Journal of Official Statistics, Vol. 20, No. 4, 2004, pp. 631–644

Rejoinder

Rebasing, Common Cycles, and Some Practical Implications of Modelling Data Revisions

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

1. Introduction

First, we would like to thank the discussants for their careful reading of our article and their thoughtful and constructive comments. We hope to have encouraged consideration of how to model the process of data revision in a multivariate context and of, at least, some of the implications of data revision for the use of data.

Almost inevitably we seemed to have generated a discussion that suggests another research agenda. Equally as inevitably there is a difference of emphasis between practising national statisticians and, to some extent, those focussed more on short-term developments in the economy, and a research framework that seeks to provide answers to questions that have concerned academic interests, for example the unit root literature. However, we argue here that a conceptualisation of the data measurement process, the DMP, is an essential part of being able to resolve or address wider issues, some of which arise from the short-term needs that often drive the timely publication of national accounts data.

The contributions from practicising statisticians from the statistics offices of Australia, the Netherlands and the United Kingdom, provide a very useful complement to our time series analysis of the data. Inevitably our interests differ, but it is our hope that the discussion arising from our article will help to focus the research agenda of academics more closely on the needs and interests of practising statisticians.

Both those who use macroeconomic data to inform their current view on the state of the economy and researchers, will find invaluable the summaries of the UK and Dutch revision processes provided by Robin Lynch and Craig Richardson (LR) and Peter van de Ven and George van Leeuwen (VL), respectively. Indeed, perhaps both summaries should be compulsory reading for students of economics and lecturers in macroeconomics. These authors are able to cite a number of recent articles (published since the acceptance of our article) that show that the topic of data revisions is clearly a very live and active one in the UK and the Netherlands.

We agree with Dennis Trewin (DT) of the Australian Bureau of Statistics that: “Good quality national accounts will have revisions.” The point is well-put by DT in that national

[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

statistical agencies face a considerable demand for timely data to inform policy decisions but unrevised data are likely to be very poor data. We hope that there is no normative indication taken from our study, and our previous work, on the merits of data revisions. Commentators and researchers alike have clearly benefited from the open process of data revision and publication followed by the statistical agencies in the UK and the U.S. (for example, Patterson 2003, and Patterson and Heravi 1991, and elsewhere). DT also raises the question of which time series to analyse; this is in essence a distinction between what we elaborate on below as revision-stage and real-time vintages.

Donald Egginton (DE) also points out that there is an alternative definition of "vintage" applied to data published in a sequential nature. DE uses data available from the Croushore and Stark (2001) data set to confirm the essence of our results applied to a different data set, but raises some questions about the implications of our finding that the evidence in our sample was against the view of common cyclical behaviour amongst the revisions.

In the sections that follow we take up in more detail some of the issues raised by the discussants; we refer to our article as PH.

2. The Definition of a "Vintage:" Organising and Analysing the Data

The question of how we view the measurement process when data is subject to revision is central to how we then proceed to analyse particular issues. Here we consider the points raised by DT on what data to analyse and on the definition of a vintage.

2.1. Revision-stage vintage and real-time vintage

We have followed the substantial literature on the empirical analysis of data revisions in defining vintage as relating to the stage of revision of the DMP; for clarification we will refer to this as the *revision-stage* definition (amongst many references see, for example, Howrey 1978, Mork 1987 and, more recently, Akritidis 2003, Richardson 2002, 2003). The revision-stage definition picks up on the (approximate) consistency of the *stages* of the measurement process. For example, LR describe the systematic stages involved in the process of measuring and publishing data on UK GDP. This description is particularly useful in showing the progressive nature of the revisions process, which is shared by many national statistical agencies. In the U.S. the stages have involved "flash" estimates, "75-day" estimates through to "benchmark" estimates, and whilst the terminology differs, the general concept does not.

Another use of the word vintage in this general context is to capture the idea of a *real-time* vintage (see, for example, Croushore and Stark 2001) rather than a *revision-stage* vintage. A real-time vintage is simply the run of data available, in a particular publication, at a particular point in time. For example, we could look at the middle of each month following the end of a quarter and list the data available (published) for a series, generically denoted y, at that point in time. We might then do the same at a quarterly or some other interval for y and keep a record of the various historical series; this would be a real-time vintage data set.

Although in the first instance a definition is what we make it, it will have relevance to the analytical framework that is constructed to explain the DMP. In order to understand the nature of the two definitions, we refer to the framework for organising the different data series used in Patterson and Heravi (1991). Therein we suggested the following conceptual basis: construct a matrix with row dimension T, indexing by t the observational period to which the data refers, and column dimension m, indexing by v the vintage or *stage* of the data measurement process.

The observational period is not the same as the publication period – or (availability) date – as there is a lag, ϑ, in producing the data. We will stylise the process somewhat to emphasise the salient points. First published (for example "preliminary" or "flash") data for observational period t is, therefore, placed in the t-th row and first column; this is referred to as y_t^1; at the next publication date this first vintage data is revised to y_t^2 and, usually simultaneously, the first vintage for observation period $t+1$ is published, and this is y_{t+1}^1. Following this process through creates a $\mathrm{T} \times m$ matrix, which we refer to below as $\mathrm{D_T}$, with an interesting feature.

When the first vintage is published for the last available observation period, here referred to as T, so that the last data is y_T^1, the data for the second and subsequent vintages for period T are not yet available; the second vintage is available for $\mathrm{T}-1$, y_{T-1}^2; the third vintage is available for $\mathrm{T}-2$, y_{T-2}^3, and so on until the last available data at $\mathrm{T}-(m-1)$ with $y_{T-(m-1)}^m$. This traces back a diagonal of data from T to $\mathrm{T}-(m-1)$; the principle applies to any of the observational periods, so that along the diagonal at any observational period t is the data available at $t+\vartheta$. (The diagonal is stepped if different vintage data is not published simultaneously.)

As the diagonal is taken from *m different* columns of the data matrix it comprises data at different stages (revision-stage vintages) of the measurement process. At time $t+\vartheta$, the data that is published by the statistical agency in a single publication source is this diagonal plus, once the last vintage has been reached for observation period $t-(m-1)$, data in the m-th column, that is "final" vintage data. It is data of this type that are referred to as "real-time" data and a set of observations at time $t+\vartheta$ as a real-time vintage. These data can be organised into a data matrix, denoted $\mathrm{R_T}$, which is a $\mathrm{T} \times \mathrm{T}$ matrix, with the row dimension indexing the observation period t to which the data relate, and the column dimension indexing the real-time vintage defined as the date on which the observations were published.

In the case of the real-time definition, there are as many "vintages" as there are availability dates $t+\vartheta$, $t = 1, \ldots, \mathrm{T}$; these real-time vintages have the characteristic of ending in the observational period t and comprise the latest available data, which is often the source of data for topical commentators and forecasters, see the comments by DT. If the data at vintage m is *actually* the final published data, then the two series of real-time vintages $t+\vartheta$ and $(t-1)+\vartheta$ will comprise the sequences

$$\text{real-time vintage } t+\vartheta: \quad \left\{y_t^1,\ y_{t-1}^2, \ldots, y_{t-(m-1)}^m,\ y_{t-m}^m, \ldots, y_1^m\right\} \tag{1}$$

$$\text{real-time vintage } (t-1)+\vartheta: \quad \left\{y_{t-1}^1,\ y_{t-2}^2, \ldots, y_{t-(m-1)}^{m-1},\ y_{t-m}^m, \ldots, y_1^m\right\} \tag{2}$$

Three points can be observed from this characterisation:

i) data prior to and including $t - m$ is common to both series;
ii) the real-time vintage for $t + \vartheta$ has one more observation than the real-time vintage for $(t - 1) + \vartheta$;
iii) given points (i) and (ii), a comparison of these two real-time vintages is, in essence, a comparison of the $(m - 1)$-length sequences $\{y^2_{t-1}, \ldots, y^m_{t-(m-1)}\}$ and $\{y^1_{t-1},\ y^2_{t-2}, \ldots, y^{m-1}_{t-(m-1)}\}$, whereas a comparison of revision-stage vintages, say vintages 1 and 2, will involve a comparison of the $(\mathrm{T} - 1)$-length sequences $\{y^1_{t-1}, \ldots, y^1_{t-(m-1)},\ y^1_{t-m}, \ldots, y^1_1\}$ and $\{y^2_{t-1}, \ldots, y^2_{t-(m-1)},\ y^2_{t-m}, \ldots, y^2_1\}$. The analysis of real-time vintages will have to bear in mind that m is, generally, not large.

We hope it is clear from this framework that the definitions associated with revision-stage vintages and real-time vintages use the data organised in different ways; whether one definition is better than another is not a relevant question unless related to the purpose of the analysis. Primarily, we have analysed the revision-stage vintages, which is consistent with our view that the data originates from an m-dimensional, not a T-dimensional, measurement process, which we implicitly take as stable over time.

In practice, there may be structural breaks in the process (and m may not be constant); for example, LR suggest the potential implementation of a feedback process that first identifies weak areas of initial data response or collection and then remedies the deficiency. Strictly this would vitiate the stability of the revision-stage vintages assumption; however, to some extent if breaks in the process are present they can be dealt with by the same methods used to rebase the data. (We consider below the problem of rebasing data, raised in part by DE's Comment 4.)

In our analysis, we chose not to focus on the real-time vintages, regarding the implications for these as following or being implied by the analysis of revision-stage vintages. Of course, if a convincing argument can be made that the data measurement process relates conceptually to the real-time vintages, not the revision stage vintages, the analytical basis of our work would have to be revised. In any case as the data matrices $\mathrm{D_T}$ and $\mathrm{R_T}$ use the same data, we observe that one is a linear transformation of the other, specifically $\mathrm{D_T} = \mathrm{R_T}\Phi$, where Φ is a $\mathrm{T} \times m$ transformation matrix of rank m. (We adopt the convention of using a 0 where data is unavailable.) Thus, although we did not consider it explicitly in our article, in principle an analysis of the data in terms of $\mathrm{D_T}$ has implications for $\mathrm{R_T}$; this has been addressed in Patterson (1995, 2003) and we hope to consider it further in future research. (See also Section 3 below.) Unless specifically indicated otherwise, our further reference to vintage means the revision-stage vintage.

2.2. *Rebasing and related issues*

The analysis of the outcomes of the revisions process often focuses on growth rates of variables rather than levels as considered in our article. Whilst interest may centre on growth rates, the measurement process relates in the first instance to levels of the variables of interest. These are often the subject of economic analysis as for example in the literature on whether the process generating U.S. GNP is better characterised as difference or trend stationary, in the building of economic models for simulation and forecasting and in a league ranking of income per capita.

Also, a consistent data set for levels enables an analysis of growth rates that should be of interest. Our analysis does not, however, focus solely on the levels to the exclusion of growth rates. The VEqCM (vector equilibrium correction mechanism) framework that we adopt in PH is based on the transformation of $I(1)$ variables into $I(0)$ variables but in such a way that information from the $I(1)$ variables is not discarded; put less technically, trended variables are reduced by transformation to nontrended variables but the underlying trend(s) still have a role to play. In our case the $I(1)$ variables are the log-levels of the revision-stage vintages of U.S. GNP, which are reduced to $I(0)$ by first differencing. This means that it is the growth rates that are the "dependent" variables in the estimation framework, and these, in the VEqCM framework, depend upon revisions to vintages provided that these revisions are stationary (not trended).

2.2.1. Rebasing

One of the reasons that many studies of the properties of revisions focus on growth rates, is the difficulty associated with the rebasing of series onto a constant price basis in order to obtain a levels series that is internally consistent over the length of the series. (However, using growth rates does not entirely avoid these problems, as we show below.) There are several changes of base in our data set (see DE Table 1, who conveniently sets out the base change years) and we have therefore had to consider how to accommodate such changes. The method we used is described in detail in Patterson and Heravi (1991), and we give a brief summary here. In particular it includes as a special case the method used by DE to convert the Croushore and Stark (2001) data for comparison with our results.

Our notation for the data now needs to distinguish the base on which data is published; thus $y^v_{t,i}$ is the observation for period t of vintage v data on base i, where $i = 1, \ldots, I$; for example base i could be data on a 1982 price basis, whereas base $i - 1$ is data on a 1972 price basis. The rebasing problem occurs because once a new price basis is established new data is published on the new price basis not the old. Thus one can observe in the data matrix D_T that, at the time a new base is introduced, say $t = b_i$, the "trace-back" diagonal, which is data for the most recent time period and previous time periods but for different vintages, is published on the *new* not the old basis. Thus, for comparison of data for the same time period for adjacent vintages but different bases, for example $y^2_{b_i-1,i}$ with $y^1_{b_i-1,i-1}$, a conversion of the data onto one base is needed. With $I = 6$ there are $I - 1 = 5$ changes of base over the sample period.

A simplifying assumption is that rebasing can be modelled as comprising a constant plus a rescaling effect; Rushbrook (1978) addresses some of the issues arising at rebasing times. That is

$$y^v_{t,i} = \alpha_{i-1} + \beta_{i-1} y^v_{t,i-1} + \omega^v_{t,i} \quad v = 1, \ldots, m \tag{3}$$

where $\omega^v_{t,i}$ is a stationary, zero mean error arising because, conceptually, the rebasing relationship is not exact. Also note that the assumption is that the rebasing coefficients are constant for each t and v. This assumption, whilst usual, requires some comment. For example, it implies that the rebasing coefficients are the same whether we are considering the relation between say $y^1_{t,i}$ and $y^1_{t,i-1}$ or $y^m_{t,i}$ and $y^m_{t,i-1}$. This assumption, whilst usual, needs to be checked, at least informally, and it is the one that led us to decide not to rebase data preceding the move to 1996 prices onto the new 1996 price base. However, for the present argument we maintain the assumption of constant coefficients.

Note that (3) is not, generally, a directly observable relationship because data for time period t is *not* issued on both base i and base $i-1$. As it stands it is just a conceptualisation of the relationship between data on different bases. Nevertheless, a simplified version of this underlies what often happens in practice when there is an overlap of data on the same observation period but not the same vintage. To illustrate assume, for simplicity, that the constant is zero, then (3) for vintage v is:

$$y^v_{t,i} = \beta_{i-1} y^v_{t,i-1} + \omega^v_{t,i} \tag{4}$$

The "splicing" method estimates β_{i-1} by:

$$b_{i-1} = y^v_{t,i} / y^{v-1}_{t,i-1} \tag{5}$$

That is, b_{i-1} is the ratio of data for t using vintage v, with base i data, to data for the same period but of vintage $v-1$ on base $i-1$. (In effect it is the ratio of adjacent horizontal cell entries in D_T, at the point where the base has changed.) This method implies the equality $y^v_{t,i} = b_{i-1} y^{v-1}_{t,i-1}$ rather than the equality in (4) obtained by setting $\omega^v_{t,i} = 0$.

Note that on the assumption that the method holds for all $v = 1, \ldots, m$, any of the adjacent, or indeed nonadjacent, vintages could be used. In practice using different values of v often results in different estimates of b_{i-1} and an average of such values might well be used.

The "splicing" method, however, confounds the rebasing and vintage effects; setting $\omega^v_{t,i}$ to its expected value of zero, (4) implies $\beta_{i-1} = y^v_{t,i} / y^v_{t,i-1}$, which differs in the denominator compared to b_{i-1}. From (5) and (3) we have:

$$b_{i-1} = \frac{y^v_{t,i}}{y^v_{t,i-1}} \frac{y^v_{t,i-1}}{y^{v-1}_{t,i-1}} = \left(\mathrm{K}^{v,v-1}_{t,i-1} \right) \beta_{i-1} \tag{6}$$

where $\mathrm{K}^{v,v-1}_{t,i-1} = y^v_{t,i-1} / y^{v-1}_{t,i-1}$; hence, only if there is no vintage effect for time t data on base $i-1$ data will it be the case that $b_{i-1} = \beta_{i-1}$. Further, as there is a tendency for upward revisions, that is $y^v_{t,i-1} > y^{v-1}_{t,i-1}$, then $b_{i-1} > \beta_{i-1}$.

In order to use (3), or without a constant (4), we require a model of the link between vintages for the same period and on the same base. Suppose this can be summarised as the linear combination of a constant and a vintage-scaling effect, thus:

$$y^v_{t,i} = a^{v-1}_0 + a^{v-1}_1 y^{v-1}_{t,i} + \varepsilon^{v-1}_t \tag{7}$$

The coefficients a^{v-1}_0 and a^{v-1}_1 are vintage dependent but are assumed not to be base dependent; ε^{v-1}_t is a stationary (vintage-specific) random error with zero mean. This relationship also holds for base $i-1$, so that:

$$y^v_{t,i-1} = a^{v-1}_0 + a^{v-1}_1 y^{v-1}_{t,i-1} + \varepsilon^{v-1}_t \tag{8}$$

Note that the scaling factor $\mathrm{K}^{v,v-1}_{t,i-1}$ in (6) can be obtained by dividing (8) by $y^{v-1}_{t,i-1}$, which, assuming that $(a_0^{v-1}/y_{t,i-1}^{v-1})$ and $(\varepsilon_t^{v-1}/y_{t,i-1}^{v-1})$ are negligible, gives:

$$b_{i-1} \approx \beta_{i-1} a_1^{v-1} \tag{9}$$

Upward revisions will be characterised by $a_1^{v-1} > 1$ and hence $b_{i-1} > \beta_{i-1}$.

Even if the error arising from linking data on an adjacent base is quite slight, it can be important in an overall sample where the data is rebased several times. To see this point consider linking data across two base changes; thus, in the simplified case, we have $y_{t,i}^v = \beta_{i-1} y_{t,i-1}^v$ and $y_{t,i-1}^v = \beta_{i-2} y_{t,i-2}^v$, which imply:

$$y_{t,i}^v = \beta_{i-1} \beta_{i-2} y_{t,i-2}^v \tag{10}$$

If the splicing method is used to estimate each linking coefficient, β_{i-1} and β_{i-2} in this case, there is the danger that a small error for each coefficient will cumulate into a large error overall.

The method we used avoids the confusion between vintage and rebasing effects. It effectively uses (3), or (4) in the simplified case, and the relationship between vintages on the same base given by (7) specified for m and v. That is, substituting (3) into (7) for m and v, we have:

$$y_{t,i}^m = a_0^v + a_1^v \left(\alpha_{i-1} + \beta_{i-1} y_{t,i-1}^v \right) + u_{t,i}^{v-1} \tag{11}$$

where $u_{t,i}^{v-1} = \varepsilon_t^{v-1} + a_1^v \omega_{t,i}^v$ and $v = 1, \ldots, (m-1)$; $t = \mathrm{T} - (m-1)$; $i = 2, \ldots, I$.

We found the rebasing constants to be unimportant and thus only the rebasing slope coefficients were retained. Estimation is efficient in the sense of using all available observations, that is $(m-1) \times [\mathrm{T} - (m-1)]$ in total, and imposing the nonlinear constraints that typical coefficients are $a_1^v \beta_{i-1}$ for different values of v and i.

This method seeks to use information in the complete sample of vintages and observations to estimate the rebasing coeffcients, rather than, for example, pick adjacent vintages at a particular observation point and use the resulting coefficient b_{i-1}. It is predicated on, what is in effect, a cointegrating relationship between the vintages as in (7) and (8); in principle, this could be directly incorporated into the estimation framework described in Section 3 of our article, rather than form a separate and prior stage.

As a check, albeit informal, on the overall rebasing procedure, we recommend examining the resulting rebased data across all vintages, rather than accepting the outcome of the estimation procedure (or splicing method if that has been used) rather mechanically whatever that may have been. If the rebasing is successful it should not be possible to tell where the underlying base changes were made. The presence of "jumps" in any of the rebased series, which for example are not present in different vintage data for the same t to $t-1$ comparison, may indicate difficulties with the underlying assumptions of the rebasing procedure. It was for this reason that we decided not to use the rebasing coefficients that included the last (in our sample) change of base. The practical outcome, therefore, was a set of vintages on one base and a conditionally final vintage on a different base; that is, $\{y_{t,I-1}^1, \ y_{t,I-1}^2, \ y_{t,I-1}^3\}$ and $\{y_{t,I}^4\}$, respectively. For simplicity the reference to the base was omitted from the notation.

One assumption we made explicit earlier was that rebasing coefficients are constant for each t and v; we have suggested that this was a difficult assumption to sustain for the last base change in our sample. In principle, we could recognise this in the rebasing relationship summarised by (3); for example we could allow the rebasing coefficients to be vintage specific so that:

$$y^v_{t,i} = \alpha^v_{i-1} + \beta^v_{i-1} y^v_{t,i-1} + \omega_t \quad v = 1, \ldots, m; \;\; i = 2, \ldots, I \tag{12}$$

In this case there are now, for example, m rebasing slope coefficients for each base change. Whether any progress could be made along these lines, and indeed what the rationale would be, was not an issue we considered.

2.2.2. Growth rates

Where the rebasing is successful, growth rates can be consistently defined for each vintage. This is of interest because much of the analysis of the characteristics of revisions, and topical discussion of developments in the economy, is carried out with growth rates rather than levels. If the levels data have been rebased then, with lower case letters denoting logs in our notation, the growth rate for vintage v is $\Delta y^v_{t,I-1} \equiv y^v_{t,I-1} - y^v_{t-1,I-1}$, where we have kept the explicit dependence on the base. If the data have *not* been rebased, then an approximation is required which again confounds vintage and rebasing effects. One possibility is to define the growth rate as $\Delta y^v_{t,i-1} \equiv y^v_{t,i-1} - y^{v+1}_{t-1,i-1}$; this suggestion involves a comparison of two adjacent vintages (generally available at the same time in the publication source) and takes advantage of the diagonal "traceback" in $\mathrm{D_T}$ at rebasing times. On this definition, the last consistently defined growth rate is $\Delta y^{m-1}_{t,i-1} \equiv y^{m-1}_{t,i-1} - y^m_{t-1,i-1}$; a final growth rate, not consistent with this definition, could be defined as $\Delta y^m_{t,i-1} \equiv y^m_{t,i-1} - y^m_{t-1,i-1}$.

2.2.3. Data redefinitions

A problem that looks similar in the data to rebasing may arise with retrospective data redefinitions. Our conceptual framework is that it is the stages of the data "production" process that serve to define the key distinctions on which to base analysis and so define the m-dimensional joint probability density function, $f(y^1_t, \ldots, y^m_t)$. However, redefinitions induce nonstationarities by way of structural breaks in $f(y^1_t, \ldots, y^m_t)$, and thus cause problems in assuming that the data on $(y^1_t, \ldots, y^m_t)$ was generated by the same process throughout the sample period. If the break points are known, this information can be accommodated by introducing dummy variables directly into the cointegration analysis, for example into (1) of PH. Alternatively, where the redefinition is retrospective and, say, occurs at $t = \mathrm{T}_1$, it is usually applied to all data issued at $\mathrm{T}_1 + \vartheta$, and later, for vintage 1 through to m. The redefinition thus traces a diagonal back along the data matrix, $\mathrm{D_T}(t - i, \;\, i + 1)$, for $i = 0, \ldots, m - 1$, where the first index refers to the observation period and the second to the vintage. If this is the case, the effect of the redefinition can be modelled in the same way as a regular rebasing of the data.

3. Multivariate Extensions

VL suggest that it may be informative to the revisions process to extend the cointegration analysis to a multivariate setting that considers the components of GDP/GNP. This is a welcome suggestion and one that we have already, in part, made some progress on, as we describe briefly below. However, one of the difficulties in extending the analysis is the rather practical one of obtaining the necessary data sets; it would undoubtedly be helpful if national statistical offices were to make available (electronically) an historical databank of vintages (either revision-stage or real-time) for the components of GNP. In previous work Patterson and Heravi (1991) considered the components of UK GDP; the analysis was made possible by the (much-appreciated) provision of the necessary data by the UK's national statistical office.

The suggestion in VL is that an analysis of data revisions should take into account the inter-relationships between variables. Patterson (2003), has developed a model based multivariate framework, illustrated therein for U.S. income and consumption, which could be combined with the state-space framework developed in a previous paper (Patterson 1995). However, one should not underestimate the demands this makes both on data and on modelling capabilities, hence the results so far tend to be bivariate illustrations of general techniques.

The "complete" model comprises two parts: a model of the data generating process, that is a model of the behavioural links between variables, and a model of the data measurement process which allows for the possibility of linked data revisions. We illustrate some of the considerations with the extension of the univariate DGP in PH to a bivariate framework, which is a special case of the general multivariate case in Patterson (2003).

For simplicity we assume that there are m vintages of data on each of two variables generically denoted y_1, and y_2, for example "income" and "consumption," respectively. The notation is:

$$y_t = (y_{1t}, y_{2t}) = \left(y_{1t}^1, y_{1t}^2, \ldots, y_{1t}^m, y_{2t}^1, y_{2t}^2, \ldots, y_{2t}^m\right) \tag{13}$$

The data vector y_t is now $2m \times 1$ comprising m vintages on the first variable and m vintages on the second variable. Apart from this change of dimension, the VeqCM of Equation (1) in PH again provides the cointegration framework.

This extension enables the potential of a more complex set of relationships: a) within the m vintages of each variable and b) between each generic variable. As to a) for each variable, full intravariable cointegration refers to the idea of a single common trend driving all m vintages. On this basis, taking the two variables together there will be $2(m-1)$ cointegrating vectors and therefore 2 common trends. As to b), if there is intervariable cointegration, then there is an additional cointegrating vector, so that the cointegrating rank is $r = 2(m-1)+1$ and hence there is one less common trend.

An alternative way of viewing this situation is that there is one common trend within each (generic) variable, hence $r_i = m - 1$ for each variable, $i = 1,\ 2$, and the 2 common trends cointegrate so that there is just one single common trend driving the complete set of $2m$ vintages of data.

Complete cointegration is said to occur when there is complete intravariable cointegration and also intervariable cointegration. Cases other than complete cointegration are interesting and may well occur. For example, suppose $m = 4$, $n = 2$ and we find $r = 6 < 2(m-1)+1$, that is there are two common trends (r is the total cointegrating rank). A possible explanation is that the final vintage for each variable has been substantially revised, breaking the cointegration bond *within* each variable. Nevertheless, intervariable cointegration is maintained across the first three vintages and separately across the final vintage.

If there is intravariable cointegration of order $r_i = m - 1$, $i = 1,\ 2$, for each variable set, but no intervariable cointegration between the common trends, then $r = 2(m - 1)$ and the two generic variables are said to be weakly separated. Other cases may also be of interest. For example, suppose that there is less than full intravariable cointegration rank for both variables but there is intervariable cointegration. A potential explanation for this is that the first (or other preliminary) vintage of data is a poor predictor of the subsequent vintages, since the revisions are not stationary, but there is intervariable cointegration between the poor predictors. In the context of the national accounts, cointegration amongst poor predictors may arise from contemporaneous joint revisions of the component aggregates in the process of reconciling the accounts.

The technical details of the state-space representation of the joint model of the DGP *and* the DMP are beyond the scope of this response. However, the essence is to use the information obtained from the multivariate (across variables and vintages) cointegration analysis and any serial correlation properties of the measurement errors to simulate or forecast *any* vintage of data, whether revision-stage or real-time. This means that we could construct a forecast of a real-time vintage not just a revision-stage vintage, which addresses the point raised by DT.

4. Subjective Assessments and Statistical Significance

DE focuses in his Section 7 and conclusion on short-run movements in GNP. In PH we noted that "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." Without further research we would not elevate the issues of locating turning points and the signs of changes to levels, to the status of facts and would not claim to have done so. The interesting question is how our suggested research on the short-run properties of the vintages might be structured.

One possibility illustrated by DE is more informal than technical and based on a practitioner's experience of what is noteworthy. That is, a data user (researcher/commentator/practitioner) decides, a priori, what difference in, say growth rates, would be considered as numerically rather than statistically significant on some measure; then counts the number of times such a difference is exceeded; and he/she finally decides whether this number is important.

In principle different data users may well choose different parameters in this assessment, so that however valid it might be to each individual it does not offer a general solution to such problems. For example, DE notes 46 (31%) instances in his data set where the differences between the first and third vintages of quarter on quarter growth rates for GDP/GNP are larger than $\pm 0.2\%$ (% point). The magnitude of these revisions seems large to us, compared to an average quarter on quarter growth rate of 0.73% (approximately 2.95% p.a, in our data set); also the proportionate number of times that the change is greater than this, at 31%, seems quite large. Nevertheless, the distinction behind an assessment of what is numerically significant compared to statistical significance may correctly draw a distinction between the probability of an outcome and the expected value of that outcome.

One possible framework is to link the empirical analysis more closely to a decision-making outcome in order that the gains and losses of different realisations can be assessed; for example, in a European context, budget contributions are linked to GNP and account for approximately 45% of the overall EU budget, which is itself set at a ceiling of 1.24% (2002) of Community Gross National Income (1.27% on the previous measure of Community Gross National Product). A revision, therefore, has a calculable effect on the change in the Member States EU budget contribution. Also of interest, in this context, is whether revisions of one sign are more likely and should be built into anticipated budgetary contributions. Also, within a country a revision to GNP implies a revision to (net) tax receipts, which may well have calculable effects.

Although the difference between statistical significance and numerical significance can well be an important one, the former is generally still of interest. Thus even if interest centres on the number of revisions/residuals larger in absolute value than, say, μ, it is likely to be helpful to know how likely such an occurrence is under a particular null hypothesis.

We illustrate with a very simple nonparametric "sign" test that can also be interpreted as a test of proportions; this is just an example, much more could be done, either nonparametric or parametric. The nonparametric description refers to the absence of assumptions regarding the distribution of the underlying quantity of interest, which we take as revisions to the growth rate of GNP based, respectively, on the first and third vintages, that is g_1 and g_3, with individual observations g_{1t} and g_{3t}.

The sign test is based on the sign of the differences between the sequence $\{g_{1t} - g_{3t}\}$ and an unknown parameter μ_c. In the simplest case $\mu_c = 0$, and the underlying hypothesis is that the medians of the population distributions of g_1 and g_3 are equal. When $\mu_c \neq 0$, the medians are equal only after a location shift of μ_c; hence, define $x_t = (g_{1t} - g_{3t})$, then the median of the distribution of $z_t = x_t - \mu_c$ is zero. Thus, under the null there is a probability of 1/2 of $z_t > 0$ and a probability of 1/2 of $z_t < 0$. The sign test just uses the number of times, T_+, in the sample that $z_t > 0$; alternatively normalising T_+ by the total number of sample observations, T, that is define $\hat{\pi} = T_+/T$, the test can be viewed as a test of proportions; either way the *p*-value is computed by comparison with the null distribution, which is binomial(T, 1/2) or a normal distribution approximation valid for large samples. The parameter μ_c is unknown, and in the sign test it is replaced by μ varying in the set $\mu \in \Lambda = [\alpha, \beta|\delta]$, where α is the lower limit, which we take to be zero (the conventional null), β is the upper limit, the increment of variation is δ and Λ should be wide enough to include μ_c. This procedure enables *p*-values to be calculated for all $\mu \in \Lambda$.

Although it is of interest to consider a test based on the sign of x_t, DE's implicit hypothesis relates to the absolute values of the differences in growth rates, that is the sequence of $|x_t| = |g_{1t} - g_{3t}|$ rather than $g_{1t} - g_{3t}$; the motivation being that a large *negative* revision is as worrying to a practitioner as a large *positive* revision. This suggests redefining the variable of interest as $z_t = |x_t| - \mu_c$; then μ_c is the median of the distribution of absolute differences and $\hat{\pi}(\mu)$ is the proportion of the sample for which $|x_t| > \mu$. A practitioner may then select a particular μ and a corresponding 'proportion of concern', $\pi(\mu)_0$, so called because if a proportion $\pi(\mu)_0$, or larger, of the sample observations exceed μ, then the practitioner expresses "concern." The hypothesis testing framework can be used to assess the statistical significance of the difference between the corresponding sample proportion $\hat{\pi}(\mu)$ and $\pi(\mu)_0$. As different practitioners may choose different $\pi(\mu)_0$ for given μ, the *p*-value information from the sample can be expressed as a function of different values of $\pi(\mu)_0$; the procedure could also be carried out for different values of μ.

To illustrate, we use annualised growth rates, from our data set, for the first and third vintages. The value of μ corresponding to $\hat{\pi}(\mu) = 0.5$ is between 1.05 and 1.1 (the slight imprecision arises from the discreteness of the data); μ is here the revision in the growth rate between the first and third vintages in units of % points p.a. The value 1.05 is an estimate of the location shift parameter μ_c for the distribution of $|x_t|$. An annual rate of 1.05% p.a corresponds to a quarterly rate of about 0.26% p.q; and by definition 50% of revisions are greater than this.

Suppose that the key numerical figure for an individual's assessment is $\pm 0.2\%$ p.q $=$ $\pm 0.8\%$ p.a (approximately), and a particular individual would be concerned if 30% of revisions exceeded this value. The null hypothesis for the individual is $\pi(\pm 0.8)_0 = 0.3$, but the data has $\hat{\pi}(\pm 0.8) = 0.61$; the test of the difference between these proportions has a *p*-value of 0. It is (completely) unlikely that the data was generated with $\pi(\pm 0.8) = 0.3$, and the practitioner should be "concerned." However, given that we are introducing a subjective element into the assessment, it is more informative to plot the *p*-values for different values of $\pi(\mu)_0$, and this is done in Figure 1a for $\pi(\pm 0.8)_0$ and in Figure 1b for $\pi(\pm 1.5)_0$, revisions of $\pm 0.8\%$ p.a. and $\pm 1.5\%$ p.a, respectively. In the former case a *p*-value of, for example, 10% corresponds to $\pi(\pm 0.8)_0 = 0.545$, whereas in the latter case for the same *p*-value $\pi(\pm 1.5)_0 = 0.286$.

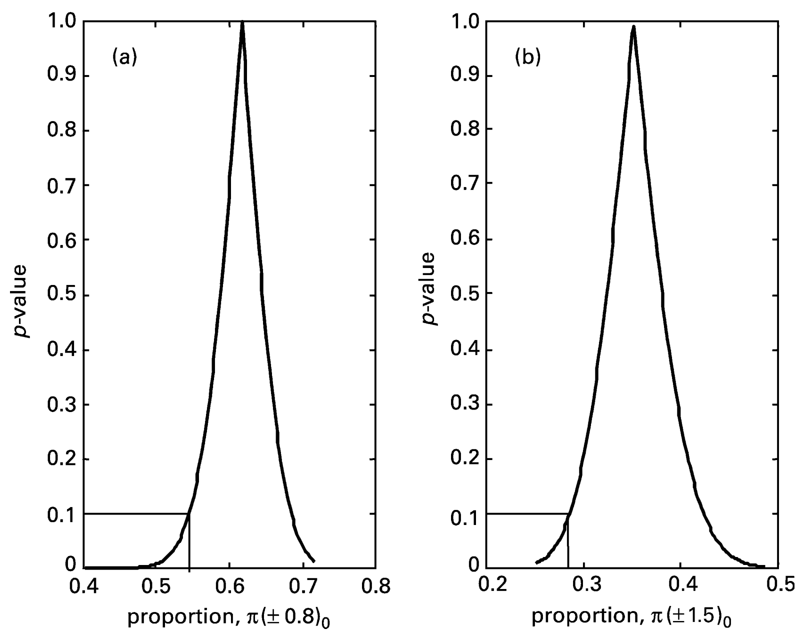

Fig. 1a and 1b. Graphing the p-values for different values of $\pi(\mu)_0$

The use of a sign or proportions test is a simple example using the statistical information in the sample. Tests could also be formulated for sequences, for example a "runs" based test would look at the number of sequences of the same sign relative to the number expected under a particular specification of the null hypothesis. Tests based on directional analysis are also likely to be useful and have attracted much interest in other areas, see, for example, Granger and Pesaran (2000), Pesaran and Timmerman (1992, 2001) and Greer (2003).

5. Concluding Remarks

The comments from the discussants have been very valuable not only in gathering together constructive views on our article but also in allowing some insight into what data agencies and practitioners regard as interesting areas of research. This response has selected some areas for elaboration and comment. Further research is planned to consider some of the other issues and some not explicitly mentioned (for example the effect of revisions to seasonal components on seasonally adjusted data, cyclical components of revisions and alternative conceptual frameworks). Central to our approach, whether it concerns the long-run or short-run properties of a particular set of revisions, is the view that the analysis of data revisions should be based on a conceptual model of the revisions process.

6. References

Akritidis, L. (2003). Revisions to Quarterly GDP Growth and Expenditure Components. Economic Trends, December, 69–85.

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

Granger, C.W.J. and Pesaran, M.H. (2000). Economic and Statistical Measures of Forecast Accuracy. Journal of Forecasting, 19, 537–560.

Greer, M. (2003). Directional Accuracy Tests of Long-term Interest Rate Forecasts. International Journal of Forecasting, 19, 291–298.

Howrey, E.P. (1978). The Use of Preliminary Data in Economic Forecasting. The Review of Economics and Statistics, 60, 193–200.

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.

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. (2003). Exploiting Information in Vintages of Time-series Data. International Journal of Forecasting, 19, 177–198.

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

Pesaran, M.H. and Timmerman, A. (1992). A Simple Non-parametric Test of Predictive Performance. Journal of Business and Economic Statistics, 10, 461–465.

Pesaran, M.H. and Timmerman, A. (2001). How Costly Is it to Ignore Breaks when Forecasting the Direction of a Time Series? Forthcoming in the International Journal of Forecasting.

Richardson, C. (2002). Revisions to GDP: A Time Profile. Economic Trends, July, 1–8.

Richardson, C. (2003). Revisions Analysis: A Time Series Approach. Economic Trends, December, 86–89.

Rushbrook, J.A. (1978). Rebasing the National Accounts – Why and How and the Effects. Economic Trends, March, 105–110.

Received October 2004