained. Figure 1 shows the estimated density functions for the standardised equation residuals, which illustrate the broad retention of symmetry but excess kurtosis relative to the normal distribution.

Thus, in summary, we do not reject the hypothesis of one common trend, and so one common factor, in the four vintages. This corresponds with the description, proposed in the Introduction, that the data measurement process is "well-behaved," despite a number of recent changes. Given this finding we can consider whether the common factor can be expressed in terms of a single vintage. We return to this and other issues of specification after considering tests for trend stationarity.

4.2. Testing for trend stationarity of real GNP

There has been a considerable debate about whether real GNP has a unit root and is, therefore, difference stationary or is stationary about a deterministic trend. Partly, the importance of this debate centres on the different implication of the persistence of shocks in the two classifications, shocks being infinitely lived in the former case but finitely lived in the latter case; see, for example, Nelson and Plosser (1982) and recently Murray and Nelson (2000).

The finding that $1 \leq r \leq (m-1)$ is rejection of stationarity for the group of vintages as a whole: from the rank deficiency of Π there must be at least one common trend. Suppose Π is of full rank, that is m, then this can only be so if each of the m variables is stationary. Thus, a test of the null hypothesis that $r = m$ against the alternative that $r \leq (m-1)$ can be viewed as a test of the null hypothesis of stationarity against the alternative that at least one of the variables is $I(1)$. In the context of testing for cointegrating rank this is just the trace test statistic for rank(Π) = m against rank(Π) = ($m - 1$), which is asymptotically distributed as χ^2 with 1 degree of freedom.

It is also possible to test the role of the deterministic trend in the common trends space for *each* vintage. (We are grateful to a referee for pointing out that there is also a joint test for trend stationarity against difference stationarity, described in Johansen (1995a, p 97).) In this case, application of the joint test leads to rejection and we are then interested in whether there is any distinction amongst the individual vintages, which is the test reported in the text. Testing for trend stationarity on individual variables in the Johansen framework can be undertaken by specifying the deterministic terms in (1) as the vector $D_t = (1, t)'$, and then allowing the trend to enter the cointegration space. Alternatively if $D_t = (1, 0)'$ we can test for stationarity but about a nonzero mean rather than nonzero trend. Although the former is more realistic given that GNP clearly has a trend, we report both approaches to indicate that the results are unambiguous.

The maintained multivariate regression is now:

$$\Delta y_t = \mu + \alpha\left(\beta' \delta'_1\right) \begin{pmatrix} y_{t-1} \\ t \end{pmatrix} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{10}$$

where δ'_1 is an $r \times 1$ vector of trend coefficients, with elements δ_{1j}, $j = 1, \ldots, r$, in the cointegration space. Then, for example, for the first vintage the hypothesis of trend stationarity amounts to asking whether space(b) $\subset$ space(β), where $b = (1, 0, 0, 0)'$, and $\delta_{11} \neq 0$. That is, does the cointegration space contain the first vintage on its own? This implies the following hypothesis:

$$H_O : \begin{pmatrix} \beta \\ \delta_1 \end{pmatrix} = (H_1 \varphi_1, \varphi_2)$$

where the columns of H_1 are $(1, 0, 0, 0, 0)'$ and $(0, 0, 0, 0, 1)'$, φ_1 is 2×1 and φ_2, of dimension $(m + 1) \times (r - 1)$, is unrestricted. The null hypothesis can be tested by means of a likelihood ratio test statistic, which is distributed as $\chi^2(g)$, where $g = (m + 1) - r$, or $J = m - r$ if $\delta_{1j} = 0$ is imposed; for general principles on testing hypotheses of this form see Johansen (1995a) and Hansen and Juselius (1995). This approach, whilst obviously differing from the usual ADF type approach in focussing on a multi-vintage framework, also differs in taking the trend stationary hypothesis as the null rather than the alternative hypothesis. The results for four vintages are summarised in Table 4.

It is evident from the results in Table 4 that, with p-values of 0 throughout, the null of trend stationarity is firmly rejected whether or not a deterministic trend is included in the cointegration space. These results are also robust to different selections of r, and therefore support the view that GNP is difference stationary rather than trend stationary.

4.3. *Testing for exclusion of variables from the system*

The uniform rejection of the hypothesis of trend stationarity is not surprising approached from the perspective of appropriate determination of the VAR specification. We could, for example, include a deterministic trend in the cointegration space and then see whether it can be excluded. Indeed, we can test directly for exclusion from the system of *any* of the

Table 4. Individual tests for trend stationarity against difference stationarity

χ^2 (g) test for stationarity about a trend of the jth vintage with $\delta_{1j} \neq 0$					
r	g: $\chi^2_{0.05}(g)$	y_t^1	y_t^2	y_t^3	y_t^4
3	$g = 2 : 5.99$	24.82[0.00]	24.83[0.00]	24.84[0.00]	25.06[0.00]
χ^2 (J) test for stationarity about a non-zero mean of the jth vintage with $\delta_{1j} = 0$					
r	J: $\chi^2_{0.05}(J)$	y_t^1	y_t^2	y_t^3	y_t^4
3	$J = 1 :$ 3.84	24.82[0.00]	24.83[0.00]	24.84[0.00]	25.06[0.00]

Notes: p-value in [.]; r is the cointegrating rank; y_t^4 is the download from the November 2001 BEA website; entries correspond to selection of $r = 3$ from cointegration results; the alternative choices of $r = 1, 2$ do not alter the conclusion.

variables. Thus, starting with a deterministic trend, which we allow in the cointegration space (but not in the data space since there is no evidence of quadratic trends in the data), we can test for exclusion of each variable in turn from the system.

The results are again unambiguous and are reported in Table 5. We can exclude the deterministic trend, where the test statistic has a p-value of 63%, but we cannot exclude any of the data vintages. Then, having excluded the time trend, we can again see if any of the vintages can be excluded. The lower part of Table 5 shows that this is not the case, as all p-values are effectively zero.

Table 5. Test statistics for exclusion of variables, distributed as χ^2 (3)

	Variable excluded				
Deterministic time trend in cointegrating space					
r	y_t^1	y_t^2	y_t^3	y_t^4	time trend
3	51.57[0.00]	49.67[0.00]	32.13[0.00]	20.57[0.00]	1.71[0.63]
No deterministic time trend in cointegrating space					
3	7.82[0.00]	50.93[0.00]	34.06[0.00]	20.63[0.00]	

4.4. *Identification and weak exogeneity*

Looking again at the three-vintage system it seems likely that this will easily sustain stationarity of the revisions (and, therefore, homogeneity of the vintages in levels), a total of 4 restrictions. With $m = 3$ and $r = 2$, two of the four restrictions are just identifying, so that the test statistic has a χ^2 distribution with 2 degrees of freedom. The test statistic is 0.16 with a p-value of 93%, so the null is not rejected.

We now consider tests for weak exogeneity; these are invariant to any choice of restrictions that just identify the cointegrating vectors. Weak exogeneity of any of the individual vintages implies that the GG permanent component is a function of that vintage alone, which together with stationarity of the revisions implies that the (log) revisions are the transitory components. See Section 3.4 for a discussion of how this underpins the distinction between the MEH and the EFH. The test statistics for weak exogeneity, distributed as $\chi^2(2)$ under the null, are 4.90 [0.09] for the first vintage, 6.43 [0.04] for the second vintage and 2.21 [0.33] for the third vintage. Only in the last case is there clear evidence for not rejecting weak exogeneity.

Thus, analysis of the first three vintages suggests the following. The (log) revisions are stationary, implying that the different vintages are homogenous and the third vintage is weakly exogenous; together these imply that in this restricted system the third vintage is the permanent component and the cumulative revisions relative to that are the transitory components. This also supports the view that the revisions are measurement errors orthogonal in the long run to the final vintage. These conclusions are conditional on excluding the final vintage; however, the results in Tables 3 and 5 suggest that an empirical analysis that includes the final vintage is warranted.

Adding the final vintage to the first three vintages, the cointegrating vectors can be just identified as sequential revisions or cumulative revisions as follows:

$$\beta' = \overset{\text{sequential revisions}}{\begin{bmatrix} 1 & -1.000 & 0 & 0 \\ 0 & 1 & -0.996 & 0 \\ 0 & 0 & 1 & -0.927 \end{bmatrix}} \quad \beta' = \overset{\text{cumulative revisions}}{\begin{bmatrix} 1 & 0 & 0 & -0.924 \\ 0 & 1 & 0 & -0.924 \\ 0 & 0 & 1 & -0.927 \end{bmatrix}} \tag{11}$$

The first of these parameterisations suggests that we can impose homogeneity across the first three vintages without loss; this is so with a $\chi^2(2)$ test statistic of 0.41 and a p-value of 82%. With this imposition the cointegrating vectors, in equivalent forms, are:

$$\beta' = \overset{\text{sequential revisions}}{\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -0.924 \end{bmatrix}} \quad \beta' = \overset{\text{cumulative revisions}}{\begin{bmatrix} 1 & 0 & 0 & -0.924 \\ 0 & 1 & 0 & -0.924 \\ 0 & 0 & 1 & -0.924 \end{bmatrix}} \tag{12}$$

We cannot, however, impose homogeneity with respect to the last vintage; the coefficient of -0.924 has an estimated standard error of 0.005, so that the test of homogeneity (that the coefficient is -1) is resoundingly rejected.

The test statistics for weak exogeneity, with a test statistic now distributed as $\chi^2(3)$ under the null, are given in Table 6. The test statistics show that the finding of weak exogeneity for the third vintage was a conditional result. Extending the system, it is only the final vintage that is weakly exogenous. We find that the hypothesis that the final vintage is weakly exogenous is not rejected with a test statistic of 1.89, distributed as $\chi^2(3)$ under the null, with a p-value of 60%. In contrast, for example, the hypothesis that the third vintage is weakly exogenous is firmly rejected with a p-value for the test statistic of zero. Thus, these tests give support to the MEH rather than EFH view of why data revisions arise.

Table 6. Test statistics for weak exogeneity of the vth vintage, $v = 1, \ldots,$ final

v	1	2	3	final
$\chi^2(3)$	12.45[0.00]	14.98[0.00]	15.46[0.00]	1.89[0.60]

4.5. Implications for the PT decomposition

The scaling factor linking the first three and "final" vintages can be read off from either of the matrices in (12); for example, the scaling factor for the first and "final" vintages is -0.924, that is $\beta_1' y_t = y_t^1 - 0.924 y_t^4$ is a stationary series. However, in order to link these different vintages we need to take account of the constant in the cointegration space.

The constants in the VEqCM, μ, can be partitioned into the cointegration space and the common trends space as: $\mu = \alpha\mu_1 + \alpha_\perp \mu_2$; premultiplying by $(\alpha'\alpha)^{-1}\alpha'$ we obtain the constants in the cointegration space, that is $\mu_1 = (\alpha'\alpha)^{-1}\alpha'\mu$. Using the maximum likelihood estimates for α and μ, the three equilibrium relationships are

$$\mu_1 + \beta' y_t = \begin{pmatrix} -0.576 \\ -0.575 \\ -0.574 \end{pmatrix} + \begin{bmatrix} 1 & 0 & 0 & -0.924 \\ 0 & 1 & 0 & -0.924 \\ 0 & 0 & 1 & -0.924 \end{bmatrix} y_t \qquad (13)$$

These parameter values enable a comparison of the first three vintages with the final vintage as they effectively reduce all the vintages to the same price units, in this case 1992 $U.S.

In Figure 2 we plot the stationary combination given by the first row of (12), which, multiplied by 100, can be interpreted as the % revision comparing the first and final vintages. Then in Figure 3 we plot the first vintage and the (scaled) final vintage, using the cointegrating coefficients from (13).

The figures show that revisions to the NIPA have not broken the cointegration bond between the final series and the first published version of GNP, even though the effect of

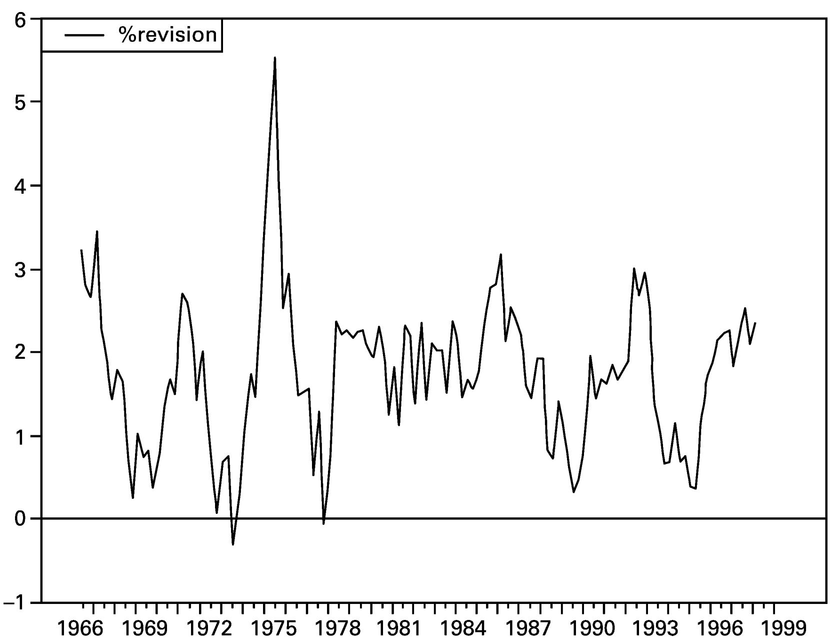

Fig. 2. % revision to GNP, v = 1 to final

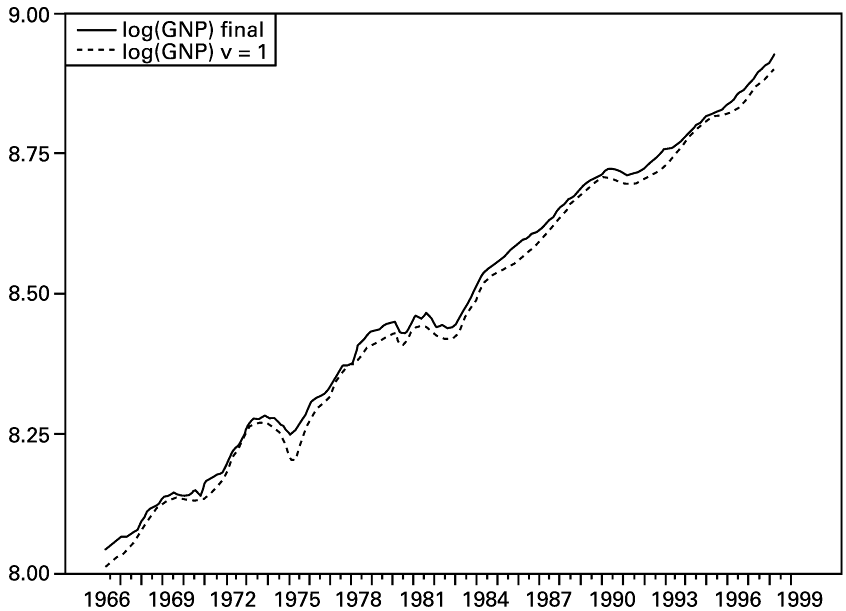

Fig. 3. 1st and final vintages of log(GNP)

recent revisions has been to increase GNP with retrospective effect. (See also Seskin and Sullivan, 2000, who note that recently revised GNP exceeds previously published versions.) One point of note, which accounts for the large revision to the early part of 1975, is how the final version of GNP has "flattened out" the previously substantial dip in GNP during that period.

4.6. Testing for serial correlation common features

In addition to testing for the dimension of the cointegrating space and hence the common trends space, we can also consider whether there are serial correlation common features that reduce the number of common cycles. In view of the results of the previous sections, with $r = m - 1 = 3$, there can be at most one such common feature.

The test statistics for serial correlation common features, reported in Table 7, lead to the rejection of all possible hypotheses; for example, the test statistic for $k > 0$ is 49.7, distributed as $\chi^2(8)$ under the null, which has a p-value of zero. This implies that there is no reduction in the number of common cycles. Thus, the short-run movements in y_t are not common. This finding is open to an interesting interpretation for which we are indebted to

Table 7. Testing for serial correlation common features

Null hypothesis	$C(p, k)$: p-value	Distribution
$k > 0$	49.7: [0.00]	χ^2 (8)
$k > 1$	112.0: [0.00]	χ^2 (18)
$k > 2$	247.2: [0.00]	χ^2 (30)

a referee and which should be the subject of further study. That is, the finding of a single common trend confirms the view that early figures give an accurate picture of the sustained movements in GNP, what commentators refer to in general terms (rather than the specific technical definitions used here) as the “trend” in the series. On the other hand, the failure to find short-run (serial correlation) common features points to some disagreement in the different vintages as to the precise timing of short-run movements, with, perhaps, differences in the timing of turning points and the signs of changes in the level of GNP. This suggests that there should be further research that examines the concordance of the different vintages for indications of common features in some key (nonparametric) measures of short-run alignment, for example agreement in identifying turning points.

5. Summary and Concluding Remarks

As much of the analysis in the previous section is necessarily quite technical, we draw out some of the practical implications in this concluding section.

1. The central idea of this study is that data revisions should be related to each other as part of an overall process, not studied in isolation. It is evident from a graph of U.S. GNP (see Figure 3) that there is a trend to the series – that is, a tendency for a sustained movement in one direction over a reasonably long period of time. This trend may be characterised as stationary deviations around a deterministic trend or, alternatively, as a stochastic trend, being the cumulative sum of past shocks that generate nonstationarity in the data. Our multivariate analysis of this distinction found in favour of the stochastic trend. Now, given that revisions to GNP effectively result in several different time series (the different vintages) for the same variable, and each has a stochastic trend, an important question is whether the various vintages have a single common trend. If not, there are potentially serious problems of interpretation since, necessarily, the different trends are *not* tied together. It is a plausible motivation for the data producers not to produce more than one trend in the different vintages. Fortunately, despite recent changes in GNP methodology, our findings support the single trend interpretation.
2. Another practical matter is whether a particular vintage adequately captures the single common stochastic trend. Our results suggest that a likely presumption of most users, that the “final” vintage is best in this respect, is warranted by our empirical findings. Whether we condition on the third vintage as the final or a later vintage, that vintage by itself can represent the long-run movements in GNP.
3. Finding a common stochastic trend is an issue that relates to the long-run tendencies in the different series, but what of the short run? We focussed on short-run patterns of correlation in the different vintages and found that the hypothesis of a common feature of this kind was rejected. So we can characterise our results as indicating that the different vintages represent the same long-run tendencies but can differ in the short-run picture.
4. Two leading explanations of the reason for revisions relate to errors of measurement and efficient forecasts. In the errors of measurement view, preliminary vintages are inherently “noisier” than later vintages that “iron out” relatively imprecise measurement in the earlier vintages, where the latter are driven by the need for timely publication.

In the efficient forecasts view, preliminary vintages are formed as forecasts of later vintages; moreover, these forecasts fully take into account information available at the time they are being made. In a lesser version of this view, preliminary vintages may be forecasts of later vintages but are not efficient. In practice, an aggregate such as GNP may incorporate both aspects and therefore be the result of measurement with error and forecast with error. With allowance for the nonstationarity of GNP, our empirical findings suggested strongly favouring the measurement error view of revisions.

5. Finally, a practical point arises out of the nature of recent methodological changes to GNP, especially the move to chained price and quantity indices. Preliminary analysis of the data for this study suggested that a method of "splicing" together the time series data from this and the preceding vintages was not going to be straightforward when the 2001 download was included. Was it possible, therefore, to relate GNP on the most recent basis back to previous published series? Part of this question has already been answered, since the presence of a single stochastic trend provides the solution. By using the estimated long-run relationship amongst the different vintages, all the time series can be put onto a single and comparable base. This was illustrated in Figure 3, which graphed vintage 1 of GNP along with the (rebased/rescaled) "final" vintage; as expected from the formal analysis the two series are tied together – a good visual illustration of cointegration.

There are some further questions that could be addressed in future research. For example, Symons (2001), Patterson (2002a) and Swanson and van Dijk (2002) have suggested linking data revisions to the stage of the business cycle. A particularly interesting idea is that the revisions process may be asymmetric with respect to upturns and downturns in the cycle. This asymmetry is one form of nonlinearity, and the analytical framework used here could be extended to explore the possibility of other forms of nonlinearity.

6. References

Abuaf, N. and Jorion, J. (1990). Purchasing Power Parity in the Long Run. Journal of Finance, 45, 157–174.

Ahn, S.K. and Reinsel, G.C. (1988). Nested Reduced-Rank Auto-Regressive Models for Multiple Time Series. Biometrika, 80, 855–868.

Balz, C. (1998). Testing the Stationarity of Interest Rates Using a SUR Approach. Economics Letters, 60, 147–150.

Bartlett, M.S. (1937). Properties of Sufficiency and Statistical Test. Proceedings of the Royal Society of London, Series A, 160, 268–282.

Caplan, D. and Daniel, D. (1992). Improving Economic Statistics. Economic Trends, 460, 87–88.

Cook, P. (1994). Improvements to Economic Statistics, Economic Trends, 483, 133–136.

Croushore, D. and Stark, T. (2001). A Real-time Data Set for Macroeconomists. Journal of Econometrics, 105, 111–130.

Engle, R.F. and Kozicki, S. (1993). Testing for Common Features. Journal of Business and Economic Statistics, 11, 369–395.

Engle, R.F. and Granger, C.W.J. (1987). Co-integration and Error Correction: Representation, Estimation and Testing. Econometrica, 55, 251–276.

Gonzalo, J. and Granger, C.W. (1995). Estimation of Common Long-memory Components in Cointegrated Systems. Journal of Business and Economic Statistics, 13, 27–35.

Granger, C.W. and Haldrup, N. (1997). Separation in Cointegrated Systems and Persistent-Transitory Decompositions. Oxford Bulletin of Economics and Statistics, 59, 449–463.

Grimm, B.T. and Parker, R.T. (1998). Reliability of the Quarterly and Annual Estimates of GDP and Gross Domestic Income. Survey of Current Business, 78, 12–21.

Hansen, H. and Juselius, K. (1995). CATS in RATS: Cointegration Analysis of Time Series. Evanston, IL: Estima.

Hansen, H. and Rahbek, A.C. (2000). Approximate Conditional Unit Root Inference. Discussion Paper, University of Copenhagen.

Hendry, D.F. (1983). Econometric Modelling: the Consumption Function in Retrospect. Scottish Journal of Political Economy, 30, 193–220.

Hendry, D.F. (1994). HUS Revisited. Oxford Review of Economic Policy, 10, 86–106.

Hendry, D.F. (1995). Dynamic Econometrics. Oxford: Oxford University Press.

Holden, K. and Peel, D.A. (1982a). Data Revisions and Time Series Models of GDP and its Components. Applied Economics, 14, 101–110.

Holden, K. and Peel, D.A. (1982b). The Relationship between Preliminary and Revised Estimates of GDP and Consumers' Expenditure in the UK. The Statistician, 31, 259–266.

Issler, J.V. and Vahid, F. (2001). Common Cycles and the Importance of Transitory Shocks to Macroeconomic Aggregates. Journal of Monetary Economics, 47, 449–473.

Jenkinson, G. and Brand, M. (2000). A Decade of Improvements to Economic Statistics. Economic Trends, 558, 45–50.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. Journal of Economic Dynamics and Control, 12, 231–254.

Johansen, S. (1992). Testing Weak Exogeneity and the Order of Cointegration in U.K Money Demand Data. Journal of Policy Modelling, 14, 313–334.

Johansen, S. (1995a). Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford: Oxford University Press.

Johansen, S. (1995b). Identifying Restrictions of Linear Equations with Applications to Simultaneous Equations and Cointegration. Journal of Econometrics, 69, 111–132.

Johansen, S. (1998). Common Trends in the Cointegrated VAR. Unpublished manuscript, May 21st.

Johansen, S. (2001). Cointegration in the VAR model. In D. Pena, G. Tiao, and R. Tsay (eds) A Course in Time Series Analysis. New York: John Wiley and Sons.

Johansen, S. (2002). A Small Sample Correction for the Test of Cointegrating Rank in the Vector Autoregressive Model. Econometrica, 70, 1929–1961.

Johansen, S. and Juselius, K. (1992). Some Structural Hypotheses in a Multivariate Cointegration Analysis of the Purchasing Power Parity and the Uncovered Interest Parity for UK. Journal of Econometrics, 53, 211–244.

Kendall, M.K. and Stuart, A. (1958). The Advanced Theory of Statistics Volume 1. London: Charles Griffin and Company.

Konishi, T. and Granger, C.W.J. (1992). Separation in Cointegrated Systems. UCSD Discussion Paper.

Landefeld, S.J. and Fraumeni, B. (2001). Measuring the New Economy. Survey of Current Business, 81, 23–40.

Mankiw, N.G. and Shapiro, M.D. (1986). News or Noise: An Analysis of GNP Revisions. Survey of Current Business, 66, 961–979.

Mankiw, N.G., Runkle, D.E., and Shapiro, M.D. (1984). Are Preliminary Announcements of the Money Stock Rational Forecasts? Journal of Monetary Economics, 14, 15–27.

Mills, T.C. and Holmes, M.J. (1999). Common Trends and Cycles in European Industrial Production: Exchange Rate Regimes and Economic Convergence. The Manchester School, 67, 557–587.

Mocan, H.N. (1994). Is There a Unit Root in Real U.S. Real GNP? Economics Letters, 45, 23–31.

Mork, R.A. (1987). Ain't Behavin: Forecast Errors and Measurement Errors in Early GNP Estimates. Journal of Business and Economic Statistics, 5, 165–175.

Moulton, B.R. (2000). Improved Estimates of the National Income and Product Accounts for 1929–1999: Results of the Comprehensive Revision. Survey of Current Business, 80, 11–17.

Moulton, B.R. and Sullivan, D.F. (1999). A Preview of the 1999 Comprehensive Revision of the National Income and Product Accounts: New and Redesigned Tables. Survey of Current Business, 79, 15–28.

Moulton, B.R., Seskin, E.P., and Sullivan, D.F. (2001). Annual Revision of the National Income and Product Accounts. Survey of Current Business, 81, 7–32.

Moulton, B.R., Parker, R.P., and Seskin, E.P. (1999). A Preview of the 1999 Comprehensive Revision of the National Income and Product Accounts. Survey of Current Business, 79, 7–20.

Murray, C.J. and Nelson, C.R. (2000). The Uncertain Trend in U.S. GDP. Journal of Monetary Economics, 46, 79–95.

Nelson, C.R. and Plosser, C.I. (1982). Trends and Random Walks in Macroeconomic Time Series. Journal of Monetary Economics, 10, 139–162.

Osterwald-Lenum, M. (1992). A Note on Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistic. Oxford Bulletin of Economics and Statistics, 54, 461–472.

Pantula, S.G. (1989). Testing for Unit Roots in Time Series Data. Econometric Theory, 5, 256–271.

Parker, R.P. (1997). Completion of the Comprehensive Revision of the National Income and Product Accounts, 1929–1996. Survey of Current Business, 77, 6–20.

Patterson, K.D. (1995). An Integrated Model of the Data Measurement and Data Generation Processes with an Application to Consumers' Expenditure. Economic Journal, 105, 54–76.

Patterson, K.D. (2000a). An Introduction to Applied Econometrics. A Time Series Approach. London: Palgrave.

Patterson, K.D. (2000b). Which Vintage of Data to Use When There Are Multiple Vintages of Data? Cointegration, Weak Exogeneity and Common Factors. Economics Letters, 69, 115–121.

Patterson, K.D. (2002a). Modelling the Data Measurement Process for the Index of Production. Journal of the Royal Statistical Society, Series A, 165, 1–18.

Patterson, K.D. (2002b). The Data Measurement Process for UK GNP: Stochastic Trends, Long Memory and Unit Roots. Journal of Forecasting, 21, 245–264.

Patterson, K.D. and Heravi, S.M. (1991a). Are Different Vintages of Data on the Components of GDP Cointegrated? Economic Letters, 35, 409–413.

Patterson, K.D. and Heravi, S.M. (1991b). Data Revisions and the Expenditure Components of GDP. Economic Journal, 101, 877–901.

Patterson, K.D. and Heravi, S.M. (1991c). Direct Estimation of Entropy and Revisions to the National Income Accounts. The Statistician, 40, 35–50.

Patterson, K.D. and Heravi, S.M. (1992). Efficient Forecasts or Measurement Errors? Some Evidence for Revisions to the United Kingdom Growth Rates. Manchester School, 60, 249–263.

Perron, P. and Phillips, P.C.B. (1987). Does GNP Have a Unit Root? A Re-evaluation. Economics Letters, 23, 139–145.

Proietti, T. (1997). Short-run Dynamics in Cointegrated Systems. Oxford Bulletin of Economics and Statistics, 59, 405–422.

Reimers, H.E. (1992). Comparisons of Tests for Multivariate Cointegration. Statistical Papers, 33, 335–359.

Rudebusch, G.D. (1992). Trends and Random Walks in Macroeconomic Time Series: A Re-examination. International Economic Review, 33, 661–680.

Rudebusch, G.D. (1993). The Uncertain Unit Root in Real GNP. American Economic Review, 83, 264–272.

Seskin, E.P. (1998). Annual Revision of the National Income and Product Accounts. Survey of Current Business, 78, 7–35.

Seskin, E.P. (1999). Improved Estimates of the National Income and Product Accounts for 1958–98: Results of the Comprehensive Revision. Survey of Current Business, 79, 15–43.

Seskin, E.P. and Sullivan, D.F. (2000). Annual Revision of the National Income and Product Accounts. Survey of Current Business, 80, 6–33.

Siklos, P.L. (1996). An Empirical Exploration of Revisions in U.S. National Income Aggregates. Applied Financial Economics, 6, 59–70.

Steel, D.G. and McLaren, C.H. (2000). The Effect of Different Rotation Patterns on the Revision of Trend Estimates. Journal of Official Statistics, 16, 61–76.

Stock, J.H. and Watson, M.W. (1988). Testing for Common Trends. Journal of the American Statistical Association, 83, 1097–1107.

Stock, J.H. and Watson, M.W. (1986). Does GNP Have a Unit Root? Economics Letters, 22, 147–151.

Swanson, N.R. and van Dijk, D. (2002). Are Statistical Reporting Agencies Getting It Right? Data Rationality and Business Cycle Asymmetry. Econometric Institute Report EI 2001–28, revised March. Erasmus University, Rotterdam.

Symons, P. (2001). Revisions Analysis of Initial Estimates of Annual Constant Price GDP and its Components. Economic Trends, 568, 40–55.

Taylor, M.P. and Sarno, L. (1998). The Behavior of Real Exchange Rates During the Post-Bretton Woods Period. Journal of International Economics, 46, 281–312.

Vahid, F. and Engle, R.F. (1993). Common Trends and Common Cycles. Journal of Applied Econometrics, 8, 341–360.

Vahid, F. and Engle, R.F. (1997). Codependent Cycles. Journal of Econometrics, 80, 199–221.

Walton, D. (1986). Does GNP Have a Unit Root? Evidence for the UK. Economics Letters, 26, 219–224.

Young, A.H. (1987). Evaluation of GNP Estimates. Survey of Current Business, 67, 18–42.

Zellner, A. (1958). A Statistical Analysis of Provisional Estimates of Gross National Product and its Components, of Selected National Income Components, and of Personal Saving. Journal of the American Statistical Association, 59, 54–65.

Öller, L.E. and Hansson, K.G. (2002). Revisions of Swedish National Accounts, 1980–1998 and an International Comparison. Report to the Swedish Commission on the Review of Economic Statistics, November.

Received January 2001
Revised November 2003