



Committee for the Coordination of Statistical Activities
Conference on Data Quality for International Organizations
Newport, Wales, United Kingdom, 27 - 28 April 2006



Session 4: Peer review methodologies

“Peer Review for International Macroeconomic Statistics: The IMF’s Standards and Codes Initiative and its Data Module”

*Claudia Dziobek*¹

A. Data ROSCs are Part of a Bigger Initiative

1. The standards and codes initiative was launched in 1999 as a prominent component of efforts to strengthen the international financial architecture. The initiative was designed to promote greater financial stability, at both the domestic and international levels, through the development, dissemination, adoption, and implementation of international standards and codes. Its three intermediate objectives include: to assist countries in making progress in strengthening their economic institutions, to inform Bank and Fund work, and to inform market participants. It covers 12 areas and associated standards, which the Bank and Fund Boards recognized as relevant for their work. These standards relate to data and policy transparency, financial sector regulation and supervision, and market integrity.
2. As of March 31, 2006, 724 reports were completed in 130 countries, thus covering 70 percent of the IMF’s membership. Participation, which is voluntary, in general has been high for emerging market countries and advanced economies, and somewhat lower for developing countries. Regional participation is somewhat uneven, with participation highest in Europe, including Eastern Europe, and lowest for East Asia and Sub-Saharan Africa. There is some evidence of “self-selection” with best performers more willing to participate than poor performers. The publication rate of ROSCs has been stable at about 75 percent and the highest publication rates are for advanced economies and for fiscal and data standards reports, with data standards reaching an over 90 percent publication rate.

¹Advisor, Statistics Department, International Monetary Fund, Washington, D.C. This paper is a work in progress and the author would welcome any comments or questions, which can be addressed to cdziobek@imf.org, or by telephone: 202 623-4874.

B. Experience with Standards and Codes

3. In mid 2005, the IMF Executive Board reviewed the experience with the standards and codes initiative and concluded that overall, the initiative has delivered substantial results in some dimensions, notably identifying vulnerability and establishing priorities for strengthening domestic institutions. The impact on actual implementations of reforms may not yet have been as substantial, but neither has it been insignificant. In this respect, it may be considered that the initiative is still relatively new given the typical timeframe of institutional reforms.
4. A somewhat unexpected result of the survey of users, which was at the heart of the review, was that the initiative fell short of its objective of informing market participants. The survey found that about half of respondents made little or no use of ROSCs and the result confirmed the outcome of an earlier survey. Market participants favored inclusions of quantitative measures of compliance, conduct of substantive updates on an annual frequency. In this respect, the data ROSC contains a summary table with quantifiable results, and it is envisioned that in future, most if not all of the ROSCs will have similar summary assessment tables that can be used for quantitative analysis. Nevertheless, the standards and codes initiative remains primarily a structured description with expert judgments of policies, processes, and transparency.

C. Focus on Emerging Market Countries and More Frequent Updates

5. Given the findings of the review, the IMF is adjusting the prioritization and country coverage of all ROSCs. Realizing that the initiative's resources are spent where they are most needed and given increasing cost constraints, future ROSC programs will be more selective and more focused. Given the initiative's origin, the review suggests that the initiative appears to have had the most beneficial impact for emerging market countries. This may be attributed to the relevance of the initiative's standards for these countries, given their level of financial development, which contrasts with the situation of most developing countries. Going forward, the IMF's medium term strategy notes that the overall ROSC program will be curtailed, but this is unlikely to affect the coverage of emerging countries. At the same time, there will be somewhat more emphasis on more frequent updates, responding in particular to the needs of market participants.

D. Data ROSCs Published to Date

6. As of end-March 2006, a total of 91 data ROSCs have been published for 82 countries, covering about 45 percent of IMF membership. More than half of these are countries that subscribe to the Special Data Dissemination Standard (SDDS). In terms of the regional distribution, using the IMF's area department grouping, about 41 percent of total data ROSCs are European, 20 percent Africa, 15 percent Middle East/Central Asia, 13 percent Western Hemisphere, and 11 percent Asia.

E. Weaknesses Identified in the Data ROSCs

7. Data ROSCs provide an in-depth analysis of a country's statistical system, based on the Data Quality Assessment Framework (DQAF). The DQAF is a structured set of criteria describing good statistical practices along the lines of six quality dimensions which are further broken down into a total of 22 elements. Since 2001, these reports include a summary table of ratings for each of the 22 elements of the DQAF, and for each dataset. A four point scale ranging from observed (O), largely observed (LO), largely not observed (LNO) to not observed (NO) is applied. Most reports cover five datasets, national accounts statistics, consumer price and producer price indices, government finance statistics, monetary and financial statistics and balance of payments statistics.² As of April 15, 2006, such ratings are available for 73 countries.³ In the following, an illustration is given of the type of information that can be obtained from these ratings. The discussion focuses on three datasets as explained below.

8. A simple average of the 22 ratings for each dataset of the 73 countries provides an overall indication of the weaknesses identified by the assessments. It shows that national accounts statistics and government finance statistics receive overall the lowest ratings. About 50 percent of ratings for national accounts and 54 percent of government finance statistics are less than "observed," of which 12 (national accounts) and 14 percent ranked as "largely not observed (LNO)" or not observed. This contrasts with other datasets, for instance the consumer price index, which has only about 5 percent in the LNO category and only about 33 percent below O. Balance of payments statistics are the third lowest ranked dataset with 46 percent of ratings below the highest grade and about 9 percent of ratings at LNO or even below at "not observed". Performing the same exercise just for 8 leading industrial countries confirms these findings with respect to government finance statistics, which receive the lowest ratings, but with balance of payments and monetary statistics receiving slightly lower marks than national accounts statistics while the ratings for the consumer price index is close to perfect score. Industrial economies generally fare better than emerging countries, but not always and there are significant differences within these groups. However, direct comparisons across countries require a considerably deeper analysis and an analysis of the ratings does not yield sufficiently meaningful results.

² A DQAF for external debt statistics is also available but it has not been applied for ROSC reports.

³ Not all data ROSCs contain ratings. In some cases, update reports were prepared without new ratings and during the experimental initial phase of the ROSC program, country reports were prepared without ratings.

9. Some further insights can be gained by considering how ratings are distributed across the various elements of the Data Quality Assessment Framework (DQAF). The ratings are broken down for national accounts, government finance statistics and for balance of payments, the three lowest ranked datasets. By breaking down the rating distributions across elements, it is possible to identify specific practices that are assessed as weaknesses.
10. A tabulation of the ratings for national accounts statistics for all 73 countries shows 10 areas, where 50 percent or more countries received ratings below “O” and five areas where 70 percent of ratings or more are below “O”. Noting the DQAF codes in brackets, the weaknesses of 70 percent of countries are concentrated in: resources (0.2), basis for recording (2.4), source data (3.1), statistical techniques (3.3), and revisions policy (4.3).
11. Concerning Government Finance Statistics, there are 13 elements with ratings less than “observed” in half of the countries and 7 areas with 70 percent or more ratings below O. Again, focusing on those where 70 percent or more are below this level, the following emerge as areas of weaknesses: resources (0.2), relevance (0.3), scope (2.2), consistency (4.1), revisions policy (4.3), and data and metadata access (5.1 and 5.2).
12. For balance of payments statistics, 10 elements have ratings where 50 percent or more are below the level of fully observed but all are above the 70 percent range. The lowest five ratings (all between 60 and 70 percent in the less than observed group) are in scope (2.2), classification (2.3), basis of recording (2.4), source data (3.1), and revision policy and practice (4.3). These are nevertheless included in the comparison presented below in Table 1.

Table 1. Weakest Areas of Macroeconomic Statistics as identified in the Reports on the Observance of Standards and Codes (ROSCs) for 73 Countries (April 2006)*

DQAF Elements	Datasets and their lowest ratings		
	NA	GFS	BOP

0.2 Resources	x	X	
0.3 Relevance		x	
2.2 Scope		x	x
2.3 Classification			x
2.4 Basis for Recording	x		x
3.1 Source Data	x		x
3.3 Techniques	x		
4.1 Consistency		x	
4.3 Revisions policy	x	x	x
5.1 Data access		x	
5.1 Metadata access		x	
Total # of Weak Areas	5	7	5

Source: IMF Country ROSCs are on the IMF External Website. [Dsbb.imf.org](http://dsbb.imf.org) (to DQRS).

NA = National Accounts Statistics, GFS = Government Finance Statistics, DQAF = Data Quality Assessment Framework, which is also on the IMF website.

Bold: Elements identified as weaknesses in all three datasets.

*Weakest areas are defined as those where 70 percent of countries receive ratings below the highest mark. For Balance of Payments Statistics, where none of the elements meet this benchmark, weakest areas refer to countries with 60 percent or more ratings below the highest grade.

Table 1 provides an overview of the weaknesses which data module ROSCs highlight for the three relatively weakest dataset, national accounts statistics, government finance and balance of payments statistics. As shown in bold, revisions policy is an area of weakness for all three datasets. This exercise shows that weak areas of national accounts and balance of payments are concentrated in just 5 of the 22 elements while government finance statistics show signs of weakness in a broader range of elements. This analysis thus yields some interesting results, although more meaningful insights could be developed by making fuller use of the data module reports and the wealth of metadata which is structured along the lines of the DQAF supporting the ratings. While remaining a rough first cut, this short analysis may motivate some further research

F. Conclusions

13. The purpose of this paper was to provide some illustrations of the findings of the IMF's standard and codes initiative, and in particular its data module. Data module ROSCs are in essence a system of peer group assessments, using the Data Quality Assessment Framework as the measuring rod. The data module reports contain ratings for each of the 22 elements of the DQAF and thus provide a wealth of detailed information about countries' practices and their strengths and weaknesses.
14. While the ratings are useful to provide a first quick identification of areas of weakness, the IMF is well aware of the limitations of such ratings. The findings presented in Table 1 might be a starting point to develop a variety of applications, including more focused technical assistance programs. It would be appropriate to

tabulate the findings by economically meaningful country groups or for countries of a region.⁴ More importantly, detailed metadata contained in the reports should be consulted to analyze the exact nature of the various weak areas in each country. Also, the recommendations contained in each of the reports should be analyzed for a fuller picture of what actions are viewed as most effective by the assessors in order to address the weaknesses. The data module reports contain a wealth of information, which is highly relevant for a variety of users, supporting national statistical plans, statistical activities on the regional and on the global level.

⁴ For example the IMF Statistics Department recently launched a study of the findings of data ROSCs for a group of European countries focusing on government finance statistics.



Committee for the Coordination of Statistical Activities
Conference on Data Quality for International Organizations
Newport, Wales, United Kingdom, 27 - 28 April 2006



Session 4: Peer review methodologies

“Peer Review for International Macroeconomic Statistics: The IMF’s Standards and Codes Initiative and its Data Module”

*Claudia Dziobek*¹

A. Data ROSCs are Part of a Bigger Initiative

1. The standards and codes initiative was launched in 1999 as a prominent component of efforts to strengthen the international financial architecture. The initiative was designed to promote greater financial stability, at both the domestic and international levels, through the development, dissemination, adoption, and implementation of international standards and codes. Its three intermediate objectives include: to assist countries in making progress in strengthening their economic institutions, to inform Bank and Fund work, and to inform market participants. It covers 12 areas and associated standards, which the Bank and Fund Boards recognized as relevant for their work. These standards relate to data and policy transparency, financial sector regulation and supervision, and market integrity.
2. As of March 31, 2006, 724 reports were completed in 130 countries, thus covering 70 percent of the IMF’s membership. Participation, which is voluntary, in general has been high for emerging market countries and advanced economies, and somewhat lower for developing countries. Regional participation is somewhat uneven, with participation highest in Europe, including Eastern Europe, and lowest for East Asia and Sub-Saharan Africa. There is some evidence of “self-selection” with best performers more willing to participate than poor performers. The publication rate of ROSCs has been stable at about 75 percent and the highest publication rates are for advanced economies and for fiscal and data standards reports, with data standards reaching an over 90 percent publication rate.

¹Advisor, Statistics Department, International Monetary Fund, Washington, D.C. This paper is a work in progress and the author would welcome any comments or questions, which can be addressed to cdziobek@imf.org, or by telephone: 202 623-4874.

B. Experience with Standards and Codes

3. In mid 2005, the IMF Executive Board reviewed the experience with the standards and codes initiative and concluded that overall, the initiative has delivered substantial results in some dimensions, notably identifying vulnerability and establishing priorities for strengthening domestic institutions. The impact on actual implementations of reforms may not yet have been as substantial, but neither has it been insignificant. In this respect, it may be considered that the initiative is still relatively new given the typical timeframe of institutional reforms.
4. A somewhat unexpected result of the survey of users, which was at the heart of the review, was that the initiative fell short of its objective of informing market participants. The survey found that about half of respondents made little or no use of ROSCs and the result confirmed the outcome of an earlier survey. Market participants favored inclusions of quantitative measures of compliance, conduct of substantive updates on an annual frequency. In this respect, the data ROSC contains a summary table with quantifiable results, and it is envisioned that in future, most if not all of the ROSCs will have similar summary assessment tables that can be used for quantitative analysis. Nevertheless, the standards and codes initiative remains primarily a structured description with expert judgments of policies, processes, and transparency.

C. Focus on Emerging Market Countries and More Frequent Updates

5. Given the findings of the review, the IMF is adjusting the prioritization and country coverage of all ROSCs. Realizing that the initiative's resources are spent where they are most needed and given increasing cost constraints, future ROSC programs will be more selective and more focused. Given the initiative's origin, the review suggests that the initiative appears to have had the most beneficial impact for emerging market countries. This may be attributed to the relevance of the initiative's standards for these countries, given their level of financial development, which contrasts with the situation of most developing countries. Going forward, the IMF's medium term strategy notes that the overall ROSC program will be curtailed, but this is unlikely to affect the coverage of emerging countries. At the same time, there will be somewhat more emphasis on more frequent updates, responding in particular to the needs of market participants.

D. Data ROSCs Published to Date

6. As of end-March 2006, a total of 91 data ROSCs have been published for 82 countries, covering about 45 percent of IMF membership. More than half of these are countries that subscribe to the Special Data Dissemination Standard (SDDS). In terms of the regional distribution, using the IMF's area department grouping, about 41 percent of total data ROSCs are European, 20 percent Africa, 15 percent Middle East/Central Asia, 13 percent Western Hemisphere, and 11 percent Asia.

E. Weaknesses Identified in the Data ROSCs

7. Data ROSCs provide an in-depth analysis of a country's statistical system, based on the Data Quality Assessment Framework (DQAF). The DQAF is a structured set of criteria describing good statistical practices along the lines of six quality dimensions which are further broken down into a total of 22 elements. Since 2001, these reports include a summary table of ratings for each of the 22 elements of the DQAF, and for each dataset. A four point scale ranging from observed (O), largely observed (LO), largely not observed (LNO) to not observed (NO) is applied. Most reports cover five datasets, national accounts statistics, consumer price and producer price indices, government finance statistics, monetary and financial statistics and balance of payments statistics.² As of April 15, 2006, such ratings are available for 73 countries.³ In the following, an illustration is given of the type of information that can be obtained from these ratings. The discussion focuses on three datasets as explained below.
8. A simple average of the 22 ratings for each dataset of the 73 countries provides an overall indication of the weaknesses identified by the assessments. It shows that national accounts statistics and government finance statistics receive overall the lowest ratings. About 50 percent of ratings for national accounts and 54 percent of government finance statistics are less than "observed," of which 12 (national accounts) and 14 percent ranked as "largely not observed (LNO)" or not observed. This contrasts with other datasets, for instance the consumer price index, which has only about 5 percent in the LNO category and only about 33 percent below O. Balance of payments statistics are the third lowest ranked dataset with 46 percent of ratings below the highest grade and about 9 percent of ratings at LNO or even below at "not observed". Performing the same exercise just for 8 leading industrial countries confirms these findings with respect to government finance statistics, which receive the lowest ratings, but with balance of payments and monetary statistics receiving slightly lower marks than national accounts statistics while the ratings for the consumer price index is close to perfect score. Industrial economies generally fare better than emerging countries, but not always and there are significant differences within these groups. However, direct comparisons across countries require a considerably deeper analysis and an analysis of the ratings does not yield sufficiently meaningful results.

² A DQAF for external debt statistics is also available but it has not been applied for ROSC reports.

³ Not all data ROSCs contain ratings. In some cases, update reports were prepared without new ratings and during the experimental initial phase of the ROSC program, country reports were prepared without ratings.

9. Some further insights can be gained by considering how ratings are distributed across the various elements of the Data Quality Assessment Framework (DQAF). The ratings are broken down for national accounts, government finance statistics and for balance of payments, the three lowest ranked datasets. By breaking down the rating distributions across elements, it is possible to identify specific practices that are assessed as weaknesses.
10. A tabulation of the ratings for national accounts statistics for all 73 countries shows 10 areas, where 50 percent or more countries received ratings below “O” and five areas where 70 percent of ratings or more are below “O”. Noting the DQAF codes in brackets, the weaknesses of 70 percent of countries are concentrated in: resources (0.2), basis for recording (2.4), source data (3.1), statistical techniques (3.3), and revisions policy (4.3).
11. Concerning Government Finance Statistics, there are 13 elements with ratings less than “observed” in half of the countries and 7 areas with 70 percent or more ratings below O. Again, focusing on those where 70 percent or more are below this level, the following emerge as areas of weaknesses: resources (0.2), relevance (0.3), scope (2.2), consistency (4.1), revisions policy (4.3), and data and metadata access (5.1 and 5.2).
12. For balance of payments statistics, 10 elements have ratings where 50 percent or more are below the level of fully observed but all are above the 70 percent range. The lowest five ratings (all between 60 and 70 percent in the less than observed group) are in scope (2.2), classification (2.3), basis of recording (2.4), source data (3.1), and revision policy and practice (4.3). These are nevertheless included in the comparison presented below in Table 1.

Table 1. Weakest Areas of Macroeconomic Statistics as identified in the Reports on the Observance of Standards and Codes (ROSCs) for 73 Countries (April 2006)*

DQAF Elements	Datasets and their lowest ratings		
	NA	GFS	BOP

0.2 Resources	x	X	
0.3 Relevance		x	
2.2 Scope		x	x
2.3 Classification			x
2.4 Basis for Recording	x		x
3.1 Source Data	x		x
3.3 Techniques	x		
4.1 Consistency		x	
4.3 Revisions policy	x	x	x
5.1 Data access		x	
5.1 Metadata access		x	
Total # of Weak Areas	5	7	5

Source: IMF Country ROSCs are on the IMF External Website. [Dsbbs.imf.org](http://dsbb.imf.org) (to DQRS).

NA = National Accounts Statistics, GFS = Government Finance Statistics, DQAF = Data Quality Assessment Framework, which is also on the IMF website.

Bold: Elements identified as weaknesses in all three datasets.

*Weakest areas are defined as those where 70 percent of countries receive ratings below the highest mark. For Balance of Payments Statistics, where none of the elements meet this benchmark, weakest areas refer to countries with 60 percent or more ratings below the highest grade.

Table 1 provides an overview of the weaknesses which data module ROSCs highlight for the three relatively weakest dataset, national accounts statistics, government finance and balance of payments statistics. As shown in bold, revisions policy is an area of weakness for all three datasets. This exercise shows that weak areas of national accounts and balance of payments are concentrated in just 5 of the 22 elements while government finance statistics show signs of weakness in a broader range of elements. This analysis thus yields some interesting results, although more meaningful insights could be developed by making fuller use of the data module reports and the wealth of metadata which is structured along the lines of the DQAF supporting the ratings. While remaining a rough first cut, this short analysis may motivate some further research

F. Conclusions

13. The purpose of this paper was to provide some illustrations of the findings of the IMF's standard and codes initiative, and in particular its data module. Data module ROSCs are in essence a system of peer group assessments, using the Data Quality Assessment Framework as the measuring rod. The data module reports contain ratings for each of the 22 elements of the DQAF and thus provide a wealth of detailed information about countries' practices and their strengths and weaknesses.
14. While the ratings are useful to provide a first quick identification of areas of weakness, the IMF is well aware of the limitations of such ratings. The findings presented in Table 1 might be a starting point to develop a variety of applications, including more focused technical assistance programs. It would be appropriate to

tabulate the findings by economically meaningful country groups or for countries of a region.⁴ More importantly, detailed metadata contained in the reports should be consulted to analyze the exact nature of the various weak areas in each country. Also, the recommendations contained in each of the reports should be analyzed for a fuller picture of what actions are viewed as most effective by the assessors in order to address the weaknesses. The data module reports contain a wealth of information, which is highly relevant for a variety of users, supporting national statistical plans, statistical activities on the regional and on the global level.

⁴ For example the IMF Statistics Department recently launched a study of the findings of data ROSCs for a group of European countries focusing on government finance statistics.