



OECD Journal

Journal of Business Cycle Measurement and Analysis

Editor-in-Chief: Michael Graff



CIRET



Volume 2010/1

OECD Journal

Journal of Business Cycle Measurement and Analysis

Editor-in-Chief: Michael Graff

Volume 2010/1

A Joint Publication of OECD and CIRET

The OECD Journal collection comprises six periodicals: *OECD Journal: Economic Studies*, *OECD Journal: Financial Market Trends*, *OECD Journal: Competition Law and Policy*, *OECD Journal: General Papers*, *OECD Journal on Budgeting*, *OECD Journal: Journal of Business Cycle Measurement and Analysis*. Each periodical within OECD Journal retains its own ISSN, volume and issue numbering. The volume numbering of each journal corresponds to the calendar year, *e.g.* Volume 2010. Volume 2010 of *Journal of Business Cycle Measurement and Analysis* contains two issues.



CIRET



KOF

Centre for International Research
on Economic Tendency Surveys

The Ciret Office is located at

KOF Swiss Economic Institute, ETH Zurich,
Switzerland

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

Periodical: OECD Journal: Journal of Business Cycle Measurement and Analysis
ISSN 1995-2880 (print)
ISSN 1995-2899 (online)

© OECD 2010

Corrigenda to OECD publications may be found on line at: www.oecd.org/publishing/corrigenda.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.

Table of Contents

Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model <i>Elena Angelini, Marta Bańbura and Gerhard Rünstler</i>	5
Chain-linking in Austrian Quarterly National Accounts and the Business Cycle <i>Marcus Scheiblecker</i>	27
Evaluating German Business Cycle Forecasts under an Asymmetric Loss Function <i>Jörg Döpke, Ulrich Fritsche and Boriss Siliverstovs</i>	39
Do Benchmark Revisions Affect the Consumption-to-Output and Investment-to-Output Ratios in Germany? <i>Thomas A. Knetsch and Hans-Eggert Reimers</i>	57
OECD and CIRET News 2010/1	71

Editors

Editor-in-Chief

Michael Graff, *ETH Zurich, Switzerland*

Associate Editor

Gyorgy Gyomai, *OECD, Paris, France*

Editorial Board

Chair

Martine Durand, *Chief Statistician, OECD, Paris, France*

Richard Curtin, *University of Michigan,
Ann Arbor, MI, USA*

Jorgen Elmeskov, *OECD, Paris, France*

Horst Entorf, *University of Frankfurt, Germany*

Hélène Erkel-Rousse, *INSEE, National
Institute for Statistics and Economic Studies,
Paris, France*

Enrico Giovannini, *ISTAT, National Institute
for Statistics, Rome, Italy*

Robert J. Gordon, *Northwestern University,
Evanston, IL, USA*

Ulrich Heilemann, *University of Leipzig,
Germany*

Kajal Lahiri, *State University of New York
at Albany, NY, USA*

Marco Lippi, *University of Rome
“La Sapienza”, Italy*

Marco Malgarini, *ISAE, Institute for Studies
and Economic Analyses, Rome, Italy*

Zbigniew Matkowski, *Warsaw School
of Economics, Warszawa, Poland*

Gian Luigi Mazzi, *Eurostat, Luxembourg*

Gernot Nerb, *Ifo Institute for Economic
Research, Munich, Germany*

Ronny Nilsson *former Associate Editor
of the JBCMA, OECD, Paris, France*

Karl Heinrich Oppenländer *former president
of IFO Institute, Munich, Germany*

Ataman Ozyildirim, *The Conference Board,
New York, NY, USA*

Günter Poser, *former Editor-in-Chief
of the JBCMA, Technische Universität Darmstadt,
Germany*

Gerd Ronning, *University of Tübingen,
Germany*

Bernd Schips, *former Editor-in-Chief
of the JBCMA, ETH Zurich, Switzerland*

Jan-Egbert Sturm, *KOF Swiss Institute for
Business Cycle Research, ETH Zurich,
Switzerland*

István P. Székely, *Research Director, DG
ECFIN European Commission, Brussels,
Belgium and Corvinus University, Budapest,
Hungary*

András Vértes, *GKI Economic Research Co.,
Budapest, Hungary*

Martin Weale, *NIESR, National Institute
of Economic and Social Research, London, UK*

Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model

by

Elena Angelini, Marta Bańbura and Gerhard Rünstler*

We estimate and forecast growth in euro area monthly GDP and its components from the dynamic factor model of Doz et al. (2006), which handles unbalanced data sets in an efficient way. We extend the model to integrate interpolation and forecasting with cross-equation accounting identities. A pseudo real-time forecasting exercise indicates that the model outperforms various benchmarks, such as quarterly time-series models and bridge equations, in forecasting growth in quarterly GDP and its components.

JEL Classification: C53, E37

Keywords: dynamic factor models, interpolation, nowcasting

* Angelini: Directorate General Research, European Central Bank, Kaiserstrasse 29, D-60316 Frankfurt am Main, Germany (e-mail: elena.angelini@ecb.europa.eu); Bańbura: Directorate General Economics, European Central Bank (e-mail: marta.banbura@ecb.europa.eu); Rünstler: Austrian Institute for Economic Research, Arsenal Objekt 20, A-1030 Vienna, Austria (e-mail: gerhard.ruenstler@wifo.ac.at). The authors would like to thank G. Camba-Mendez, D. Giannone, L. Reichlin, C. Schumacher and two anonymous referees for useful discussions and suggestions. The opinions in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank.

Given the delays in the publication of national accounts, economic policy making in real time faces the difficulty of uncovering the actual state of the economy. For the euro area, a flash estimate of GDP is not available until six weeks after the end of the respective quarter. The first national accounts estimate is released about four weeks later. Meanwhile, observers must rely on high-frequency indicators that arrive within the quarter, such as industrial production, surveys and financial market data. However, the large number of available indicators, the noise present in many of the series and the different delays in their publication make the efficient exploitation of this information a difficult task. Under these circumstances, various approaches have been proposed to obtain measures of the economic status from the monthly data. Those include projections of quarterly GDP for the current and, possibly, next quarter, estimates of monthly GDP and monthly coincident indicators of economic activity.

In this paper, we investigate a unified approach to the interpolation and forecasting of GDP and of its demand and value-added components. The approach is based on a dynamic factor model (DFM) for a large monthly data set, and is suitable for dealing with asynchronous data releases in real-time application. Our objective is to obtain estimates and forecasts that satisfy temporal aggregation constraints with respect to quarterly data, as well as appropriate accounting identities.

We build on the DFM of Doz *et al.* (2006a), which differs from other approaches (*e.g.*, those of Stock and Watson, 2002; Forni *et al.*, 2000, 2005) in modelling factor dynamics in an explicit manner. From a state-space representation of the model, forecasts are obtained through application of the Kalman smoother. As a consequence, the model can deal with those irregular patterns of data availability at the end of the sample, which arise in real-time data sets because of differences in publication delays (Giannone *et al.*, 2008). In addition, the state-space framework allows for a comprehensive analysis of the contributions of individual data to the forecasts (Bańbura and Rünstler, 2010).

We combine the model with forecasting equations for monthly GDP and its demand components with appropriate temporal aggregation rules and the relevant accounting identities. Hence, our approach provides monthly estimates and predictions of quarterly GDP growth and its components, which are mutually consistent. This greatly facilitates communicating the results to policy makers, as it clarifies the implications of quarterly predictions for intra-quarter dynamics, and vice versa.

In the empirical part of the paper we evaluate forecasts for GDP and its demand and value components in terms of out-of-sample forecast performance against various alternative models. Banerjee *et al.* (2005), Bańbura and Rünstler (2010) and Angelini *et al.* (2008) report good forecasting performance of factor models for euro area GDP growth. Alternatively, GDP growth has been forecast from bridge equations using a small set of selected monthly indicators, notably measures of production and sales (*e.g.*, Rünstler and Sédillot, 2003; Baffigi *et al.*, 2004; Diron, 2008).

Thus far, estimates of monthly GDP have been derived primarily from bottom-up approaches based on estimates of monthly components (e.g., Mitchell et al., 2005a, 2005b; Proietti and Frale, 2007), which again are based on selected indicators. The bottom-up approach suffers, however, from the potential weakness that poor interpolates for certain components may hamper the aggregate GDP interpolate. Breitung and Schumacher (2008) have employed a diffusion index model for estimating monthly GDP directly.

We find that the factor model forecasts for euro area GDP beat the forecasts from alternative models, such as quarterly time series models and bridge equations. The same applies to forecasts for the demand and value-added components, with the exception of private and public consumption, for which none of the models does well. (For the latter series, further research is required in order to detect informative monthly indicators.) We also compare the monthly interpolates of GDP delivered by our model to those obtained from standard interpolation methods based on a small number of indicators. The resulting in-sample monthly estimates are similar. We argue, however, that this might not be the case for real-time GDP interpolates of the most recent periods.

The paper is organised as follows. After a brief review of the DFM of Doz et al. (2006a), the integrated DFM version used to forecast and interpolate the national accounts is presented in Section 1. Section 2 reports the results of a pseudo real-time exercise to compare the performance of the model with various alternative models. Section 3 shows estimates of monthly GDP and compares them to results from standard interpolation methods. Section 4 concludes.

1. The model

1.1. A dynamic factor model

DFMs are designed to explain the dynamics in a panel of series by a few common sources of variation. Consider a vector of n stationary monthly series $x_t = (x_{1t}, \dots, x_{nt})'$, $t = 1, \dots, T$, which have been standardised to mean zero and variance one. The DFM by Doz et al. (2006a) is given by the equations

$$x_t = \Lambda f_t + \xi_t, \quad \xi_t \sim \mathbb{N}(0, \Sigma_\xi), \quad (1)$$

$$f_{t+1} = \sum_{s=1}^p A_s f_{t-s+1} + B \eta_{t+1}, \quad \eta_{t+1} \sim \mathbb{N}(0, I_q). \quad (2)$$

Equation (1) relates the monthly series x_t to a $r \times 1$ vector of latent factors $f_t = (f_{1,t}, \dots, f_{r,t})'$ from a matrix of factor loadings Λ , plus an idiosyncratic component $\xi_t = (\xi_{1,t}, \dots, \xi_{n,t})'$. The latter is assumed to be multivariate white noise with covariance matrix Σ_ξ . Equation (2) describes the law of motion for the latent factors f_t , which are driven by q -dimensional white noise η_{t+1} , where B is a $r \times q$ matrix. The stochastic process for f_t is assumed to be stationary.

Stock and Watson (2005) refer to the specification given by (1) and (2) as the static representation of a DFM, since series x_t load only on current values of factors f_t . However, the specification can be derived from a restricted version of a more general DFM

with q so-called dynamic factors. Static factors f_t then contain current and lagged values of dynamic factors, while series x_t may load on current and lagged values of the latter. In this case, it holds that $q < r$.¹

The specification above differs from the representation of, *e.g.*, Stock and Watson (2002) in that the dynamics of the factors is explicitly modelled through equation (2). Exploiting the dynamics may lead to efficiency improvements in real-time forecasting (Doz *et al.*, 2006a; Rünstler *et al.*, 2009).

1.2. Interpolation

Following Banbura and Rünstler (2010), we use a mixed-frequency approach to combine the monthly factor model with equations to model monthly GDP growth within a single state-space form. For this purpose, we introduce monthly growth in GDP and its components as latent variables.²

More precisely, denote with $Y_t = (Y_{1,t}, \dots, Y_{m,t})'$ the $m \times 1$ vector of monthly national accounts, which satisfy the accounting identity

$$Y_{1,t} = \sum_{i=2}^m Y_{i,t} + K_t,$$

where K_t denotes a possible error term. At the end of each quarter, $t = 3k$, we find the quarterly national accounts as the sum of the respective monthly values,

$$Y_{i,3k}^Q = \sum_{s=0}^2 Y_{i,3k-s},$$

where $k = 1, 2, \dots, \lfloor T/3 \rfloor$.³ We form monthly growth rates $\gamma_{i,t} = \ln(Y_{i,t}) - \ln(Y_{i,t-1})$ and three-month growth rates $\gamma_{i,t}^{(3)} = \ln(Y_{i,t}) - \ln(Y_{i,t-3})$. Let $\gamma_t = (\gamma_{1,t}, \dots, \gamma_{m,t})$. From logarithmic approximation, the following familiar aggregation rules apply to the growth rates of the monthly national accounts (see also, *e.g.*, Mariano and Murasawa, 2003; Breitung and Schumacher, 2008):

$$\gamma_t^{(3)} = \gamma_t + \gamma_{t-1} + \gamma_{t-2}, \quad (3)$$

$$\gamma_{3k}^Q = \frac{1}{3} (\gamma_{3k}^{(3)} + \gamma_{3k-1}^{(3)} + \gamma_{3k-2}^{(3)}). \quad (4)$$

The national accounts identity for monthly growth rates becomes

$$\chi'_{t-1} \gamma_t = \kappa_t, \quad \kappa_t \sim \mathbb{N}(0, \sigma_\kappa^2), \quad (5)$$

where weights χ_{t-1} represent the shares of components in GDP, *i.e.*, $\chi_{1,t-1} = -1$, and $\chi_{i,t-1} = Y_{i,t-1} / Y_{1,t-1}$ otherwise.

Finally, monthly growth rates γ_t are related to the common factors by the static equation

$$\gamma_{t+1} = \mu + \Lambda_y f_{t+1} + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}(0, \Sigma_\varepsilon), \quad (6)$$

where μ is a $m \times 1$ constant term and Λ_y is the $m \times r$ matrix of factor loadings for the national accounts variables. The idiosyncratic component $\varepsilon_t = (\varepsilon_{1,t}, \dots, \varepsilon_{m,t})'$ is assumed to be multivariate white noise with covariance matrix Σ_ε . We further assume that the vectors of innovations $\xi_t, \eta_t, \varepsilon_t$, and κ_t are mutually uncorrelated.

Constraints (3) to (5) are log-linear approximations of the original identities and ignore issues related to chain-linking. In a recent paper, Proietti (2008) proposes an iterative non-linear estimator to interpolate the national accounts, which satisfies the national accounts constraints exactly and also allows for implementing exact (non-linear) temporal constraints for chain-linked data. The approach requires the data set to be balanced and, hence, is not applicable to forecasting in real time. Proietti and Frale (2007) have shown that ignoring chain-linking has only a very small impact on the interpolates. The log-linear approach, in turn, requires a small error term κ_t , but has the advantage that it allows for handling unbalanced real-time data sets.

1.3. State space form

Equations (1) to (6) can be cast in a single state-space form, which is illustrated below for the case of $p = 1$. The transition equation contains the dynamic law of motion for the state vector $\alpha'_t = (f'_t, \gamma'_t, \gamma'_{t-1}, \gamma^{(3)}_t, Q'_t)$ comprising factor dynamics (2), temporal aggregation rules (3) and (4), and forecasting equations (6) for the monthly national accounts. In the state space form below, aggregation rule (4) is implemented in a recursive way from

$$Q_t = \Xi_{t-1} Q_{t-1} + \frac{1}{3} \gamma^{(3)}_t,$$

where $\Xi_{t-1} = 0_{m \times m}$ in the first month and $\Xi_{t-1} = I_m$ otherwise (see Harvey, 1989: 309ff). As a result, expression (4) holds in the third month of the quarter, $\gamma^{(3)}_{3k} = Q_{3k}$. The recursive implementation of the aggregation through Q_t reduces the size of the state vector, and thereby computation time, to a considerable extent.

The transition equation is given by

$$\begin{bmatrix} I_r & 0 & 0 & 0 & 0 \\ -\Lambda_y & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & -I & -I & I & 0 \\ 0 & 0 & 0 & -\frac{1}{3}I & I \end{bmatrix} \begin{bmatrix} f_{t+1} \\ \gamma_{t+1} \\ \gamma_t \\ \gamma^{(3)}_{t+1} \\ Q_{t+1} \end{bmatrix} = \begin{bmatrix} 0 \\ \mu \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} A_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & \Xi_t \end{bmatrix} \begin{bmatrix} f_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma^{(3)}_t \\ Q_t \end{bmatrix} + \begin{bmatrix} B\eta_{t+1} \\ \varepsilon_{t+1} \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

where I denotes the $m \times m$ identity matrix. The equation is to be pre-multiplied by the inverse of the left-hand matrix to achieve the standard state-space form.

The observation equation is

$$\begin{bmatrix} x_t \\ 0 \\ \gamma_t^Q \end{bmatrix} = \begin{bmatrix} \Lambda & 0 & 0 & 0 & 0 \\ 0 & \chi'_{t-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_m \end{bmatrix} \begin{bmatrix} f_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_t^{(3)} \\ Q_t \end{bmatrix} + \begin{bmatrix} \xi_t \\ \kappa_t \\ 0 \end{bmatrix}.$$

The final rows of the observation equation, related to γ_t^Q , are defined only for the third month of the quarter and otherwise skipped in application.

1.4. Estimation, interpolation and forecasting

As shown by Doz *et al.* (2006a), under certain regularity conditions consistent estimates of the model parameters can be obtained as follows:

1. Apply principal components analysis to x_t to estimate the first r common factors $\hat{f}_t$, together with factor loadings $\hat{\Lambda}$ and variances of idiosyncratic components $\hat{\Sigma}_\xi$.
2. Estimate the VAR $\hat{f}_t = \sum_{s=1}^p A_s \hat{f}_{t-s} + \hat{\zeta}_t$ to obtain estimates $\hat{A}_s$ and $\hat{\Sigma}_\zeta$. Further, apply principal components to the estimated covariance matrix $\hat{\Sigma}_\zeta$ of residuals $\hat{\zeta}_t$ and extract the first q components to obtain $\hat{B}$.
3. Obtain quarterly aggregates $\hat{f}_t^Q$ of estimates $\hat{f}_t$ as from equations (3) and (4). Estimate a quarterly version of (6),

$$\gamma_t^Q = \mu^Q + \Lambda_\gamma \hat{f}_t^Q + \varepsilon_t^Q, \quad \varepsilon_t^Q \sim \mathbb{N}(0, \sigma_\varepsilon^Q),$$

by ordinary least squares (OLS). As equation (6) is static, the quarterly aggregates give consistent estimates of Λ_γ , $\sigma_\varepsilon = 3 / \sqrt{19} \sigma_\varepsilon^Q$ and $\mu = 1 / 3 \mu^Q$. Similarly, $\sigma_\kappa = 3 / \sqrt{19} \sigma_\kappa^Q$, where σ_κ^Q is estimated from the quarterly data.⁴

We now turn to the application of the model to interpolation and forecasting in real time. Real-time data sets typically contain missing observations at the end of the sample because of publication lags. Moreover, the amount of missing data differs across series due to the different timing of data releases. In our forecast exercise we will apply *pseudo-real-time* data sets $\mathcal{Z}_t$, which use the final data releases but take account of the timing of data releases. This is achieved by shifting the pattern of publication lags embodied in $\mathcal{Z}_T$ recursively back in time. That is, monthly observation $z_{i,t-s}$, $s \geq 0$ is eliminated in $\mathcal{Z}_t$, if and only if observation $z_{i,T-s}$ is missing in $\mathcal{Z}_T$. The quarterly national accounts are treated in a similar fashion.

To obtain efficient estimates and forecasts of GDP growth from unbalanced data sets, Kalman filter and smoother recursions can be applied. For state-space form,

$$\begin{aligned} z_t &= W_t \alpha_t + u_t & u_t &\sim \mathbb{N}(0, \Sigma_u), \\ \alpha_{t+1} &= c + T_t \alpha_t + v_t, & v_t &\sim \mathbb{N}(0, \Sigma_v), \end{aligned} \tag{7}$$

and any unbalanced data set $\mathcal{Z}_t$, the Kalman filter and smoother provide minimum mean square error (MMSE) estimates $a_{t+h|t}$ of the state vector and their precision, $P_{t+h|t}$,

$$a_{t+h|t} = \mathbb{E}[\alpha_{t+h} | \mathcal{Z}_t], \quad (8)$$

$$P_{t+h|t} = \mathbb{E}\left[a_{t+h|t} - \alpha_{t+h} \right] \left[a_{t+h|t} - \alpha_{t+h} \right]', \quad (9)$$

for any $h > -t$. To handle missing observations, the rows in equation (7) corresponding to missing observations in z_t are simply skipped when applying the Kalman filter recursions (Durbin and Koopman, 2003: 92f). In the case of forecasting, $h > 0$, it is sufficient to run the Kalman filter, whereas *ex post* estimates of monthly national accounts are derived from the smoother.

Further, Bańbura and Rünstler (2010) have proposed using an algorithm by Harvey and Koopman (2003) to obtain the Kalman filter and smoother weights of individual series in forecasts and monthly estimates of national accounts. This allows expressing estimates $a_{t+h|t}$ as

$$a_{t+h|t} = \sum_{s=0}^{t-1} \omega_s(h) z_{t-s}. \quad (10)$$

As data sets $\mathcal{Z}_t$ embody fixed data release patterns, the $1 \times n$ vector of weights $\omega_s(h)$ is independent of time t , once the Kalman filter has approached its steady state (see Bańbura and Rünstler, 2010). We will consider the cumulative smoother weights $\sum_{s=0}^{t-1} \omega_{s,j}(h)$ for series j , where $\omega_{s,j}(h)$ is the j^{th} element of $\omega_s(h)$, $j = 1, \dots, n$. The contribution of series j to estimate $a_{t+h|t}$ is calculated as $\sum_{s=0}^{t-1} \omega_{s,j}(h) z_{j,t-s}$.

2. Forecast evaluation

In this section we present a forecast exercise to evaluate the historical forecast performance of the dynamic factor model against various rival models, including univariate time series models and bridge equations. We consider forecasts over the period of 2000Q1 to 2006Q2. We address the following questions. First, how well can the components of GDP be forecast? While a number of studies have inspected forecasts for GDP, components have been largely neglected. Second, how does the DFM compare to benchmark models? Third, does constraint (5) help in forecasting?

2.1. Data, publication lags and forecast design

Our euro area data set ($\mathcal{Z}_T$) begins in January 1993 and was downloaded on 20 February 2007. It contains 85 monthly series including official data on economic activity, surveys and financial market data. Our choice of series is based on earlier studies by, *e.g.*, Stock and Watson (2002) and Giannone *et al.* (2008) for the United States, and Bańbura and Rünstler (2010) for the euro area.⁵ The monthly series of data on euro area economic activity contains components of industrial production (17), employment and unemployment data (5), extra euro area trade values from the balance of payments (4), retail sales (1) and new passenger

car registrations (1). We use 24 series from the European Commission business, consumer, retail and construction surveys. Financial data comprise 17 series including exchange rates (6), interest rates (7) and equity price indices (4). The data contain monetary aggregates and loans (5) and 11 series on the international economy including raw material prices (5) and key macroeconomic indicators for the United States (6). The series list is given in the annex, together with the data transformations we use for all models in this study. The annex also reports the publication lags of the individual series and Augmented Dickey-Fuller (ADF) tests against unit roots in the transformed data.

The monthly data are published at different times. Surveys, financial data and raw material prices are available at the end of the respective month. By contrast, most of the official data on euro area economic activity, such as industrial production, employment and retail sales, are published with a delay of six to eight weeks after the end of the month. The same applies to the euro area monetary aggregates. This implies that in our data, surveys and financial data are available for January 2007, but most of the real activity data for December 2006 only.

Our euro area quarterly national accounts data include GDP and the major demand components, with inventories subsumed under the statistical discrepancy. In addition, we consider value added and its two major components, industry (including construction and agriculture) and services. The national accounts are published about ten weeks after the end of the respective quarter, while a flash estimate of GDP is available about one month earlier. Hence, our data contain the GDP flash estimate for 2006Q4, but first releases of the demand components and value added only for 2006Q3.

With our forecast design, we aim to replicate the real-time application of the models as closely as possible. Although we do not have real-time data sets at hand, following Rünstler and Sédillot (2003) and Giannone *et al.* (2008), we take account of publication lags in the series and use pseudo real-time data sets $\mathcal{Z}_t$ as defined in Section 1.3. In addition, we re-estimate the models at each point in time based on the available data at the time the forecast is made. Since our data were downloaded on 20 February 2007, our forecasts will replicate the data availability situation on the twentieth day of the month.

We inspect *eight* forecasts for growth in GDP and its components in a certain quarter. These forecasts are obtained in consecutive months. We start with forecasting in the first month of the previous quarter and stop in the second month of the subsequent (next) quarter, one month before the first estimate of national accounts is released by Eurostat. The design will be illustrated in the following section.

2.2. Forecast evaluation

All forecasts are evaluated over the period of 2000Q1 to 2006Q2, with recursive estimation starting in 1993Q1.⁶ We consider the following models:

- As benchmarks, we use naive (random walk) forecasts and first-order autoregressive processes (AR(1)) for quarterly GDP and its components. The naive forecast is simply the unconditional mean of the growth rate in each quarterly series, which amounts to a random walk with drift forecast in the level of the series. Again, both forecasts are calculated recursively, i.e., each forecast is based on the available data at the time the forecast is made.
- Bridge equations are widely used for the short-term forecasting of GDP and its components (*e.g.*, Baffigi *et al.*, 2004; Rünstler and Sédillot, 2003; Diron, 2008), as they

employ intraquarter information from the individual indicators. Following Kitchen and Monaco (2006), we obtain forecasts for quarterly series $\gamma_{i,t}^Q$ from each indicator $x_{j,t}$, and average across those forecasts. We forecast the individual monthly indicators from monthly AR(p) models, $\varphi(L)x_{j,t}^{(3)} = e_{j,t}$ over the desired horizon, where we use three-month rates $x_{j,t}^{(3)}$, as this tends to give better forecasts. We use the Bayesian information criterion (BIC) to determine lag length p . Forecasts, $\hat{x}_{j,t+h}^{(3)}$, are then aggregated to quarterly frequency, $\hat{x}_{j,t+h}^Q$, and the target series $\gamma_{i,t}^Q$ is predicted from the “bridge” equation

$$\hat{\gamma}_{i,t+h}^{Q(j)} = c_{ij} + \beta_{ij}\hat{x}_{j,t+h}^Q.$$

We estimate parameters c_{ij} and β_{ij} by OLS. The final forecast $\hat{\gamma}_{i,t+h}^Q$ is obtained as the average of the forecasts

$$\hat{\gamma}_{i,t+h}^Q = n^{-1} \sum_{j=1}^n \hat{\gamma}_{i,t+h}^{Q(j)}. \quad (11)$$

- As for the DFM, we consider the multivariate model both with and without constraint (5). We apply the model to two sets of national accounts data. The first data set includes GDP and its demand components, i.e., private and public consumption, gross fixed capital formation (GFCF), export and imports and the statistical discrepancy. The second version contains total value added (VAD) plus its breakdown into VAD industry and VAD services.⁷

As to the specification of the DFM, we determine the number of static factors r from the information criterion developed by Bai and Ng (2002), which gives $r = 4$, while the number of lags in factor dynamics is found from the BIC with $p = 3$. Studies have argued that two shocks are sufficient to model economic activity, and we therefore set $q = 2$ (Giannone et al., 2008). Compared to specification selection based on forecast performance, this approach has the advantage that the specification choice is independent of the target series, as we want to evaluate our model across a set of target series.

Table 1 shows the root mean square error (RMSE) from the naive quarterly forecast over the period of 2000Q1 to 2006Q2 (26 observations). As noted above, we inspect *eight* forecasts for growth in GDP and its components in a certain quarter, which are obtained in consecutive months. We start with forecasting in the first month of the previous quarter and stop in the second month of the subsequent (next) quarter, one month before the first estimate of national accounts is released by Eurostat. As an example, the table illustrates the timing of the forecasts and data releases for the second quarter of the year. We run the first forecast for the second quarter on 20 January and the final (eighth) on 20 August. Note that the last two “forecasts” are actually backcasts, whereas forecasts 4 to 6 amount to nowcasting the current quarter.

The timing of forecasts for the other quarters is equivalent. Since the naive forecast is based on the quarterly data, the RMSE shifts every three months. The timing of these shifts reflects publication dates. New observations for GDP become available in the second month of the quarter, those for components one month later.

Table 1. **RMSE of naive forecast**

Fcst	Example 2 nd quarter	GDP	Priv cons	Gov cons	GFCF	Export	Import	Stat discr	VAD total	VAD ind	VAD serv
8	August	0.000	0.314	0.356	0.798	1.509	1.438	0.320	0.364	0.629	0.301
7	July	0.332	0.314	0.356	0.798	1.509	1.438	0.320	0.364	0.629	0.301
6	June	0.332	0.314	0.356	0.798	1.509	1.438	0.320	0.364	0.629	0.301
5	May	0.332	0.316	0.354	0.807	1.526	1.460	0.317	0.372	0.642	0.307
4	April	0.338	0.316	0.354	0.807	1.526	1.460	0.317	0.372	0.642	0.307
3	March	0.338	0.316	0.354	0.807	1.526	1.460	0.317	0.372	0.642	0.307
2	Feb	0.338	0.317	0.358	0.819	1.538	1.474	0.316	0.380	0.652	0.314
1	Jan	0.347	0.317	0.358	0.819	1.538	1.474	0.316	0.380	0.652	0.314

Table 2 presents the results for the AR(1) and the bridge equations in terms of the RMSE relative to the RMSE of the naive forecast. For the bridge equations, we construct the GDP forecast from a limited set of 8 series, which gives a smaller RMSE as compared to forecasts from the entire set of 87 series. The forecasts for components are based on the full set of series.⁸ As regards GDP, both models improve upon the naive forecast for the very short horizons, i.e., forecasts 5 to 8 (the backcasts and late nowcasts). For the one-quarter-ahead forecasts (1 to 3), only bridge equations outperform the naive forecast, but the gains remain below 10%.

For the components of VAD, the bridge equations beat the naive forecast. Among the demand components, some gains emerge for exports and imports, but these rarely exceed 10%. For private and public consumption as well as GFCF, the benchmark models give largely uninformative forecasts.

Table 2. **RMSE of benchmark models (relative to naive forecast)**

Quarterly AR(1)										
Fcst	GDP	Priv cons	Gov cons	GFCF	Exp	Imp	Stat discr	VAD total	VAD ind	VAD serv
8		1.04	0.97	1.10	0.95	0.91	0.91	0.89	0.94	0.96
7	0.82	1.04	0.97	1.10	0.95	0.91	0.91	0.89	0.94	0.96
6	0.82	1.04	0.97	1.10	0.95	0.91	0.91	0.89	0.94	0.96
5	0.82	1.07	0.96	1.01	1.00	1.01	1.01	0.97	1.00	0.97
4	0.98	1.07	0.96	1.01	1.00	1.01	1.01	0.97	1.00	0.97
3	0.98	1.07	0.96	1.01	1.00	1.01	1.01	0.97	1.00	0.97
2	0.98	1.07	0.99	1.03	1.01	1.05	1.01	1.02	1.01	1.03
1	1.03	1.07	0.99	1.03	1.01	1.05	1.01	1.02	1.01	1.03
Bridge equations										
Fcst	GDP	Priv cons	Gov cons	GFCF	Exp	Imp	Stat discr	VAD total	VAD ind	VAD serv
8		1.02	0.97	1.06	0.86	0.94	0.99	0.80	0.76	0.83
7	0.86	1.02	0.97	1.06	0.86	0.94	0.99	0.80	0.76	0.83
6	0.87	1.04	1.00	1.06	0.87	0.94	1.00	0.81	0.77	0.84
5	0.88	1.04	1.04	1.08	0.87	0.95	0.99	0.79	0.79	0.86
4	0.87	1.04	1.01	1.09	0.89	0.96	0.99	0.79	0.79	0.87
3	0.89	1.03	1.00	1.08	0.92	0.99	0.99	0.82	0.82	0.88
2	0.92	1.03	1.00	1.10	0.98	1.02	0.99	0.86	0.87	0.89
1	0.94	1.04	0.99	1.09	0.99	1.02	0.99	0.85	0.90	0.88

The results for the DFM version without constraint (5) are shown in the upper panel of Table 3. The model shows substantial improvements upon the alternative models for GDP and most demand and value-added components. For the short horizons, the RMSE of the GDP forecast is now 30% lower as compared with the naive forecast. For the first forecast, eight months ahead of the data release, the improvement still amounts to 20%. Similar gains occur for gross fixed capital formation, exports and imports, and the components of value added. While the absence of any gains for private and public consumption may reflect a lack of forecastability of the series *per se*, it also may be a consequence of a lack of appropriate monthly indicators in our data set. It has been argued that private consumption follows a random walk (Hall, 1988), and this also seems plausible for government consumption. However, the lack of forecastability does not necessarily preclude informative *nowcasts* of the series based on intra-quarter information.

Table 3. **RMSE of dynamic factor models (relative to naive forecast)**

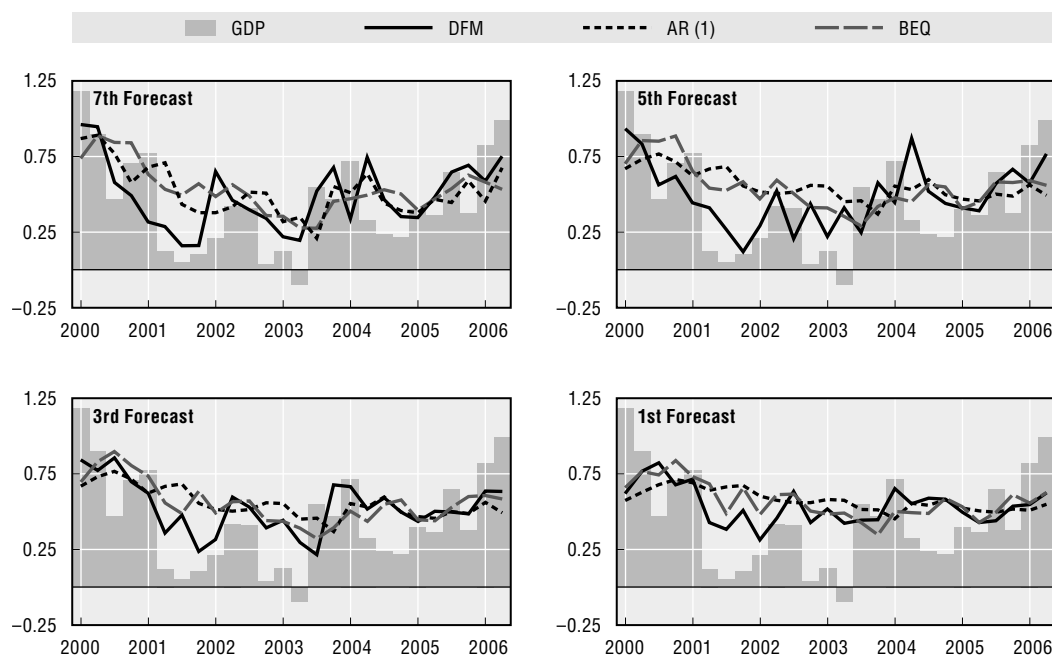
Without constraint (5)										
Fcst	GDP	Priv cons	Gov cons	GFCF	Exp	Imp	Stat discr	VAD total	VAD ind	VAD serv
8		1.01	1.13	0.81	0.77	0.69	1.13	0.72	0.64	0.85
7	0.70	0.98	1.15	0.81	0.77	0.72	1.13	0.70	0.64	0.84
6	0.72	0.98	1.20	0.84	0.77	0.70	1.11	0.75	0.70	0.86
5	0.74	0.98	1.04	0.81	0.72	0.74	0.99	0.75	0.75	0.86
4	0.73	0.97	1.02	0.82	0.70	0.68	0.98	0.77	0.76	0.87
3	0.73	0.96	1.01	0.81	0.82	0.80	0.98	0.76	0.70	0.88
2	0.80	0.96	1.01	0.82	0.86	0.86	0.99	0.76	0.77	0.85
1	0.81	0.98	1.01	0.86	0.90	0.90	0.98	0.80	0.84	0.86
Including constraint (5)										
Fcst	GDP	Priv cons	Gov cons	GFCF	Exp	Imp	Stat discr	VAD total	VAD ind	VAD serv
8		0.94	1.13	0.78	0.74	0.69	1.20	0.72	0.62	0.87
7	0.67	0.97	1.16	0.82	0.78	0.72	1.13	0.72	0.63	0.86
6	0.70	0.96	1.21	0.85	0.77	0.70	1.12	0.75	0.68	0.88
5	0.76	0.97	1.04	0.78	0.73	0.74	0.99	0.76	0.73	0.87
4	0.74	0.96	1.03	0.80	0.70	0.68	0.99	0.76	0.73	0.88
3	0.73	0.96	1.01	0.78	0.82	0.80	0.98	0.74	0.68	0.87
2	0.79	0.95	1.02	0.80	0.85	0.87	0.99	0.78	0.77	0.87
1	0.79	0.97	1.01	0.84	0.89	0.91	0.98	0.84	0.84	0.91

The lower panel of Table 3 shows the results for the DFM using constraint (5). For the demand components, the inclusion of the constraint tends to improve the RMSE for the eighth forecast, but leaves it unchanged otherwise. This appears to be related to the fact that the flash estimate of quarterly GDP is already available for the eighth forecast, as it is published about four weeks before the full national accounts. In this situation, the information contained in the flash estimate contributes to forecasting the demand components.

Figure 1 shows forecasts from the DFM and the benchmarks. The graphs visualize the higher precision of the DFM forecasts compared to the AR(1) and the bridge equations. They indicate a better performance of the DFM compared to the other models, in particular for the 2001–2003 period. As shown in Table A.2 in the annex, however, the DFM also improves upon the alternative models in subsequent years, although to a smaller extent.

Finally, Table A.3 in the annex reports forecasts encompassing and Diebold-Mariano tests of the DFM using constraint (5) against the bridge equations. They indicate that, for

Figure 1. GDP forecasts



GDP and most of the components, forecasts from bridge equations do not add information to those from the DFM and are significantly less precise over short horizons. More precisely, we run the encompassing regression

$$y_{i,t+h}^Q = \lambda \hat{y}_{i,t+h}^{Q,DFM} + (1 - \lambda) \hat{y}_{i,t+h}^{Q,BEQ} + u_t \quad (12)$$

to explain observations $y_{i,t+h}^Q$ by the two forecasts (Clements and Hendry, 1998: 228ff). A value of $\lambda = 1$ indicates that forecasts $\hat{y}_{i,t+h}^{Q,BEQ}$ from bridge equations do not add information to forecasts $\hat{y}_{i,t+h}^{Q,DFM}$ from the DFM, and the same holds in an opposite manner for a value of $\lambda = 0$.

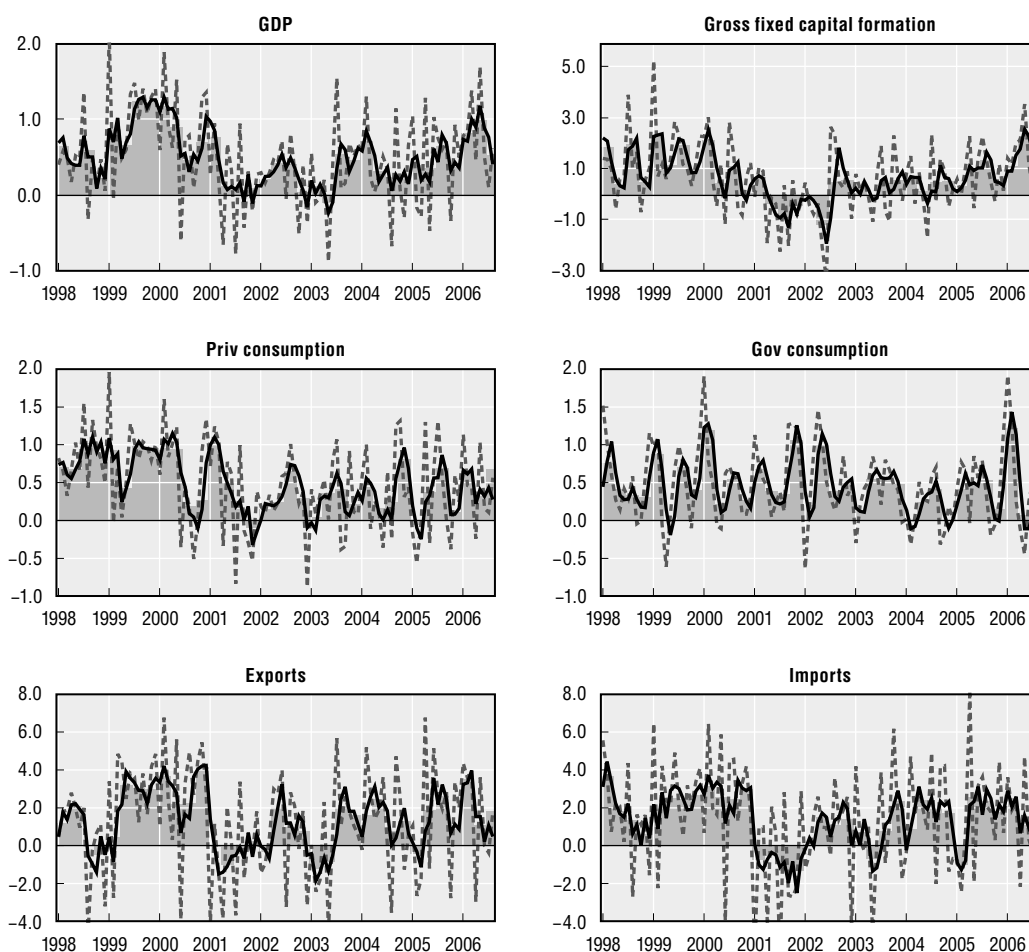
Table A.3 shows that, with the exception of government consumption, estimates of λ are generally close to one or at least higher than 0.5. The Diebold-Mariano tests against the null hypothesis of equal forecast efficiency find the efficiency gains from the DFM over part of the horizon significant at the 10% level for GDP, consumption, exports, imports and VAD services.

3. Estimates of monthly national accounts

The smoothed estimates of growth in monthly GDP and its components from the DFM using constraint (5) are shown in Figure 2. The graph contains estimates of both three-month and month-on-month rates, multiplied by three, together with the observed quarterly rates. Note that these estimates are obtained from the Kalman smoother based on the entire data set $\mathcal{Z}_T$.

Angelini et al. (2006) compare factor-based interpolation methods with the traditional method by Chow and Lin (1971), and conclude that both methods fare well. We therefore inspect estimates of monthly GDP growth from applying the Chow-Lin method to a single equation. Following existing studies on estimating euro area monthly GDP (Mitchell et al.,

Figure 2. Estimates of monthly growth in GDP and demand components



Notes: The grey bars show quarterly growth in the component, while the bold and thin lines show estimates of three-month and month-on-month growth rates, respectively. The latter are multiplied by three.

2005a, 2005b; Proietti and Frale, 2007), we choose euro area industrial production in manufacturing, total employment, the business confidence indicator and retail sales as explanatory variables.⁹

Figure 3 demonstrates high correspondence among estimates of monthly growth rates from the two methods, with a contemporaneous correlation of 0.86 among the monthly series over the period of 1998M1 to 2006M6. This reflects the fact that the Kalman smoother attaches high weights to items of industrial production and, to a lesser extent, business surveys, when backcasting monthly growth rates. In this case, the DFM effectively uses information similar to what has been chosen in the aforementioned studies.

Bañura and Rünstler (2010) have shown, however, that the weights of individual series in quarterly GDP forecasts may change considerably with the forecast horizon. From contribution analysis, as in equation (10), it can be shown that the same applies to estimates and forecasts of monthly growth. Table 4 presents the mean absolute contributions (MACs) of individual data groups to the forecasts of monthly GDP growth. Sample contributions have been estimated from the same pseudo real-time forecast design over the period 2000Q1 to 2006Q2 as used in Section 2.

Figure 3. Estimates of monthly GDP growth

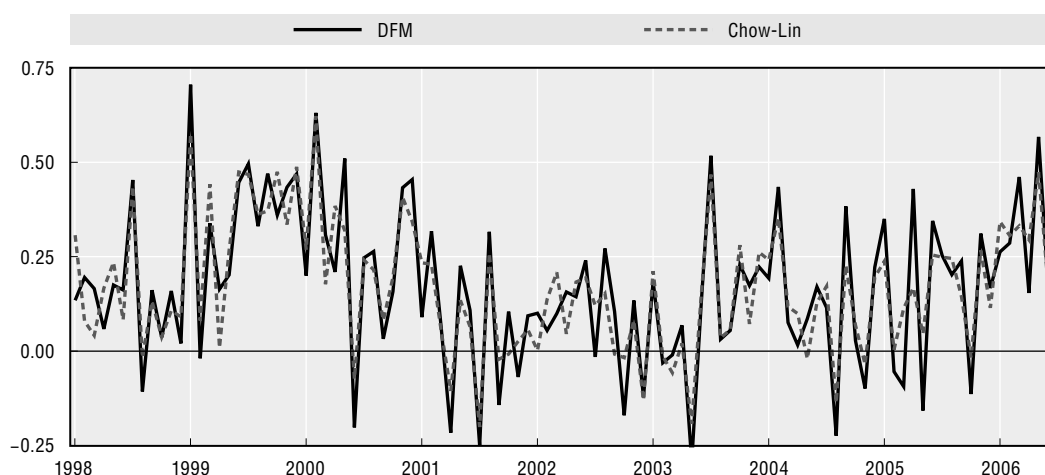


Table 4. MAC to forecasts of monthly GDP growth

Fcst	Industr Prod	Surveys	Financial	Int'l	Labour	Money	GDP
9	0.85	0.18	0.18	0.02	0.01	0.01	0.18
8	0.90	0.18	0.19	0.02	0.00	0.00	0.18
7	0.90	0.22	0.30	0.05	0.01	0.01	0.11
6	0.44	0.63	0.35	0.06	0.01	0.01	0.11
5	0.21	0.32	0.33	0.10	0.01	0.01	0.11
4	0.10	0.21	0.20	0.04	0.00	0.00	0.00
3	0.05	0.14	0.21	0.03	0.00	0.00	0.00
2	0.06	0.07	0.12	0.05	0.00	0.00	0.00
1	0.04	0.11	0.10	0.02	0.00	0.00	0.00

Notes: MACs are calculated as the sum of contributions from individual series belonging to a given group (see Table A.1). They are divided by the mean absolute deviation of monthly GDP growth. The GDP column shows the MAC of quarterly GDP growth. The sum of MACs across data groups exceeds one because in some periods contributions are of conflicting sign.

Table 4 shows the mean absolute values of the contributions of data groups as defined in Table A.1. The table demonstrates the shifts in the contributions of the individual data groups over the forecast horizon. Forecasts 8 and 9 are actually estimates of monthly GDP, where quarterly GDP is already known. In this case, the model attaches very high weights to industrial production data. Note, also, that the relative weight of quarterly GDP growth itself is small, even though temporal aggregation constraints (3) and (4) are relevant. As the horizon increases, survey and financial data gain more weight relative to industrial production and (quarterly) GDP growth. For nowcasts of monthly GDP in the current quarter (i.e., forecasts 4 to 6), the weight of survey data already exceeds that of industrial production series.

Overall, these findings parallel those of Bańbura and Rünstler (2010) for quarterly GDP forecasts. They indicate that equations designed to estimate historical monthly growth in GDP, and therefore that rely heavily on industrial production data, are not necessarily optimal for the purpose of assessing the economic stance in real time.

4. Conclusions

The paper has combined the dynamic factor model of Doz *et al.* (2005) with equations to obtain monthly estimates and short-term forecasts of quarterly growth in the national accounts. The model contains the necessary temporal aggregation and national accounting identities. Hence, monthly estimates and short-term forecasts of the quarterly accounts are mutually consistent, which has advantages when the model is used for monitoring economic developments in real time.

For GDP and a number of components, the model beats forecasts from quarterly models-based time series and from bridge equations using selected indicators. One exception is public and private consumption, for which all forecasts remain largely uninformative.

Ex post estimates of monthly GDP growth are similar to those derived from single equation methods that employ monthly indicators, such as industrial production and confidence indicators. Our findings suggest, however, that equations that have been designed to estimate historical monthly growth in GDP are not necessarily optimal for assessing the economic status in real time.

Our forecast exercise seeks to replicate the real-time application of the models as closely as possible. However, we did not have a real-time data set at hand. Clearly, the properties of short-term forecasts based on real-time data are important for future research. The findings of Diron (2008) from forecasting euro area GDP with a small real-time data set indicate that this may increase the relevance of survey and financial data.

Notes

1. In the dynamic representation of a DFM, series x_t are related to dynamic factors g_t from $x_t = \lambda(L)g_t + \xi_t$ with some finite lag polynomial $\lambda(L)$. Defining $f'_t = (g'_t, g'_{t-1}, \dots)$ as the stacked vector of current and lagged values of g_t gives equation (1) (see Stock and Watson 2005).
2. Giannone *et al.* (2008) use a two-step approach to forecast quarterly GDP growth from the factor model. In a first step, they obtain forecasts of the latent factors, as from the state-space model given by equations (1) and (2). In a second step, quarterly GDP is predicted from quarterly aggregates of forecasts by means of a static regression.
3. This notation assumes that $t = 1$ represents the first month of a certain quarter.
4. With ε_t being white noise, ε_t^Q follows an MA(1) process with coefficient 4/19. This does not affect the consistency of estimates from the quarterly version of equation (6). Doz *et al.* (2006b) present an expectation-maximisation (EM) algorithm to obtain maximum likelihood estimates, but report few gains in forecasting performance.
5. Boivin and Ng (2006) have argued that using smaller data sets may improve the forecasts. However, series selection methods for unbalanced data sets have not been established thus far. Since our aim is to forecast not only GDP but also demand components, we opt for a reasonably large data set.
6. We have value added data beginning only in 1995Q1. Hence, estimates of equation (6) for these series start in this period.
7. The two model versions, hence, contain complete breakdowns of GDP and total value added, respectively, and κ_t in equation (5) reflects only approximation errors. The model without the constraint is almost equivalent to running models that contain a single quarterly series (the case of $m = 1$).
8. The series used for GDP are: industrial production in manufacturing, retail sales, new car registrations, the unemployment rate, and the European Commission business, consumer, building and retail confidence indices. They are used in bridge equations for euro area GDP

proposed by Rünstler and Sédillot (2003) and Diron (2008). We have also experimented with some specific equations proposed in these studies. They did not outperform our approach.

9. The aforementioned studies actually use more series as they derive monthly GDP as the sum of interpolates of its value-added components. For the most part, estimates use sectoral equivalents of our series. Other equations give similar results, as long as major industrial production items are included.

References

- Angelini, E. J. Henry, and M. Marcellino, 2006, "Interpolation and backdating with a large information set", *Journal of Economic Dynamics and Control* 30(12), 2693-2724.
- Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin and G. Rünstler, 2008, "Short-term forecasts of euro area GDP growth", *ECB working paper* 949.
- Baffigi, A., R. Golinelli, and G. Parigi, 2004, "Bridge models to forecast the euro area GDP", *International Journal of Forecasting* 20(3), 447-460.
- Bai, J. and S. Ng, 2002, "Determining the number of factors in approximate factor models", *Econometrica*, 70(1), 191-221.
- Bañbura, M. and G. Rünstler, 2010, "A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP", *International Journal of Forecasting*, forthcoming.
- Banerjee, A., M. Marcellino, and I. Masten, 2005, "Leading indicators for euro-area inflation and GDP growth", *Oxford Bulletin of Economics and Statistics*, 67, 785-813.
- Boivin, J. and S. Ng, 2006, "Are more data always better for factor analysis?", *Journal of Econometrics* 132(1), 169-194.
- Breitung, J. and C. Schumacher, 2008, "Real-time forecasting of GDP based on a large factor model with monthly and quarterly data", *International Journal of Forecasting* 24, 368-398.
- Clements, M. and D. Hendry, 1998, "Forecasting Economic Time Series", Cambridge University Press: Cambridge.
- Chow, G. and A. Lin, 1971, "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series", *Review of Economics and Statistics* 53, 372-375.
- Diron, M., 2008, "Short-term forecasts of euro area real GDP growth: an assessment of real-time performance based on vintage data", *Journal of Forecasting*, 27, 371-390.
- Doz, C., D. Giannone, and L. Reichlin, 2006a, "A two-step estimator for large approximate dynamic factor models based on Kalman filtering", *THEMA Working Papers* 2006-23, Université de Cergy-Pontoise.
- Doz, C., D. Giannone, and L. Reichlin, 2006b, "A quasi maximum likelihood approach for large approximate dynamic factor models", *CEPR Discussion Paper* No. 5724.
- Durbin, J. and S.J. Koopman, 2001, "Time Series Analysis By State Space Methods", Oxford University Press: Oxford.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin, 2000, "The generalized dynamic factor model: identification and estimation", *The Review of Economics and Statistics*, 82, 540-554.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin, 2005, "The generalized factor model: one-sided estimation and forecasting", *Journal of the American Statistical Association* 100, 830-40.
- Giannone, D., L. Reichlin, and D. Small, 2008, "Nowcasting: the real-time informational content of macroeconomic data", *Journal of Monetary Economics*, 55(4), 665-676.
- Hall, R. 1988, "Intertemporal substitution in consumption", *Journal of the Political Economy* 96(2): 339-352.
- Harvey, A.C., 1989, "Forecasting, Structural Time Series Analysis, and the Kalman Filter", Cambridge University Press: Cambridge.
- Harvey, A.C. and S.J. Koopman, 2003, "Computing observation weights for signal extraction and filtering", *Journal of Economic Dynamics & Control*, 27, 1317-1333.
- Harvey, D., Leybourne, S. and P. Newbold, 1997, "Testing the equality of prediction mean squared errors", *International Journal of Forecasting* 13, 281-291.

- Kitchen, J. and R. Monaco, 2003, "Real-time forecasting in practice: The U.S. Treasury staff's real-time GDP forecast system", *Business Economics* 38(4), 10-19.
- Mariano, R.S., and Y. Murasawa, 2003, "A new coincident index of business cycles based on monthly and quarterly series", *Journal of Applied Econometrics*, 18, 427-443.
- McKinnon, J.G., 1991, "Critical values for co-integration tests", in: R.F. Engle and C.W.J. Granger (eds.), *Long-run Economic Relationships* Oxford University Press: Oxford.
- Mitchell, J., R. Smith, M. Weale, S. Wright and E. Salazar, 2005a, "An indicator of monthly GDP and an early estimate of quarterly GDP Growth", *Economic Journal* 115, 108-129.
- Mitchell, J. and M. Weale, 2005b, "A monthly indicator of GDP for the euro area", paper presented to the Workshop on frontiers in benchmarking techniques and their application to official statistics, Eurostat.
- Proietti, T. and C. Frale, 2007, "A monthly indicator of euro area GDP", *Econometrics Journal* 9(3), 357-372.
- Proietti, T., 2008, "Estimation of Common Factors under Cross-Sectional and Temporal Aggregation Constraints: Nowcasting Monthly GDP and its Main Components.", mimeo.
- Rünstler, G. and F. Sédillot, 2003, "Short-term estimates of GDP by means of monthly data", ECB Working Paper No. 176.
- Rünstler, G., K. Barhoumi, S. Benk, R. Cristadoro, A. den Reijer, A. Jakataine, P. Jelonek, A. Rua, K. Ruth, C. van Nieuwenhuyze, 2010, "Short.term forecasting of GDP using large data sets", *Journal of Forecasting* 28(7), 595-61.
- Stock, J. H., and M. W. Watson, 2002, "Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economics Statistics*, 20, 147-162.
- Stock, J. H. and M. W. Watson, 2005, "Implications of dynamic factor models for VAR analysis", Princeton University, mimeo.

ANNEX

Table A.1. **Data**

No.	Series	Group	Publication lag (months)	Transformation code	Dickey-Fuller test
1	IP-Total industry	IndProd	3	2	-2.930
2	IP-Total Industry (excl construction)	IndProd	2	2	-2.911
3	IP-Manufacturing	IndProd	2	2	-2.962
4	IP-Construction	IndProd	3	2	-5.343
5	IP-Total Industry excl construction and MIG Energy	IndProd	2	2	-2.880
6	IP-Energy	IndProd	2	2	-6.274
7	IP-MIG Capital Goods Industry	IndProd	2	2	-2.985
8	IP-MIG Durable Consumer Goods Industry	IndProd	2	2	-3.657
9	IP-MIG Energy	IndProd	3	2	-7.054
10	IP-MIG Intermediate Goods Industry	IndProd	2	2	-3.246
11	IP-MIG Non-durable Consumer Goods Industry	IndProd	2	2	-4.942
12	IP-Manufacture of basic metals	IndProd	2	2	-3.192
13	IP-Manufacture of chemicals and chemical products	IndProd	2	2	-4.555
14	IP-Manufacture of electrical machinery and apparatus	IndProd	2	2	-3.207
15	IP-Manufacture of machinery and equipment	IndProd	2	2	-3.137
16	IP-Manufacture of pulp, paper and paper products	IndProd	2	2	-4.476
17	IP-Manufacture of rubber and plastic products	IndProd	2	2	-3.591
18	Retail trade, except of motor vehicles and motorcycles	IndProd	2	2	-3.712
19	New passenger car registrations	IndProd	1	2	-4.893

Notes: Transformation code: 1 = 3-month difference, 2 = 3-month growth rate.

Augmented Dickey-Fuller tests: 5% and 10% critical values for the t-statistics are -2.86 and -2.56, respectively (McKinnon, 1991).

Table A.1. **Data** (cont.)

No.	Series	Group	Publication lag (months)	Transformation code	Dickey-Fuller test
20	Unemployment rate, total	Labour	2	3	-4.551
21	Index of Employment, Construction	Labour	3	2	-2.299
22	Index of Employment, Manufacturing	Labour	3	2	-1.946
23	Index of Employment, Total Industry	Labour	3	2	-3.422
24	Index of Employment, Total Industry (excluding construction)	Labour	3	2	-1.995
25	Industry Survey: Industrial Confidence Indicator	Surveys	0	1	-3.414
26	Industry Survey: Production trend observed in recent months	Surveys	0	1	-3.501
27	Industry Survey: Assessment of order-book levels	Surveys	0	1	-3.461
28	Industry Survey: Assessment of export order-book levels	Surveys	0	1	-3.637
29	Industry Survey: Assessment of stocks of finished products	Surveys	0	1	-3.940
30	Industry Survey: Production expectations for the months ahead	Surveys	0	1	-3.558
31	Industry Survey: Employment expectations for the months ahead	Surveys	0	1	-3.526
32	Industry Survey: Selling price expectations for the months ahead	Surveys	0	1	-3.059
33	Consumer Survey: Consumer Confidence Indicator	Surveys	0	1	-3.363
34	Consumer Survey: General economic situation over last 12 months	Surveys	0	1	-2.680
35	Consumer Survey: General economic situation over next 12 months	Surveys	0	1	-3.688
36	Consumer Survey: Price trends over last 12 months	Surveys	0	1	-3.094
37	Consumer Survey: Price trends over next 12 months	Surveys	0	1	-3.485

Notes: Transformation code: 1 = 3-month difference, 2 = 3-month growth rate.

Augmented Dickey-Fuller tests: 5% and 10% critical values for the t-statistics are -2.86 and -2.56, respectively (McKinnon, 1991).

Table A.1. **Data** (cont.)

No.	Series	Group	Publication lag (months)	Transformation code	Dickey-Fuller test
38	Consumer Survey: Unemployment expectations over next 12 months	Surveys	0	1	-3.317
39	Construction Survey: Construction Confidence Indicator	Surveys	0	1	-3.191
40	Construction Survey: Trend of activity compared with preceding months	Surveys	0	1	-4.188
41	Construction Survey: Assessment of order books	Surveys	0	1	-3.013
42	Construction Survey: Employment expectations for the months ahead	Surveys	0	1	-4.033
43	Construction Survey: Selling price expectations for the months ahead	Surveys	0	1	-3.877
44	Retail Trade Survey: Retail Confidence Indicator	Surveys	0	1	-3.487
45	Retail Trade Survey: Present business situation	Surveys	0	1	-3.635
46	Retail Trade Survey: Assessment of stocks	Surveys	0	1	-6.208
47	Retail Trade Survey: Expected business situation	Surveys	0	1	-3.777
48	Retail Trade Survey: Employment expectations	Surveys	0	1	-4.257
49	Total trade - Intra Euro 12 trade, Export Value	Int'l	2	2	-2.708
50	Total trade - Extra Euro 12 trade, Export Value	Int'l	2	2	-3.527
51	Total trade - Intra Euro 12 trade, Import Value	Int'l	2	2	-3.130
52	Total trade - Extra Euro 12 trade, Import Value	Int'l	2	2	-3.043
53	US, Unemployment rate	Int'l	1	1	-2.748
54	US, IP total excl construction	Int'l	1	2	-2.621
55	US, Employment, civilian	Int'l	1	2	-3.482
56	US, Retail trade	Int'l	1	2	-3.351
57	US, Production expectations in manufacturing	Int'l	0	1	-5.429

Notes: Transformation code: 1 = 3-month difference, 2 = 3-month growth rate.

Augmented Dickey-Fuller tests: 5% and 10% critical values for the t-statistics are -2.86 and -2.56, respectively (McKinnon, 1991).

Table A.1. **Data** (cont.)

No.	Series	Group	Publication lag (months)	Transformation code	Dickey-Fuller test
58	US, Consumer expectations index	Int'l	0	1	-5.082
59	World market prices of raw materials in Euro, total, HWWA	Int'l	0	2	-3.574
60	World market prices of raw materials in Euro, total, excl energy, HWWA	Int'l	0	2	-3.503
61	World market prices, crude oil, USD, HWWA	Int'l	1	2	-4.064
62	Gold price, USD, fine ounce	Int'l	0	2	-2.596
63	Brent Crude, 1 month fwd, USD/BBL converted in euro	Int'l	0	2	-4.356
64	ECB Nominal effective exch. rate	Financial	0	2	-3.117
65	ECB Real effective exch. rate CPI deflated	Financial	0	2	-3.083
66	ECB Real effective exch. rate producer prices deflated	Financial	0	2	-3.151
67	Exch. Rate: USD/EUR	Financial	0	2	-3.348
68	Exch. Rate: GBP/EUR	Financial	0	2	-3.306
69	Exch. Rate: YEN/EUR	Financial	0	2	-3.088
70	Eurostoxx 500	Financial	0	2	-3.138
71	Eurostoxx 325	Financial	0	2	-3.226
72	US S&P 500 composite index	Financial	0	2	-2.782
73	US, Dow Jones, industrial average	Financial	0	2	-3.383
74	US, Treasury Bill rate, 3-month	Financial	0	1	-2.072
75	US Treasury notes & bonds yield, 10 years	Financial	0	1	-3.527
76	10-year government bond yield	Financial	0	1	-3.537
77	3-month interest rate, Euribor	Financial	0	1	-3.020
78	1-year government bond yield	Financial	0	1	-2.665
79	2-year government bond yield	Financial	0	1	-2.864
80	5-year government bond yield	Financial	0	1	-3.286
81	Index of notional stock - Money M1	Money	2	2	-3.013
82	Index of notional stock - Money M2	Money	2	2	-2.834
83	Index of notional stock - Money M3	Money	2	2	-2.636
84	Index of Loans	Money	2	2	-2.031
85	Money M2 in the U.S.	Money	2	2	-2.087

Notes: Transformation code: 1 = 3-month difference, 2 = 3-month growth rate.

Augmented Dickey-Fuller tests: 5% and 10% critical values for the t-statistics are -2.86 and -2.56, respectively (McKinnon, 1991).

Table A.2. **RMSE for GDP for subsamples (relative to the naive forecast)**

Fcst	2000Q1 - 2003Q3					2003Q4 - 2006Q2				
	Naive	AR(1)	QVAR	BEQ	DFM	Naive	AR(1)	QVAR	BEQ	DFM
8										
7	0.375	0.77	0.74	0.89	0.59	0.282	0.91	0.90	0.99	0.80
6	0.375	0.77	0.74	0.91	0.60	0.282	0.91	0.90	1.04	0.80
5	0.375	0.77	0.75	0.94	0.60	0.282	0.91	0.90	1.05	0.97
4	0.385	0.99	0.98	0.94	0.60	0.284	0.96	1.00	1.04	0.88
3	0.385	0.99	0.98	0.96	0.64	0.284	0.96	1.00	1.07	0.85
2	0.385	0.99	0.98	1.00	0.71	0.284	0.96	1.00	1.06	0.84
1	0.392	1.03	1.04	1.01	0.74	0.287	1.05	1.11	1.06	0.85

Note: The table shows the RMSE of the naive forecast and, for the remaining models, the RMSE relative to the naive forecast.

Table A.3. **Encompassing and Diebold-Mariano tests**

(2000Q1 - 2006Q4)									
Fcst	GDP	Priv cons	Gov cons	GFCF	Exp	Imp	VAD total	VAD Ind	VAD serv
<i>Encompassing tests</i>									
7	0.939	0.899	0.055	0.630	0.683	0.922	0.961	0.612	1.094
4	0.952	0.151	-0.237	0.783	0.945	0.978	0.908	0.558	1.254
1	0.892	0.716	-0.111	0.720	0.808	0.827	0.813	0.512	1.186
<i>Diebold-Mariano tests</i>									
7	* 1.538	** 2.327	-1.113	0.309	* 1.446	1.080	1.216	0.527	0.957
4	0.952	0.505	-0.628	0.534	** 2.767	* 1.583	1.211	0.281	* 1.282
1	0.259	0.582	-0.442	0.262	0.932	0.526	1.204	0.511	* 1.281

Notes: The upper panel of the table shows coefficient λ from encompassing regression (13). The lower panel shows the Diebold-Mariano statistics using the small-sample correction by Harvey et al. (1997).

* Significant at the 5% level.

** Significant at the 10% level.

Chain-linking in Austrian Quarterly National Accounts and the Business Cycle

by
Marcus Scheiblecker*

In 2005, European Union member countries began to calculate national account volume estimates using prices from the previous year, rather than from a fixed base year. For quarterly national accounts, the average of the total previous year – and not of the previous quarter – began to serve as the price basis. This allows for the use of a Laspeyres-type quantity index.

In order to obtain a time series of absolute values of volume estimates, it is necessary to chain-link growth rates. This is straightforward when calculating annual figures, but when calculating quarterly figures, EU countries can choose from one of three methods. This results in different outputs, time-series properties and, possibly, price-adjusted quarterly national account figures.

The current study demonstrates the different results obtained using the three methods, when applied to Austrian quarterly GDP data. I observe the consequences of consecutive time-series-based processing, such as seasonal adjustment and business cycle analysis. Although dating turning points are rather robust using all three methods, seasonal and workday adjustment and detection of outliers based on time-series modelling can be negatively affected, as can business cycle dating.

JEL Classification: C43, E01, E32

Keywords: quarterly national accounts, chain-linking

* Austrian Institute of Economic Research (WIFO), Arsenal Objekt 20, 1103 Vienna, Austria (e-mail: marcus.scheiblecker@wifo.ac.at).

It is conventional wisdom that using a distant base year for price adjustment may result in bias.¹ For this reason, in the area of national accounts, countries have moved from price adjustment related to a fixed base year to one based on previous year prices.² This allows calculating growth rates in real terms – but not absolute values. As absolute values are strongly in demand, these growth rates are chained together to construct either an index series or absolute values related to a reference period.

Whereas this is straightforward for annual figures, it is more complicated for sub-annual series. In quarterly national accounts, either the most recent quarter or the entire previous year can be used as the price basis. If the first approach is used – as is the case in the United States – a Fisher index is usually applied in order to avoid distortions from seasonal fluctuations in prices and quantities. EU member countries, instead, use the entire previous year as the basis for calculating a Laspeyres-type index.

Whereas chain-linking using the previous quarter as a basis for price adjustment can be done in the same way as for annual figures, there exist three possible methods for linking quarterly figures if the entire previous year serves as the basis. This can result in different outputs, time-series properties and price-adjusted quarterly national account figures. Nevertheless, countries within the EU are free to choose among them and currently, all three types of approaches are used.

This study demonstrates the different results obtained using the three methods, when applied to Austrian quarterly GDP data. I observe the consequences of consecutive time-series-based processing, such as seasonal adjustment and business cycle analysis.

Section 1 briefly presents the three chain-linking methods (the over-the-year method, the annual overlap method and the quarterly overlap method³). The magnitude of time inconsistency based on Austrian quarterly national accounts (QNA) data for all three techniques is presented in Section 2, where a further method for benchmarking that adjusts time series to annually chain-linked totals is also given. Section 3 addresses how the different chain-linking techniques can interfere with time-series modelling. The sensitivity of dating business cycle turning points with the popular Bry and Boschan (1971) method to the different chain-linking methods is discussed in Section 4.

1. Different methods of chain-linking for quarterly data

The choice of the proper base period is less straightforward for sub-annual studies than for annual data. This problem has been solved in the EU by determining that the average (over all four quarters) has to be used for the construction of series at the previous year's prices. However, countries can decide how to calculate growth rates and how to chain them together for producing series in absolute values.⁴

To construct the three different methods – the over-the-year method (OTY), the annual overlap method (AO) and the quarterly overlap method (QO) – two “series”⁵ are necessary: one that calculates each quarter at average prices of the current year, and one that calculates each quarter at average prices of the previous year.

The OTY method calculates growth rates by comparing a certain quarter at average prices of the previous year to the respective quarter at average prices of the previous year. These growth rates are chained together analogously to an index series:

$$I_{y,s}^{OTY} = \frac{\bar{p}^{-y-1} q^{y,s}}{\bar{p}^{-y-1} q^{y-1,s}} I_{y-1,s}^{OTY} \quad (1)$$

where y represents the year, s the respective quarter ($s = 1, \dots, 4$), q the quantities and $\bar{p}^{y-1}$ the average price level of the previous year.

Index chains constructed this way allow for a comparison of quantities not over consecutive quarters, but over the year.⁶ Furthermore, it can be shown that the quarters do not add up to independent annual chained years, which is very often requested by users of national accounts data.

The AO method calculates growth rates by comparing a certain quarter calculated at average prices of the previous year to the average of the quarters of the previous year calculated at average prices of this respective year. These growth rates are chained together analogously to an index series:

$$I_{y,s}^{AO} = \frac{\bar{p}^{-y-1} q^{y,s}}{\sum \bar{p}^{-y-1} q^{y-1,s}} I_{y-1}^{AO} \quad (2)$$

4

As all quarters of the current year are related to the entire previous year, the index series constitutes a consistent time series until a new year can serve as the next price basis. This happens each first quarter of a new year, causing a break in the series. The size of this break depends on the change in the fraction of the fourth quarter quantities and/or prices relative to the total year between two successive years.⁷ In practice, this break will be rather small, unless there is a large change in the price or quantity structure relative to the entire year. In respect of the criterion of time consistency, the annual overlap method is the only one of all three chain-linking approaches in which the sum of the quarters gives the exact respective annually chained totals.

The QO method calculates all growth rates by comparing the current quarter at average prices of the previous year to the fourth quarter of the previous year valued at average prices of the respective year. Once again, these growth rates are chain-linked in the same manner. The index series is calculated as

$$I_{y,s}^{QO} = \frac{\bar{p}^{-y-1} q^{y,s}}{\bar{p}^{-y-1} q^{y-1,4}} I_{y-1,4}^{QO} \quad (3)$$

Of all three methods, it is the only one that allows a comparison of quantities over consecutive quarters without breaks.⁸ This is probably why the IMF⁹ strongly recommends this approach for setting up quarterly national accounts, despite the fact that this kind of quarterly series does not add up to annual totals, derived by independently chaining the years. Lippe and Küter (2005) correctly note that the QO method, in this respect, performs even worse than the OTY method.

When deciding in favour of one of these three methods, statistical offices have to choose whether to give more emphasis to the time-consistency criterion (summation of the quarters to annual totals) or to the break-free development of the series over time. As the OTY method fulfils neither the time consistency nor the time-series criterion,¹⁰ the decision usually has to be made between the AO and the QO method.

As time-consistent quarterly series are strongly demanded by users,¹¹ a further way to achieve this is by applying methods that distribute the differences between the sum of the quarters and the respective annuals in a second step. In national accounts, procedures that force sub-annual data to be consistent with annual totals are called benchmarking techniques. These techniques range from basic forms, like the distribution of differences equally among all quarters or proportionally (*pro rata*), to more sophisticated mathematical or statistical forms.

It is interesting to note that index chains constructed by the QO method produce exactly the same results as the AO where time consistency is achieved, by following implicit *pro rata* distribution of differences.¹² This may be viewed as somewhat problematic, as the IMF regards the explicit application of a *pro rata* benchmarking technique as “unacceptable” in its *Quarterly National Accounts Manual*.¹³

2. Empirical differences due to time inconsistency

In this section, the differences among the three chain-linking methods discussed in Section 1 are given for Austrian GDP data. They range from the first quarter of 1988 to the end of 2006. Table 1 presents the differences from the time inconsistency of all three chain-linking methods.

Table 1. Differences between quarters and annual totals

	OTY		AO		QO	
	Million EUR	Per cent	Million EUR	Per cent	Million EUR	Per cent
1988	12.02	0.01	0.00	0.00	-128.41	-0.09
1989	12.94	0.01	0.00	0.00	-229.65	-0.15
1990	11.94	0.01	0.00	0.00	422.48	0.26
1991	12.43	0.01	0.00	0.00	761.13	0.45
1992	4.99	0.00	0.00	0.00	730.39	0.42
1993	10.23	0.01	0.00	0.00	839.55	0.48
1994	4.91	0.00	0.00	0.00	497.05	0.28
1995	7.97	0.00	0.00	0.00	266.50	0.15
1996	7.60	0.00	0.00	0.00	98.26	0.05
1997	9.34	0.00	0.00	0.00	-266.37	-0.14
1998	11.24	0.01	0.00	0.00	-721.54	-0.37
1999	9.50	0.00	0.00	0.00	-1 587.38	-0.78
2000	0.00	0.00	0.00	0.00	0.00	0.00
2001	-19.63	-0.01	0.00	0.00	591.10	0.28
2002	-20.52	-0.01	0.00	0.00	1 006.43	0.47
2003	-22.55	-0.01	0.00	0.00	1 458.10	0.67
2004	5.57	0.00	0.00	0.00	1 457.55	0.66
2005	19.59	0.01	0.00	0.00	636.15	0.28
2006	23.59	0.01	0.00	0.00	286.39	0.12

The results given in Table 1 are consistent with the expectations formed on the basis of the theoretical considerations of Section 1. There are no differences between the sum of the quarters – derived by the AO method – and the independent annually chained years. As regards the OTY approach, only small differences in absolute numbers and percentages occur. It is clear that the difference is zero only for the reference year 2000. The largest differences between quarterly and annually chain-linked data result from the QO method. Negative as well as positive deviations have been observed. In absolute values, they range

from EUR –1.6 billion (–0.78 per cent of annually chain-linked GDP) to EUR 1.5 billion (0.67 per cent). There does not seem to be a trend, but rather the differences look as if they evolve smoothly.

The size of the differences derived from the QO method may suggest that this method is inferior to others. But it does show the extent of the amount distributed implicitly *pro rata* in the AO chain-linking method. As time consistency may be regarded as an important feature for some users, the question arises if there is an alternative benchmarking technique that can be applied to the series which disturbs the quantitative evolution over time less than the implicit *pro rata* distribution included the AO method. A possible alternative could be the approach taken by Denton (1971), which is strongly recommended by the IMF.¹⁴ In this method, the allocation of the annual difference is distributed over the quarters by minimizing their relative change from quarter to quarter. A drawback of this method is, of course, that each new annual observation requires a recalculation of the minimization process that leads to regular revisions of the entire time series historically.

Figure 1 shows the differences between the quarters calculated by the AO method (which implies a *pro rata* distribution of annual differences) and the QO method after benchmarking with the proportional Denton procedure. Differences are given as percentages of the arithmetic mean of the quarters calculated by both methods. They range from –0.31 per cent (EUR 155 million) in the first quarter of 2000 to +0.22 per cent (EUR 120 million) in the fourth quarter of this year. It may seem odd that in the reference year in which all methods show no deviations from annual totals, the largest differences between the AO method and the benchmarked QO (B-QO) occur. This is because there were very large differences to be distributed over the quarters for 1999 and 2001. The smooth distribution mechanism of the proportional Denton approach requires an adjustment of the quarters of 2000 as well. Figure 2 shows the adjustments relative to the unadjusted QO method, which have to be made in order to reach annual totals.

Figure 1. **Relative differences between the AO and the benchmarked QO method**

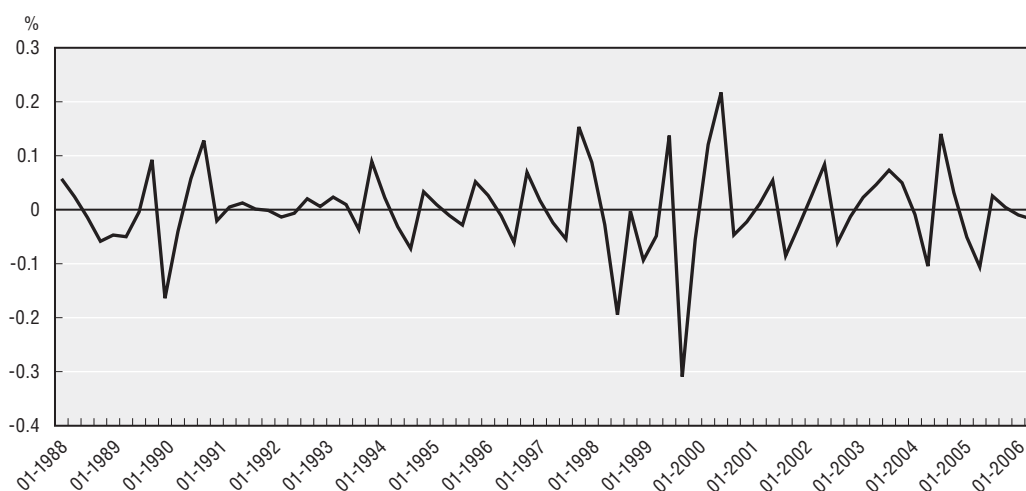
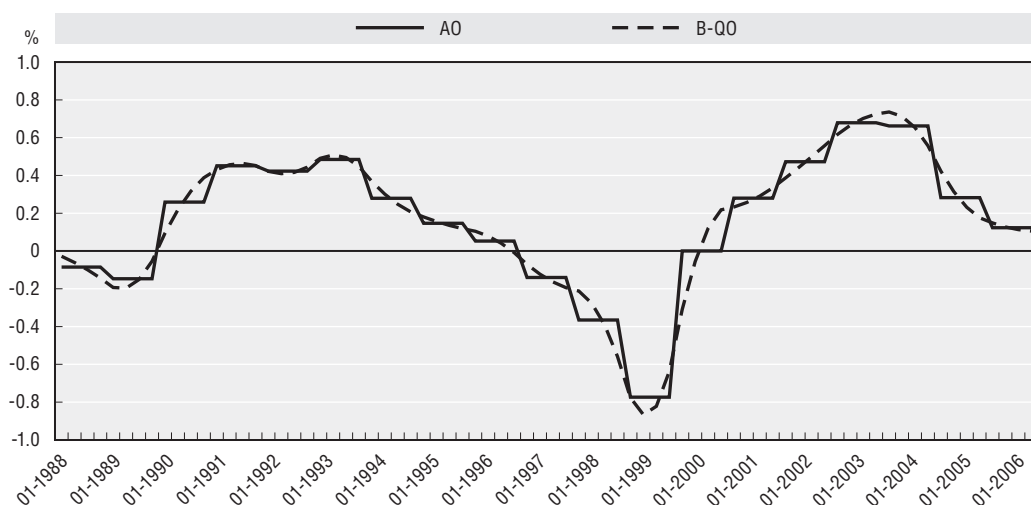


Figure 2. **Distribution of annual differences by the AO and the Denton benchmarked QO method**



3. Consequences for time-series modelling

To show possible consequences for modelling time series, an automatic time-series modelling software has been applied to the output of the three quarterly chain-linking methods, as well as to the one with the benchmarked QO results. This software fits seasonal ARIMA models to economic time series in order to adjust for seasonal and workday effects, as well as for possible outliers. To this end, the software package DEMETRA,¹⁵ which includes the TRAMO-SEATS¹⁶ module, has been used.

For the OTY method, a model has been proposed that represents the seasonal part as an ARIMA (0,1,1) and the regular (trend-cycle) part as a (0,1,0) model. Furthermore, an outlier in the form of a transitory component type was detected in the first quarter of 1993 and removed automatically.

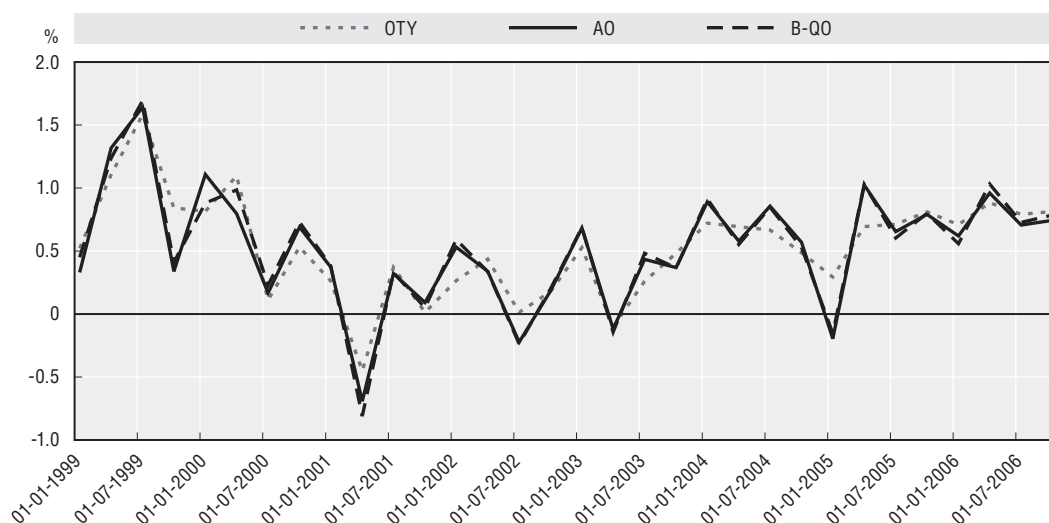
In the case of the AO method, a model that represents the seasonal part as an (0,1,1), but the trend-cycle part as an (0,1,1) process, has been selected. Again an outlier was found in the first quarter of 1993, but this time it was an additive outlier affecting only the quarter.

For the unadjusted QO method, the same type of model has been chosen. While the MA coefficient for the seasonal part is nearly the same as in the AO case, the MA coefficient of the regular part is somewhat lower. In this case, an outlier has also been detected for the same period, but this time a transitory component like in the OTY case.

The QO series adjusted by the proportional Denton procedure has been modelled as an ARIMA (0,1,0)(0,1,1), which is quite different from all other chain-linking methods. The parameter of the trend-cycle MA term is the smallest of all models. As in the case of the AO approach, an additive outlier was detected in the first quarter of 1993.

Figure 3 shows the quarter-on-quarter growth rates of the seasonally and workday-adjusted series for the OTY, the AO and the Denton benchmarked QO method as they were produced by the different time-series models. It may be observed that the ones derived by the AO and benchmarked QO (B-QO) methods show nearly the same pattern, with coinciding local peaks and troughs. However, the model based on the OTY chain-linked data results in growth rates that behave quite differently in that they seem to have a lower

Figure 3. **Quarter-on-quarter growth rates on seasonally and workday-adjusted series**



variance. The OTY series does not seem to have a generally leading or lagging property when compared to the others, but the local maxima at the beginning of 2000 and 2002 were reached one quarter later than in the other series.

If the AO and the B-QO series are compared, the only difference in the local maxima of growth rates can be observed for the beginning of 2000 with the B-QO method showing the peak one quarter later.

4. Extracting the business cycle

Looking at the growth rates over the various methods may suggest that the different chain-linking models come to more or less the same conclusions. But differences in the ARIMA structure of the trend-cycle part may entail severe consequences for business cycle analysis. Fukuda (2008) has shown that existing MA or AR terms in the non-seasonal ARIMA component are indicators of whether the trend can be statistically separated from the cycle. Ignoring this can lead to spurious cycle detection.¹⁷

For the sake of simplicity, it is assumed here that trend and cycle can be separated from each other. Furthermore, it is done in this case by non-parametric filtering approaches. Here, only the most popular ones – the famous Hodrick-Prescott (HP) and Baxter-King (BK) filters – are applied in order to show possible differences in results.

Figure 4 gives the result in log differences between consecutive quarters of the different series for which a HP-1600 trend has been eliminated. The output is quite similar to the ones in which the trend has been eliminated by first-order differences of logs, which correspond approximately to the calculation of growth rates shown in Figure 3.

Table 2 gives the corresponding dates of the business cycle turning points as detected by the routine proposed by Bry and Boschan (1971), by demanding a minimum cycle length of five quarters and a minimum length of three quarters. The unbenchmarked QO series serves as the reference cycle to which all other results are compared.¹⁸

Figure 4. Differences in log of HP-1600 transformed series

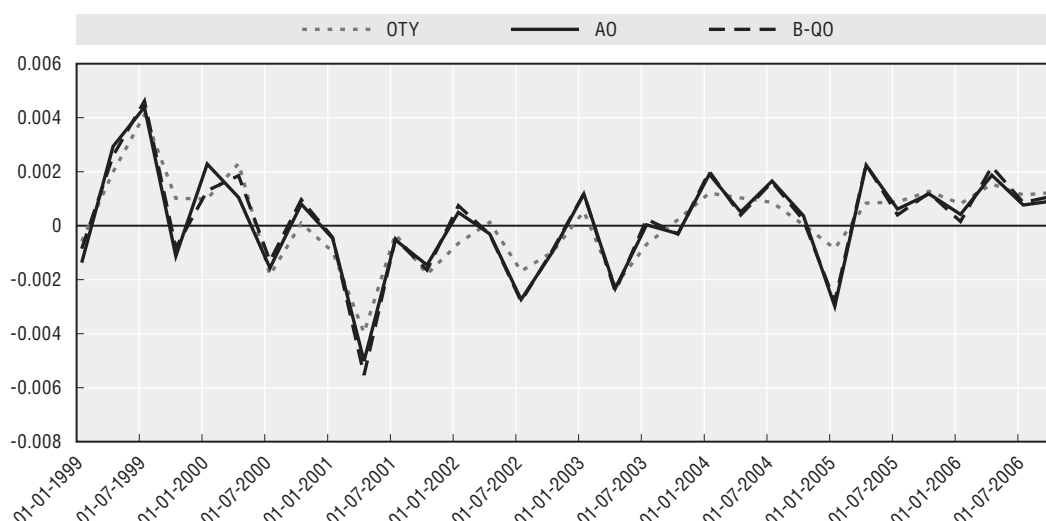


Table 2. Turning point analysis for HP-1600 filtered series

Series	Peak	Trough	Peak	Trough	Peak	Trough	Peak	Trough	# of extra cycles
	Q3-1991	Q1-1993	Q1-1994	Q1-1995	Q1-1996	Q2-1997	Q2-2000	Q4-2003	
QO	0	0	0	0	0	0	0	0	0
AO	0	None	None	0	0	-1	0	0	-1
B-QO	0	None	None	0	0	-1	0	0	-1
OTY	0	None	None	0	0	0	0	-1	-1

Note: + (-) denotes a lag (lead) with respect to the reference series.

The reference series (QO) given in the first line shows the largest number of completed cycles between 1988 and 2006. It counts four cycles, while all chain-linking methods show only three. The one starting with a trough in the first quarter of 1993 and showing its peak in the first quarter of 1994 has not been recognised by all other chain-linked series. A possible explanation for this is that the outlier detected in all series has not been removed sufficiently by the QO method or has been removed too much by the other methods.¹⁹

There is some evidence in the literature that the period around 1993 was indeed a cyclical trough. For the Austrian case, Artis, Krolzig and Toro (2004) reported a trough in the third quarter of 1993 for the Austrian economy.* For the euro area, Mönch and Uhlig (2004) also identify a trough in the first quarter of 1993. The CEPR Business Cycle Dating Committee (2003) places it in the fourth quarter of 1992, Artis, Krolzig and Toro (2004) in the second quarter of 1993, Artis, Marcellino and Proietti (2004) in the fourth quarter of 1993 and Forni *et al.* (2004) in the first quarter of 1994.

From the above, one can conclude that the benchmarking process – done either explicitly or implicitly – can potentially interfere with the appropriate detection and estimation of outliers, which may in turn hamper the detection of business cycle turning points.

Further results indicate that the trough in the second quarter of 1997 was indicated by the AO and the benchmarked QO method one period earlier than by the other methods.

* Scheiblecker (2008) dated it to the third or fourth quarter of 1993, respectively.

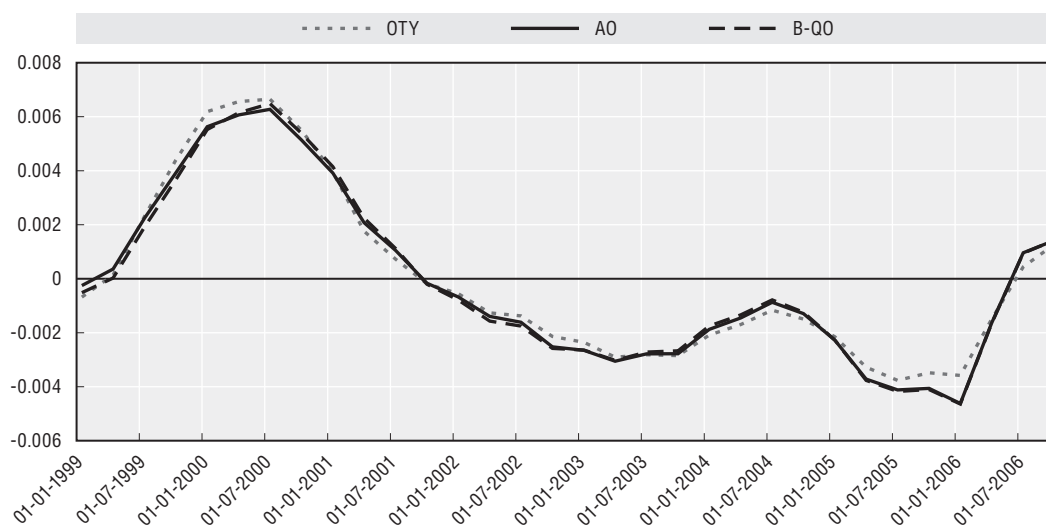
The same goes for the trough in the fourth quarter of 2003, which was detected by the OTY method one quarter earlier.

A further method for business cycle extraction was used to explore whether the differences in business cycle analysis stem from the differences in time series modelling – used for outlier detection and seasonal and workday adjustment – or from the chain-linking process itself. The band-pass filter proposed by Baxter and King (1995) was applied to all four chain-linking methods as an alternative. As this filter removes not only the trend but also frequencies higher than the business cycle like seasonal variations, it can be applied directly to the series without prior adjustment for seasonal fluctuations.²⁰

To make sure that seasonal variations are properly extracted, the filter was set to retain only variations with a frequency of between 8 and 2 years (32 and 8 quarters, respectively) instead of the common 8 and 1½ years. Additionally, the filter length has been expanded to 8 quarters in order to reduce leakage, but this advantage comes at the expense of a higher loss of observations at either end of the series (which are usually extended by an ARIMA forecast). This is not, however, a drawback in this study.

Figure 5 shows the cyclical component after applying the BK-filter to the logs of the series of OTY, AO and the Denton benchmarked QO chain-linked series. They look quite similar in their ups and downs between 1999 and 2006. The only remarkable difference is that the OTY series seems to reach its lowest point two quarters before the AO and the B-QO series.

Figure 5. **Baxter-King 8-2 filtered series**



This result can be confirmed by looking at the turning points as detected by the Bry-Boschan dating procedure given in Table 3. The trough found by the three other series in the first quarter of 2006 is dated by the OTY series two quarters earlier.²¹ This time, five complete cycles have been found which can be attributed to dropping a prior outlier detection procedure. As a consequence, the second quarter of 1993 has been dated as a trough, followed by a peak in the first quarter of 1996.

Generally, the dating calendar of the BK-8-2 transformed series looks quite different from the one for the HP-1600 seasonally filtered ones. However, it is conventional wisdom that the Bry and Boschan (1971) procedure detects fewer turning points in HP-filtered series

Table 3. **Turning point analysis for BK-8-2 filtered series**

Series	Peak	Trough	Peak	Trough	Peak	Trough	Peak	Trough	Peak	Trough	# of extra cycles
	Q4-1988	Q1-1990	Q4-1991	Q2-1993	Q1-1996	Q2-1997	Q1-2000	Q2-2003	Q3-2004	Q1-2006	
QO	0	0	0	0	0	0	0	0	0	0	0
AO	0	0	0	0	0	0	2	0	0	0	0
B-QO	0	0	0	0	0	0	2	0	0	0	0
OTY	0	0	0	0	0	0	2	0	0	-2	0

Note: + (-) denotes a lag (lead) with respect to the reference series.

that are still plagued by high frequency components. This makes it difficult for this dating procedure to find turning points in such highly erratic series.

From Tables 2 and 3, one can conclude that the Denton benchmarked QO method and the AO method give the same business cycle turning points for the Austrian GDP between 1988 and 2006. However, the unadjusted QO method, which represents the only breakless quantity index, and the OTY method produce different dates and number of cycles, with the latter sometimes showing some leading properties when compared to the QO method.

This analysis has demonstrated that the different methods of quarterly chain-linking give quite robust results for business cycle analysis based on non-parametric filtering. However, modelling such data by time-series models can give quite different results, not only in respect of the model type but also concerning the detection of outliers. The analysis was based on data drawn from a period that did not have significant changes in prices and weights. In times of strong economic change, results concerning differences between quarterly chain-linking methods can be more pronounced.

Notes

1. See, e.g. Boskin *et al.* (1998).
2. In 2003 EU member countries were obliged to deliver national accounts data from 2005 onwards on a previous-year price basis.
3. The quarterly overlap method is frequently called the “one-quarter overlap method”.
4. Unlike annual figures, QNA figures are usually not published as index series.
5. These series cannot be regarded as time series in a narrower sense, as they show breaks at the beginning of each new year.
6. See, e.g. Bikker (2005, p. 9) or Kirchner (2007).
7. For a more formal explanation, the reader is referred to Kirchner (2007, p. 5).
8. This is apart from the usual ones due to permanent weight changes. See also Bikker (2005).
9. See Bloem, Dippelsman and Maehle (2001).
10. Therefore, the EU statistical office EUROSTAT does not recommend this method.
11. Hence the EU is demanding the production of benchmarked chained national accounts data.
12. See Bikker (2005).
13. See Bloem, Dippelsman and Maehle (2001, p. 84).
14. See Bloem, Dippelsman and Maehle (2001, p. 87).
15. DEMETRA 2.1 © European Communities, 1999-2007.
16. The TRAMO-SEATS software package used here has been developed by Gomez and Maravall (1992) and is strongly recommended by the EC for seasonal adjustment in QNA.

17. Per Harvey (1989, p. 46), a distinction must be made between a trend-plus-cycle and a cyclical-trend model.
18. This is done on practical grounds, as here one more cycle has been found and is not based on some priority reasoning.
19. The detection of a trough makes the Bry-Boschan routine look for a following peak in the series, as changing signs for business variations is a precondition for this procedure.
20. It has to be noted that, for a sensible business cycle analysis, a prior adjustment for working days and outliers is indispensable, but here the focus is just on the difference between methods and not on finding the correct business cycle.
21. It must be noted that this difference has to be disregarded, as it concerns the margin of the time series, which has been enlarged by the ARIMA forecast. In Section 4, it was shown that the different chain-linking methods interfere quite strongly with this kind of time-series models.

References

- Artis, M., H.M. Krolzig and J. Toro (2004), "The European Business Cycle", Oxford Economic Papers, No. 56, Oxford University Press.
- Artis, M., M. Marcellino and T. Proietti (2004), "Characterizing the Business Cycle for Accession Countries", CEPR Discussion Paper, No. 4457.
- Baxter, M. and R.G. King (1995), "Measuring Business Cycles Approximate Band-Pass Filters for Economic Time Series", NBER Working Paper Series, No. 5022.
- Bikker, R. (2005), "Chained Indices for Quarterly National Accounts", Statistics Netherlands, Division of Macro-economic Statistics and Dissemination, Development and Support Department, available under www.cbs.nl/NR/rdonlyres/6D35BF36-0DF8-4FE8-AB22-64418D37512E/0/2005Chainedindices050422art.pdf.
- Bloem, A., R. Dippelsman and N. Maehle (2001), "Quarterly National Accounts Manual – Concepts, Data Sources and Compilations", IMF, Washington DC.
- Boskin, M.J., E. Dulberger, R. Gordon, Z. Griliches and D. Jorgenson (1996), "Toward a More Accurate Measure of the Cost of Living. Final Report to the Senate Finance Committee", in *The Journal of Economic Perspectives*, 12(1), pp. 3-26.
- Bry, G. and C. Boschan (1971), "Cyclical Analysis of Time Series: Selected Procedures and Computer Programs", Technical Paper, No. 20, NBER New York.
- Center for Economic Policy Research (2003), *Business Cycle Dating Committee of the Center for Economic Policy Research*, CEPR, September 2003, available as download under www.cepr.org/press/dating.pdf.
- Denton, F.T. (1971), "Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization", *Journal of the American Statistical Society*, Vol. 66, No. 333, pp. 99-102.
- Forni, M., D. Giannone, F. Lippi and L. Reichlin (2004), "Opening the Black Box: Structural Factor Models versus Structural VARs", Université Libre de Bruxelles, mimeo.
- Fukuda, K. (2008), "Differentiating between Business Cycles and Growth Cycles: Evidence from 15 Developed Countries", *Applied Economics*, 40:7, pp. 875-883.
- Gomez, V. and A. Maravall (1992), "Time Series Regression with ARIMA Noise and Missing Observations: Program TRAMO", Working Paper, No. 92/81, European University Institute, Florence.
- Harvey, A.C. (1989), "Forecasting, Structural Models and the Kalman Filter", Cambridge University Press, Cambridge.
- Kirchner, R. (2007), "Chain-linking and Seasonal Adjustment, Properties of Unadjusted and Adjusted Time Series", available under www.cmfb.org/pdf/TF-SA%20QNA%20-%20Final%20Report.pdf.
- Lippe von der, P. and J. Küter (2005), "Methoden der Verkettung von Quartalsergebnissen für deflationierte Aggregate der Volkswirtschaftlichen Gesamtrechnung", Diskussionsbeiträge aus dem Fachbereich Wirtschaftswissenschaften der Universität Duisburg-Essen, No. 146.
- Mönch, E. and H. Uhlig (2004), "Towards a Monthly Business Cycle Chronology for the Euro Area", CEPR Discussion Paper, No. 4377.
- Scheiblecker, M. (2008), "The Austrian Business Cycle in the European Context", Forschungsergebnisse der Wirtschaftsuniversität Wien, Peter Lang Verlag, Frankfurt am Main.

Evaluating German Business Cycle Forecasts under an Asymmetric Loss Function

by

Jörg Döpke, Ulrich Fritsche and Boriss Siliverstovs*

Based on annual data for growth and inflation forecasts for Germany covering the 1970-2007 period and up to 17 different forecasts per year, we test for a possible asymmetry of the forecasters' loss function and estimate the degree of asymmetry for each forecasting institution using the approach of Elliot et al. (2005). Furthermore, we test for the rationality of the forecasts under the assumption of a possibly asymmetric loss function and for the features of an optimal forecast under the assumption of a generalised loss function. We find evidence of the existence of an asymmetric loss function of German forecasters only in the case of pooled data and a quad-quad loss function. We can reject the hypothesis of rationality of the growth forecasts based on a pooled dataset, but not on data for single institutions. The rationality of inflation forecasts is frequently rejected in the case of single institutions, and also for pooled data.

JEL Classification: C53, E42.

Keywords: business cycle forecast evaluation, asymmetric loss function, rational expectations.

* Döpke: University of Applied Sciences Merseburg, D-06217 Merseburg, Germany (e-mail: joerg.doepke@hs-merseburg.de); Fritsche: Faculty of Economics and Social Sciences, Hamburg University, von-Melle-Park 9, D-20146 Hamburg, Germany (e-mail: ulrich-fritsche@wiso.uni-hamburg.de); Siliverstovs: KOF Swiss Economic Institute, Eidgenössische Technische Hochschule Zürich (ETHZ) (Federal Institute of Technology Zürich), Weinbergstrasse 35, CH-8092 Zürich, Switzerland (e-mail: siliverstovs@kof.ethz.ch). The authors thank Gerit Vogt; participants at the Third Dresden Workshop on Business Cycles and Macroeconomics 2008, the Ninth Macroeconometric Workshop in Halle, and the BICEPS seminar in Riga; and two anonymous referees for helpful comments on a previous draft of this paper.

A central assumption in economics and finance is that economic agents behave rationally when they form their expectations. Consequently, a large body of literature has investigated the accuracy and rationality of forecasts, including several studies regarding German business cycle forecasts (see, *e.g.*, Fildes and Stekler, 2002, for a survey, and Döpke and Fritsche, 2006, for an overview of related papers for German data). While a large body of research has supported the hypothesis of rationality in German business cycle forecasts, some papers have recently challenged these findings. For example, Osterloh (2008) argues that forecasts with a horizon beyond one year do not fulfil the requirements for rational forecasts. A similar argument is made by Dovern and Weissner (2008), who base their argument on a methodological variation of tests for rationality using pooled data. Ager, Kappler and Osterloh (2009) not only reject the idea of information efficiency, but also find biased forecasts in some cases.

Virtually all of these studies, however, regardless of whether the point is made explicitly or implicitly, analyse the issue under the assumption of a symmetric loss function, *i.e.*, the notion that over- and underestimations are equally costly to the respective forecaster. While this assumption has been more or less undisputed for a long period, it may be criticised for very good economic reasons.

Consider possible customers of business cycle forecasters. For a single firm, for example, there is *a priori* no reason to assume that the costs of underpredicting demand in terms of a loss of sales or reputation should be exactly equal to the costs of overpredicting demand in terms of additional costs and storage (Elliot *et al.*, 2005, 2008). On a macroeconomic level, it is very likely that central banks have asymmetric preferences regarding inflation, perhaps leaning toward more caution against inflation acceleration. Alan Blinder summarises his experience as a central bank officer, claiming that a central bank “take(s) far more political heat if it tightens preemptively to avoid higher inflation than if it eases preemptively to avoid higher unemployment” (1998). Furthermore, while an overestimation of a budget deficit may help the career of a finance minister, an underestimation may end it. As the famous German economist and politician Ludwig Erhardt once put it: “If it gets better than expected, even the false prophet will be forgiven” (quoted according to Miersch, 2008). Furthermore, international or supranational institutions like the International Monetary Fund (IMF), World Bank or European Commission face principal-agent problems regarding their relationship with clients or member states – which, in turn, could justify asymmetric loss functions (Artis and Marcellino, 2001; Elliott *et al.*, 2005; Christodoulakis and Mamatzakis, 2008, 2009). An additional line of argument, which may point to the possibility of an asymmetric loss, refers to the political economy of business cycle forecasts (see Döpke, 2000, for related arguments). In this view, individual forecasters represent competing political points of view and use the forecasts as instruments to achieve their political goals. Hence, under- and overestimations of growth and inflation are likely to be unequally costly in the eyes of the forecaster, since they give different incentives for good or bad policies. Additional reasoning might be found in the model of Laster *et al.* (1999) who argue that forecasters face different incentives. For example, a private institution might be very interested in garnering public attention for its

forecasts. All in all, a certain scepticism regarding the symmetry assumption is well justified. Therefore, we analyse signs of asymmetric loss functions for those institutions publishing regular forecasts for the German economy.

Consequently, several approaches have been developed to incorporate more general loss function into forecasting evaluations. Based on influential work by Chistofferson and Diebold (1997), Granger (1999) and Batchelor and Peel (1998), among others, Elliott *et al.* (2005, 2008) have proposed estimating the degree of asymmetry of the loss function and testing for a significant degree of asymmetry. Moreover, Patton and Timmerman (2007) analyse the properties of an optimal forecast under a generalised loss function and discuss how to test for these properties. We make use of these approaches to re-evaluate the issue of rationality of German business cycle forecasts, namely, growth and inflation forecasts covering the 1970-2007 time span, and up to 17 different forecasts.

In our results, we find only limited evidence of asymmetric loss functions of German business cycle forecasters. For example, for a piecewise linear (lin-lin) loss function, as in the approach of Patton and Timmerman (2007), we are not able to reject the null hypothesis of a symmetric loss function at the usual significance levels in all but two cases for each growth and inflation rate forecast. This conclusion, based on individual data, is further reinforced by evidence based on pooled data, where we also cannot reject the null hypothesis of an asymmetric lin-lin loss function at the usual significance levels, for both growth and inflation rate forecasts. In the case of the piecewise quadratic (quad-quad) loss function, the evidence against a symmetric loss function is stronger, as we can reject the corresponding null hypothesis for up to five institutions for each growth and inflation forecast, depending on a set of instrumental variables. Furthermore, in contrast to the results obtained on pooled data for the lin-lin loss function, in the case of a quad-quad loss function for growth forecasts we are able to reject the null hypothesis even at the 1% level for all considered sets of instruments, indicating a general tendency of forecasting institutions to produce overly optimistic forecasts of GDP growth rates. We also find some evidence of asymmetry in pooled inflation forecasts, but the test outcome appears not to be fully robust *vis-à-vis* the choice of instruments.

Furthermore, we check whether the usual results concerning the rationality of the forecasts still hold when the assumption regarding the loss function is relaxed. In a nutshell, we find that neither a specifically asymmetric loss function nor the assumption of a generalised loss function alters the findings obtained under a symmetric loss function by very much, although the results of the test proposed by Elliot *et al.* (2005) give some contrary results for inflation forecasts.

The remainder of the paper is organised as follows. Section 1 describes the data and the econometric method proposed by Elliot *et al.* (2005) to back out the parameter of asymmetry of a loss function and statistical testing for the existence of asymmetry, and discusses the results for the data set at hand. Section 2 tests for the rationality of the forecasts under different assumptions: a symmetric loss function, a specific asymmetric loss function and a generalised loss function. Section 3 summarises and concludes.

1. Estimating loss function asymmetry parameters and testing for asymmetry

We evaluate the forecasts of several institutions that deliver macroeconomic forecasts regarding the German economy. Further details on the data set can be found in Döpke and Fritsche (2006). For all institutions, we have collected growth and inflation forecasts. The

growth forecast is the predicted growth rate of real GNP (for the time span 1983 to 1989) and of real GDP (for all other years). In the case of published interval forecasts, the average is used. The numbers refer to West Germany up to 1992, and to the whole of Germany from 1993 to the present. As a measure of the inflation forecast, we use the predicted change of the deflator of private consumption, when this figure is available. In some cases, however, no explicit reference was given to whether an inflation forecast referred to the consumption deflator or to the CPI/HICP. In such cases we assume that no distinction between the figures was intended by the forecaster, and we use the available inflation forecast. As regards the actual outcome, it is possible to refer to the latest available revised data or to the first published ("real-time") data. As it is common in the analysis of business cycle forecasts, we make use of the latter type of numbers, i.e., we compare the forecasts made at the end of a certain year "x" or at the beginning of the following year "x + 1" with the first published figure for the year "x + 1".

The analysis by Elliot *et al.* (2005) starts from the general loss function:¹

$$L(p, \alpha, \Theta) = [\alpha + (1 - 2\alpha) \cdot \mathbf{1}(y_{t+1} - \hat{y}_{t+1})] \cdot |y_{t+1} - \hat{y}_{t+1}|^p \quad (1)$$

In this loss function the parameter p represents the underlying assumption of the subsequent analysis. In particular, $p = 1$ stands for a linear-linear (lin-lin) loss function, while in case of $p = 2$ the calculations are based on a quadratic-quadratic (quad-quad) loss function. Furthermore, the loss function contains the parameter α representing the degree of asymmetry of the loss function. In particular, $\alpha = 0.5$ yields a symmetric loss function, while $\alpha > 0.5$ represents the case of forecasters' incentives to issue optimistic forecasts. Finally, $\alpha < 0.5$ stands for the case of forecasts that are too pessimistic. Thus, a particular set of parameters leads to a well-known loss function. For example, $L(2, 1/2, \Theta) = (y_{t+1} - \hat{y}_{t+1})^2$ yields a symmetric quadratic loss function (Elliot *et al.*, 2005, p. 1110). The key problem they address is, of course, that the value of this parameter is unknown and has to be estimated from the data.

Elliot *et al.* (2005) establish conditions for optimality of forecasts, which, in turn, deliver the necessary moment conditions. By observing the sequence of forecasts, the authors propose a generalised method of moments (GMM) estimator that yields the following expression to estimate the asymmetry parameter of the loss function out of the moment conditions:

$$\hat{\alpha}_T = \frac{\left[\frac{1}{T} \sum_{t=\tau}^{T+\tau+1} v_t |y_{t+1} - \hat{y}_{t+1}|^{p_0-1} \right] \hat{S}^{-1} \left[\frac{1}{T} \sum_{t=\tau}^{T+\tau+1} v_t \mathbf{1}(y_{t+1} - \hat{y}_{t+1} < 0) |y_{t+1} - \hat{y}_{t+1}|^{p_0-1} \right]}{\left[\frac{1}{T} \sum_{t=\tau}^{T+\tau+1} v_t |y_{t+1} - \hat{y}_{t+1}|^{p_0-1} \right] \hat{S}^{-1} \left[\frac{1}{T} \sum_{t=\tau}^{T+\tau+1} v_t |y_{t+1} - \hat{y}_{t+1}|^{p_0-1} \right]} \quad (2)$$

with $\hat{S} = \frac{1}{T} \nu \nu' (\mathbf{1}(y_{t+1} - \hat{y}_{t+1} < 0) - \alpha_t)^2 |y_{t+1} - \hat{y}_{t+1}|^{2p_0-2}$ as a weighting matrix. Since S depends on α_T , estimation has to be performed iteratively, assuming $S = 1$ in the first round, since the identity matrix is a consistent starting point, and using v_t as instrument(s). The estimation is based on considerations that have led to the GMM estimator proposed by Hansen (see Hansen and West, 2002, for a survey and a discussion of its relation to macroeconomic applications). Elliot *et al.* (2005) show that the estimator of α_T is asymptotically normal and, hence, renders it possible to test for the hypothesis $\alpha = 0.5$, i.e., for loss function symmetry.

Instruments are warranted for the proposed GMM estimator. Following Elliot *et al.* (2005, p. 461), our instruments are: i) a constant; ii) a constant and a lagged forecast error; iii) a

constant and the lagged variable to be predicted (i.e., the growth and inflation rate, respectively); and iv) a constant, the lagged forecast error and the lagged variable to be predicted. We use the terms Model 1 to 4 in the course of the paper for estimations based on each of the four sets of instruments, respectively. These are the sets of instruments proposed in the literature and that also have been used in a similar context (Elliott *et al.*, 2005; Christodoulakis and Mamatzakis, 2008, 2009). The estimation results for the data set under investigation are given for lin-lin loss functions in Table 1 and quad-quad loss functions in Table 2.

As regards growth forecasts and calculations based on the assumption of a lin-lin loss function, the findings presented in Table 1 suggest only very limited evidence of

Table 1. **Evidence for an asymmetric loss function, lin-lin function**

	Model 1	s.e.	p-value	Model 2	s.e.	p-value	Model 3	s.e.	p-value	Model 4	s.e.	p-value
Growth forecasts												
Berlin Institute	0.320	0.077	0.022	0.319	0.077	0.018	0.320	0.077	0.019	0.318	0.077	0.017
Council of Economic Advisers	0.595	0.081	0.241	0.595	0.081	0.241	0.607	0.080	0.182	0.639	0.079	0.079
Employers Institute	0.571	0.084	0.393	0.575	0.084	0.372	0.578	0.083	0.352	0.578	0.083	0.351
Essen Institute	0.568	0.081	0.407	0.568	0.081	0.407	0.573	0.081	0.372	0.578	0.081	0.339
European Commission, autumn	0.541	0.082	0.621	0.542	0.082	0.611	0.547	0.082	0.569	0.550	0.082	0.545
European Commission, spring	0.486	0.082	0.869	0.486	0.082	0.869	0.486	0.082	0.869	0.486	0.082	0.869
Government's Economic Report	0.514	0.082	0.869	0.514	0.082	0.866	0.515	0.082	0.853	0.516	0.082	0.850
Halle Institute	0.571	0.132	0.589	0.572	0.132	0.588	0.576	0.132	0.567	0.586	0.132	0.515
Hamburg Institute	0.432	0.081	0.407	0.432	0.081	0.403	0.428	0.081	0.377	0.427	0.081	0.370
IMF, autumn	0.588	0.084	0.296	0.588	0.084	0.296	0.592	0.084	0.274	0.624	0.083	0.136
IMF, spring	0.444	0.083	0.502	0.441	0.083	0.475	0.444	0.083	0.501	0.441	0.083	0.472
Joint Forecast, spring	0.514	0.082	0.869	0.514	0.082	0.869	0.514	0.082	0.869	0.514	0.082	0.869
Joint Forecast, autumn	0.486	0.082	0.869	0.486	0.082	0.866	0.485	0.082	0.855	0.485	0.082	0.851
Kiel Institute	0.486	0.082	0.869	0.486	0.082	0.869	0.486	0.082	0.863	0.486	0.082	0.861
Munich Institute	0.459	0.082	0.621	0.459	0.082	0.620	0.459	0.082	0.618	0.459	0.082	0.618
OECD	0.568	0.081	0.407	0.571	0.081	0.383	0.581	0.081	0.318	0.585	0.081	0.294
Trade Union Institute	0.486	0.082	0.869	0.486	0.082	0.866	0.485	0.082	0.851	0.484	0.082	0.849
Pooled data	0.505	0.020	0.807	0.505	0.020	0.806	0.505	0.020	0.797	0.505	0.020	0.793
Inflation forecasts												
Berlin Institute	0.378	0.080	0.127	0.333	0.077	0.031	0.374	0.080	0.114	0.327	0.077	0.025
Council of Economic Advisers	0.649	0.078	0.058	0.704	0.075	0.007	0.705	0.075	0.006	0.714	0.074	0.004
Employers Institute	0.514	0.084	0.866	0.515	0.084	0.861	0.515	0.084	0.857	0.515	0.084	0.856
Essen Institute	0.514	0.082	0.869	0.533	0.082	0.688	0.525	0.082	0.763	0.535	0.082	0.671
European Commission, autumn	0.568	0.081	0.407	0.602	0.080	0.206	0.576	0.081	0.353	0.606	0.080	0.189
European Commission, spring	0.432	0.081	0.407	0.428	0.081	0.374	0.432	0.081	0.401	0.428	0.081	0.374
Government's Economic Report	0.432	0.081	0.407	0.416	0.081	0.297	0.430	0.081	0.390	0.414	0.081	0.291
Halle Institute												
Hamburg Institute	0.568	0.081	0.407	0.588	0.081	0.274	0.575	0.081	0.354	0.588	0.081	0.274
IMF, autumn	0.441	0.085	0.490	0.427	0.085	0.392	0.439	0.085	0.472	0.427	0.085	0.389
IMF, spring	0.528	0.083	0.738	0.531	0.083	0.711	0.528	0.083	0.735	0.531	0.083	0.711
Join Forecast, spring	0.405	0.081	0.241	0.388	0.080	0.160	0.390	0.080	0.169	0.382	0.080	0.139
Joint Forecast, autumn	0.649	0.078	0.058	0.702	0.075	0.007	0.670	0.077	0.028	0.710	0.075	0.005
Kiel Institute	0.432	0.081	0.407	0.412	0.081	0.276	0.432	0.081	0.401	0.392	0.080	0.178
Munich Institute	0.541	0.082	0.621	0.546	0.082	0.578	0.543	0.082	0.596	0.546	0.082	0.575
OECD	0.600	0.089	0.264	0.622	0.089	0.169	0.600	0.089	0.261	0.630	0.088	0.141
Trade Union Institute	0.514	0.082	0.869	0.519	0.082	0.819	0.514	0.082	0.869	0.523	0.082	0.780
Pooled data	0.509	0.021	0.651	0.512	0.021	0.567	0.509	0.021	0.628	0.512	0.020	0.567

Legend: Model 1: Constant. Model 2: Constant and lagged forecast errors. Model 3: Constant and lagged variable. Model 4: Constant, lagged forecast error, lagged variable.

Table 2. **Evidence for an asymmetric loss function, quad-quad loss function**

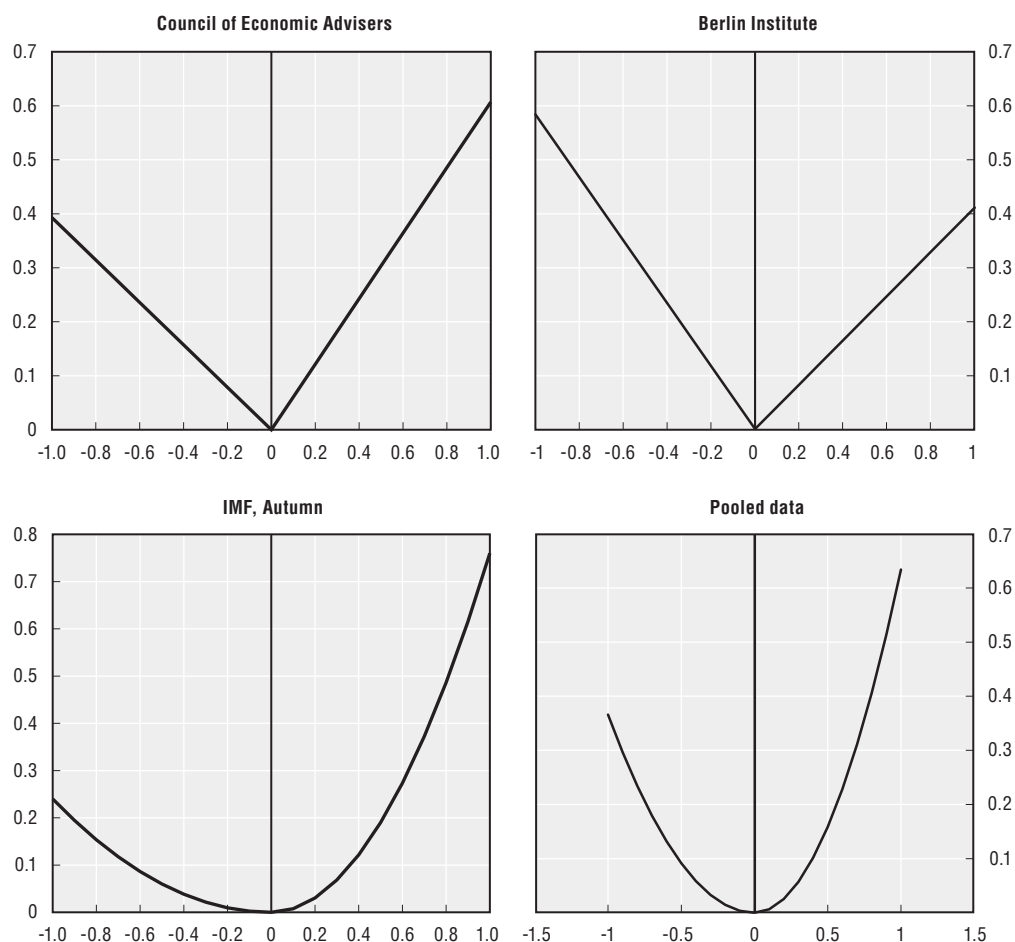
	Model 1	s.e.	p-value	Model 2	s.e.	p-value	Model3	s.e.	p-value	Model 4	s.e.	p-value
Growth forecasts												
Berlin Institute	0.492	0.120	0.947	0.470	0.112	0.790	0.508	0.118	0.948	0.413	0.109	0.426
Council of Economic Advisers	0.647	0.104	0.159	0.645	0.103	0.157	0.674	0.100	0.080	0.686	0.099	0.060
Employers Institute	0.639	0.108	0.199	0.661	0.105	0.124	0.658	0.104	0.127	0.648	0.104	0.153
Essen Institute	0.611	0.099	0.265	0.599	0.097	0.309	0.622	0.098	0.212	0.604	0.097	0.287
European Commission, autumn	0.685	0.094	0.050	0.699	0.089	0.025	0.717	0.087	0.012	0.715	0.087	0.013
European Commission, spring	0.608	0.103	0.298	0.598	0.104	0.344	0.612	0.100	0.265	0.609	0.101	0.280
Government's Economic Report	0.614	0.106	0.279	0.616	0.105	0.270	0.628	0.103	0.214	0.614	0.102	0.263
Halle Institute	0.674	0.175	0.320	0.877	0.092	0.000	0.869	0.093	0.000	0.872	0.092	0.000
Hamburg Institute	0.592	0.108	0.392	0.597	0.107	0.366	0.606	0.106	0.318	0.608	0.106	0.309
IMF, autumn	0.745	0.085	0.004	0.748	0.084	0.003	0.760	0.082	0.001	0.881	0.058	0.000
IMF, spring	0.562	0.111	0.576	0.551	0.111	0.646	0.563	0.109	0.565	0.567	0.108	0.537
Join Forecast, spring	0.624	0.107	0.245	0.632	0.099	0.185	0.654	0.099	0.119	0.645	0.098	0.138
Joint Forecast, autumn	0.644	0.097	0.136	0.647	0.095	0.123	0.655	0.095	0.103	0.641	0.095	0.140
Kiel Institute	0.609	0.108	0.313	0.609	0.106	0.302	0.627	0.105	0.228	0.599	0.106	0.351
Munich Institute	0.555	0.110	0.620	0.562	0.109	0.568	0.571	0.108	0.514	0.571	0.108	0.510
OECD	0.645	0.103	0.162	0.668	0.096	0.080	0.685	0.095	0.051	0.676	0.095	0.063
Trade Union Institute	0.534	0.109	0.752	0.538	0.108	0.727	0.545	0.107	0.673	0.544	0.108	0.684
Pooled data	0.617	0.025	0.000	0.620	0.025	0.000	0.626	0.026	0.000	0.634	0.025	0.000
Inflation forecasts												
Berlin Institute	0.383	0.103	0.259	0.352	0.102	0.146	0.372	0.103	0.214	0.351	0.102	0.140
Council of Economic Advisers	0.496	0.111	0.972	0.589	0.110	0.420	0.572	0.111	0.520	0.598	0.110	0.374
Employers Institute	0.568	0.121	0.575	0.567	0.121	0.582	0.603	0.118	0.384	0.602	0.116	0.380
Essen Institute	0.443	0.112	0.613	0.500	0.114	0.999	0.205	0.081	0.000	0.155	0.071	0.000
European Commission, autumn	0.474	0.112	0.814	0.627	0.109	0.244	0.575	0.112	0.504	0.633	0.108	0.220
European Commission, spring	0.544	0.105	0.675	0.601	0.098	0.303	0.557	0.102	0.576	0.602	0.098	0.296
Government's Economic Report	0.388	0.116	0.336	0.384	0.116	0.319	0.388	0.116	0.335	0.381	0.116	0.306
Hamburg Institute	0.501	0.119	0.993	0.679	0.108	0.098	0.624	0.113	0.274	0.677	0.107	0.099
IMF, autumn	0.626	0.106	0.234	0.649	0.104	0.151	0.644	0.104	0.163	0.634	0.104	0.198
IMF, spring	0.618	0.108	0.274	0.607	0.109	0.327	0.608	0.109	0.323	0.605	0.109	0.336
Joint Forecast, autumn	0.518	0.111	0.872	0.773	0.087	0.002	0.704	0.098	0.038	0.773	0.087	0.002
Joint Forecast, spring	0.444	0.107	0.605	0.448	0.107	0.629	0.470	0.105	0.775	0.462	0.104	0.713
Kiel Institute	0.414	0.119	0.473	0.070	0.060	0.000	0.235	0.094	0.005	0.078	0.061	0.000
Munich Institute	0.552	0.119	0.659	0.670	0.105	0.104	0.622	0.112	0.276	0.669	0.104	0.105
OECD	0.556	0.125	0.654	0.571	0.125	0.568	0.588	0.120	0.462	0.577	0.119	0.515
Trade Union Institute	0.509	0.112	0.938	0.681	0.098	0.065	0.546	0.110	0.676	0.698	0.097	0.040
Pooled data	0.497	0.029	0.926	0.574	0.028	0.009	0.543	0.029	0.137	0.578	0.028	0.006

Legend: Model 1: Constant. Model 2: Constant and lagged forecast errors. Model 3: Constant and lagged variable. Model 4: Constant, lagged forecast error, lagged variable.

asymmetric loss functions. Only in the case of the Berlin Institute do the individual results point to a loss function giving incentives for overly pessimistic forecasts. Depending on the number of instruments, there are also some weak (significant at the 10% level) results for a loss function of the Council of Economic Advisers fostering overly optimistic forecasts. The results may support conventional wisdom regarding these institutions: the Berlin Institute has long been seen as the most Keynesian of German institutes. Thus, being pessimistic might achieve a more activist economic policy. By contrast, the Council of Economic Advisers has widely been seen as supply-side oriented, and the opposite behaviour may be seen as plausible. The emerging general picture, however, is not completely in line with such an explanation, since other institutes with strong opinions (the Trade Union Institute

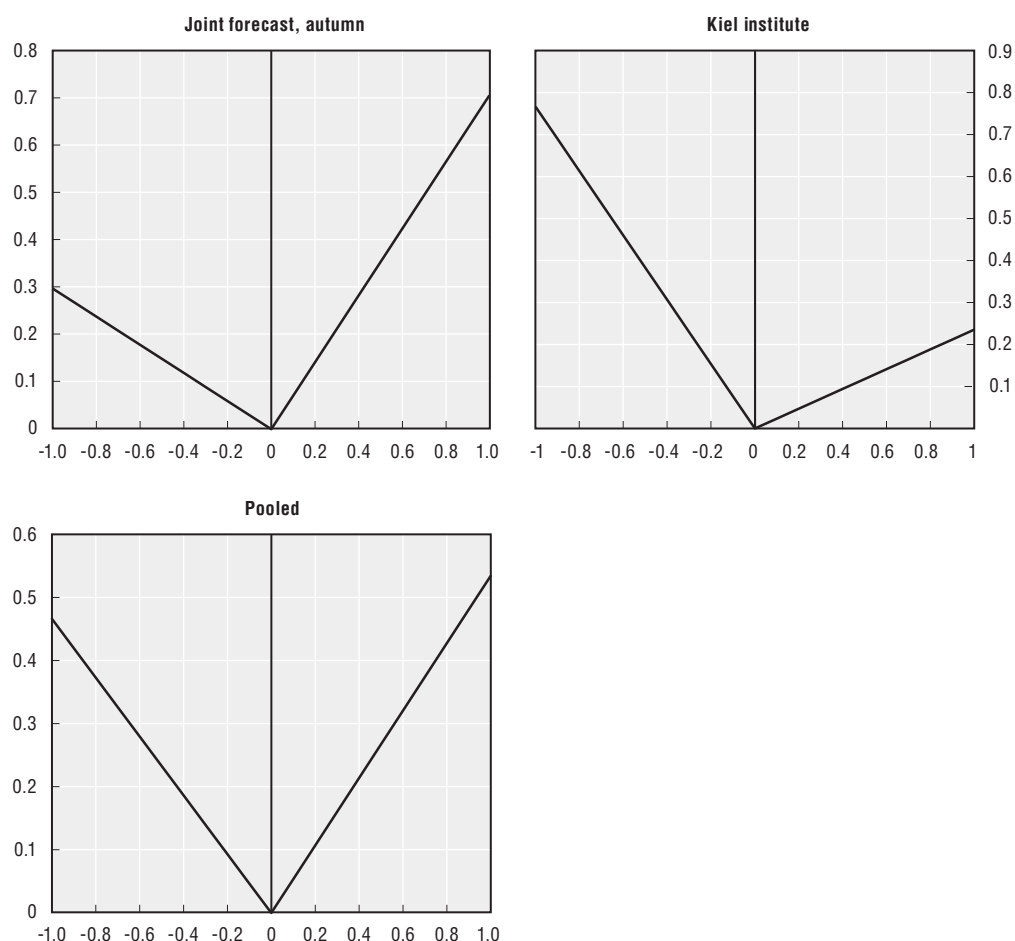
or Employers Institute, for example) show no such results. The test results are illustrated by visual inspection of the estimated loss functions. Some pronounced cases of these functions are given in Figure 1. (The others are available upon request from the authors.)

Figure 1. **Selected estimated loss functions, growth forecasts**



Most estimated loss functions look quite symmetric, representing the fact that virtually all estimated α parameters are very close to 0.5. This is also supported by the estimators based on pooled data – which could be interpreted as a joint test under no severe heterogeneity. For example, in Table 1 the pooled estimates of asymmetry parameter for a lin-lin function are close to 0.5 for both the growth and inflation forecasts, implying that, on average, all institutions overpredict growth rates as often as they underpredict them.²

Turning to the inflation forecasts, there is slightly more hindsight to asymmetric loss functions under the lin-lin loss function. The Joint Forecast as well as the Council of Economic Advisers have overestimated inflation systematically in the past, while the Berlin Institute has more likely underestimated it. Again, visual inspection of the estimated loss functions (Figure 2) confirms the picture given by the formal statistical tests.

Figure 2. **Selected estimated loss functions, inflation forecasts**

Based on the assumption of a quad-quad loss function for growth forecasts, the broad picture remains more or less unchanged, i.e., there is hardly any convincing evidence of a significant degree of asymmetry across the board of the forecasting institutions (see Table 2). Nevertheless, it is quite a remarkable fact that in all but one case the point estimate of the asymmetry parameter exceeds 0.5, indicating a general tendency of forecasting institutions to issue overly optimistic forecasts of GDP growth rates. As a reflection of this finding, the test based on pooled data decisively rejects the null hypothesis of a symmetric loss function. By pooling, we are able to estimate the parameter of interest with greater precision – that is, observe a substantial decrease in reported standard errors in pooled estimation compared to individual estimation results – and hence a significant gain in testing power is achieved, allowing us to reject the null hypothesis of $\alpha = 0.5$.³

As regards the inflation forecasts, we are able to reject the null hypothesis of symmetry more often under a quad-quad loss function than under a lin-lin loss function. Depending on a set of instrumental variables for up to five institutions – the Essen Institute, the Joint Forecast in autumn, the Kiel Institute, the Hamburg Institute and the Trade Unions Institute – the corresponding null hypothesis is rejected.

The overall impression is that the outcome of tests for loss function symmetry depends on a choice of the underlying loss function, such that under a quad-quad loss

function we could reject the null hypothesis of symmetry more often than under a lin-lin loss function. This seems to happen for tests based on both individual and pooled data. This is unsurprising, as under different loss functions different features of forecast errors are taken into account. Thus, under the lin-lin loss function, only the proportion of negative forecast errors in total number of forecasts is relevant for estimation of the asymmetry parameter. At the same time, under the quad-quad loss function, a share of the sum of negative forecast errors in total sum of absolute values of forecast errors is taken into account. This implies that in a situation where a forecaster, on average, overestimates as often as underestimates a variable of interest, the asymmetry test based on the lin-lin loss function will not reject the null hypothesis. If, however, in this situation the size of negative forecast errors is on average larger than the size of positive forecast errors, the test based on the quad-quad loss function will reject the corresponding null hypothesis. It is interesting to observe that we find this pattern in our data set of forecast errors, implying that, on average, the number of overpredictions and underpredictions is balanced, but when overpredictions happen they tend to be of a larger magnitude.

2. Testing for rationality and optimality of a forecast under different loss functions

2.1. Testing for rationality under a symmetric loss function

Testing the rationality of a forecast under a symmetric loss function is typically based on two requirements. First, the forecast should be unbiased, i.e., no systematic errors should occur (the expected value of the forecast error should not be different from zero). Second, the forecast should make efficient use of all information available at the forecasting date, i.e., for an optimal forecast one should be unable to find any variables, which helps to forecast the errors. In a nutshell, former studies of the rationality of German business cycle forecasts have typically found them unbiased, but not necessarily efficient.

To obtain a first insight into the rationality of the forecasts under investigation, we present rationality tests based on a version of the Mincer-Zarnowitz equation (Batchelor and Peel, 1998). In particular, a standard rationality test can be based on estimating the equation

$$(y_{t+n} - \hat{y}_{t+n,t}) = \alpha_0 + \alpha_1(y_{t+n-1} - \hat{y}_t) + u_{t+n} \quad (3)$$

As Batchelor and Peel (1998) (referring to Christofferson and Diebold, 1997) argue, under the null hypothesis of rationality, and assuming a symmetric loss function, forecast errors should be orthogonal to all information known at t , and in particular to the expected change in y . Thus, if the forecast is rational, $\alpha_0 = 0$, $\alpha_1 = 0$ holds. This is tested with a standard F -test. The results of this task are given in Table 3.

The results for growth forecasts, documented in Table 3, give few hints of departures from rationality. In fact, in all but one case (IMF, autumn) we cannot reject the corresponding null hypothesis of forecast rationality. This is also confirmed by the estimation and testing based on pooled data. Turning to inflation forecasts, rationality can be rejected at the usual significance levels in 10 out of 17 cases. Moreover, the null hypothesis is also rejected strongly in the case of pooled data. This is in line with the results reported in Döpke and Fritsche (2006) and possibly due to a high degree of serial correlation in the forecast errors, especially in the 1970s and 1980s. In general, we find very limited evidence for non-rationality under symmetric loss for growth forecasts, and overwhelming evidence for non-rationality for inflation forecasts.

Table 3. **Test for rationality of the forecasts under a symmetric loss function, 1970 to 2007**

	Constant	Slope	Test for rationality (F-value)	Test for rationality (p-value)
Growth forecasts				
Berlin Institute	0.015 (0.06)	0.081 (0.48)	0.118	0.89
Council of Economic Advisors	-0.293 (-1.20)	0.031 (0.18)	0.825	0.45
Employers Institute ¹	-0.398 (-1.41)	-0.247 (-1.47)	1.758	0.19
Essen Institute	-0.199 (-0.89)	0.107 (0.62)	0.727	0.49
European Commission, autumn	-0.456 (-1.64)	-0.027 (-0.16)	1.388	0.26
European Commission, spring	-0.140 (-0.77)	0.142 (0.86)	0.808	0.46
Government's Economic Report	-0.253 (-1.09)	-0.142 (-0.84)	0.818	0.45
Halle Institute ²	-0.332 (-1.35)	-0.352 (-1.37)	1.397	0.28
Hamburg Institute	-0.190 (-0.83)	-0.075 (-0.45)	0.408	0.67
IMF, autumn ³	-0.660 (-2.08)	0.007 (0.04)	2.55	0.09
IMF, spring ⁴	-0.067 (-0.34)	0.295 (1.79)	1.756	0.19
Joint Forecast, autumn	-0.349 (-1.27)	-0.009 (-0.05)	0.829	0.45
Joint Forecast, spring	-0.202 (-1.04)	-0.005 (-0.03)	0.594	0.58
Kiel Institute	-0.221 (-0.87)	0.027 (0.16)	0.416	0.66
Munich Institute	-0.116 (-0.53)	-0.127 (0.76)	0.399	0.67
OECD	-0.345 (-1.32)	-0.082 (-0.49)	0.895	0.42
Trade Union Institute	-0.086 (-0.33)	-0.055 (-0.32)	0.098	0.91
Pooled data	-0.051 (-1.18)	-0.062 (-1.47)	1.780	0.17
Inflation forecasts				
Berlin Institute	0.096 (0.67)	0.356 (2.31)	3.296	0.049
Council of Economic Advisors	-0.012 (-0.86)	0.459 (3.12)	4.874	0.01
Employers Institute ¹	-0.061 (-0.50)	0.196 (1.20)	0.865	0.43
Essen Institute	0.012 (0.10)	0.536 (3.86)	7.617	0.002
European Commission, autumn	0.001 (0.01)	0.547 (4.00)	8.030	0.001
European Commission, spring	-0.022 (-0.28)	0.259 (1.56)	1.302	0.28
Government's Economic Report	0.073 (0.53)	0.376 (2.45)	3.459	0.04
Halle Institute ²	0.046 (0.23)	0.219 (0.76)	0.321	0.73
Hamburg Institute	-0.016 (-0.11)	0.458 (3.12)	4.875	0.014

Table 3. **Test for rationality of the forecasts under a symmetric loss function, 1970 to 2007 (cont.)**

	Constant	Slope	Test for rationality (F-value)	Test for rationality (p-value)
IMF, autumn ³	-0.127 (-0.89)	0.289 (1.86)	2.344	0.11
IMF, spring ⁴	-0.068 (-0.52)	0.413 (2.66)	4.11	0.025
Joint Forecast, autumn	-0.038 (-0.27)	0.563 (4.23)	8.973	0.001
Joint Forecast, spring	0.042 (0.43)	0.221 (1.33)	1.011	0.374
Kiel Institute	0.015 (0.10)	0.581 (4.67)	11.278	0.0001
Munich Institute	-0.054 (-0.39)	0.231 (1.44)	1.134	0.33
OECD	-0.041 (-0.30)	0.194 (1.08)	0.656	0.524
Trade Union Institute	-0.014 (-0.09)	0.547 (3.92)	7.96	0.002
Pooled data	-0.106 (-3.18)	-0.302 (-7.12)	37.78	0.000

1. 1972 to 2007.

2. 1993 to 2007.

3. 1973 to 2007.

4. 1971 to 2007 (t-values in parentheses).

2.2. Rationality testing under an asymmetric loss function

2.2.1. The Batchelor/Peel (1998) approach

One approach to test for forecast rationality under an asymmetric loss function has been proposed by Batchelor and Peel (1998). They start from a so-called linex loss function, which takes the form:

$$L = \frac{\beta}{\delta^2} [\exp(\delta e_t - \delta e_t - 1)],$$

where δ and β are constants and e is the forecast error as described above. The parameter δ determines the degree of asymmetry, while β is a scaling factor. For $\delta > 0$, losses are approximately exponential for $e > 0$ and approximately linear for $e < 0$. If the forecast error is defined as in our case, this presents a situation where underestimations are more costly than overestimations. Conversely, with $\delta < 0$ the function is exponential to the left of the origin of e , and linear to the right. Asymptotically for $\delta = 0$, the function coincides with the standard quadratic case.

The standard rationality test may be extended as follows: Batchelor and Peel (1998) argue that under the linex loss function the optimal forecast has a clearly defined bias. This bias, in turn, depends on the volatility of the time series to be forecasted and has an analytical expression for the linex loss function. Thus, to test for rationality, an additional term in the test equation is warranted which reflects the expected value of the conditional error variance:

$$(y_{t+1} - \hat{y}_{t+1,t}) = \alpha_0 + \alpha_1(\hat{y}_{t+1,t} - y_t) + \frac{\delta}{2} E_t(\sigma_{t+1}^2) + u_t \quad (4)$$

Again, the null of a rational forecast is represented by the parameter restriction $\alpha_0 = 0$, $\alpha_1 = 0$. Thus, in empirical testing, Batchelor and Peel (1998) suggest estimating an

ARCH-in-Mean model, from Engle, Lilian and Roberts (1987). In their original paper, they suggest a GARCH(1,1) model, but argue that the test for rationality does not depend on a specific form of the ARCH-in-Mean term. Hence, in our case, we start with the presumably most simple GARCH(1,1) and use other models only in cases where this model does not fit well with the data. It turns out that, in most cases, using the log for the ARCH-in-Mean term helps to achieve convergence. All in all, the test is performed by estimating the following equations:

$$\begin{aligned}(y_{t+1} - \hat{y}_{t+1,t}) &= \alpha_0 + \alpha_1(\hat{y}_{t+1,t} - y_t) + \alpha_2 E_t(\sigma_{t+1}^2) + u_{t+1} \\ u_t &\sim N(0, \sigma_{t+1}^2) \\ \sigma_{t+1}^2 &= c_1 + c_2 u_{t+1}^2 + c_3 \sigma_t^2\end{aligned}\tag{5}$$

As in the original contribution of Batchelor and Peel (1998), the ARCH-in-Mean terms turn out to be insignificant in most of cases here. The presence of this term might, however, alter the estimates of the other coefficients in the equation and, thus, the results of testing for rationality, namely, $a_0 = 0$, $a_1 = 0$. The results presented in Table 4 suggest that in all but one cases (the Council of Economic Advisers), the null of rationality cannot be rejected for the growth forecasts considered in this paper. The results change when considering the inflation forecast errors; in this case, at least four institutions produced forecasts that could be interpreted as non-rational – even under asymmetric loss. Compared to the results for a symmetric loss reported in Table 3 – where ten institutions could be labeled as those departing from rationality assumption – we observe a remarkable drop in the number of non-rational forecasts. This finding emphasizes the importance of allowing for asymmetric loss function of forecasters when testing for forecast rationality.

Table 4. Test for rationality of the forecasts under an asymmetric loss function (Batchelor/Peel approach), 1970 to 2007

		Constant	Slope	Test for rationality (F-value)	Test for rationality (p-value)
Growth forecasts					
Berlin Institute	Coefficient	0.08	0.40	1.47	0.24
	t-value	0.14	2.02		
Council of Economic Advisers	Coefficient	2.11	-0.33	5.22	0.01
	t-value	0.52	-1.65		
Employers Institute ¹	Coefficient	-0.47	-0.10	0.04	0.96
	t-value	-1.23	-0.41		
Essen Institute	Coefficient	-0.61	-0.14	0.55	0.58
	t-value	-2.70	-0.47		
European Commission, autumn	Coefficient	-0.03	-0.14	0.92	0.41
	t-value	-0.12	-0.62		
European Commission, spring ²	Coefficient	-0.37	0.42	0.59	0.56
	t-value	-1.38	2.12		
Government's Economic Report	Coefficient	-0.23	0.06	0.02	0.98
	t-value	-0.56	0.27		
Halle Institute ³	Coefficient	0.02	-0.39	1.45	0.29
	t-value	0.16	-1.11		
Hamburg Institute	Coefficient	-0.70	-0.21	0.36	0.70
	t-value	-1.15	-0.71		
IMF, autumn ⁴	Coefficient	-1.15	-0.04	1.38	0.27
	t-value	-1.69	-0.08		
IMF, spring ⁵	Coefficient	0.06	0.16	1.38	0.27
	t-value	0.27	0.55		

Table 4. **Test for rationality of the forecasts under an asymmetric loss function (Batchelor/Peel approach), 1970 to 2007 (cont.)**

		Constant	Slope	Test for rationality (F-value)	Test for rationality (p-value)
Joint Forecast, autumn	Coefficient	1.44	-0.46	0.37	0.69
	t-value	0.50	-0.48		
Joint Forecast, spring	Coefficient	1.44	-0.46	0.37	0.69
	t-value	0.50	-0.48		
Kiel Institute	Coefficient	-2.19	0.20	1.22	0.31
	t-value	-0.36	0.36		
Munich Institute	Coefficient	-0.42	0.03	0.24	0.78
	t-value	-0.44	0.25		
OECD	Coefficient	-0.64	-0.30	0.03	0.97
	t-value	-1.46	-1.53		
Trade Union Institute	Coefficient	0.15	0.04	1.68	0.20
	t-value	0.91	0.16		
Inflation forecasts					
Berlin Institute	Coefficient	0.08	0.40	8.66	0.00
	t-value	0.14	2.02		
Council of Economic Advisors	Coefficient	2.11	-0.33	1.15	0.33
	t-value	0.52	-1.65		
Employers Institute ²	Coefficient	-0.47	-0.10	2.31	0.12
	t-value	-1.23	-0.41		
Essen Institute	Coefficient	-0.61	-0.14	0.53	0.60
	t-value	-2.70	-0.47		
European Commission, autumn	Coefficient	-0.03	-0.14	1.65	0.21
	t-value	-0.12	-0.62		
European Commission, spring	Coefficient	-0.10	0.08	0.79	0.46
	t-value	-0.50	0.43		
Government's Economic Report	Coefficient	-0.23	0.06	0.21	0.81
	t-value	-0.56	0.27		
Halle Institute ³	Coefficient	Convergence failed			
Hamburg Institute	Coefficient	-0.70	-0.21	0.27	0.76
	t-value	-1.15	-0.71		
IMF, autumn ¹	Coefficient	-1.15	-0.04	2.28	0.12
	t-value	-1.69	-0.08		
IMF, spring ⁴	Coefficient	0.06	0.16	2.28	0.12
	t-value	0.27	0.55		
Joint Forecast, autumn	Coefficient	1.44	-0.46	11.30	0.00
	t-value	0.50	-0.48		
Joint Forecast, spring	Coefficient	1.44	-0.46	0.56	0.57
	t-value	0.50	-0.48		
Kiel Institute	Coefficient	-2.19	0.20	11.66	0.00
	t-value	-0.36	0.36		
Munich Institute	Coefficient	-0.42	0.03	0.58	0.57
	t-value	-0.44	0.25		
OECD	Coefficient	-0.96	-0.19	0.69	0.51
	t-value	-1.03	-0.46		
Trade Union Institute	Coefficient	0.15	0.04	2.67	0.08
	t-value	0.91	0.16		

1. 1972 to 2007.

2. 1993 to 2007.

3. 1973 to 2007.

4. 1971 to 2007.

5. Convergence could be achieved only after eliminating the outlier in year 1975 by a dummy variable in the mean equation.

2.2.2. The Elliot et al. (2005) approach

Elliot et al. (2005) suggest a test of the joint null hypothesis of forecast rationality and the more flexible loss function. Under the null hypothesis, the test statistic is:

$$J = \frac{1}{T} \left(\sum_{t=r}^{T+r+1} [\nu_t \mathbf{1}(y_{t+1} - \hat{y}_{t+1} - \alpha_T)] |y_{t+1} - \hat{y}_{t+1}|^{p-1} \right)' \hat{S}^{-1} \left(\sum_{t=r}^{T+r+1} [\nu_t \mathbf{1}(y_{t+1} - \hat{y}_{t+1} - \alpha_T)] |y_{t+1} - \hat{y}_{t+1}|^{p-1} \right) \sim X_{d-1}^2 \quad (6)$$

Hence, a rejection of the null hypothesis might be due to irrationality of the forecast or to the rejection of the functional form of the loss function (1). The results for our data at hand are given in Table 5.

Table 5. **Joint test for forecast rationality and loss function (J-test), lin-lin function, 1970 to 2007**

	J-test (Model 2)	p-value	J-test (Model 3)	p-value	J-test (Model 4)	p-value
Growth forecasts						
Berlin Institute	0.57	0.45	0.49	0.48	0.69	0.71
Council of Economic Advisers	0.01	0.98	2.18	0.14	5.90	0.05
Employers Institute	0.76	0.39	1.42	0.23	1.44	0.49
Essen Institute	0.00	0.98	1.27	0.26	2.39	0.30
European Commission, autumn	0.51	0.48	2.42	0.12	3.35	0.19
European Commission, spring	0.05	0.83	0.00	0.95	0.08	0.96
Government's Economic Report	0.48	0.49	2.11	0.15	2.45	0.29
Halle Institute	0.01	0.91	0.39	0.53	1.16	0.56
Hamburg Institute	0.13	0.72	1.11	0.29	1.36	0.51
IMF, autumn	0.00	0.98	0.72	0.40	4.89	0.09
IMF, spring	1.08	0.30	0.04	0.84	1.19	0.55
Joint Forecast, autumn	0.45	0.50	1.89	0.17	2.25	0.33
Joint Forecast, spring	0.02	0.89	0.04	0.85	0.04	0.98
Kiel Institute	0.04	0.84	0.90	0.34	1.19	0.55
Munich Institute	0.03	0.88	0.12	0.73	0.13	0.94
OECD	0.90	0.34	3.07	0.08	3.80	0.15
Trade Union Institute	0.43	0.51	2.26	0.13	2.50	0.29
Pooled data	0.99	0.32	14.38	0.00	19.53	0.00
Inflation forecasts						
Berlin Institute	5.02	0.03	0.61	0.44	5.52	0.06
Council of Economic Advisers	4.99	0.03	5.06	0.02	5.62	0.06
Employer's Institute	0.57	0.45	1.12	0.29	1.14	0.56
Essen Institute	10.90	0.00	8.39	0.00	11.32	0.00
European Commission, autumn	6.21	0.01	1.94	0.16	6.66	0.04
European Commission, spring	1.21	0.27	0.20	0.65	1.21	0.55
Government's Economic Report	3.69	0.06	0.65	0.42	3.88	0.14
Hamburg Institute	4.37	0.04	1.91	0.17	4.37	0.11
IMF, autumn	3.22	0.07	0.65	0.42	3.32	0.19
IMF, spring	1.75	0.19	0.24	0.62	1.77	0.41
Joint Forecast, autumn	4.88	0.03	2.31	0.13	5.43	0.07
Joint Forecast, spring	2.93	0.09	2.63	0.11	3.70	0.16
Kiel Institute	4.33	0.04	0.23	0.63	6.94	0.03
Munich Institute	2.03	0.16	1.20	0.27	2.18	0.34
OECD	2.69	0.10	0.07	0.79	3.44	0.18
Trade Union Institute	5.23	0.02	0.03	0.86	7.59	0.02
Pooled data	62.23	0.00	19.85	0.00	62.42	0.00

Table 5. **Joint test for forecast rationality and loss function (J-test), lin-lin function, 1970 to 2007 (cont.)**

	J-test (Model 2)	p-value	J-test (Model 3)	p-value	J-test (Model 4)	p-value
Growth forecasts						
Berlin Institute	0.20	0.66	0.97	0.32	2.48	0.29
Council of Economic Advisers	0.01	0.93	0.94	0.33	2.43	0.30
Employers Institute	1.62	0.20	0.60	0.44	1.96	0.38
Essen Institute	0.28	0.60	0.58	0.45	1.81	0.40
European Commission, autumn	0.22	0.64	0.98	0.32	1.40	0.50
European Commission, spring	0.93	0.33	0.04	0.85	1.12	0.57
Government's Economic Report	0.91	0.34	0.49	0.48	0.92	0.63
Halle Institute	1.49	0.22	1.39	0.24	1.58	0.45
Hamburg Institute	0.24	0.63	0.67	0.41	0.74	0.69
IMF, autumn	0.07	0.79	0.65	0.42	5.59	0.06
IMF, spring	2.59	0.11	0.00	0.97	2.89	0.24
Joint Forecast, autumn	0.03	0.87	0.81	0.37	2.17	0.34
Joint Forecast, spring	0.04	0.83	0.70	0.40	0.89	0.64
Kiel Institute	0.00	0.99	1.39	0.24	3.72	0.16
Munich Institute	0.49	0.49	0.67	0.42	0.81	0.67
OECD	0.43	0.51	1.35	0.24	2.07	0.36
Trade Union Institute	0.09	0.77	0.82	0.37	2.08	0.35
Pooled data	0.40	0.53	11.33	0.00	20.42	0.00
Inflation forecasts						
Berlin Institute	5.02	0.03	2.79	0.09	5.04	0.08
Council of Economic Advisers	5.68	0.02	6.62	0.01	6.39	0.04
Employers Institute	1.33	0.25	2.02	0.16	2.02	0.36
Essen Institute	6.82	0.00	4.11	0.04	11.32	0.00
European Commission, autumn	6.39	0.01	5.52	0.02	6.38	0.04
European Commission, spring	2.16	0.14	0.24	0.62	2.28	0.24
Government's Economic Report	3.46	0.06	0.77	0.38	4.67	0.09
Hamburg Institute	5.11	0.02	4.23	0.04	5.14	0.07
IMF, autumn	2.96	0.08	1.06	0.30	3.38	0.21
IMF, spring	2.66	0.10	2.98	0.08	3.15	0.21
Joint Forecast, autumn	6.21	0.01	5.50	0.02	6.21	0.04
Joint Forecast, spring	1.35	0.24	1.27	0.26	1.57	0.46
Kiel Institute	8.70	0.00	6.37	0.01	10.37	0.01
Munich Institute	3.21	0.07	2.19	0.13	3.22	0.20
OECD	0.92	0.33	0.59	0.44	0.92	0.63
Trade Union Institute	5.29	0.02	1.89	0.17	6.25	0.04
Pooled data	60.84	0.00	47.98	0.00	61.94	0.00

Legend: Model 2: Constant and lagged forecast errors. Model 3: Constant and lagged variable. Model 4: Constant, lagged forecast error, lagged variable.

In the case of growth forecasts and the lin-lin setting, the null hypothesis has to be rejected in only very few cases. In particular, for the IMF autumn forecast, the OECD and the Council of Economic Advisers, the composite null hypothesis is not supported by the data. In none of the three cases, however, does the result appear to be robust with respect to the choice of the instruments. Thus, these suggestions of either irrationality of the forecasts or the necessity of a different loss function are not convincing. By contrast, the results for the inflation forecasts lead to a rejection of the null hypothesis for virtually any of the institutions under investigation. Given the point estimates of the asymmetry parameter reported in Section 2, one might suspect that the rejection is due to the failure of the rationality hypothesis rather than the assumption of a particular loss function, but formally

the test does not tell anything about this. However, the results reported for similar tests based on the assumption of a quad-quad loss function yield a similar picture: again, there are very few results, if any at all, pointing to the rejection of the null for the growth forecasts, but the inflation forecasts fail to achieve rationality under this particular loss function. Hence, the rationality of growth forecasts is generally supported by the *J*-test while the rationality of the inflation forecasts is much more in doubt. It is noteworthy that the null of rationality of inflation forecasts is more frequently rejected when the lagged forecast errors are used as instruments (i.e., Models 2 and 4), implying that the orthogonality condition between actual and lagged forecast errors does not hold. This finding corresponds to the high positive autocorrelation of the inflation forecast errors frequently reported in the literature (see Döpke and Fritsche, 2006, and the papers cited therein).

Turning to the results based on pooled data, we find that growth forecast errors appear to be orthogonal to own past forecast errors (Model 2), for both pooled and individual forecasts. Furthermore, the pooled estimates provide decisive evidence against the forecast rationality when we include lagged GDP in the IV set, contradicting the results based on individual GMM estimations.

In the case of inflation, we have decisive rejection of forecast rationality for all IV sets, corroborating most evidence from individual GMM. Observe that the magnitude of pooled *J*-test statistics for inflation by far exceeds that of growth forecasts, which is also in line with the stronger evidence against forecast rationality observed from individual *J*-tests.

3. Conclusion

The paper analyses the degree of asymmetry of German business cycle forecasts, namely, growth and inflation forecasts covering the 1970-2007 period and up to 17 different forecasting institutions. We find the forecasts to be mostly symmetric, with only few exceptions. The point estimates of the degree of asymmetry are not systematic in any respect: some forecasters seem to have incentives for overly pessimistic forecasts, others for overly optimistic forecasts. There is a general tendency for inflation forecasts to be biased. The results appear to be not fully robust against the choice of the instruments warranted to estimate the loss function with a GMM approach. We also investigate the rationality of the forecasts at hand. To this end, we rely on the assumption of a symmetric loss function, but also make use of approaches based on an asymmetric or even flexible loss function. In a nutshell, we find that neither a specifically asymmetric loss function nor the assumption of a generalised loss function alters, by any significant amount, the findings for growth forecasts obtained under a symmetric loss function. However, the results of the test proposed by Elliot *et al.* (2005) obtain contrary results for inflation forecasts – which indicates that it matters under which loss functions we test. We can detect decisively an asymmetric loss function in case of GDP forecasts (tendency to produce overly optimistic forecasts) only for a quad-quad loss function. Hence, using a different metric in forecast assessment leads to different conclusions.

As regards the question of forecast (ir-)rationality, we find virtually no hint of irrationality of growth forecasts as long as we refer to single institutions. By pooling the data, however, we obtain the opposite conclusion, i.e., growth forecasts appear to be irrational as long as lagged GDP growth is included in the set of instruments. For inflation forecasts, the conclusions based on individual *J*-tests mostly agree with those obtained by pooling the data. Our results extend, therefore, previous research, which found growth

forecasts to be inefficient in the case of a more general assumption regarding the underlying loss function and of inflation forecasts (see, e.g., Dovern and Weissner, 2008; Osterloh, 2008; and Ager et al., 2009).

Given the results of this paper, some further research may be required. First, it should be verified whether data with a higher frequency may alter results. Having more data may help estimate the asymmetry parameter with greater precision and, hence, lead to more cases of a significant degree of asymmetry. Second, it may be worthwhile to estimate the asymmetry parameter for government in order to compare it with the values for the forecasters. It is plausible to assume that political entities have different loss functions from forecasters which may, in turn, explain some of the negative reputation of business cycle forecasts in public opinion.

Notes

1. Realisations are denoted by y and forecasts by $\hat{y}$.
2. For Model 1 there exists a closed-form solution to equation (2), implying that under the lin-lin function the estimate of the asymmetry parameter equals the share of negative forecast errors.
3. In case of the quad-quad loss function, the share of sum of negative forecast errors in total sum of both positive and negative forecast errors is relevant. For example, in the case of growth forecasts, we have a sum of negative forecast errors of -379.85 and a sum of positive forecasts errors of 235.80 , which gives $379.85/(379.85 + 235.80) = 0.617$ in the case of Model 1. The respective numbers for the inflation forecasts are -179.50 (sum of negative forecast errors) and 181.45 (sum of positive forecast errors), which leads to $179.5/(179.5 + 181.45) = 0.497$ for Model 1.

References

- Artis, M. and M. Marcellino (2001), "Fiscal Forecasting: The Track Record of the IMF", OECD and EC, *Econometrics Journal*, No. 4, pp. 20-36.
- Ager, P., M. Kappler and S. Osterloh (2009), "The Accuracy and Efficiency of the Consensus Forecasts: A Further Application and Extension of the Pooled Approach", *International Journal of Forecasting*, No. 25, pp. 167-181.
- Batchelor, R. and D. Peel (1998), "Rationality Testing under Asymmetric Loss", *Economics Letters*, No. 61, pp. 49-54.
- Blinder, A.S. (1998), "Central Banking in Theory and Practice", Cambridge, MIT Press.
- Christodoulakis, G. and M. Mamatzakis (2008), "An Assessment of the EU Growth Forecasts under Asymmetric Preferences", *Journal of Forecasting*, No. 27, pp. 483-492.
- Christodoulakis, G. and M. Mamatzakis (2009), "Assessing the Prudence of Economic Forecasts in the EU", *Journal of Applied Econometrics*, No. 24, pp. 583-606.
- Christoffersen, P.F. and F.X. Diebold (1997), "Optimal Prediction under Asymmetric Loss", *Econometric Theory*, No. 13, pp. 808-817.
- Dovern, J. and J. Weisser (2008), "Are They Really Rational? Assessing Professional Macro-Economic Forecasts from the G7-Countries", *Kiel Working Papers*, No. 1447, Kiel Institute for the World Economy, Kiel.
- Döpke, J. (2000), "Haben Konjunkturprognosen einen politischen Bias?", *Schmollers Jahrbuch – Zeitschrift für Wirtschafts- und Sozialwissenschaften*, No. 120, pp. 587-620.
- Döpke, J. and U. Fritsche (2006), "Growth and Inflation Forecasts in Germany: A Panel-based Assessment of Accuracy and Dispersion", *Empirical Economics*, No. 31, pp. 777-798.
- Elliot, G., I. Komunjer and A. Timmermann (2005), "Estimation and Testing of Forecast Rationality under Flexible Loss", *Review of Economic Studies*, No. 72, pp. 1107-1125.
- Elliot, G., I. Komunjer and A. Timmermann (2008), "Biases in Macroeconomic Forecasts: Irrationality or Asymmetric Loss?", *Journal of European Economic Association*, No. 6, pp. 122-15.

- Engle, R.F., D.M. Lilien and R.P. Robins (1987), "Estimating Time-varying Premia in the Term Structure: The ARCH-M Model", *Econometrica*, No. 55, pp. 391-407.
- Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation", *Econometrica*, No. 50, pp. 987-1008.
- Fildes, R. and H.O. Stekler (2002), "The State of Macroeconomic Forecasting", *Journal of Macroeconomics*, No. 24, pp. 435-468.
- Granger, C.W.J. (1999), "Outline of Forecast Theory Using Generalized Cost Functions", *Spanish Economic Review*, No. 1, pp. 161-173.
- Hansen, B.E. and K.D. West (2002), "Generalized Methods of Moments and Macroeconomics", *Journal of Business and Economics Statistics*, No. 20, pp. 460-490.
- Laster, D., P. Bennet and I.S. Geoum (1999), "Rational Bias in Macroeconomic Forecasts", *Quarterly Journal of Economics*, No. 114, pp. 293-318.
- Miersch, M. (2008). "Wir prognostizieren uns zu Tode", <http://debatte.welt.de/kommentare/103709/wir+prognostizieren+uns+zu+tode>.
- Osterloh, S. (2008), "Accuracy and Properties of German Business Cycle Forecasts", *Applied Economics Quarterly*, No. 54, pp. 27-57.
- Patton, A.J. and A. Timmermann (2007), "Testing Forecast Optimality under Unknown Loss", *Journal of American Statistical Association*, No. 102, pp. 1172-1184.

Do Benchmark Revisions Affect the Consumption-to-Output and Investment-to-Output Ratios in Germany?

by

Thomas A. Knetsch and Hans-Eggert Reimers*

The balanced growth and stochastic growth theory implies stable consumption-to-output and investment-to-output ratios. It is tested by cointegration techniques for three different German data vintages. Systems cointegration tests are helpful in revealing inconsistencies across vintages. Differencing and rebasing, often used to adjust for benchmark revisions, are generally not sufficient to ensure consistent real-time macroeconomic data. Vintage transformation functions estimated by cointegrating regressions are more flexible. Empirically, the cointegrating property between consumption and output, as well as between investment and output, is often found, whereas the one-to-one relationship is mostly rejected. Moreover, the linear transformation function is helpful in describing the relation between two older final vintages. This function seems to be insufficient if the most recent data collection framework is involved.

JEL Classification: C22, C32, C82, O40

Keywords: investment-output ratio, consumption-output ratio, real-time data, benchmark revisions

* Knetsch: Deutsche Bundesbank, Economics Department, Wilhelm-Epstein-Str. 14, 60431 Frankfurt am Main, Germany (e-mail: thomas.knetsch@bundesbank.de); Reimers: Hochschule Wismar, Business Department, Postfach 1210, 23952 Wismar, Germany (e-mail: hans-eggert.reimers@hs-wismar.de). The authors wish to thank two anonymous referees for useful comments and suggestions. The authors are fully responsible for all remaining shortcomings. The paper expresses the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank.

Whether growth is viewed as exogenous or endogenous, standard growth models imply that economies are converging towards a steady-state equilibrium path in which capital and output grow at the same rate. Given that shocks to productivity are the main source driving the economy in the long run, the consumption-to-output and the investment-to-output ratios, usually called “big ratios”, are constant along the balanced growth path. Hence, economic fluctuations cause big ratios to oscillate without altering their long-run equilibrium values, which are determined by (time-invariant) structural parameters. By modelling permanent productivity shocks as random walks with drift, King *et al.* (1991) show that the time series of consumption, investment and output (in logs) can be represented by stochastic processes which are integrated and cointegrated. In this setup, the big ratios are long-run economic relationships estimated by cointegrating vectors.

In general, cointegration methodology allows the testing of long-run implications of the model. The analysis should be based on long time series in order to obtain meaningful empirical results from an economic point of view. The data requirement, however, may be at odds with the revision practice of statistical offices (for a survey of real-time data analysis, see Croushore, 2010). McKenzie (2006) analyses revisions of gross domestic product (GDP), index of industrial production and retail trade volume for OECD countries and selected non-member economies. He shows that revisions to growth rates are substantial in all countries and become larger the longer the period from the first published estimate. One reason may be that every five years or so, national accounts are subject to substantial revisions, altering measurement principles and survey methods (henceforth called benchmark revisions). Even a pure base year shift may alter the growth rates of real variables, if measured under the fixed-weight methodology, because the price structure of the base year defines the weighting scheme of the aggregates (see Croushore and Stark, 2001).

In order to mitigate the effects of benchmark revisions, most authors use growth rates. To avoid statistical breaks due to benchmark revisions, economic indicators are usually back-calculated on the basis of new accounting standards. By using growth rates, however, valuable information will be lost. On the one hand, Croushore and Stark (2003) use spectral analysis to investigate benchmark revisions. In general, they conclude that they cannot characterize these revisions as white noise or as a particular autoregressive integrated moving average (ARIMA) process. They tend to be idiosyncratic and are hard to predict. On the other hand, Jacobs and Sturm (2008) decompose the whole revision into its components. Benchmark revisions are an important part of revisions. Moreover, Siklos (2008) examines the relation between data releases for US macroeconomic aggregates, which is tested by cointegrating relationships. He shows that cointegration breaks down in the last two years before benchmark revisions. Nevertheless, it is not widely known that benchmark revisions can also change the parameter estimates of long-run economic relationships (see Knetsch and Reimers, 2006, 2009). This unpleasant result implies that the cointegration analysis performed to determine whether the big-ratio hypothesis of balanced growth models holds may principally be affected by pure statistical measurement issues.

In 1999, the pan-European accounting standards set out in the European System of Accounts (ESA) 1995 were generally implemented in German national accounts (see Bleses *et al.*, 1999). This benchmark revision adopted, *inter alia*, the NACE rev.* classification of economic sectors, substantively altered the definition of investment categories and referred real economic indicators to the new base year, 1995. The next benchmark revision took place in 2005. Further ESA 1995 requirements were fulfilled in measuring real economic variables in the previous year's prices (instead of constant prices of a given base year), assigning financial intermediation services to end users and expanding the application of hedonic approaches to quality adjustment in price measurement (see Braakmann *et al.*, 2005). In the course of both benchmark revisions, the Statistische Bundesamt not only recalculated the time series back to the first quarter of 1991, but also provided statistically (almost) corresponding data for western Germany between the first quarter of 1970 and the fourth quarter of 1991. The adjustment, however, was not complete in every detail. Owing to lack of primary information, it was not possible to construct hedonic price indices before 1991, for instance.

The paper examines whether these benchmark revisions affect the empirical measurement of big ratios in German national accounts. While cointegration analysis gives support to the existence of a stable consumption-to-output ratio under either accounting regime considered, this is not the case for the investment-to-output ratio. For one thing, business investment and GDP turn out to be cointegrated, but the estimated cointegrating vector is far from meeting the big-ratio restriction. For another, estimates differ greatly when performed with data of the various measurement regimes. In the terminology introduced by Knetsch and Reimers (2009), the benchmark revisions are not innocuous with respect to the long-run economic relationship between business investment and GDP.

The reason for the conflicting evidence on consumption and investment may be caused by the fact that, among the broad expenditure aggregates, the concept and measurement of investment was most strongly revised by the benchmark revisions under consideration. It might also be of relevance that the consumption share in output is considerably larger than the corresponding investment share. This suggests that, by a pure arithmetic argument, methodological changes affecting private consumption influence GDP quite similarly, while a corresponding impact through investment is limited.

The remainder of the paper is organized as follows. In Section 1, the economic background and an overview of the literature are given. In Section 2, the empirical analysis is presented. Section 3 concludes.

1. Theoretical background and existing evidence

Within the basic neoclassical growth theory of a one-sector economy following Solow (1956), output (Y_t) is generated from a constant returns to scale Cobb-Douglas production function

$$Y_t = \lambda_t K_t^{1-\gamma} L_t^\gamma, \quad (1)$$

where K_t and L_t denote the capital stock and labour input, respectively, and λ_t is the total factor productivity. The assumption of constant returns to scale allows us to employ the production function in intensive form or in the corresponding per capita version.

* NACE is the acronym for *Nomenclature général des activités économiques dans les Communautés européennes*.

Considering a deterministic trend results in steady-state growth if some assumptions concerning preferences, capital accumulation and resource constraints hold. To form the neoclassical growth model under uncertainty, the total factor productivity is characterized as a random walk process with deterministic drift term (Brock and Mirman, 1972; Donaldson and Mehra, 1983). Under uncertainty, the neoclassical model results in balanced growth, implying a common growth rate. Since factor productivity has a random walk with drift representation, the logarithms of output, consumption and investment are integrated of order one and share a common stochastic trend (King *et al.*, 1991). The logarithmic growth ratios $c_t - y_t$ and $i_t - y_t$ follow stationary stochastic processes. The variables c_t and y_t (i_t and y_t) are cointegrated. The growth ratios imply a linear restriction that can be tested within a cointegration framework. If $c_t + \beta_1 y_t$ and $i_t + \beta_2 y_t$ are stationary, the two hypotheses of interest are $H_0 : \beta_1 = -1$ and $H_0 : \beta_2 = -1$, respectively.

Adopting the suggestion of King *et al.* (1991), researchers have analyzed the hypothesis for other countries. Overall, the evidence for stationarity is mixed (see Attfield and Temple, 2006, and their cited literature). For example, Serletis and Krichel (1995) examine ten OECD countries and often reject the null hypothesis. More positive is Harvey *et al.* (1998), who provide evidence in favor of cointegration for a few countries, whereas Clemente *et al.* (1999) often reject the hypothesis. This evidence changes when structural breaks are included in the model specification. Then, the number of rejections increases. Accounting for the German unification as a structural break, Trenkler, Saikkonen and Lütkepohl (2006) find evidence of a cointegrating relationship when analysing German data for the period 1960Q1 to 2005Q4. Herwartz and Reimers (2002) study the G7 countries for the period 1970Q1 to 1997Q4 and find evidence for cointegration but reject the one-to-one relationship. Attfield (2003) presents some evidence for a balanced growth hypothesis in some 15 EU countries. In a study on the three largest euro area countries, Knetsch (2006) finds evidence of long-run economic relationships between private consumption and GDP, as well as business investment and GDP for Germany and France, even though the big-ratio hypothesis does not appear to be met in any case.

Attfield and Temple (2006) argue that the great ratios depend on structural parameters. In their model, these parameters are the rate of technical progress, the depreciation rate, the intertemporal elasticity of substitution and the discount factor. Changes in these parameters can have substantial effects on the big ratios. For example, the investment ratio is increasing in the discount factor and the depreciation rate. The broader use of hedonic price measures may shift real investment upward and raise the depreciation rate. Therefore, they suggest using nominal variables to conduct the analysis.

Systems cointegration tests are helpful in revealing inconsistencies across vintages. Differencing and rebasing, often used to adjust for benchmark revisions, are generally not sufficient to ensure consistent real-time macroeconomic data. Vintage transformation functions estimated by cointegrating regressions are more flexible. Inappropriate conversion may cause observed revision statistics to be affected by nuisance parameters (see Knetsch and Reimers, 2009).

2. Empirical analysis

2.1. The data

The German national account data are collected by the Federal Statistical Office, Wiesbaden. They are updated on a quarterly basis. This consists of adding a new figure at

the end of the series six to eight weeks after the end of the reporting quarter and revising some figures that have already been published. The revisions may be necessary for the last quarters of the past two to three years. The vintages of the series may be stored next to each other in chronological order, creating matrices that contain the real-time data sets of the economic variables (see Table 1). In this structure, the reporting quarters are in rows, while the vintages are in columns. In contrast to the term “vintage”, which measures the time series at a specific publication date, we use the term “release” to denote the issue of the quantity ascribed to a specific reporting period. In the matrix scheme, the time series collecting specific releases are found in diagonals (see Golinelli and Parigi, 2005, pp. 3-4, for an overview of the terminology).

Table 1. **Structure of real-time data of GDP**

Reporting quarter	Publishing quarter																
	03:1	03:2	03:3	03:4	04:1	04:2	04:3	04:4	05:1	05:2	05:3	05:4	06:1	06:2	06:3	06:4	07:1
01:1	8	9	.	.	.	.	.	.	F	B	F	F	F	F	F	F	F
01:2	7	8	9	.	.	.	.	.	F	B	F	F	F	F	F	F	F
01:3	6	7	8	9	.	.	.	.	F	B	F	F	F	F	F	F	F
01:4	5	6	7	8	9	.	.	.	F	B	F	F	F	F	F	F	F
02:1	4	5	6	7	8	9	.	.	F	B	.	.	F	F	F	F	F
02:2	3	4	5	6	7	8	9	.	F	B	.	.	F	F	F	F	F
02:3	2	3	4	5	6	7	8	9	F	B	.	.	F	F	F	F	F
02:4	1	2	3	4	5	6	7	8	9	B	.	.	F	F	F	F	F
03:1		1	2	3	4	5	6	7	8	B	.	.	.	.	.	.	F
03:2			1	2	3	4	5	6	7	B	9	.	.	.	.	.	F
03:3				1	2	3	4	5	6	B	8	9	.	.	.	.	F
03:4					1	2	3	4	5	B	7	8	9	.	.	.	F
04:1						1	2	3	4	5	6	7	8	9	.	.	.
04:2							1	2	3	4	5	6	7	8	9	.	.
04:3								1	2	3	4	5	6	7	8	9	.
04:4									1	2	3	4	5	6	7	8	9
05:1										1	2	3	4	5	6	7	8
05:2											1	2	3	4	5	6	7
05:3												1	2	3	4	5	6
05:4													1	2	3	4	5
06:1														1	2	3	4
06:2															1	2	3
06:3																1	2
06:4																	1

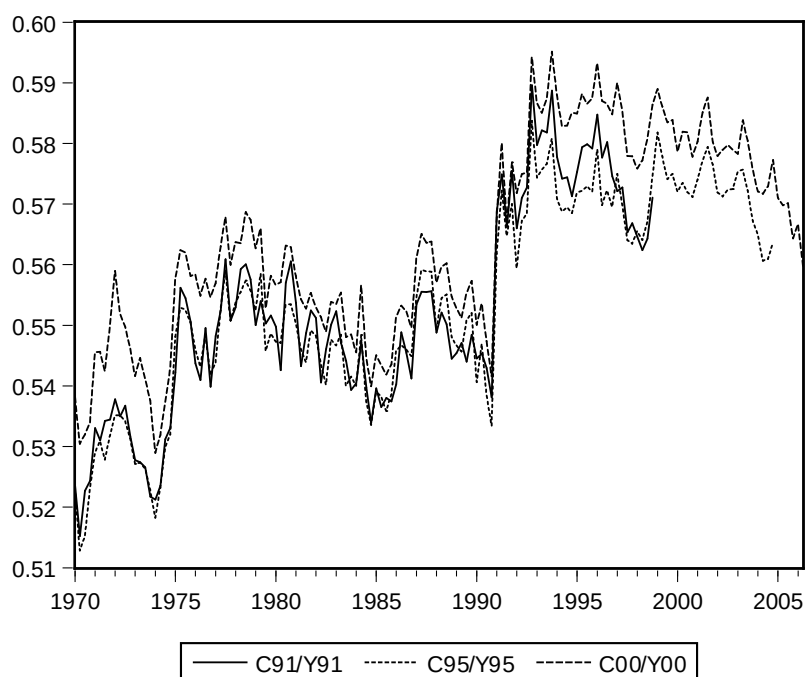
In Table 1, “F” denotes final values and “B” indicates benchmark revisions. The benchmark revision partitions the matrix. From a “B” onwards up to the next “B”, the real-time data considered in the analysis are homogenous with regard to the benchmark revisions. This characteristic allows us to analyze a transformation method which may adjust for this statistical break in the vintage dimension. It is worth noting that Hecq and Jacobs (2009) propose investigating the diagonals of this matrix. In our paper we stick to relations of special verticals.

To analyze the great ratio hypothesis for Germany, three vintages are investigated. These final data series from the database of the Deutsche Bundesbank – real gross domestic product (Y), real consumption (C) and real business investment (I) – are seasonally adjusted. Moreover, the series are recalculated so that the implied deflators are 100 in the year 1991.

To eliminate the structural break at the beginning of the 1970s due to first oil price shock, the data start in 1975Q1. The first data set ends in 1998Q4, which is indexed by 91; the second (third) data set covers 2004Q4 (2006Q2) and is indexed by 95 (00), respectively.

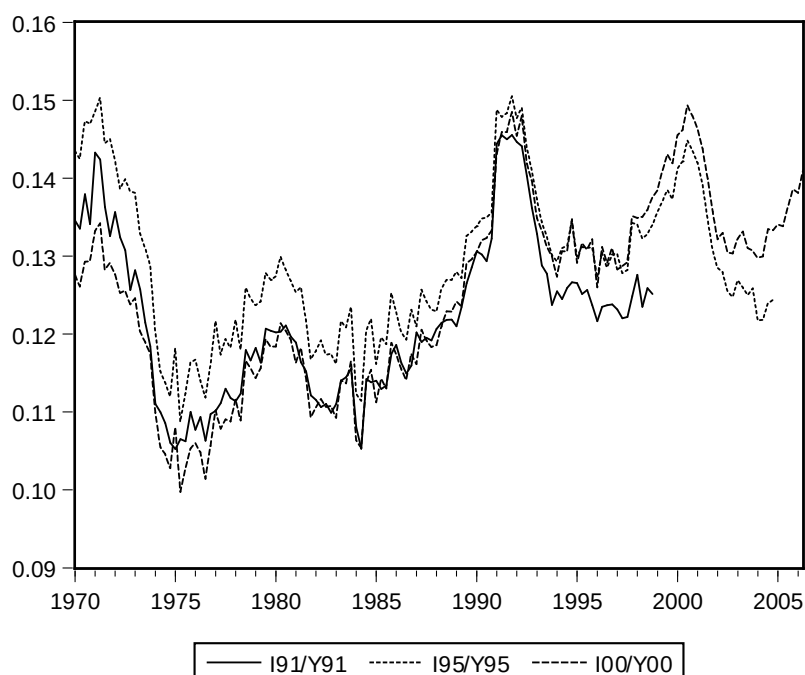
Figure 1 (2) shows the consumption-output (CY) ratio (investment-output (IY) ratio), where abbreviations are completed by the indicator of the base year. The three graphs of the CY ratios include a jump in 1991Q1 due to German unification. This jump seems to be a level shift of the series. The difference between the CY_{91} and CY_{95} series seems to be small, whereas the CY_{00} is always higher than the two other series. This may be explained by the change in accounting for financial intermediation services indirectly measured (FISIM). In the CY_{00} series this change raises consumption and, as an arithmetical consequence, also GDP, but by a smaller factor (see Eichmann, 2005). In the past 15 years, the consumption ratio fluctuates around 58%. In the period from 2004 to 2006, it falls to below 56%. The graphs of the IO ratios also exhibit a clear jump in 1991Q1 due to German unification. However, there is a strong decline of the ratios in 1993. The IY_{95} is always higher than the IY_{91} , whereas the IY_{00} is higher than the IY_{95} after 1996 and lower before 1990.

Figure 1. **The consumption-output ratios of different final data sets**



2.2. Cointegrating property of final data sets

To test the big ratio hypothesis for the different final data sets, the Saikkonen and Lütkepohl (2000) procedure is used. This approach provides a framework in which it is possible to test directly the hypothesis of interest that a cointegrating relationship exists with a mean shift in the cointegrating relation. The break point is known as 1991Q1. It is modelled by a shift dummy, which is unity up to the first quarter of 1991 and zero elsewhere. Because the residuals of the systems show an outlier in the first quarter of 1991, the tests are

Figure 2. **The investment-output ratios of different final data sets**

additionally presented for systems including a shift dummy and an unrestricted impulse dummy, where the first quarter of 1991 is unity and zero elsewhere. To be consistent with a final data set, the samples used for tests and estimates are shortened in all cases. The first, second and third data sets end at 1995Q4, 2001Q4 and 2004Q4, respectively. Before testing the cointegrating property of the system, its lag order has to be determined. We apply information criteria such as the Akaike (AIC), the Hannan-Quinn (HQ) and the Schwarz criterion (SC) (see Lütkepohl, 2005, Chap. 4). Table 2 and 3 report the estimates of the lag orders which minimize the criteria for unrestricted VARs in levels taking the lag orders $p = 1, \dots, 10$. The unrestricted VARs include a constant and a shift dummy. The lag order criteria suggest a lag order of one. In all cases, the null hypothesis of no cointegration is rejected at the 1% level. These results are also obtained if the systems additionally contain an unrestricted impulse dummy (see Table 2 for consumption-output ratios, CY).

To estimate the parameters of the cointegrating relationships, the shift dummy and an unrestricted impulse dummy, a simple two-step procedure is implemented using JMulti software (see Lütkepohl, 2004, pp. 103-05). Table 2 presents the long-run estimates of the system in equation form and some diagnostic statistics of the consumption-output relation. Standard errors are given in brackets below the parameter estimates. The long-run coefficient of the first system (CY_{91}) is not significantly below unity at the 5% level. Moreover, the shift dummy is significantly different from zero. German unification raises consumption. The null hypothesis of $H_0 : \beta_i = -1$ is checked in the system via a Wald test. The test statistic is asymptotically χ^2 -distributed, where the number of restrictions determines the degree of freedom. The null hypothesis is not rejected at the 5% level. The same evidence is found for the second system (CY_{95}). However, for the third system (CY_{00}) the long-run coefficient is found to be less than one. This result is at odds with the big-ratio hypothesis.

Table 2. **Test results of vector autoregressive models for consumption and output**

System	Information base	Lag order	SL-test p -value	Cointegrating relation	LB(16) p -value	LM(5) p -value	$H_0 : \beta_i = -1$
CY_{91}	1975Q1-	AIC = 1	32.84 (0.000)	$C_{91t} = 0.953Y_{91t} + 0.070S_t$ (0.026) (0.010)	0.685	0.328	3.464 (0.063)
	1995Q4	HQ = 1					
		SC = 1					
CY_{95}	1975Q1-	AIC = 1	40.87 (0.000)	$C_{95t} = 0.970Y_{95t} + 0.061S_t$ (0.026) (0.012)	1.000	0.792	1.425 (0.233)
	2001Q4	HQ = 1					
		SC = 1					
CY_{00}	1975Q1-	AIC = 1	44.71 (0.000)	$C_{00t} = 0.936Y_{00t} + 0.078S_t$ (0.024) (0.011)	1.000	0.669	7.012 (0.008)
	2004Q4	HQ = 1					
		SC = 1					

Notes: SL-test: Saikkonen-Lütkepohl test for cointegration including shift dummy and impulse dummy (s.dum + i.dum). Below the estimates are standard errors in parentheses. LB(16): Adjusted Portmanteau test for autocorrelation up to 16 lags. LM(5): Lagrange-Multiplier-test up to 5 lags. Null hypothesis $H_0 : \beta_i = -1$ using a Wald-test. p -values in parentheses below the test statistic.

Table 3. **Test results of vector autoregressive models for investment and output**

System	Information base	Lag order	SL-test p -value	Cointegrating relation	LB(16) p -value	LM(5) p -value	$H_0 : \beta_i = -1$
IY_{91}	1975Q1-	AIC = 6	17.67 (0.005)	$I_{91t} = 1.642Y_{91t} - 0.182S_t$ (0.185) (0.078)	1.000	0.982	12.63 (0.000)
	1995Q4	HQ = 1	16.57 (0.008)	$I_{91t} = 1.700Y_{91t} - 0.214S_t$ (0.330) (0.132)	1.000	0.577	4.360 (0.035)
		SC = 1					
IY_{95}	1975Q1-	AIC = 6	22.60 (0.001)	$I_{95t} = 1.342Y_{95t} - 0.134S_t$ (0.179) (0.077)	1.000	0.998	4.241 (0.039)
	2001Q4	HQ = 1	26.27 (0.000)	$I_{95t} = 1.503Y_{95t} - 0.226S_t$ (0.316) (0.142)	0.999	0.812	3.036 (0.081)
		SC = 1					
IY_{00}	1975Q1-	AIC = 6	18.85 (0.003)	$I_{00t} = 1.699Y_{00t} - 0.170S_t$ (0.170) (0.075)	0.999	0.932	17.06 (0.000)
	2004Q4	HQ = 1	25.40 (0.000)	$I_{00t} = 1.826Y_{00t} - 0.240S_t$ (0.308) (0.143)	0.959	0.186	6.572 (0.010)
		SC = 1					

Note: See Table 2.

Turning to the investment-output relation, we consider two cases for each system (see Table 3). While the SC and HQ suggest a lag order of one, the AIC opts for a lag order of six. In all cases, the null hypothesis of no cointegration is rejected at the 1% level. For the first system (IY_{91}), the long-run coefficient is greater than unity. The null hypothesis of a unity coefficient is rejected at the 5% test level for both cases. The shift dummy shows a reduction in the investment share and this coefficient is significant at the 5% level for a VAR lag order of one. For the system IY_{95} with the parsimonious lag cycles choice, the unity hypothesis for β_i is not rejected at the 5% level. Using the higher lag order, the null hypothesis is rejected at the 5% level. The two-standard-error confidence interval of the coefficient of the shift dummy includes zero in this case. For the last system (IY_{00}), the null hypothesis of $\beta_i = -1$ in the system is clearly rejected. Analyzing the higher lag order, the

null hypothesis is rejected at the 5% level. This result is in line with the shown investment-output ratio of Figure 2, which is much lower on average in the mid-1970s than in the more recent period. This finding contradicts the big-ratio hypothesis. The result is consistent with evidence in Herwartz and Reimers (2002). Following Temple (1999) or Mankiw, Romer and Weil (1992), the rejection of growth ratio hypotheses may be addressed to the definition of investment applied in national accounts, which does not take human capital into account.

Because the last final data set incorporates to a greater degree the application of hedonic approaches to capture quality changes in price measurement (see Braakmann *et al.*, 2005), it improves the data quality. This implies, on the one hand, that the old data set does not reflect the actual economic knowledge. There are features in the evolving economy that are increasingly poorly reflected by the old methods. On the other hand, it was not possible to construct hedonic price indices before 1992. The data improvement sticks to one part of the data set. Therefore, the new data set is a mixture of old and new measurement techniques.

In general, we see that the evidence for the big ratio hypothesis seems also to depend on the final data set used. Therefore, it is necessary to check empirical results against alternative vintages.

2.3. Vintage relationship

In real-time data sets, benchmark revisions may induce a statistical break between vintages before and after the date of conversion (see Siklos 2008). Hecq and Jacobs (2009) diagnose a co-break in the series of revisions. To eliminate this property, differencing and rebasing are well-known strategies. We follow Knetsch and Reimers (2009) to estimate vintage transformation functions between different vintages. The transformation function is modelled as the affine relationship

$$x(t) = a_0 + a_1 \bar{x}(t) + \xi(t), \quad (2)$$

where $x(t)$ is the series after the benchmark revision and $\bar{x}(t)$ is the series before the benchmark revision; a_0 and a_1 are scalars; and $\xi(t)$ is a residual process with zero mean and variance σ_ξ^2 . It is worth stressing that (2) is a cointegrating regression because the series of interest are $I(1)$. The series $x(t)$ and $\bar{x}(t)$ are both final releases, only differing with respect to the base year. The rebasing method implies unit elasticity of the transformation function $a_1 = 1$. This implication is tested empirically. Given that the long-run relationships between consumption and output, as well as between investment and output, are subject to a structural break in 1991Q1, we include the shift dummy. The new relationship is

$$x(t) = a'_0 + a'_0 S_t + a'_1 \bar{x}(t) + \zeta'_t, \quad (3)$$

where S_t is a break dummy or shift dummy.

We start the analysis by determining the lag orders of the underlying VARs. The AIC chooses comparably long lag lengths, leading to overparametrized models, especially in the case of investment ($I_{91} - I_{95}$). In the consumption-output case, the HQ and the SC, however, opt for the lag order 1. The latter choice is confirmed by diagnostic checks, which indicate the absence of residual autocorrelation of order 5 (see Table 4). The SL test

Table 4. Results of the cointegrating relationship of the transformation relationship

System	Information base	Lag order	S & L <i>p</i> -value	Estimates	LB(16) <i>p</i> -value	LM(5) <i>p</i> -value	$H_0 : \alpha'_1 = 1$
C_{91}	1975Q1	AIC = 6	3.970 (0.715)	$C_{95t} = 1.039C_{91t} + 0.005S_t$ (0.011) (0.004)	0.964	0.501	13.65 (0.000)
C_{95}	1995Q4	HQ = SC = 1	29.30 (0.000)	$C_{95t} = 1.038C_{91t} + 0.005S_t$ (0.007) (0.003)	0.999	0.574	25.50 (0.000)
C_{95}	1975Q1	AIC = 3	21.66 (0.001)	$C_{00t} = 0.997C_{95t} - 0.002S_t$ (0.015) (0.007)	0.999	0.063	0.034 (0.854)
C_{00}	2001Q4	HQ = SC = 1	23.65 (0.000)	$C_{00t} = 1.008C_{95t} - 0.007S_t$ (0.019) (0.009)	0.966	0.029	0.169 (0.681)
Y_{91}	1975Q1	AIC = 1	41.85 (0.000)	$Y_{95t} = 1.025Y_{91t} + 0.017S_t$ (0.004) (0.002)	1.000	0.812	40.55 (0.000)
Y_{95}	1995Q4	HQ = SC = 1					
Y_{95}	1975Q1	AIC = 5	12.09 (0.054)	$Y_{00t} = 1.003Y_{95t} - 0.002S_t$ (0.038) (0.016)	1.000	0.000	0.544 (0.461)
Y_{00}	2001Q4	HQ = SC = 1	29.34 (0.000)	$Y_{00t} = 1.031Y_{95t} - 0.019S_t$ (0.053) (0.024)	0.999	0.094	0.422 (0.516)
I_{91}	1975Q1	AIC = 8	9.68 (0.134)	$I_{95t} = 0.826I_{91t} - 0.017S_t$ (0.055) (0.026)	0.817	0.652	9.745 (0.002)
I_{95}	1995Q4	HQ = SC = 4	34.51 (0.000)	$I_{95t} = 0.849I_{91t} - 0.043S_t$ (0.056) (0.032)	1.000	0.990	21.42 (0.000)
I_{95}	1975Q1	AIC = 6	12.32 (0.049)	$I_{00t} = 1.382I_{95t} - 0.061S_t$ (0.073) (0.035)	0.952	0.000	23.43 (0.000)
I_{00}	2001Q4	HQ = SC = 2	38.28 (0.000)	$I_{00t} = 1.190I_{95t} + 0.016S_t$ (0.067) (0.035)	1.000	0.604	5.716 (0.017)

Notes: Cointegration test according to Saikkonen and Lütkepohl (2000) with known break 1991Q1. Break modelled by a shift dummy and impulse dummy (s.dum + i.dum). Estimated long-run relationship, where the system includes a shift dummy up to 1991Q1 in the long-run relation and an unrestricted impulse dummy for 1991Q1. Standard errors in parentheses. Other notes: see Table 2.

statistics are reported in Table 5. The null hypothesis of no cointegration can be rejected for the lag order 1. Power erosion might be the reason the SL test is unable to reject the null for the long AIC lag lengths. It is worth noting that the inclusion of an unrestricted impulse dummy does not change the evidence in favour of one cointegrating relationship.

The transformation functions are estimated by the cointegration procedure of Saikkonen and Lütkepohl (2000). The estimates are given in Table 5. For the consumption case, the estimates α'_1 are close to unity. Nevertheless, the null hypothesis of unity elasticity is rejected for the C_{91} , C_{95} case, but not for the C_{95} , C_{00} case. For the output series, the results point in the same direction. The results of the investment series show a different picture. In the case I_{91} , I_{95} the coefficient α'_1 is less than unity. The null hypothesis of unity is rejected at the 1% level. This hypothesis is also rejected for the case I_{95} , I_{00} . However, the coefficient is greater than unity. First, it indicates that the more recent data collection framework may imply a higher real investment than the former framework due to the use of hedonic price indices. Second, these results indicate that rebasing or backwards calculation using former growth rates is inappropriate for investment series from a statistical point of view.

Knetsch and Reimers (2009) extended the vintage analysis by the concept of innocuousness, which they define as follows: A benchmark revision is said to be innocuous with regard to long-run economic relationships estimated by cointegration regressions if

the affine vintage transformation functions of the economic variables involved are equivalent (Knetsch and Reimers, 2009, p. 218). The innocuousness hypothesis requires that the vintage transformation functions share the same coefficients, where deterministic elements restricted to the cointegration space are ignored at this moment. Therefore, two bivariate systems are put together to obtain a system of four variables. Let the long-run economic relationship be described by the cointegrating relation $x(t) + cz(t) \sim I(0)$ and the vintage transformation functions be modelled by (3) and

$$z(t) = b_0 + b_0^s S_t + b_1' \bar{z}(t) + \eta(t), \quad (4)$$

where the two versions $z(t)$ and $\bar{z}(t)$ of the second variable, the scalars b_0 , b_0^s and b_1' and the residuals $\eta(t) \sim (0, \sigma_\eta^2)$ are defined by analogy to (3). The cointegration space of the enlarged system comprises three cointegrating relations, represented by

$$\begin{bmatrix} 1 & 0 & c & 0 & d \\ 1 & a_1' & 0 & 0 & a_0^s \\ 0 & 0 & 1 & b_1' & b_0^s \end{bmatrix} \begin{bmatrix} x(t) \\ \bar{x}(t) \\ z(t) \\ \bar{z}(t) \\ S(t) \end{bmatrix}$$

in matrix notation. This structure is exactly identified, while the innocuousness hypothesis $a_1' = b_1'$ and $a_0^s = b_0^s$, which is denoted as strong innocuousness, imposes two over-identifying restrictions in this case. The weak innocuousness hypothesis is given as $a_1' = b_1'$. Linear restrictions on the cointegration space can be tested by an LR test, which is asymptotically χ^2 -distributed. In the example at hand, two (one) overidentifying restrictions are involved. Therefore the limiting distribution is $\chi^2(2)(\chi^2(1))$. As the restrictions span cointegrating vectors, Boswijk's switching algorithm has to be applied (see Boswijk and Doornik, 2004, Section 4.4).

Table 5. Results for the innocuousness hypothesis

System	Information base	Order	S&L test	Rang spec.	Test of innocuousness hypotheses	
					Strong	Weak
$C_{91}, Y_{91}, C_{95}, Y_{95}$	1975Q1-1995Q4	AIC = HQ = SC = 1	3	3	9.368 (0.009)	3.258 (0.071)
$C_{95}, Y_{95}, C_{00}, Y_{00}$	1975Q1-2001Q4	HQ = SC = 1 AIC = 5	1 1			
$I_{91}, Y_{91}, I_{95}, Y_{95}$	1975Q1-1995Q4	AIC = HQ = SC = 1	3	3	6.855 (0.032)	8.603 (0.003)
$I_{95}, Y_{95}, I_{00}, Y_{00}$	1975Q1-2001Q4	SC = HQ = 1 AIC = 7	2 2	3 3	31.065 (0.000) 9.845 (0.007)	10.828 (0.000) 13.292 (0.000)

Notes: S&L: Cointegration test of Saikkonen and Lütkepohl (2000), where the 5% test level is applied; Rang spec.: Specification of cointegration rang to test the innocuousness hypotheses.

The tests of the innocuousness hypothesis are shown in Table 5. The table gives the lag order of the VARs in the third column and the results of the cointegration test of Saikkonen and Lütkepohl (2000) in the fourth column. For the first system, the

cointegration rank of three is confirmed. The strong version of innocuousness is rejected, whereas the weak version is not rejected at the 5% level. For the second system, the cointegration rank is not confirmed. The S&L-test indicates a rank of 1% at the 5% level. This gives a hint that the results of the bivariate processes are not found in the comprehensive model. Hypothesis tests for cointegrating space of rank 3 are not reasonable. Turning to the system involving investment, the cointegration rank of three is confirmed. The innocuousness hypothesis is rejected in both versions at the 5% level. In the fourth system, the S&L test indicates a cointegration rank of two. Conducting the test for innocuousness, which relies on a cointegrating rank of three, the null hypothesis is rejected at the 1% level. This indicates that the linear transformation functions are not sufficient to describe the changes between the vintages.

To illustrate the implications of the distinct vintage transformation functions, they can be used to connect the consumption-output and investment-output relations of the different final data sets. The results of such an exercise are given in Table 6. The transformation functions are substituted in subsequent consumption-output (investment-output) relationships in order to obtain the former cointegrating relationship. The results show that the shift from the base year 1991 to base year 1995 for the consumption-output relation can be explained to a very large extent by the transformation functions. In contrast to the four-dimensional system, the linear transformation function helps to explain the difference in the investment-output relation. Turning to the second case, introducing the linear transformation functions results in long-run coefficients, which diverge substantially from the former estimated values. Linear transformation functions do not seem to be sufficient to explain the coefficient shifts. The evidence fits the results of the four-dimensional analysis.

Table 6. Importance of the linear transformation for the linear relationships

	Base year 1991 against base year 1995	Base year 1995 against base year 2000
Consumption-output relation		
Target – estimated:	$C_{91} = 0.953 \cdot Y_{91}$	$C_{95} = 0.970 \cdot Y_{95}$
Estimated:	$C_{95} = 0.970 \cdot Y_{95}$	$C_{00} = 0.936 \cdot Y_{00}$
Estimated:	$C_{95} = 1.038 \cdot C_{91}$	$C_{00} = 1.008 \cdot C_{95}$
Estimated:	$Y_{95} = 1.025 \cdot Y_{91}$	$Y_{00} = 1.031 \cdot Y_{95}$
	$C_{91} = \frac{0.970}{1.038} 1.025 \cdot Y_{91}$	$C_{95} = \frac{0.936}{1.031} 1.008 \cdot Y_{95}$
Implicit:	$C_{91} = 0.958 \cdot Y_{91}$	$C_{95} = 0.915 \cdot Y_{95}$
Interpretation:	Helps	Does not help
Investment-output relation		
Target – estimated:	$I_{91} = 1.700 \cdot Y_{91}$	$I_{95} = 1.503 \cdot Y_{95}$
Estimated:	$I_{95} = 1.503 \cdot Y_{95}$	$I_{00} = 1.826 \cdot Y_{00}$
Estimated:	$I_{95} = 0.849 \cdot I_{91}$	$I_{PY} = 1.190 \cdot I_{95}$
Estimated:	$Y_{95} = 1.025 \cdot Y_{91}$	$Y_{00} = 1.031 \cdot Y_{95}$
	$I_{91} = \frac{1.503}{0.849} 1.025 \cdot Y_{91}$	$I_{95} = \frac{1.826}{1.031} 1.190 \cdot Y_{95}$
Implicit:	$I_{91} = 1.815 \cdot Y_{91}$	$I_{95} = 2.108 \cdot Y_{95}$
Interpretation	Helps	Does not help

3. Conclusion

In this paper, we have tested empirically the implications of balanced growth and stochastic growth theory, using three different final data vintages for Germany. We find the cointegration property of all three consumption output systems and the investment output systems. For the first two data collection regimes, the big-ratio hypothesis of consumption and output is confirmed. The implied long-run coefficient restriction is rejected for the most recent vintage and for all investment output systems. The evidence shows that the results depend on special data sets. Therefore, researchers should use alternative vintages to check their empirical results.

Adopting an affine linear transformation function explains changes in the consumption-output and investment-output relation from the first final data vintage to the second final data vintage. However, this model is not sufficient to explain the relation between the second and third final vintage. The new data collection framework makes more intensive use of the hedonic price index, which may demand a nonlinear transformation function. Moreover, the chain price index may also explain the difference from a statistical point of view. In general, our results indicate that rebasing or backwards calculation using former growth rates is inappropriate for investment series to construct long historical time series.

The more general conclusion of our study may be that real-time data analysis should incorporate an initial step consisting of tests of benchmark revisions if a backward calculation is necessary to describe the implications of the benchmark revisions for the series. With respect to econometric modelling, understanding them may prevent misspecifications and, with regard to the analysis of regular revisions, may protect the researcher from the risk that results are adversely affected by an incongruent real-time data set.

References

- Attfield, C.L.F. (2003), *Balanced Growth and Output Convergence in Europe*, Working Paper No. 02/547, Department of Economics, University of Bristol.
- Attfield, C.L.F., J.R.W. Temple (2006), *Balanced Growth and the Great Ratios: New Evidence for the US and UK*, Discussion Paper Series No. 074, Centre for Growth & Business Cycle Research, University of Manchester.
- Bleses, P., H. Essig, N. Hartmann, W. Strohm (1999), Revision der Volkswirtschaftlichen Gesamtrechnungen 1999 – Anlass, Konzeptänderungen und neue Begriffe, *Wirtschaft und Statistik* pp. 257-462.
- Boswijk, H.P., J.A. Doornik (2004), "Identifying, Estimating and Testing Restricted Cointegrated Systems: An Overview", *Statistica Neerlandica*, 58, 440-465.
- Braakmann, A., N. Hartmann, N. Räth, W. Strohm (2005), Revision der Volkswirtschaftlichen Gesamtrechnungen 2005 für den Zeitraum 1991 bis 2004, *Wirtschaft und Statistik*, 425-462.
- Brock, W.A., L.J. Mirman (1972), "Optimal Economic Growth and Uncertainty: the Discounted Case", *Journal of Economic Theory*, 4, 479-513.
- Clemente, J., A. Montanes, M. Ponz (1999), "Are the Consumption/Output and Investment/Output Ratios Stationary? An International Analysis", *Applied Economics Letters*, 6, 687-691.
- Croushore, D. (2010), "Frontiers of Real-Time Data Analysis", *Journal of Economic Literature* (forthcoming).
- Croushore, D., T. Stark (2001), "A Real-Time Data Set for Macroeconomists", *Journal of Econometrics*, 105, 111-130.

- Croushore, D., T. Stark (2003), "A Real Time Data Set for Macroeconomists: Does the Data Vintage Matter?", *Review of Economics and Statistics*, 85, 605-617.
- Donaldson, J.B., B. Mehra (1983), "Stochastic Growth with Correlated Productivity Shocks", *Journal of Economic Theory*, 29, 282-312.
- Eichmann, W. (2005), "Finanzserviceleistung, indirekte Messung (FISM)", *Wirtschaft und Statistik*, 7, 710-716.
- Golinelli, R., G. Parigi (2005), *Short-Run Italian GDP Forecasting and Real-Time Data*, CEPR Discussion Paper 5302.
- Harvey, D.I., S.J. Leybourne, P. Newbold (1998), *How Great are the Great Ratios?* Discussion Paper, Department of Economics, University of Nottingham.
- Hecq, A., J.P.A.M Jacobs (2009), *On the VAR-VECM Presentation of Real Time Data* Paper presented at the EC²-Meeting in Aarhus.
- Herwartz, H., H.-E. Reimers (2002), "Testing Growth Ratios via Pooled Error Correction Models", *Economics Bulletin*, 3, 1-11.
- Jacobs, J.P.A.M., J.E. Sturm (2008), "The Information Content of KOF Indicators on Swiss Current Account Data Revisions", *Journal of Business Cycle Measurement and Analysis*, 4, 163-183.
- King, R.G., C.I. Plosser, H.J. Stock, M.W. Watson (1991), "Stochastic Trends and Economic Fluctuations", *American Economic Review*, 81, 819-840.
- Knetsch, T.A. (2006), "Short-Run and Long-Run Comovement of GDP and Some Expenditure Aggregates in Germany, France and Italy", in: O. de Bandt, H. Herrmann, G. Parigi (eds.) *Convergence or Divergence in Europe*, Springer Berlin-Heidelberg, pp. 209-249.
- Knetsch, T.A., H.-E. Reimers (2006), *How to Treat Benchmark revisions? The Case of German Production and Orders Statistics*, Discussion paper Research Centre Deutsche Bundesbank, No. 38/2006, Frankfurt am Main.
- Knetsch, T.A., H.-E. Reimers (2009), "Dealing with Benchmark Revisions in Real-Time Data: The Case of German Production and Orders Statistics", *Oxford Bulletin of Economics & Statistics*, 71, pp. 209-235.
- Lütkepohl, H. (2004), "Vector Autoregressive and Vector Error Correction Models", in: H. Lütkepohl and M. Kräzig (eds.), *Applied Time Series Econometrics*, Chapter 3, pp. 86-158, Cambridge University Press, Cambridge.
- Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer Verlag, Berlin.
- Mankiw, N.G., D. Romer, D.N. Weil (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, 107, 407-437.
- McKenzie, R. (2006), *Undertaking Revisions and Real-Time Data Analysis Using the OECD Main Economic Indicators Original Release Data and Revisions Database*, Statistics Working Paper STD/DOC(2006)2, Organisation for Economic Co-operation and Development.
- Saikkonen, P., H. Lütkepohl (2000), "Testing for the Cointegrating Rank of a VAR Process With Structural Shifts", *Journal of Business and Economic Statistics*, 18, 451-464.
- Serletis, A., T. Krichel (1995), *International Evidence on the Long-Run Implications to the Neoclassical Growth Model*.
- Siklos, P.L. (2008), "What Can we Learn from Comprehensive Data Revisions for Forecasting Inflation? Some U.S. Evidence", in: D. Rapach and M.E. Wohar (eds.), *Forecasting in the Presence of Structural Breaks and Model Uncertainty* Elsevier, Amsterdam.
- Solow, R.M. (1956), "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70, 65-94.
- Temple, J. (1999), "The New Growth Evidence", *Journal of Economic Literature*, 37, 112-156.
- Trenkler, C., P. Saikkonen, H. Lütkepohl (2006), *Testing for the Cointegrating Rank of a VAR Process with Level Shift and Trend Break*, SFB 649 Discussion Paper 2006-067, Humboldt-University Berlin.

OECD and CIRET News 2010/1

Conferences

30th CIRET Conference

Wednesday, 13 October – Saturday, 16 October 2010, New York

Economic Tendency Surveys and the Services Sector

Special Topic

Economic Tendency Surveys and Financial Markets

Hosted by

The Conference Board, New York, NY

The overall aim of CIRET conferences is to encourage and improve communication, exchange and co-operation between academics and practitioners who conduct economic surveys, analyse survey data and develop or make use of cyclical indicators. CIRET, the Centre for International Research on Economic Tendency Surveys, is also a forum for discussion and application of new methodological developments and their results.

Economic Tendency Surveys and the Services Sector

Manufacturing and industrial activity are generally regarded as more cyclical than services. However, the role of the service sector and its relationship to different parts of the

economy, including the industrial sector, is important and has been growing in advanced and developing countries alike. The impact of services activity on the dynamics of business cycles, labor markets, and economic growth is of interest to many researchers and producers of business tendency surveys. The relative importance of and the relationship between the manufacturing and service sectors within the context of business cycles is not very well understood. Addressing these issues requires new data on services. The lack of timely data on services activity is a significant challenge, but it can be overcome not only through new business tendency surveys, but also by innovative research based on existing surveys of industrial sectors.

Topics

I. Short-term Economic Surveys and Indicators

- Business Tendency Surveys (including Investment Surveys)
- Consumer Tendency Surveys
- Composite and Leading Indicators
- Ad hoc Surveys
- New Methods
- Real-time monitoring and forecasting

III. Special Topic

● Economic Tendency Surveys and Financial Markets

The special topic invites papers on the evaluation of the usefulness of business and consumer survey (BCS) data in analyzing both the recent global recession and the subsequent recovery.

- How well have BCS data performed in predicting the financial crisis and the global recession?

II. Surveys Related to Structural Aspects of the Economy

- Innovative Activity
- Information and Communication Technology (ICT)
- Organisational Change
- Labour Market Analysis

Other related questions using BCS data could be:

- What role does money and credit play in modern economies?
- Has this role changed over time?
- What are the real time forecasting aspects of the use of BCS data?

CIRET Conference website:

www.ciret.org/conferences/newyork_2010.

Isaac Kerstenetzky Award

sponsored by FGV, for the best papers presented at the CIRET Conference in distinct categories (see details on the CIRET Conference website).

Papers

Some 220 papers were submitted for consideration, and of these 125 were accepted.

Publications

Authors are kindly invited to submit their ready-to-publish papers to the joint CIRET and OECD Journal of Business Cycle Measurement and Analysis (JBCMA) (www.ciret.org/jbcma).

Submissions accepted by the Editor-in-Chief will be peer reviewed by two referees.

Contact

CIRET Office
c/o ETH Zurich
KOF Swiss Economic Institute
WEH D4
8092 Zurich, Switzerland

Phone: +41 44 632 42 38

Fax: +41 44 632 11 50

www.ciret.org

info@ciret.org

Aims and Scope

The *Journal of Business Cycle Measurement and Analysis* is jointly published by the OECD and CIRET to promote the exchange of knowledge and information on theoretical and operational aspects of economic cycle research, involving both measurement and analysis.

The OECD has developed various cyclical indicators using quantitative and qualitative information from its member countries. Information about the work of OECD in the field of short-term economic statistics may be found at www.oecd.org/std.

CIRET, the Centre for International Research on Economic Tendency Surveys, is a global forum for economists and institutions, which conduct and analyse business and consumer surveys. CIRET holds international conferences every other year. Information about CIRET may be found at www.ciret.org.

The scope of the *Journal of Business Cycle Measurement and Analysis* encompasses:

- Analysis of cyclical fluctuations
- Business cycle specification, definition and classification
- Statistical approaches to the development of short-term economic statistics and indicators
- Business tendency, investment and consumer surveys
- Use of survey data or cyclical indicators for business cycle analysis

Researchers in the field of business cycle measurement and analysis are encouraged to submit their papers to the Editor-in-Chief, Michael Graff (editor-in-chief@ciret.org). All papers undergo a double-blind review process.

Under the heading “Report” additional contributions, including topics such as sample design, data collection and validation, dissemination and evaluation of survey results, are published. Reports are reviewed by members of the Editorial Board.

Other correspondence related to the journal may also be addressed to György Gyomai, at the OECD (Gyorgy.GYOMAI@oecd.org).

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(33 2010 01 1 P) ISSN 1995-2880 – No. 57241 2010

OECD Journal

Journal of Business Cycle Measurement and Analysis

The *Journal of Business Cycle Measurement and Analysis* is jointly published by the OECD and the Centre for International Research on Economic Tendency Surveys (CIRET) to promote the exchange of knowledge and information on theoretical and operational aspects of economic cycle research, involving both measurement and analysis.

Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model

Elena Angelini, Marta Bańbura and Gerhard Rünstler

Chain-Linking in Austrian Quarterly National Accounts and the Business Cycle

Marcus Scheiblecker

Evaluating German Business Cycle Forecasts under an Asymmetric Loss Function

Jörg Döpke, Ulrich Fritzsche and Boriss Siliverstovs

Do Benchmark Revisions Affect the Consumption-to-Output and Investment-to-Output Ratios in Germany?

Thomas A. Knetsch and Hans-Eggert Reimers

OECD and CIRET News 2010/1

www.oecd.org

www.ciret.org

www.kof.ethz.ch

The *OECD Journal* collection comprises six periodicals: *OECD Journal: Economic Studies*, *OECD Journal: Financial Market Trends*, *OECD Journal: Competition Law and Policy*, *OECD Journal: General Papers*, *OECD Journal on Budgeting*, *OECD Journal: Journal of Business Cycle Measurement and Analysis*. Each periodical within *OECD Journal* retains its own ISSN, volume and issue numbering. The volume numbering of each journal corresponds to the calendar year, e.g. Volume 2009. Volume 2009 of *Journal of Business Cycle Measurement and Analysis* contains two regular issues.

Subscribers to this printed periodical are entitled to free online access:

www.sourceoecd.org/OECDjournal

If you do not yet have online access via your institution's network, contact your librarian or send an e-mail to SourceOECD@oecd.org.

Volume 2010/1

OECD publishing

www.oecd.org/publishing

ISSN 1995-2880

VOLUME 2010
(2 ISSUES)

33 2010 01 1 P 9



771995 288001