EURONA

EUROSTAT REVIEW ON NATIONAL ACCOUNTS AND MACROECONOMIC INDICATORS

1/2016

eurostat

EUROSTAT REVIEW ON NATIONAL ACCOUNTS AND MACROECONOMIC INDICATORS

1/2016

Europe Direct is a service to help you find answers to your questions about the European Union

Freephone number (*):

00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on the Internet (http://europa.eu).

Luxembourg: Publications Office of the European Union, 2016

PDF: ISSN 1977-978X
Cat. No: KS-GP-16-001-EN-N

Print: ISSN 2443-7832
Cat. No: KS-GP-16-001-EN-C

Theme: Economy and finance
Collection: Statistical books

Printed in Luxembourg

Contents

Aims and scope

EURONA is an open access, peer-reviewed, scholarly journal dedicated to National Accounts and Macroeconomic Indicators. EURONA aims at providing a platform for researchers, scholars, producers and users of macroeconomic statistics to exchange their research findings, thereby facilitating and promoting the advancement of National Accounts and Macroeconomic Indicators.

EURONA publishes empirical and theoretical articles within the scope of National Accounts and Macroeconomic Indicators, as well as articles on important policy uses of these statistics. They may relate to both users' and producers' interests, present subjects of general relevance or investigate specific topics.

EURONA is non-partisan and applies the highest standards to its content, by emphasising research integrity, high ethical standards, validity of the findings and cutting edge results. EURONA gives room to all viewpoints.

The articles published in EURONA do not necessarily reflect the views or policies of the European Commission.

Website: http://ec.europa.eu/eurostat/web/national-accounts/publications/eurona

Contact: ESTAT-EURONA@ec.europa.eu

Editors

Editorial board

([1]) Paulus.Konijn@ec.europa.eu
([2]) sartore@unive.it

Editorial

The production of European statistics is governed by a Code of Practice that includes the principles defining their quality. The articles in this issue of EURONA address several of these principles — relevance, accuracy, timeliness and coherence — and thus contribute to improving the quality of statistics.

In the first article, Jan van Tongeren and Ruud Picavet argue that the relevance of national accounts and other integrated statistics could be improved by bringing their compilation closer to the analytical use for which they are designed. They discuss methods to integrate different types of data into a unifying framework like the national accounts. They propose a Bayesian estimation approach that takes the reliabilities of the data, the indicators and ratios explicitly into account. They demonstrate the method using a detailed dataset from the Netherlands. One advantage of the method is to enhance the transparency and reproducibility of national accounts compilation methods. But the authors show that the methods can be extended beyond national accounting.

In the second article, Antonio Chessa discusses a very topical subject in price statistics, which is the use of scanner data to compile consumer price indices (CPIs). The use of 'big data' like scanner data could contribute to CPIs with higher accuracy due to bigger samples. A new method is proposed, the quality-adjusted unit value approach, which is similar to the Geary-Khamis method in international comparisons. The method is applied to several products using actual data and is already used by Statistics Netherlands for the CPI on mobile phones.

In April 2016, Eurostat started publishing the first euro area and EU GDP estimates at 30 days after the reference quarter. In the third article, Arto Kokkinen and Hans Wouters describe the development and methodology of these flash estimates. They elaborate the results of the tests that were undertaken during the development, which focussed on the revisions that were made to the flash estimates in subsequent releases. This work has made a great contribution towards improving the timeliness of national accounts data.

In the fourth article, Robert Obrzut analyses the coherence between the national accounts and balance of payments statistics in Europe. He shows that, although the methodological standards of the two areas are fully consistent, in practice there remain smaller and larger differences. A closer coordination between the different actors in the compilation of these statistics might improve the consistency.

The quality of EURONA is to a large extent assured by the use of the double blind peer review system. Without the peer reviewers, this journal would not be possible. The Editorial Board of EURONA would like to thank the following persons that have reviewed one or more articles that appeared, or not appeared, in the first four issues of the journal: R. Akers, E. Angelini, M. Bloom, D. Buono, A. Cañada, J.F. Divay, L. Ferrara, J. Fischer, C.M. Garcia, L. Fontagné, K. Hayes, S. Humphries, R. Inklaar, C. Ivan-Ungureanu, M. Kellaway, R. Lynch, A. Maravall, M. Marini, C. Mathieu, C. Obst, E. Oltmanns, M. Prud'homme, E. Quilis, F. Ravazzolo, H. Shank, M. Scheiblecker, P. van de Ven, M. Weale and J-L. Weber.

Paul Konijn

Editor of EURONA

1 Bayesian estimation approach in frameworks; integration of compilation and analysis

JAN W. VAN TONGEREN (1) AND RUUD PICAVET (2)

Abstract: The paper shows how formalization of national accounts compilation through explicit design of the national accounts framework and the use of a Bayesian estimation approach with explicit prior reliabilities of data and indicator (ratio) values could bring national accounting closer to analysis. Through formalization of the national accounts procedure it is also possible to extend the approach to projections of national accounts estimates to the future and to apply it to satellite accounting. The paper illustrates the Bayesian estimation in frameworks with help of Netherlands data. It shows how a complete set of estimates could be produced if only a limited data set is available, how this affects the posterior reliabilities of the estimates and how projections could be obtained from the complete set of variables of national accounts while using a limited number of projected values of exogenous variables under different scenarios.

JEL codes: C11, E01, E27

Keywords: Bayesian Estimation, Frameworks, National Economic Accounts, Compilation, Analysis, Projections.

(1) Jan W. van Tongeren is ex-Chief of the National Accounts Branch, UN Statistics Division in New York
(2) Ruud Picavet is ex-Assistant Professor of Development Economics at Tilburg University, Netherlands

1. Introduction

The present paper demonstrates a newly developed methodology in descriptive statistics, in which analysis guides the compilation and not the other way around, as it is commonly the practice in nowadays' national economic and satellite accounting. Analysis referred to here are the methods of using processed data and estimates to arrive at quantitative conclusions about economic, social and environmental issues. Compilation, on the other hand, concerns the methods by which basic data are transformed into processed data, estimates and indicators, so that they can be used in analysis. The link between analysis and compilation is achieved by explicitly defining frameworks of tables, matrices and vectors for use in specific types of analysis. The methodology has the potential to be extended from frameworks of national accounts and satellite accounts to other frameworks in response to desired analyses. The analysis features of the methodology are incorporated in particular by including in the framework analytical indicators that are in most cases defined as binary ratios based on two variables. But also composite ratios, such as the Human Development Indicator (HDI) developed by UNDP (see United Nations Development Programme (2015)) could be used in principle in the methodology. The indicators or ratios that are used in the *compilation* (estimation of variables and indicators), at the same time play an important role in the *analysis*, and also in the *projection* of data and indicator ratios to the future.

The two essential elements of (what may be called) a *systems approach to national accounting* (see United Nations Statistics Division (1999)) are the design of a framework of variables that can be used in compilation, analysis and projections of variables, and the application of a Bayesian estimation approach that is used to estimate the variables in the framework. The design of the frameworks is the most essential element. To our knowledge, the present attempt is the first one in which the Bayesian estimation methodology is applied to sets of variables and analytical indicators incorporated in a framework of matrices (= tables), that explicitly describes the analytical structures of an economy or other phenomena to be studied (see Van Tongeren (2011)) (3).Mathematical details of the Bayesian estimation approach are presented in Danilov and Magnus (2007) and (2008). A similar approach to the Bayesian estimation used here, and based on discussions about the present methodology, was developed by Statistics Netherlands (Bikker et al. (2008, 2013) and an alternative method, called the entropy estimation method, was developed by Robilliard and Robinson (2003) for Madagascar; both alternatives were applied to the compilation of national accounts. Use of the Bayesian methodology in econometrics and modelling is quite common but generally refers to estimates of functional relationships between a limited number of variables based on *time series analysis*. Examples of the use of this methodology in econometrics can be found in Zellner (1996). The *time series analysis* used in econometrics should be distinguished from what might be called here *structural analysis*, in which relations between variables at selected points in time are the essence of the estimation approach as used in the present paper.

Through the framework approach, the variables that serve analysis are identified whilst, through ratios and identities, the relations between the variables are defined and the analysis is being integrated. As data underlying variables and ratios (indicators) are very often inconsistent, the Bayesian estimation approach is used to obtain consistent data and indicator sets. The scope of analysis may be larger than available data on variables and indicators are, so the Bayesian approach is used to complete the set of data on variables and indicators with indirect estimates.

(3) The Bayesian estimation methodology applied to frameworks was part of a cooperation led by Van Tongeren, in which participated Magda Ascues, Dimitri Danilov, Arthur Giesberts, Jan Magnus, Ruud Picavet, Bertha Vallejo, Tugrul Temel.

The approach recommended in this paper has several advantages over the manual methods used until now mainly in national economic accounting. The method is flexible, as it can be extended to frameworks other than national economic accounts and is also not limited to data in monetary units only. It can thus be applied to detailed accounts such as i-o and supply and use tables, satellite accounts and any other data frameworks underlying analytical studies. The method is transparent, as all conditions of identities, ratios and reliabilities are explicitly presented as part of the framework and links between basic data and final estimates are not lost. Updating is easy and quick, as in existing frameworks, indirect estimates can easily be replaced by newly available data and updated estimates, taking this new information into account. This considerably improves the timeliness of estimates and makes updating less costly. The approach is not limited to compilation of frameworks only, but also can be used in projections, as will be shown later on. Experience shows that nonsense results, as a consequence of faulty conditions of identities, ratios and reliabilities can be easily detected. Finally, results cannot be easily manipulated, as final estimates have to satisfy a very large number of conditions, and thus some data or conditions manipulated can be identified through the checks of the very large number of other conditions. The large number of checks (conditions) contribute to the quality of the estimates.

The paper is organized as follows: Framework and Bayesian estimation approach are dealt with in sections 2 and 3. Section 2.1 shows how analytical objectives determine the selection of the variables and indicators that are included in the framework. Section 2.2 applies these principles to the design of the framework of variables and analytical indicators that is used in this paper. Section 3 presents the Bayesian estimation approach with help of a simple example. In the subsequent sections the approach will be applied in the compilation of the framework, i.e. estimation of variables and indicators (section 4), in the analysis of variables of the framework and their relations with help of analytical indicators (section 5) and also in making projections of variables to the future (section 6). In the concluding section (7) are presented views on how the framework approach presented in the present paper could benefit many areas of interdisciplinary analysis and how international standardization could be made more flexible and respond more effectively to the needs of policy analysis in individual countries.

The paper is based on 2010 and 2013 national accounts data from Statistics Netherlands. A very detailed data set has been derived from their website ([4]). The detailed data have been aggregated into the data categories used in the present paper. Links between the detailed data and those used in the present paper have been maintained, but not reflected in the present paper. Those who are interested can download the detailed data set on this location: http://ec.europa.eu/eurostat/product?code=KS-GP-16-001.

2. Analytical frameworks

An example of an analytical framework is presented in table 1. Before entering into the details of this framework, attention may first be paid to details of the role of analytical frameworks in the proposed methodology.

([4]) https://www.cbs.nl/en-gb/figures/statline/make-your-own-tables

2.1. General principles of analytical frameworks

The origin of the framework approach are national accounts, in which many economic variables are incorporated in a comprehensive framework of the national economy in order to be used jointly in analysis of the macroeconomic structure and functioning of that economy. Contrary to national accounting, however, frameworks of variables referred to in this paper are not standardized (internationally, as are national accounts and some satellite accounts) but are designed prior to and in accordance with the research problem to be investigated and/or the analysis intended. Concepts in economic frameworks therefore may differ from concepts in national economic and satellite accounting, if the analysis intended requires this. Also, frameworks may not necessarily focus only on economic analysis and be extended to other fields. Thus, they may also include variables that are not measured only in monetary terms.

In order to reflect an analysis in a framework, it is necessary that the framework includes all relevant variables that are used in the formulation of the research question and/or analysis and that it can be measured directly or indirectly on the basis of existing statistics. It is thus a research framework for quantitative analysis.

As a consequence, if a *productivity analysis* were the objective of the analysis, then the framework should include variables on output, value added, investments in current and constant prices, and employment. A further breakdown of these variables should also be done by economic activities such as agriculture, industry, transport and trade, and services. Or, if analysts prefer carrying out such analysis in more details, they may use further breakdowns of variables based on ISIC and CPC and other international classifications (5).

If *consumption patterns* were studied in the analysis, the framework should include variables on household final consumption and disposable income by income groups, by income groups and regions, or by other household groups. Final consumption should also be broken down by functions.

If the *impact of population growth* on economic development were studied, there would be a need to include demographic variables in addition to economic variables. If *economic impacts of tourism* were studied, variables would include the number of tourists in addition to economic variables measuring the expenditures per tourist and economic impact variables of tourism. Or if the *environment impacts of economic development* were the objective of the study, the framework would include variables on emissions into air, water and land, in addition to measures of the economic costs of those emissions, production related variables measuring the causes of the emissions and also the production and employment variables aimed at reducing or avoiding the consequences (impacts) of the emissions.

Selecting the relevant variables is decisive, but not a sufficient condition for defining a coherent framework of analysis. For that purpose, relations should also be defined between the variables, so that the analysis of single variables can be integrated into a more comprehensive analysis in which the interactions between them are taken into account.

The two types of relations that are most suitable for integration of single variables into a comprehensive analysis are *identities* and *ratios* between variables.

In the above productivity analysis, *identities* define for instance the relation between output, intermediate consumption and value added, or the sum-relation between output and value added for the economy as a whole and for individual economic activities. In the analysis of consumption, patterns identities that define the relation between consumption of income groups

(5) See United Nations Statistical Division (UNSD) registry of classification: http://unstats.un.org/unsd/class/default.asp

and total consumption, and the population groups of different labour income levels who add up to the total working population may be included. In studies of the impact of changes in demographics on economic development, additional identities would define the relation between subgroups of the population on the one hand and the total or working population on the other. In tourism studies, additional identities would define the relation between the total flow of tourists as the sum-total of tourists visiting different regions, staying in hotels, visiting tourist attractions, etc. In the case of economic-environmental analysis, sum-totals would be defined between the total use of water and the use of water by different population groups and industries, CO2 and other emissions caused by different economic activities and the total emissions into water, air and land resulting from these partial impacts.

Identities defined between variables are a means of arriving at consistent values of the variables. In order to improve the overall consistency of values of variables, it is important that a comprehensive set of identities is defined in the framework. If some of the identities are missing, data and estimates would not satisfy those identities and this would affect the consistency of the values included in the framework. If identities cannot be defined in a framework due to missing groups of the population or missing income or expenditure items, the framework must be extended so that it includes these missing variables.

Ratios are generally analytical indicators that are used in analysis. They may include ratios between value added and output for the economy as a whole and by economic activities, reflecting productivity, the ratio between value added and investments in each activity, emissions per unit of output of different industries, expenditures per tourist, or compensation per employee. Ratios define the structure of the set of variables included in the framework. Ratios are less precise than identities, as identities should hold exactly while ratio values may vary within limits.

A second element of the approach is *the use of a Bayesian estimation methodology* in order to obtain values of the variables and also ratios in the framework. The Bayesian estimation methodology, which will be explained in the next section (3) is the mathematical formalization of the estimation approach used in national accounting. This methodology, as in national accounting, is needed in order to reconcile conflicting values of variables and ratios (=indicators) and also to derive indirectly values of variables and ratios for which no direct information is available. When using the Bayesian estimation approach, another element is introduced into the frameworks and this is the *reliability of data and of values of indicator ratios*. By using the framework approach together with the Bayesian estimation, the present paper will show that it is possible to closely link a comprehensive compilation (i.e. estimation) of variables with a comprehensive analysis using the values of the variables.

Frameworks of national accounts, as other frameworks referred to here, are called logical frameworks, as the variables included are selected depending on their relevance for the formulation and response to research questions, together with data, analytical ratios (indicators) and identities defined between the variables in order to serve pre-established types of analysis. The framework also includes *prior* reliability indicators of data and ratio values. The Bayesian estimation approach uses these prior elements in order to arrive at (improved) *posterior* estimates within the framework of variables, ratios and posterior reliability of variables and ratios.

Table 1: Simple SNA Framework, including, ratios, identitites and reliabilities, 2010 Netherlands data
(million EUR)

Part 1		National economy	NFC & FC	GOV	HH & NPI	ROW	
Output/imports		1 241 940				401 585	3
Intermediate consumption/exports		611 167				454 398	4
GDP		631 512					5
Labour income (compensation of employees and mixed income)		338 844			334 411	4 433	6
Product and production taxes less subsidies		62 036		60 963		1 812	7
Difference VAT imputed and received by GOV		739					8
Consumption of fixed capital		106 982	59 969	20 588	26 425		9
Operating surplus, net (other than mixed income)		123 650	123 650	0	0		10
Property income, receipts less payments		41 129	– 31 631	5 357	35 772	– 9 498	11
Taxes on income and wealth			– 13644	70 134	– 54 587	– 1 903	12
Social transfers (incl. social transfers in kind)			16 151	– 93 769	76 969	649	13
Other current and capital transfers		29 034	– 15 688	– 13 346	16 293	12 741	14
Disposable income, net (adjusted for (i) transfers in kind, (ii) capital transfers)			78 838	29 339	408 858		15
Final consumption		449 742		56 378	393 364		16
Saving, net (changes in net worth due to net saving plus capital transfers)			78 838	– 27 039	15 494	– 45 318	17
Gross capital formation (incl. acquisitions less disposals of non–produced assets)		128 957	69 738	25 415	33 804		18
Net lending			69 069	– 31 866	8 115	– 45318	19
Population *1000					16 575		20
Employment *1000		7 391					21
Vertical identities	GDP	– 739					23
	Value added components	0					24
	Disposable income		0	0	0		25
	Saving–disposable income		0	0	0	739	26
	Net lending–saving		0	0	0	0	27
Vertical behavioristic and technical, ratios	Exports as % of output					36.6	28
	Propensity to consume of HH's (%)				96.2		29
	GDP as % of output	50.8					30
	Labour income as % of GDP	53.7					31
	Product and production taxes less subsidies as % of GDP	9.8					32
	Consumption of fixed capital as % of GDP	16.9					33
	Operating surplus as % of GDP	19.6					34
Distribution taxes less subsidies within GOV	Product and production taxes less subsidies (%)			46.5			35
	Taxes on Income and Wealth (%)			53.5			36
Per capita/ worker	Per capita final consumption				23.732		37
	Labour income per worker (*1000 euro)	45.845					38
	Employment share in population (%)	44.6					39
A	B	C	D	E	F	G	

Source: Statistics Netherlands.

Table 1 (Cont.)

Previous year values for national economy	Horizontal identities	Coefficients of distribution of GDP by			illustrative selection of comments in cells: including reliability indicators (@...@), values of indicator ratios (… = #...#) identity definitions (= … + …–)
		Components of disposable Income (%)	Expenditure components (%)		
1 201 256	0		63.6	3	C3: @S@
			72.0	4	I3: = C5 – C16 – C18 – (G4 – G3)
617 650				5	
	0	53.7		6	J6: @P@ C6/C5 = #0.537#
	– 739	9.8		7	E7: @F@
				8	G7: @H@
	0			9	I9: = C9 – (D9 + E9 + F9)
	0			10	
0	0	6.5		11	
	0	11.1		12	
	0	14.8		13	J13: @P@ E13/C5 = #0.148#
0	0	4.6		14	
	– 739			15	
	0		71.2	16	
	0			17	
127 772	0		20.4	18	K18: @P@ C18/C5 = #0.204#
	0		20.7	19	K19: Share of Capital Formation in GDP, previous year
				20	F20: @F@
				21	
				23	
				24	
				25	D25: = D15– (D10 + D11 + D12 + D13 + D14)
				26	
				27	E27: = (E9 + E17) – (E18 + E19)
				28	
	Variables with basic data			29	F29: @H@ F16/F1 = #0.962#
	Variables without basic data			30	C30: @H@ C5/C3 = #0.508#
	Independent Identities			31	
	Dependent Identities			32	
	Ratios with values			33	
	Ratios without values			34	
				35	
				36	
				37	
				38	C38: @L@ C6/21 = #45.845#
				39	
H	I	J	K		L

Table 1 (cont.)

Part 2		National economy	NFC & FC	GOV	HH & NPI	ROW	
Distribution between sectors	Distribution of supply between output and imports (%)	75.6				24.4	40
	% of supply used for capital formation	7.8					41
	Distribution of capital formation (%)		54.1	19.7	26.2		42
	Labour income (compensation of employees and mixed income)(%)				98.7	1.3	43
	Product and production taxes less subsidies (%)			98.3		2.9	44
	Difference VAT imputed and received by GOV, as % of product and production taxes less subsidies	1.2					45
	Consumption of fixed capital (%)		56.1	19.2	24.7		46
	Operating surplus, net (other than mixed income)		0	0.000	0.000		47
	Property income, receipts less payments (%)		76.9	13.0	87.0	23.1	48
	Taxes on Income and wealth (%)		19.5		77.8	2.7	49
	Social transfers (incl. social transfers in kind) (%)		17.2		82.1	0.7	50
	Other current and capital transfers (%)		54.0	46.0	56.1	43.9	51
	Final consumption (%)			12.5	87.5		52
	Disposable Income, net (adjusted for (i) transfers in kind, (ii) capital transfers) (%)		15.2	5.7	79.1		53
	Saving, net (changes in net worth due to net saving plus capital transfers) (%)		83.6	37.4	16.4	62.6	54
	Net lending (%)		89.5	41.3	10.5	58.7	55
Distribution of disposable income	Labour income (compensation of employees and mixed income) (%)				81.8		56
	Product and production taxes less subsidies (%)			207.8			57
				12.7			58
	Operating surplus, net (other than mixed income) (%)		156.8	0.000	0.000		59
	Property income, receipts less payments (%)		– 40.1	18.3	8.7		60
	Taxes on Income and wealth (%)		– 17.3	239.0	– 13.4		61
	Social transfers (incl. social transfers in kind) (%)		20.5	– 319.6	18.8		62
	Other current and capital transfers (%)		– 19.9	– 45.5	4.0		63
GOV deficit as % of GDP				5.0			64
Policy indicators	Output/capital ratio (ypinv) (%)	10.8					65
	GDP growth rate (Δy) (%)	2.2					66
	Growth rate output (Δp) (%)	3.4					67
	Growth rate investments (Δinv) (%)	0.9					68
A	B	C	D	E	F	G	

Source: Statistics Netherlands.

Table 1 (cont.)

Previous year values for national economy		**Coefficients of distribution of GDP by**			**illustrative selection of comments in cells: including reliability indicators (@...@), values of indicator ratios (…. = #...#) identity definitions (= … + … –)**
		Components of disposable Income	**Expenditure components**		
				40	G40: @L@ C3/(C3 + G3) = #0.244#
				41	
				42	
				43	
				44	
	Variables with basic data			**45**	
	Variables without basic data			**46**	
	Independent Identities			**47**	
	Dependent Identities			**48**	
	Ratios with values			**49**	
	Ratios without values			**50**	
				51	
				52	
				53	
				54	
				55	
				56	
				57	
				58	
				59	
				60	
Variables to be estimated			**58**	**61**	
				62	
Information items			**112**	**63**	
	Data		27	**64**	
	Ratios		62	**65**	
	Identities (independent)		23	**66**	
				67	
Information factor			**1.93**	**68**	
H	**I**	**J**	**K**		**L**

2.2. Analytical framework in this paper

The analysis embedded in the present framework of table 1 (parts 1 and 2) is close to the economic analysis embedded in the System of National Accounts (SNA) (see United Nations (2009)). However, there are some differences in analytical emphasis. SNA abbreviations of sectors have been used in table 1 and throughout the paper: NFC for Non-Financial Corporations, FC for Financial Corporations, HH for Households, NPI for Non-Profit Institutions serving households, and ROW for the Rest Of the World. Its main aim is to show how GDP (cell *C5* ([6]) in the framework of table 1) is generated and thereafter distributed to disposable income of sectors (HH's and NPI's (*F15*), NFC's and FC's (*D15*), and GOV (*E15*)) through different instruments. When distributing GDP by expenditures (final consumption (*C16*), capital formation (*C18*), exports (*G4*) and imports (*G3*)), the discrepancy between the distributions of disposable income and expenditures results in deficits and surpluses of sectors (net lending: *D19*, *E19*, *F19*) and of the ROW (*G19*) and thus in financial flows between sectors. The financial flows are not recorded in detail in the present framework to avoid unnecessary complications at this point.

The framework in table 1 includes five elements, which are used in the Bayesian estimation approach and which will be explained below in section 3 with help of a simple example. They are: (i) variables, (ii) data, (iii) ratios, (iv) identities and (v) reliabilities of data and ratio values. Variables and ratios (or analytical indicators) define the analysis; identities and values of ratios define the consistency of the framework; and data, ratios, identities and reliabilities determine the basis of the compilation. Data, ratio values and identities are called here information items: they are used in the estimation of the variables. The minimum number of information items is equal to the number of variables. If there are fewer information items than variables, no estimates of variables can be made. On the other hand, the number of information items may be much larger than the number of variables, as data and ratio values do not have fixed values, but are defined within reliability intervals. Availability of more information items generally implies that more checks are available, which may enhance the accuracy of the estimates.

The *variables* are presented in the upper part of the framework, i.e. *C3:H21*. They define the scope of the analysis. If no variables are included for specific phenomena, no analysis of these phenomena is possible.

Data are the values that are available for variables. In general, there are fewer available data than variables that should be estimated (compiled). The variables for which no data are available are estimated with help of ratio values and identities. In the framework as presented in table 1, data are available for variables in dark purple cells and no data are available for variables in light purple cells (as indicated in the legend). Data without any coloring are not used in the estimation of the variables of the framework. This applies in particular to the 'Difference VAT imputed and received by GOV' in cell *C8*. This difference is assumed to be equal to zero in the present framework. This cell and other zero values in *E10* and *F10* are not treated as variables in the framework.

Values and definitions of *ratios* are presented in the cells *J3:K21* to the right of the variables and in cells *C28:G68* below the variables. Ratios are binary relations defined between the variables of the framework. They are typically indicators that could be used in the analysis and also in the compilation of the variables. Some of the ratio functions and values are presented as examples in comments. A distinction is made between ratios with and without values; the distinction is made with help of different shades of green as indicated in the legend to the table. For the first ones there are prior values, while for the second ones there are not. Values for the latter

([6]) EXCEL Cell references are made throughout the paper. They refer to the row/column location of EXCEL cells in tables 1 and 1A, which include Netherlands data. The EXCEL rows/ column locations are included in those tables.

are derived with help of the Bayesian estimation. Ratios that are not shaded, such as the one in K19, defined as the share of capital formation in GDP of the previous year, are not used in the compilation, neither as ratios with or without values.

Identities are definitions and other consistency relations between the variables. These cells, identified with different shades of blue in the legend of the table, include the definition of the identities, which are formulated in such a manner that they are equal to zero. They may not hold between available prior data, but should hold between posterior estimates of the variables. They are presented below the variables in cells *C23:G27*, and to the right of the variables in cells *I3:I19*. Some of the identity functions are presented as examples in comments.

Prior reliabilities of data and ratio values are based on the assignment of reliabilities by those (e.g. national accountants) compiling the data and assigning the ratio values. They are subjective values based on experience. The reliabilities are expressed with help of symbols P(poor), L(low), M(medium), H(high), S(superior), F1(near-fixed) and F(fixed), meaning that the posterior estimates could deviate from the prior values with resp. 24 %, 12 %, 6 %, 3 %, 1 %, 0.1 % and 0 %, based on the subjective experience of those familiar with the data and ratio values. Reliabilities of some of the data and ratio values are presented as examples in comments (see column L of table 1). F-reliabilities for data and F1-reliabilities for ratios are generally assigned when posterior values should not differ from the prior values. This may apply to data of GOV budgets or other administrative data, such as customs data on imports and BOP data, as well as social security data. Such data are not necessarily highly reliable, but they cannot be changed as they are the result of accepted administrative processes. Ratios with F1 reliabilities may be tax rates that are administratively fixed, or GOV deficits as percent of GDP, which are also the result of an administrative process. The percentage deviations, while illustrative, do reflect in general implicit assessments made by national accountants when they reconcile national accounts data. As only a limited number of prior standard deviations are used, namely six (F, F1, S, M, L, P), it is sufficient to classify the data and ratio values by groupings of those standard deviations.

For the purpose of explaining the framework and the methodology, it is important to know how many information items there are in the framework of table 1 and how they relate to the number of variables to be estimated. The number of information items, as shown at the bottom of the table, is 112, which is the sum of the number of data (27), ratios (62) and independent identities (23). This number is much larger than the number of variables in the framework (58). As only 58 information items are needed to estimate all variables, a large number of checks are available to arrive at estimates of the variables. The ratio between the number of information items divided by the number of variables is 1.93; it is called the information factor. The minimum number of information items is equal to the number of variables to be estimated; in this case the information factor is equal to 1. The higher the information factor (>1) the more checks are available. Whether these checks have a significant impact on the precision (posterior reliability) of the posterior estimates of the variables depends on whether these checks are of a high quality (high prior reliability). More on the relation between the number of information items and the number of variables will be mentioned at the end of section 3, when dealing with the estimation method.

When implementing this analysis, several indicators and supporting variables — some of which are referred to above — are used and made explicit in the framework. The indicators and supporting variables are the following:

1. Propensity to consume of HH's (*F29*), defined as the ratio between HH consumption (*F16*) and HH disposable income (*F15*).
2. Tax rates for product and production taxes as percent of GDP (*C32*) on the one hand and income and wealth taxes as percent of disposable income of NFC's and FC's (*D61*) and HH's (*F61*). For GOV as a recipient sector, *E61* stands for the ratio of taxes on income and wealth received as percent of disposable income of the GOV sector.
3. I-O technical coefficients (*C30*), that are defined in this paper as the ratio between GDP (*C5*) and Output (*C3*).
4. Exports as percent of output (*G28 = G4/C3*) and imports as percent of total supply (*C40 = C3/(C3 + G3*).
5. Average labour income per worker, including employees and own account workers (*C38*), which is defined as the ratio between labour income (*C6*) and the number of workers (*C21*).
6. Average final consumption per individual (*F37*), which is defined as the ratio between *F16/F20*.
7. GDP (*C5*) is distributed to arrive at disposable income per sector (*D15:F15*), using five different instruments, i.e. (i) Labour income (i.e. compensation of employees plus mixed income) received by HH's and ROW (*J6 = C6/C5*), (ii) Taxes (product and production taxes less subsidies and income and wealth taxes) received by GOV from HH's, NFC's and FC's, and ROW (*J7 = C7/C5* and *J12 = E12/C5*), (iii) Property income, mainly received by HH's and NFC's and FC's (*J11 = C11/C5*), (iv) Social transfers, received by HH's from GOV and to a minor extent by the ROW, including transfers in kind (*J13 = E13/C5*), and (v) Other current and capital transfers, mainly received by HH's and the ROW (*J14 = C14/C5*). The value of each instrument is expressed as a percent of GDP.
8. Percent distribution of disposable income between sectors (*D53:F53*).
9. Percent distribution of GDP by expenditures (*K3:K18*).
10. GOV tax receipts distributed among income and wealth taxes, and product and production taxes (*E35:E36*).
11. Percent distribution between collective consumption of GOV and individual consumption of HH's (*E52, F52*).
12. GOV deficits as percent of GDP, which is a policy indicator used in the European Union (*E64 = E19/C5*).
13. Percent distribution of net lending, firstly between the lending sectors of HH's and NFC's and FC's (*D55, F55*), and secondly between the borrowing sectors and GOV and ROW (*E55, G55*).

It is important to note that the framework of the present paper only includes data in current prices, and no information on prices and variables expressed in constant prices. Therefore, growth rates of output and GDP and i-o coefficients are in current prices, which deviates from the practices in national accounting, where such indicators are generally expressed in constant prices.

Important analytical indicators are those mentioned under point 7. They are not only used in the analysis of section 3 and the projections of section 4 but also in the compilation described in section 6. It has been assumed here that the instruments used to distribute GDP to disposable income add up to disposable income for the total economy, except for the part to and from the ROW. In the case of the Netherlands data, this assumption is approximately correct; i.e. the total measured for the instruments is approximately equal to GDP and national disposable income. Conceptually, however, the relations between these instruments and GDP and national disposable income are more complex, as the transfer instruments include two types: first, direct distributions of GDP to income components of disposable income, such as labour income and part of property income that are directly distributed from production units to HH's, and product and production taxes that are directly distributed from production units to the GOV sector. Second, re-distributions of income, such as income taxes that are levied on labour and property income, and social transfers that are paid from taxes received by the GOV sector. In particular, the inclusion of the second type of transfer instruments may involve some double counting. This needs to be further examined.

Some of the mentioned indicators may be used as policy instruments as a means to influence developments in the short or long term, including tax rates (2) and GOV tax receipts distributed between product and production taxes on the one hand, and income and wealth taxes on the other (point 10). This is also the case for some of the instruments through which GDP is distributed to disposable income per sector (point 7), such as social transfers, taxes and compensation of employees. Others are technical, behaviouristic and other relations that reflect the functioning of the economy.

In order to carry out the analysis effectively, some SNA concepts have been adapted to the kind of analysis pursued here:

1. Saving has been replaced by the SNA concept of changes in net worth due to capital transfers and saving. This adjusted SNA concept of saving is close to the concept of saving in economic theory (S), which is equal to capital formation (I).
2. For the same reason capital transfers have been included with other current transfers, so that they are reflected in the adjusted concept of saving, mentioned in the previous point.
3. Capital formation includes the purchase and sale of non-produced non-financial assets, which is different from the concept in the SNA. This affects capital formation by sector, but not for the total economy. Purchases and sales of non-produced non-financial assets for the total economy is zero, as no sales/purchases of non-produced non-financial assets are included in the ROW (see next point).
4. Sale of patents and licenses to the ROW have been treated as negative capital formation of NFC's and FC's, and included in exports. No correction has been introduced for acquisition of non-produced assets by the ROW, as it is presently done in the Netherlands national accounts. As a consequence, capital formation reflects all purchases and sales of capital goods (produced and non-produced) by resident sectors and no purchases of non-produced assets by the ROW.
5. Disposable income has been replaced by the SNA concept of adjusted disposable income. This includes transfers in kind and also reflects capital transfers mentioned in point 1.

3. Bayesian estimation in frameworks; illustration of the methodology

The Bayesian estimation method is a mathematical formulation of the compilation methods conventionally used in frameworks of national accounting. With help of the mathematics, some software has been developed, including the estimation program as well as an interface between the framework (defined in EXCEL) and the program. As in conventional national accounting, the Bayesian estimation methodology is applied to frameworks of variables and analytical indicators. The frameworks are predesigned for specific types of analysis and values of variables and indicators are included, if available. Estimates are made using the Bayesian estimation methodology, in which prior values of data and ratios are not fixed but subject to the reliability of the indicators that are used in the compilation. We explain the Bayesian methodology later on in this section.

The mathematics of the compilation method are complex, but may be illustrated with a simple estimation of a few variables and indicators extracted from the framework of table 1 and presented in table 1a. The details and results of the estimation method are described in tables 2a and 2c and a graphical presentation in graph 2b. All tables are based on the same 2010 Netherlands data as in table 1.

Table 1a: Illustrative simple framework
(million EUR)

	GDP (Y)	Gross capital formation	GOV final consump-tion	HH/NPI final consumption (C_{HH})	ROW					
GDP & expenditures	631 512	128 957	56 378	393 364	401 585	Imports	0	SUT identity		73
Consumption of fixed capital	106 982				454 398	Exports				74
Difference VAT imputed and received by GOV	739									75
Disposable income by sectors (*R*)		78 838	29 339	408 858	8 234		0	GDP-Disposable Income identity		76
Population (*POP*)				16 575		Variables with basic data				77
						Variables without basic data				78
Per capita HH final consumption (C_{HH}/POP)				23.732		Independent Identities				79
	National Economy	NFC & FC	GOV	HH & NPI	ROW	Ratios with values				80
B	C	D	E	F	G	H	I	J	K	

Source: Statistics Netherlands

As a first step, and in order to make a two-dimensional graphical presentation possible, reliabilities will be interpreted as percent intervals of data and ratios. It will be assumed that all data and ratio values are fixed (reliability F), except for prior reliabilities of C_{HH} (HH final consumption, cell *F71* in table 1a) and *Y* (GDP, in cell *C72*), which are L(low) and M(medium) respectively, and indicator C_{HH}/POP (per capita final consumption of HH's, cell *F79*), with M prior reliability. Thus, if a national accountant believes that a data item *X* has a prior reliability of 6 %, its posterior value would be between *X*(1 – 6 %)* and *X*(1 + 6 %)*. The prior reliability is thus interpreted as a fixed interval, within which posterior estimates would be located. The identities are the supply-use identity $Y = C_{HH} + constant\ (1)$ in *I73* and the GDP-disposable income identity $Y = R_{HH} + constant\ (2)$ in *I76*. The *constant (1)* in *I73* is the sum of capital formation (cell *D73*), GOV consumption (*E73*) and exports (*G74*) minus imports (*G73*). The *constant (2)* in *I76* is the sum of disposable income of NFC and FC's (*D76*), HH's (*F76*) and GOV (*E76*), net factor income and transfers, from abroad (*G76*) plus depreciation (*C74*) and plus the difference between product taxes imputed and transferred to the GOV (*C75*). All elements included in *constants (1)* and *(2)* of *I73* and *I76* have fixed values (prior standard deviation = 0). In practice, prior reliabilities are set by national accountants on the basis of their assessment of the reliability of the data and ratio values. In the present example, they are set illustratively. How this works out quantitatively in the estimation process is illustrated in tables and graphs 2a-c below.

At this point, the objective is to arrive at posterior values of C_{HH} and *Y* within the prior reliability intervals of all data and ratio values, and satisfy zero values of identities. This means that the posterior values of variables are consistent within the framework. It should be noted, however, that consistency only implies that posterior values of variables and ratios satisfy the prior conditions of data, ratios, identities and prior reliability intervals of data and ratio values. It does not necessarily imply that posterior values are close to reality. That fully depends on the prior conditions. If those prior conditions are not realistic, posterior values, even though consistent, are also not realistic. If, however, the prior conditions are realistic, the posterior estimates are also realistic and thus close to reality.

Table 2a: Prior values with conditions of identities, ratios and reliability intervals. Linear two-dimensional model

			Prior Reliability		Minimum value	Maximum value	
(1)		(2)	(3)	(4)	(5)	(6)	
Data		Prior Value	Reliability Indicator	+/– % Interval			
Final consumption HH's (C_{HH})		393 364	L	12.0	346 160	440 568	C_{HH}
GDP (*Y*)		631 512	M	6.0	593 621	669 403	*Y*
Ratios	Definition (formulas)						
Final consumption per capita	C_{HH}/POP	23.732	M	6.0	369 762	416 966	C_{HH}
Identities							
GDP-disposable income	$Y = R_{HH} + (R_{NFC\&FC} + R_{GOV} + Deprec - T_{xprod-diff} + R_{ROW})$				631 512	631 512	*Y*
Supply-use	$Y = C_{HH} + (C_{GOV} + K + X - M)$				584 308	678 716	*Y*
	$Y = C_{HH,POP} + (C_{GOV} + K + X - M)$				607 910	655 114	*Y*
	$C_{HH} = Y - (C_{GOV} + K + X - M)$				355 473	431 255	C_{HH}

Source: Author's calculations

Table 2a ([7]) presents the prior minima and maxima of *Y* and C_{HH} in the last two columns. In the first two lines final consumption of HH's (C_{HH}) and GDP (*Y*) have a prior reliability interval

([7]) Please note that the figures on minima and maxima of C_{HH} and *Y* in the last two columns cannot be derived accurately from the figures in the columns before. The reason is that all figures are rounded off to the decimals shown, but calculations have been carried out at the level of more decimals.

of 12 % and 6 %. Accordingly, minima and maxima of the two data items have been calculated by adding to and deducting from the prior values, 12 % and 6 % respectively. The ratio of final consumption per capita (C_{HH}/POP) also generates a minimum and maximum value of C_{HH}, because the *POP* value in the denominator is fixed. This ratio has a reliability of 6 % and as *POP* has a fixed value (16 575 thousand inhabitants), the minima and maxima of C_{HH} have been derived by adding to or deducting from the prior value of the ratio 6 % and multiplying the result by the *POP* figure. The GDP-disposable income identity results in only one figure for *Y* (631 512 = minimum and maximum), as *Y* is equal to the fixed values in table 1a of disposable income of all sectors ($R_{NFC\&FC}$, R_{GOV}, R_{HH}) plus factor income and transfers from abroad (R_{ROW}), depreciation (*Deprec*) and a difference between imputed and actual taxes on products paid ($Tx_{prod-diff}$). The supply-use identity in table 2a has two sets of minima and maxima for *Y*. The first one results from using the minima and maxima of C_{HH} in the first line of the table and adding those to the fixed values in table 1a of GOV consumption (C_{GOV}), capital formation (*K*) and exports (*X*) minus imports (*M*). The second one results from using the minima and maxima of C_{HH}/POP in the third line of the table. The supply-use identity results in one minimum and maximum for C_{HH}/POP by deducting from the minimum and maximum of Y in the second line of the table the fixed values in table 1a of C_{GOV}, *K*, *X* and minus *M*.

The above minima and maxima are reflected in graph 2b in a graphical presentation of the prior conditions. Maxima and minima of *Y* are represented by vertical lines and of *C* by horizontal lines. The C_{HH}/POP condition is represented by two horizontal lines. And the identity $Y - C_{HH} = C_{Gov} + K + X - M$ by a line with an upward slope. The points where the lines coincide ($\alpha, \beta, \gamma, \delta, \varepsilon$) are the vectors Y, C_{HH} that determine the posterior minima and maxima of *Y* and C_{HH}.

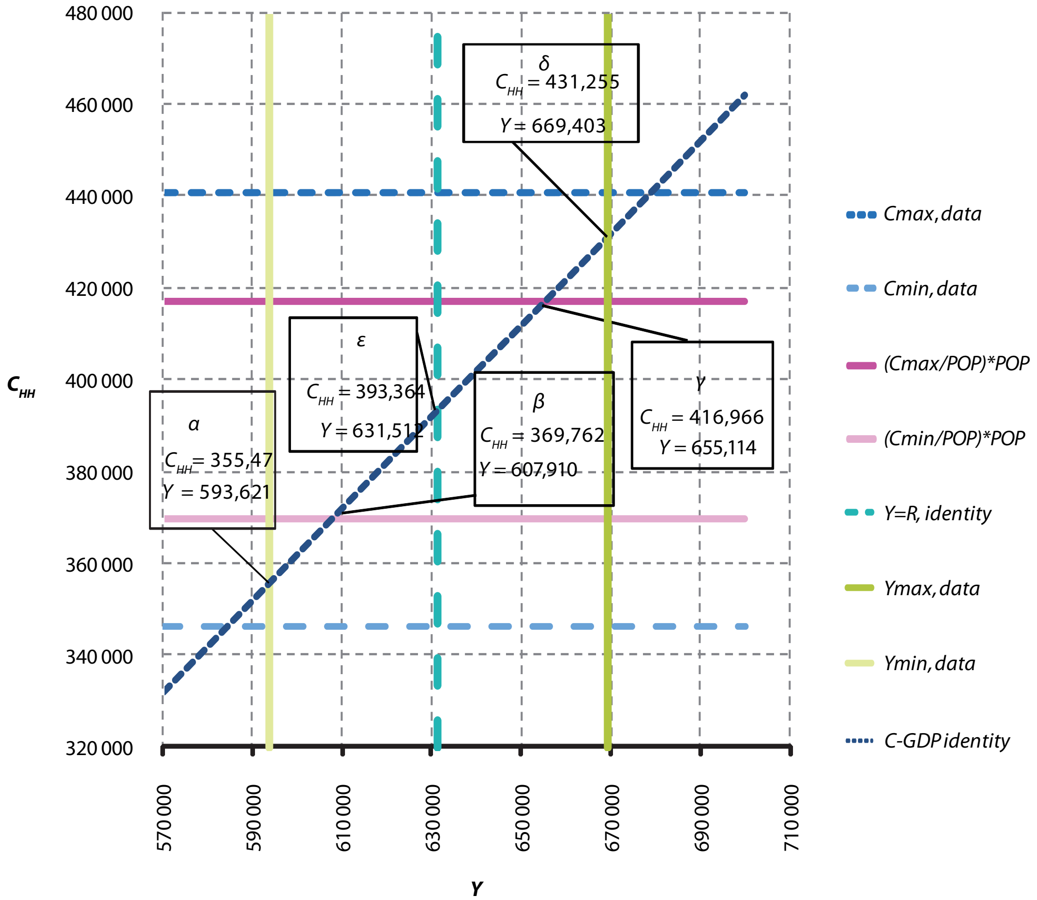
Three options are presented in table 2c and the minima and maxima are reflected in the graphical presentation of graph 2b.

1. In option (i), posterior estimates are based on three conditions, i.e. the minima and maxima of C_{HH} and *Y*, the minima and maxima of C_{HH}/POP, and the supply use identity. The posterior minima and maxima are then determined by vectors β and γ.
2. In option (ii), it has been assumed that values of C_{HH} and *Y* are not available and thus, only the minima and maxima of C_{HH}/POP and supply-use identity apply. This has no impact on the posterior minima and maxima of C_{HH} and *Y*. The same posterior values of C_{HH} and *Y* result in vectors β and γ, as in option (i).
3. In option (iii), all conditions hold, including the supply-use (C-GDP) identity and this results in only one value of the vector (C_{HH}, *Y*), i.e. ε.

The vectors α and δ are relevant in none of the three options, but actually apply to a fourth option, i.e. when only the GDP (*Y*) minima and maxima would be used as restrictions and the supply-use (C-GDP) identity holds.

In table 2c, columns (3), (7) and (11) display the posterior estimates of the minimum and maximum values *Y* and C_{HH}, which satisfy the conditions of the three options mentioned above. They are referred to as LP (*linear programming*) estimates, in which fixed prior intervals are assumed; the LP approach was used in Van Tongeren (1986). The *Bayesian estimates* in columns (4), (8) and (12) will be explained below. The minima and maxima of *Y* and C_{HH} correspond to points (vectors) β, γ and ε in graph 2b. The three LP columns also show the posterior central values of C_{HH} and *Y*, and of the ratio C_{HH}/POP, and posterior percent intervals. The latter ratio uses the posterior minimum and maximum values of C_{HH} and *Y* in each option and the fixed value of *POP* in table 1a. The central values are calculated as the averages of posterior minima and maxima for each variable and ratio in each option. The posterior percent intervals are calculated as the difference between posterior minimum and maximum, divided by 2, as a percent of the central value.

Graph 2b: Graphical presentation of the linear two-dimensional model



Source: Author's calculations

A few conclusions may be drawn from the figures in table 2c:

1. When comparing prior values with posterior values (column (2) with column (3) in option (1)), it can be observed that the maxima of Y and C_{HH} decrease and the minima increase, so that the posterior interval for C_{HH} decreases from 12 % to 6 % and that of Y from 6 % to 3.7 %. The minima and maxima C_{HH}/POP ratios remain the same between prior and posterior estimates; the reason is that these C_{HH}/POP ratio minima and maxima determine the effective minima and maxima of C_{HH} and therefore, are the cause of the changing posterior intervals of Y and C_{HH}. This can be easily verified in the graph of figure 2b, where points β and γ represent the posterior values of the vectors, which are located at the lines of C_{HH}/POP max and min.

2. There are no differences between the posterior results, whether C_{HH} and Y data are available or not in options (1) and (2). This was also made clear above in the graphical presentation 2b; points β and γ are the posterior vectors of both options (1) and (2).
3. In option (3), only one set of posterior estimates are obtained for Y and C_{HH}, without any difference between minimum and maximum. The posterior estimates of this option are in-between values of the estimates in options (1) and (2). In graph 2b, these estimates are represented by vector ε. The latter is obvious, as in option (3) all conditions are set and as in options (1) and (2), only part of the conditions apply.

It should be noted in table 2c that all three options have the same central values and that the central value of prior and posterior values is also equal. This holds in this particular simple case, but is not necessarily valid in other more complex frameworks.

Table 2c: Prior and posterior values of variables and ratios in linear (LP) and Bayesian estimation (million EUR)

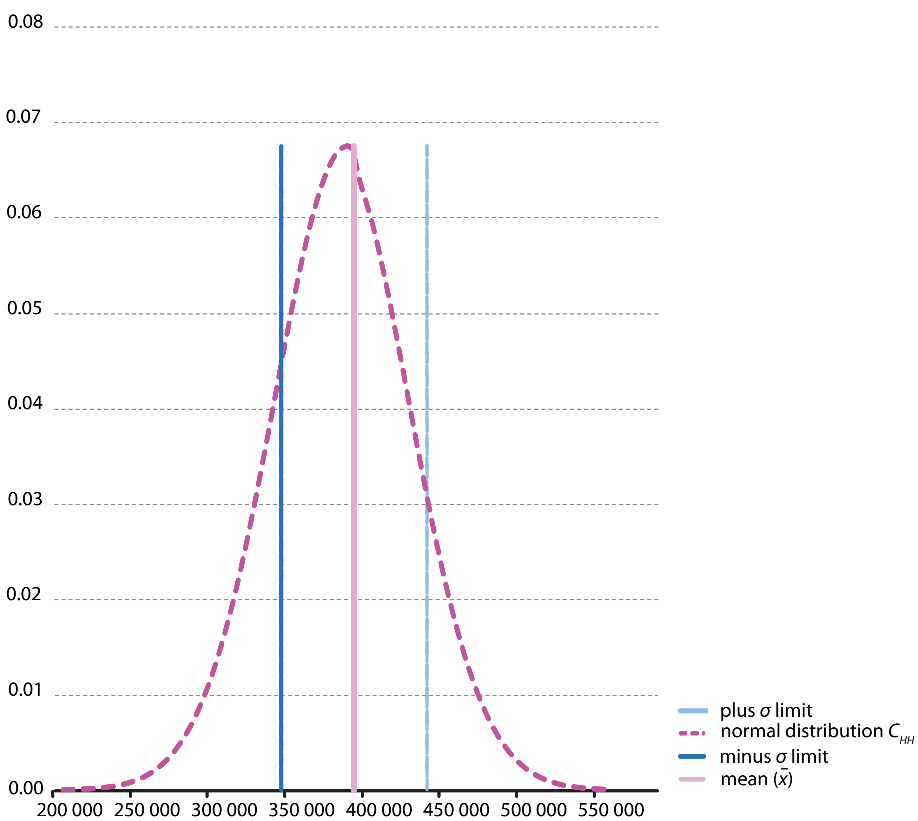
		Prior values	Posterior values					
			$C_{min/max}$, $Y_{min/max}$, $Y = C_{HH} + (C_{GOV} + K + X - M)$, $C/POP_{min/max}$		$Y = C_{HH} + (C_{GOV} + K + X - M)$, $C/POP_{min/max}$		$Y = C_{HH}$, $C/POP_{min/max}$, $Y = R_{HH} + (R_{NFC\&FC} + R_{GOV} + Deprec - Tx_{prod-diff} + R_{ROW})$	
			option (1)		option (2)		option (3)	
			LP (β and γ in graph 2b)	Bayesian	LP (β and γ in graph 2b)	Bayesian	LP (ε in graph 2b)	Bayesian
	(1)	(2)	(3)	(4)	(7)	(8)	(11)	(12)
HH final consumption	$C_{HH, central}$	393 364	393 364	393 364	393 364	393 364	393 364	393 364
	$C_{HH, max}$	440 568	416 966	411 805	416 966	416 966	393 364	393 372
	$C_{HH, min}$	346 160	369 762	374 923	369 762	369 762	393 364	393 356
	$\% C_{HH\ interval}$	12.0	6.0	4.7	6.0	6.0	0.0	0.0
GDP	$Y_{Central}$	631 512	631 512	631 512	631 512	631 512	631 512	631 512
	Y_{max}	669 403	655 114	649 953	655 114	655 114	631 512	631 520
	Y_{min}	593 621	607 910	613 071	607 910	607 910	631 512	631 504
	% Y interval	6.0	3.7	2.9	3.7	3.7	0.0	0.0
Per capita HH consumption	$C_{HH}/POP_{Central}$	23.732	23.732	23.732	23.732	23.732	23.732	23.732
	C_{HHmax}/POP	25.156	25.156	24.845	25.156	25.156	23.732	23.733
	C_{HHmin}/POP	22.308	22.308	22.620	22.308	22.308	23.732	23.732
	$\% C_{HH}/POP$ interval	6.0	6.0	4.7	6.0	6.0	0.0	0.0

Source: Author's calculations.

Until now, the prior reliabilities have been defined in terms of fixed intervals. Though similar, it is not the approach followed in this paper. In the Bayesian approach, it will be assumed that the reliabilities of data and ratio values have a normal distribution and that the percentages should be interpreted as the percent value of a variation coefficient. As in the LP approach, %'s apply to data and ratio values and may differ between different data items and between different ratio values. The variation coefficient (σ_v) is defined as the ratio of one standard deviation (σ) divided by a value of x ($\sigma_v = \sigma/x$). Thus one standard deviation $\sigma = \sigma_v * x$. For example, if C_{HH} is the variable x, the reliability of C_{HH} (*F72* in table 1a) is 12 % (L) and the value is 393 364, as in table 1a, one standard deviation is $\sigma = \sigma_v * x = 0.12*393\,364 = 47\,204$. The standard deviation is a parameter in a normal distribution. The probability of x values is determined by the function: $P(x) = \frac{1}{\sqrt{(2\pi\sigma)}} e^{-(x-\bar{x})^2/2\sigma^2}$.

The values of variables (x) are between $-\infty$ and $+\infty$, $\bar{x}$ is the prior value of C_{HH}, and σ is the standard deviation. The functional relation of the normal distribution of C_{HH} is presented graphically in figure 1. The value of $C_{HH} = 393\,364$ is located in the centre of the graph and the $-\sigma$ and $+\sigma$ interval is also indicated in the graph. The shape of the graph is determined by the value of $\sigma = 47\,204$. If the variation coefficient were higher and thus the standard deviation too, the curve would be wider and the distance between $-\sigma$ and $+\sigma$ would be larger. If smaller, the distance between $-\sigma$ and $+\sigma$ would be smaller.

Figure 1: Normal distribution of values of C_{HH} (x) with mean ($\bar{x}$) and standard deviation (σ)



Source: Author's calculations.

As in the LP approach described above, prior intervals of data and ratio values determine the posterior intervals. This is not different in the Bayesian estimation approach. The consequences, however, differ somewhat between the LP and Bayesian estimation approaches. In the LP approach, posterior values (x) are always located within their prior fixed intervals, when other conditions are taken into account, as was shown in the example above. Their posterior value can never be higher than the prior maximum or lower than their prior minimum. In the Bayesian approach, however, the value (x) is located in a normal distribution, which varies between $-\infty$ and $+\infty$. As a consequence, the posterior value (x) may be located outside the one–σ interval. As in the LP approach, the posterior value of σ may never be higher than the prior value. But as the posterior value of x may be higher or lower within the $-\infty$ and $+\infty$ interval of the normal distribution, the posterior value of the variation coefficient (σ_v) could be higher, in principle, than the prior value. In practice it is found that this occurs very seldom, though.

The posterior value of X in the Bayesian method is determined on the basis of an optimization procedure, which may be described in simplified format as follows. If $\bar{X}_i$ is the prior value and X_i the posterior value of a variable and $\bar{\sigma}_i$ is the prior standard deviation of $\bar{X}_i$, the Bayesian estimation method minimizes the following weighted least square expression $\sum_i \frac{1}{\bar{\sigma}_i^2} * (X_i - \bar{X}_i)^2$.

The minimization is done under the conditions of a set of identities $A*X$, in which X is the vector of variables and A is the matrix of coefficients (0, 1 or – 1) of each variable in the identities. Ratios are linearized in realistic values of the nominator, so that they can also be included in the identities and in the expression that should be minimized.

Through the use of prior *weights* $1/\bar{\sigma}_i^2$, the minimization procedure allows larger differences between posterior and prior values of X_i, with larger values of $\bar{\sigma}_i$ than those with smaller values. More accurate data and ratio values are less changed during the optimization than values that are less precise according to prior information. Posteriors are derived of X_i, the standard deviation σ_i and of the coefficient of variation.

As in the case of the LP approach described above, not only posterior values and standard deviations of variables with data are estimated, but also of variables for which no prior information is available. The latter estimates are made with help of the ratios and identities that are used in the estimation process.

It was also shown above in the LP approach how prior intervals influence and generally reduce the posterior reliability intervals as compared to the prior intervals. The same holds in the Bayesian approach, when comparing prior and posterior values of σ. This means that the use of frameworks generally improves the reliability of the estimates, as many restrictions are imposed and as data and ratio values with high prior reliability improve estimates with those of lower reliability. Our experience is that the more conditions are imposed, the more will be reduced the posterior intervals.

Based on the example above-presented, more must be said about the relation between the number of variables and the number of information items (identities, ratios and data):

1. The number of identities (n_i) can never be higher than the number of variables for which no data are available. Thus, if the total number of variables is n_x and the number of variables with data is n_d, $(n_i) \leq (n_x - n_d)$. If the number of identities were higher, there would be two possibilities. One is that some identities are dependent identities, which can be derived from linear combinations of other identities. As the dependent identities do not contribute to the conditions of the framework, they are not counted as information items. Dependent and independent identities are identified in the table with help of light and dark coloring (as indicated in the legend to the table). The other possibility is that the identities are conflicting with each other, in which case only posterior values could be obtained by adapting/correcting or removing the conflicting identities.

2. If the number of identities is equal to the number of variables less data after removing the conflicting and/or dependent identities, the values of all variables could be estimated as precise numbers, provided no ratios and reliabilities were used. In the case of table 1, the maximum of n_i is $(n_x - n_d) = 58 - 27 = 31$. The actual number of identities are lower, i.e. $n_i = 23$. The difference between the number of variables minus the number of data and identities $(n_x - n_d - n_i) = 58 - 27 - 23 = 8$ is compensated for by the number of ratios (n_r) that are defined in the framework. In the table, the number of ratios ($n_r = 62$) are much larger and this contributes to the number of data checks (referred to above in section 2.1) that is available in the framework of table 1.
3. There is no limit to the number of ratios as no ratio value is fixed, but all are subject to standard deviations in a normal distribution. In the framework of table 1, the number of ratios is 62, which is even larger than the difference between the number of variables to be estimated (58) and the total number of information items (112), including ratios, that are available.
4. If the number of information items are less than the number of variables, it is not possible to arrive at posterior values of variables and ratios. It is only possible to arrive at posterior estimates if the number of information items is equal or larger than the number of variables, i.e. $n_d + n_r + n_i \geq n_x$.
5. In general terms, the larger the number of information items, the lower the posterior variation coefficients of variables and ratios are and thus, the more precise the posterior estimates. This may be stated more correctly as follows: if an identity is added to the existing information items and this identity only holds for ranges of values of the variables, that are more limited than their posterior ranges based on the existing information items, such additional identity will decrease the variables' posterior coefficients of variation, and will thus contribute to more precise estimates. In reverse, if the additional identity holds for values of the corresponding variables outside their posterior ranges based on existing information items it does not contribute to the precision of posterior estimates.
6. If additional data or ratios are added to the existing set of information items, it decreases the posterior coefficients of variation of corresponding variables and ratios only if the prior coefficients of variation of the additional data or ratios are smaller than the posterior ones that were derived without the additional data or ratio values.

4. Use of frameworks in compilation

In table 3 two periods, 2010 and 2013, are selected for inclusion in the present paper, in order to quantify the impact on posterior values of variables and ratios of a Bayesian estimation approach to frameworks.

The table includes three groups of columns, i.e. one for 2010 for which a complete data set is available (2010C), another for 2010 in which a partial data set has been assumed (2010P), and a third group of columns for 2013 with also a partial data set (2013P). The table only includes a selected set of variables and ratios from the framework of table 1. Information on all variables and ratios of the framework is available, but has not been presented in order not to overcrowd the table. In the first two columns of the table the cell references of the selected variables and ratios in the framework of table 1 are indicated. Prior coefficients of variation (σ_v) for variables and ratios are listed in the next two columns when available. The percentages presented refer to the prior reliability indicators set by national accountants, i.e. P (poor, σ_v = 24 %), L (low 12 %), M (medium 6 %), H (high 3 %), S (superior 1 %), F1 (semi-fixed.01 %) and F (fixed 0 %). Prior values of variables and ratios for 2010 are presented in the next two columns. They apply both to the framework with a complete data set (2010C) and also to the partial data sets for 2010 (2010P) and 2013 (2013P). For variables without data, tentative values that are either based on actual data available from Statistics Netherlands, or on assumed values derived with help of selected ratios and identities are included. Those values, together with the corresponding coefficients of variation, however, are not used in the Bayesian estimation procedure.

The next two columns for each data set refer to posterior estimates of variables and ratios, derived with help of the Bayesian estimation methodology. Thereafter the percentage difference between posterior and prior values is calculated for each variable and ratio and for each of the three frameworks 2010C, 2010P and 2013P. In the last two columns for each of the three data sets, the posterior values of the variation coefficients have been presented after application of the Bayesian estimation methodology.

Hereafter follows an assessment of the Bayesian (posterior) estimates and their reliability for each of the three data sets. Only a selection of the most important variables and analytical indicators is included in the table. Extreme values have been coloured in the table.

1. The 2010C data set has been included merely to check the proper functioning of the Bayesian estimation. It includes a complete data set for all variables for 2010 and also for all ratios. A major obstacle in the use of the Bayesian estimation, which is checked with this data set, is whether all identities that should hold, are included. If some are omitted, nonsensical results may be obtained, and further checks are needed in order to identify the identities that were omitted. For the set of data, ratios and identities presented for 2010C this check has been carried out and corrections have already been made. As a result, posterior estimates of variables and ratios are very close to the prior values. This can be seen in the two columns (L and M) presenting the % difference between prior and posterior values of variables and ratios. All differences are 0.0 %, meaning that the Bayesian estimation is working properly and thus can be applied to the partial data sets of 2010P and 2013P.

   What is important, however, is the significant reduction that is observed in the posterior coefficients of variation of most variables and ratios, as compared with the prior values (columns N and O). This result shows how estimation in frameworks improves the reliability of the posterior results after using the many checks of ratios and identities. The quantification of this improvement can only be observed with help of the Bayesian estimation methodology as it derives values of posterior coefficients of variation, which cannot be obtained through the conventional compilation methodology used in national accounting. In particular, considerable reductions between prior and posterior values of variation coefficients could be observed for the variables and ratios of final consumption and capital formation from resp. 12 % and 24 % to 0.0 % and 0.1 %. The considerable reduction of the variation coefficients is partly caused by the inclusion of variables with fixed prior values (F) in the GOV sector. One exception should be noted, and that is the growth rate of output, the posterior coefficient of variation of which remains high (17.7 %). Past experience shows that it is more difficult to estimate growth rates over time. In the present case, this applies to output growth and, in the case of 2010P and 21013P, also to investment growth. It would also apply to GDP growth, if it were not fixed as was done for all three cases of table 4.

2. The data set for 2010P was included in order to check the impact of having only partial data for this year. The partial data refer to those that are generally available early on in the compila-

tion process and therefore are considered as good priors that could be used in the compilation before comprehensive data become available. From the textbox on information items in table 1 it can be seen that of the 58 variables in the framework, only 27 have data, but this lack of data is compensated for by 62 ratios and 23 independent identities. The information factor (number of information items/the number of variables to be estimated) is thus quite high, i.e. 1.93. From table 3 (columns T and U), it can be seen that for the data set the difference between posterior and prior values of variables and ratios is still very low (though slightly higher than in the case of 2010C), and that the posterior coefficients of variation (columns V and W) are considerably reduced (though a bit less than in the case of 2010C). The coefficients of variation of final consumption and capital formation, and property income and current transfers are less reduced as was the case for 2010C. Coefficients of variation of the balancing items and their corresponding ratios of net lending of HH's and NPI's, and NFC's and FC's on the other hand even increased (considerably) as compared to the prior values of the coefficients of variation. The latter can be justified as the prior values and coefficients of variation of these balancing items were not taken into account in the Bayesian estimation procedure. The closeness of the prior and posterior values for 2010P is understandable, as the data that are assumed to be available for 2010 are the same and the structure of the framework, reflected in identities, prior values of ratios and reliabilities is also the same.

3. The last case concerns 2013P, which is a year with tentative data that may be revised in the future. Some data are not available (employment data) and data for other variables will be assumed to be not available ('missing data'). Identities and ratios will then be used to estimate the variables with missing data through the Bayesian estimation approach. The ratio values used in this approach will be based on 2010 data, which simulates an approach that is commonly used in national accounting practices. As it is a partial data set with 2010 structures, it is obvious that posterior values (columns AJ and AK) deviate more from the prior values than for 2010P. The differences, as compared to 2010P values, though higher for most variables are not significantly higher. The posterior coefficients of variation are still significantly reduced when compared to the prior values, except for some ratios, i.e. the income tax rate for FC's and NFC's, and the growth rates of output and investments.

In particular, it should be noted that final consumption does not significantly deviate from its prior value, but that capital formation deviates much more. Also property income and current transfers show significant deviations. Balancing items of disposable income of NFC's and FC's, and net lending of HH's and NPI's and NFC's and FC's also show significant deviations from the prior values.

The reason is that those variables are derived from other variables and therefore accumulate the deviations of each of the component variables. The same applies to ratios and coefficients, which are derived from two other variables.

What is surprising is that the reductions in the variation coefficient of 2013P are very similar to those of 2010P, except for some transactions that were already noted for 2010P. One further deviation that is observed for 2013P is that the income and wealth tax rates for NFC's and FC's deviate considerably from the prior value whilst the posterior coefficient of variation remains high. The reason behind the latter could lie in the insufficient checks available for these ratios in the framework.

Table 3: Assessment of estimates resulting from the Bayesian estimation methodology

		Variables	Ratios	Prior reliability (coeff. of variation)		2010C						
						Prior value		Posterior estimate		% Difference from prior value		
Extreme values				Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	
Output		C3		1.0		1 241 940		1 241 969		0.0		4
Output, previous year		H3		0.0		1 201 256		1 201 256		0.0		5
GDP		C5		0.0		631 512		631 512		0.0		6
GDP, previous year		H5		0.0		617 650		617 650		0.0		7
Expenditures distribution GDP	Final consumption (ratio y_c)	C16	K16	12.0	24.0	449 742	71.2	449 742	71.2	0.0	0.0	8
	Gross capital formation (ratio y_{inv})	C18	K18	12.0	24.0	128 957	20.4	128 957	20.4	0.0	0.0	9
	Gross capital formation (ratio y_{inv}), previous year	H18	K19	0.0		127 772	20.7	127 772	20.7	0.0	0.0	10
	Consumption of fixed capital	C9		6.0		106 982		106 982		0.0		11
	Exports (ratio y_x)	G4	K4	1.0	24.0	454 398	72.0	454 399	72.0	0.0	0.0	12
	Imports (ratio y_m)	G3	K3	1.0	24.0	– 401 585	– 63.6	– 401 585	– 63.6	0.0	0.0	13
Instruments distributing GDP to disposable income	Labour income	C6	J6	3.0	24.0	338 844	53.7	338 844	53.7	0.0	0.0	14
	Product and production taxes less subsidies	C7	J7	0.0	24.0	62 036	9.8	62 036	9.8	0.0	0.0	15
	Taxes on income and wealth	E12	J12	0.0	24.0	70 134	11.1	70 134	11.1	0.0	0.0	16
	Property income, receipts less payments	C11	J11	12.0	24.0	41 129	6.5	41 129	6.5	0.0	0.0	17
	Social transfers	E13	J13	0.0	24.0	93 769	14.8	93 769	14.8	0.0	0.0	18
	Other transfers	C14	J14	12.0	24.0	29 034	4.6	29 034	4.6	0.0	0.0	19
Distribution of disposable income by sectors	NFC' and FC's	D15	D53	6.0	6.0	78 838	15.2	78 838	15.2	0.0	0.0	20
	GOV	E15	E53	0.0	6.0	29 339	5.7	29 339	5.7	0.0	0.0	21
	HH's and NPI's	F15	F53	12.0	6.0	408 858	79.1	408 858	79.1	0.0	0.0	22
Distribution of Net Lending (+) & borrowing (–) by sectors	NFC' and FC's	D19	D55	1.0	1.0	69 069	89.5	69 069	89.5	0.0	0.0	23
	GOV	E19	E55	0.0	1.0	– 31 866	– 41.3	– 31 866	– 41.3	0.0	0.0	24
	HH's and NPI's	F19	F55	1.0	1.0	8 115	10.5	8 115	10.5	0.0	0.0	25
	ROW	G19	G55	3.0	1.0	– 45 318	– 58.7	– 45 318	– 58.7	0.0	0.0	26
Employment		C21	C39	1.0	3.0	7 391	44.6	7 391	44.6	0.0	0.0	27
Output ratios	i-o coefficient (io)		C30		3.0		50.8		50.8		0.0	28
	Output as % of total supply		C40		12.0		75.6		75.6		0.0	29
	Exports as % of Output		G28		3.0		36.6		36.6		0.0	30
A	BC	D	E	F	G	H	I	J	K	L	M	

Source: Statistics Netherlands

Table 3 (Cont.)

2010C		2010P				2013P, with 2010 structures								
Posterior reliability (coeff. of variation)		% Difference from prior value		Posterior reliability (coeff. of variation)		Prior value		Posterior estimate		% Difference from prior value		Posterior reliability (coeff. of variation)		
Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	
0.6		0.0		0.9		1 299 292		1 306 440		0.6		0.9		**4**
0.0		0.0		0.0		1 305 062		1 305 062		0.0		0.0		**5**
0.0		0.0		0.0		642 851		642 851		0.0		0.0		**6**
0.0		0.0		0.0		640 644		640 644		0.0		0.0		**7**
0.1	0.1	– 0.1	– 0.1	1.3	1.3	458 869	71.4	462 479	71.9	0.8	0.8	1.1	1.1	**8**
0.0	0.0	– 0.1	– 0.1	4.6	4.6	117 572	18.3	139 471	21.7	18.6	18.6	3.6	3.6	**9**
0.0		0.0	0.0	0.0		124 146		124 146		0.0		0.0		**10**
1.9		– 0.2		6.6		108 625		87 128		– 19.8		6.8		**11**
0.6	0.6	0.1	0.1	0.7	0.7	533 186	82.9	516 208	80.3	– 3.2	– 3.2	0.7	0.6	**12**
0.7	0.7	– 0.1	– 0.1	0.7	0.7	– 466 776	– 72.6	– 475 307	– 73.9	1.8	1.8	0.7	0.7	**13**
0.5	0.5	– 0.1	– 0.1	2.2	2.2	354 081	55.1	355 889	55.4	0.5	0.5	1.5	1.5	**14**
0.0	0.0	1.2	1.2	0.1	0.1	64 990	10.1	66 337	10.3	2.1	2.1	0.1	0.1	**15**
0.0	0.0	0.0	0.0	0.0	0.0	65 873	10.2	65 873	10.2	0.0	0.0	0.0	0.0	**16**
3.8	3.8	0.0	0.0	5.2	5.2	53 588	8.3	45 276	7.0	– 15.5	– 15.5	4.8	4.8	**17**
0.0	0.0	0.0	0.0	0.0	0.0	89 001	13.8	89 001	13.8	0.0	0.0	0.0	0.0	**18**
2.9	2.9	0.0	0.0	4.0	4.0	35 785	5.6	29 487	4.6	– 17.6	– 17.6	3.7	3.7	**19**
2.2	1.8	– 0.1	0.0	6.3	6.0	55 337	10.6	81 164	15.0	46.7	41.1	4.1	3.8	**20**
0.0	0.4	0.0	0.1	0.0	1.4	38 634	7.4	38 634	7.1	0.0	– 3.8	0.0	0.9	**21**
0.2	0.3	– 0.1	0.0	1.7	1.1	427 628	82.0	422 287	77.9	– 1.2	– 5.0	1.2	0.7	**22**
0.5	0.1	– 0.2	– 0.2	11.2	11.2	46 324	67.7	37 488	89.5	– 19.1	32.2	1.1	0.1	**23**
0.0	0.5	0.0	0.0	0.0	1.2	– 14 629	– 21.4	– 14 629	– 34.9	0.0	63.3	0.0	0.9	**24**
0.8	0.7	1.5	1.5	94.1	94.0	22 088	32.3	4 403	10.5	– 80.1	– 67.4	2.0	1.0	**25**
0.9	0.4	0.0	0.0	2.1	0.9	– 53 783	– 78.6	– 27 262	– 65.1	– 49.3	– 17.2	1.7	0.4	**26**
0.9	0.9	0.0	0.0	2.9	2.9	7 387	44.0	7 498	44.7	1.5	1.5	2.9	2.9	**27**
	0.6		0.0		0.9		49.5		49.2		– 0.5		0.9	**28**
	0.2		0.0		0.3		73.6		73.3		– 0.3		0.3	**29**
	0.9		0.1		1.1		41.0		39.5		– 3.7		1.0	**30**
N	**O**	**T**	**U**	**V**	**W**	**AF**	**AG**	**AH**	**AI**	**AJ**	**AK**	**AL**	**AM**	

Table 3 (Cont.)

			Variables	Ratios	Prior reliability (coeff. of variation)		2010C						
							Prior value		Posterior estimate		% Difference from prior value		
Extreme values					Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	
GOV taxes, consumption, investments and deficits	Income and Wealth Tax rates as % of disposable income	NFC's & FC's		D61		12.0		17.3		17.3		0.0	31
		HH's & NPI's		F61		12.0		13.4		13.4		0.0	32
	Product and Production Tax rates as % of GOV tax revenues			E35		1.0		46.5		46.5		0.0	33
	Income taxes as % of total GOV tax revenues			E36		1.0		53.5		53.5		0.0	34
	GOV consumption as % of total final consumption			E52		12.0		12.5		12.5		0.0	35
	GOV capital formation as % of total capital formation			E42		12.0		19.7		19.7		0.0	36
	GOV deficit as % of GDP (p_x)			E64		24.0		5.0		5.0		0.0	37
Per capita and per worker values	Labour income per worker			C38		12.0		45 845		45 845		0.0	38
	Final consumption per individual			F37		12.0		23 732		23 732		0.0	39
	HH propensity to consume			F29		3.0		96.2		0.9		0.0	40
Rate of return and growth rates	Output/capital ratio (p_{inv})			C65		24.0		10.8		10.8		0.0	41
	Growth rate of GDP (Δy)			C66		5.0		2.2		2.2		0.0	42
	Growth rate of output (Δp)			C67		5.0		3.4		3.4		0.1	43
	Growth rate of investments (Δ_{inv})			C68		24.0		0.9		0.9		0.0	44
A	BC		D	E	F	G	H	I	J	K	L	M	

Source: Statistics Netherlands

In the above the emphasis is on the estimation of variables. However, the Bayesian estimation methodology also estimates the values of analytical ratios or indicators. As in the case of values for variables, there are also differences between priors and posteriors for ratio values. The tendencies for posterior values of ratios to differ from prior values are similar to those of variables, as can be seen from the lower part of the table.

As a consequence of using the Bayesian estimation approach, the characteristics of analytical ratios in frameworks that are estimated with help of the Bayesian estimation methodology are different from those based on the conventional methodology used by national accountants. In the conventional methodology, ratios are generally fixed either on the basis of external information or more frequently, through the ratio values of a previous base year. In the course of the compilation and using the subjective considerations of national accountants, those ratios may be somewhat adjusted, in case they are inconsistent with other data in the framework. Basically the ratios are fixed by national accountants and, as a consequence, cannot be used any more effectively by analysts. In the Bayesian approach this is different. Values of ratios, either presumed, based on previous year's values or external information, are introduced with a prior reliability. In other words, the external information is not accepted as such, but is taken into account (used as priors), when making estimates. Given other information on data and

Table 3 (Cont.)

2010C		2010P				2013P, with 2010 structures								
Posterior reliability (coeff. of variation)		% Difference from prior value		Posterior reliability (coeff. of variation)		Prior value		Posterior estimate		% Difference from prior value		Posterior reliability (coeff. of variation)		
Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	Variables	% Ratios	
	2.2		0.0		8.9		23.9		16.6		– 30.7		10.1	**31**
	0.2		0.1		2.6		11.9		12.0		0.8		2.3	**32**
	0.0		0.0		0.0		49.1		49.1		0.0		0.0	**33**
	0.0		0.0		0.0		50.9		50.9		0.0		0.0	**34**
	0.1		0.1		1.3		12.2		12.1		– 0.8		1.1	**35**
	0.0		0.1		4.6		16.0		13.5		– 15.7		6.3	**36**
	0.0		0.0		0.0		2.3		2.3		0.0		0.0	**37**
	1.1		– 0.1		3.6		47 933		47 466		– 1.0		3.1	**38**
	0.2		– 0.2		1.5		24 001		24 217		0.9		1.2	**39**
	0.1		0.0		2.0		94.2		96.2		2.2		0.2	**40**
	0.0		0.0		0.0		1.8		1.8		0.0		0.0	**41**
	0.0		0.0		0.0		0.3		0.3		0.0		0.0	**42**
	17.7		0.2		27.5		– 0.4		0.1		– 123.9		850.8	**43**
	5.0		– 8.0		548.5		– 5.3		12.3		– 333.1		32.6	**44**
N	**O**	**T**	**U**	**V**	**W**	**AF**	**AG**	**AH**	**AI**	**AJ**	**AK**	**AL**	**AM**	

Source: Statistics Netherlands

ratio values, all with their own prior reliabilities, prior values of the ratios may be adjusted. The final outcome would be ratio values that may be based on prior information, but could still be used in analysis. This is another important advantage of the Bayesian estimation approach, in addition to the ones mentioned earlier, which is not available in the traditional manual compilation approaches, in which prior values of ratios are fixed and thefore cannot be used in analysis.

5. Analysis of framework estimates after Bayesian compilation

The ratios or analytical indicators used in the framework are an integral part of and define the analysis of this framework. The ones included in table 1 were identified when designing the framework. They were all used in the compilation, as described in the previous section. Only a selection of the analytical indicators of table 1 was included when assessing the results of the compilation in table 3 above, but many others are available. The same indicators are presented in table 4, which deals with the analysis of the compilation estimates and mainly focuses on analytical indicators. The scope and orientation of table 4 is different from that of table 3.

Table 4: Analysis of Bayesian estimates

		Ratios	2010C Posterior estimate % Ratios	2013C Posterior estimate % Ratios
(1) i-o coefficient, rate of return, growth rates	i-o coefficient (io1)	C30	50.8	49.5
	i-o coefficient, previous year (io0)	H5/H3	51.4	49.1
	Output as % of total supply (pm1)	C40	75.6	73.6
	Output/capital ratio (ypinv)	C65	10.8	1.8
	Growth rate of GDP (Δy)	C66	2.2	0.3
	Growth rate of output (Δp)	C67	3.4	– 0.4
	Growth rate of investments (Δinv)	C68	0.9	– 5.3
(2) Shares of GDP and output	Final consumption, share of GDP (yc1)	K16	71.2	71.4
	Gross capital formation, share of GDP (yinv1)	K18	20.4	18.3
	Gross capital formation, share of GDP, previous year (yinv0)	K19	20.7	19.4
	Exports, share of GDP (yx1)	K4	72.0	82.9
	Exports as % of Output (px1)	G28	36.6	41.0
	Imports, share of GDP (ym1)	K3	– 63.6	– 72.6
(3) Shares of instruments distributing GDP to disposable income	Labour income	J6	53.7	55.1
	Product and production taxes less subsidies	J7	9.8	10.1
	Taxes on income and wealth	J12	11.1	10.2
	Property income, receipts less payments	J11	6.5	8.3
	Social transfers	J13	14.8	13.8
	Other current and capital transfers	J14	4.6	5.6
(4) Distribution of disposable income by sectors	NFC' and FC's	D53	15.2	10.6
	GOV	E53	5.7	7.4
	HH's and NPI's	F53	79.1	82.0
(5) GOV taxes, consumption, investments and deficits	Income and wealth tax rates as % of disposable income — NFC's & FC's	D61	17.3	23.9
	Income and wealth tax rates as % of disposable income — HH's & NPI's	F61	13.4	11.9
	Product and production tax rates as % of GOV tax revenues	E35	46.5	49.1
	Income taxes as % of total GOV tax revenues	E36	53.5	50.9
	GOV consumption as % of total final consumption	E52	12.5	12.2
	GOV capital formation as % of total capital formation	E42	19.7	16.0
(6) Distribution of Net Lending (+) & borrowing (–) by sectors	NFC' and FC's	D55	89.5	67.7
	GOV	E55	– 41.3	– 21.4
	GOV deficit as % of GDP	E64	5.0	2.3
	HH's and NPI's	F55	10.5	32.3
	ROW	G55	– 58.7	– 78.6
(7) Per capita and per worker values, share of employment in population	Share of employment in population	C39	44.6	44.0
	Labour income per worker	C38	45 845	47 933
	Final consumption per individual	F37	23 732	24 001
	HH propensity to consume	F29	96.2	94.2

Source: Statistics Netherlands

It includes only posterior Bayesian estimates of indicators of the 2010C and 2013C frameworks, for which complete data sets were available and made consistent through the Bayesian estimation approach ([8]). It does not include any information on reliabilities. Information on priors and reliabilities is available, however. The emphasis of the table is on the analysis of analytical indicators, which are represented by the ratios. In the examples of analysis presented below it is shown how indicators could be used jointly in different types of analysis and how they interact.

The indicators are grouped together in 7 analytical segments, and are used jointly to carry out comprehensive analyses. Comparisons of the values of the indicators between 2010C and 2013C are also shown in table 4. The indicators (and in one case variables *H5/H3*) are referred to in the first column by the cells in which they are presented in table 1. Symbols other than the cell references are used below and in the table, to simplify the presentations.

Below are presented five examples of joint analyses, which illustrate the interaction between the indicators. The indicators used here are binary ratios between two variables. All variables used in defining the indicators are included in the framework. This is an essential feature of the framework which would also apply, if not only binary, but also composite indicators were used. An example of this is the composite Human Development Indicator (HDI index), developed by UNDP, referred to above. In its original format, this indicator includes three variables, i.e. GNI (Gross National Income per capita) life expectancy at birth and education (years of schooling). Yet, it has not been included in the present framework, as life expectancy and education are not incorporated in the framework for the time being. This could be done, however, by expanding the number of variables in the framework.

Table 4a: i-o, growth, output-capital ratio, shares of GDP, segments (1) and (2) of table 4
(%)

		2010C	2013C	Assumed
io_1	C30	50.8	49.5	50.0
	H5/H3	51.4	49.1	49.1
pm_1	C40	75.6	73.6	80.0
y_{inv1}	C65	10.8	1.8	11.0
Δy_1	C66	2.2	0.3	2.4
Δp_1	C67	3.4	– 0.4	0.5
	K16	71.2	71.4	58.0
y_{inv1}		20.4	18.3	22.0
y_{in0}	K19	20.7	19.4	19.4
	K4	72.0	82.9	70.0
y_{x1}	G28	36.6	41.0	35.0
	K3	– 63.6	– 72.6	– 50.0
$y_x – y_m$	K4– K3	8.4	10.3	20.0

Source: Statistics Netherlands

([8]) The estimates resulting from the Bayesian estimation of 2010C and also 2013C used in table 4 may not be exactly equal to the Statistics Netherlands' data set. One reason is that in making the estimates, the figures for the 'Difference VAT imputed and received by GOV' have been assumed to be equal to zero. The other reason is that prior reliabilities assigned to variables and ratio values are different from those implicit in the Statistics Netherlands' data, and this may have a small impact on posterior Bayesian estimates, even if the data set is complete. Posterior Bayesian estimates have been used in the analysis of table 4 and not the Statistics Netherlands' figures.

Here follow a few examples of the analytical indicators and their interactions:

1. The first example deals with analytical indicators described in segments (1) and (2) of table 4: GDP growth, i-o coefficients and the GDP expenditure distribution. It is shown how (i) GDP growth ($\Delta y_1 = C66$), (ii) the distribution of GDP by expenditures ($y_{inv1} = K18$, $y_{c1} = K16$, $y_{x1} = K4$ and $y_{m1} = K3$) and (iii) the import-export balance as percent of GDP ($y_{x1} - y_{m1}$) are affected by the values of five other indicators, i.e. changes in the (iv) i-o coefficient ($io_1 = Y_1 / P_1 = C30$), changes in the (v) export-output ratio ($P_x = X_1 / P_1 = G28$) and changes in the (vi) composition of supply between output and imports ($P_{m_1} = P_1 / (P_1 + M_1) = C40$), (vii) the incremental output-capital ratio yinv1 $= \Delta y_1 / y_{inv_0} = C65$, and changes in the (viii) investment share of GDP $y_{inv_0} = inv0 / Y_0 = K19$. ($Y$) refers to GDP, ($P$) to output, ($Inv$) refers to investments, (X) to exports and (M) to imports; the subscripts 0 and 1 relate to the previous and present year respectively. The first groups of ratios (i-iii) may be called endogenous ratios and the second group of five ratios (iv-viii) are exogenous ratios. The values of the endogenous ratios are dependent on the values of the exogenous ratios. The exogenous ratios are coloured in table 4a.

   The incremental output-capital ratio used here is the inverse of the incremental capital-output ratio (ICOR), as incorporated in the Harrod-Domar growth model (see Domar (1946)) and quantified by Kuznets (1960) over time and for several countries. These indicators could be extended with another indicator developed by Piketty (2014) called the Rate of Return on Capital (Piketty, page 52). By adding this indicator to the relation between GDP growth and investments, the effect of GDP growth to distribution between labour and capital income can be explained. This could well be used in the present framework, as all relevant variables, including property income, are part of the framework. However, as property income has conceptual characteristics that are not compatible with those used in the Piketty indicator, it would require some changes in the concepts of the framework in order to incorporate the Piketty indicator in a future framework version.

   The identities used to establish the relations between the indicators above mentioned include the shares of expenditures in GDP that are equal to 1 and the share of the external balance, i.e. the difference between the shares of exports and imports, as percent of GDP:

   $Y_{c1} + Y_{inv1} + Y_{x1} - y_{m1} = 1$

   or $K16 + K18 + K4 - K3 = 1$

   and $Y_{extbalance} = y_x - y_m$

   or $Y_{extbalance} = K4 - K3$

   The conceptual and value relations are presented in table 4a, which is an extract from the segments (1) and (2) of table 4. The figures used below are those of the 2013C column. Similar calculations can be made for 2010C.

   GDP growth (Δy) is derived by multiplying the output-capital ratio ($ypinv_1$) by the share of investment in GDP in the previous period (Y_{inv0}). This relation is derived as follows:

$$ypinv_1 = \frac{Y_1 - Y_0}{Inv_0} = \frac{Y_0 * (\frac{Y_1}{Y_0} - 1)}{Inv_0} = \frac{\frac{Y_1}{Y_0} - 1}{\frac{Inv_0}{Y_0}}$$

The numerator of this expression is the growth rate of GDP and the denominator is the share of investment in GDP in period 0.

Thus $ypinv_1 = \Delta y_1 / y_{inv_0}$

and thus $\Delta y_1 = ypinv_1 * y_{inv0}$

or, in terms of cell references of the framework, *C66 = C65*K19*;

and in terms of 2013C values, 0.344 % = 1.778 %*19.4 % (see column 2013C of table 4a).

From this, the growth of output $(\Delta p_1 = (P_1 / P_0) - 1)$ can be derived with help of io-coefficients of the present period (*C30*) and the previous period (*H5/H3*) as follows:

$\Delta y_1 = (Y_1 / Y_0) - 1)$

Or, in terms of *P*, $\Delta p_1 = (P_1 / P_0)^* (io_1 / io_0) - 1$.

Substituting in this expression $(P_1 / P_0) = \Delta p_1 + 1$,

results in $\Delta y_1 = \frac{io_1}{io_0} * (\Delta p_1 - 1) - 1$

and thus, $p1 = \frac{\Delta y_1 + 1}{io_1/io_0} - 1$.

This could be rewritten in terms of cell references of the framework as

$$C67 = \frac{C66 + 1}{\frac{C30}{H5/H3}} - 1$$

and in 2013C values as $-$ 0.442 % = $\frac{0.344\ \% + 1}{49.5\ \%/49.1\ \%} - 1$.

Each of the expenditure shares of GDP could be expressed in terms of the three indicators mentioned. The export share in GDP is a function of the export share in output divided by the i-o-coefficient,

i.e. $P_x = \frac{X_1}{P_1} = \frac{X_1/Y_1}{P_1/Y_1} = y_{x_1} * io_1$,

or $y_{x_1} = \frac{p_{x1}}{io_1}$,

i.e. *K4 = G28/C30*

and in 2013C values: 82.9 % = 41.0 %/49.5 % [9].

The import share of GDP is a function of the import share and the i-o-coefficient,

$$i.e.\ y_{m1} = -\frac{M_1}{Y_1} = -\frac{M_1/P_1}{Y_1/P_1} = \frac{\frac{M_1}{P_1 + M_1} * \frac{P_1 + M_1}{P_1}}{io_1} = -(1/io_1)^* \frac{(1 - p_{m1})}{p_{m1}}.$$

Thus $y_{m1} = -\frac{1}{io_1} * \left(\frac{1}{p_{m1}} - 1\right)$

(9) In this and the next section, percent values are used in multiplications, divisions and other functional relationships. In the example presented here *82.9 % = 41.0 %/49.5 %* may be read as equivalent to *0.829 = 0.41/0.495*

Or $K3 = -\frac{1}{C30} * \left(\frac{1}{C40} - 1\right)$

(values in 2013C: – 72.6 % = 1/49.5 %*(1/73.6 % – 1))([10]).

The share of the trade deficit in GDP is equal to the difference between the export and the import share of GDP,

(5) i.e. $Y_{extbalance} = K4 - K3$

(values in 2013C: 82.9 % – 72.6 % = 10.3 %),

and the share of final consumption is the difference between 1 and the sum of the shares of investment, exports minus imports,

(6) i.e. $Y_{c1} = 1 - (y_{inv1} + Y_{x1} - Y_{m1})$ or $K16 = 1 - (K18+K4\text{-}K3)$

(values in 2013C: 71.4 % = 1 – (18.3 % + 82.9 % – 72.6 %))

The values of the exogenous indicators in 2010C are not much different from those in 2013C and as a result, the GDP shares do not differ very much between both years too. However, if we assume much more substantial changes in the indicator values of 2013C, the impacts of changes in the exogenous indicators may be clearer (see column 'assumed' in table 4a). For instance, it may be assumed that there is an increase in investment share of GDP to, say 22 %, an increase in the output share of supply to, say 80 %, an i-o-coefficient of 50 % that remains approximately the same, an increase in the output-capital ratio to 11 %, and a reduction of the export share to, say 35 %. Under these conditions, GDP growth will be 2.420 % (using expression (1), Δy_1 = 22.0 %*11.0 % = 2.420 %), instead of 0.345 % in 2013C. Output will grow, according to expression (2), by 0.554 %(=(2.42 % + 1)/(50.0 %/49.1 %)– 1). When using the expressions (3)-(6), the GDP expenditures shares can be calculated. The export share of GDP would decrease to 70.0 %(= 35.0 %/50.0 %), the import share to – 50.0 %(= 1/50.0 %*(1/80.0 % – 1), and the consumption share of GDP will reduce to 58.0 %(= 1 – (22.0 % + 70.0 % – 50.0 %). The trade balance as percent of GDP would increase to 20.0 %(= 70.0 % – 50.0 %).

2. The next two segments (3) and (4) of table 4a deal with the distribution of GDP to disposable income of sectors.

   The framework includes *six instruments that distribute GDP generated in production to disposable income of sectors*. In segment (3) they include five indicators, i.e. the percent of GDP distributed through labour income (*J6*), through production and income taxes (*J7 + J12*), through property income (*J11*), through social transfers (*J13*) and through other current and capital transfers (*J14*). The highest percentages are observed for labour income, incl. mixed income (2013C: 55.1 %), taxes (20.4 %) and social transfers (13.8 %). They are expressed as percent points of GDP, but could be approximately also expressed as percent points of gross disposable income, as the values of GDP and gross disposable income are close to each other: the value of GDP in 2013C (not shown in table 4) is 642 851 and the gross disposable income (also not shown in table 4) is 630 224 (i.e. net disposable income of NFC's and FC's, GOV, and HH's and NPI's plus consumption of fixed capital: i.e. (55 337 + 38 634 + 427 628) + 108 625 = 630 224).

   In 2013C, as a result of the use of transfer instruments, disposable income is distributed (see segment (4) in 2013C)to NFC's and FC's by 10.6 % (*D53*), to GOV by 7.4 % (*E53*) and to HH's and NPI's by 82.0 % (*F53*). The percent for HH's and NPI's is the highest, because labour income

([10]) Calculations are carried out in more decimals than shown in the paper. As a consequence, calculations with the figures presented in the text may result in figures that slightly deviate from the results presented in the text.

and social transfers mainly benefit HH's. As the sum of the percentages for these instruments, however, does not add up to the disposable income share of HH's, a large percent of taxes should be added to it and distributed to HH's (social transfers in kind) through GOV consumption on education, health and housing. As a consequence, the percent distribution of disposable income to GOV is relatively low as compared to the percentages of taxes received.

3. How the GOV influences the economy is reflected by the indicators in table 4, segment (4), together with indicators in segments (2) and (3). The values of the indicators used will be those of 2013C.

   Product and production taxes, and income and wealth taxes are 10.1 % + 10.2 % = 20.4 % (indicators *J7* + *J12*) of GDP. On the expenditure side of the GOV account, GOV consumption is 12.2 % (*E52*) of total final consumption and, as total final consumption is 71.4 % (*K16*) of GDP, GOV consumption is 12.2 %*71.4 % = 8.7 % of GDP. The same can be measured for GOV capital formation, which accounts for 16.0 % (*E42*) of total capital formation in the country and, as the latter is 18.3 % (*K18*) of GDP, GOV investments are 16 %*18.3 % = 2.9 % of GDP. GOV consumption and capital formation are two of the three main expenditure items of the GOV budget. The third component is social transfers: they amount 13.8 % (*J13*) of GDP. The total of these expenditure components are thus 25.5 % = 8.7 % + 2.9 % + 13.8 % of GDP, which is higher than the GDP share of taxes mentioned above. Although this only gives a partial picture of GOV contributions to the National Economy (GOV has other revenues and also other expenditures), it does give a fair indication of why the GOV accounts incur a deficit (see also point 4 below).

   Among the GOV tax revenues, 49.1 % (*E35*) are product and production taxes and 50.9 % (*E36*) are income and wealth taxes. With regard to income and wealth taxes, NFC's and FC's have an average tax rate of 23.9 % (*D61*) and HH's have an average tax rate of 11.9 % (*F61*). As HH and NPI disposable income is 82.0 % (*F53*) of total disposable income, and NFC and FC disposable income is 10.6 % (*D53*) of total disposable income, the income tax rates of two sectors can be converted to percent tax rates of total disposable income: for NFC's and FC's, and for HH's and NPI's they amount respectively 23.9 %*10.6 % = 2.5 % and 11.9 %*82.0 % = 9.8 %. As Gross Disposable Income and GDP are approximately equal, as mentioned above, the income tax rates to total disposable income could be converted to GDP. The income tax contributions of HH's are relatively low, as this sector is also charged a large part of product and production taxes, particularly those that are related to the largest component of product taxes, i.e. VAT. If we were to add the total product and production taxes to the contribution of HH's, their tax burden of GDP would increase to 9.8 % + 10.1 % (*J7*) = 19.9 %.

4. After deducting final consumption and capital formation from disposable income, net lending/borrowing results for each sector. When reviewing the distribution of net lending/borrowing, it may be observed that two sectors are lenders and two are borrowers. The two lending sectors are NFC's and FC's (*D55* = 67.7 %, mainly FC's), and HH's and NPI's (*F55* = 32.3 %, mainly HH's) and the borrowing sectors are GOV (*E55* = – 21.4 %) and ROW (*G55* = – 78.6 %). It should be noted that both for the lending and borrowing sectors, the distribution changes considerably between 2010C and 2013C, as these percentages depend heavily on changes from one year to the other for all transactions that result in net lending and borrowing figures.

5. Labour income per worker, employment as percent of population and final consumption per individual are interrelated analytically in the framework. The labour income per worker is 47 933. As employment is 44.7 % of the population, labour income per capita, i.e. available on average to each member of the population, would be 44.7 % of 47 933, which is 21 102.

This is related to the HH consumption per capita of 24 001. The fact that HH consumption is higher than labour income per capita is because HH's also receives property income as well as social and other transfers. This is taken into account in the framework when making the estimates.

The above examples illustrate the joint use of the indicators in analysis. However, in each of these cases, they are only partial analyses, which do not fully reflect the whole complex structure of the economy and how a large number of indicators would interact with each other. How such complex interrelationships would work out was shown above in section 3 on the use of Bayesian estimation in compilation and will also be illustrated in section 6, where projections have been made using all indicators jointly with exogenous changes in some of the variables.

6. Use of frameworks in projections

It will be shown here how the framework and the Bayesian methodology can also be used for projections to the near future, i.e. 2020. Three scenarios are used in the projections. In each of them exogenous variables and ratios that could be used as policy instruments are selected and alternative values that reflect alternative policies are assigned to them. The projected values of the exogenous variables are fixed without any possibility of change in the Bayesian estimation of the framework variables. It must be pointed out that the exogenous variables are not necessarily the ones for which priors were established in the compilation of section 3. The reason for it is that data that become available at an early stage of the compilation are not necessarily exogenous in projections. The number of exogenous variables and ratios is small, much smaller than the number of data for which prior information becomes available early on in the compilation. Thus, in the projections there will be much more reliance on the ratios and identities of the framework, in order to establish the values of other (endogenous) variables and ratios. Posterior values, coefficients of variation and average annual growth rates between 2013 (last year of available data) and 2020 of selected variables, and posterior values and coefficients of variation of selected indicators (ratios) are presented in table 5. The selected variables are grouped together into themes in column A of the table whilst the variables and ratios themselves together with the cell references in the framework of table 1, are presented in columns B and C. The cells with exogenous variables and ratios for the three scenarios are coloured in the table.

In order to arrive at projected values of exogenous variables and ratios, their values over time have been observed between 2009 and 2013. Values and growth rates have been presented in columns D to L of table 5, and average growth rates and average values of indicators have been presented in column M. The average values are used as reference in making assumptions about the values of exogenous variables and indicators in the three projection scenarios.

Using three options, the projected figures for 2020 resulting from the Bayesian estimation have been presented in three groups of columns (N-V) for each scenario. Each group of columns of projections include a column for posterior values in 2020 (cols. N, Q, T), coefficients of variation (cols. O, R, U) and columns for average annual growth rates of variables between 2013 and 2020 (cols. P, S, V). The cells for exogenous variables and ratios are coloured for each scenario. The exogenous values can be recognized, as their posterior coefficients of variation are 0.0 %, and as for each variable and ratio, they are equal to the prior coefficient of variation that was assumed to be equal to 0.0 %.

Starting from the prior 2013 data of variables and ratios, three projections are made for the year 2020. In these projections, projected values of exogenous variables are considered as prior data with fixed values. With help of the same identities and ratios highlighted in section 2.2, posterior values of the 'missing data' are estimated. The ratio values are based on the ones for 2013. *Two of the projections (scenarios B and C) simulate the ongoing discussion about a 'monetary' and 'economic' approach stimulating economic growth*. The projected values of exogenous variables will be derived by using assumed growth rates of selected variables such as GDP (in scenario C), Exports (in all three scenarios), GOV consumption and capital formation (scenarios A and B), and GOV social transfers (also in scenarios A and B).

Exogenous indicators include i-o coefficients (scenarios A and B) and investment as percent of GDP (in all three scenarios), representing productivity and technological progress, and GOV deficits as percent of GDP, which is a fiscal policy variable in the EU (in all three scenarios). Population is also an exogenous variable. In all three scenarios it is fixed, meaning that an annual average growth of 0.4 % is assumed (as in the recent past) and that *no account is taken of increased immigration*.

Exogenous variables and indicators have coefficients of variation that are equal to zero both for priors and posteriors. In order to arrive at a complete set of projected estimates of the framework, including variables and indicators for which no data are available, values of exogenous variables and indicators are used together with *conceptual identities* and prior values of endogenous ratios. The conceptual identities include among others (i) the identity between supply (output and imports) and use (intermediate and final consumption, capital formation and exports minus imports), (ii) the value added identity between output and intermediate consumption, (iii) identities between items (property income, transfers, etc.) for the National Economy and the sum of those by sectors. *Endogenous structural ratios* include among others (i) i-o coefficients, (ii) the coefficients of redistribution of GDP to disposable income through the instruments of labour income, property income, taxes, and social and other transfers; and also ratios, such as (iii) tax rates, (iv) labour income per worker, (v) coefficients of distribution of GDP between expenditures, etc. Endogenous indicators used in the projections have higher prior coefficients of variation, so that they can change in order to arrive at projected values that are consistent and satisfy all identities. As the number of exogenous data in all projections is small, posterior variation coefficients of estimates of endogenous variables are generally larger than was the case in the compilation (as discussed in section 4) and thus, projections are less precise than estimates of the compilation.

The present paper could not determine how the options would influence the creation of financial assets and liabilities (debts), as the present framework does NOT include details for financial assets and liabilities.

Table 5: Analysis of impacts (posterior estimates) of 'assumptions' (growth rates, other coefficients and values, based on prior values of variables measured in the period 2009–13 (and averages over time) of projection scenarios

			2009	2010		2011		2012		2013		
			Value	Value	Annual growth (%)	Value	Annual growth (%)	Value	Annual growth (%)	Value	Annual growth (%)	
GDP	GDP	C5	617 650	631 512	2.2	642 929	1.8	640 644	– 0.4	642 851	0.3	**5**
	I-O coefficient (%)	C30		50.8		49.5		49.1		49.5		**6**
	Output/capital ratio (%)	C65		10.8		8.9		– 1.7		1.8		**7**
Coefficients of Distribution GDP to Disposable Income (%)	Labour income (compensation of employees & mixed income)	J6		53.7		54.5		54.9		55.1		**8**
	Product and production taxes less subsidies	J7		9.8		9.5		9.5		10.1		**9**
	Property income, receipts less payments	J11		6.5		6.3		7.3		8.3		**10**
	Taxes on Income and Wealth	J12		11.1		10.7		10.2		10.2		**11**
	Social transfers (incl. social transfers in kind)	J13		14.8		14.2		14.0		13.8		**12**
	Other current & capital transfers	J14		4.6		4.6		5.1		5.6		**13**
Expenditures GDP	Final consumption	C16	443 586	449 742	1.4	456 097	1.4	458 274	0.5	458 869	0.1	**14**
	Final consumption GOV	E16	56 118	56 378	0.5	55 499	– 1.6	55 412	– 0.2	56 136	1.3	**15**
	Gross capital formation	C18	127 772	128 957	0.9	131 928	2.3	124 146	– 5.9	117 572	– 5.3	**16**
	Gross capital formation GOV	E18	26 456	25 415	– 3.9	24 157	– 4.9	22 976	– 4.9	18 787	– 18.2	**17**
	Investment as % of GDP	K18		20.4		20.5		19.4		18.3		**18**
	Exports	G4	394 714	454 398	15.1	497 347	9.5	525 559	5.7	533 186	1.5	**19**
	Imports	G3	– 348 422	– 401 585	15.3	– 442 443	10.2	– 467 335	5.6	– 466 776	– 0.1	**20**
Labour and Population	Labour income received by HH's	F6	337 502	334 411	– 0.9	345 939	3.4	346 569	0.2	349 486	0.8	**21**
	Employment	C21		7 391		7 392	0.0	7 387	– 0.1	7 387	0.0	**22**
	Labour income per worker (*1000 euro)	C38		45 845		47 441	3.5	47 573	0.3	47 933	0.8	**23**
	Population	F20		16 575		16 656	0.5	16 730	0.4	16 780	0.3	**24**
Property income received by residents		C11		41 129		40 687	– 1.1	47 053	15.6	53 588	13.9	**25**
Taxes	Taxes on production and products, received by GOV	E7	58 036	60 963	5.0	60 011	– 1.6	59 552	– 0.8	63 568	6.7	**26**
	Product and production taxes less subsidies as % of GDP	C32	9.4	9.8		9.5		9.5		10.1		**27**
	Taxes on income and wealth received by GOV	E12	67 693	70 134	3.6	68 629	– 2.1	65 633	– 4.4	65 873	0.4	**28**
	Taxes on income and wealth as % of Disposable income, HH's	F61		13.4		12.8		– 12.0		– 11.9		**29**
	Taxes on income and wealth as % of disposable income, NFC's and FC's	D61		– 17.3		– 17.6		– 16.9		– 23.9		**30**
Social transfers (incl. social transfers in kind)		E13	90 650	93 769	3.4	90 998	– 3.0	89 431	– 1.7	89 001	– 0.5	**31**
Net lending	NFC's & FC's	D19	54 671	69 069	26.3	61 269	– 11.3	65 885	7.5	46 324	– 29.7	**32**
	GOV	E19	– 33 718	– 31 866	– 5.5	– 27 835	– 12.6	– 25 330	– 9.0	– 14 629	– 42.2	**33**
	GOV deficits as % of GDP	E64		5.0		4.3		4.0		2.3		**34**
	HH's & NPI's	F19	12 555	8 115	– 35.4	11 463	41.3	15 384	34.2	22 088	43.6	**35**
	ROW	G19	– 33 508	– 45 318	35.2	– 44 897	– 0.9	– 55 939	24.6	– 53 783	– 3.9	**36**
A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	

Source: Statistics Netherlands

Table 5 (cont.)

Average 2010–2013 (%)	Posterior values 2020A Projection , scenario A			Posterior values 2020B Projection , scenario B			Posterior values 2020C Projection , scenario C			
	Posterior Estimate	Variation coefficient	Annual growth (%)	Posterior Estimate	Variation coefficient	Annual growth (%)	Posterior Estimate	Variation coefficient	Annual growth (%)	
1.0	638 366	0.0	– 0.10	689 225	0.0	1.00	689 223	0.0	1.00	**5**
49.7	50.0	0.0		60.0	0.0		49.2	1.0		**6**
4.9	– 0.5			4.2			6.4			**7**
54.5	55.1	0.0		55.1	0.0		55.5	1.9		**8**
9.7	10.1	0.0		10.1	0.0		10.3	1.9		**9**
7.1	8.3	0.0		8.3	0.0		8.3	2.8		**10**
10.6	10.2	0.0		10.2	0.0		10.3	2.0		**11**
14.2	13.8	0.0		13.8	0.0		13.6	2.8		**12**
5.0	5.6	0.0		5.6	0.0		5.6	2.9		**13**
	457 848	1.7	– 0.03	489 516	2.3	0.93	492 075	1.9	1.00	**14**
0.0	60 185	0.0	1.00	64 483	0.0	2.00	49 830	8.9	– 1.69	**15**
	121 290	0.0	0.45	172 306	0.0	5.61	131 167	1.0	1.58	**16**
– 8.0	18 787	0.0	0.00	21 580	0.0	2.00	19 002	11.9	0.16	**17**
19.4	19.0	0.0		25.0	0.0		19.0	1.0		**18**
7.9	591 754	0.0	1.50	655 752	0.0	3.00	655 752	0.0	3.00	**19**
	– 532 525	1.5	1.90	– 628 349	1.8	4.34	– 589 770	1.6	3.40	**20**
	347 048	0.2	– 0.10	374 713	0.2	1.00	377 506	1.9	1.11	**21**
	7 583	2.9	0.38	7 606	3.0	0.42	7 617	2.9	0.44	**22**
	46 366	2.8	– 0.47	49 913	2.0	0.58	50 214	3.2	0.67	**23**
0.4	17 255	0.0	0.40	17 255	0.0	0.40	17 255	0.0	0.40	**24**
	53 214	0.0	– 0.10	57 454	0.0	1.00	57 549	2.8	1.02	**25**
	61 970	0.6	– 0.36	66 835	0.7	0.72	68 058	1.9	0.98	**26**
	10.1	0.0		10.1	0.0		10.3	1.9		**27**
	65 413	0.0	– 0.10	70 625	0.0	1.00	70 709	2.0	1.02	**28**
	– 12.0	2.5		– 11.9	2.5		– 11.9	3.6		**29**
	– 23.8	9.0		– 23.7	9.3		– 24.3	8.7		**30**
– 0.4	88 380	0.0	– 0.10	95 421	0.0	1.00	93 696	2.8	0.74	**31**
	39 858	11.6	– 2.12	20 240	34.1	– 11.16	37 052	16.9	– 3.14	**32**
	– 14 682	0.0	0.05	– 24 123	0.0	7.41	– 1	99.9	– 75.91	**33**
3.9	2.3	0.0		3.5	0.0		0.0	100.0		**34**
	20 474	31.9	– 1.08	16 614	62.6	– 3.99	13 775	44.5	– 6.52	**35**
	– 45 649	17.0	– 2.32	– 12 731	87.0	– 18.60	– 50 826	18.2	– 0.80	**36**
M	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	

Following is a description of the details of the projections resulting from the three scenarios:

1. In the first scenario A (columns N-P in table 5), present growth rates and structural ratio values will be used to project the economy to the future without any major change. Average growth rates of the period 2009–2013 for selected exogenous variables have been generally extended to the future; they are assumed to have a prior coefficient of variation of 0 %. This applies to the exogenous variables of GOV final consumption (*E16*), GOV capital formation (*E18*), exports (*G4*) and also social transfers (*E13*). Regarding the latter, it has been assumed that they will be reduced slightly over time, as is the present policy. The growth of population (*F20*) has been assumed to remain stable over time, not explicitly taking into account the effect of immigration. The ratios with values that are maintained at the 2013C level (with 0-value for the prior coefficient of variation) are in particular the i-o coefficient (*C30* = 50 %), the GOV deficit as percent of GDP (*E64*) that remains at the 2013C level of 2.3 %, the coefficients of distribution of GDP to disposable income (*J6* = 55.1 %, *J7* = 10.1 %, *J11* = 8.3 %, *J12* = 10.2 %, *J13* = 13.8 %, *J14* = 5.6 %), and also the investment coefficient (*K18* = 19.0 %) as percent of GDP.

   The results of the projections of option A with the Bayesian method have been included in columns P to R (P: projected values, Q: posterior coefficients of variation, R: annualized growth rates between 2013 and 2020). As a result, the annual growth rates of most variables decrease, except for employment which increases slowly (0.38 %). Imports (*G3*) also increase by 1.9 %. On the other hand, GDP growth is negative (– 0.1 %), GOV income from production and income taxes decreases (annually respectively – 0.36 % and – 0.10 %). Labour income also decreases annually by – 0.10 %, labour income per worker decreases by – 0.47 % and property income by – 0.10 %. Net lending of NFC's and FC's (*D19*) and HH's (*F19*) decreases (by – 2.12 % and – 1.08 % respectively). GOV net borrowing (*E19*) increases slightly by 0.05 % and net borrowing of the ROW (*G19*) decreases by – 2.32 %.

2. The second scenario B (columns Q to S) reflects the economic approach to development. Exogenous variables and ratios (columns S and T) are: (i) Increases in productivity, i.e. increases in the i-o-coefficient between GDP and output (*C30*) to 60 %, (ii) increases in the GOV budget to stimulate growth, i.e. increase of GOV consumption (*E16*) and capital formation (*E18*), each with 2.00 %, (iii) an increase in total investments as percent of GDP (*K18*) to 25 %, (iv) exports (*G4*) increase with 3 %, (v) social transfers (*E13*) do not decrease, but increase with 1 %, and (vi) the GOV budget deficit increases to 3.5 % of GDP (*E64*). Also the coefficients of distribution of GDP to disposable income remain the same as in 2013C, i.e. *J6* = 55.1 %, *J7* = 10.1 %, *J11* = 8.3 %, *J12* = 10.2 %, *J13* = 13.8 %, *J14* = 5.6 %. As in scenario A, the population (*F20*) increase is held stable at 0.4 %.

   As a consequence, the growth rate of GDP (*C5*) increases to 1.0 % and employment increases by 0.42 %. Imports (*G3*) increase considerably with 4.34 %. GOV income from taxes on production (*E7*) and income (*E12*) increase considerably more than in option A, i.e. by 0.72 % and 1.00 % respectively. Labour income (*F6*) increases annually much more than in option A, i.e. by 1.00 %, labour income per worker (*C38*) also increases more by 0.58 % and also property income (*C11*) increases faster by 1.0 %. On the other hand, net lending of NFC's and FC's (*D19*), and HH's (*F19*) decrease considerably by – 11.16 % and – 3.99 %. Net borrowing of GOV (*E19*) increases by 7.41 % as compared to 2013C, and net borrowing of the ROW (*G19*) decreases considerably with – 18.60 %. Thus, despite GOV borrowing more, the financial support of the economy decreases considerably, but the ROW borrows much less than in 2013C.

3. In the third scenario C (columns T to V), GDP growth (*C30*) is fixed at the same level as in scenario B, i.e. 1.0 % growth, while there is no GOV deficit (*E19* = – *1* and *E64* = *0.0 %*). The

objective of this scenario is to determine how such GDP growth with zero GOV deficit would affect other variables in the economy. There are two other exogenous restrictions: Population (*F20*) growth remains fixed at 0.4 % and export (*G4*) growth at 3 %. Thus, in this scenario there are very few prior values, while most projected values are derived with help of identities and structural ratios through the Bayesian estimation approach. Therefore, this scenario would show what changes would be needed in order to arrive at growth rates similar to those under option B, but with a balanced GOV budget.

The consequences of this scenario are that: (i) labour income (*F6*) increases by 1.11 %, (ii) employment (*C22*) rises by 0.44 %, and (iii) labour income per worker (*C38*) augments by 0.67 %; these percent increases regarding labour are the same or slightly higher than in option B. (iv) Taxes on production (*E7*) and income (*E12*) needed to arrive at a balanced budget increase more than in option B, i.e. respectively by 0.98 % and 1.02 %. (v) GOV consumption (*E16*) decreases by – 1.69 % and GOV capital formation (*E18*) increases slightly by 0.16 %, but total final consumption (*C16*) and capital formation (*C18*) increase both respectively by 1.00 % and 1.58 %, which for total capital formation (*C18*) is much less than in scenario B. (vii) Social transfers (*E13*) increase by 0.74 %, which is less than in option B. (viii) GOV net lending (*E19*) decreases considerably by – 75.9 %, as compared to scenario B. On the other hand, ROW net borrowing (*G19*) decreases by – 0.80 %, much less than in scenario B. NFC and FC net lending (*D19*) decrease by – 3.14 %, which is less than in scenario B. Also, net lending by HH's (*F19*) decrease by – 6.52 %, which is more than in scenario B. Overall, financial support is lesser in the scenario than in scenario B.

It should be noted that in all three scenarios the posterior coefficients of variation are generally less than 10 %, and that many of those even have coefficients of variation that are less than 2 %.

Larger coefficients of variation are generally found for the estimates of net lending and net borrowing, which are estimated as residuals, and thus reflect coefficients of variation of other elements of sectors. This means that in the majority of the cases, projections are fairly reliable within the conditions defined by the framework, even though relatively few exogenous variables and ratios are used.

Whether the projections are realistic and their quality is high is determined by two main factors. First, of course, is the prior reliability of the values of the exogenous variables and ratios. The second factor is implicit in the design of the framework: Does it include all exogenous and also endogenous variables that should be taken into account? If the latter should be expanded, should then additional identities be defined? Are the ratio values and their prior reliabilities realistic and/or should additional ratios be defined? An answer to all these questions would make the evaluation of the projections' quality more realistic. It should also be noted that since the number of exogenous variables in the present exercise is very limited, final estimates are heavily dependent on the prior conditions defined in the framework.

7. Conclusions

While the proposed methodology originates in national accounting, it can be used effectively as an instrument for quantitative (policy) analysis in general. The present paper is still mainly national accounts-oriented, but includes elements of a wider discussion, in which frameworks of variables could be designed in response to a variety of integrated analyses that cover many variables and indicators from many different fields of study. They do not only include variables with monetary values, but also environmental variables with physical values and socio-demographic variables based on counts of the population and population segments with particular characteristics. Demographic variables were recently included in a National Transfer Accounts (NTA) framework, as presented in United Nations (2013) and linked to SNA in appendix C of that publication. Thus, frameworks would support very effectively interdisciplinary approaches in quantitative studies of phenomena, which require disciplines to work together in a coordinated manner. A further application of the Bayesian framework approach could be to Big Data (Buelens et al. (2014)), in which case the number of variables and the complexity of the relations between those increases significantly.

At present, frameworks of national economic accounts (SNA) and satellite accounts (Environmental Accounts SEEA, Health Accounts, Tourism Accounts, Social Accounting Matrices SAM's and most recently, National Transfer Accounts NTA) are generally standardized internationally and therefore their analysis is standardized too. In the study, we started by defining the desired analysis through the design of a framework and then by determining how to compile the framework's statistics. Frameworks thus designed could deviate from internationally standardized frameworks and their analysis. By emphasizing analysis in the framework approach, international standardization should not be abandoned, but made more flexible and respond more effectively to the needs of policy analysis in individual countries.

References

Bikker, R., J. Daalmans and N. Mushkudiani (2008), 'Macro-integratie, Deelthema: Inpassen', *Statistische methoden* (08003), Statistics Netherlands.

Bikker, R., J. Daalmans and N. Mushkudiani (2013), 'Benchmarking large accounting frameworks: a generalized multivariate model', *Economic Systems Research*, 25(4), pp. 390–408.

Buelens, B., P. Daas, J. Burger, M. Puts, and J. van den Brakel (2014) 'Selectivity of big data', Discussion Paper, Statistics Netherlands, The Hague/Heerlen.

Danilov, D. and J.R. Magnus (2007), 'Some equivalences in linear estimation' (in Russian), Quantile, 3, 83–90. (English version can be downloaded from http://cdata4.uvt.nl/websitefiles/magnus/paper77b.pdf).

Danilov, D. and J.R. Magnus (2008), 'On the estimation of a large sparse Bayesian system: The SNAER program', *Computational Statistics & Data Analysis*, 52, pp. 4203–4224.

Domar, D. (1946), 'Capital expansion, rate of growth and employment', *Econometrica*, Vol.14, No. 2, pp. 137–147.

Kuznets S. (1960), 'Quantitative aspects of the economic growth of nations: V. Capital formation proportions: International comparisons for recent years', *Economic Development and Cultural Change*, Vol. 8, No.4, Part 2, pp. 1–96.

Piketty, T. (2014), *Capital in The Twenty-First Century*, Harvard University Press, Cambridge, Massachusetts, London, England.

Robilliard, A-S. and S. Robinson (2003), 'Reconciling household surveys and national accounts data using a cross entropy estimation method', *The Review of Income and Wealth*, 49, Number 3,.

United Nations Statistics Division (1999), 'A systems approach to national accounts compilation — A technical report', *Studies in Methods*, Series F No. 77.

United Nations (Population Division, Department of Economic and Social Affairs) (2013), 'National Transfer Accounts Manual: Measuring and Analysing the Generational Economy', E.13. XII.6, New York.

United Nations Development Programme (UNDP) (2015), 'Human Development Report 2015', Work for Human Development, New York.

Van Tongeren, J.W. (1986), 'Development of an algorithm for the compilation of national accounts and related systems of statistics', *The Review of Income and Wealth*, series 22.

Van Tongeren, J.W. (2011), *From National Accounting to the Design, Compilation and Use of Bayesian Policy Analysis Frameworks*, Tilburg School of Economics and Management, CentER Dissertation Series No. 295, Tilburg.

Zellner, A. (1996), *An Introduction to Bayesian Inference in Econometrics*, John Wiley & Sons, New York.

2 A new methodology for processing scanner data in the Dutch CPI

ANTONIO G. CHESSA [1]

Abstract: This paper presents a new methodology for processing electronic transaction data and for calculating price indices, with the aim of reducing the differences across the methods used for different retailers and consumer goods in the Dutch CPI.

Meaningful price indices can only be computed for homogeneously defined products. GTINs (barcodes) contain the highest degree of homogeneity. However, their use may be hampered by the occurrence of so-called 'relaunches'. This term refers to existing items that re-enter the stores with a new GTIN. The old and new GTINs need to be linked in order to capture possible price increases after relaunches. This may be achieved through retailers' own product codes (Stock Keeping Units), or otherwise through item characteristics. A sensitivity analysis is proposed for selecting item attributes, which quantifies the additional impact of attributes on price change.

The new index method calculates price indices as a ratio of a turnover index and a weighted quantity index. The method is in fact the Geary-Khamis method applied to the time domain. Quantity weights of homogeneous products are calculated from prices and quantities of each month in the current year. The weights, which are used to calculate indices with respect to a fixed base month, are updated each month. The method does not lead to chain drift as the price indices coincide with a transitive version of the method at the end of each year. Comparisons with two weighting variants show that the results are robust.

The new methodology has replaced the current sample-based method in the CPI for mobile phones in January 2016. The paper concludes with future plans with the methodology for scanner data of different retailers.

JEL codes: C43, E31.

Keywords: CPI, scanner data, GTIN, relaunch, product homogeneity, chain drift, Geary-Khamis method.

(1) Statistics Netherlands, Team CPI.

1. Introduction

Scanner data have clear advantages over traditional survey data collection, notably because such data sets offer a better coverage of items sold and they contain complete transaction information (prices and quantities). In spite of their potential, scanner data are used by a small number of statistical agencies in their CPI, but the number is gradually increasing (2).

By scanner data we mean transaction data that specify turnover and numbers of items sold by GTIN (barcode). At the time of introduction in the Dutch CPI in 2002, scanner data involved two supermarket chains. In January 2010, the number of supermarket chains was extended to six, as part of a re-design of the CPI (de Haan (2006), van der Grient and de Haan (2010), de Haan and van der Grient (2011)). At present, scanner data of 10 supermarket chains are used and surveys are not carried out anymore for supermarkets since January 2013. Meanwhile, also scanner data from other retailers are used in the CPI (e.g., do-it-yourself stores). Other forms of electronic data containing both price and quantity information that are used in the CPI include data from travel agencies, for fuel prices and for mobile phones. More than 20 % of the Dutch CPI is based on electronic transaction data (in terms of household expenditures of 2015).

Electronic transaction data sets offer possibilities of refining index methods compared with survey data. Expenditure shares can be used as weights at the most detailed product level and complete sets of sold GTINs can be processed instead of samples of items (3). With thousands of GTINs per retailer, the question is how to find efficient and satisfactory solutions. This has turned out to be a complex quest over the years, which is reflected in a range of different methods across retailers and consumer goods in the Dutch CPI (Walschots (2016)). The current method for supermarket scanner data processes 'all' GTINs (i.e., those that satisfy various filters), while the methods for other retailers still make use of samples of items.

As the search for new electronic data continues, the question was put forward whether a generic index method could be developed that is applicable to different types of consumer goods and services and that is capable of handling issues that are not resolved in a satisfactory way so far in current methods. One of these issues is the so-called relaunch problem. A 'relaunch' could be defined as a return to the stores of an existing item after undergoing changes, which typically concern packaging shape and formulation. The GTIN changes as well and the item may have a higher price than before the relaunch, which must be captured by a price index.

Section 2 gives a global outline of a new methodology for processing electronic transaction data. This section shows how the new methodology fits within the CPI system. The aim of the methodology is to process all GTINs, thus abandoning the traditional approach of selecting samples of products, and to have an index method that deals with assortment changes over time, that timely includes new products and that efficiently handles relaunches.

Sections 3 and 4 elaborate the essential components of the new methodology: product homogeneity and price index calculation. When relaunches occur, GTINs are not appropriate as unique identifiers of homogeneous products. Product homogeneity should then be achieved at a broader level, at which GTINs are combined into groups. Homogeneous products could be defined by combining GTINs that share the same set of characteristics. These have to be selected in some way. Section 3 describes and illustrates a method for this purpose.

Turnover and sold quantities of items are summed and used to calculate unit values for homogeneous products. These are subsequently combined into 'quality-adjusted unit value indices'

(2) Until 2015, four countries were using scanner data in Europe (the Netherlands, Norway, Sweden, Switzerland). In April 2016, the number of countries had increased to seven (Belgium, Denmark and Iceland joined meanwhile).
(3) In this paper, the term 'item' and GTIN are used interchangeably.

('QU-indices') for 'consumption segments', which consist of comparable homogeneous products (e.g., a segment T-shirts with products that are described by different characteristics). The index method is described in Section 4.

The price index method uses prices and quantities of each month in the current year for calculating and monthly updating the adjustment factors of the quantities sold of each product (quantity weights). This means that relatively little information is used in the first months (two months in January, three in February, etc., with December as base month). In order to investigate the impact of this smaller amount of information, price indices are compared with a transitive version of the method that uses all 13 months for calculating the quantity weights. The results of an empirical study are presented in Section 5.2.

A second issue concerns the weighting scheme used for calculating the quantity weights of the products. The product prices of each month are deflated by the price indices and weighted according to the share of the quantities sold in a month. In international price comparisons, this method is known as the Geary-Khamis (GK) method. The GK-method has been the subject of criticism because of the 'substitution bias'. When time is replaced by country, the product prices of large countries receive larger weights than those of smaller countries. If the larger countries exhibit higher prices, then a situation can arise that is felt to contradict with economic theories (consumers tend to buy more of some good when prices decrease). In order to gain some insight into the extent of the substitution bias in intertemporal comparisons, two versions of the QU-method with different weighting schemes are considered and compared with the QU-GK method. The results of an empirical study are presented in Section 5.3.

Section 6 summarises the first experiences with the methodology in the Dutch CPI for mobile phones. Final remarks are made in Section 7.

2. Outline of a new processing framework

The development of different methods for different retailers in the Dutch CPI has made the system increasingly complex over time. Different choices were made when a new data set was added to the production system. The current index method for supermarkets uses a monthly chained Jevons index for consumption segments. Different types of price and turnover filters are built within the Jevons index. Because of the equal weighting of GTINs (taken as homogeneous products), items with monthly turnover shares below a certain threshold are excluded. Old and new GTINs of relaunched items are not linked. A 'dump price filter' is applied to outgoing GTINs in order to limit downward biases of the index.

The methods for data of other retailers select samples of GTINs in order to have a better grip on relaunches. The method for do-it-yourself stores handles relaunches by manually replacing outgoing items by new ones. This method may be labour intensive during monthly production. The sample-based methods for a Dutch department store and for drugstores link GTINs automatically by making use of item characteristics.

Possibilities of developing a generic method have been studied in order to reduce the current methodological differences and issues in the Dutch CPI. These differences and issues cover three aspects: data processing, product homogeneity and price index calculation. The differences with regard to these aspects apply to consumption segments in the CPI. This is a product group level below the most detailed ECOICOP[4] level at which CPI figures are published. Con-

(4) ECOICOP = European Classification of Individual Consumption according to Purpose, defined in Annex 1 of Regulation (EU) 2016/792 (http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:JOL_2016_135_R_0002&from=EN).

sumption segments are created in order to link GTINs to ECOICOP in an efficient way. Consumption segments are derived from the retailers' own classification of GTINs (called 'ESBAs' in the Dutch CPI ([5])).

Aggregation of price indices for consumption segments to higher levels is carried out with Laspeyres type methods for different scanner data sets, so the methods become uniform at ECOICOP and higher levels. This section sets the methodological targets for a generic methodology for consumption segments. The methods suggested for reaching these targets are described in the next two sections.

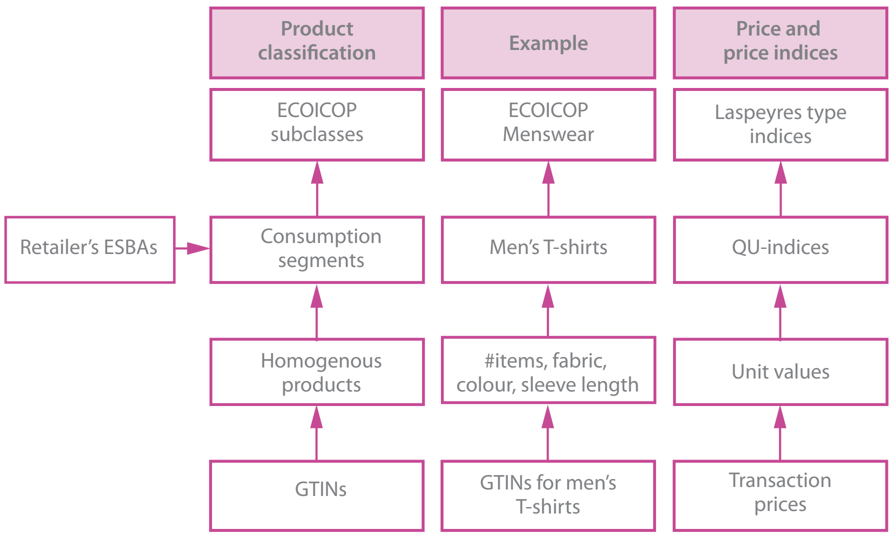
Data processing

The new methodology aims at processing all GTINs, thus abandoning the traditional idea of calculating price indices for a sample of goods. With regard to index calculation this implies that assortment changes are directly included and the use of price and turnover filters is limited to identifying extremes in the data, such as implausible price changes.

Product homogeneity

In principle, the GTIN level is the most detailed level of homogeneity. However, a less detailed level of product differentiation is pursued, since GTINs of relaunched items have to be linked. Two linking possibilities are described in Section 3. For example, GTINs in a consumption segment Men's T-shirts can be combined when items have the same number of T-shirts per package, the same sleeve length, fabric and colour. These GTINs form a homogeneous product (Figure 1).

Figure 1: Nested product classifications, prices and price indices at different levels in the new methodology



([5]) ESBA stands for 'Externe Scannerdata Berichtgever Aggregaat', or 'External Scanner data Retailer Aggregate'

Price index calculation

For each homogeneous product, a unit value is calculated by summing turnover and quantities sold over GTINs that belong to the same product. Unit values and quantities sold for homogeneous products are subsequently used to calculate a price index for each consumption segment. The QU-method is developed for this purpose, which is described in Section 4. This method is able to process all transactions, which means that it includes new products directly into the index. The new methodology is summarised in Figure 1.

3. Consumption segments and product homogeneity

In order to make choices about consumption segments and homogeneous products, statistical agencies require information about item characteristics and classifications used by retailers (ESBAs). Our experiences with electronic data sets show that this information may be supplied in varying formats by different retailers. For instance, the variables in drugstore scanner data are all contained in separate columns (Chessa (2013)). On the other hand, information about item characteristics is only contained in text strings of GTIN descriptions in scanner data of a Dutch department store.

The first example is clearly the preferred data format. The second example requires some form of text mining in order to retrieve and place information about item characteristics in separate columns. Text mining falls outside the scope of this paper and will therefore not be treated further (for details, see Chessa et al. (2015)).

Consumption segments consist of comparable homogeneous products. Different degrees of comparability can be chosen; for instance, different types of socks combined into one segment, or sports, thermal and walking socks as separate segments. We have decided to define consumption segments broadly (i.e., 'men's socks' rather than 'men's sports socks', 'men's thermal socks' and so on). This results in fewer segments, which provides a more manageable system during monthly validation.

When consumption segments have been defined, the question is how to define homogeneous products. Before proceeding, we introduce the following terminology: by 'characteristic' of an item we refer to a specific value that an 'attribute' can take. For example, 'white' is a characteristic of a T-shirt that belongs to the attribute 'colour'.

GTINs can be linked and combined into homogeneous products by using one of the following two approaches:

- Old and new GTINs can be matched through the retailer's internal product codes or SKUs (Stock Keeping Units), which retailers use for their stock administration;
- If SKUs are not available, or cannot be used for some reason, then GTINs can be matched by making use of item characteristics.

SKUs provide direct GTIN matches. If SKUs are not available, then the question is which attributes should be selected for linking GTINs. Item characteristics may have to be extracted from GTIN descriptions, which may be time consuming. Such data sets require maintenance of lists of keywords for identifying characteristics, for example, for new items and because retailers can modify the coding of a characteristic in a text string. Not all attributes may be equally important to product homogeneity, so it is useful to find out whether a subset of item attributes suffices to describe a price index to some degree of accuracy (to be defined).

The traditional work of the consumer analyst, which involves setting up product descriptions for surveys, is combined with a sensitivity analysis that quantifies the impact of attributes on a price index. The suggested approach for selecting item attributes consists of the following steps:

1. For a given consumption segment, the consumer analyst selects an initial set of item attributes that (s)he finds to be relevant;
2. A price index is calculated for the consumption segment according to the method of Section 4. GTINs that share the characteristics chosen in step 1 are combined into the same product;
3. A sensitivity analysis is performed: an attribute that was not selected in step 1 is now added and the price index is re-calculated. If the price index changes 'significantly', then the attribute is added. This step can be repeated with other attributes. Attributes may also be omitted when their impact on the price index is negligible.

An example of a sensitivity analysis is given below. Historical scanner data of a Dutch department store from the period February 2009 until March 2013 are used for this purpose. The selection of item attributes is illustrated for men's and ladies' wear. These two ECOICOP categories are subdivided into four and eight consumption segments, respectively:

- Menswear: socks, underwear, T-shirts, and sweaters and pullovers;
- Ladies' wear: socks, stockings, tights, nightwear, bras, underwear, T-shirts, and sweaters and pullovers.

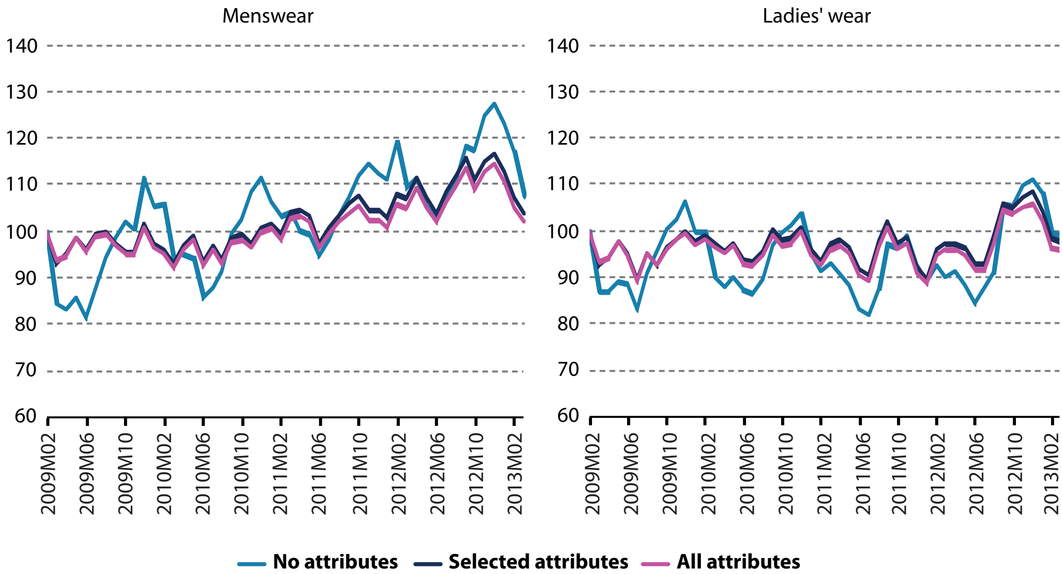
The consumer analyst selected the following item attributes (step 1): type of garment, number of items in a package, fabric, seasonality (e.g., sleeve length) and colour. Some attributes only apply to specific consumption segments (e.g., seasonality applies to T-shirts, pullovers and tights, but not to nightwear and bras).

Price indices were calculated for products differentiated by the above list of attributes. The QU-index method of Section 4 was applied to each consumption segment. These price indices were subsequently aggregated to ECOICOP by calculating Laspeyres type indices, with the turnover shares of the consumption segments of the preceding year as weights.

The price indices for men's and ladies' wear are shown in Figure 2. The indices are compared to price indices based on unit value indices for the consumption segments (thus ignoring all attributes) and the price indices when selecting all attributes from the data. The price indices that are based on the attributes selected by the consumer analyst (dark blue lines in Figure 2) are found to be satisfactory. The differences with the price indices when all attributes would be selected (purple lines) are small at ECOICOP level and can be ignored at overall CPI level. The differences between the year-on-year price indices would affect the CPI by slightly more than 0.001 percentage point in 2010 and even less than 0.001 percentage point in 2011 and 2012. Fabric and colour could even be omitted without changing these findings.

The margin of accuracy mentioned is expected to hold for the entire assortment of the department store. Within the context of a revision programme, which may involve different retailers, a margin of 0.001 percentage point per retailer is more than acceptable. In a situation with, say, five retailers, a margin of 0.01 percentage point per retailer could be set. The impact of unselected attributes over the five retailers together would not become visible at CPI level, since figures are published up to the tenth percentage point. These ideas could serve as a guideline of how the problem of attribute selection may be dealt with in practice. Not only when it comes to defining products before introducing a new method into the CPI, but also when monitoring attributes during production. Experience has to be gained with these issues from this year onward in the Dutch CPI.

Figure 2: Price indices for men's and ladies' wear of a department store, compared with unit value based indices (no attributes selected for consumption segments) and price indices when all attributes are selected
(Feb 2009 = 100)



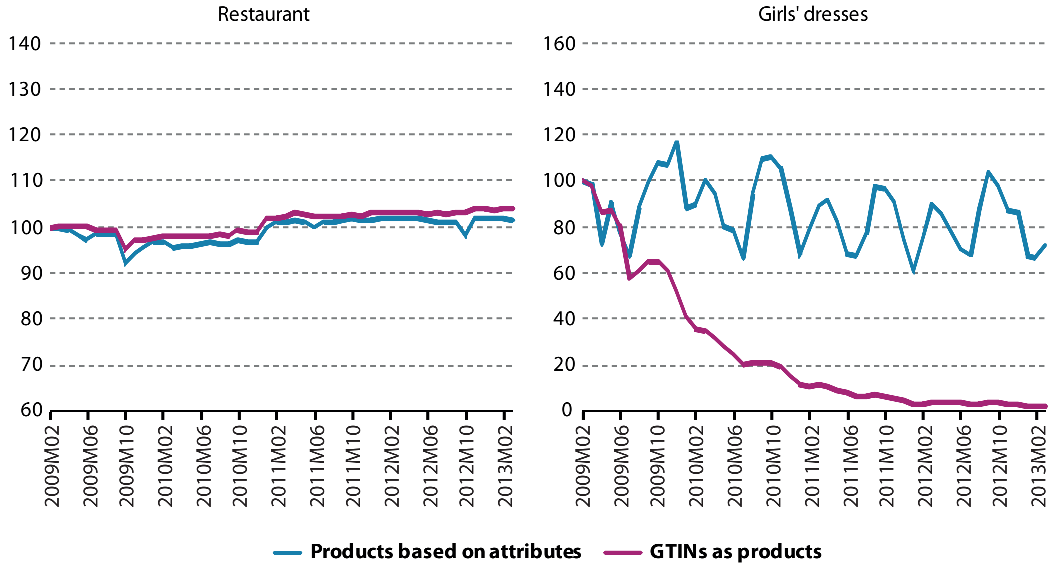
Source: Author's calculations

If SKUs are not available and if information on item characteristics is limited to a few attributes (6), then GTINs could be taken as homogeneous products in specific situations (stable assortments, no relaunches). Figure 3 shows the QU-index with GTINs as products for restaurant services of the department store. SKUs are not available. Three attributes were derived with some effort from the short GTIN descriptions (type of meal or drink, size and taste). The price index with GTINs as products shows a behaviour that is comparable with the price index for products that are characterised by the three attributes. This is an example of a stable assortment, so there is no issue with choosing GTINs as products in this case.

The graph on the right-hand side of Figure 3 shows the two price indices for girls' dresses. The price indices differ considerably. The collection is renewed each year. New items with the same characteristics enter the store at high introduction prices, which rapidly decrease in subsequent months. This explains why the index for GTINs as products shows a fast decline (it gets close to zero in about three years' time). SKUs are not available, so the only option was to match GTINs through item characteristics (type of dress, fabric and colour; size and type of fit do not affect the index). The resulting price index shows a more plausible behaviour, which exhibits a seasonal pattern.

(6) Limited information about item characteristics could be complemented by making use of web scrapers. Web scraped prices and information about item characteristics are currently used for clothing stores in the Dutch CPI (Griffioen and ten Bosch, 2016), but are not (yet) combined with scanner data.

Figure 3: QU-indices for a restaurant and for girls' dresses of a Dutch department store, for two choices of product differentiation
(Feb 2009 = 100)



Source: Author's calculations

4. An index method for consumption segments

4.1. Price index formula

Once consumption segments and the homogeneous products in each segment have been defined, the question is which price indices could be computed. The following aspects were considered in our choice of index method:

- Since the aim is to process complete sets of electronic transaction data, new products should be incorporated into the index calculations during the current month instead of waiting until the next base month;
- The method must be free of chain drift;
- The price index simplifies to a unit value index for homogeneous consumption segments.

This section only presents the general form of the index formula. Specific choices on the quantity weights v_i in expression (1) below have to be made in order to handle the above first two points. These will be described in sections 4.2 and 4.3 and will be illustrated with examples in sections 5.1 and 5.2. The third point follows as a corollary, as will be argued later in the present section.

We introduce the following notation. Let G_0 and G_t denote sets of homogeneous products for some consumption segment G in months 0 and t. The sets of homogeneous products in months 0 and t may be different. Let $p_{i,t}$ and $q_{i,t}$ denote the prices and quantities sold for product $i \in G_t$, respectively, in month t (7).

We denote the price index in month t with respect to, say, a base month 0 by P_t. The following formula is proposed for calculating price indices:

$$(1)\ P_t = \frac{\sum_{i \in G_t} p_{i,t} q_{i,t} / \sum_{i \in G_0} p_{i,0} q_{i,0}}{\sum_{i \in G_t} v_i q_{i,t} / \sum_{i \in G_0} v_i q_{i,0}}.$$

The numerator is a turnover index, while the denominator is a weighted quantity index. The quantity weights v_i are the only unknown factors in formula (1). Choices concerning the definition and calculation of the v_i are described in sections 4.2 and 4.3.

Price index formula (1) can be written in the following compact form:

$$(2)\ P_t = \frac{\bar{p}_t / \bar{p}_0}{\bar{v}_t / \bar{v}_0},$$

where $\bar{p}_t$ and $\bar{v}_t$ denote weighted arithmetic averages of the prices and the v_i, respectively, over the set of products in month t, that is,

$$(3)\ \bar{p}_t = \frac{\sum_{i \in G_t} p_{i,t} q_{i,t}}{\sum_{i \in G_t} q_{i,t}},$$

$$(4)\ \bar{v}_t = \frac{\sum_{i \in G_t} v_i q_{i,t}}{\sum_{i \in G_t} q_{i,t}}.$$

If a consumption segment is homogeneous, then product quantities are summed (equal v_i for all products) so that (1) simplifies to a unit value index, which is the numerator of (2). If consumption segments are not homogeneous, then the unit value index must be adjusted. The adjustment term is the denominator of (2). This term captures shifts in consumption patterns among different products between two months. A shift towards products with higher quantity weights (i.e., more expensive, higher quality products) increases the denominator of (2), which has a complementary downward effect on the price index. We call index (1)-(2) a 'quality adjusted unit value index' ('QU-index'). Such methods have been studied and suggested previously in the price index literature (Silver (2010), von Auer (2014)).

4.2. Definition of the quantity weights

Formula (1) can be considered as a family of price indices. Different choices for the v_i lead to different price index formulas, which also include well-known bilateral indices. For instance, if we set the v_i equal to the product prices from the current month t, then (1) simplifies to a Laspeyres index. If the v_i are set equal to the prices in the base month 0, then (1) turns into a Paasche price index.

(7) A different notation is used in this paper from the commonly accepted notation of time as a superscript in prices, quantities and indices. In this paper, preference is given to the notation of both product and time indices as subscripts. This was done in order to reserve the superscript for other purposes (see Chessa (2015), Section 3).

Bilateral indices have several shortcomings. Monthly chained indices are not able to deal with price changes for products that are temporarily unavailable, while direct indices cannot include new products until the next base month, unless prices are imputed. However, price imputations are not needed when the product prices from each month in the current year are included in the v_i. This allows us to directly compute quantity weights for both existing and new products.

We define the v_i by including the product prices and quantities from each month in the current year and the base month December of the previous year (i.e., for 13 months) ([8]). We define v_i for product *i* as follows:

$$(5)\ v_i = \sum_{z=0}^{12} \varphi_{i,z} \frac{p_{i,z}}{P_z},$$

where

$$(6)\ \varphi_{i,z} = \frac{q_{i,z}}{\sum_{s=0}^{12} q_{i,s}}$$

denotes the share of month z in the total amount of quantities sold for product *i* over the 13 months. The v_i are defined as a weighted average of deflated prices. Price change is thus removed from the product prices in order to yield weights v_i in the weighted quantity index of (1).

Notice that price index (1) is transitive. However, in practice we can use prices and quantities of all 13 months only in the final month of a year, so that some updating method is needed for computing the v_i each month. Such a method is suggested in Section 4.3.

The index method described by formulas (1), (5) and (6) is known as the Geary-Khamis (GK) method in international price comparisons, with time replaced by country (Geary (1958), Khamis (1972), Balk (1996, 2001, 2012)). We denote the QU-method with choices (5) and (6) for the quantity weights by the 'QU-GK method'. The GK-method has been the subject of debates in the literature concerning substitution effects, to which we will return in Section 5.3.

4.3. Computation of price indices in practice

As was pointed out in the previous section, it is not possible in practice to fix the quantity weights at the same values each month when the product prices and quantities from each month in the current year are used to compute the v_i. This raises the question how the quantity weights could be updated each month, such that the resulting indices give good approximations of the transitive index of Section 4.2. We refer to the transitive index as the 'benchmark index'. The index with monthly updated quantity weights is called 'real time index'.

Different methods exist for updating the quantity weights. Krsinich (2014) suggests a rolling 13-month window, which is used to calculate chained year-on-year indices each month. However, this method has given quite volatile and biased results (Chessa (2015)). We propose the following approach for calculating real time price indices:

- A time window is used with December of the previous year as fixed base month. The first window consists of the base month and January of the present year. The window is enlarged each month with the current month;

([8]) Periods of 13 months are preferred to periods with other lengths, as was argued in Chessa (2015).

- The v_i are updated with product prices and quantities of the current month;
- The price index of the current month is calculated with respect to the base month, by making use of the updated v_i;
- The above three steps are repeated until December of the current year;
- The base month is shifted to December of the current year and the above procedure is repeated in the subsequent year. The 13-month indices of each year are chained to the index of December of the previous year.

The third point ensures that the benchmark and real time indices are equal at the end of each year, so that real time indices are free of chain drift as well. This is an essential property of the method. The question is how the two price indices compare in previous months. This will be investigated in Section 5.2 for the Dutch department store.

We now describe a method for calculating the price index of a current month *t*. Price indices cannot be calculated directly, since the v_i depend on the price indices. We propose a method that is easy to implement, which follows an iterative scheme:

1. Choose initial values P_z for the price indices from base month 0 up to month $t \geq z \geq 0$, with $P_0 = 1$,
2. Calculate the v_i for each product sold between the base month and month *t* by making use of product prices and quantities up to month *t*:

   $$(7)\ v_i = \sum_{z=0}^{t} \varphi_{i,z} \frac{p_{i,z}}{P_z},$$

   where

   $$(8)\ \varphi_{i,z} = \frac{q_{i,z}}{\sum_{s=0}^{t} q_{i,s}}.$$

3. Substitute the v_i obtained in step 2 into expression (1) and calculate updated price indices up to month *t*;
4. Repeat steps 2 and 3 until the differences between the price indices obtained in the last two iterations are 'small', according to some pre-defined difference measure.

We conclude with the following remarks:

- The initial indices in step 1 can be chosen arbitrarily. But computation times can be reduced by choosing a constructive approach. A method for calculating initial indices is described in Chessa (2015), which has given very good approximations of the transitive benchmark indices;
- The above algorithm converges to a unique solution. Such a solution exists under mild conditions (Khamis (1972), p. 101)([9]);
- The price indices of all previous months are re-calculated. However, the price indices of previous months cannot be revised in the CPI (apart from exceptional cases). This means that

([9]) Translated into CPI practice, this boils down to checking each month whether a product exists that has been sold both in the current month and in one of the previous months. If this is not the case, then the price index of the consumption segment will be imputed in the current month (e.g., from the corresponding ECOICOP).

only the price index for the current month will be retained from the calculations, which itself will not be modified in successive months.

5. Results and discussion of some issues

5.1. Contribution of new products

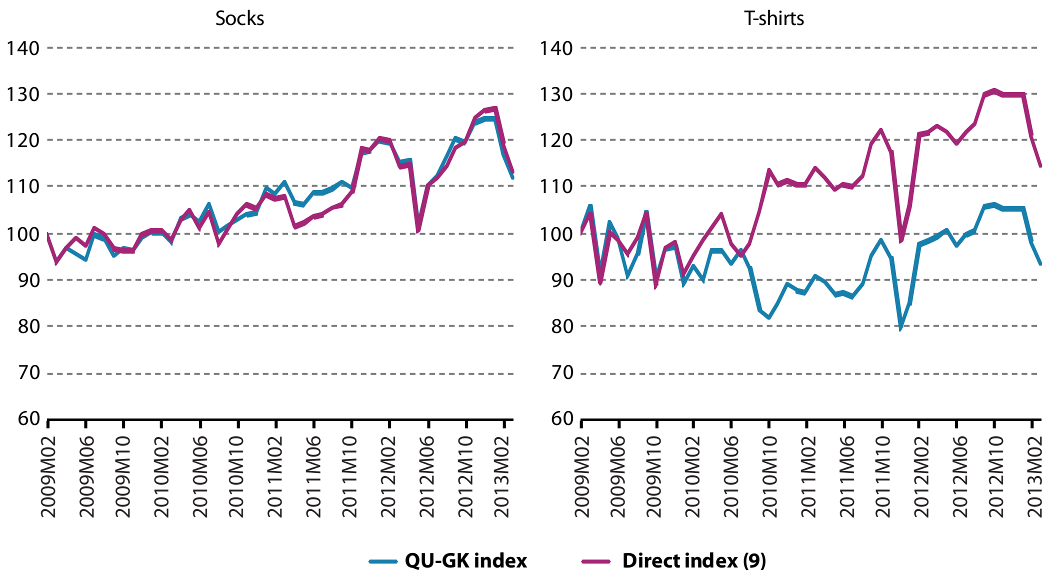
One of the targets in the quest for a more generic index method is the processing of complete data sets. This involves direct inclusion of new products into the index calculations. This section gives an example that shows the extent to which new products may contribute to a price index. QU-GK indices are compared to a direct bilateral index, in which new products are included only in next year's base month.

The bilateral index follows from the QU-GK index with the v_i calculated from prices and quantities of the base month 0 (December of each year) and current month *t*. Solving (1), (5) and (6) gives the following expression:

$$(9)\ P_t = \frac{\sum_{i \in G_0 \cap G_t} p_{i,t}\, h(q_{i,0}, q_{i,t})}{\sum_{i \in G_0 \cap G_t} p_{i,0}\, h(q_{i,0}, q_{i,t})},$$

where *h* is the harmonic mean of the quantities sold in the two months (see also Khamis (1972), p. 102).

Figure 4. Real time QU-GK indices compared with direct bilateral index (9) for men's socks and T-shirts, based on scanner data of a department store
(Feb. 2009 = 100)



Source: Author's calculations

Figure 4 compares the QU-GK index with direct index (9) for men's socks and T-shirts ([10]). The results show large differences for T-shirts. The direct index does not capture the contribution from new products to price change in the year of introduction into the assortment. New types of T-shirts, made of organic cotton, were introduced in 2010 at high initial prices, which started to decrease after a few months. This price decrease is captured by the QU-GK index. The direct index only evidences the price behaviour of the existing part of the assortment, which, in contrast to the new items, mainly shows a price increase in 2010.

The examples show that it is important to have an index method in which not only existing items enter the calculations, but in which also new items are timely included. This means that the v_i should be calculated for new products as soon as these appear in an assortment ([11]).

5.2. Comparisons with the transitive benchmark index

Price indices for consumption segments are calculated according to the algorithm in Section 4.3, which computes real time indices. The resulting indices are free of chain drift by construction, as they are equal to the transitive benchmark indices at the end of each year. The question is how the two indices compare throughout a year.

Real time and benchmark indices were calculated for a large sample from the scanner data of the Dutch department store, which covers almost 60 per cent of the total 4-year turnover in the period February 2009–March 2013. Seven ECOICOP groups were included in the comparison: food and non-alcoholic beverages, menswear, ladies' wear, children and baby clothing, household textiles, products for personal care, and restaurants. QU-GK indices were calculated for each consumption segment in the seven groups, which were aggregated to ECOICOP and overall indices according to Laspeyres type indices, with turnover shares from the preceding year as weights.

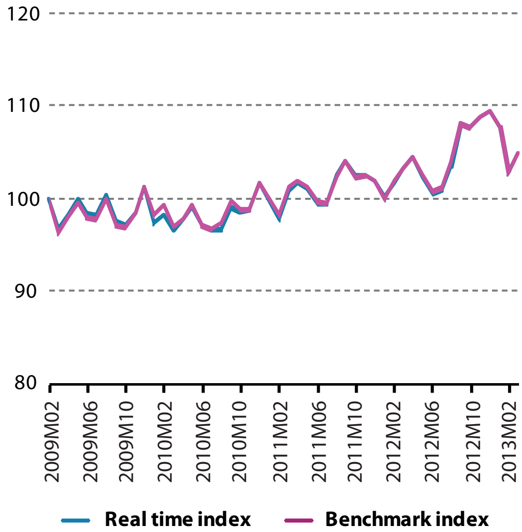
The overall price indices are shown in Figure 5, that is, for the seven groups combined. The differences throughout the 4-year period are small. The year-on-year indices differ by one to several tenths of a percentage point. Averaged over the whole 4-year period, the price indices are equal up to the tenth percentage point.

The differences at ECOICOP level are small as well. For 6 out of 28 cases (7 groups times 4 years), year-on-year differences were larger than 0.5 percentage point. Figure 6 shows both price indices for men's and ladies' wear, which are the groups with the largest differences. But also in these two examples the real time and benchmark indices show small differences. The figures evidence only some local differences, which mainly occur in the first months of a year. Real time indices are calculated with price and quantity data from a few months in these situations, which explains the larger differences found in the first months of some years. However, these differences practically disappear at overall level (Figure 5).

([10]) The results for T-shirts differ from those in Chessa (2015) and Chessa et al. (2015), since a subset of the data was used in the cited papers. However, the findings are the same as in the present study.

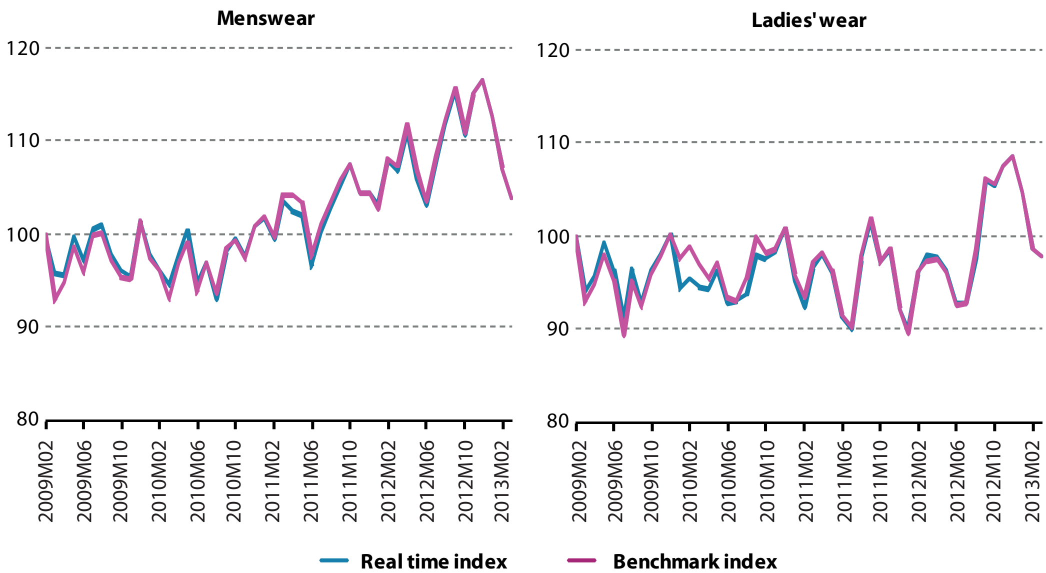
([11]) However, note that new products contribute to the index from the second month in which they are sold.

Figure 5. Real time and transitive benchmark indices for seven ECOICOP groups combined for a Dutch department store
(Feb. 2009 = 100)



Source: Author's calculations

Figure 6. Real time and benchmark indices for men's and ladies' wear for a department store
(Feb. 2009 = 100)



Source: Author's calculations

5.3. The impact of weighting in the quantity weights

The GK-method has been the subject of some debate because of the use of quantity shares as weights in the v_i. International reference prices of commodities (i.e. the v_i, with time replaced by country) will be influenced most by the prices of larger countries. If the larger countries exhibit higher prices than the other countries in the comparison, then the reference prices are felt to contradict with economic theory, as consumers tend to buy more of some good when prices decrease (Balk (1996), Hill (1997), Diewert (2011)). This so-called 'substitution bias' is considered to be a potential problem for all average price methods.

The question is to what extent this bias plays a role in the time domain. This section presents an empirical study, which is aimed at quantifying the impact of different weighting schemes in the quantity weights on the price index ([12]). The QU-GK method is compared with two variants of the QU-method:

- In the first variant, the deflated prices are weighted according to a month's share in the sum of turnover shares of a product over different months. We refer to this variant as the 'QU-TS method';
- In the second variant, each month receives equal weight. This variant will be referred to as the 'QU-EW method'.

The first variant was proposed by Hill (2000) as an alternative to the GK-method in international price comparisons. If we denote the turnover share of product i in month t by $w_{i,t}$, then the weights of the deflated prices in the v_i become:

$$(10)\ \varphi_{i,z} = \frac{w_{i,z}}{\sum_{s=0}^{t} w_{i,s}},$$

where $t \geq z \geq 0$ denotes the current month. Expression (10) is used in the calculation of real time indices, which replaces expression (8) in the algorithm of Section 4.3 ([13]).

The second variant applies the following weighting in the v_i:

$$(11)\ \varphi_{i,z} = \frac{\delta_{i,z}}{\sum_{s=0}^{t} \delta_{i,s}},$$

where $\delta_{i,z} = 1$ if $q_{i,t} > 0$, and $\delta_{i,t} = 0$ otherwise. In other words, deflated prices in months with sales receive the same weight.

At first sight, it might seem odd to ignore the actual sales figures in the quantity weights and only include the information whether a product has been sold or not in a month. However, a deeper analysis shows that weighting scheme (11) leads to an interesting variant of the QU-method: under certain conditions, the price index in the bilateral case is equal to the Fisher index ([14]). For this observation alone it is interesting to include the second variant in the comparison. But also because weighting scheme (11) may lead to completely different weights com-

([12]) Chessa (2015) investigated non-linear forms for the v_i. The price indices hardly differ from the QU-GK method.

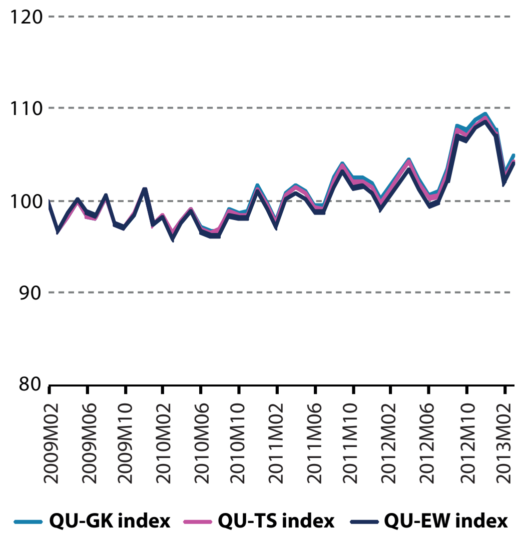
([13]) It should be noted that summing turnover shares over different time periods is not allowed from the viewpoint of the theory of measurement scales. Shares in different periods represent measurements from ratio scales with different scaling factors. The first variant is included in the comparison because it has been suggested as an alternative to the GK-method in PPP-studies. Turnover share based weighting is also used in other multilateral price index methods, like the time product dummy index.

([14]) This holds in the situation where the turnover share of matched products is the same in both months, and in the case where the prices of all unmatched products are imputed. The general expression of the index formula is more complex. Details are left out in this study.

pared to (8). The question then is to what extent the differences between schemes (8) and (11) affect the price indices.

Real time price indices were computed for both variants according to the algorithm of Section 4.3, which are compared below with the QU-GK method. The same data as in Section 5.2 were used. It must be stressed that a threshold was applied in the QU-EW method. This was done in order to prevent out-of-season prices and dump prices of outgoing GTINs to receive disproportionally large weights compared with the usually very low sales. However, a very mild threshold was set: prices were excluded from the index calculations if quantities sold decrease more than 90 per cent with respect to 'regular sales' (quantities sold averaged over past months in which the threshold is satisfied).

Figure 7. Overall price indices for the three QU-methods for a department store (Feb. 2009 = 100)



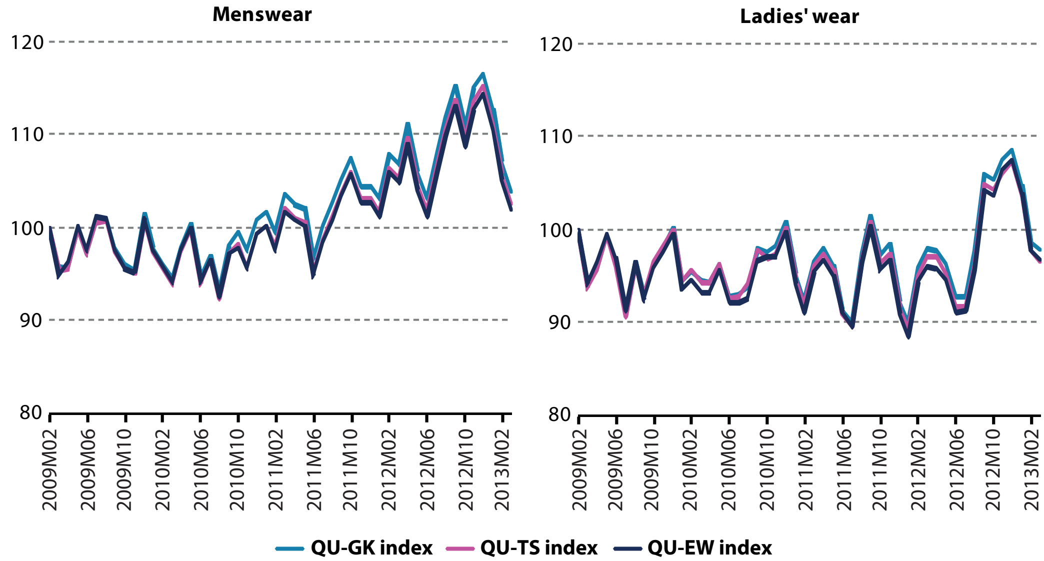
Source: Author's calculations

Figure 7 shows the real time indices for the three QU-methods for the Dutch department store, for the seven ECOICOP groups combined. The three price indices can hardly be distinguished. The QU-EW index lies remarkably close to the QU-GK index. The year-on-year QU-GK indices are 0.25 percentage point higher on average. The differences between the QU-GK and the QU-TS indices are 0.1 percentage point smaller on average.

The differences for the underlying seven ECOICOP groups are small as well. The largest differences are found for clothing. Figure 8 shows the three price indices for menswear and ladies' wear. The differences between the year-on-year indices lie within 0.5 percentage point in most years.

Although the comparative study in this section is empirical, it is important to emphasise the small differences found between the QU-GK method and the two variants. The results show that the substitution bias, if present at all, is very small, not only at overall level, but also for the ECOICOP groups. The differences at consumption segment level are somewhat larger, but are small at this most detailed level as well.

Figure 8. The three QU-indices for men's and ladies' wear for a department store (Feb. 2009 = 100)



Source: Author's calculations

6. First experiences in the Dutch CPI

The methodology has been implemented last year and has been tested on scanner data of a department store and on electronic transaction data of mobile phones. A summary of different research and test phases for the department store data, from text mining and attribute selection to index calculation and validation of the results, is described in Chessa et al. (2015).

The methodology is part of the CPI for mobile phones since January 2016. This section gives a summary of different process stages, from data analysis and attribute selection until price index calculation.

Data analysis

Each month, Statistics Netherlands receives transaction data on mobile phones from market research company GfK. The data analysed cover the period December 2013 until December 2015. For every device, the data include transaction prices, numbers of devices sold and information on item attributes, either as a separate variable or contained in the item description. All variables were reported for each device in every month. Outliers were not detected among prices of devices. A number of potentially relevant item attributes are not included in the data, amongst which all processor characteristics (e.g., speed, number of cores), working memory and screen resolution.

Attribute selection and homogeneous products

Additional attributes were collected from a web site for a smaller set of 70 devices, which together cover about 75 % of the 2-year turnover. The attributes were analysed by applying a sensitivity analysis as described in Section 3. The first step in this analysis was to quantify the impact of each attribute separately on the unit value index. Next, the most influential attribute was selected and others were added in order to quantify their additional contribution to the year-on-year index. Five attributes completely determine the index. Most attributes seem to be correlated, in the sense that, for instance, devices with a higher screen resolution usually have a more powerful processor.

From the set of 5 attributes, Near Field Communication was left out since paying by smartphone is still in a pilot phase in the Netherlands. This may change in the coming years, in which case NFC could be added as a relevant attribute. Long Term Evolution (LTE/4G) adds less than 0.01 percentage point to the year-on-year index, so that LTE was omitted as well. Moreover, the majority of the smartphones is currently equipped with LTE. This share is still growing, so we do not expect this attribute to contribute much to product differentiation.

Three attributes were eventually selected: brand, internal storage capacity and 'performance'. The latter is measured by a benchmark test score ([15]), which indicates how different components of a device act together when performing CPU and GPU tasks (processor type/model, number of cores, working memory). Benchmark scores are obviously not included in the data, so we collected scores from the internet. This has been done for more than 130 devices, which cover 88 per cent of the total 2-year turnover. Benchmark scores are subdivided into three performance segments (high, medium and lower performing devices). Refinements to 4 or 5 segments did not affect the price index significantly.

Price index

Mobile phones are defined as a consumption segment, but are also an ECOICOP subclass in the CPI. Mobile phones are differentiated into homogeneous products according to the three attributes mentioned above. An example of a product is a high performance, 16GB storage Apple product. This product contains the iPhones 5s, 6 and 6s. Another example is a medium performance, 16GB storage Samsung product, which contains the devices Galaxy S5 Mini, S3 and S4.

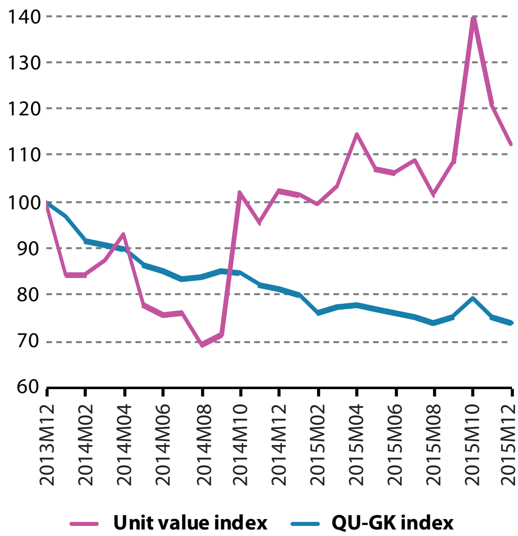
Figure 9 shows the real time QU-GK index for mobile phones, which is compared with the unit value index. The unit value index strongly deviates from the price index. Its volatile behaviour can be explained to a large extent by the introduction of new high-end devices. For instance, the introduction of the iPhone 6 in October 2014 and the iPhone 6s in October 2015 can be easily singled out. The price index lies above the unit value index before the introduction of the iPhone 6. This means that the consumption pattern initially shifted towards devices of lower quality. The introduction of the iPhones 6 and 6s attracted many consumers, which led to shifts towards higher quality devices. The quality-adjustment term of the QU-index, that is, the denominator of expression (2), can be used to quantify these shifts. The quality-adjustment term becomes greater than 1 in October 2014 and increases afterwards, as can be seen as well in Figure 9.

In order to carry out the monthly index calculations, benchmark scores for new devices have to be collected first (this requires a small amount of time). In addition to this, the set of attributes

([15]) The benchmark Geekbench is used, which is developed by Primate Labs. Geekbench is widely used for testing smartphones and notebooks, and test scores are reported in reviews. Benchmark scores can be found on https://www.primatelabs.com/ and http://www.notebookcheck.net/

needs to be monitored. The intention is to do this twice a year. Attributes that were not selected might become relevant. Technological developments may lead to the need to include new attributes. The approach described in Section 3 will be applied for this purpose.

Figure 9. QU-GK index for mobile phones compared with the unit value index (Dec. 2013 = 100)



Source: Author's calculations

7. Final remarks

The differences across index methods that make use of electronic data in the Dutch CPI, in conjunction with the increased use of such data, motivated a search towards a more generic index method. To this end, a methodology for characterising homogeneous products and for calculating price indices has been developed. The question whether the methodology can be applied to data sets of different retailers and consumer goods can be answered in a positive way.

The methodology allows to process complete data sets, without applying turnover or dump price filters, irrespective of retailer and type of consumer good. The methodology has been applied to a broad range of consumer goods, which include the broad assortment of a department store, mobile phones, items sold by do-it-yourself (DIY) stores and drugstores.

The QU-method makes it possible to include new items into the index calculations as soon as items are introduced into an assortment. This is a desirable property, since postponing the inclusion of new items until the next base month may have a big impact on a price index (Section 5.1). The quantity weights are exclusively based on consumption patterns of the current year. This is a major difference and contribution compared to traditional methods.

The quantity weights are updated every month, so that the weights are time dependent. However, the resulting real time price indices are free of chain drift by construction. The differences with transitive benchmark indices appear to be very small throughout the year and vanish at the end of each year. The smaller amount of price and quantity information in the first months of a year does not seem to be an issue. Also a recent study hardly showed any difference between the real time and benchmark indices for drugstore scanner data (Chessa (forthcoming)). The index method of Section 4.3 thus seems to be an adequate updating method.

The empirical study of Section 5.3 suggests that the possible impact of the substitution bias is very small or can even be ignored (see also Chessa (forthcoming)). One of the two variants of the QU-GK method turned out to be very interesting, which is the one with equal weights applied to the deflated prices. The price indices for this variant show small differences with the QU-GK indices. This is an intriguing result, which deserves further study.

The comparable results obtained for the equal weights variant may reveal to be a very useful finding in view of the rapidly increasing popularity of using web scraped internet prices in price index methods (Griffioen and ten Bosch, 2016), because quantities sold are not available. This makes the QU-EW variant an interesting method for future studies with internet prices.

Application of the QU-GK method to the department store scanner data in the CPI is expected to be realised within several months. Application of the methodology to other scanner data sets is being investigated, in particular for DIY-stores. A test data set containing SKUs and information about additional attributes for paint and electrical equipment has been analysed. The use of SKUs for linking GTINs looks promising.

A study that compares the QU-GK method with the current method for supermarket scanner data and a comparative study that involves different methods for consumer electronics are planned at a later stage. These studies are part of a 4-year research programme at Statistics Netherlands (de Haan et al. (2016)).

Acknowledgments

The author wants to thank two anonymous reviewers for their valuable comments and his colleagues at Statistics Netherlands for the many discussions. The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

References

Auer, L. von (2014), 'The generalized unit value index family', *Review of Income and Wealth*, 60, pp. 843–861

Balk, B.M. (1996), 'A comparison of ten methods for multilateral international price and volume comparison', *Journal of Official Statistics*, 12, pp. 199–222.

Balk, B.M. (2001), 'Aggregation methods in international comparisons: What have we learned?', Paper originally prepared for the Joint World Bank — OECD Seminar on Purchasing Power Parities, 30 January–2 February 2001, Washington DC, USA.

Balk, B.M. (2012), *Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference*, Cambridge University Press, UK.

Chessa, A.G. (2013), 'Comparing scanner data and survey data for measuring price change of drugstore articles', Paper presented at the Workshop on Scanner Data for HICP, 26–27 September 2013, Lisbon, Portugal.

Chessa, A.G. (2015), 'Towards a generic price index method for scanner data in the Dutch CPI', Room document for the Ottawa Group Meeting, 20–22 May 2015, Urayasu City, Japan.

Chessa, A.G. (forthcoming), 'The QU-method: A new methodology for processing scanner data', to appear in the Proceedings of Statistics Canada 2016 International Methodology Symposium.

Chessa, A.G., Boumans, S., and Walschots, J. (2015), 'Towards a new methodology for processing scanner data in the Dutch CPI', Paper presented at the Workshop on Scanner Data, 1–2 October 2015, Rome, Italy.

Diewert, W.E. (2011), 'Methods of aggregation above the basic heading level within regions', Chapter 5 in *Measuring the Real Size of the World Economy*, ICP Book, World Bank, Washington DC, USA.

Geary, R. C. (1958), 'A note on the comparison of exchange rates and purchasing power between countries', *Journal of the Royal Statistical Society* A, 121, pp. 97–99.

Grient, H.A., van der and de Haan (2010), 'The use of supermarket scanner data in the Dutch CPI', Paper presented at the Joint ECE/ILO Workshop on Scanner Data, 10 May 2010, Geneva, Switzerland.

Griffioen, R., and ten Bosch, O. (2016), 'On the use of Internet data for the Dutch CPI', Paper presented at the Meeting of the Group of Experts on Consumer Price Indices, 2–4 May 2016, Geneva, Switzerland.

Haan, J. de (2006), 'The re-design of the Dutch CPI', S*tatistical Journal of the United Nations Economic Commission for Europe*, 23, pp. 101–118.

Haan, J., de and van der Grient, H.A. (2011), 'Eliminating chain drift in price indexes based on scanner data', *Journal of Econometrics*, 161, pp. 36–46.

Haan, J. de, Willenborg, L., and Chessa, A.G. (2016), 'An overview of price index methods for scanner data', Room document for the *Meeting of the Group of Experts on Consumer Price Indices*, 2–4 May 2016, Geneva, Switzerland.

Hill, R.J. (1997), 'A taxonomy of multilateral methods for making international comparisons of prices and quantities', *Review of Income and Wealth*, 43, pp. 49–69.

Hill, R.J. (2000), 'Measuring substitution bias in international comparisons based on additive purchasing power parity methods', *European Economic Review*, 44, pp. 145–162.

Khamis, S. H. (1972), 'A new system of index numbers for national and international purposes', *Journal of the Royal Statistical Society* A, 135, pp. 96–121.

Krsinich, F. (2014), 'The FEWS Index: Fixed Effects with a Window Splice — Non-revisable quality-adjusted price indexes with no characteristic information', Paper presented at the Meeting of the Group of Experts on Consumer Price Indices, 26–28 May 2014, Geneva, Switzerland.

Silver, M. (2010), 'The wrongs and rights of unit value indices', *Review of Income and Wealth*, 56 , pp. S206–S223.

Walschots, J. (2016), 'Fifteen years of progress in the collection of prices data in the Netherlands', Paper presented at the Meeting of the Group of Experts on Consumer Price Indices, 2–4 May 2016, Geneva, Switzerland.

3 EA and EU GDP flash estimates at 30 days

ARTO KOKKINEN AND HANS WOUTERS [1]

Abstract: GDP growth is one of the most followed National Accounts variables. At best, economic decision making would need the data real-time. This, though, is not possible because of the need for collecting and analysing the basic statistical information. Therefore, a reasonable request from the decision-makers is to have the data available as soon as possible after the end of the quarter. However, there is a delicate balance between the speed of publication and accuracy of the early estimate.

This paper shows how Eurostat with a group of EU Member States has managed meeting the challenge to produce high quality data on GDP growth for the EU and the euro area at an advanced timeliness of 30 days after the quarter-end. The aim of the paper is to address the request of the statistical and scientific world to provide transparency and to better understand the development and test process of new statistical products in official statistics, more concretely regarding the new GDP t+30 flash estimation.

The paper elaborates the statistical methods that Eurostat developed for the early GDP estimates and discusses the user needs motivating the advancement of GDP estimates. It analyses the results of the estimates for 16 test quarters that are based on the national estimates of a group of Member States and that were prepared in close cooperation with them. Finally, the article discusses the a priori set quality acceptance criteria for the test estimates, assesses the test results against those criteria, and concludes that the test results passed all predefined quality criteria.

JEL Codes: C82, E01, E32

Keywords: Gross domestic product, economic growth, early estimates, monitoring economic growth and business cycle

[1] Eurostat, Unit C2 National accounts — Production.

1. Introduction

1.1. Research question

The research question addressed in this article is the following: *Estimating reliably European GDP at 30 days after the quarter-end — is that feasible?*

Since 2003, the first quarterly GDP estimates for the euro area and the EU have been produced at 45 days after the quarter-end. In comparison, one of the main EU trading partners, the US, compiles its first quarterly GDP estimate at 30 days since the late 1970s. Some European countries — United Kingdom, Spain, Belgium, Lithuania and more recently Austria, Latvia and France — publish GDP flash estimates at 30 days after the quarter-end as well.

Eurostat started to investigate the possibility of advancing the first GDP estimate for the euro area and the EU with a group of Member States in 2013. The article discusses whether and how Eurostat, in close cooperation with Member States, has met the challenge of the above research question.

1.2. Background

Eurostat began publishing quarterly GDP flash estimates in 2003. Since then, these estimates have been released every quarter, 45 days after the quarter-end. The introduction of the GDP t+45 flash estimates was described in the Eurostat flash project (Eurostat, 2003) as the project's 'first practical objective'. Reducing the time needed to produce a flash estimate of GDP to 30 days was cited as a 'next step' for the project. Ten years later, in May 2013, a task force was established with the mandate of assessing whether it would be possible to publish a sufficiently reliable flash estimate of GDP for the euro area and the EU based on the information available at t+30 days, including national estimates produced by the Member States.

The work on the GDP t+30 days project covered:

- sharing information on and best practices for preparing early quarterly GDP estimates between Member States;
- producing a country-by-country overview of the sources available for GDP estimates between t+25 and t+60 days;
- developing a guidance document on methods and national estimation techniques;
- developing a method for producing GDP t+30 flash estimates for the euro area and the EU;
- developing quality acceptance criteria to be used for assessing the test results;
- preparing test estimates; and
- developing a communication plan.

1.3. Content of the article

The paper is organised as follows: section 2 discusses the question of why the GDP t+30 flash estimates for the euro area and the EU are being introduced and elaborates user needs. Section 3 examines the estimation methodology developed by Eurostat. This methodology was used for preparing the test estimates and will continue to be used for producing the official estimates. The methodology for GDP t+30 is placed in a broader perspective by also addressing the methods used for the GDP t+45 and the 'regular' annual and quarterly national accounts estimates. Section 4 discusses the quality acceptance criteria and how they were developed for assessing the results of the test estimates. Section 5 presents the results of the 16 quarterly test estimates, and the assessment of these results against the quality acceptance criteria. The test estimates are also compared with the US results for the corresponding quarters. Section 6 concludes the article.

2. Why introduce GDP t+30 estimations ?

2.1. Introduction, user needs and developments in official statistics

The question as to why GDP t+30 flash estimations are being introduced can be answered by considering three issues: user needs and developments in official statistics, the background of the GDP t+45 flash estimates, and the link between the GDP estimates and the Principal European Economic Indicators (PEEIs).

User needs, economic and financial developments and technological improvements inevitably trigger changes and further progress in the area of official statistics. In view of this, users' demand for more detailed information and for information to be available sooner plays an important role.

At a strategic level, in the ESS Vision 2020 the European Statistical System (ESS) has committed to 'respond to the need to provide policymakers with reliable, comparable and timely statistics to execute economic and financial policy' (European Statistical System (ESS), 2013). The work on the GDP t+30 flash estimates is one aspect of the action being taken by the ESS to fulfil this commitment.

Since May 2003, Eurostat has been publishing a flash estimate of quarterly seasonally and calendar-adjusted GDP growth 45 days after the end of the reference quarter (t+45). This has been greatly appreciated by the users of statistics and in particular by policymaking organisations such as the European Commission and the European Central Bank (ECB). The economic, financial and technological developments seen in recent years mean, however, that policymakers increasingly need to analyse economic activity faster, i.e. sooner after the end of a reference period. Sooner available high quality, reliable published statistical data is therefore essential. As a result, policymakers and other main users of statistics were keen to see quarterly GDP estimates available earlier.

MAIN USERS: THE EUROPEAN COMMISSION AND THE EUROPEAN CENTRAL BANK

The European Commission's Directorate-General for Economic and Financial Affairs (DG ECFIN) and the ECB rely heavily on the macroeconomic statistics produced and disseminated by Eurostat, and in particular on the quarterly national accounts. They therefore supported the ESS's work on producing GDP t+30 flash estimates for the euro area and EU. Bringing the GDP release date forward is seen as a significant improvement in data availability, and it also meets users' long-standing demand to have flash estimates available for the euro area and the EU at a similar date to that at which data is available for the other main economies, and in particular the euro area's main trading partners, such as the United States and the United Kingdom.

Having GDP estimates available earlier will be useful for the following purposes:

A. For policy decision-making

The ECB is responsible for monetary policy in the euro area. The ECB's Governing Council aims to keep inflation below, but close to, 2 % over the medium term. In order to achieve this objective, the Governing Council bases its decisions on a two-pillar monetary policy strategy, and implements them via its operational framework. The two pillars are economic analysis and monetary analysis — these form the basis of the Governing Council's overall assessment of the risks to price stability, and thus of its decisions on monetary policy. The ECB's economic analysis focuses on real economic activity, and therefore takes into account the fact that price developments over those horizons are largely determined by the interplay of supply and demand in the goods, services and factor markets. The GDP flash estimate is thus one of the most important indicators used when setting monetary policy, in particular because policy decisions influence developments in economic variables such as output and prices. Bringing the quarterly GDP flash estimate forward by two weeks is therefore valuable to the ECB for its monetary policy decision-making.

B. For monitoring the economic situation

- DG ECFIN provides a continuously updated assessment of the overall economic situation in the euro area and in the EU. This is used in discussions by all groups and committees working in areas covered by DG ECFIN, including particularly the Eurogroup. DG ECFIN and other Commission departments are also in close and frequent contact with other international organisations.
- The ECB (specifically the Directorate General Economics) produces updated assessments of the overall economic and monetary situation in the euro area and briefing notes that include the latest statistics on a regular basis. These are then used by the ECB's Executive Board and its Governing Council to inform their discussions, and by other EU bodies and international fora.

Having a sound understanding of the current developments in the economy, at an aggregate level, is critical in the contexts mentioned above. This understanding is informed by all the available statistics and indicators, and earlier availability of GDP estimates would therefore be of great benefit, GDP being the standard accepted aggregate measure of economic activity.

A reliable, objective and publicly available measure, such as GDP growth in the euro area and the EU, can help to establish a common understanding in international fora, and having this information available earlier would therefore be very valuable in this context also.

C. For forecasting/projection purposes

- The Commission's macroeconomic forecasts for the euro area and the EU are built bottom-up by aggregating data from individual country forecasts produced by country desk officers in DG ECFIN, but also include top-down elements, such as information provided by European aggregates. These forecasts are produced three times a year (winter, spring and autumn) and each forecast involves several iterations. Estimates of GDP for the euro area and the EU released 30 days after the end of the reference period could be of great value for forecasting. At least in the current forecast set-up, the flash estimate at t+30 days would still be released after the internal 'cut-off date' for information for each of the three forecasts, and could not therefore be taken into account when producing the forecasts. Nonetheless, in contrast to the t+45 releases, these estimates would be available prior to the *publication* of the forecast (the Q4 estimate would be available before the publication of the winter forecast in early February, the Q1 estimate before the publication of the spring forecast in early May, the Q3 estimate before the publication of the autumn forecast in early November). This would enable the communication issued at the time of the forecast release to be adapted to the latest economic developments.

Whether the earlier availability of the European aggregates could help to ensure and improve the quality of the forecast in future depends on when a t+30 days release would fall in future forecast calendars. If the GDP estimate were available at an early stage in the forecasting calendar, it could help to shape the central guidance issued to country desks at the beginning of the forecast exercise. This would thus ensure that national forecasts reflected the available aggregated figures (top-down approach) and that country forecasts were produced on the basis of a common interpretation of the current economic situation. If available at a later stage of a forecast exercise, the GDP t+30 flash estimates could be used for checking the plausibility of the aggregated national forecasts.

- The ECB carries out a number of forecasting and projection exercises as part of its monetary policy strategy. Two distinct types of projections can be identified.

  The first are the forecasted 'early' estimates of quarterly GDP growth, which are produced twice a month. The forecasting model is a complex tool that combines backcasting and nowcasting techniques, focusing on quarterly developments. The earlier availability of a GDP flash estimate for the euro area will be quite significant for this type of forecasting. Having the GDP estimate available sooner would also reduce the time-gap between the release of this and other economic indicators used in the forecasting process.

  The second type of forecasting is the macro-projection exercise, which is performed on a quarterly basis in two streams (one of which also involves all EU national central banks) and is similar to the forecasting carried out by DG ECFIN. The results of the ECB's macro-projections are made publicly available to users via the ECB's website and are also discussed in a dedicated publication. There is a 'cut-off' date for data to be included in these projections and it is important both for euro-area aggregates, such as GDP, and for the euro-area and EU Member States individually, that all statistics are taken into account and that the most recent data are included. Bringing the flash estimate of GDP for the euro area and the EU forward is therefore also useful for this type of projection.

OTHER USERS

In addition to the main users discussed here, there are also many other users who will benefit from data on GDP growth being available sooner. These include other Commission departments, such as the Directorate-General for Employment, Social Affairs and Inclusion and the Directorate-General for Financial Stability, Financial Services and Capital Markets Union. In addition, having GDP data available earlier will be of relevance for institutional users such as the European Council and the European Parliament. Earlier information on the GDP growth rate in the euro area and in the EU will also be useful for national policymakers and governments.

Finally, the general public is better served by providing earlier, but still reliable estimates of economic development.

2.2. Background to the GDP t+45 days estimates

The aim of the GDP t+45 estimates was (and still is) to provide quarter-on-quarter and year-on-year seasonally and calendar adjusted growth rates for quarterly GDP. Results for the euro area, the EU and for individual Member States that publish data at a national level are published in Eurostat's news release at (or slightly before) 45 days after the end of the reference quarter.

Eurostat started producing and releasing the t+45 GDP estimates in 2003. At the time, the euro area and the EU comprised, respectively, 12 and 15 countries. At the beginning of the project, five countries (Germany, Greece, Italy, the Netherlands and the United Kingdom) provided their national estimates to Eurostat. These countries together accounted for more than 50 % of the EU GDP. For the large Member States that could not provide national estimates, related indicators were used. The national estimates (for the countries listed) and related indicators (for other large countries) provided the input for a regression model that estimated the GDP for the euro area and the EU. The launch of the GDP t+45 releases in 2003 was viewed as an intermediate stage that would eventually lead to estimates of quarterly GDP being released 30 days after the end of the reference quarter.

Gradually, more Member States started producing estimates of national GDP at t+45 days. By 2006Q1, for example, the countries contributing to the GDP flash project represented 74 % and 72 % respectively of the euro-area and the EU GDP. By 2015Q2, this had risen to 97 % for the euro area and 95 % for the EU. The revisions made to the GDP t+45 estimates at the first regular release at t+65 days are very small. The estimates of the European aggregates are now almost entirely based on Member States' input. Estimation techniques and modelling are no longer used. Section 3.2 discusses the methodology used to produce the GDP t+45 estimates for the euro area and the EU in more detail. Table 1 in section 4.3 describes in turn the good revision performance of the euro area t+45 estimates.

2.3. Link with Principal European Economic Indicators (PEEIs)

Users of EU economic statistics have, for a long time, been looking forward to see releases of European national accounts data set to a calendar of 30, 60 and 90 days after the end of the reference period. This would match the established practices for data releases of other critical monthly and quarterly macroeconomic indicators as well as those of the EU main economic partners such as the United States. The largest euro area and EU Member States are also already following this schedule for data releases.

A first attempt at assessing the feasibility of compiling a GDP t+30 flash estimate for the euro area and the EU was made in 2008, as part of a work on the Principal European Economic Indicators (PEEIs), which includes the GDP flash estimate. In a workshop organised by the ECB Directorate General Statistics, the practices of a number of national statistical institutes, the users' needs of ECB's statistics, and the Eurostat plan to develop GDP t+30 flash estimates for the euro area and the EU were discussed. The schedule for data releases proposed as a result of this initiative was as follows:

- 30 days: advanced estimates ('flash estimates') of GDP, possibly employment, and ideally some expenditure/output components;

- 60 days: preliminary estimates of GDP and the main expenditure and output components; a flash estimate of the balance of payments;
- 90 days: comprehensive quarterly data estimates of GDP and main expenditure and output components, plus quarterly financial and non-financial accounts for institutional sectors and balance of payments statistics.

Although this proposal was not put into practice at the time, the users' and policymakers' need for earlier estimates remained the same.

The main advantage of the proposed schedule, with data releases at 30, 60 and 90 days was (and still is) that it allows a comprehensive and consistent approach to be taken to data releases across the different areas within national accounts. The proposed GDP t+30 flash estimates are an important part of this schedule of releases.

3. Eurostat's estimation methods

This section discusses the method used by Eurostat for compiling the GDP t+30 flash estimates for the euro area and the EU. This method was developed by Eurostat and used to produce the test estimates and it is used to compile the regular GDP t+30 estimates that are released as of the 1st quarter 2016.

Section 3.1 introduces the terminology related to flash estimates. Section 3.2 discusses first the overall methods used for preparing the annual and quarterly national accounts. As the method for preparing the GDP t+30 flash estimate was designed to be as close as possible to that used for the GDP t+45 flash estimate, section 3.2 elaborates secondly the method used to compile this latter estimate. Finally, the section concludes by presenting the method developed for the GDP t+30 estimates.

3.1. Flash estimate for a quarterly national accounts variable: terminology

The users of national accounts data have been calling for reliable, early quarterly European national accounts figures for some time. The short-term indicators currently available (e.g. production and other short term statistics indices, statistics on prices and foreign trade and business surveys) help users to form a picture of economic developments in a particular variable and/or industrial activity. In order to obtain a more complete picture of developments at macroeconomic level, however, there needs to be a system in which each partial variable can be assessed with respect to other variables, and in which the breakdown of a variable is shown as proportions of an aggregate figure.

Quarterly national accounts form such a system. Users would like to have at their disposal preliminary flash estimates available for all the main quarterly national accounts variables. Ideally, all the variables would be published soon after the end of the quarter and would be reliable. Experience has, however, shown that it is easier to produce a reasonably reliable early estimate of aggregate GDP than of its components. On top of that, quarterly GDP growth is the headline indicator of the quarterly national accounts.

The flash estimate of quarterly GDP:

- gives an early picture of economic development;
- can be used in conjunction with other national accounts variables;
- is produced and published as soon as possible after the end of the quarter; and
- is based on a less complete set of information than is the GDP estimate given in the traditional quarterly national accounts.

Flash estimates differ equally from forecasts, pure nowcasts and leading indicators. The aim of producing GDP flash estimates at EU level is to estimate GDP growth in a way that is, as far as possible, similar to that in which later regular estimates are produced but by using less complete data sources.

Flash estimates and traditional estimates of quarterly accounts differ in a number of their characteristics:

- Release date: flash estimates are available earlier than traditional estimates (typically within 30–45 days).
- Accuracy: there is a trade-off between timeliness and accuracy. Flash estimates are in general less accurate than traditional estimates. The loss of accuracy is, however, kept as small as possible ([2]).
- Reference period: flash estimates are typically only produced for the latest quarter. The data for the preceding and earlier quarters are not usually revised.
- Coverage: the number of breakdowns of variables included in flash estimates is usually more limited.
- Information available: flash estimates are based on a more limited set of information. Information from surveys covering the whole quarter, for example, is often not available.
- Use of estimates: due to the lack of direct information, flash estimates may include components that are estimated using statistical methods. For example, there may be only two months of short-term statistics available. Other available related indicators might thus be used for the third month. The third month estimate, included in the flash, would typically be produced by combining all the back information on past observations with a current quarterly indicator by using autoregressive distributed lag (ADL) models (see e.g. Stock & Watson, 2007), autoregressive integrated moving average with exogenous variables (ARIMAX) models and time series regression techniques. (On ARIMA and ARIMAX models and time series regression techniques, see e.g. Box & Jenkins, 1976; Hamilton, 1994; Brockwell & Davis, 2003; Eurostat, 2016b.)

3.2. Compilation methodology

The European Statistical System (ESS) is a network formed by the national statistical institutes in the Member States, plus Eurostat, the European Commission's statistical office. The national accounts figures for the euro area and the EU are compiled by Eurostat using the national accounts figures submitted by the Member States' national statistical institutes. Therefore, the European aggregates, despite being calculated by Eurostat, also depend on input from the whole ESS.

([2]) For more discussion on timeliness and accuracy, see e.g. Fixler, 2007.

Regarding the national accounts main aggregates, the same approach is used for data on all variables, irrespective of the timing. This goes from the supply and use tables, sent 36 months after the end of the reference period, and which include the most detailed breakdowns, to the quarterly GDP flash estimates, sent 45 days after the quarter-end. The national accounts data sent by Member States are based on primary statistical data gathered by each of the national statistical institutes. The European aggregates are therefore produced using an indirect approach, i.e. on the basis of data gathered at Member State level, rather than by surveying the variables directly at EU level.

COMPILATION OF ANNUAL AND QUARTERLY EUROPEAN NATIONAL ACCOUNTS

The European System of Accounts 2010 (ESA 2010, see Eurostat (2013)) provides guidance with the aim of achieving consistency across Member States' national accounts data. Country data need to be comparable in order to be suitable for compiling European national accounts. The ESA 2010 transmission programme is included as an Annex to the ESA 2010 Regulation. The programme sets out the deadlines for submitting national data to Eurostat. Eurostat thus receives the required data from most Member States. Its first step in producing the European annual national accounts is to aggregate the national data to give the totals for the euro area and the EU. This first aggregation is done in current prices, and is thus a simple addition, as data in current prices are additive. To the contrary, the national figures for chain-linked volumes with moving base years are not additive. Therefore, the country data have to be totalled in the previous year's prices to obtain the aggregate variables for the euro area and the EU. Calculating the chain-linked volume figures for the European aggregate variables in year *t* requires both of the following to be available: figures for the European aggregate variables for year *t* in the previous year's prices, and figures for year *t–1* in current prices. Furthermore, after having carried out the first figures aggregation, the European-level national accounts data in current and previous year's prices have to be balanced, so that, e.g. the components of (final) supply and (final) use will add up to the same figure.

For the quarterly national accounts data transmissions, for example at 60 days after the quarter-end, some Member States have derogations allowing them to send data later than specified in the requirements. Data for the vast majority of, but not all, Member States are therefore available for compiling the European quarterly national accounts. As a result, the quarterly European level figures (in current prices and in chain-linked volumes) cannot be directly calculated due to missing data for certain countries. To solve this problem, another indirect approach is applied using the time dimension: the annual European national accounts level figures are temporally distributed to the quarters based on the quarterly national accounts figures of those Member States that sent quarterly data to Eurostat. The temporal disaggregation of the annual figures is performed using the Chow-Lin (Chow & Lin (1971)) regression-based method. A linear model is specified to describe the relationship of the quarterly GDP indicator with the annual GDP. This model can also be used for estimating quarterly changes in GDP for the most recent quarter(s) for which annual GDP is not available. The process of disaggregating the annual total over time also ensures that the quarterly figures add up to the annual ones: the non-adjusted quarterly figures add up to the non-adjusted annual figure, and the seasonally and calendar-adjusted quarterly figures add up to the calendar-adjusted annual figures.

METHOD FOR PRODUCING THE GDP T+45 FLASH ESTIMATE

The ESA 2010 transmission programme requires Member States to send the first set of quarterly national accounts data two months after the end of the quarter. In view of users' need to have data on GDP available earlier than this, the majority of Member States have also agreed to provide a flash estimate of quarterly GDP to be used for compiling the GDP t+45 flash estimates for the euro area and the EU. In most cases, Member States and Eurostat currently publish their GDP t+45 estimates on the same day.

Given that 45 days after the quarter-end is a relatively short deadline for producing macro-economic statistics, it was agreed when Member States first started providing the GDP t+45 estimates that they would only be expected to transmit the seasonally and calendar-adjusted GDP growth rates. The growth figures they send are: i) the reference quarter compared with the previous quarter (quarter-on-quarter growth); and ii) the reference quarter compared with the same quarter of the previous year (year-on-year growth). Providing information on the seasonally adjusted quarter-on-quarter growth (including calendar adjustment where relevant) was seen as the main target of the GDP t+45 flash estimates. Whenever these data are available for a Member State, Eurostat therefore uses the quarter-on-quarter national GDP growth as its primary information.

The GDP t+45 flash estimates for the euro area and the EU have been published since 2003Q1. Eurostat developed the method for producing the flash estimate and tested its performance before the first data release. The GDP t+30 flash estimate has been developed following a similar procedure, and the aim was to make the methodology for producing this estimate as similar as possible to that already being used for the GDP t+45 flash estimate. It is therefore useful to consider briefly the methodological decisions that were taken during the development of the GDP t+45 flash estimates.

The first regular GDP t+45 flash estimate for the euro area and the EU was already being prepared at that time using the Chow-Lin regression-based temporal disaggregation technique, in a manner similar to that described above. To calculate an estimate of GDP for the most recent quarter(s), the specified Chow-Lin regression model relation between quarterly GDP indicator (deviating from the annual level) and annual GDP was applied for obtaining a European quarterly GDP that is in line with the evolution at annual level.

During the testing phase of the GDP t+45 estimates for the euro area and the EU, there were five Member States producing GDP estimates before or at t+45 days: Germany (t+45), Greece (t+45), Italy (t+45), the Netherlands (t+45) and the United Kingdom (t+25). For two other large economies (France and Spain) for which GDP t+45 estimates were not yet available, the industrial production index of the reference quarter was available before t+45. For these two countries, GDP growth was therefore estimated using a regular regression model for which the inputs were one or more available indicators (typically including the industrial production index).

Next, the temporal disaggregation was prepared for quarters up to the quarter preceding the reference quarter (*Q–1*):

1. by aggregating the annual GDP of all countries up to the most recent year available; and
2. by aggregating the quarterly GDP of all countries on which it was available for quarters up to Q–1 (GDP data for Ireland and Luxembourg were missing at that time).

In order to complete the quarterly GDP indicator time series, i.e. to have a figure for the GDP indicator for the reference quarter (*Q*), the weighted average ([3]) of the quarterly growth rates of the countries providing flash estimates was considered as the growth estimate for the EU (for which there were seven such countries) and for the euro area (for which there were six country estimates). The quarterly GDP indicator series (as described in point 2 above) was then extended to include the reference quarter *Q* by applying the calculated estimates of the euro area and EU growth rates to the series.

It should be remembered that the quarterly indicator series (referred to in point 2 above) was not at the same level as the annual series due to quarterly GDP data being missing for two countries. As a last step, temporal disaggregation was therefore applied using the Chow-Lin regression model, to align the annual and quarterly GDP series.

The method for producing the GDP t+45 flash estimates described in this section was mainly used until 2013, at which time the methodology was simplified. Since 2013, the quarterly GDP levels for the euro area and the EU have been directly derived by applying the respective quarter-on-quarter growth rates for quarter *Q* to the GDP levels for quarter *Q–1*. These growth rates are obtained by aggregating the countries' GDP t+45 flash growth rates using the weights of their GDP in current price euros for year *y–1*. The main factors that made this simplification possible were, first, that countries that had not been providing quarterly GDP data in 2003 had started to do so, and second, that by 2013, the number of Member States producing GDP t+45 flash estimates had grown so that they then accounted for over 90 % of euro-area and EU GDP ([4]). The GDP t+45 flash estimates for the euro area and the EU are therefore now almost entirely based on the input from Member States, with estimation techniques and modelling no longer being used.

METHOD FOR PRODUCING THE GDP T+30 FLASH ESTIMATE

When the GDP t+30 project started its work in May 2013, four Member States — Belgium, Lithuania, Spain and the United Kingdom — were already publishing a GDP flash estimate at t+30 days or earlier (the UK publishes at t+25 days). The aim was to document the experience and best practices learnt by Member States that already publish a flash estimate at t+30 days, those that are testing or have tested its feasibility and those who would be willing to start testing the possibility. Another, more specific objective of the project was to assess whether it would be possible to publish a sufficiently reliable flash estimate of GDP for the euro area and the EU, based on the information available at t+30 days and including national estimates produced by the Member States ([5]). It should, however, be emphasised that the project did not aim to push Member States to publish national quarterly GDP data at t+30. The decision as to whether or when to publish data is left entirely to the national statistical institutes.

The choice to use an indirect approach had already been made at the beginning of the project, and before starting to produce the GDP t+30 test estimates for the euro area and the EU. The flash estimates produced by Member States, whether published or not for internal use only, and/or the indicators available at t+29, are therefore used as the main data sources, rather than indicators at euro area or EU level. The exact definition of the statistics to be provided for the GDP t+30 estimates is the same as that for the GDP t+45 estimates: seasonally and calendar-

([3]) The proportions of Member States' current price GDP (in euro) in the EU GDP and euro area GDP were used as weights.

([4]) For the t+45 flash estimates for 2013Q1, for example, these Member States accounted for 97 % of euro-area GDP and 95 % of EU GDP. In September 2015, five Member States did not provide GDP t+45 estimates: Croatia, Denmark, Malta, Slovenia and Sweden. Sweden only provides a GDP t+30 estimate for the second quarter.

([5]) More information on Member States' estimation methods particularly concerning missing data for t+30 estimates can be found in Eurostat Statistical Working Paper 'Overview of GDP flash estimation methods' (Eurostat, 2016b).

adjusted quarter-on-quarter and year-on-year growth estimates for the euro area and for the EU. When developing the method by which to produce the GDP t+30 flash estimate, the approach taken was therefore to follow the method already being used for the GDP t+45 flash estimates as closely as possible.

The process for producing the GDP t+30 flash estimates for the euro area and the EU can be presented as five steps:

- A group of Member States send their t+30 quarter-on-quarter and year-on-year seasonally and calendar-adjusted GDP growth estimates 29 days after the end of the reference quarter. This group includes Member States that already publish their t+30 estimates ([6]) and Member States that send confidential estimates to Eurostat.
- Eurostat aggregates the countries' quarter-on-quarter growth rates using the weights of their respective annual GDP in current price euros for year y−1, to produce aggregate growth rates for the euro area and the EU. (See equation A.5 in Annex A for further details. Annex A also provides a more extensive discussion on the method used for aggregating the volume growth rates of the components of GDP).
- Optional step: estimating the growth rate for a major EU country in case such country data are missing. If the GDP estimate for a major Member State is not available, its GDP growth is estimated using other available data. The other available data would typically consist of indicators such as the industrial production index, deflated retail trade and the turnover of services, for which two months' data would be available and the third month could be nowcast. Rather than relying only on a one-point estimate, the aim is to have a number of nowcast estimates, in order to be able to compare them. Nowcast estimates of national data would typically be produced in the following ways:
    1. using the national accounts weights of the country to weight the growth rates of the industrial production index, deflated retail trade and turnover of services;
    2. using all or some of the indicators mentioned above in an ADL or ARIMAX model, together with the GDP;
    3. using the economic sentiment indicator (ESI) in one of the models mentioned in point 2 ([7]).

  The above estimates can then be compared with forecasts issued by official and/or commercial forecasters. Once an estimate has been determined for a country for which data were missing, it can then be included in the aggregation of the country growth rates, along with the other GDP estimates provided directly by the Member States.
- Optional step: using the economic sentiment indicator to top-up the European estimate in a difficult rounding case. If there is a situation where rounding poses a potential problem, i.e. if the second decimal of the GDP t+30 estimate for either the euro area or the EU is close to 5 (i.e. a growth rate of ~ x.x5 %), this estimate can be topped up by using the ESI for the part of the European aggregate for which country estimates are missing. A weighted aggregate of the ESIs of the countries for which GDP estimates are missing would be calculated. This would then be entered into an ARIMAX model, together with the aggregated GDP growth

([6]) By the end of 2015, six Member States were publishing quarterly GDP estimates before or at t+30: Austria, Belgium, Latvia, Lithuania, Spain and the United Kingdom. In addition, France began to publish its GDP t+30 estimate in January 2016.

([7]) The European Commission Economic Sentiment Indicator (ESI) is a leading indicator that combines consumer confidence and business confidence indicators of European countries (European Commission (2016)). Economic sentiment indicator values are freely available for the European countries, for the euro area and European Union in Eurostat data base, http://ec.europa.eu/eurostat/data/database?node_code=ei_bssi_m_r2.

rate for previous quarters for the same set of countries, in order to nowcast the development of the reference quarter. In the most recent test estimates, Member States that were not providing GDP flash estimates accounted for approximately 10 % of EU GDP and 6 % of euro-area GDP. The contribution of these countries to the developments seen at euro-area and EU level is therefore typically minor. More extensive information about how ESI can be used is given below in a separate section.

- Once a final quarter-on-quarter GDP growth estimate has been obtained for the euro area and the EU, the GDP level estimates in chain-linked volumes for both areas for the reference quarter (Q) are then derived by applying the quarter-on-quarter GDP growth rate to the seasonally and calendar-adjusted GDP figures for quarter Q–1. The seasonally and calendar-adjusted estimate of GDP in the reference quarter (Q) is used to calculate an estimate of the year-on-year GDP growth for both the euro area and the EU.

TOP-UP OF THE T+30 GROWTH ESTIMATES WITH THE ECONOMIC SENTIMENT INDICATOR

As described above, the core of the Eurostat GDP t+30 flash estimates is based on a sample to population estimator. The growth rates of those Member States providing estimates are weighted together to form an estimate for the euro area and the EU. In the last test estimate (for 2015Q4) the coverage of these 17 countries formed 90.6 % of the EU28 GDP in the previous year; the euro area countries (11) in this group covered 94.3 % of the EA19 GDP, respectively. This section discusses in more detail, whether the development of the GDP for the missing part, the 'residual' GDP growth, could be estimated with existing economic sentiment indicator (ESI). As given above, the top-up procedure discussed below may be used when the sample to population estimator gives a value in which the (up or down) rounding to one decimal growth rate is on the border (i.e. a growth rate of ~ x.x5 %). We use the top-up of the sample to population estimator for the EU here as an example, but the same procedure applies for the EA19 as well.

In order to be able to achieve the goal, we need to divide the EU Member States into two groups. The first group, named 'exogenous' group, includes the countries already providing their estimates. The second group, named 'endogenous' group, contains the countries (9 for 2015Q4) for which estimates are missing, i.e. the 'residual' GDP growth of the EU28 not yet estimated. To test whether we could estimate the 'residual' GDP growth, the GDPs and the ESIs were aggregated both for the exogenous and endogenous group from 2003Q1 onwards.

The rationale why ESI could be used to estimate the 'residual' GDP after the sample to population estimate (exogenous group), is given in Figure 3 of Annex D. The figure plots the seasonally adjusted year-on-year GDP growth rates for the EU28, the endogenous group and the exogenous group together with the seasonally adjusted ESI index points of the same economic areas. The highest correlation with the seasonally adjusted ESI is found with the year-on-year GDP growth of the EU (correlation coefficient for the EU $r_{EU} = 0.93$ and for the euro area $r_{EA} = 0.94$ with data for 2003Q1–2015Q4), and therefore the year-on-year seasonally adjusted GDP growth of the endogenous group is estimated.

The top-up of the t+30 estimate with ESI for the EU is performed in three steps:

1. The lacking year-on-year growth rate of the last quarter for the 'residual' (or endogenous group's) GDP is estimated by using an AR(I)MAX model with aggregated national ESI used

as the exogenous input ([8]). The basis for the use of the AR(I)MAX model is given by the fact that both GDP growth and the ESI time series were tested to be stationary. In addition, the AR(I)MAX model proved to provide best nowcasts for the 'residual' GDP. The model has the following structure: 'residual' year-on-year GDP growth is regressed with the ESI of the endogenous ('residual') group together with autoregressive and/or moving average term. The number of lagged autoregressive and/or moving average terms is identified by using Bayesian (BIC, also known as Schwarz's information criteria) and Akaike's (AIC) information criteria ([9]). Alternative 'residual' GDP growth estimates for the last quarter are produced by using the ESI of the residual group, the ESI of the whole EU or the ESI of the exogenous group as the exogenous input. The final estimate for the 'residual' GDP may be an average of three to five different model estimates. An example of a specified model is given in Table 10 and Figure 4 of the Annex D.

2. The previously received sample to population year-on-year growth estimate (exogenous group) and the year-on-year growth rate estimate for the 'residual' (endogenous group) is weighed together. For 2015Q4 the weights are 90.6 % and 9.4 %.

3. The year-on-year GDP growth rate for the whole EU28 is applied to the GDP level of the same quarter a year ago for obtaining a level estimate for the quarter. The quarter-on-quarter GDP growth estimate for the EU28 is then to be calculated between the GDP level figures of the two last quarters.

4. Quality acceptance criteria

4.1. Introduction

An important part of the GDP t+30 project was to assess the feasibility of producing a flash estimate of GDP growth at 30 days after the end of the reference quarter for both the euro area and the EU ([10]). The EU Member States that participated in the project were therefore asked to provide test data in real time from 2014Q1 to 2015Q4 (with delivery of the data no later than 29 days after the end of the quarter). In addition, data were to be 'reconstructed' for all quarters going back to 2012Q1. Member States that were already producing GDP flash estimates at t+30 were able to simply supply Eurostat with these estimates. The countries that were not already producing GDP flash estimates at t+30 days were asked to calculate their national estimate retrospectively (i.e. using the data that would have been available to them 29 days after the end of the quarter).

National data were therefore available for 16 quarters (eight quarters in real time and eight quarters as a mixture of retrospective and real time results) in order to test the quality of the GDP t+30 flash estimates produced for the euro area and the EU.

([8]) On ARIMAX models, see e.g. Box & Jenkins, 1976; Hamilton, 1994; Brockwell & Davis, 2003.

([9]) The basic principle in both information criteria is that one aims to choose a model with as small residual variance as possible with as little as possible number of model parameters. The model with the lowest information criteria value is chosen. See e.g. Akaike, 1973; Schwarz, 1978; Brockwell & Davis, 2003; Stock & Watson, 2007.

([10]) In 2016, the EU is made up of 28 Member States, 19 of which are in the euro area. The test estimates for the period 2012–2015 were carried out using the 'changing composition' of the euro area and the EU, i.e. on the basis of the countries actually in the euro area and the EU at any particular point in time. The test estimates for the EU produced between 2012Q1 and 2013Q2 therefore relate to 27 Member States, whilst those for later quarters are based on 28 Member States. The number of euro area countries was 17 in 2012/2013, 18 in 2014 and 19 in 2015.

In order to assess the feasibility of producing a flash estimate of GDP growth in the euro area and the EU at 30 days after the end of the reference quarter, it was necessary to set a series of criteria that would have to be fulfilled by the results obtained by the test estimates. These 'quality acceptance criteria' and how they were determined is the subject of this section.

Section 4.2 examines the definition of quality. Section 4.3 discusses some of the considerations that were taken into account when setting the quality criteria. Section 4.4 presents the criteria to be used for assessing the quality of the test estimates.

4.2. Defining quality

Quality is a term frequently used in statistics, but its definition may vary. It can be defined in a narrow or in a broader way. In the former case, the focus is purely on statistical accuracy. In the latter, the definition may also include accessibility and timeliness, i.e. how soon statistics are released after a reference period. It is clear that quality is important to users of official statistics and that it should be a main consideration when developing statistics for publication.

The ESS recommends that quality assessment of statistics should be based on nine areas:

- relevance, assessment of user needs and perceptions;
- accuracy and reliability;
- timeliness and punctuality;
- coherence and comparability;
- accessibility and clarity;
- cost and burden;
- confidentiality;
- statistical processing.

These different aspects of quality should all be considered when developing statistics for publication. For the GDP t+30 flash estimate, however, at least in the early stages of development, the main considerations were accuracy and timeliness. These two dimensions are discussed in more details below.

The accuracy of the GDP t+30 flash estimate is probably the most important quality indicator. The aim is to produce an estimate that gives a good indication of GDP growth but at an earlier point in time than the current GDP t+45 estimate. The quality of later estimates is assessed regularly. A 'good' GDP t+30 flash estimate is therefore one that consistently produces an estimate of growth 'similar' to the ones released later (at t+45, t+65 and a later estimate such as t+100 or the estimate for current quarter when the regular t+65 estimate is published for the next quarter, i.e. at t + 3 months + 65 days ([11])). The revisions needed to the GDP t+30 flash estimates will form the basis of one of the acceptance criteria.

([11]) Eurostat published the European aggregates at t+100 for quarters up to 2014Q3. As t+100 estimates have not been published since 2014Q3, in section 5 the t+30 test estimates are compared with t+45, t+65 and with the t+65 publication concerning the following quarter — i.e. at t+3 months + 65 days after the reference quarter under review. In section 4.3, when discussing what t+30 revisions could be, the t+45 flash GDP growth rates since 2003Q1 are still compared with Eurostat t+65 and t+100 estimates which is enabled by the longer back history of t+45 estimates. Correspondingly, the US t+30 estimates average revision are compared with t+60 and t+90 US estimates.

It may also be of interest to examine the revisions needed for individual Member States' GDP flash estimates. However, assessing the quality of the Member States' estimates goes beyond the scope of this paper. The only revisions considered for developing the quality acceptance criteria will be those needed for the GDP t+30 flash estimates for the euro area and the EU ([12]).

The timeliness — the time between the end of the reference period and the publication — is set in the case of the GDP t+30 flash estimates by definition. It is important to ensure that the EU Member States involved consistently send their data by the t+29 deadline ([13]) in order to avoid any risk of the European aggregates publication being delayed.

4.3. Considerations when setting the acceptance criteria

A number of issues had to be considered before setting the quality acceptance criteria. One critical consideration was to determine if enough observations would be available to identify whether the aggregate has a sufficient quality. Using retrospective back data from Member States enabled to extend the time series but they still amount to 16 quarters of data only. Such short time span will obviously increase the risk in results' quality giving a misleading picture of the statistics' accuracy and thus an incorrect assessment.

The data for quarters from 2012Q1 to 2013Q4 were, for the majority of Member States, produced on a retrospective basis. During initial discussions about the possible methods for producing flash estimates, a number of Member States indicated that their normal methods were not entirely theoretical, and that expert judgement also played a role. Retrospective data can only be produced using a purely theoretical approach, which may potentially lead to inferior results compared to those that would have been produced had there been an element of expert judgment (although there is no way to prove this). Consideration should also be given to the fact that Member States are at the start of a process of development. It would be expected that Member States will refine their methods during the next two years, leading to a gradual improvement in results. Accordingly, later quarters may provide a better barometer for accuracy than the earlier estimates. This, however, is very much dependent on the Member States' approach to developing their methodology for flash estimates.

Table 1: Revisions to quarter-on-quarter GDP t+45 growth of the euro area (EA-12) at 65 and 100 days, in 2003Q1–2013Q4
(percentage points of GDP growth)

	Average revision	Average absolute revision
t+45 to t+65	0.00	0.02
t+45 to t+100	0.00	0.03

Source: Eurostat calculations

As stated in the previous section, the preferred measure of accuracy is the revision made to the GDP t+30 flash estimate. A useful starting point for determining what is an acceptable level of revision is to consider the revisions made to the current GDP t+45 flash estimate. Eurostat publishes data on the revisions of the euro area (EA-12 , including the 12 first members of the euro area) on its website ([14]). At the time of preparing this article, data were available from 2003Q1 to 2013Q4 for the flash (t+45), second (t+65) and third (t+100) estimates. Table 1 below shows

([12]) For instance, (Zwijnenburg, 2015) provides an update on OECD country revision performances.
([13]) In order to allow publication at t+30 days in 2016, national GDP flash data will need to be provided before 16.00h one day before.
([14]) http://ec.europa.eu/eurostat/web/national-accounts/data/other.

the average and absolute average revision made to the t+45 estimates compared to the t+65 and t+100 estimates.

Over this 11-year period, the revision performance of the GDP t+45 flash estimate for the EA-12 was very good. The revision performance of the t+45 estimates for the EU was similar to that for the EA-12, shown in table 1.

To put these data into context, and also as a starting point for fixing the criteria to be met by the GDP t+30 flash estimate, it is worth considering the revisions made to the US GDP flash estimate, which is published at t+30. These were as shown in table 2, over the period from 2003Q1 to 2014Q3:

Table 2: Revisions to quarter-on-quarter GDP t+30 growth of the United States at 60 and 90 days, in 2003Q1–2013Q4
(percentage points of GDP growth)

	Average revision	Average absolute revision
t+30 to t+60	0.02	0.13
t+30 to t+90	0.02	0.16

Source: Eurostat calculations

As can be seen from the tables 1 and 2, the US appears to revise its estimate more than has been necessary for the EA-12 estimate, with the absolute revision in particular being significantly larger. There is of course a variety of reasons why this might be the case, including the fact that the US flash estimate is produced 15 days earlier. The revisions made to the US estimates will be taken into consideration when setting the acceptance criteria for accuracy for the GDP t+30 flash estimates for the euro area and the EU.

As a minimum, the GDP t+30 flash estimate needs to be an unbiased estimate of the GDP t+45 figure. In other words, the frequency of upwards and downwards revisions should be broadly equal, and the average revision should be not statistically significantly different from 0. Performing a statistical test on such a short run of observations would, unfortunately, be difficult. Using a student's t-test on a sample of this size would result in a high probability of a 'type II error', leading to the null hypothesis being accepted on false grounds, i.e. concluding that there is no bias when there is. See Annex B for more details.

In addition, as mentioned in section 4.2, it is necessary to set deadlines by which Member States are required to submit their data. This will ensure that Eurostat always receives the data with enough time to produce the GDP estimates for the euro area and the EU by the t+30 deadline. As the first publication was scheduled for April 2016, it was considered sufficient to test whether the four quarters of 2015 are transmitted to Eurostat in time (i.e. at t+29 days).

4.4. The acceptance criteria

The previous sections have looked at the meaning of the quality of statistics and the factors that need to be taken into consideration when setting the acceptance criteria. Given the options available as criteria and the limitations of the data, the following three main acceptance criteria were adopted.

UNBIASED GDP T+30 FLASH ESTIMATE

The GDP t+30 flash estimate should be an unbiased estimate of the t+45 estimate of GDP growth, with an average revision between – 0.05 and +0.05 percentage points, and no more than 66.7 % of revisions in the same direction.

This criterion has been set to test for bias. The ideal way of testing this bias would be to perform a statistical test. As the sample is too small to perform this type of test (see Annex B), the recommendation is therefore to use the boundaries – 0.05 and +0.05. In order to avoid a scenario where a small number of large revisions in one direction (plus or minus) offset a large number of smaller revisions in the other direction, the criterion also includes a test on the percentage of revisions in each direction (tested at publication level, i.e. to one decimal place).

This criterion should be tested on all the available test data (from the first flash estimate in 2012Q1 onwards) and kept under review once the estimates start to be published.

LIMITED AVERAGE ABSOLUTE REVISION

The average absolute revision made to the t+30 estimate of GDP growth should be: i) less than or equal to 0.10 percentage points when the t+45 estimate is published; and ii) less than or equal to 0.13 percentage points when the t+65 estimate is published.

This criterion has been set to ensure that the level of revisions made to the GDP t+30 flash estimate is acceptable. In theory, the first part of criterion 1 (that the average revision must be between – 0.05 and +0.05 percentage points) could be met if there were large offsetting revisions in both directions, which would be undesirable. Criterion 2 will guard against this. When setting these boundaries, particular consideration was given to the revisions made to the US early estimates of GDP over the period 2012Q1 to 2014Q3. An overview of the revisions to early estimates of US GDP growth is attached as Annex C ([15]). Although it could be argued that the revisions made to GDP estimates for the euro area and the EU could be expected to be smaller than those for the US statistics, it should also be kept in mind that the GDP t+30 flash estimates for the euro area and the EU are developing statistics. As mentioned in section 4.3, a substantial number of Member States produced the data for the quarters between 2012Q1 and 2013Q4 on a purely theoretical basis, without having the opportunity to make use of expert judgement. It can be assumed that expert judgement would improve these early estimates, and therefore that the revisions needed should decrease over time. On the basis of these considerations, the limits that have been set are quite similar to the revisions made to US statistics.

This criterion should be tested on data from all test quarters.

SUFFICIENT COVERAGE

For the final two quarters of 2015, Eurostat needs to receive estimates from Member States whose combined GDP in the previous year accounted for 70 % of total GDP for the euro area and the EU respectively by the t+29 deadline.

This criterion will ensure sufficient coverage. It has also been set to safeguard against the publication of a GDP t+30 flash estimate before it has been proven that it will be possible to deliver the estimate consistently by this date. This criterion will reduce the likelihood of insufficient data being available to produce good quality GDP t+30 flash estimates for the euro area and the EU once data releases begin. In addition to meeting this criterion, there should also be no known reasons why the criteria would not be met for every quarter in 2016. This should be ensured by obtaining commitment from the representatives of the Member States involved.

The acceptance criteria described above have been applied to the GDP t+30 test estimates for the euro area and the EU. The next section assesses the GDP t+30 test estimates against the quality acceptance criteria.

([15]) Annex C provides data on revisions to US estimates over a slightly longer period: 2012Q1–2015Q2.

5. Results and assessment of the GDP t+30 test estimates

5.1. Introduction

The previous sections discussed the method used to prepare the GDP t+30 test estimates for the euro area and the EU and explained the quality acceptance criteria to be used for assessing these test estimates. This section will present the results for the 16 test quarters and compare these results with the GDP estimates released at t+45 days, t+65 days and at t+3 months + 65 days (the t+45 flash estimate, the 'regular' t+65 estimate and the estimate for the current quarter when publishing the 'regular' t+65 estimate for the next quarter). It analyses the revisions made to the GDP t+30 estimates at each of these dates (i.e. the difference between the GDP t+30 estimate and each of the above-mentioned later estimates) and assesses the GDP t+30 test estimates against the quality acceptance criteria.

Section 5.2 briefly presents the test estimates and section 5.3 assesses these results against the quality acceptance criteria, as set out in section 4.

5.2. Results of the GDP t+30 test estimates

To assess the feasibility of estimating the quarter-on-quarter and year-on-year GDP growth in the euro area and the EU at 30 days after the quarter-end, Eurostat, in cooperation with the participating Member States, compiled 16 GDP t+30 test flash estimates. Eight of the test estimates were produced retrospectively for past quarters from 2012 and 2013 and eight estimates were compiled in real-time during 2014 and 2015.

For the retrospectively produced quarters, Member States were asked to provide estimates calculated using only primary source data that would have been available 29 days after the end of the reference quarter. Real-time t+30 estimates could be used for the UK, Belgium, Spain and Lithuania as these countries were already publishing GDP t+30 estimates at the time.

For the real-time estimated quarters, the participating Member States were asked to provide their estimates to Eurostat 29 days after the end of the reference quarter. The data provided by the various Member States included published estimates (as mentioned above) for the UK, Belgium, Spain and Lithuania, and, from 2014Q3 onwards, also for Austria and Latvia. Using the estimates provided by all participating Member States, Eurostat then compiled the real-time GDP t+30 flash estimates for the euro area and the EU.

Table 3 below compares the test t+30 and the t+45 flash estimates for quarter-on-quarter GDP growth in the euro area. Table 4 shows the corresponding data for the EU ([16]).

5.3. Assessment of GDP t+30 test estimates

Table 3 shows that the revisions to the published data (growth rate given to one decimal place) for the GDP t+30 estimates for the euro area would have been 0.0 six times (38 %), +0.1 five times (31 %), – 0.1 four times (25 %) and – 0.2 on one occasion (6 %), meaning there would have been five downward revisions (31 %), and also five upward revisions (31 %). Table 4 shows

([16]) All the t+30 test estimates in tables 3 and 4 were produced by using only the Member States' test estimates. Here no top-up e.g. with the ESI was used.

that revisions to the published data for the EU would have been 0.0 nine times (56 %), – 0.1 four times (25 %) and +0.1 three times (19 %), meaning there would have been four downward revisions (25 %) and three upward revisions (19 %). The average revisions, when comparing figures at t+30 and t+45 given to two decimal places, are 0.00 for the euro area and – 0.01 for the EU. Both the number of revisions in each direction and the average revision to be made between the t+30 and the t+45 estimate are therefore within the limits allowed by the first quality criterion as specified in section 4.4 (an unbiased GDP t+30 flash estimate for the euro area and the EU).

As can be seen from Tables 3 and 4 above, the coverage of the GDP t+30 estimates varied between 63 % and 94 % of total GDP for the euro area, and between 65 % and 92 % of total GDP for the EU. During the final two quarters of 2015, the coverage has been at an average of 94 % of GDP for the euro area and 91 % for the EU. The test estimates therefore clearly meet the coverage criterion as specified in section 4.4 (coverage should be more than 70 % of total GDP for the euro area and the EU respectively).

The performance of the t+30 quarter-on-quarter GDP growth estimates is further analysed in Tables 5 (for the euro area) and 6 (for the EU). A similar analysis for the t+30 estimates of year-on-year GDP growth is given in Tables 7 (for the euro area) and 8 (for the EU).

The four tables above compare the GDP growth rate at t+30 days with the estimates published at t+45 days (first estimate), at t+65 days (second estimate), and after three months plus 65 days. The last growth vintage allows for assessing how the GDP t+30 estimate for the reference quarter has changed by the time of publication of the 'regular' 65 days' estimate for the next quarter. The revisions are calculated as the later estimate minus the t+30 estimate, e.g. the estimate produced at t+45 minus the estimate at t+30. The average revisions, average absolute revisions and the root of the mean squared errors (RMSQE) are given in each table.

As can be seen in Tables 5 and 6, the average absolute revisions made to the t+30 estimate of the quarter-on-quarter GDP growth rates at t+45 and t+65 were 0.06 percentage points for the euro area and 0.05 percentage points for the EU (each at both points in time). The second quality criterion given in section 4.4 requires that the difference between the t+45 and t+30 estimates be less than or equal to 0.10 percentage points and the difference between the t+65 and t+30 estimates be less than or equal to 0.13 percentage points. The GDP t+30 test estimates for the euro area and the EU therefore also fulfil the second quality criterion.

It should be noted that the average absolute revisions made to the t+30 estimates of quarter-on-quarter GDP growth for the euro area only slightly increased when the 'regular' 65 days' estimate for the subsequent quarter was available (the absolute average revision at this point being 0.08 percentage points, where it was 0.06 percentage points at 45 and 65 days after the end of the reference quarter). The average absolute revision for the EU was 0.06 percentage points on releasing the 65 days' estimate for the following quarter, which was only 0.01 percentage points higher than it had been at t+45 and t+65.

As can be seen in Tables 7 and 8, the average absolute revisions to the GDP t+30 year-on-year growth are only slightly larger, for both the euro area and the EU. The same pattern has been seen for the t+45 flash.

Overall, the GDP t+30 test estimates for the euro area and the EU met all the predefined quality acceptance criteria.

Table 3: Euro area, test t+30 and t+45 flash estimates of GDP growth
(quarter-on-quarter GDP growth rates, %; revisions to growth rates, percentage points)

Quarter	Composition euro area	Compilation level			Publication level			Coverage	
		t+30 test	t+45	Revision t+45 – t+30	t+30 test	t+45	Revision t+45 – t+30	% of GDP	Number of Member States
2012Q1	EA– 17	– 0.06	0.02	0.08	– 0.1	0.0	0.1	70	7
2012Q2	EA– 17	– 0.07	– 0.18	– 0.11	– 0.1	– 0.2	– 0.1	70	7
2012Q3	EA– 17	0.07	– 0.05	– 0.13	0.1	– 0.1	– 0.2	91	8
2012Q4	EA– 17	– 0.51	– 0.59	– 0.08	– 0.5	– 0.6	– 0.1	91	8
2013Q1	EA– 17	– 0.21	– 0.22	– 0.01	– 0.2	– 0.2	0.0	70	7
2013Q2	EA– 17	0.31	0.27	– 0.04	0.3	0.3	0.0	91	8
2013Q3	EA– 17	0.11	0.09	– 0.02	0.1	0.1	0.0	91	8
2013Q4	EA– 17	0.26	0.28	0.02	0.3	0.3	0.0	91	8
2014Q1	EA– 18	0.28	0.19	– 0.10	0.3	0.2	– 0.1	70	8
2014Q2	EA– 18	0.07	0.05	– 0.03	0.1	0.0	– 0.1	63	8
2014Q3	EA– 18	0.13	0.16	0.03	0.1	0.2	0.1	88	9
2014Q4	EA– 18	0.22	0.34	0.12	0.2	0.3	0.1	94	10
2015Q1	EA– 19	0.29	0.40	0.11	0.3	0.4	0.1	73	10
2015Q2	EA– 19	0.27	0.31	0.04	0.3	0.3	0.0	94	11
2015Q3	EA– 19	0.26	0.30	0.04	0.3	0.3	0.0	94	11
2015Q4	EA– 19	0.24	0.27	0.03	0.2	0.3	0.1	94	11

Source: Eurostat calculations

Table 4: EU, test t+30 and t+45 flash estimates of GDP growth
(quarter-on-quarter GDP growth rates, %; revisions to growth rates, percentage points)

Quarter	Composition euro area	Compilation level			Publication level			Coverage	
		t+30 test	t+45	Revision t+45 – t+30	t+30 test	t+45	Revision t+45 – t+30	% of GDP	Number of Member States
2012Q1	EU– 27	– 0.03	– 0.05	– 0.01	0.0	0.0	0.0	70	12
2012Q2	EU– 27	– 0.10	– 0.16	– 0.06	– 0.1	– 0.2	– 0.1	73	13
2012Q3	EU– 27	0.25	0.13	– 0.12	0.2	0.1	– 0.1	86	13
2012Q4	EU– 27	– 0.48	– 0.51	– 0.04	– 0.5	– 0.5	0.0	86	13
2013Q1	EU– 27	– 0.08	– 0.11	– 0.03	– 0.1	– 0.1	0.0	70	12
2013Q2	EU– 27	0.35	0.32	– 0.03	0.3	0.3	0.0	89	14
2013Q3	EU– 28	0.26	0.24	– 0.02	0.3	0.2	– 0.1	86	13
2013Q4	EU– 28	0.37	0.39	0.03	0.4	0.4	0.0	86	13
2014Q1	EU– 28	0.42	0.32	– 0.09	0.4	0.3	– 0.1	71	14
2014Q2	EU– 28	0.22	0.18	– 0.04	0.2	0.2	0.0	65	13
2014Q3	EU– 28	0.27	0.30	0.03	0.3	0.3	0.0	85	15
2014Q4	EU– 28	0.30	0.38	0.08	0.3	0.4	0.1	90	17
2015Q1	EU– 28	0.34	0.42	0.08	0.3	0.4	0.1	75	16
2015Q2	EU– 28	0.38	0.41	0.03	0.4	0.4	0.0	92	17
2015Q3	EU– 28	0.31	0.37	0.06	0.3	0.4	0.1	90	16
2015Q4	EU– 28	0.31	0.35	0.03	0.3	0.3	0.0	91	17

Source: Eurostat calculations

Table 5: GDP t+30 test estimates and revisions for the euro area
(quarter-on-quarter GDP growth rates, %; revisions to growth rates, percentage points)

	GDP growth rates				Revisions to t+30 estimates at		
	t+30	t+45	t+65	t+3m+65d	t+45	t+65	t+3m+65d
2012Q1	– 0.06	0.02	– 0.01	– 0.02	0.08	0.05	0.04
2012Q2	– 0.07	– 0.18	– 0.17	– 0.17	– 0.11	– 0.09	– 0.10
2012Q3	0.07	– 0.05	– 0.05	– 0.07	– 0.13	– 0.12	– 0.14
2012Q4	– 0.51	– 0.59	– 0.59	– 0.59	– 0.08	– 0.09	– 0.08
2013Q1	– 0.21	– 0.22	– 0.21	– 0.16	– 0.01	0.00	0.05
2013Q2	0.31	0.27	0.29	0.30	– 0.04	– 0.01	– 0.01
2013Q3	0.11	0.09	0.08	0.15	– 0.02	– 0.03	0.04
2013Q4	0.26	0.28	0.28	0.26	0.02	0.02	0.00
2014Q1	0.28	0.19	0.18	0.22	– 0.10	– 0.10	– 0.06
2014Q2	0.07	0.05	0.03	0.07	– 0.03	– 0.04	0.00
2014Q3	0.13	0.16	0.16	0.18	0.03	0.03	0.06
2014Q4	0.22	0.34	0.33	0.36	0.12	0.11	0.14
2015Q1	0.29	0.40	0.37	0.52	0.11	0.08	0.23
2015Q2	0.27	0.31	0.36	0.39	0.04	0.09	0.12
2015Q3	0.26	0.30	0.29	:	0.04	0.03	:
2015Q4	0.24	0.27	:	:	0.03		:
Average revision					0.00	0.00	0.02
Average absolute revision					0.06	0.06	0.08
Root of the mean squared errors (RMSQE)					0.07	0.07	0.10

Source: Eurostat calculations

Table 6: GDP t+30 test estimates and revisions for the EU
(quarter-on-quarter GDP growth rates, %; revisions to growth rates, percentage points)

	GDP growth rates				Revisions to t+30 estimates at		
	t+30	t+45	t+65	t+3m+65d	t+45	t+65	t+3m+65d
2012Q1	– 0.03	– 0.05	– 0.02	– 0.02	– 0.01	0.01	0.01
2012Q2	– 0.10	– 0.16	– 0.14	– 0.15	– 0.06	– 0.04	– 0.05
2012Q3	0.25	0.13	0.12	0.11	– 0.12	– 0.13	– 0.14
2012Q4	– 0.48	– 0.51	– 0.49	– 0.50	– 0.04	– 0.01	– 0.02
2013Q1	– 0.08	– 0.11	– 0.06	– 0.05	– 0.03	0.02	0.03
2013Q2	0.35	0.32	0.36	0.36	– 0.03	0.01	0.01
2013Q3	0.26	0.24	0.23	0.31	– 0.02	– 0.03	0.05
2013Q4	0.37	0.39	0.42	0.39	0.03	0.05	0.03
2014Q1	0.42	0.32	0.30	0.33	– 0.09	– 0.11	– 0.08
2014Q2	0.22	0.18	0.16	0.23	– 0.04	– 0.06	0.01
2014Q3	0.27	0.30	0.30	0.32	0.03	0.02	0.05
2014Q4	0.30	0.38	0.41	0.43	0.08	0.11	0.14
2015Q1	0.34	0.42	0.44	0.54	0.08	0.10	0.20
2015Q2	0.38	0.41	0.45	0.46	0.03	0.06	0.08
2015Q3	0.31	0.37	0.37	:	0.06	0.06	:
2015Q4	0.31	0.35	:	:	0.03	:	:
Average revision					– 0.01	0.00	0.02
Average absolute revision					0.05	0.05	0.06
Root of the mean squared errors (RMSQE)					0.06	0.07	0.08

Source: Eurostat calculations

Table 7: **GDP t+30 test estimates and revisions for the euro area**
(year-on-year GDP growth rates, %; and revisions to growth rates, percentage points)

	GDP growth rates				Revisions to t+30 estimates at		
	t+30	t+45	t+65	t+3m+65d	t+45	t+65	t+3m+65d
2012Q1	– 0.07	0.01	– 0.05	– 0.05	0.08	0.02	0.02
2012Q2	– 0.28	– 0.39	– 0.46	– 0.46	– 0.11	– 0.18	– 0.19
2012Q3	– 0.46	– 0.58	– 0.61	– 0.61	– 0.12	– 0.15	– 0.15
2012Q4	– 0.78	– 0.86	– 0.89	– 0.96	– 0.08	– 0.11	– 0.18
2013Q1	– 1.04	– 1.05	– 1.11	– 1.02	– 0.01	– 0.07	0.01
2013Q2	– 0.66	– 0.70	– 0.45	– 0.57	– 0.04	0.21	0.09
2013Q3	– 0.35	– 0.37	– 0.36	– 0.29	– 0.02	0.00	0.07
2013Q4	0.48	0.50	0.52	0.50	0.02	0.04	0.02
2014Q1	0.97	0.88	0.90	0.95	– 0.09	– 0.08	– 0.02
2014Q2	0.68	0.66	0.67	0.80	– 0.02	0.00	0.12
2014Q3	0.75	0.78	0.79	0.80	0.03	0.04	0.05
2014Q4	0.76	0.88	0.86	0.87	0.12	0.10	0.11
2015Q1	0.88	0.99	1.01	1.24	0.11	0.14	0.37
2015Q2	1.18	1.22	1.53	1.59	0.04	0.35	0.41
2015Q3	1.53	1.57	1.59		0.04	0.06	:
2015Q4	1.47	1.49	:	:	0.03	:	:
Average revision					0.00	0.02	0.05
Average absolute revision					0.06	0.10	0.13
Root of the mean squared errors (RMSQE)					0.07	0.14	0.18

Source: Eurostat calculations

Table 8: **GDP t+30 test estimates and revisions for the EU**
(year-on-year GDP growth rates, %; revisions to growth rates, percentage points)

	GDP growth rates				Revisions to t+30 estimates at		
	t+30	t+45	t+65	t+3m+65d	t+45	t+65	t+3m+65d
2012Q1	0.15	0.14	0.11	0.10	– 0.01	– 0.04	– 0.05
2012Q2	– 0.18	– 0.24	– 0.29	– 0.32	– 0.06	– 0.11	– 0.13
2012Q3	– 0.23	– 0.35	– 0.41	– 0.36	– 0.13	– 0.18	– 0.14
2012Q4	– 0.58	– 0.61	– 0.59	– 0.66	– 0.03	– 0.01	– 0.08
2013Q1	– 0.64	– 0.67	– 0.68	– 0.65	– 0.03	– 0.04	– 0.01
2013Q2	– 0.21	– 0.24	– 0.03	– 0.14	– 0.03	0.17	0.06
2013Q3	0.08	0.06	0.08	0.21	– 0.02	0.00	0.14
2013Q4	0.97	0.99	1.06	1.05	0.02	0.09	0.08
2014Q1	1.51	1.41	1.40	1.43	– 0.10	– 0.11	– 0.08
2014Q2	1.25	1.21	1.20	1.34	– 0.04	– 0.05	0.10
2014Q3	1.26	1.29	1.30	1.25	0.03	0.04	– 0.02
2014Q4	1.23	1.32	1.30	1.30	0.08	0.07	0.07
2015Q1	1.28	1.36	1.45	1.68	0.08	0.17	0.39
2015Q2	1.58	1.61	1.87	1.69	0.03	0.29	0.11
2015Q3	1.80	1.86	1.89	:	0.06	0.09	:
2015Q4	1.72	1.76		:	0.03	:	:
Average revision					– 0.01	0.03	0.03
Average absolute revision					0.05	0.10	0.10
Root of the mean squared errors (RMSQE)					0.06	0.12	0.14

Source: Eurostat calculations

6. Conclusions

The research question for this article was: Estimating reliably European GDP at 30 days after the quarter-end — is that feasible? The core of this article concerns the description of the estimation method and the assessment of the estimates for the test quarters against the predefined acceptance criteria.

The method developed for producing preliminary flash estimates of quarterly GDP for the euro area and the EU 30 days after the end of the quarter is in many respects similar to the method currently used for the euro area and EU GDP t+45 flash estimates. The main differences are that the t+30 estimate has a slightly limited country coverage and that greater use is made of preliminary country estimates (data for the third month often having to be estimated or partially estimated by the Member State at this earlier date). The methodology for producing GDP t+30 (test) flash estimates for the euro area and the EU is characterised by the following features:

- The flash estimates for the EU and the euro area are produced using the indirect approach, i.e. Member States flash estimates and/or indicators at t+29 ([17]) are used as the main data source, rather than indicators at a European level.
- All available information is used as far as possible:
  - — All GDP t+30 flash estimates that are available from the Member States for the reference quarter will be used. This is the case irrespective of whether the country's GDP t+30 estimate has been published at t+30 days or whether it is unpublished and is provided confidentially to Eurostat exclusively for the purpose of producing the GDP t+30 flash estimate of the European aggregates.
  - — Other relevant indicators may be used for a (major) Member State if its GDP t+30 estimate is not available. These alternative indicators are mainly short-term economic statistics, such as the industrial production index, deflated real trade and the turnover of services, for which two months' data are available and the third month can be nowcast. The economic sentiment indicator (ESI) may also be used. Furthermore, the forecasts issued by forecasters may be considered as secondary information in case a decision is needed on a growth estimate for a missing data (major) country.
- The decision to start publishing the GDP t+30 flash estimates for the euro area and the EU was based on the results of a testing phase carried out over the period 2012–2015. The GDP t+30 flash estimates for the euro area and the EU for the quarters from 2012Q1to 2013Q4 were calculated retrospectively, and the Member States involved were asked to use only the data that would have been available within 30 days of the end of the reference quarter when supplying their national estimates. The estimates for the quarters from 2014Q1 to 2015Q4 were calculated in real time.
- Flash estimates are available earlier than traditional estimates. There is a trade-off between timeliness and accuracy. Based on the 16 test estimates, the expected revision of the GDP t+30 growth rates at t+45 would be ± 0.1 percentage points for the euro area at publication level (i.e. considering the published figures at the two points in time, which are given to one decimal place). For the EU, the revision was 0.0 percentage points in nine of the 16 quarters

([17]) Member States are requested to transmit their national estimates to Eurostat one working day before publication of euro area and EU GDP t+30 estimates.

and ± 0.1 percentage points in the other seven. If there were to be a sudden dramatic change in economic trends, as was seen in 2008Q2 when the financial crisis broke, larger revisions than those mentioned above might be expected.

- The revisions have been tested against quality criteria set *a priori*. Over the 16 test quarters, the average revision made to the GDP t+30 flash estimates when publishing the t+45 estimates was 0.00 percentage points for the euro area and – 0.01 percentage points for the EU (the quality criterion requires these values to be between – 0.05 and 0.05 percentage points). The average absolute revision made to the GDP t+30 flash estimates when publishing the t+45 estimates was 0.06 percentage points for the euro area and 0.05 percentage points for the European Union (the quality criterion requires these values to be less than 0.10 percentage points). Comparing the GDP t+30 estimates with those produced at t+65, the average absolute revision would also have been 0.06 percentage points in the euro area and 0.05 percentage points in the European Union. The quality criterion requires these values to be less than or equal to 0.13 percentage points, a level set on the basis of the revision typically made between US t+30 and t+60 GDP estimates. The test results therefore fulfil the *a priori* determined quality criteria.

The good results of the GDP t+30 project (with test estimates showing that only limited revisions have been necessary, and coverage of all test estimates already at a higher level than achieved for the GDP t+45 estimates when they were first released) indicated that the release of GDP estimates for the euro area and the EU can be brought forward to 30 days after the end of the reference period. Therefore, Eurostat decided in December 2015 to start releasing the GDP estimates for the euro area and the EU 30 days after the quarter-end, beginning with the first quarter of 2016. The revisions to the first officially published GDP t+30 estimates for 2016Q1 continued to be small and to meet well all the quality acceptance criteria.

A large number of Member States contribute to the GDP t+30 estimates for the euro area and the EU by providing their national estimates. The launching of the preliminary GDP flash estimates around 30 days is seen as an achievement of the whole European Statistical System (ESS).

Based on all above, this paper can be concluded: Estimating reliably (enough) the European GDP at 30 days after the quarter-end has proven to be feasible.

Acknowledgements

The authors would like to thank all those who have contributed to the work of the GDP at t+30 project and to the preparation of this article. In the early part of the development work substantial support was received from Ingo Kuhnert (DG ECFIN). Particular thanks also go to Robert Doody (UK), who coordinated the task force work for section 4 and Appendix B, and to Ben Gardiner, Wojtek Szewczyk and Adam Brown (all three from Cambridge Econometrics) on their work with the possible top-up of the sample to population estimator by the economic sentiment indicator. Thanks are also due to the editor Paul Konijn and to two anonymous reviewers for their comments on the text of this article.

Last but not least the authors would like to thank all members of the task force that contributed to the GDP at t+30 project: Milen Kolev (Bulgaria), Kornel Mahlstein/Stefan Leist (Switzerland), Michal Siroky/Karel Safr/Tereza Košťáková (Czech Republic), Erich Oltmanns/Christian Mueller/Tanja Mucha (Germany), Samu Hakala (Finland), Hélène Soual/Alexis Louvel/Thomas Laurent/Philippe Gallot (France), Anikó Száraz (Hungary), Filippo Moauro (Italy), Kristina Kiriliauskaite (Lithuania), Alina Sadovska/Andris Kalnins/Gita Kinkevska (Latvia), Nico van Stokrom

(The Netherlands), Piotr Markuszewski/Pawel Markocki/Joanna Kozicka/Irmina Cerling (Poland), Carla Grosa (Portugal), Alina Predescu (Romania), Anže Podnar (Slovenia), Robert Doody (United Kingdom), Stanimira Kosekova (ECB), Wojtek Szewczyk/Shalini Mittal/Adam Brown/Ben Gardiner (Cambridge Econometrics), and Silke Stapel-Weber/Ani Todorova/Arto Kokkinen/Hans Wouters (Eurostat). The Member State representatives contributed to the work of the task force by providing information on their data sources, by developing methods for producing early estimates and/or by providing quarterly national GDP estimates to be used for producing the test estimates for the euro area and the EU. Without their contributions, the project could not have been concluded successfully.

References

Akaike, H. (1973), 'Information theory and an extension of the maximum likelihood', in B. N. Petrov, & F. Csaki, *2nd International Symposium on Information Theory* (pp. 267-281). Akademiai Kiado, Budapest.

Box, G., & Jenkins, G. (1976), *Time Series Analysis: Forecasting and control, Revised edition,* San Francisco: Holden Day.

Brockwell, P. J., & Davis, R. A. (2003), *Introduction to Time Series and Forecasting,* New York: Springer Texts in Statistics. Springer, 2nd edition.

Chow, G., & Lin, A. (1971), 'Best linear unbiased interpolation, distribution and extrapolation of time series by related series', *The Review of Economics and Statistics, 53*, 372–375.

European Commission (2016), *The joint harmonised EU programme of business and consumer surveys, user guide,* http://ec.europa.eu/economy_finance/db_indicators/surveys/method_guides/index_en.htm.

Eurostat (2003), 'Flash estimation of the quarterly Gross Domestic Product of the euro-zone and the European Union', *Eurostat methodology*, section 5, Publications Office of the European Union, Luxembourg

Eurostat (2013), *Handbook on quarterly national accounts,* Publications Office of the European Union, Luxembourg.

European Statistical System (ESS) (2013), *The ESS Vision 2020,* http://ec.europa.eu/eurostat/documents/10186/756730/ESS-Vision-2020.pdf/8d97506b-b802-439e-9ea4-303e905f4255.

Eurostat (2013), *European System of Accounts (ESA 2010), Publications Office of the European Union,* Luxembourg, see also http://ec.europa.eu/eurostat/web/esa-2010.

Eurostat (2016a), 'Euro Area and European Union GDP flash estimates at 30 days', *Eurostat Statistical working papers*, Publications Office of the European Union, Luxembourg.

Eurostat (2016b), 'Overview of GDP flash estimation methods', *Eurostat Statistical working papers,* Publications Office of the European Union, Luxembourg.

Fixler, D. (2007), 'Timeliness and Accuracy', in M. Boumans, *Measurement in Economics, A Handbook,* London: Elsevier, Academic Press.

Hamilton, J. D. (1994), *Time Series Analysis,* Princeton University Press.

Schwarz, G. E. (1978), 'Estimating the dimension of a model', *Annals of Statistics 6 (2)*, 461–464.

Stock, J. H., & Watson, M. W. (2007), *Introduction to Econometrics, 2nd edition,* Pearon Education Inc.

Zwijnenburg, J. (2015), 'Revisions of quarterly GDP in selected OECD countries', *OECD Statistics Brief, July 2015* (No. 22).

Annex A: Aggregation of the country growth rates

This Annex explains how the Member States' chain-linked volume growth rates are aggregated to compile the EU and euro-area growth rates.

The following notation will be used (reflecting the notation used in the Handbook on quarterly national accounts (Eurostat 2013, pp. 190-191)):

- The aggregate variable *X* at quarter *c* of year *y* in current prices is denoted by $X_{CP}^{c,y}$, in prices of the previous year by $X_{PPY}^{c,y}$, and for chain-linked volumes by $X_{CV}^{c,y}$.
- The implicit price deflator of the chain-linked volume indices is denoted by $P_x^{c,y} = \frac{X_{CP}^{c,y}}{X_{CP}^{c,y}}$.
- Measures indexed by a year only denote annual data, e.g. X_{CP}^{y} (the current price value of the aggregate *X* in year *y*) and $P_x^y = \frac{X_{CP}^y}{X_{CV}^y}$ is the implicit price deflator of *X* in the same year.
- Chain-linked volume indices have to be referenced to a value in a particular period, usually at 100 or the current price value in the reference year, thus $X_{CV}^0 = 100$ or $X_{CV}^0 = X_{CP}^0$.
- The annual chain-linked volume growth between *y* and *y*-1 can be calculated as $\frac{X_{CV}^y}{X_{CV}^{y-1}} = \frac{X_{PPY}^y}{X_{CP}^{y-1}}$ (where the latter is the Laspeyres type volume index $L_Q^y = \frac{X_{PPY}^y}{X_{CP}^{y-1}}$).
- An annual chain Laspeyres volume index referenced to the current price value in year 0 can therefore be expressed as

$$X_{CV}^y = \prod_{m=1}^{y} \frac{X_{PPY}^m}{X_{CP}^{m-1}} X_{CP}^0 .$$

In the case of annual overlap quarterly chain-linking method (used by Eurostat and the common practice in most EU countries) the quarterly previous year price figures refer to average prices of the previous year. This means that the quotient of a variable given in the previous year's prices and the corresponding chain-linked volume estimate is equal to the annual implicit price deflator of the previous year

$$\frac{X_{PPY}^{c,y}}{X_{CV}^{c,y}} = \frac{X_{PPY}^{y}}{X_{CV}^{y}} = \frac{X_{CP}^{y-1}}{X_{CV}^{y-1}} = P_X^{y-1}.$$

The following section first explains the background to the aggregation of annual country (or component) growth rates to give the annual growth rate of the aggregate. It then discusses the use of the annual current price weights for the year y-1 in aggregating Member States' chained quarter-on-quarter GDP growth rates to produce aggregates for the euro area and the EU.

Annual case

The relative growth rate of the chain-linked volumes (CV) for an individual country *i* is calculated by

A.1 $\frac{x^{y}_{CV,i} - x^{y-1}_{CV,i}}{x^{y-1}_{CV,i}} = \frac{x^{y}_{PPY,i} - x^{y-1}_{CP,i}}{x^{y-1}_{CP,i}}$. Multiplying both sides of the equation by $x^{y-1}_{CP,i}$ gives

A.2 $x^{y-1}_{CP,i} . \left(\frac{x^{y}_{CV,i}}{x^{y-1}_{CV,i}} - 1 \right) = x^{y}_{PPY,i} - x^{y-1}_{CP,i}$. This equation shows that, when multiplying the annual chain-linked volume relative growth by the current price figure at *y-1*, the result is the absolute volume growth between year y in the previous year's prices and year *y-1* in current prices. Dividing both sides of formula A.2 above for country i by the same formula expressed for the European aggregate gives

A.3 $$\frac{x^{y-1}_{CP,i} . \left(\frac{x^{y}_{CV,i}}{x^{y-1}_{CV,i}} - 1 \right)}{X^{y-1}_{CP} . \left(\frac{X^{y}_{CV}}{X^{y-1}_{CV}} - 1 \right)} = \frac{x^{y}_{PPY,i} - x^{y-1}_{CP,i}}{X^{y}_{PPY} - X^{y-1}_{CP}}.$$

Now, multiplying both sides of the above by the chain-linked volume growth rate of the aggregate, i.e. by $\left(\frac{X^{y}_{CV}}{X^{y-1}_{CV}} - 1 \right)$, gives:

A.4 $$\left(\frac{x^{y-1}_{CP,i}}{X^{y-1}_{CP}} \right) . \left(\frac{x^{y-1}_{CV,i}}{x^{y-1}_{CV,i}} - 1 \right) = \left(\frac{x^{y}_{PPY} - x^{y-1}_{CP,i}}{X^{y}_{PPY} - X^{y-1}_{CP}} \right) . \left(\frac{X^{y-1}_{CV}}{X^{y-1}_{CV}} - 1 \right).$$

The right-hand side of equation A.4 shows that the chain-linked relative volume growth rate of the aggregate is multiplied by a term in which the absolute volume growth of component x_i is expressed as a proportion of the absolute volume growth of the aggregate *X*. The result is thus the contribution made by the volume growth of component x_i to the relative volume growth of the aggregate *X*.

The left-hand side of the equation A.4 shows that the contribution made by the same component x_i to the relative volume growth of the aggregate *X* can also be calculated by multiplying the relative growth rate of the component x_i by the current price level of the component x_i over the current price level of the aggregate *X* in year y-1.

Quarterly case

The Handbook on quarterly national accounts (Eurostat 2013, pp. 200-204) discusses different methods for calculating the growth contributions of the components to the quarter-on-quarter chain-linked volume growth of the aggregate variable, such as GDP. The method discussed above, applied to the quarter-on-quarter growth rate of the component x_i of the aggregate *X*, is called the weighted growth rates (WGR) method. In our example, using annual current prices as the weights, the formula is:

A.5 $$Contrib.WGR^{A}(x_i,X)^{c,y} = \left(\frac{X^{y-1}_{CP,i}}{X^{y-1}_{CP}}\right).\left(\frac{X^{c,y}_{CV,i} - X^{c-1,y}_{CV,i}}{X^{c-1,y}_{CV,i}}\right) = \left(\frac{X^{y-1}_{CP,i}}{X^{y-1}_{CP}}\right).\left(\frac{X^{c,y}_{CV,i}}{X^{c-1,y}_{CV,i}} - 1\right).$$

In the case of European aggregates, when the Laspeyres chained- volume index is used, together with the annual overlap quarterly chain-linking methodology, the current price figure for year y-1 forms the moving base period for the volume calculation. A natural choice is therefore to use the weighted growth rates method described above with annual weights. It is worth noting that, as mentioned in the Handbook on quarterly national accounts (Eurostat 2013, p. 204), the attractions of the WGR method are that it is simple to use and easy to understand and it is independent of the quarterly linking method. The Handbook on quarterly national accounts (Eurostat 2013, pp. 203-204) suggests it as one method for calculating the contribution made by the component x_i to the quarter-on-quarter volume growth of the aggregate X. This method is used in calculating the t+30 and t+45 flash estimates of the quarterly GDP for the euro area and the EU. The quarter-on-quarter growth estimates provided by the Member States are aggregated using the equation A.5 above.

It is worth noting, though, that the contributions calculated using this method are not exactly additive in the quarterly case. This is most visible in the case of the quarter-on-quarter growth rate between Q1 and Q4, as the former has year *y-1* as a base year and the latter year *y-2*. The Handbook on quarterly national accounts (Eurostat 2013, pp. 200-204) discusses another additive volume data method, which includes a correction for the first quarter, the effect of which is that the quarterly contributions will add up exactly to the same contribution of the component in the annual data. Finally, the Handbook discusses how a correction factor can be introduced into the WGR method (using annual or quarterly weights), in order to obtain results for the contributions of the components that are exactly additive:

A.6 $$Contrib.WGR^{A}(x_i,X)^{c,y} . \left(\frac{X^{c-1,y}_{CV,i} / X^{c-1,y}_{CV}}{X^{y-1}_{CV,i} / X^{y-1}_{CV}}\right) = Contrib.AVD(x_i,X)^{c,y}.$$

In practice, when aggregating the country growth rates to give the euro-area and EU totals, the size of the correction that would need to be made to the aggregate relative growth rate is quite small, e.g. between 2010Q1 and 2015Q2, the average absolute value of the correction was 0.0004 %. The proportions of the Member States' GDPs in the aggregate GDP are rather stable, both between quarters and between years.

Annex B: Sample size for testing revisions for bias

Introduction

This Annex discusses the effect on the results of a student's t-test in the case of having a small number of observations. It examines the method used by the UK to test whether the revisions to economic indicators would be biased. Finally, the Annex considers whether the limited number of observations available from the testing phase of the GDP t+30 flash estimate rules out the use of a statistical test in assessing the possible bias in revisions.

Terminology

- Type 1 error: rejecting the null hypothesis when it is true (i.e. saying there is bias when there is not).
- Significance level: the probability of a type 1 error (i.e. the probability of saying there is bias when there is not).
- Type 2 error: accepting the null hypothesis when the alternative hypothesis is true (i.e. saying there is no bias when there is).
- Power: 1 minus the probability of a type 2 error (i.e. the probability of correctly identifying bias).

Method

The current method for testing for bias in the revisions to early estimates of economic indicators in the UK is to use t-tests to test for bias. This method was recommended in the paper "An examination of the way CSO [the Chief Scientist Office] tests for bias in the initial estimates of major economic indicators; some suggested improvements". (Dr C. Chatfield, CSO Research Project, University of Bath).

For the annual series, the t-test assumes that revisions are independent.

The test statistic is calculated as $t = \frac{\bar{x}}{s_{\bar{x}}}$, where:

$\bar{x}$ is the average revision

$s_{\bar{x}}^2 = \frac{s^2}{n}$, where:

s^2 is the population variance in the revisions (estimated using the sample)

n is the sample size.

The test statistic is compared to a t-distribution with n degrees of freedom. The significance level used is 0.05, meaning that the probability of a type 1 error is 0.05 or 5 %.

For the quarterly series, an adjusted t-test is used, which assumes that revisions follow an AR(1) process (meaning that the revision at one time point depends on the revision made at the previous time point).

The test statistic is calculated as $t = \frac{\bar{x}}{s_{\bar{x}}^{*2}}$, where:

$\bar{x}$ is the average revision.

$s_{\bar{x}}^{*2} = \frac{s^2(1+\alpha)}{n(1-\alpha)}$, where:

s^2 is the population variance in the revisions (estimated using the sample)

n is the sample size

α is an estimate of the correlation between revisions at time t and time t-1.

The test statistic is compared to a t-distribution with $n\frac{(1-\alpha^2)}{(1+\alpha^2)}$ degrees of freedom.

The significance level used is 0.05, meaning that the probability of a type 1 error is 0.05 or 5 %.

Impact of reducing the sample size on type 1 errors

If the sample size were to be reduced, then the critical value used to compare the test statistic should be changed to reflect the change in the degrees of freedom. If the significance level remained unchanged, then the probability of a type 1 error would also remain unchanged.

The critical value for a two-sided t-test with a significance level of 0.05 and a sample size of 14 is 2.145.

Impact of reducing the sample size on the power of the test

The problem created by a smaller sample size is that it reduces the power of the test. The power of the test depends on the sample size, the effect size and the significance level.

The effect size puts into context how large the average revision is in relation to the spread of the revisions. It is calculated as the mean of the revisions divided by the standard deviation of the revisions (i.e. the t-test statistic divided by the square root of the number of observations).

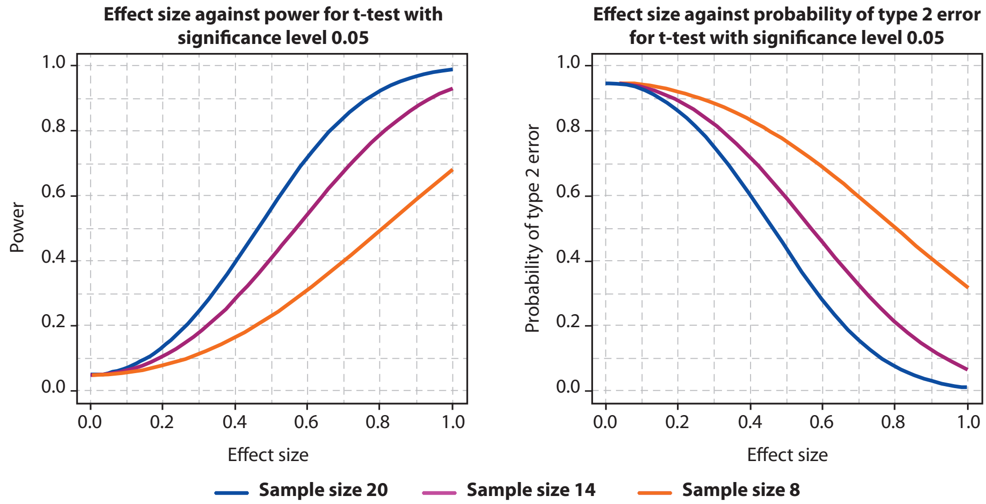
$d = \frac{\bar{x}}{s}$, where:

$\bar{x}$ is the average revision

s^2 is the population variance in the revisions (estimated using the sample)

Figure 1 below shows the power of the test and the probability of type 2 errors for different effect sizes when using a t-test at significance level 0.05 for sample sizes of 14 (the number of quarters likely to be available for testing) and 20 (i.e. five complete years of test quarters) and 8. The sample size of 8 has been chosen because it would be the degrees of freedom used in the modified t-test if alpha was 0.5 and the sample size was 14. If alpha is close to 0, then the power will be close to that of a test with sample size 14. If alpha is between 0 and 0.5, then the power will be between the power of the tests carried out with sample sizes 8 and 14. If alpha is greater than 0.5, then the power will be lower than the power of the test carried out using a sample size of 8.

Figure 1: **Power of the t-test for given sample sizes and effect sizes and a significance level of 0.05.**



Source: UK Office for National Statistics

To get an idea of the effect size in the current revisions hypothesis test, effect sizes based on the current revisions analysis are:

- Annual comparisons:
  - Preliminary GDP estimate (at t+25d) compared with the third GDP estimate (at t+89d): – 0.7 (same as 0.7),
  - Second GDP estimate (t+58 d) compared with the third GDP estimate (t+89 d): – 0.3 (same as 0.3),
  - Third GDP estimate (t+89 d) compared with the estimate after 3 years: 0.4;
- Quarterly comparisons:
  - Preliminary GDP estimate (at t+25d) compared with the third GDP estimate (at t+89d): 0.3,
  - Second GDP estimate (t+58 d) compared with the third GDP estimate (t+89 d): 0.1,
  - Third GDP estimate (t+89 d) compared with the estimate after 3 years: (same as 0.1).

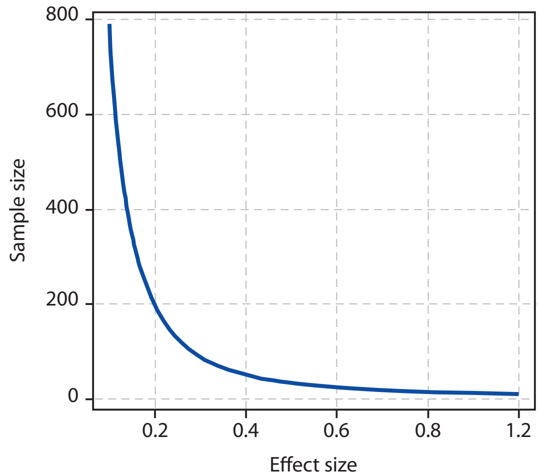
To get an idea of alpha in the current revisions hypothesis test, alpha based on the current revisions analysis are:

- Annual comparisons:
  - Preliminary GDP estimate (at t+25d) compared with the third GDP estimate (at t+89d): – 0.1 (same as 0.1),
  - Second GDP estimate (t+58 d) compared with the third GDP estimate (t+89 d): – 0.2 (same as 0.2),
  - Month 3 estimate (t+89 d) compared with the estimate after 3 years: 0.0.

- Quarterly comparisons:
  - — Preliminary GDP estimate (at t+25d) compared with the third GDP estimate (at t+89d): 0.0,
  - — Second GDP estimate (t+58 d) compared with the third GDP estimate (t+89 d): – 0.5 (same as 0.5),
  - — Third GDP estimate (t+89 d) compared with the estimate after 3 years: (same as 0.1) 0.4.

A common threshold for acceptable power is 80 %. Figure 2 shows the sample size that would be required to achieve a power of 80 % for given effect sizes and a significance level of 0.05.

Figure 2: Sample size required to achieve 80 % power for different effect sizes and a significance level of 0.05.



Source: UK Office for National Statistics

RECOMMENDATION

The initial results suggest that the GDP t+30 flash estimate has a low effect size. The recommendation is therefore that a student's t-test would not be used as a quality acceptance criterion, due to the risk of a type II error. Instead, it is recommended that appropriate boundaries for the size of the revision should be set.

Annex C: US t+30, t+60 and t+90 estimates of quarter-on-quarter GDP growth, and the associated revisions to the t+30 estimates

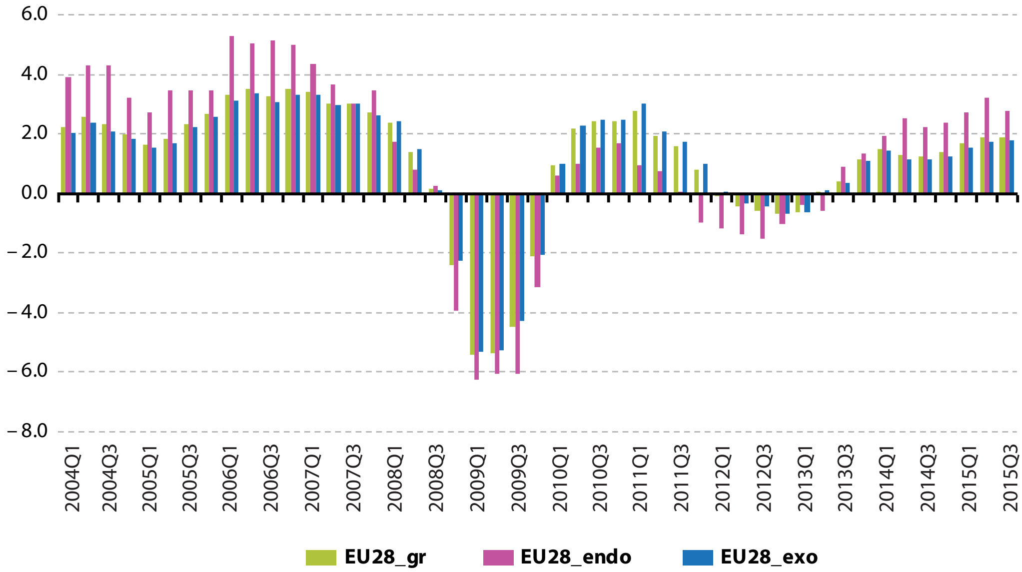
Table 9: **GDP t+30 estimates and revisions for the US**
(quarter-on-quarter GDP growth rates, %; revisions to growth rates, percentage points)

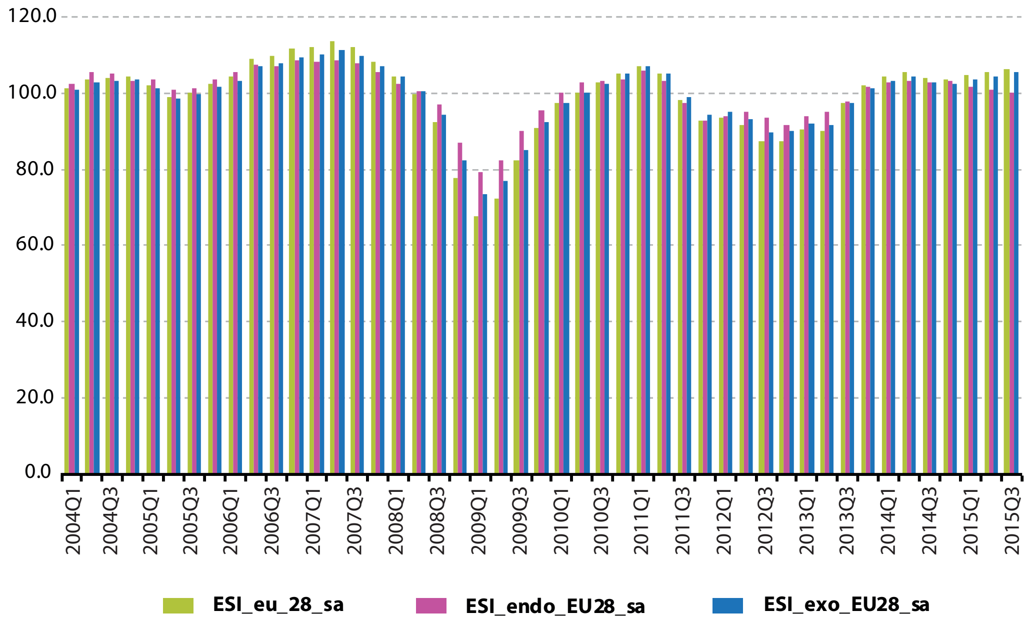
	GDP growth rates			Revisions to t+30 estimates at	
	t+30	t+60	t+90	t+60	t+90
2012Q1	0.55	0.46	0.46	– 0.08	– 0.08
2012Q2	0.38	0.43	0.42	0.05	0.04
2012Q3	0.50	0.66	0.77	0.16	0.27
2012Q4	– 0.04	0.03	0.09	0.07	0.13
2013Q1	0.62	0.59	0.44	– 0.03	– 0.18
2013Q2	0.71	0.92	0.91	0.21	0.20
2013Q3	0.71	0.89	1.02	0.18	0.31
2013Q4	0.80	0.59	0.65	– 0.21	– 0.15
2014Q1	0.03	– 0.25	– 0.74	– 0.27	– 0.77
2014Q2	0.97	1.03	1.13	0.05	0.16
2014Q3	0.88	0.96	1.22	0.08	0.34
2014Q4	0.65	0.54	0.55	– 0.11	– 0.10
2015Q1	0.06	– 0.19	– 0.04	– 0.25	– 0.10
2015Q2	0.58	0.91	0.97	0.33	0.39
Average revision				0.01	0.03
Average absolute revision				0.15	0.23
Root of the mean squared errors (RMSQE)				0.18	0.29

Note: NB: US publishes annualised quarter-on-quarter growth rates. For comparison here the annualised growth rates have been transformed back to plain quarter-on-quarter growth rates which are comparable with the ones given for the euro area and the EU.
Source: US Bureau of Economic Analysis

Annex D: The EU28 GDP growth and economic sentiment indicator

Figure 3: The year-on-year seasonally adjusted GDP growth rates (upper graph) and the seasonally adjusted economic sentiment index (ESI, lower graph) for the EU, for the endogenous group and the exogenous group





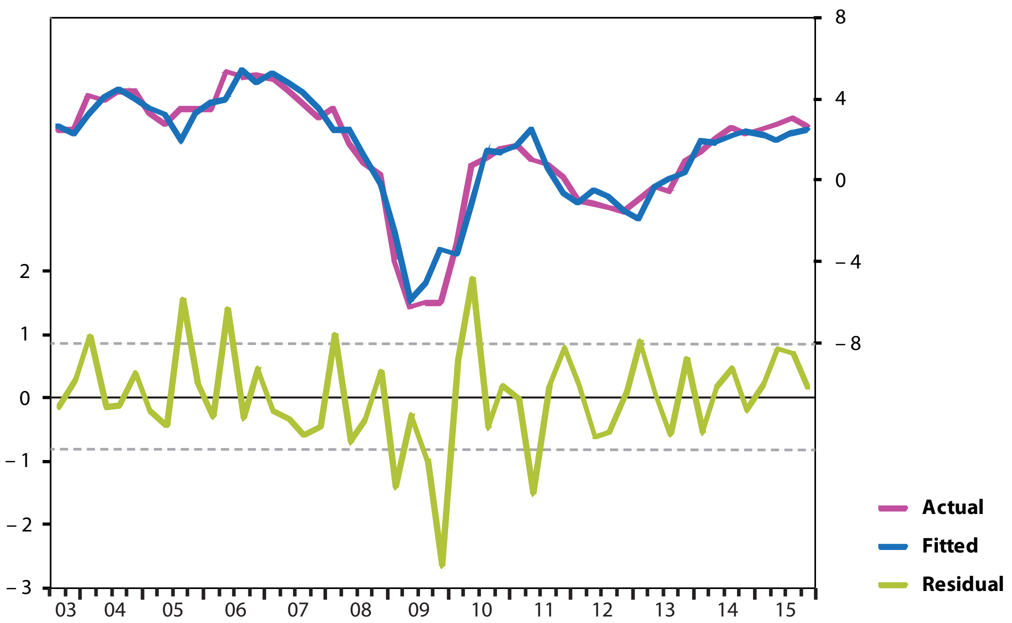
Source: Author's calculations

Table 10: An ARIMAX model for estimating the 'residual' GDP for the EU28 in 2015Q4

Dependent Variable: EU28_ENDO				
Method: Least Squares				
Date: 01/28/16 Time: 18:21				
Sample (adjusted): 2003Q2 2015Q3				
Included observations: 50 after adjustments				
Convergence achieved after 10 iterations				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
ESI_ENDO_EU28_SA	0.292212	0.038018	7.686121	0.0000
C	-28.00812	3.865700	-7.245292	0.0000
AR(1)	0.823405	0.077576	10.61418	0.0000
R-squared	0.925903	**Mean dependent var**		1.416924
Adjusted R-squared	0.922750	**S.D. dependent var**		2.867556
S.E. of regression	0.797004	**Akaike info criterion**		2.442210
Sum squared resid	29.85511	**Schwarz criterion**		2.556931
Log likelihood	-58.05525	**Hannan-Quinn criter.**		2.485897
F-statistic	293.6533	**Durbin-Watson stat**		2.026317
Prob(F-statistic)	0.000000			
Inverted AR Roots	0.82			

Source: Author's calculations

Figure 4: The fit of the model against the original values and the model residuals



Source: Author's calculations

4 Consistency between national accounts and balance of payments statistics

ROBERT OBRZUT (1)

Abstract: In order to achieve full consistency of the methodological standards in national accounts and balance of payments statistics, an assessment of this consistency between the two appears justified. This paper presents the results of a study on consistency between the quarterly sector accounts and quarterly balance of payments statistics, and is complemented with aspects of internal consistency, which address statistical discrepancies in balance of payments statistics. For better interpretation, explanations made available from a survey exercise in 2015 conducted by Eurostat ('the BOP/ROW survey') were incorporated. It will be shown that although methodological standards are consistent, both statistics are still exposed to discrepancies to a greater or lesser extent. This rigid development results from the heterogeneous nature of compilation processes in European statistics and from the lack of sharing information and data sources at national as well as international level. At the same time it also can be shown that in some specific areas a better coordination among national compilers has actually improved consistency. The paper further points at the conceptual differences in both statistics, which also contribute to this situation, and would require to be addressed by international bodies for follow-up. It is concluded that by maintaining strictly autonomous compilation practices in European statistics, European compilers will hamper their ability to adapt to the challenges of compiling statistics in a globalised world. Further work will follow as regards the analysis of external asymmetries and their impact on national consistency.

Keywords: balance of payments, sector accounts, international trade, data consistency.

JEL codes: E01, F40.

(1) Eurostat, unit C5, Integrated global accounts and Balance of Payments.

1. Introduction

The new standard methodology BPM6/ESA 2010 requires full consistency (2). An assessment of the consistency between the balance of payments (BOP) and the rest-of-the-world (ROW) account in the national accounts thus appears justified, in order to conclude how far these two statistics have been reconciled with each other over the past years. Since September 2014 the statistical compilation practice of European BOP statistics has migrated to the new standard BPM6 (3), following the earlier implementation of the ESA 2010 standard in national accounts statistics (4).

This analysis complements a survey exercise launched by Eurostat earlier in 2015 ('the BOP/ROW survey') where Member States provided explanations for their observed discrepancies, regarding data from 1999 to 2013 for both the current/capital and the financial account. It aims at the current state of consistency between BOP/International Investment Positions (IIP) and national accounts statistics based on an analysis of the most recent data available from the corresponding production rounds in both statistics during April 2016. It starts with a description of the parameters of this exercise and the underlying principles of reconciliation, and hence elaborates on the subject in regard to the current/capital and the financial account. Then, it concludes with the technical challenges experienced, the perceived gaps between the existing concepts and the availability of data, and the organisational setting of data compilation processes for the two statistics.

The analysis is based on quarterly data for (mostly gross) transactions from the BOP current/capital and the financial account (QBOP) and the ROW account of the quarterly sector accounts (QSA). The time span of the analysis (2010–2015) was subjected to the availability of time series according to the BPM6 standard.

2. Parameters of the data confrontation

The analysis is based on quarterly data from the BOP current/capital and the financial account (QBOP) and the ROW account from the quarterly sector accounts (QSA). The time span of the analysis (2010–2015) was subjected to the availability of time series according to the BPM6. Currently data as per 2010 are available from all Member States. For easier interpretation all quarterly data have been annualised. The analysis focused on (mostly gross) transaction data, but is extended to positions data in the financial account in order to develop further conclusions. A confrontation of data was applied to the current/capital and the financial account of BOP separately with the corresponding items of the ROW account (5) directly after production date, in order to minimize vintage and revision effects (6). For missing data no measuring of consistency could be applied (7). Discrepancies were measured as total absolute differences in gross transactions of the BOP current account items and the corresponding ROW items. These discrepancies were evaluated for each component account for the EU-aggregate, which is calculated and

(2) BPM6 — Appendix 7, ESA 2010 — Chapter 18.
(3) Balance of Payments and International Investment Position Manual, 6th edition (BPM6), International Monetary Fund (2009).
(4) European System of Accounts 2010 (ESA 2010), see Eurostat (2013).
(5) Emphasising the sign convention of the ROW account.
(6) With revision effects having either a smoothening or escalating impact on the measured discrepancies.
(7) Luxembourg does not report data on primary and secondary income, as well as the capital account to the QSA. To a lesser extent gaps applied also to the QSA data of Malta (secondary income account). Quarterly time series for 2015 were still incomplete in Belgium and Bulgaria at the time of this analysis.

published by Eurostat, and national data for the 28 Member States, which were reported to Eurostat. Where deemed useful, national measures are also presented in relation to their GDP.

3. Applied principles of reconciliation

3.1. Reconciling the non-financial accounts

The data confrontation of the BOP accounts with the respective ROW account of the QSA requires a clear convention on their reconciliation, in order to gain a common understanding of discrepancies. Based on the available data, the non-financial accounts appear directly comparable (Table 1).

Table 1: Reconciling current/capital (BOP) and the ROW non-financial account (QSA)

BOP component	ROW NA item	Description
Goods	P61	Exports of goods
	P71	Imports of goods
Services	P62	Exports of services
	P72	Imports of services
Primary income	D1	Compensation of employees
	D2	Taxes on production and imports
	D3	Subsidies
	D4	Property income
Secondary income	D5	Current taxes on income & wealth
	D6	Social contributions and benefits
	D7	Other current transfers
	D8	Adjustment for the change in pension entitlements
Capital account	D9	Capital transfers
	NP	Acquisition less disposal of non-financial non-produced assets

Note: BOP and NA items according to BPM6 and ESA 2010.

The credit flows of the BOP current account with the rest of the world directly compare to transactions payable in the ROW account, debit flows of the BOP current account compare to transactions receivable in the ROW account (equations 1 and 2). As one capital account item is only reported on a net basis in the QSA ([8]), this study compares only balances under this item (equation 3).

(1) $BOP_{Credit} = ROW_{Paid}$

(2) $BOP_{Debit} = ROW_{Received}$

(3) $BOP_{(Credit-Debit)} = ROW_{Net}$

In our study we focused on the current account components, so that consistency issues related to the respective items, which would be otherwise offset in total accounts, could be identified.

([8]) Gross acquisitions/disposals of non-produced non-financial assets (NP).

3.2. Reconciling the financial accounts

In the BOP financial account in principle net transactions in assets compare to net transactions in liabilities of the ROW account, while vice versa net transactions in liabilities compare to net transactions in assets to the ROW account. As a consequence net lending/net borrowing in the BOP financial account compares to net acquisition of assets/net incurrence of liabilities in the ROW account as follows:

(4) $BOP_{Net} = BOP_{Assets} - BOP_{Liabilities}$

(5) $ROW_{Net} = ROW_{Liabilities} - ROW_{Assets}$

(6) $BOP_{Net} = ROW_{Net}$

The consequence of (6) is that both net values should follow the same sign convention. The same reasoning applies to positions data, when comparing the IIP with the financial balance sheet of the ROW account.

Regarding the financial account however, we encountered also conceptual differences which put an additional strain on the correct selection of BOP breakdown data. Appendix 7 in the BPM6 establishes a direct linkage to the SNA 2008 accounts for the rest of the world (ROW)(9). Table A7.1 propagates full comparability, consolidating the two concepts - the functional category of the BPM6 and the financial instrument category of the ESA 2010. Similarly, ESA 2010 — Table 18.14 (10) depicts the possible reconciliation between the two categories. Table 2 below illustrates the bridging in concepts between the two standards in a simplified form.

Table 2: Reconciling ROW financial account (QSA) and financial account (BOP)

NA item / BOP component		Direct investment	Portfolio investment	Financial derivatives	Other investment	Reserve assets
Monetary gold and SDRs	F.1				x	x
Currency and deposits	F.2	x			x	x
Debt securities	F.3	x	x			x
Loans	F.4	x			x	x
Equity	F.5	x	x		x	x
Insurance	F.6	x			x	
Financial derivatives	F.7			x		x
Trade credits and other accounts	F.8	x			x	

Note: Simplified version for illustration purposes.

Source: ESA 2010, Table 18.14.

Two patterns appear noteworthy in this context: firstly, the financial instrument categories of ESA 2010 have a multidimensional (one-to-many) relationship to the functional categories of the BPM6. For example, the comparison of ESA 2010 category F.2 (Currency and deposits) would require the corresponding BOP items recorded under direct investment, other investment and reserve assets. Secondly, some BPM6 components (e.g. direct investment debt instrument) are of composite character, requiring availability of all breakdowns, before becoming comparable. Omission of the above could result in imprecisions with considerable overestimation of discrepancies, depending on the relevance of the respective items in the compiling economy.

(9) System of National Accounts (SNA) 2008, see United Nations et al. (2009).
(10) Also: Table 26.6 (SNA 2008).

3.3. Limitations of the data comparison

The BPM6 Annex 9 shows the standard components of BOP, which can be considered as the fundamental representation of the BOP data framework. The aforementioned tables A7.1 of the BPM6 and table 18.14 of ESA 2010 incorporate however some items, which are not mentioned in the standard representation. While the functional category of direct investment proposes a breakdown for equity/investment fund shares and debt instruments, all direct investment activities relating to the insurance sector and in regards to trade credits appear omitted. This concerns in particular the following activities conducted by direct investment enterprises:

- non-life insurance technical reserves,
- life insurance and annuity entitlements,
- claims of pension funds on pension managers and
- provisions for calls under standardised guarantees; as well as
- trade credits and advances,
- other accounts receivable/payable.

According to the BPM6, direct investment debt instruments (in BPM5: other capital) consist of operations in Special Drawing Rights (SDRs), currency and deposits, debt securities, loans (incl. intercompany loans), insurance technical reserves, etc. (BPM6, paragraph 5.31). The composite character of debt instruments recorded under direct investment (FL) in BPM6 consequently does not allow for a straightforward comparison, for example with the ROW component F.3 (Debt securities), or F.4 (Loans). The definition of debt instruments in direct investment incorporates not only negotiable securities (debt securities, certificates) but also intercompany loans, technical reserves, etc. with a reference to different financial instrument categories of ESA 2010. A breakdown for direct investment debt instruments is currently not part of the standard representation of the BPM6. A similar reasoning could apply to reserve assets, as they consist of operations in monetary gold, SDRs, currency and deposits (incl. interbank positions), debt securities, loans, equities/investment fund shares and financial derivatives. However, in the latter case the BPM6 standard representation considers them sufficiently, in contrast to debt instruments.

Due to the different objectives in both statistics, the concepts of direct investment and reserve assets are not subject to the financial instrument category of national accounts. As a consequence data confrontations would exaggerate discrepancies and cannot lead to meaningful conclusions.

Trade credits and advances and insurances technical reserves are both components of direct investment according to ESA 2010 (Table 18.14), which is explicitly not represented in the BPM6 standard representation. The standard presentation of BPM6-Appendix 9 omitted some subcomponents of direct investment, while focusing on financial transactions in equity/investment fund shares and debt instruments, and their breakdowns by chains of control and ownership. Thus, the presentation of BPM6 appears not fully comparable to ESA 2010.

We had to conclude finally, that these breakdowns are either collected in Member States but are not subject of the regular reporting to Eurostat ([11]), or they do not exist in the standard representation of BPM6 ([12]). As a consequence data confrontations for the financial accounts between the two concepts have to be based on the available data sets and restricted to their total net

([11]) Specified in Eurostat (2016a), Update February 2016.
([12]) BPM6, Appendix 9.

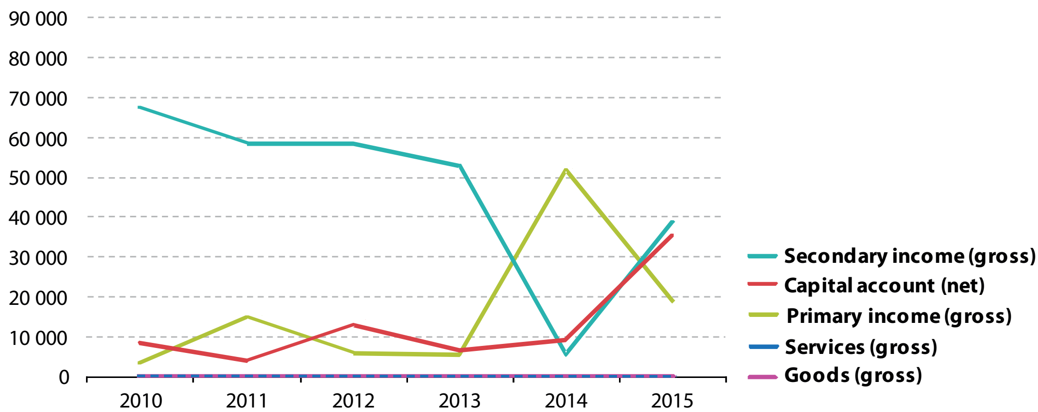
values ([13]), as they cannot be fully reconciled in the component accounts with the data available to Eurostat. It should however be noted that Member State compilers have access to the required breakdowns, and such comparisons should be possible at national levels.

4. Levels of discrepancy in a multiannual context

4.1. Discrepancies in the EU-28 aggregate

Eurostat compiles the consolidated EU-28 aggregates of the non-financial accounts of BOP and QSA, based on the national data reported by all Member States. When comparing the corresponding non-financial accounts for the consolidated EU-28 aggregate the full consistency in goods and services ([14]) is more than evident (see Figure 1). However, secondary income shows elevated levels of discrepancies until 2013, and the primary income remains stable at low levels, but reveals an outlier in 2014. In general the secondary income and capital accounts show an increase of discrepancies in the most recent reference year.

Figure 1: Discrepancies in the EU-28 non-financial accounts, EU-28 aggregate, 2010–2015
(million EUR)



Note: Line for Goods and Services overlap with a zero value for the entire period, representing full consistency. April 2016 data.
Source: Eurostat.

Given the different production calendars a bias for revision and vintage effects can never be completely excluded. We tried to avoid it by strictly choosing the comparable vintages from the data as published in April 2016. Comparing with discrepancies observed in earlier data confrontations ([15]) we take note that overall levels of discrepancies in the EU-28 aggregate rigidly remain at around 0.6 % of GDP over the time span.

([13]) Financial account net lending/borrowing (BOP/IIP) and ROW net acquisition of assets/net incurrence of liabilities (QSA), i.e. B9F.
([14]) Eurostat uses the available BOP data for goods and services for the compilation of the EU-28 aggregate in the QSA.
([15]) Eurostat (2016b), p.8.

In 2015 currently about EUR 93.3 billion were measured in absolute differences in the EU-28 current/capital account ([16]), which compare to EUR 66.7 billion for the previous year. Total levels of discrepancies remain stable but reached an annual average of EUR 76.7 billion over the 6-years period. In comparison, the sum of discrepancies occurring to the current/capital account based on national data of the 28 Member States amounted to EUR 264.8 billion in 2015 (EUR 316.8 billion in 2014). This illustrates the limited scope of interpretation of discrepancies occurring in the consolidated EU-28 aggregate, where offsetting effects occur which compensate the dynamics in discrepancies at national level. Hence, an analysis of national data appears more conclusive in the study of discrepancies arising from Member States' compilation processes.

4.2. Discrepancies in national data in the non-financial accounts

Since the latest publication of data in April 2016, quarterly data covering the entire time span of 2015 have become available in most Member States. This allows the previously conducted analysis (Eurostat (2016b)) to be extended to 6 years. However, data for 2015 still appear provisional in some countries, or even missing ([17]). As a consequence, we should not over-interpret the measures for 2015 at this moment, as subsequent revisions in the national data of both statistics can be expected henceforth.

Table 3: Absolute exposure to discrepancies, sum of national data, current/capital account, by BOP item, 2010–2015
(million EUR)

	2010	2011	2012	2013	2014	2015 (p)
Total	187 288	208 028	207 868	223 739	316 819	301 630
% of GDP	1.5	1.6	1.5	1.7	2.3	2.1
Goods	28 138	27 729	27 874	30 335	32 888	59 715
Services	69 077	71 645	77 975	82 449	107 012	103 075
Primary income	45 281	55 081	40 296	53 377	122 477	76 999
Secondary income	34 736	38 114	50 865	45 699	42 169	46 808
Capital account	10 055	15 460	10 858	11 879	12 273	15 033

Notes: Discrepancies = absolute differences BOP minus ROW items in gross transactions, for capital account net transactions. Measures for 2015 provisional (p).

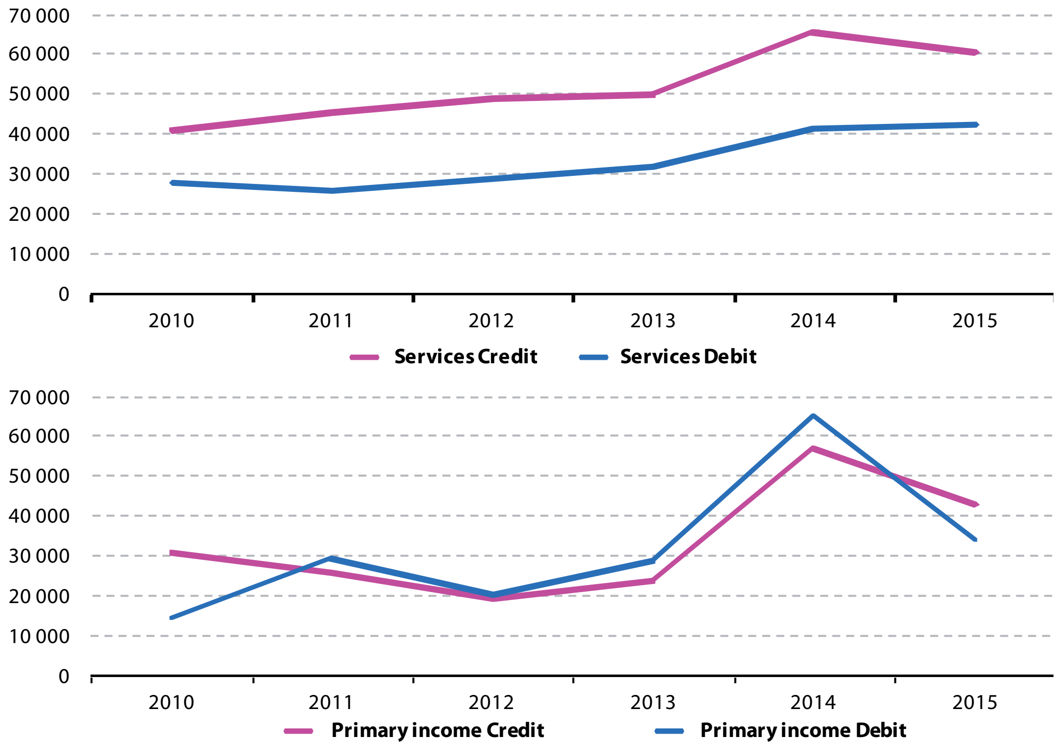
Source: Eurostat.

According to table 3, levels of discrepancies in Member States' non-financial accounts have surpassed the EUR 300 billion mark since 2014. In general, levels appear elevated although the structure of discrepancies has changed over the 6-years period. The major contributor in explaining discrepancies in Member States' non-financial accounts has been traditionally the services account, which temporarily has been overtaken by an outlier value in the primary income account in 2014. On the other hand, the goods, the secondary income and the capital account show stable, but elevated levels throughout the observed period. While we would expect future revision practice to 'smoothen' discrepancies occurring to the more recent observations of 2015, the persistently elevated levels of discrepancies in a multiannual context appear however remarkable — particularly in the light of full consistency of the methodological standards.

([16]) Discrepancies for the current/capital account are the sum of discrepancies occurring in the respective component accounts.
([17]) QSA data for Bulgaria are currently missing as from 2nd quarter 2015.

In absolute terms the services and primary income accounts appear to be most affected by discrepancies. While services showed a discrepancy of EUR 103.1 billion in 2015 (EUR 107.0 billion in 2014), primary income dramatically went up from EUR 53.4 billion in 2013 to EUR 122.5 billion in 2014 (EUR 103.1 billion in 2015), explaining together more than 72 % of total discrepancies occurring in 2014. The patterns of inconsistency in these two accounts appear however very different: while the primary income account shows high levels of discrepancies both for credit and debit flows, debit flows seem generally less exposed to discrepancies than credit flows (thus relating to exports of services) in the services account (see Figure 2).

Figure 2: Discrepancies in the services and primary income accounts, sum of national data, 2010–2015
(million EUR)



Note: Discrepancies = differences BOP minus ROW items, 2015 provisional.

Source: Eurostat.

Relative to the annual transactions in the respective component accounts, exposure to discrepancies appears important in the capital [18] and to a lesser extent in the secondary income account (Table 4). This is due to the smaller transaction volumes applying to these two component accounts. For services, the relative exposure amounted to 3.4 % of annual credit flows and 2.8 % of annual debit flows in 2015, while for the primary income account the relative exposure was 3.2 % and 2.5 % respectively. For goods it can be concluded that the relative exposure appears minor, given the high volumes of this component's average transactions.

(18) In order to exclude a bias arising from the applied net recording of transactions in QSA for the component acquisition less disposals of non-produced non-financial assets (NP) discrepancies have been measured on the capital account's net transactions.

Table 4: Relative exposure to discrepancies, sum of national data, current/capital account, by BOP item, 2010–2015
(% of annual transactions)

	2010	2011	2012	2013	2014	2015 (p)
Goods - Credits	0.5	0.4	0.3	0.4	0.5	0.8
Goods - Debits	0.3	0.3	0.3	0.4	0.3	0.5
Services - Credits	3.2	3.3	3.3	3.2	4.0	3.4
Services - Debits	2.4	2.2	2.3	2.4	2.9	2.8
Primary income - Credits	2.3	1.8	1.4	1.8	4.2	3.2
Primary income - Debits	1.1	2.1	1.5	2.2	4.9	2.5
Secondary income - Credits	11.8	10.2	13.0	10.8	9.5	9.5
Secondary income - Debits	4.3	5.5	6.8	5.9	5.6	6.4
Capital account - Net	24.7	35.1	20.1	21.2	24.9	18.8

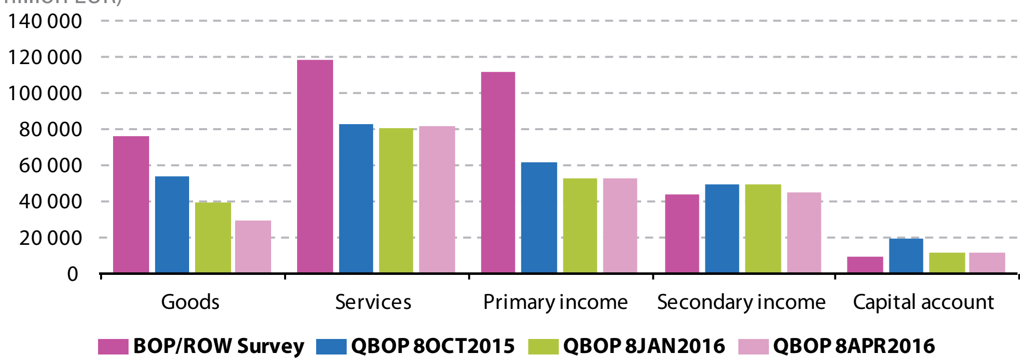
Note: Total discrepancies by credits and debits relating to average annual transactions; 2015 provisional.

Source: Eurostat.

As mentioned above, we would expect constant revisions of time series to 'smoothen' the levels of discrepancies in the long run, as compilers would compare their statistical products with each other in due course and apply corrective measures for the sake of improved data quality. We can show this expected process of converging statistics by comparing the results of earlier data confrontations with the current exercise. Eurostat runs these exercises on a regular basis since October 2015, following a pilot survey among compilers ('BOP/ROW survey') in April 2015, when compilers were asked to explain the observed discrepancies for the first time after the synchronisation of the methodological standards in both statistics (BPM6 and ESA 2010) by end-2014. Clearly this does not support any long-term comparisons, but could give an indication of the current situation.

When comparing the genesis of discrepancies occurring to the 2013-vintage in both statistics (Figure 3), we can see a distinct course of convergence only in the goods account, while improvements in all other accounts appear only hesitantly. So far we can conclude that after the 'shock' of methodological synchronisation, levels in discrepancy have been reduced, but remain elevated in most components of the non-financial accounts.

Figure 3: The genesis of measured discrepancies over time, sum of national data, 2013
(million EUR)



Source: Eurostat.

4.3. Discrepancies in national data in the financial accounts

The sum of national data in total absolute discrepancies of net financial account transactions (BOP) amounted to EUR 164.5 billion in 2015 (EUR 302.7 billion in 2014), showing a volatile development over the 6-years period (Table 5). Given the earlier described limitations in the analysis of the financial account components, only total net values could be compared ([19]), which are generally more volatile due to their composite nature, and allow little conclusions. Generally we may content ourselves in observing higher discrepancies for 2014, but much lower for 2015, but it will have to be seen whether this positive development becomes a trend in future. Under any circumstances compilers have become more aware of the consistency issue, which the current figures seem to confirm ([20]). On the contrary, the total net value of the international investment position (IIP) and the corresponding financial balance sheet of the ROW account record increases in discrepancies. In 2015 EUR 826.0 billion were measured at constantly elevated levels in a multiannual context (EUR 610.1 billion in 2014)([21]).

Table 5: Financial account discrepancies, sum of national data, 2010–2015
(million EUR)

	2010	2011	2012	2013	2014	2015 (p)
Transactions (BOP/ROW), net	221 123	96 302	305 250	206 292	302 744	164 464
In % of GDP	1.7	0.7	2.3	1.5	2.2	1.1
Positions (IIP/ROW), net	547 968	301 765	374 119	591 481	610 134	825 998
In % of GDP	4.3	2.3	2.8	4.4	4.4	5.6

Note: Discrepancies = differences BOP/IIP net lending/borrowing (financial account; net = assets-liabilities) minus net acquisition of assets/net incurrence of liabilities (ROW account; net = liabilities-assets); 2015 provisional.

Source: Eurostat.

4.4. Contagion effect of the financial account

In countries where compilation processes of financial account statistics in BOP are based on positions data (IIP/financial balance sheets), financial transactions (BOP) are derived under consideration of price/foreign exchange effects and other changes ([22]).

As financial account positions are used by compilers for their estimations on income and financial intermediation services indirectly measured (FISIM), inconsistencies could effectively transmit into the current account/non-financial ROW account via these estimations, by being based on inconsistent positions data. Consequently, a potential contagion effect applies implicitly, transmitting inconsistencies not only from the IIP/financial balance sheets into the BOP/ROW financial account (stock-

([19]) Comparing net lending/net borrowing in the BOP financial account with net acquisition of assets/net incurrence of liabilities in ROW (B9F).

([20]) Germany, France and Italy contributed to this with higher discrepancy of EUR 29.3 billion, 25.2 billion and 24.5 billion respectively in 2015. Germany explains these measured differences by the specific needs of the respective statistical domains. While the sector accounts primarily focus on compiling sound data for the domestic sectors, BOP concentrates on external relationships. Initiatives to address the observed discrepancies and differences in source data will be reinforced in Germany, once BPM6/ESA 2010 production processes have been stabilised. This will also entail a closer cooperation between the national counterparts.

([21]) Missing data in QSA were replaced by corresponding net values from the Annual Sector Accounts. This could have blurred the picture by vintage effects in the earlier years. Missing financial data occurred in 8 countries for 2010 and 2011 (Bulgaria, Denmark, Ireland, Greece, Croatia, Italy, Cyprus, Slovakia).

([22]) For example, featuring data collection model 5 — monthly stocks securities-by-securities, derived monthly flows securities-by-securities, European Central Bank (2002) — Annex to the executive summary.

flow inconsistencies) but also into the current account/ROW non-financial account. This gives a new dimension to the analysis of discrepancies, as not only item exposure and the use of different estimation practices or data sources are relevant, but also the contagion effect that arises from internal inconsistencies.

5. Reasons for discrepancies according to the BOP/ROW survey

In order to obtain a better understanding of the nature of discrepancies in general, we will resort to the explanations given by Member States in the earlier mentioned BOP/ROW survey (23). This survey was launched by Eurostat in the aftermath of the methodological change in BOP statistics by end-2014. Instead of purely monitoring, compilers were asked to provide explanations for observed discrepancies in their jurisdictions. The exercise was organised in two cycles — one dealing with discrepancies in the non-financial accounts (April 2015), and one with the financial accounts (November 2015). The survey helped to get a first picture about the consistency between the two statistics after the synchronisation of methodologies in BOP and national account statistics, enriched with a qualitative self-assessment by national compilers. The latter helped to identify the major parameters which explain statistical inconsistencies in the current data comparison.

5.1. Uncoordinated compilation practices and inconsistent use of data sources

According to Table 6, 44.1 % of all explained discrepancies occurring in the current/capital account were attributed to the use of different compilation practices and data sources in the compiling economy by Member States. Although not necessarily of homogenous character, this clearly is the major contributing factor to inconsistency in the two statistics.

Inconsistent compilation practices applied predominantly to the primary income account. The compilation of its major component property income, and to a lesser extent the compensation of employees, seemed to be hardly coordinated. While some Member States remained unspecific, where the inconsistency applies, others explicitly attributed the observed differences to specific components in the primary income account (i.e. property income or compensation of employees). Compilation problems for property income are commonly known when it comes to estimating income flows generated from resident investments abroad and computing a breakdown of distributed and reinvested cross-border income flows (e.g. income distributed and reinvested from shares held by residents in non-resident mutual funds). Similar information asymmetries apply to compilers, when estimating compensation of employees in border or seasonal work contracts (e.g. residents working abroad or unregistered/illegal work by non-residents in domestic labour markets). Different approaches and parallel estimation practice could result in coordination issues among national counterparts. Further, the use of different data sources in the computation of the two statistics appears relevant.

(23) Eurostat (2016b), p.12 ff.

Table 6: Explained discrepancies by stated reason, current account
(% of total)

Reason of discrepancy	%
Different compilation practices (unspecified)	30.1
Different data sources	14.0
Vintage and revision differences	13.1
Reclassification Services and Goods	9.9
FISIM allocation	6.6
Differences property income	6.3
Transit trade (natural gas)	4.8
Errors detected	4.8
Balancing of accounts	3.3
Illegal economy	1.8
CIF-FOB adjustment	1.6
Differences compensation of employees	0.7
Other	3.1

Note: BOP/ROW survey.

Source: Eurostat.

5.2. Other common reasons for discrepancies in the current/capital account

Furthermore, the survey specifies more common reasons for discrepancies, although with less quantitative impact than the previously mentioned. The uncoordinated reclassification practice between the goods and services account (or vice versa) occurred in the wake of the new compilation methodology (e.g. goods under merchanting, goods for processing), and the inconsistent allocation of FISIM. It applied to the services account (where FISIM is registered [24]) and concerned especially Luxembourg and Greece. The main known problems rely on the calculation of sector-specific reference rates for loans and deposits which are negotiated with the resident financial sector outside the regulated banking system (e.g. financial leasing), and in the availability of external reference rates, where information on sector-specific interest rates appears difficult to obtain. These difficulties could have resulted in coordination problems among national counterparts.

Regarding the required practice of CIF/FOB adjustments Eurostat has assumed bilateral contacts with the concerned Member States. Sweden in particular seemed to be effected by this shortcoming.

Vintage and revision differences arose from the different production cycles in national accounts and BOP. A careful choice of vintages and a better synchronisation of production cycles (including revisions) could minimise, but never neutralise these effects [25]. This clearly refers to the need to put more emphasis on the correct implementation of the harmonised EU revision policy for national accounts and BOP, in order to support better comparability [26].

[24] 'In BPM6, interest flows are measured on exactly the same basis as in the SNA with FISIM separated and treated as an import or export of financial services.' (SNA 2008, paragraph 26.60).

[25] An implicit bias for example occurs in the present study, when in the lack of quarterly data we had to refer to equivalent data from the Annual Sector Accounts.

[26] While QBOP is compiled at t+85 days, comparable data in QSA refer to a more heterogeneous production calendar — euro area countries produce at t+85 days, all other Member States at t+3 months.

5.3. Reasons for discrepancies in the financial account

Although we were facing considerable challenges in analysing financial transactions due to the above mentioned conceptual differences in both statistics, resorting to the explanations received from Member States in the BOP/ROW survey could be helpful to attain a better understanding of patterns applying to the financial account (see Table 7). Beside the already mentioned different compilation practices — which clearly also apply to the financial account (21.0 % of all explained differences) — different net recording practices in regard to financial derivatives were mentioned as the most prominent contributing factor to inconsistencies (36.6 % of all explained differences). Hence the different net recording practices of financial derivatives and employee stock options appear of particular interest.

Table 7: Explained discrepancies by stated reason, financial account
(% of total)

Reason of discrepancy	%
Net recording (financial derivatives)	36.6
Vintage and revision differences	22.8
Different compilation practices (unspecified)	21.0
Different concepts (direct investment, reserve assets)	7.7
Different data sources	5.9
Other	5.9

Note: BOP/ROW survey 2015.

Source: Eurostat.

While the ROW account shows assets and liabilities (F.7) according to the balance sheet approach, the BOP financial account records only net values (assets minus liabilities). The reason for this relies on the interpretation of BPM6, paragraph 8.7 on net recording, and more specifically paragraph 8.34 on financial derivatives. While changes in financial assets should generally not be netted against changes in liabilities, the BPM6 allows an exception from this rule for financial derivatives, wherever gross recording is impractical. This should however lead to comparable net values, when abstracting from financial derivatives in reserve assets ([27]).

Table 8: Net financial derivatives, countries with large discrepancies or opposite signs, 2014 and 2015
(million EUR)

	2014		2015	
	BOP	ROW	BOP	ROW
Germany	31 769	– 25 022	25 796	7 337
Ireland	24 317	12 852	32 733	21 125
Italy	– 3 581	5 450	3 358	6 752
Greece	373	– 338	331	– 187
Slovenia	– 2	4	29	25

Note: This comparison excludes reserve assets components.

Source: Eurostat.

We measured EUR 46.1 billion in absolute discrepancies for 2015 (after EUR 100.4 billion in 2014). While Germany and Ireland showed outliers (EUR 18.5 billion, and EUR 11.6 billion respectively), all other Member States recorded minor discrepa-

([27]) We assume financial derivatives have a minor importance in reserve assets. Thus, a comparison of the respective items in BOP and national accounts appears justified and comparable.

cies ([28]). We also noticed that some Member States, which show opposite signs for the respective net values in 2014 ([29]), have changed this in their 2015-figures (Table 8). We may conclude that certain confusion about the sign convention that applied earlier could have been clarified. Under any circumstances, clarifications on the latter are subject to the provisions of SNA 2008, paragraph 26.10 in regard to the sign convention of the ROW account, which leaves little room for interpretation ([30]).

6. The geographical representation of discrepancies

The analysis of the geographical breakdown of discrepancies appears useful in order to illustrate the originators of discrepancies occurring in both statistics. It will be shown that the situation is concentrated around a few Member States only. In this light Eurostat's approach to tackle inconsistencies in Europe is focused on the major originator countries, in order to address absolute volumes of discrepancies with highest impact to the measured totals in the EU-28. The raised discrepancies in absolute volumes play very often a minor role in the respective countries in regard to their GDPs, a circumstance which may have led to the persistently elevated levels of discrepancy in both statistics. The following data represent multiannual mean values for the period 2010–2015 ([31]), both in absolute terms and relative to GDP, which should assist in identifying long-term exposure by country and provide a starting point for further investigations at national level.

6.1. Current and capital account

A geographical breakdown of discrepancies occurring in the current/capital account (Figure 4a) show a high concentration around a group of five countries, with France as the major contributor (35.4 % of total discrepancies), the Netherlands (12.6 %), Belgium (6.8 %), Luxembourg (6.2 %), and Germany (5.9 %). Together these countries already cover two thirds of all observed discrepancies in the current and capital account in 2010–2015 ([32]). A comparison of the multiannual mean discrepancies in the specific component accounts clearly shows the different exposure of countries. While in France and Luxembourg discrepancies particularly apply to the services account, Belgium and the Netherlands show an exposure in the primary income account ([33]), and Germany in the secondary income account ([34]).

([28]) Among those, France showed full consistency in 2015, and the United Kingdom, which appears highly exposed to this financial instrument category recorded only minor differences (EUR 0.1 billion).

([29]) Eurostat (2016b), p.14.

([30]) '…the rest of the world is drawn up from the perspective of the rest of the world. BPM6 looks at the same stocks and flows from the point of view of the domestic economy. Thus the BPM6 entries are the mirror image of the SNA entries relating to the rest of the world.' (SNA 2008, paragraph 26.10).

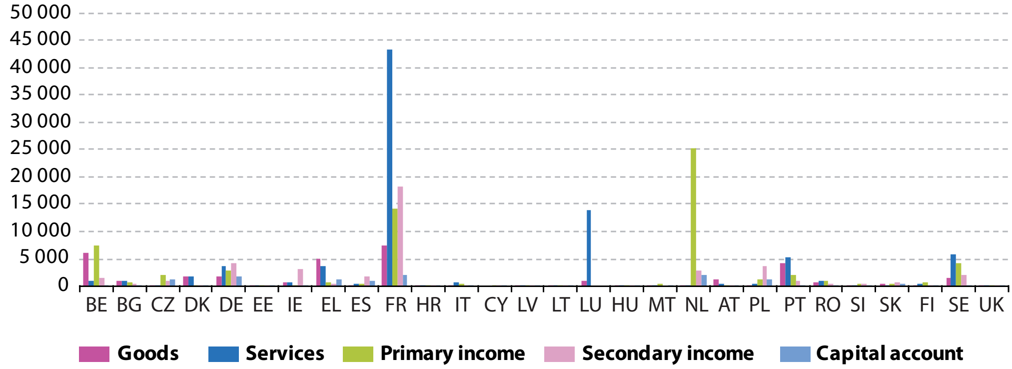
([31]) We have not weighted the multiannual mean further, as we postulate that all time-series which were compiled according to ESA 2010 and BPM6 should be entirely comparable.

([32]) Mean annual discrepancies 2010–2015: France EUR 85.6 billion, Netherlands EUR 30.4 billion, Belgium EUR 16.4 billion, Luxembourg EUR 15.0 billion).

([33]) The discrepancies have been calculated on the basis of gross income flows. It should be noted that gross income flows are rather large in the Netherlands related to the presence of special purpose entities (SPEs). Their activities do not influence the income balance. However, due to still different revision policies gross income flows of SPEs may temporarily differ between national accounts and BOP. On a net basis the discrepancies are much smaller.

([34]) The discrepancies observed in Germany seem to relate to a great extent to revision effects arising from social benefits and contributions.

Figure 4a: Mean annual discrepancies, current/capital account, by Member State, 2010–2015 (million EUR)



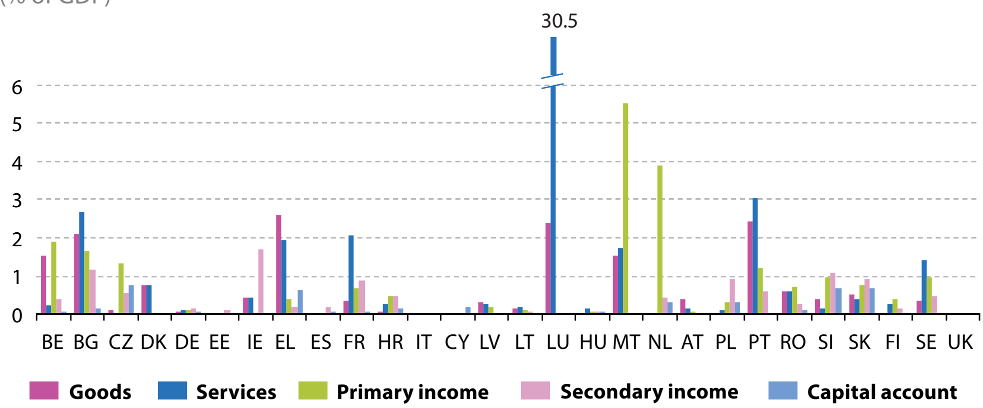
Note: Absolute discrepancies = differences BOP minus ROW items.

Source: Eurostat.

In relative terms (Figure 4b), the highest exposure to discrepancies in the current/capital account in relation to their GDP was observed in Luxembourg for services (30.5 %) and in Malta and the Netherlands for primary income (5.5 % and 3.9 % respectively).

Finally, it can be concluded that the geographical representation of discrepancies in the non-financial account shows a high concentration on only a few countries. A strategy to address the discrepancies occurring in the above mentioned 5 countries could already significantly reduce levels of inconsistency in both statistics.

Figure 4b: Mean annual discrepancies, current/capital account, by Member State, 2010–2015 (% of GDP)



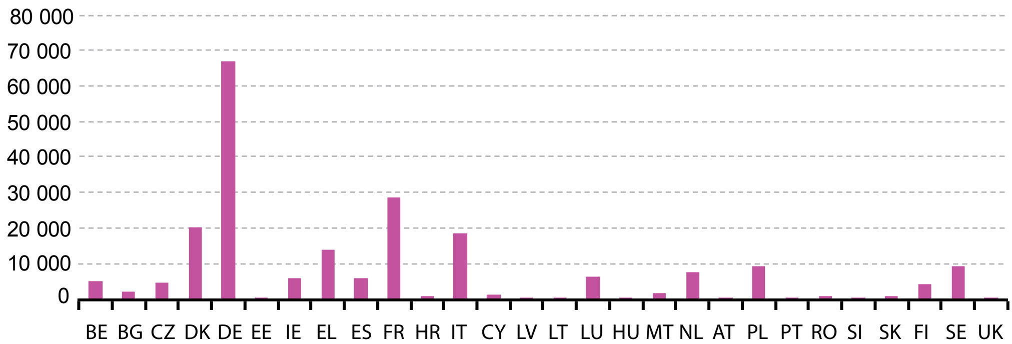
Note: Relative discrepancies = absolute discrepancies in % of GDP (mean 2010–2015).

Source: Eurostat.

6.2. Financial account

Discrepancies in the financial account appear even more concentrated around some countries. The current data confrontation revealed at least 5 Member States appearing as a source of discrepancies in the multiannual context, explaining more than 90 % of the observed discrepancies (Figure 5a). Most prominently Germany contributed to this (31.1 % of total discrepancies)([35]), followed by France (13.3 %), Denmark (9.3 %)([36]), Italy (8.7 %) and Greece (6.5 %).

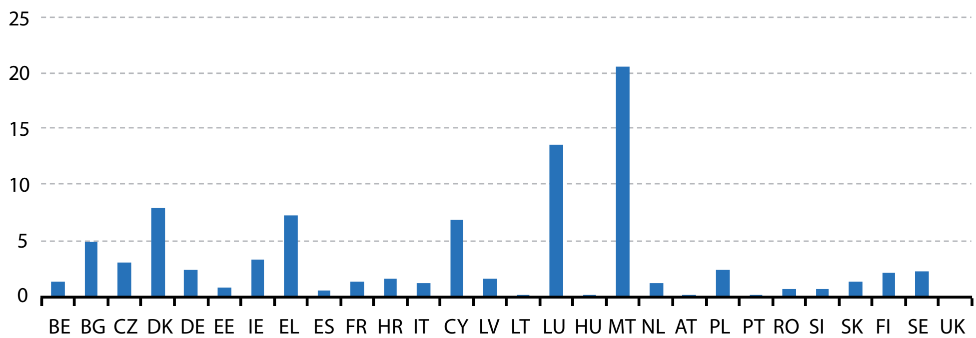
Figure 5a: Mean absolute discrepancies, financial account, total net, by Member State, 2010–2015
(million EUR)



Note: Absolute discrepancies=differences BOP minus ROW items. Comparing net lending/net borrowing in the BOP financial account with net acquisition of assets/net incurrence of liabilities in the ROW account

Source: Eurostat

Figure 5b: Mean relative discrepancies, financial account, total net, by Member State, 2010–2015
(% of GDP)



Note: Discrepancies = differences BOP minus ROW items. Comparing net lending/net borrowing in the BOP financial account with net acquisition of assets/net incurrence of liabilities in the ROW account.

Source: Eurostat.

In relative terms (Figure 5b), the highest exposure to discrepancies in relation to their GDP was observed in Malta (20.7 %) and Luxembourg (13.6 %). Although the concentration of discrepancies in the total net of the financial accounts around a group of 5 countries is considerably high, further studies would be necessary by Member States, in order to identify those components of the financial account which are mostly affected by discrepancies. Due to the limited data, a deeper analysis cannot be performed at this moment.

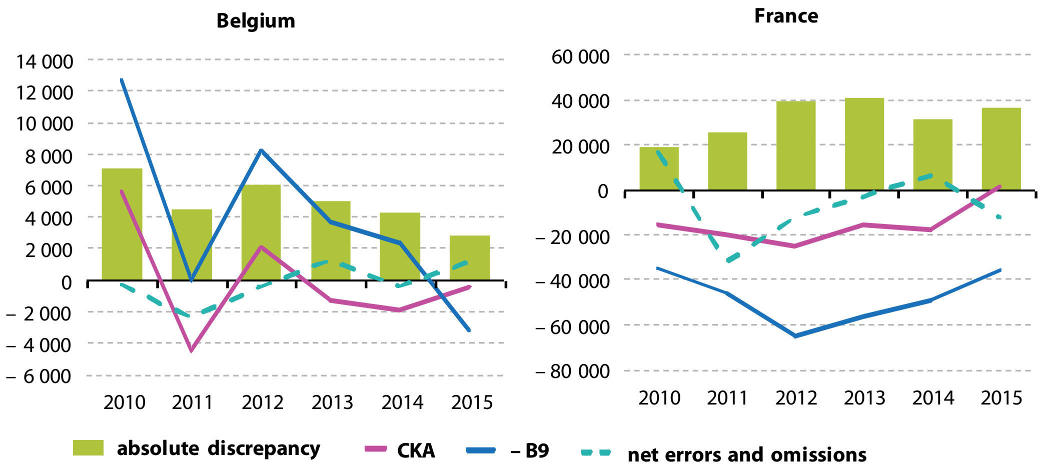
([35]) Particularly based on an outlier in the corresponding QSA time series for 2014.
([36]) Danish inconsistencies seem to refer rather to older data of 2010-2013, which suggest further revisions for these periods. On the contrary recent years (2014-2015) appear more balanced.

6.3. Studying measures of comparability

An indicative measure to analysts for the comparability of BOP and national accounts statistics is to study the behaviour of the current/capital account balance and the mirror net lending/net borrowing (B9) of the ROW account. In general, little or no differences would indicate good comparability. Parallel behaviour of the curves could refer to systematic differences in one or the other, but strong and erratic deviations would indicate rather poor comparability of both statistics.

In the line with our earlier findings, the situation among Member States appears quite heterogeneous. Countries like Belgium and France (Figure 6a) show sharp deviations between the two measures, thus indicating poor or little comparability in general, while countries like Germany or the United Kingdom (Figure 6b) show little or no deviations in their measures, thus indicating better comparability. However, this should not lead us to premature conclusions. The above conclusions are based on the assumption that current/capital account and financial account are balanced with each other, and thus there are no errors and omissions ([37]) in BOP statistics. On the other hand it is assumed that net lending/net borrowing of the non-financial account (B9) balances with the net financial account (B9F) of the ROW account, i.e. no vertical discrepancy occurs in the ROW. Unfortunately both assumptions usually are not supported by statistical evidence, and thus blur our conclusions on comparability. Consequently, the above presented measures always have to be studied in the context of the arising net errors and omissions in the BOP — in general, good convergence of both measures (CKA, B9) over time under limited net errors and omissions indicate good comparability of the two statistics.

Figure 6a: Current/capital account balance and net lending/net borrowing, Belgium and France, 2010–2015
(million EUR)

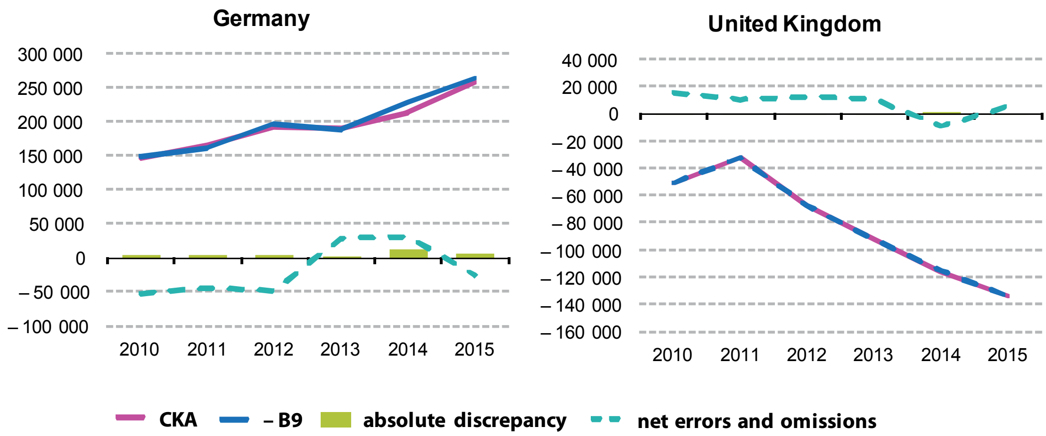


Note: CKA = current and capital account balance. B9=net lending/net borrowing of the ROW account. Absolute discrepancy = differences CKA minus B9. Net errors and omissions = CKA minus net financial account.

Source: Eurostat.

([37]) Net errors and omissions equal the difference between the current/capital account balance and the net financial account.

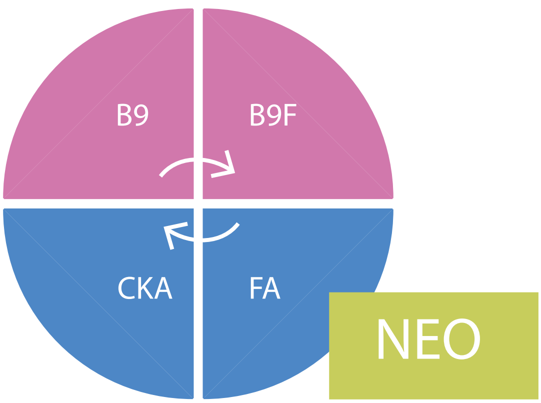
Figure 6b: Current/capital account balance and net lending/net borrowing, Germany and United Kingdom, 2010–2015
(million EUR)



Note: CKA = current and capital account balance. B9 = net lending/net borrowing of the ROW account. Absolute discrepancy=differences CKA minus B9. Net errors and omissions = CKA minus net financial account. Identical lines of CKA and B9.

Source: Eurostat.

Figure 7: The impact of consistency to net errors and omissions



Note: BOP: CKA = current and capital account balance. FA = net financial account ROW: B9 = net lending/net borrowing of the non-financial sector accounts. B9F = net lending/net borrowing of the financial sector accounts. NEO = net errors and omissions.

Source: Eurostat.

As illustrated in Figure 6b, the occurrence of net errors and omissions limits the interpretation of our comparability measures. Further, due to their residual character they allow little conclusions about the source of inconsistency in BOP statistics ([38]). A similar reasoning arises from discrepancies occurring between the net values of non-financial (B9) and financial accounts (B9F). As a consequence, compilers ambitions are to keep internal inconsistencies low, in order to maintain good interpretability of their respective

([38]) Net errors and omissions could derive from data gaps occurring either in the current/capital or the financial account, or from both.

statistics. Quality processes which focus on increasing internal consistency in both statistics would therefore aim at lower errors and omissions and internal discrepancies between the non-financial and financial accounts of the QSA (Figure 7). As a consequence, the study of internal consistency aspects appears instrumental in order to better understand the dynamics of inconsistencies in both statistics.

7. Aspects of internal consistency in BOP and national accounts statistics

Internal consistency measures provide a picture about systematic discrepancies arising from each statistics autonomously, before being compared with each other. The focus is hereby on discrepancies which contravene the accounting framework of balances and accounts in both statistics. Internally consistent statistics would be the ideal starting base for comparisons between different statistics, in order to neutralise potential contagion effects which would be imported into the analysis of external consistency. Unfortunately this is rarely the case in practice. Therefore it appears justified to incorporate aspects of internal consistency into this analysis as well. A prominent measure for internal consistency is the analysis of the residual item net errors and omission in BOP statistics, and the vertical discrepancy of the net values of financial (B9F) and non-financial (B9) accounts in the ROW. Recorded discrepancies somewhat challenge the interpretation of the accounts, and bear a potential of importing inconsistencies into the presented analysis, when appearing large or volatile. In order to stay consistent with the earlier analysis, we will continue to measuring these discrepancies in absolute values.

In respect of the overall accounting framework of BOP statistics, net errors and omissions measure the extent of internal discrepancies occurring between total recorded inflows and total recorded outflows (both financial and non-financial). Due to the double accounting framework of BOP they should balance each other in principle, although in practice this appears more difficult, given the fact that generally the 'legs' of the transactions are not collected by the compiler simultaneously but usually from different sources of information ([39]).

Comparing net errors and omissions with the gap between the non-financial and the financial accounts (i.e. the vertical discrepancy) as absolute difference of net acquisition of financial assets/net incurrence of financial liabilities (B9F) and net lending/net borrowing (B9), would intuitively imply some relationship between the two measures. However, this cannot be assumed reasonably without oversimplifying the complex underlying compilation processes. In countries where BOP is the main source for the compilation of the ROW account, net errors and omissions may play a prominent (but not exclusive) role in explaining vertical discrepancies in the ROW. This justifies (although with some caveats) a closer look at net errors and omission occurring in BOP statistics.

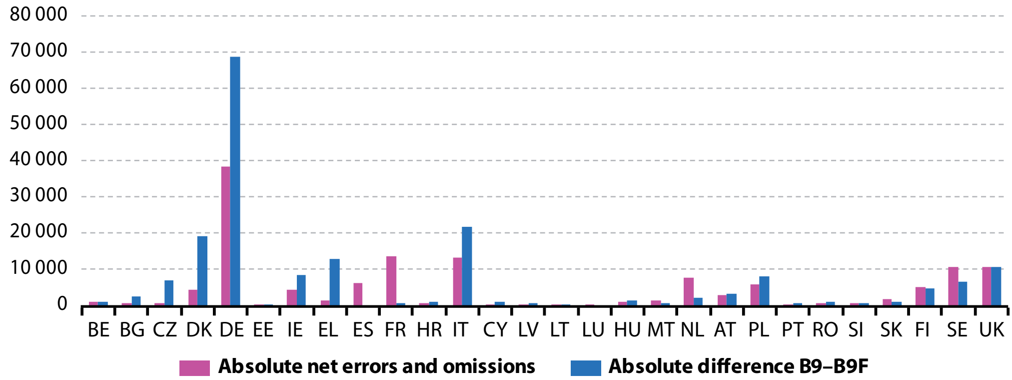
When comparing the multiannual mean of 2010–2015, we noticed the following patterns (Figure 8):

- countries showing elevated levels in both net errors and omissions and vertical discrepancy (e.g. Germany, Italy, United Kingdom, Sweden, Finland);

([39]) A useful explanation among others provides Banque de France (2015), part 2.

- countries showing prominent levels exclusively in net errors and omissions (e.g. the Netherlands, France, Spain);
- countries showing prominent levels exclusively in the vertical discrepancy of the ROW accounts (e.g. Czech Republic, Denmark, Greece).

Figure 8: Internal discrepancies in balance of payments and national accounts, by country, absolute values, mean 2010–2015
(million EUR)



Note: Absolute net errors and omissions=absolute difference of net acquisition of assets/net incurrence of liabilities in financial account minus net lending/net borrowing in current/capital account; Absolute difference B9F-B9=absolute difference of net acquisition of financial assets/ net incurrence of liabilities of financial ROW account (B9F) minus net lending/net borrowing of non-financial ROW account minus (B9). – ROW: values for Luxembourg missing.

Source: Eurostat.

Germany and to a lesser extent the United Kingdom show elevated levels in internal discrepancies in both statistics. This puts the earlier findings on good comparability of both statistics into perspective, and illustrates that in some countries the issue of vertical discrepancy is crowded out by horizontal consistency in the domestic sector accounts ('balancing the accounts').

Keeping in mind the above mentioned situation, the BPM6 offers some basic interpretation for the study of net errors and omissions: negative values would indicate that either credit flows in the current/capital account and net increases in financial account liabilities are too high, or alternatively debit flows in the current/capital account and net increases in financial account assets are too low ([40]).

After all, the exposure to internal accounting discrepancies in European BOP statistics appears more moderate in recent years. This promising trend since 2013 could have arisen from new data sources and extended data collections that have become available based on micro data sources, which provided a more complete and consistent picture to the compiler, particularly on financial assets and liabilities ([41]), and Member States becoming more actively aware of data gaps in their jurisdictions in the aftermath of the financial crisis, possibly addressing the bias in the non-financial accounts ([42]).

([40]) For the interpretation of net errors and omissions, see: BPM6, paragraphs 2.24 ff. Also very helpful the BPM6 Compilation Guide (International Monetary Fund (2014)), paragraphs 8.93 to 8.96 on analysis over time.
([41]) Introduction of securities-by-securities reporting based on balance-sheet data and custodian reports on securities holdings (incl. households), and extension of statistical coverage (special purpose entities, financial holdings, insurances).
([42]) In the current and capital accounts credit flows exceeded debit flows significantly for many years, which might have been based on general information asymmetries concerning imports into the compiling economy, and their resulting underrepresentation of debit flows in this context.

Table 9: Net errors and omissions, by country, 2010–2015
(million EUR)

	2010	2011	2012	2013	2014	2015 (p)
EU-28	– 93 126	– 114 479	– 104 248	18 127	8 695	– 15 819
Belgium	– 332	– 2 404	– 374	1 274	– 428	1 180
Bulgaria	1 061	711	769	– 118	– 1 604	764
Czech Republic	– 740	– 101	883	330	842	1 708
Denmark	– 14 722	– 981	– 1 141	– 8 105	– 2 478	– 79
Germany	– 53 352	– 45 366	– 48 379	29 055	30 416	– 24 664
Estonia	53	109	335	– 40	– 113	228
Ireland	– 9 328	– 12 138	– 238	1 940	– 11	3 941
Greece	1 761	32	1 322	3 208	1 894	1 196
Spain	– 5 439	257	– 1 024	12 789	6 419	12 301
France	16 689	– 32 296	– 12 388	– 2 699	6 649	– 12 321
Croatia	– 886	– 1 089	– 398	– 888	– 545	– 426
Italy	– 29 948	– 17 175	– 9 090	– 3 271	10 428	10 949
Cyprus	227	552	503	– 52	383	677
Latvia	328	63	220	193	511	– 536
Lithuania	– 270	– 5	– 115	– 403	– 1 636	90
Luxembourg	– 20	79	0	2	0	– 40
Hungary	– 1 015	– 2 478	407	– 1 370	– 1 123	– 1 280
Malta	122	22	674	– 664	7 924	– 111
Netherlands	– 3 632	– 8 947	– 13 919	– 1 455	– 13 238	5 176
Austria	– 5 254	25	947	5 117	– 4 454	– 1 235
Poland	– 10 246	– 7 281	– 2 990	– 8 457	– 4 979	– 1 511
Portugal	410	– 239	125	320	359	192
Romania	498	1 168	1 181	92	– 352	729
Slovenia	– 1 469	– 739	– 1 113	– 743	– 92	– 943
Slovakia	– 261	– 723	– 1 779	– 3 723	– 2 492	– 2 652
Finland	689	– 4 983	– 13 291	– 4 184	– 5 656	– 2 902
Sweden	6 211	8 722	– 17 909	– 10 851	– 8 575	– 11 277
United Kingdom	15 740	10 727	12 533	10 829	– 9 353	5 028

Note: Net errors and omissions (NEO) = net acquisition of assets/net incurrence of liabilities in financial account minus net lending/net borrowing in current/capital account — EU-28 = sum of national net errors and omissions. 2015 provisional.

Source: Eurostat.

However, the situation was very different in the Member States (Table 9) — in 2015 14 Member States showed negative errors and omissions, while the other 14 recorded positive values ([43]). The highest absolute values were recorded in Spain (EUR + 12.3 billion), while Germany and France showed negative values (EUR – 24.7 and – 12.3 billion, respectively). Compared to their total average current account gross transactions Sweden and Cyprus showed highest relative exposure ([44]), while compared to their total average financial account stocks, Slovakia and Slovenia measured highest

([43]) Among those, Luxembourg showed full internal consistency since 2012 with measures close to zero.
([44]) Sweden 4.8 %, Cyprus 4.0 %.

relative exposure ([45])([46]). Generally, net errors and omissions for all 28 Member States showed a negative value of EUR – 15.8 billion after EUR + 8.7 billion in 2014. This appears moderate compared to earlier outliers at around EUR – 100 billion or less.

8. Conclusions

Since the synchronisation of the methodological standards in BOP and national accounts statistics consistency between the two has moved into the focus of interest among compilers and statistics users. We showed that improvements were made in the course of the past year with particularly encouraging results in some of the non-financial account components (goods, secondary income, capital account) of quarterly statistics. However discrepancies in services and the primary income still appear elevated and persistent over the time. It has been also shown that discrepancies occurred with a high concentration at a group of only 5 countries in the EU-28. By addressing relevant issues in those jurisdiction would already significantly reduce levels of discrepancies in European statistics.

The analysis of inconsistencies in the financial accounts however is more problematic due to lack of breakdown data, and should be supported by compilers' own detailed analysis of the component accounts. However, it has also been emphasised that inconsistencies in financial account stocks could easily transmit into the above mentioned non-financial account components ('contagion effect') via estimation practices being based on inconsistent financial data (FISIM, property income).

The explanations received from compilers in the BOP/ROW survey have further illustrated that the complex nature of certain statistical items in BOP and national accounts require the application of shared or common estimation methods by national counterparts. National coordination issues and the different use of databases appear as the most prominent explanation for measured discrepancies. A way forward (if not already achieved) is to enhance the technical cooperation between national counterpart organisations, dealing with both statistics. In some Member States such initiatives have shown productive results and helped to detect and correct errors in the respective data. Further, an integrated approach in data compilation could be envisaged between national counterparts (as is foreseen in the Netherlands([47])). However, the latter clearly requires a national consent and could be hampered by institutional rigidities in some Member States. In order to honour reconciliation efforts mediation by the European institutions may be appropriate. A continued close cooperation between Eurostat and ECB appears instrumental to a coordinated mediation of country-specific inconsistency issues, for example through the ongoing practice of country visits by both partners.

Currently production cycles in BOP (QBOP) and national accounts (QSA) are not entirely synchronised, which gives rise to discrepancies due to vintage and revision effects. While QBOP is generally produced at t+85 days ([48]), quarterly non-financial sector accounts data become available by t+85 days for Member States whose currency is the euro, but t+3 months for Member

([45]) Slovakia 3.4 %, Slovenia 2.0 %.
([46]) The indicator of average relative errors (ARE) is usually related to the average current account gross transactions (total annual current account credits + debits, divided by 2). Alternatively the average financial account stocks (total annual IIP assets + IIP liabilities, divided by 2) are proposed as a base for a relative quality measure.
([47]) Bieleveldt, E. and Claasen, P. (2015).
([48]) This deadline will be shortened in 2017.

States whose currency is not the euro (49). This situation leads to inhomogeneous production cycles of country data within QSA and gives a bias to data comparisons. Consequent synchronisation of the production and revision calendars for the two statistics would not only facilitate quality monitoring, but also pave the way towards more harmonised compilation processes in the EU-28. This clearly refers to a need for a harmonised implementation of EU revision policy.

The presented data confrontation of financial transactions in particular revealed some conceptual inconsistencies in both statistics, which require further follow-up by the international institutions. In this context there is a clear need for a consistent mapping of the BPM6 functional categories with the ESA 2010 financial instrument categories, with a particular impact on the reporting of the direct investment components in BOP as regards debt instrument components (FL), insurance, pensions and standardised guarantee schemes (F.7) and other accounts receivable/payable (F.9). Further clarification appears necessarily for the net recording of financial derivatives and the applied sign convention in the ROW account, although recent improvements were noticed.

An analysis of the comparability of both statistics cannot avoid taking into account internal consistency aspects. In this context net errors and omissions have observed promising trends of more internally consistent data in BOP statistics during the recent years, although the impact on the vertical discrepancy in the ROW accounts is not conclusive. This proofs that further investigation at country level is necessary, given the complex relationship between internal consistency issues in both statistics.

Acknowledgements

The author likes to thank Mr. Richard Venniker (De Nederlandsche Bank) and Mr. Jens Walter (Deutsche Bundesbank) for their explanations and critical remarks provided.

References

Banque de France (2015), *Méthodologie — La balance des paiements et la position extérieure de la France*, Paris, France.

Bieleveldt, E. and Claasen, P. (2015), 'Integrating Balance of Payments and Sectoral Accounts in the Netherlands', 60th World Statistics Congress, Rio de Janeiro, Brazil, https://www.bis.org/ifc/events/ifc_isi_2015/008_bieleveldt_claassen_paper.pdf.

European Central Bank (2002), 'Task Force on Portfolio Investment Collection Systems — Final Report', Frankfurt, Germany, https://www.ecb.europa.eu/pub/pdf/other/portfolioinvestmenttaskforce200206en.pdf.

Eurostat (2013), *European System of accounts (ESA 2010)*, Luxembourg, http://ec.europa.eu/eurostat/web/esa-2010/overview.

Eurostat (2016a), *Balance of Payments Vademecum*, Luxembourg.

(49) For quarterly b.o.p. see Commission Regulation (EU) No. 555/2012 of 22 June 2012; for quarterly national accounts see Regulation (EU) No. 549/2013 of the European Parliament and of the Council of 21 May 2013.

Eurostat (2016b), 'Consistency between national accounts and balance of payments statistics', *Eurostat Working Paper*, Luxembourg.

International Monetary Fund (2009), *Balance of payments and international investment position manual (BPM6)*, 6th edition, Washington D.C., USA.

International Monetary Fund (2014), *BPM6 Compilation Guide* — Companion document to the sixth edition of the Balance of payments and international investment position manual, Washington D.C., USA.

United Nations et al. (2009), *System of National Accounts 2008* (SNA 2008), New York, USA.

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy: via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm); from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm); by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

Priced subscriptions:

- via one of the sales agents of the Publications Office of the European Union (http://publications.europa.eu/others/agents/index_en.htm).

PDF: KS-GP-16-001-EN-N
PRINT: KS-GP-16-001-EN-C

Publications Office