Testing for news and noise in non-stationary time series subject to multiple historical revisions*

Alain Hecq

Department of Quantitative Economics, School of Business and Economics,

Jan P.A.M.Jacobs†

Faculty of Economics and Business, University of Groningen,

University of Tasmania, Australia, CAMA, Australia, CIRANO, Canada

Michalis P. Stamatogiannis

Management School, University of Liverpool,

This version: February 2019

Abstract

This paper focuses on testing non-stationary real-time data for forecastability, i.e., whether data revisions reduce *noise* or are *news*, by putting data releases in vector-error correction forms. To deal with historical revisions which affect the whole vintage of time series due to redefinitions, methodological innovations etc., we employ the recently developed *impulse indicator saturation* approach, which involves potentially adding an indicator dummy for each observation to the model. We illustrate our procedures with the U.S. real GNP/GDP series of the Federal Reserve Bank of Philadelphia and find that revisions to this series neither reduce noise nor can be considered as news.

Keywords: data revision, cointegration, news-noise tests, outlier detection
JEL-code: C32, C82, E01

*The article has benefited from comments received following presentations of previous versions at the Inaugural Society for Economic Measurement Conference, Chicago, the 16th OxMetrics User Conference, Aix-Marseille University, Aix-en-Provence, France and the 9th International Conference on Computational and Financial Econometrics (CFE2015), London. In particular we would like to thank Neil Ericsson, Jean-Pierre Urbain, Simon van Norden, and the Editor and two referees for very useful comments. The usual disclaimer applies.

†Corresponding author at: Faculty of Economics and Business, University of Groningen, PO Box 80 0, 970 0 AV Groningen, The Netherlands. Tel.: + 31 50 363 3681; fax: + 31 50 363 7337. E-mail address: j.p.a.m.jacobs@rug.nl (J.P.A.M. Jacobs).

1 Introduction

Before being considered definitive, data currently produced by statistical agencies typically undergo a recurrent revision process resulting in different releases of the same phenomenon. The collection of all these vintages is referred to as a *real-time data set*. In the recent past, economists and econometricians have come to realize the importance of this type of information for economic modeling, forecasting and policy formulation. Consequently there exists a growing interest for investigating this type of data (see *inter alia* Croushore and Stark, 2001, Orphanides and van Norden, 2002, and Croushore, 2011a, 2011b).

Several aspects of real-time data can be investigated: (i) structural or trend breaks (see Jacobs and van Norden (2016) for a summary of the reliability of productivity growth rate trends); (ii) forecastability, i.e., whether revisions reduce *noise* or are *news* (the literature is briefly reviewed in Section 3.1); (iii) *historical revisions*, which affect the whole vintage of time series due to redefinitions, methodological innovations, etc., make testing difficult. The standard approach to dealing with historical revisions is either to employ growth rates to mitigate the effects of historical revisions, or to 'clean' the series in an attempt to get rid of the effects of historical revisions. The former approach has been criticized by Knetsch and Reimers (2009). Götz, Hecq and Urbain (2016) illustrate that growth rates can also be affected by large revisions.

Whereas the tests and the procedures to deal with historical revisions are well-documented for stationary time series (e.g., using Mincer-Zarnowitz type tests), the situation is less clear for non-stationary time series. The paper aims at filling this void, building upon Hecq and Jacobs (2009). We focus on testing forecastability for non-stationary real-time data, putting data releases in vector-error correcting forms (VECMs hereafter). To deal with forecastability under historical revisions at unknown dates, we estimate VECMs using an automatic modelling method for selecting conditional mean parameters (the Autometrics algorithm, see Hendry and

Doornik, 2014) together with the Impulse Indicator Saturation approach (IIS hereafter, see e.g., Hendry and Santos, 2005). Briefly, IIS involves adding an indicator dummy[1] for potentially each observation to the model and hence is able to determine a parsimonious model that fits model requirements in terms of misspecification. We illustrate our procedures with the U.S. real GNP/GDP series of the Federal Reserve Bank of Philadelphia and find that, in general, revisions neither reduce noise nor can be considered as news. Conclusions would have been different without the IIS approach.

An alternative strategy to the IIS algorithm consists of introducing dummy variables for each historical revision. This operation is less obvious than one might think at first glance and can be very tedious and time consuming for an external researcher who does not have complete information on thousands of economic variables for different countries. While one can easily find the description of the modifications for the main aggregates for the U.S. or the European Union for instance, this task is much more demanding when the information about data revisions is for instance not in English or not available online on national statistical agencies websites. Using IIS helps in investigating those time series within a few seconds. Secondly, one can also notice that the date at which vintages are released might differ from the date at which the series has been theoretically modified. As an example, books may describe that there is a new definition of an economic indicator in January but the series published on, say the 10th of January, still applies the old definition. It might be for this latter example that a second vintage is available at the end of the month such that we observe multiple vintages for one particular month, a situation that adds difficulties for the researcher. Third, IIS can also capture smaller revisions (e.g., annual or seasonal revisions due to e.g., the change of seasonal factors) that would have been ignored based on historical revisions only. Finally, many real-time databases have been build manually, either by merging files or using manpower for

[1]We leave for further research the use of additional step dummies in the IIS framework.

scanning or copying figures from statistical reports. Those operations can also introduce errors. So identifying historical revisions is not always straightforward and as a result "dummying-out" specific dates without using a statistical method can be considered as a subjective practice.

The remainder of the paper is structured as follows. After a brief introduction of data revisions and notations in Section 2, Section 3 describes news-noise tests for stationary and non-stationary real-time data as well as the intuition underlying the IIS approach. Section 4 illustrates our procedure with the U.S. Real GNP/GDP series. Section 5 concludes.

2 Data revisions and notation

Real-time data are typically displayed in the form of a real-time data trapezoid as in Figure 1. We move to later vintages as we move across columns from left to right and we move to later points in time when we move down the rows. Note that the frequency of vintages needs not necessarily correspond to the unit of observation; for example, in our illustration below the statistical agency publishes monthly vintages of quarterly observations. In this paper we investigate the releases, namely the diagonals of the data trapezoid. We use the following notation: superscripts refer to releases $i = 1, ..., v$, while subscripts refer to periods; y_t^1 denotes the first release for variable y in period t, whereas the sequence $\{y_t^1\}_{t=1}^T$ or simply y_t^1, $t = 1, \ldots, T$ refers to as the whole time series for the first release, namely the first diagonal in Figure 1.

Data revisions may be conveniently categorized into three types:

1. initial revisions in the first few vintages,

2. annual (seasonal) revisions due to updated seasonal factors and the confrontation of quarterly with annual information, and

Figure 1: The real-time data trapezoid

$$\begin{bmatrix} y_1^1 & \cdots & y_1^i & \cdots & y_1^v \\ & \ddots & \vdots & \ddots & \vdots \\ & & y_{t-l}^1 & \cdots & y_{t-l}^i \\ & & & \ddots & \vdots \\ & & & & y_t^1 \end{bmatrix}$$

3. historical or comprehensive revisions, related to changes in statistical methodology, etc.

The distinction of revisions into these types requires careful handling of the real-time data and in many cases direct access to the officials of the statistical agency. Initial and seasonal revisions are regular and recurring, and can in principle be modeled and forecast. As an example, Eurostat releases its first estimate of e.g., real GDP 45 days after the end of the corresponding quarter (flash estimate); the next release is 15 days later. Historical revisions are much more difficult to handle. Redefinitions like changes of base years do not cause many difficulties, however changes in definitions changes do.

Whatever their origins, data revisions imply the existence of measurement errors. The modeling of measurement errors has two main traditions that are surveyed in the next section.

3 Method

3.1 News-noise tests

Stationary data

The older tradition, which is still widespread among statisticians, is that measurement errors should be thought of as noise. Data are measured with errors which are orthogonal to true values ($\widetilde{y}_t$). This implies for a stationary time series y_t that for all releases i we have

$$y_t^i = \widetilde{y}_t + \zeta_t^i, \qquad \text{cov}(\widetilde{y}_t, \zeta_t^i) = 0. \tag{1}$$

One implication of this is that revisions will generally be forecastable by taking weighted averages of previous releases. To test whether measurement errors reduce noise, the Mincer-Zarnowitz (1969) test can be used, which regresses the revision $y_t^{CV} - y_t^i$ on a constant and the most recent, i.e., the current (last observed column) vintage y_t^{CV}, taken as measure of the unobserved true value $\widetilde{y}_t$

$$y_t^{CV} - y_t^i = \delta_1 + \beta_1 y_t^{CV} + \zeta_t^i. \tag{2}$$

The null hypothesis that measurement errors are independent of true values ($\delta_1 = 0,\ \beta_1 = 0$) may be tested with a Wald test; since the errors may suffer from heteroskedasticity and autocorrelation, robust HAC standard errors are typically used.

The newer tradition, motivated by Mankiw, Runkle and Shapiro (1984), Mankiw and Shapiro (1986) and de Jong (1987), describes measurement errors as news.[2] News errors imply that published data are optimal forecasts, so revisions are orthogonal to earlier releases and are not forecastable. More precisely,

$$y_t^{CV} = y_t^i + \nu_t^i, \qquad \text{cov}(y_t^i, \nu_t^i) = 0. \tag{3}$$

The analogous test of the "news" model regresses the revision ($y_t^{CV} - y_t^i$) on a constant and the ith-release

$$y_t^{CV} - y_t^i = \delta_2 + \beta_2 y_t^i + \nu_t^i. \tag{4}$$

[2]See also the more recent contributions of Faust, Rogers and Wright (2005), Swanson and van Dijk (2006) and Aruoba (2008). More references are in Jacobs and van Norden (2011).

A similar null hypothesis ($\delta_2 = 0, \ \beta_2 = 0$) now tests whether data revisions are predictable. The two null hypotheses are mutually exclusive but they are not collectively exhaustive, i.e., we may be able to reject both hypotheses, particularly when the constant in both test equations differs from zero (see Aruoba, 2008, Appendix A.2). The main conclusion of the empirical literature on characteristics of real-time data is that macroeconomic time series are in principle not well-behaved. Revisions can be substantial and reduce noise or add news at different horizons.

An alternative way to test for news and noise is to estimate the Jacobs and van Norden (2011) data revision model, a state-space form in which measurement errors are decomposed into news and noise with the possibility of spillovers. Recently, Clements and Galvão (2013) extended the Jacobs and van Norden (2011) framework by allowing for revision bias. The alternative state-space forms of Cunningham et al. (2012) and Kishor and Koenig (2012) should in principle be able to do the same. Fixler and Nalewaik (2009) propose an alternative test, whose properties still have to be explored. Finally, the multi-period survey approach of Patton and Timmermann (2011) belongs in this category too.

Non-stationary data

Testing measurement errors in case of non-stationary variables is more complicated even when a single time series, like gross national product, is considered. Indeed, the existence of cointegration between different releases hampers the application of Mincer-Zarnowitz tests explained above for two reasons. First, the presence of cointegration implies that there exists a long-run relationship between different releases and hence news/noise tests would be subject to the usual omitted variable problem if we estimate (4) or (2) on the growth rates of time series only. Second, assuming that we correctly account for cointegrated I(1) series in VECM systems, the issue still remains that we cannot establish the direction of causality, i.e., whether the first release is explained by the final release, or the other way around. However, weak exogeneity tests in cointegrated systems (see Urbain, 1992, 1995) can be helpful here.

Cointegration between time series of different releases, or *intra-variable* cointegration, can be modeled in two ways.[3] The approach most frequently adopted in the literature looks at releases on an observation basis, for example first and second releases of the non-stationary variable y_t observed on $T+1$ data points. The Observation Balanced System (OBS hereafter) tests for cointegration between series y_t^1 and y_t^2, $t = 1, \ldots, T$. Superscripts denote respectively the first and the second released diagonals. It must be understood though that we take the first two releases as a convenient explanatory example but that we investigate the relationships between several releases in this paper.[4]

The alternative approach compares the releases on a vintage basis, i.e., the two most recent observations of vintages. In the Vintage Balanced System (VBS hereafter) cointegration between y_t^1 and y_{t-1}^2, $t = 1, \ldots, T$, is considered. Patterson (2000) is a typical example of the OBS approach, whereas Garratt et al. (2008, 2009) adopt VBS.

Note that, if y_t is integrated of order one, OBS and VBS are equivalent in terms of the cointegration property because of the identity

$$y_t^1 - y_{t-1}^2 \equiv (y_t^1 - y_{t-1}^1) + (y_{t-1}^1 - y_{t-1}^2)$$

and such that $(y_t^1 - y_{t-1}^1)$ is I(0).

Weak exogeneity tests in OBS can reveal whether revisions reduce noise or add news. By exploiting the Gonzalo-Granger (1995) permanent-transitory decomposition, Patterson (2002, 2003) shows that if the final release is weakly exogenous for the parameters of the system then measurement errors in OBS cointegration are noise. For an alternative way to see this consider a bivariate VECM of order one in

[3] The remainder of this section draws upon Hecq and Jacobs (2009).

[4] We leave the multivariate investigation of the whole set of releases for further investigations. In this paper we only look at pairwise tests.

first differences for OBS releases as

$$\begin{pmatrix} \Delta y_t^2 \\ \Delta y_t^1 \end{pmatrix} = \alpha\beta' \begin{pmatrix} y_{t-1}^2 \\ y_{t-1}^1 \end{pmatrix} + \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} \begin{pmatrix} \Delta y_{t-1}^2 \\ \Delta y_{t-1}^1 \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \tag{5}$$

with no deterministic terms to simplify the notations and where the errors are i.i.d. Gaussian with zero mean and variance matrix

$$\Omega = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}.$$

Vectors α and β are respectively 2×1 loading and cointegrating vectors; short-run dynamic matrices Φ_1 and Φ_2 are of dimension 1×2. If the final release Δy_t^2 can be treated as weakly exogenous, i.e., $\alpha_1 = 0$, valid inference in the OBS VECM can proceed in the conditional model of Δy_t^1 given Δy_t^2 and the past, i.e., the second equation of the VECM becomes

$$\Delta y_t^1 = \lambda \Delta y_t^2 + \alpha_2 \left(y_{t-1}^2 - y_{t-1}^1 \right) + \tilde{\Phi}_2 \begin{pmatrix} \Delta y_{t-1}^2 \\ \Delta y_{t-1}^1 \end{pmatrix} + \tilde{\varepsilon}_{2t},$$

where $\lambda = \sigma_{21}/\sigma_1^2$, $\tilde{\Phi}_2 = \Phi_2 - \lambda\Phi_1$ is a 1×2 vector and $\tilde{\varepsilon}_{2t} = \varepsilon_{2t} - \lambda\varepsilon_{1t}$, see Johansen (1995, Chapter 8). This equation can be interpreted as a 'noise' equation, because the final release enters as regressor in the equation of the first release. If this requirement holds, i.e., the hypothesis $\alpha_1 = 0$ is not rejected, the noise null hypothesis $H_0: \ \lambda = 1$ can be investigated.

Analogously, we can rearrange the bivariate VECM as a news equation if the first release is weakly exogenous, i.e., $\alpha_2 = 0$. The 'news' equation becomes

$$\Delta y_t^2 = \mu \Delta y_t^1 + \alpha_1 \left(y_{t-1}^2 - y_{t-1}^1 \right) + \tilde{\Phi}_1 \begin{pmatrix} \Delta y_{t-1}^2 \\ \Delta y_{t-1}^1 \end{pmatrix} + \tilde{\varepsilon}_{1t},$$

where $\mu = \sigma_{12}/\sigma_2^2$, $\tilde{\Phi}_1 = \Phi_1 - \mu\Phi_2$ and $\tilde{\varepsilon}_{1t} = \varepsilon_{1t} - \mu\varepsilon_{2t}$, and if the null hypothesis that $\alpha_2 = 0$ is not rejected, we can test the news null hypothesis $H_0 : \mu = 1$.

3.2 Impulse-indicator saturation

From the previous subsection it emerged that the hypothesis

$$\alpha_1 = 0 \text{ or } \alpha_2 = 0,$$

i.e., that the loading coefficients are zero, is crucial in our setting. From the outcome of those tests we determine the news/noise prevalence of the revision process, although additional restrictions on coefficients have to be tested. By construction however, the diagonals of the OBS in its VECM representation are going to be affected by the presence of outliers due to historical revisions, the modification of base dates and redefinitions of the data, updated seasonal factors, etc.

It is not possible—and often not even feasible—without a very good understanding of the series under study and without an insider knowledge of the exact effect of the revision process to rebase the entire series at each vintage dates before constructing the diagonals. The presence of such aberrant values that one creates in the diagonals are going to seriously affect the behavior of our test statistics of the loadings αs.

There exist several ways (either parametric or non-parametric) to identify and to robustify a regression for the presence of such outliers. In this paper we rely on the recent literature on IIS (see Castle, Doornik, and Hendry, 2008; Santos, Hendry and Johansen, 2008; Johansen and Nielsen, 2009; Ericsson and Kamin, 2009; Ericsson and Reisman, 2012). IIS involves adding an indicator dummy variable for each observation to the model. In the simplest case, namely a regression for the $I(0)$ univariate time series Δy_t without any additional explanatory variables nor step dummies this leads to

$$\Delta y_t = \delta_0 + \delta_1 I_1 + \delta_2 I_2 + \ldots + \delta_T I_T + u_t,$$

a model with $T+1$ parameters for T observations (and more generally a model with $T+K$ parameters for T observations), which cannot be estimated.

However, in the essence of the IIS approach the dummies can be added in blocks. In general IIS splits the sample in blocks of $T/2$ observations each and adds an impulse dummy for every observation in that block of $T/2$ observations; significant outcomes for a chosen significance level, say 5%, are retained. Then one drops that set of impulse indicators and proceeds similarly on the other half of the sample, with the significant outcomes retained. Finally one combines the recorded impulse indicators obtained in both parts and those that remain significant when both dummies from both parts are added, are selected. This procedure is implemented in *Autometrics* (Doornik and Hendry, 2013), where the algorithm makes it possible to estimate such a model, performing a joint selection over dummy variables and other regressors.

We apply this approach to VECM systems for $(\Delta y_t^i, \Delta y_t^{i+l})$ and compare the selection of the parsimonious systems with and without IIS.

4 Illustration

We consider the real GNP/GDP (ROUTPUT) series, seasonally adjusted, of the Federal Reserve Bank of Philadelphia Real-Time Data Set. This indicator of economic activity for the U.S. is available quarterly since 1965:Q4 with recorded vintages starting in 1947:Q1. Historical (comprehensive) revisions are documented for the vintages 1976:Q1, 1981:Q1, 1986:Q1, 1992:Q1, 1993:Q1, 1996:Q1, 1997:Q2, 1999:Q4, 2000:Q2, 2004:Q1, 2009:Q3 and 2013:Q3. See specific notes on data collected for real GNP/GDP, which is available under the header documentation at the ROUTPUT internet page. Information on the most recent historical revisions is provided by the Bureau of Labour Statistics (BLS) at Information on previous updates of the NIPA accounts.

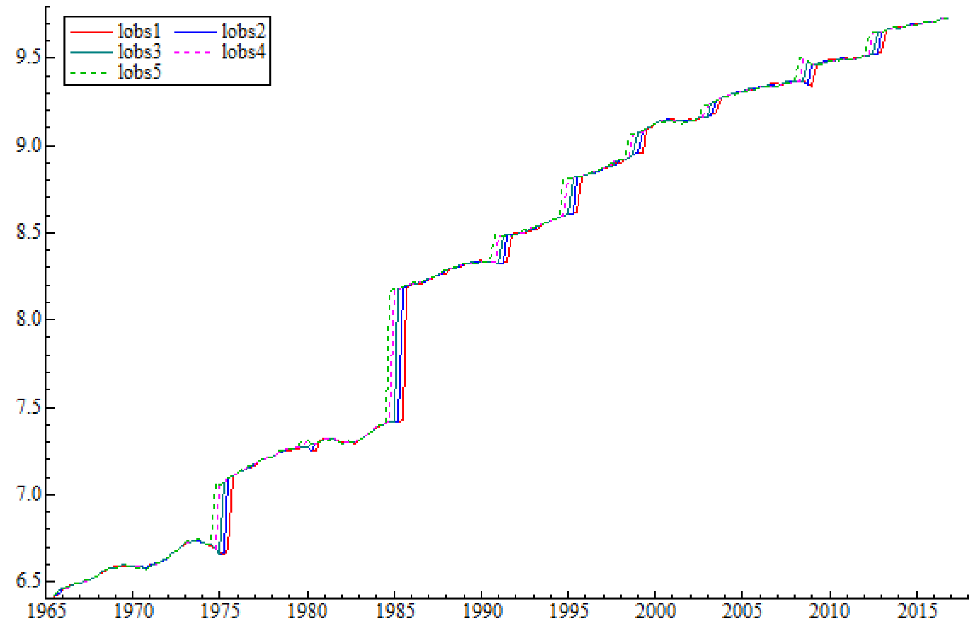
We observe four vintages per year and clean the data such that we have a nice trapezoid with regular steps. To do so we delete columns associated with vintages which are not first releases and when the next observation has the same value. We perform the analysis with vintages 1965:Q4 up to and including 2018:Q1. Hence we work with a maximum of $T = 210$ quarterly observations for the period 1965:Q3–2018:Q1. Note that the data is published with a lag of one quarter. We apply the logarithmic transformation on the data.[5]

Figure 2 displays the first five releases, i.e. diagonals of our real-time data trapezoid, corresponding in our case to the quarterly vintages. The presence of breaks (jumps) around the modification of bases is obvious. We also observe that historical revisions associated with the vintages 1981:Q1, 1993:Q1, 1997:Q2 and 2000:Q2 are small. Historical revisions affect the series of releases in different periods. The most recent historical revision for example, the one associated with vintage 2013:Q3, results in a step-wise jump in 2013:Q2 in the series of first releases (because of the publication lag), in 2013:Q1 in the series of second releases, etc. First differences of the first releases show a spike in 2013:Q2, of the second releases in 2013:Q1, etc.

We begin with a detailed analysis of the relationship between the first and the second releases y_t^1 and y_t^2 of U.S. real GNP/GDP. We show how the outcomes can be different while using or not the IIS approach. We first assume that the cointegrating vector between them[6] is $(1, -1)$ and estimate bivariate VECM models with $z_t = y_t^1 - y_t^2$. We use Autometrics without IIS to determine the lag length and the significance of the loading of each equation in the VECM. The estimation results are presented in equations (6) and (7) along with standard diagnostic tests. Values in (.) correspond to t-ratios, whereas values in [.] contain the p-values of the corresponding

[5] We also ran the analyses for levels instead of log levels. Results are available upon request.
[6] Note that this assumption is subsequently tested in Johansen's maximum likelihood context.

Figure 2: First five OBS diagonals (releases)



statistics. The results support rejection of $H_0 : \alpha_1 = 0$ (significance of the first loading) but not $H_0 : \alpha_2 = 0$, suggesting that the equation can be used for a noise test.

$$\Delta y_t^1 = \underset{(4.5)}{0.004} - \underset{(-6.03)}{0.69}\, z_{t-1} + \underset{(1.55)}{0.167} \Delta y_t^1, \tag{6}$$

$$R^2 = 0.99$$

$$\text{AR[1-5]:}F(5,94) = 2.495[0.033];\ \text{ARCH[1-4]:}\ F(4,94) = 4.730[0.001]$$

$$\text{Normality:}\chi^2_{(2)} = 29.863[0.000];\ \text{RESET}: F(2,97) = 0.227[0.797]$$

$$\Delta y_t^2 = \underset{(3.41)}{0.016}, \tag{7}$$

$$\text{AR[1-5]:}F(5,177) = 0.111[0.990];\ \text{ARCH[1-4]}: F(4,189) = 0.019]$$

$$\text{Normality:}\chi^2_{(2)} = 6721.9[0.000]$$

The diagnostics of Equation (6) suggest that the equation is well-specified, in contrast to Equation (7), except for the normality of the residuals which is strongly rejected in both equations. Figures 3 and 4 present the estimated scaled residuals of equations (6) and (7), respectively. The existence of outliers corresponding to historical revisions is clearly evident, especially in the second equation of the system, and motivates the use of the IIS method.

Figure 3: Scaled residuals of Equation (6); Autometrics is used only for the selection of lag length.

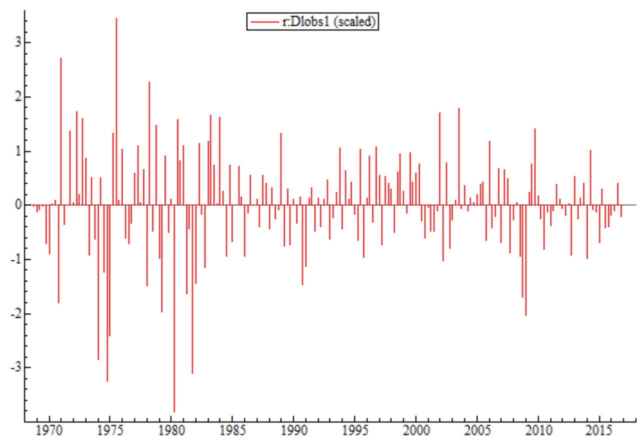
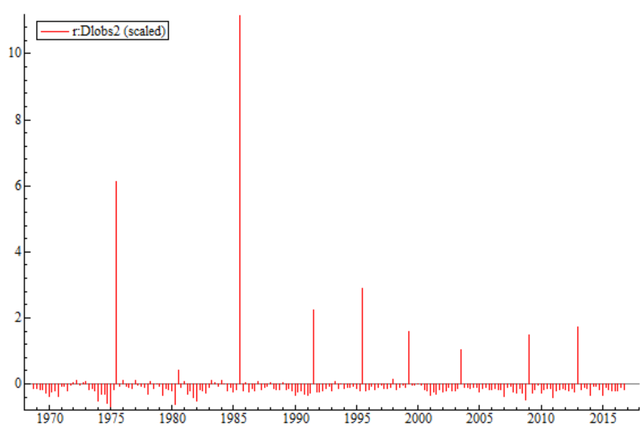


Figure 4: Scaled residuals of Equation (7); Autometrics is used only for the selection of lag length.



We proceed with estimating the VECM with the use of Autometrics and IIS. The results are presented in Equations (8) and (9).

$$\begin{aligned} \Delta y_t^1 =& \underset{(7.89)}{0.004} \underset{(-10.9)}{-0.562}\, z_{t-1} + \underset{(8.57)}{0.434} \Delta y_{t-1}^2 + 11 \text{ dummies}, \\ R^2 =& 0.99 \end{aligned} \tag{8}$$

AR[1-5]: $F(5, 175) = 0.703[0.622]$; ARCH[1-4]: $F(4, 186) = 1.597[0.177]$

Normality: $\chi^2_{(2)} = 4.375[0.112]$; RESET: $F(2, 178) = 0.268[0.765]$

$$\begin{aligned} \Delta y_t^2 =& \underset{(5.43}{0.382} z_{t-1} + \underset{(5.43)}{0.376} \Delta y_{t-1}^2 + 19 \text{ dummies}, \\ R^2 =& 0.99 \end{aligned} \tag{9}$$

AR[1-5]: $F(5, 165) = 1.461[0.205]$; ARCH[1-4]: $F(4, 186) = 0.944[0.440]$

Normality: $\chi^2_{(2)} = 5.807[0.055]$; RESET: $F(2, 168) = 4.301[0.015]$

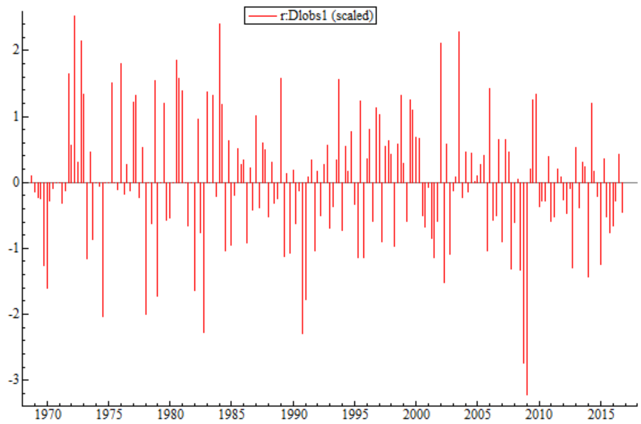
Using IIS with a small significance level[7] of 1%, we identify 11 dummies in the estimated equation for the first release Δy_t^1 (1970:Q4, 1971:Q1, 1974:Q1, 1974:Q4, 1975:Q1, 1975:Q3, 1978:Q2, 1979:Q2, 1980:Q2, 1981:Q2, 1981:Q4) and 19 dummies for the equation for the second release Δy_t^2 (1974:Q1, 1974:Q4, 1975:Q1, 1975:Q3, 1978:Q2, 1980:Q2, 1980:Q3, 1982:Q1, 1983:Q2, 1984:Q1, 1985:Q3, 1991:Q3, 1995:Q3, 1998:Q1, 1999:Q2, 2003:Q3, 2008:Q4, 2009:Q1, 2013:Q1). The dummies for the first equation capture relatively small outliers in the beginning of the sample. Effects of historical revisions are captured by the error-correction term z_{t-1} which has spikes in the same periods as Δy_t^1. The first half of the 19 dummies of the second equation more or less correspond to the dummies of the first equation. Most of the the other dummies are clearly linked to historical revisions, taking into account the one-quarter publication lag of the vintages and the fact that second releases are affected

[7] A higher significance level (such as 5%) yields a much higher number of significant dummies, resulting in a severe reduction in degrees of freedom of each equation.

by historical revisions with a lag. For example, the historical revision in vintage 2013:Q3 yields the 2013:Q1 dummy in the equation for Δy_t^2.

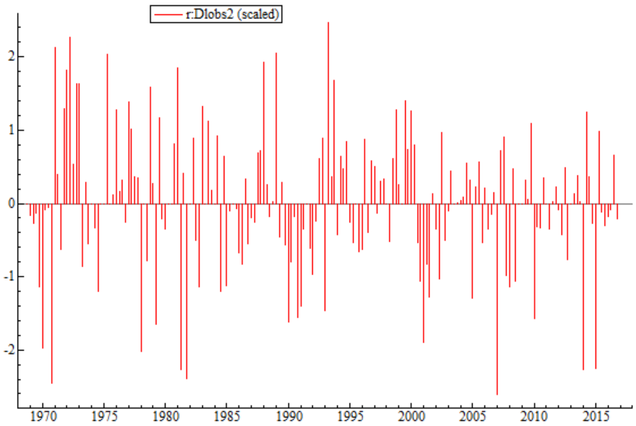
The introduction of the IIS substantially improves the estimation results of the VECM, as there is much less evidence of misspecification. Figures 5 and 6 which present the estimated residuals of Equations (8) and (9) respectively, and do not give any evidence of outliers related to historical revisions. It is obvious that the use of IIS has greatly improved the statistical adequacy of the model. However, inference regarding the hypotheses of interest changes: both hypotheses $\alpha_1 = 0$ and $\alpha_2 = 0$ are rejected which implies that the bivariate VECM for releases one and two can neither be rearranged as a pure noise equation nor as a pure news equation.

Figure 5: Scaled residuals of Equation (8) with both Autometrics and IIS for the detection of outliers.



Next we perform the same analysis for all possible combinations of the five releases that we have considered in this illustration, taking lower releases in the first equation of the system and higher releases as second. In each case the lag length of the VECMs and the significance of the loadings is determined by Autometrics. We run the estimation of each combination, with and without IIS in order to see the effect of this method to the inference. The dummies found significant by the IIS method are listed in Table 1. The first two columns with header $\ln(y_t^1, y_t^2)$ show the dummies for the combination of the first releases and second releases analysed

Figure 6: Scaled residuals of Equation (9) with both Autometrics and IIS for the detection of outliers



above. The inclusion of dummies for the equations of the other combinations follow a similar pattern. In all cases the number of outliers is higher for the second equations of the combinations. Dummies in first equations of the VECM are typically not linked to historical revisions. Dummies for second equations more or less include the IIS dummies obtained for first equations. The large majority of the remainder match dates of historical revisions, see for example the dummy 2013:Q1 appearing in the combination (y_t^1, y_t^2), the dummy 2012:Q4 in (y_t^1, y_t^3), the dummy 2012:Q3 in (y_t^1, y_t^4), and the dummy 2012:Q2 (y_t^1, y_t^5), which are related to the historical revision in 2013:Q3. There is one exception. The IIS method does not pick up the comprehensive revision in 1981:Q1 in the combinations (y_t^1, y_t^3) and (y_t^2, y_t^3).

Table 1: Dummies resulting from the IIS method. The first column of each case refers to the dates for which the impulse dummies are found significant for the first equation of the VECM, whereas the second column refers to the dummy variables found significant for the second equation of the VECM. Dates printed in bold are clearly linked to historical revisions.

$\ln(y_t^1, y_t^2)$		$\ln(y_t^1, y_t^3)$		$\ln(y_t^1, y_t^4)$		$\ln(y_t^1, y_t^5)$		$\ln(y_t^2, y_t^3)$	
1970(4)			1970(4)				1970(4)		1970(4)
1971(1)			1971(1)					1971(1)	1971(1)
		1972(2)		1972(2)			1972(2)		
1974(1)	1974(1)	1974(1)	1974(1)		1974(1)	1974(1)	1974(1)	1974(1)	1974(1)
1974(4)	1974(4)	1974(4)	1974(4)		1974(4)	1974(4)	**1974(4)**	1974(4)	1974(4)
1975(1)	1975(1)	1975(1)	1975(1)	1975(1)	**1975(1)**		1975(1)	1975(1)	1975(1)
			1975(2)						**1975(2)**
1975(3)	**1975(3)**	1975(3)	1975(3)	1975(3)	1975(3)		1975(3)	1975(3)	1975(3)
						1975(4)			
						1976(2)			
						1976(3)			
		1977(2)							
		1978(1)							
1978(2)	1978(2)						1978(2)	1978(2)	1978(2)
									1978(4)
1979(2)		1979(2)				1979(2)			
							1979(4)		
					1980(1)				
1980(2)	1980(2)	1980(2)		1980(2)	1980(2)	1980(2)	1980(2)	1980(2)	
	1980(3)								
		1980(4)							
				1981(1)					
1981(2)							1981(2)		1981(2)
		1981(3)							
1981(4)			1981(4)		1981(4)	1981(4)	1981(4)		1981(4)
	1982(1)	1982(1)			1982(1)	1982(1)	1982(1)		
	1983(2)		1983(2)						1983(2)
	1984(1)								
							1984(4)		
					1985(1)				
			1985(2)						**1985(2)**
	1985(3)								
						1985(4)			
						1986(1)			
							1990(4)		
					1991(1)				
			1991(2)						**1991(2)**
	1991(3)								
							1994(4)		
					1995(1)				
			1995(2)						**1995(2)**
	1995(3)								
	1998(1)								
							1998(3)		
					1998(4)				
			1999(1)						**1999(1)**
	1999(2)								
							2001(2)		
					2001(3)				
							2002(4)		
					2003(1)				
			2003(2)						**2003(2)**
	2003(3)								
							2008(2)		
					2008(3)				
	2008(4)	2008(4)	**2008(4)**		2008(4)		2008(4)		**2008(4)**
	2009(1)	2009(1)	2009(1)		2009(1)	2009(1)	2009(1)		
							2010(2)		
							2012(2)		
					2012(3)				
			2012(4)						**2012(4)**
	2013(1)								

Table 1 continued.

$\ln(y_t^2, y_t^4)$		$\ln(y_t^2, y_t^5)$		$\ln(y_t^3, y_t^4)$		$\ln(y_t^3, y_t^5)$		$\ln(y_t^4, y_t^5)$	
1970(1)									
1970(4)				1970(4)	1970(4)	1970(4)		1970(4)	
			1971(1)	1971(1)	1971(1)	1971(1)	1971(1)	1971(1)	1971(1)
1972(2)									
1974(1)	1974(1)	1974(1)	1974(1)	1974(1)	1974(1)	1974(1)	1974(1)	1974(1)	1974(1)
1974(4)	1974(4)		**1974(4)**	1974(4)	1974(4)	1974(4)	**1974(4)**	1974(4)	**1974(4)**
1975(1)	**1975(1)**	1975(1)	1975(1)	1975(1)	**1975(1)**	1975(1)			
			1975(2)	1975(2)		1975(2)	1975(2)		1975(2)
1975(3)	1975(3)	1975(3)	1975(3)	1975(3)	1975(3)	1975(3)	1975(3)	1975(3)	1975(3)
1977(2)						1977(2)			
1978(1)						1978(1)			
1978(2)			1978(2)	1978(2)	1978(2)	1978(2)	1978(2)	1978(2)	1978(2)
		1978(4)							
			1979(4)				**1979(4)**		**1979(4)**
	1980(1)				**1980(1)**				
1980(2)	1980(2)	1980(2)	1980(2)	1980(2)	1980(2)	1980(2)	1980(2)	1980(2)	1980(2)
									1980(3)
		1981(1)	1981(1)						
			1981(2)	1981(2)	1981(2)		1981(2)	1981(2)	1981(2)
	1981(4)	1981(4)	1981(4)	1981(4)	1981(4)			1981(4)	1981(4)
1982(1)	1982(1)					1982(1)			
							1982(2)		1982(2)
		1983(2)	1983(2)	1983(2)			1983(2)	1983(2)	1983(2)
1984(1)				1984(1)		1984(1)			
			1984(4)				**1984(4)**		**1984(4)**
	1985(1)				**1985(1)**			1985(1)	
						1985(2)			
		1985(3)							
				1986(1)		1986(1)			
			1990(4)				**1990(4)**		**1990(4)**
	1991(1)				**1991(1)**				
									1992(3)
					1992(4)				
			1994(4)				**1994(4)**		**1994(4)**
	1995(1)				**1995(1)**				
			1998(3)				**1998(3)**		**1998(3)**
	1998(4)				**1998(4)**				
			2001(2)				2001(2)		2001(2)
	2001(3)				2001(3)				
			2002(4)				**2002(4)**		**2002(4)**
	2003(1)				**2003(1)**				
			2008(2)				**2008(2)**		**2008(2)**
	2008(3)				**2008(3)**				
2008(4)	2008(4)		2008(4)		2008(4)				
2009(1)	2009(1)		2009(1)		2009(1)				
			2012(2)				**2012(2)**		**2012(2)**
	2012(3)				**2012(3)**				

Table 2 displays the VECM results for the growth rates of the real GNP/GDP series, the cointegrating vectors being fixed to $(1, -1)$ for the logs of the variables. Note that the without IIS and with IIS outcomes in the first row of Table 2 reproduce the outcomes of the detailed analysis on the first two releases (given in Equations (6)-(9)). An entry with a $0*$ denotes that the cointegrating vector has not been included by the algorithm in the final model using *Autometrics*. This is a specification for which either $\alpha_1 = 0$ or $\alpha_2 = 0$ is not rejected. The first panel displays the outcomes without IIS, the middle panel displays the outcomes obtained with IIS, while the third panel shows the outcomes if we include 9 dummies explicitly linked to the historical revisions.

We conclude that the use of the IIS method for the detection of the outliers is crucial for the inference regarding news-noise tests. In the absence of this method we would (falsely) conclude that there is strong evidence for the case of noise, as in all cases the second loading appears to be insignificant. Accounting for the existence of outliers allows us to test for noise only in three cases: between the first and the fourth release $(y_t^1, y_t^4,)$, between the first and the fifth release (y_t^1, y_t^5), and between the second and the fourth release (y_t^1, y_t^5). We further conclude that we obtain the same conclusions if we use explicit historical revision dummies instead of IIS.[8]

[8] We thank reviewer 1 for raising this issue, as it gives us the opportunity to confirm our argument that IIS can be used to capture historical revisions.

Table 2: Testing for zero loadings in the logs of OBS diagonals.

	Without IIS		With IIS				Historical Revision Dummies	
Variables	1st load.	2nd load.	1st load.	2nd load.	# dum. 1	# dum. 2	2nd load.	# dum. 2
$\ln(y_t^1, y_t^2)$	0.003	0*	−0.56	0.38	11	19	0.56	9
	(4.50)		(−10.90)	(5.43)			(7.05)	
$\ln(y_t^1, y_t^3)$	−0.53	0*	−0.37	0.23	14	17	0.27	9
	(−8.54)		(−8.11)	(5.60)			(5.85)	
$\ln(y_t^1, y_t^4)$	−0.33	0*	−0.29	0*	5	18	0*	9
	(−8.34)		(−7.72)					
$\ln(y_t^1, y_t^5)$	−0.28	0*	−0.26	0*	12	23	0*	9
	(−7.56)		(−8.69)					
$\ln(y_t^2, y_t^3)$	−1.02	0*	−0.62	0.43	7	19	0.47	9
	(−127.00)		(−11.20)	(6.87)			(6.43)	
$\ln(y_t^2, y_t^4)$	−0.41	0*	−0.54	0*	15	18	0*	9
	(−6.16)		(−11.00)					
$\ln(y_t^2, y_t^5)$	−0.34	0*	−0.30	0.18	9	23	0.13	9
	(−6.82)		(−7.98)	(6.30)			(3.79)	
$\ln(y_t^3, y_t^4)$	−1.01	0*	−0.60	0.33	14	22	0.38	9
	(−123.00)		(−11.40)	(5.70)			(5.02)	
$\ln(y_t^3, y_t^5)$	−0.42	0*	−0.43	0.32	15	19	0.19	9
	(−6.75)		(−7.87)	(8.12)			(4.35)	
$\ln(y_t^4, y_t^5)$	−0.99	0*	−0.57	0.59	11	22	0.38	9
	(−122.00)		(−11.20)	(9.64)			(5.23)	

Note: Values in (.) correspond to t-ratios for the null of insignificance. The Historical Revisions case is using dummies only for the dates that correspond to historical revisions. These dates are presented in Table 1 with bold font.

To estimate the cointegrating relationship for each of the pairs, we use the union of the impulse dummies given in Table 1 and run the Johansen approach. The results are presented in Table 3. We observe that the cointegrating relationships estimated with the Johansen approach are very close to $(1, -1)$. The estimates of the loadings are similar to the ones resulting from the VECM for the known cointegrating vector in Table 2.

Table 3: Estimation of the bivariate system using the Johansen's approach.

Variables	β_1	β_2	α_1	α_2	$H_0 : \beta_1 = -\beta_2$
$\ln(y_t^1, y_t^2)$	1	-1.0002	-0.6969	-0.5103	0.0836
	NA	(0.0006)	(0.1174)	(1.2124)	[0.7725]
$\ln(y_t^1, y_t^3)$	1	-1.0003	-0.4230	-0.0577	0.0582
	NA	(0.0014)	((0.0710)	(0.5696)	[0.8094]
$\ln(y_t^1, y_t^4)$	1	-1.0006	-0.3340	0.3119	0.1140
	NA	(0.0018)	(0.0482)	(0.3604)	[0.7357]
$\ln(y_t^1, y_t^5)$	1	-1.0007	-0.2368	0.3865	0.0691
	NA	(0.0026)	(0.0378)	(0.2785)	[0.7927]
$\ln(y_t^2, y_t^3)$	1	-1.0001	-0.7030	-0.4740	0.0132
	NA	(0.0007)	(0.1214)	(1.117)	[0.9086]
$\ln(y_t^2, y_t^4)$	1	-1.0003	-0.4218	0.5026	0.0602
	NA	(0.0013)	(0.0681)	(0.5468)	[0.8062]
$\ln(y_t^2, y_t^5)$	1	-1.0005	-0.3242	0.5059	0.0713
	NA	(0.0019)	(0.0482)	(0.3583)	[0.7894]
$\ln(y_t^3, y_t^4)$	1	-1.0002	-0.7164	1.0933	0.0956
	NA	(0.0007)	(0.1235)	(1.1056)	[0.7572]
$\ln(y_t^3, y_t^5)$	1	-1.0003	-0.4062	0.7575	0.0543
	NA	(0.0013)	(0.0671)	(0.5531)	[0.8157]
$\ln(y_t^4, y_t^5)$	1	-1.0001	-0.6835	1.3727	0.0226
	NA	(0.0007)	(0.1235)	(1.1214)	[0.8806]

Notes. The cointegrating vector is denoted by $\begin{pmatrix} \beta_1 & \beta_2 \end{pmatrix}$ and the loading vector by $(\alpha_1, \alpha_2)'$. Values in (.) are the standard errors of the estimates. The last column presents the likelihood ratio statistic for the null hypothesis $H_0 : \beta_1 = -\beta_2$. The values in [.] present the p-value for each test.

In the context of the Johansen maximum likelihood approach we test the null hypothesis that the cointegrating vector is $(1, -1)$ for each case. Evidence presented in Table 3 support the use of the restriction in all cases. So, we proceed to test for noise ($\lambda = 1$) in the three possible cases (y_t^1, y_t^4), (y_t^1, y_t^5) and (y_t^2, y_t^4). Table 4 lists the outcomes for the "noise" regression for the purposes of estimating λ and testing the null hypothesis that $\lambda = 1$. for each pair. The null-hypothesis that $\lambda = 1$ is rejected in all three cases. So, the null of revisions reducing noise is strongly rejected.

Table 4: Noise tests.

Variables	$\widehat{\lambda}$	t-test for $H_0 : \lambda = 1$	# dum.
$\ln(y_t^1, y_t^4)$	0.004	-105.195	18
	(0.009)		
$\ln(y_t^1, y_t^5)$	0.515	-6.071	25
	(0.080)		
$\ln(y_t^2, y_t^4)$	0.043	-40.170	22
	(0.024)		

5 Conclusion

This paper considers news-noise testing of univariate non-stationary real-time series. Standard Mincer-Zarnowitz tests are typically used for stationary time series. We describe an alternative for non-stationary series in which we have to test for weak exogeneity in so-called observation based cointegrated systems. If the first release is weakly exogenous, we can condition on it to set up a news equation. Alternatively, if the final release is weakly exogenous, a noise equation can be obtained.

Real-time data suffer from historical revisions. Roughly once every five years redefinitions or methodological innovations affect a data vintage from the beginning to the end, which hampers modeling and testing. Rather than taking growth rates or 'cleaning' the data as a first step of the empirical analysis, we propose to employ the IIS approach in our regressions, which involves adding indicator dummies for each observation to the model, and a general-to-specific selection process to test equations down to the preferred specification.

Our illustration with the U.S. real GNP/GDP series of the Federal Reserve Bank of Philadelphia shows that the IIS dummies are linked to the dates of historical revisions. We find that if we do not include indicator dummies, there is some evidence that we can rearrange the GNP/GDP releases we analyse as noise equations. However allowing for historical revisions by means of indicator dummies implies that we can cast our cointegrated systems into noise equations for three of the pairs we analysed, but have to reject the null of data revisions reducing noise.

References

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

Castle, J.L., J.A. Doornik, and D.F. Hendry (2012), "Model selection when there are multiple breaks", *Journal of Econometrics*, **169**, 239–246.

Clements, M. and A.B. Galvão (2013), "Real-time forecasting of inflation and output growth with autoregressive models in the presence of data revisions", *Journal of Applied Econometrics*, **28**, 458–477.

Croushore, D. (2011a), "Frontiers of real-time analysis", *Journal of Economic Literature*, **49**, 72–100.

Croushore, D. (2011b), "Forecasting with real-time data vintages", in M.P. Clements and D.F. Hendry, editors, *The Oxford Handbook of Economic Forecasting*, Oxford University Press, Oxford, New York, etc., chapter 9, 247–267.

Cunningham, A., J. Eklund, C. Jeffery, G. Kapetanios, and V. Labhard (2012), "A state space approach to extracting the signal from uncertain data", *Journal of Business & Economic Statistics*, **30**, 173–180.

De Jong, P. (1987), "Rational economic data revisions", *Journal of Business & Economic Statistics*, **5**, 539–548.

Doornik, J.A. and D.F. Hendry (2013), *Empirical Econometric Modelling—PcGiveTM 14: Volume I*, Timberlake Consultants Press, London.

Ericsson, N.R. and S.B. Kamin (2009), "Constructive data mining: Modelling argentine broad money demand", in J.L. Castle and N. Shephard, editors, *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, chapter 17, 412–439.

Ericsson, N.R. and E.L. Reisman (2012), "Evaluating a Global Vector Autoregression for forecasting", *International Advances in Economic Research*, **8**, 247–258.

Faust, J., J.H. Rogers, and J.H. Wright (2005), "News and noise in G-7 GDP announcements", *Journal of Money, Credit and Banking*, **37**, 403–419.

Fixler, D.J. and J. Nalewaik (2009), "News, noise, and estimates of the "true" unobserved state of the economy", Mimeo, U.S. Bureau of Economic Analysis.

Garratt, A., K. Lee, E. Mise, and K. Shields (2008), "Real time representations of the output gap", *Review of Economics and Statistics*, **90**, 792–804.

Garratt, A., K. Lee, E. Mise, and K. Shields (2009), "Real time representations of the UK output gap in the presence of model uncertainty", *International Journal of Forecasting*, **25**, 81–102.

Gonzalo, J. and C.W.J. Granger (1995), "Estimation of common long-memory components in cointegrated systems", *Journal of Business & Economics Statistics*, **13**, 27–35.

Götz, T., A. Hecq, and J.-P. Urbain (2016), "Combining forecasts from successive data vintages: An application to U.S. growth", *International Journal of Forecasting*, **32**, 61–74.

Hecq, A. and J.P.A.M. Jacobs (2009), "On the VAR-VECM representation of Real Time data", Mimeo, University of Maastricht.

Hendry, D.F. and J.A. Doornik (2014), *Empirical Model Discovery and Theory Evaluation. Automatic Selection Methods in Econometrics*, MIT Press, Cambridge MA.

Hendry, D.F. and C. Santos (2005), "Regression models with data-based indicator variables", *Oxford Bulletin of Economics and Statistics*, **67**, 571–595.

Jacobs, J.P.A.M and S. van Norden (2011), "Modeling data revisions: Measurement error and dynamics of "true" values", *Journal of Econometrics*, **161**, 101–109.

Jacobs, J.P.A.M and S. van Norden (2016), "Why are initial estimates of productivity growth so unreliable?", *Journal of Macroeconomics*, **47**.

Johansen, S. (1995), *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford University Press, Oxford.

Johansen, S. and B. Nielsen (2009), "Saturation by indicators in regression models", in J.L. Castle and N. Shephard, editors, *The Methodology and Practice of Econometrics. A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, New York, chapter 1.

Kishor, N.K. and E.F. Koenig (2012), "VAR estimation and forecasting when data are subject to revision", *Journal of Business & Economic Statistics*, **30**, 181–190.

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

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

Mankiw, N.G. and M.D. Shapiro (1986), "News or noise: An analysis of GNP revisions", *Survey of Current Business*, **66**, 20–25.

Mincer, J. and V. Zarnowitz (1969), "The evaluation of economic forecasts", in J. Mincer, editor, *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance*, National Bureau of Economic Research, New York, chapter 1, 3–46.

Orphanides, A. and S. van Norden (2002), "The unreliability of output gap estimates in real time", *Review of Economics and Statistics*, **84**, 569–583.

Patterson, K.D. (2000), "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**, 279–296.

Patterson, K.D. (2003), “Exploiting information in vintages of time-series data”, *International Journal of Forecasting*, **19**, 177–197.

Patton, A.J. and A. Timmermann (2011), “Predictability of output growth and inflation: A multi-horizon survey approach”, *Journal of Business & Economic Statistics*, **29**, 397–410.

Santos, C., D. F. Hendry, and S. Johansen (2008), “Automatic selection of indicators in a fully saturated regression”, *Computational Statistics*, **23**, 317–335.

Swanson, N.R. and D. van Dijk (2006), “Are statistical reporting agencies getting it right? Data rationality and business cycle asymmetry”, *Journal of Business & Economic Statistics*, **24**, 24–42.

Urbain, J.-P. (1992), “On weak exogeneity in error correction models”, *Oxford Bulletin of Economics and Statistics*, **54**, 187–207.

Urbain, J.-P. (1995), “Partial versus full system modelling of cointegrated systems—an empirical illustration”, *Journal of Econometrics*, **69**, 177–210.