

BALANCE OF PAYMENTS STATISTICS AND EXTERNAL RESERVES DIVISION Judith Pfluger **2**: 7389

20 December 2001 ST/IMF/BP/QUALITY.DOC

ASSESSING THE QUALITY OF THE EURO AREA B.O.P./I.I.P. STATISTICS

For information of the IMF BOP Committee

INTRODUCTION

- 1. Useful input on the classification of quality dimensions has come from the work undertaken by the International Monetary Fund (IMF) in this field, leading to the recent seminar on data quality (held on 6-8 December 2000 in Korea¹). In particular, the IMF's cascading approach to structuring quality features into dimensions, elements and indicators, which is broadly supported by senior statisticians, has been considered.²
- 2. This paper follows some background work of the Working Group on Balance of Payments and External Reserves Statistics (WG-BP&ER) and takes on board the IMF's work above mentioned. It presents a framework for assessing the quality of euro area balance of payments (b.o.p.)/international investment position (i.i.p.) statistics. The paper is structured as follows: Part One recalls the list and definition of quality dimensions according to the cascading approach devised by the IMF; Part Two elaborates further on the qualitative assessment and quantitative measurement of quality, including a possible trade-off between some elements; finally, Part Three describes a possible steady-state for assessing the quality of euro area b.o.p./i.i.p. statistics, giving both a short-term and a long-term horizon.

For more information on the outcome of the seminar, please see: http://www.nso.go.kr/sqs2000/.

See the documents entitled "A Framework for Assessing the Quality of Balance of Payments Statistics", IMF, Statistics Department, presented at the October 2000 IMF Committee on Balance of Payments, and "Further steps Toward a Framework for Assessing Data Quality", presented at the International Conference on Quality in Official Statistics (May 2001)

Part One: List and definition of quality dimensions and elements

- 3. Quality is a multidimensional concept. The IMF defines five quality dimensions in its cascading structure, seen both from the <u>input</u> (process / institutional framework) and the <u>output</u> (product) side.³ The quality dimensions are further subdivided into elements and indicators. Each dimension or element requires a clear conceptual definition; to be effective, it should allow qualitative assessment and/or quantitative measurement. Moreover, addressing the issue of quality also implies an evaluation of best practices among countries. This part is general, and does not only refer to the euro area b.o.p. /i.i.p.
- 4. The following dimensions and elements are considered suitable to assess the quality of euro area b.o.p./i.i.p. statistics. This list has been agreed by the WG-BP&ER. It encompasses:⁴
- relevance, timeliness, stability and consistency as elements of the serviceability dimension;
- accuracy as a dimension, with plausibility as a sub-item;
- integrity as a dimension;
- accessibility as a dimension, with transparency as a sub-item.

1.1 The serviceability dimension

5. Serviceability can be defined as the usefulness of a data element and the overall statistical system to which it belongs. Usefulness means at the same time, the existence of defined users' needs (in terms of object) and the requirement to allow an interpretation of the fact depicted.

1.1.1 Relevance

6. Relevance refers directly to the use of statistical information and describes the degree to which the data fulfil the needs expressed by users. Relevance may vary over time depending on the scope of the data and policy needs. Changes in the overall economic environment (in terms of features and constraints) may have implications for the information which is useful to collect.

1.1.2 Timeliness

7. Timeliness refers to the time lag between the period or event in question and the availability of the statistical data that correspond to it. Adequate timeliness is when policy-makers can take decisions in time to achieve the target results.

1.1.3 Stability

8. Data are stable when they can be trusted over time, i.e. when necessary revisions are not significant compared with the phenomenon observed. Users perceive as stable those data upon which decisions can be

In this paper emphasis will be put on the product side of statistics. However, one aim of the paper is to start a discussion on benchmarking the process side of statistics to a framework of best practices, in particular for those b.o.p. items where greater coordination in collection/compilation is seen as a prerequisite for achieving quality.

⁴ The grouping corresponds closely to the IMF approach.

taken as soon as the data are released, since the first assessment is made on the basis of sufficient information to produce a good picture. Stability has to be assessed in connection with accuracy.

1.1.4 Consistency⁵

- 9. In its cascading approach (see footnote 3), the IMF describes the conceptual consistency as when "the scope of the statistics [...] is comprehensive and in accordance with internationally accepted standards". In particular "classification and sectorisation are in accordance with internationally recommended systems." This definition may appear too limited; in addition, with respect to a subset of statistics like b.o.p./i.i.p., this needs further clarification. Consistency may be (i) over time; (ii) between data collected at different frequencies; (iii) internationally; (iv) across variables; in the latter it may be vertically (across transactions), horizontally (across institutional sectors), and/or between flows and stocks.
- 10. Consistency allows comparability of statistics by users. The following are regarded as main features of consistency:
- to allow consistency through time, published data should follow the conceptual framework and present homogeneous coverage;
- ii) data collected at different frequencies are expected to reflect the same reality though level of breakdown may not be the same;
- iii) the conceptual consistency highlighted by the IMF fosters international comparability of statistics, even when compiled by different institutions,⁶ In addition, different measurements of the same phenomenon should not result in unreasonably different data⁷;
- iv) numerical consistency: across variables consistency requires the adherence to the basic arithmetic constraint in terms of coherence of items and related sub-items.

1.2 The accuracy dimension

1.2.1 Accuracy

11. Accuracy is the degree to which data correctly describe or quantitatively assess the phenomenon that the corresponding statistics were designed to measure. More precisely, it can be defined as the closeness of the presented value (directly collected or estimated) to the (unknown) true population number. Perceived accuracy, notably if reinforced by consistency through an integrated statistical information system, generates confidence among users At the origin of the statistical compilation process, accuracy results from a sound data collection system, sound compilation procedures and sound estimation methods.

In the IMF cascading approach, consistency is an element of serviceability.

Oifferences may still arise from different practices regarding the publication of revisions. To the extent that different institutions aim at integrity and accuracy, those differences should not be so high as to produce a different picture of the reality described.

⁷ For instance, consistency between regional b.o.p. statistics compiled by different international organisations.

1.2.2 **Plausibility**

- 12. Plausibility describes the (internal) likelihood of the data. Although it may be difficult to predefine a stable and constant or just "plausible" pattern in statistics and it is unlikely that the relative structures of the b.o.p. of different countries/zones will remain unchanged over time, in principle unexpected sizeable outliers nonetheless deserve attention.8 Significant outliers or sudden and unexpected changes in trend need investigation, especially when there is virtually no economic and/or methodological explanation for them. In practice, the assessment of plausibility takes into account that certain statistics are by nature more predictable (and within a shorter time horizon) than others.
- 13. Revisions as an assessment of plausibility are tracked and mined for the information they may provide. Analysis and studies of revisions are carried out regularly in the statistical process and should be made public and used to guide line users.

1.3 The integrity dimension

- 14. Integrity is achieved through transparency of the procedures and practices by which statistics are collected, compiled and disseminated. Thus this concept underlines the relationship between user confidence in the statistics produced and the perceived objectivity and professionalism of the agency producing the statistics.
- 15. Professionalism, adherence to high ethical standards and independence of agencies and their staff indicate that the institutional framework in force (with regard to the relationship between the statistical agencies and the decision-making bodies, in particular policy-makers) is not detrimental to the data (and metadata) produced.

1.4 The accessibility dimension

1.4.1 Accessibility

16.

From the users' point of view, accessibility reflects the ease of obtaining the information disseminated by a statistical agency, the suitability of the form in which it is shown, the media of dissemination and the availability of metadata.

17. In this context, accessibility is defined as a direct dimension of quality. Information that is not accessible to intended users is *ipso facto* of poor quality, regardless of its accuracy.

As an example, the simplest way to formalise a plausibility check is to standardise the variable x under observation. The average (\bar{x}) and the variance (σ_x^2) of the variable can be calculated over a specific historical time-range. The plausibility check is therefore performed on the observation x_i with the formula $|s_i| = \left| \frac{x_i - \overline{x}}{\sigma_x} \right| > c$, where c is an arbitrary limit over which the observation is considered implausible.

1.4.2 Transparency

18. Transparency indicates an aspect of data quality at the dissemination stage according to which statistics are provided with a complete set of information (the "metadata") on the real content of the statistics themselves. This element is a subset of the dimension accessibility, since it clarifies for users the real content of the data disseminated. This information should include: i) a description of the data; ii) its coverage; iii) its compliance with international methodological standards and relevant breakdowns (in order to foster comparability); iv) the main cases for departure from agreed standards; v) a description of the main estimation procedures applied for missing data, and confronting data from various sources.

1.4.3 Assistance to users

19. A rapid and knowledgeable support service must be available for helping the users in any difficulty. This support service should include: i) a publicised contact person for each subject field, and ii) catalogues of publications, documents, and other services, including information on any charges, widely available.

Part Two: General framework for assessing the data quality of the euro area b.o.p./i.i.p. data

- Statistics on external transactions/positions of the euro area compiled by the ECB can be considered as being under the responsibility of the WG-BP&ER, that is under the joint responsibility of the Balance of Payments Statistics and External Reserves Division (BP&ERD) and the WG-BP&ER members. The quality of euro area statistics is based on i) a common understanding of the quality dimensions and elements, and a shared approach to assessing and measuring quality; ii) a common vision as regards the priorities to be given to quality standards, especially taking into account the possible trade-offs; and iii) a clear perception of the institutional framework in which the assessment of data quality takes place. Within the IMF approach, these elements are seen as pre-requisites for quality in statistics.
- 21. At the September 2000 WG-BP&ER thematic meeting on the future of the euro area b.o.p./i.i.p. collection system a fact-finding exercise was carried out to investigate Member States' views as regards the prioritisation of the quality dimensions. A summary of the main findings (including a second round of surveying the Member States) is given in Annex 1. The exercise produced the following results:
- In the financial account (and particularly in portfolio investment) *accuracy* (with emphasis on the exhaustiveness of the source) is regarded as a key issue; while *regular revisions* cannot be avoided in the framework of the euro area monthly b.o.p., keeping *stability* is a key requirement.
- For goods there is less emphasis on accuracy, exhaustiveness and stability. Consistency across frequencies and regular revisions is needed.
- Investment income follows the quality requirements for the b.o.p. financial account.

- For other items of the current and capital accounts, Member States do not seek the same levels of exhaustiveness and accuracy.
- For all items, *consistency* with similar/parallel data sets is an important requisite.
- 22. The fact-finding exercise indicates that **accuracy**, and **consistency** in frequencies and between alternative sources and interrelated statistical fields are regarded as priorities in long-term quality-oriented data production and assessment. **Stability** is seen as an intermediate quality element. Although the questionnaire did not explicitly mention **timeliness** and **plausibility**, timeliness is a key element confirmed in the note "Reassessing the scope for the euro area b.o.p." and plausibility is a feature present (or lacking) in all data.

2.1 Quality assessment/measurement

Assessing or measuring quality may imply a certain degree of subjective evaluation, which might differ from one statistical agency to another. However, this section is written with a view to reaching a common understanding as regards the assessment and measurement of the quality dimensions/elements relevant to euro area b.o.p./i.i.p. statistics. The assessment/measurement proposed focuses mostly on the product side of quality, although some reference to the process side is also made, when necessary

2.1.1 Relevance: qualitative assessment

- 24. In principle, a periodic review of the scope for data collection should be carried out by polling users, to discover any new perceived requirements and priority for improvements, etc. This can be covered by surveys of internal (ECB/ESCB) and some external users.
- 25. Users' needs should be stated and documented clearly for the compiler. The legal basis for data collection should be given, and general internal documentation defining the focus, use and analysis of the data compiled should be supplied.
- 26. To increase relevance, national compilers should be kept informed of whose needs are covered by the euro area b.o.p./i.i.p. data.

2.1.2 Timeliness: quantitative assessment

27. A quantitative assessment involves recording the number of days and recurrence of delayed delivery. Close monitoring is needed to discover the reasons behind any persistent delays, i.e. possible structural deficiencies in the data collection process at respondent level, and/or in Member States' data compilation/estimation procedures.⁹

B.o.p./i.i.p. statistics for the euro area are derived by aggregating each MS' transactions outside the euro area. A precondition for achieving timely delivery of data is that a clear timetable is set out for the provision of MSs' extra-euro area contributions to the ECB. Although estimations are accepted to cope with the first assessment for the euro area b.o.p., an active strategy notably toward reporting agents and the allocation of specific resources is expected to help, at country level, to limit the effects of insufficient real data on the quality of the euro area b.o.p., in particular its financial account and i.i.p. statistics.

28. Documented tools and procedures may assist in prompt checking of the consistency of the b.o.p. and i.i.p. statistics with all relevant sources. Information stemming from relevant sources should be checked regularly. Differences in figures between similar time series compiled by alternative sources should be monitored and commented on.

2.1.3 Stability: quantitative assessment

- 29. The analysis of stability involves a statistical analysis of the difference between first release and successive assessments, with special focus on first updates. Proper procedures, together with assigned responsibilities and adequate resources, are important for monitoring the compilation process. Member States could consider implementing a regular assessment of the stability of their output (provided to the ECB). Systematic revisions could be analysed and new procedures and/or redefined estimation methods considered, as regards: i) general data collection procedures (from respondents to compilers); ii) internal data compilation procedures; and iii) estimation procedures for missing, incorrect or incomplete data.
- 30. A more structured proposal on ways for assessing the stability is addressed in the note "Proposal for a standardised approach to measure stability of balance of payment series".

2.1.4 Consistency: quantitative assessment ¹⁰

- 31. Consistency is assessed by means of a comparison of: i) statistics from other independent fields measuring the same phenomenon; or ii) statistics belonging to the same field but compiled and/or disseminated by different sources. The analysis compares both the absolute and the relative sizes of the differences¹¹.
- 32. Under i) analysis of consistency may be undertaken as follows:
- between b.o.p. goods and external trade statistics;
- asymmetries observed with mirror statistics from other countries/zones;
- analysis of b.o.p. flows as a component (together with prices, foreign exchange and other adjustments) of the reconciliation table for changes in i.i.p. stocks;
- all b.o.p./i.i.p. items as rest-of-the-world sector for the (national) economic and financial accounts which may give feedback with, for example, other sources pertaining to resident sectors, and
- between b.o.p. "other investment"/monetary financial institution (MFI) sector and the balance sheet reported by the MFIs for the purpose of compiling money and banking statistics;
- the regular and documented reconciliation between b.o.p. flows and stocks.

Consistency, which is here defined with a focus on the product side (how to measure whether statistical outputs are consistent), can be approached from process side, with particular emphasis on the co-ordination of the updates and revision policies for interrelated statistics (co-ordination is a basic precondition for consistency).

In principle, consistency is reached by construction when statistics related to the same institutional sector are compiled using the same source, e.g. balance sheet data (besides the different revision practices). Conceptual consistency needs to be further assessed, as well as integrity of the compiling agency.

- 33. Under ii) analysis of consistency may mean comparing data with those disseminated by certain other organisations (IMF, Organisation for Economic Co-operation and Development (OECD), European Commission) bearing in mind the regional coverage they have in focus (e.g. European Union or the euro area) and/or the given methodology (e.g. for the compilation of an aggregate). ¹².
- 34. In addition, monthly, quarterly and annual data are expected to be coherent, though departures from agreed methodology are accepted for the provision of timely monthly b.o.p. data.

2.1.5 Accuracy: qualitative assessment

- 35. The assessment of accuracy derives from an evaluation of the soundness of the entire data production procedure and the exhaustiveness of the coverage. It covers the level and type of thresholds, the adjustments for non-response or late response, the regular updating (through census) of grossing-up for any sampling procedure and its confidence interval, etc. Documented adherence to internationally agreed statistical standards constitute an important indication of accuracy.
- 36. Persistently high errors and omissions cast some doubts on the quality of the disseminated data, though the opposite is not necessarily a sign of quality as many factors could work in different directions leading to a netting of opposite errors. Hence, size and sign of the errors and omissions in the published statistics can be deemed as an indicator in assessing the accuracy of the b.o.p. statistics.

2.1.6 Plausibility: quantitative assessment

- Procedures and tools should be set up to perform an ongoing statistical analysis of the trend of the b.o.p. data. This should encompass: i) analysis of the trend of individual items of the b.o.p.; ii) analysis, for gross flows, of the trend of the share of the extra euro area data in the national figures; iii) analysis of the pattern of single items' shares in global b.o.p., in particular within the current and capital accounts, and in direct investment.
- A practical solution would be: when an observation records a change over the previous reference period (month/quarter/year) under i) and ii) higher than "x" % in absolute value for current and capital account items and for direct investment (depending on the item), and higher than "y" % for the rest of the financial account, an investigation should be carried out to support the observation with proper explanation. The same applies for iii). When time series contain numerous outliers, quality assessment would better be based on cumulative figures.

For instance, statistics on the EU/euro area b.o.p. were disseminated at the beginning of 1999 by different international organisations (the IMF, the OECD) using a different approach (the so-called "step 1", addition of national data) from the one agreed between the ECB and the Commission (Eurostat) (the so-called "step 2", aggregation of extra-transactions). Similarly, national b.o.p. data are disseminated both by the countries' compiling institutions and by international organisations in which the countries take part. Owing to – for example – different revision practices, data may differ significantly. Given the increased accessibility of the data, world-wide users usually perform several cross-checks of this kind using different sources.

2.1.7 Integrity: qualitative assessment

39. Users' confidence in the integrity of ESCB statistical authorities is enhanced by the following: i) the dissemination of the terms and conditions under which statistics are produced, including those relating to the confidentiality of individually identifiable information; ii) the absence or disclosure of any access by a central government agency to data before release; iii) the disclosure of ministerial commentary prior to statistical releases.

2.1.8 Accessibility: qualitative assessment

40. The dissemination of statistics in electronic form on the ECB and NCB official websites, with a user-friendly browser and interfaces, is a way to improve accessibility. A printed publication may no longer suffice when markets need easy and equal access to information. The accessibility of data may be seen as a dynamic dimension, given the developments in information systems and technologies. The interaction with users should also be given extensive consideration, in particular as concerns feedback on data presentation and content quality. The ECB's Directorate General Statistics is engaged in a project to improve the accessibility of data on the ECB's website.

2.1.9 Transparency: qualitative assessment

41. Transparency is so far assured, thanks to the publication of the "B.o.p. Book" (by the ECB) and the gentleman's agreements (by Eurostat). In addition, the IMF's website provides useful metadata under the dissemination standard.

2.2 The interrelation between quality dimensions / elements

- 42. Up to now, the quality dimensions have been considered in isolation. In practice, these dimensions are interrelated, and in some cases there is a trade-off between them. It is important to have a clear picture of this interrelation, in order to approach the quality issue in a systematic manner. Some trade-off cases are listed below.
- Between timeliness and accuracy: the common understanding is that the shorter the deadline, the more challenging it is to achieve accuracy. However, it is possible to underline several aspects where (to some extent) improvements in timeliness can be achieved without being detrimental to accuracy. Amongst others: i) a reorganisation aiming at giving higher priority by statistical agencies to euro area data; ii) exchange of information on compiling methods (aiming to target towards best practices and fostering consistent compilation/estimation methods); iii) the review of practical (and legal) constraints contributing to the late compilation of first results; iv) a review of all stages of national collection and processing systems (benchmarking of the group of best performing Member States).
- 44. Between stability and accuracy: although stable data are appreciated by users, stability could indicate that additional (more comprehensive) information is not being used to enhance the picture given in the first assessment; even more worryingly, it could suggest that deficiencies in the first compilation of the observation are being kept undisclosed for some time. In this sense, stable data might not show a picture of

reality. An additional case would be where a methodological change is made but back observations are not revised, so little information is available on the trend. At the same time high instability (whether disclosed or not) is an indication of the potential inappropriateness of the data collection and compilation procedure to cope with the required timeliness. Similarly, if the errors and omissions item, in part or in full, were hidden or incorporated elsewhere in the accounts, then a low value would not necessarily mean high quality b.o.p. statistics.

- 45. Between stability and integrity: any statistical agency which publishes data is expected (in particular for countries subscribing to the IMF's Special Data Dissemination Standard) to deliver an accurate picture of recent as well as previous periods, according to an advanced release calendar. Despite all effort to produce relevant statistical information, it is unlikely that data are compiled at the first assessment using the complete set of information needed to provide the most accurate figures. Thus, the more timely the data, the more they are subject to subsequent revisions. However, revisions of a large magnitude would be a sign of lack of accuracy in the data collection and/or compilation process, and hence, in their integrity. In special cases, significant revisions may lead to the data collection and compilation system being reviewed and improved. A balance is needed between "frozen" data and a Boolean motion; it may rely on professional standards, experience gained and proactive attitude (for example, vis-à-vis users) on the part of the statistical agency.
- 46. Meeting the various quality dimensions and elements generally leads to significant burden on respondents and compiling agencies. In particular accuracy of detailed (e.g. geographical) breakdowns may become burdensome in the case of sampling surveys as it may swiftly increase the size of the sample. It has also been observed that the number of questions in a survey may decrease the accuracy and directly conflict with timeliness. While these issues are important, they do not enter into the scope of this work and are, instead, considered within the merits and costs approach for any development in existing statistics or new statistical needs.

Part Three: time horizon for implementation of minimum quality standards and monitoring of euro area b.o.p./i.i.p. statistics

- 47. In order to approach the implementation of quality standards, it may be useful to distinguish between two time horizons: for a transitional approach and for the steady-state. The framework underlying this distinction could be as follows:
- the <u>transitional approach</u> refers to the time frame in which the Member States' current data collection systems are still in place. Within this time horizon, **quality implementation is mainly limited to the output side**. However, Member States with similar data collection systems should aim at best practice, which might already imply improvements on the input side;
- the <u>steady-state</u> refers to the time when revised and more harmonised data collection systems will be in place. To the extent that some accounts of the b.o.p./i.i.p. would require a higher degree of

harmonisation (in particular portfolio investment flows, stocks and related income), the minimum quality standards may cover both the product and process.

- 48. The joint ECB/Commission paper on the future of EU/euro area b.o.p./i.i.p. data collection systems proposes the setting up of minimum quality standards to be developed in the WG-BP&ER and BOP WP to assess contributions by Member States to euro area/EU aggregates. The aim is to make clear the criteria to be met when designing new data collection systems. Indeed, the collection and compilation methods applied may significantly vary depending on the size and characteristics of each country.
- 49. Therefore the successful implementation of (or the move towards) new systems depends on the provision of quality standards for the Member States' contributions, showing what is necessary to compile timely and reliable euro area/EU aggregates. In setting up this long-term objective, the European Commission (Eurostat) and the ECB (Directorate General Statistics) will co-operate to achieve convergence in the definition of quality and in benchmarking.
- 50. To this end both product and process aspects have to be considered. The common understanding as regards the minimum quality standards for euro area aggregates may be that
- the higher the envisaged degree of harmonisation in data collection (for instance, particularly in the field of portfolio investment and related income) the more aspects related to the data *production process* would be involved. By its nature a common product, collected and compiled using harmonised methods, might call for a specific quality process, done under agreed and recognised terms;
- for other items, where no or almost no harmonisation on the input side is envisaged, quality will be assessed mainly on the *product side* by assessing and measuring, where relevant, the quality dimensions and elements aforementioned.

Summary of EU Member States' views on the importance of the quality dimensions for the b.o.p. items¹³

Dimensions of quality		Exhaustivenes	Consistency	Degree	Regular	Monthly -
	of Accurac	s of sources	with other statistics	of stability	Revision	quarterly consistency
B.o.p. items	Y			Stability		consistency
Goods			Trade figures, SNA		•	
Services			travel / tourism statistics.		•	
			National Accounts			
Compensation of employees			National Accounts,			
			Bilateral data, EU			
Investment Income			b.o.p. National Accounts.			
investment income			I.I.P.			
Current transfers			National Accounts,			
			EU b.o.p.			
			EU transfers, Govt budgets			
Capital account			National Accounts,			
-			EU b.o.p.			
			EU transfers, Govt			
Direct investment income			budget I.I.P., SNA,			
Biredi investment income			National Accounts			
Direct investment			I.I.P. survey, MUFA,		•	
			bilateral data			
Portfolio investment income						
Portfolio investment	•				•	
Financial derivatives			MFI BS, I.I.P., MUFA		•	
011 1 11			stock exchange			
Other investment income			I.I.P., National Accounts,		•	
			SNA			
Other investment			MFI balance sheets,		•	
			MUFA			

•	high importance					
	medium importance					

o low importance

Refers to the outcome of the fact-finding exercise on quality priority assessment, held during the September 2000 ECB Working Group on BP&ER thematic meeting on the future of data collection systems.