Australian Bureau of Statistics

RESERVE BANK OF AUSTRALIA

APRA

REPORTING PRACTICE GUIDE

RPG 702.0 ABS/RBA Data Quality for the EFS Collection

March 2018

Contents

About this guide

Reporting practice guides (RPGs) provide guidance on sound practice in particular areas. This RPG provides guidance on managing data quality for entities reporting under the economic and financial statistics (EFS) data collection.

Terms that are defined in *Reporting Standard ARS 701.0 ABS/RBA Definitions for the EFS Collection* or in this RPG appear in bold italics.

This guide should be read in conjunction with:

- the ***EFS collection***, including *Reporting Standard ARS 701.0 ABS/RBA Definitions for the EFS Collection* and *Reporting Practice Guide RPG 701.0 ABS/RBA Reporting Concepts for the EFS Collection*, which contains definitions of, and guidance about, the data to be reported; and

- *Prudential Practice Guide CPG 235 Managing Data Risk*.

This guide does not seek to provide an all-encompassing framework, or to replace or endorse existing industry standards and guidelines.

Subject to reporting requirements set out in the ***EFS reporting standards***, an EFS reporting entity has the flexibility to manage its reporting for the ***EFS collection*** in a manner that is best suited to its business. Not all of the practices outlined in this RPG will be relevant for every EFS reporting entity and some aspects may vary depending upon the size, complexity and systems configuration of the EFS reporting entity.

Glossary

In this Reporting Guidance:

ABS	The Australian Bureau of Statistics established under the *Australian Bureau of Statistics Act 1975*.
ADI	An authorised deposit-taking institution within the meaning of the *Banking Act 1959*.
Agencies	The ABS and RBA.
APRA	The Australian Prudential Regulation Authority established under the *Australian Prudential Regulation Authority Act 1998*.
CPG 235	*Prudential Practice Guide CPG 235 Managing Data Risk*.
Data item	The information required to be entered in a specific cell of a form.
EFS	The economic and financial statistics collected by APRA through the ***EFS collection***.
EFS collection	The ***EFS reporting standards*** and data collected under the ***EFS reporting standards***.
EFS reporting standard(s)	Has the meaning given in *Reporting Standard ARS 701.0 ABS/RBA Definitions*.
Flow	A ***data item*** with a reporting basis of 'during' the reporting period, as specified in the instructions for the relevant reporting standard.
High priority	A ***data item*** identified as such in the EFS Priority Listing for Data Items.
Large institution	An ***ADI*** or ***RFC*** with greater than or equal to $200 billion in total assets measured on a ***domestic books*** basis.
RBA	The Reserve Bank of Australia established under the *Reserve Bank Act 1959*.
Registered Financial Corporations (RFCs)	Corporations that are registered entities under the *Financial Sector (Collection of Data) Act 2001)*.

Reporting error	A difference between the data reported to ***APRA*** and the data required to be reported to ***APRA*** under ***EFS reporting standards*** that is outside the ***agencies***' expectations for data quality. A ***reporting error*** may arise at any point in the data's life cycle, including, but not limited to, data capture, processing, retention, preparation and submission of reports.
RPG 701.0	*Reporting Practice Guide RPG 701.0 ABS/RBA Reporting Concepts for the EFS Collection*.
Standard priority	A ***data item*** in an ***EFS reporting standard*** that is not ***a very high priority data item*** nor a ***high priority data item***.
Stock	A ***data item*** with a reporting basis of 'as at the end of' the reporting period, as specified in the instructions for the relevant reporting standard.
Very high priority	A ***data item*** identified as such in the EFS Priority Listing for Data Items.

Introduction

1. ***ADIs*** and ***RFCs*** required to report under the ***EFS reporting standards*** (reporting entities) must meet quality control requirements that require them to have in place systems, processes and controls to assure the reliability of reported information. These requirements are supported by *Prudential Practice Guide CPG 235 Managing Data Risk* (CPG 235), which provides guidance on steps entities can take in managing data risk, including assessing data quality by reference to fitness for use—the degree to which data is relevant and appropriate for the intended purpose.

2. This RPG provides guidance to assist reporting entities to meet quality control obligations and data risk management expectations in relation to the ***EFS reporting standards***. To assist, it outlines the intended purposes of the data collection and the ***agencies'*** assessment of the degree to which ***data items*** are fit for purpose.

Purpose of the EFS Collection

3. Data collected under ***EFS reporting standards*** are primarily used by the ***ABS*** and ***RBA*** for analysis, publication and policy-making purposes. The ***ABS*** uses the data to compile and publish key macroeconomic series, including Australia's National Accounts and leading indicators of lending activity, which are widely used to monitor Australia's economic growth. The ***RBA*** uses the data to construct and publish Australia's monetary and credit aggregates, and for analytical and policy purposes. The data published by the ***RBA*** are used by other policy makers and the wider public for research, analysis and policy making.

Managing data quality

4. The ***agencies*** expect reporting entities to place high importance on the quality of data being submitted in the ***EFS collection***. In particular, reporting entities are expected to give consideration to the design and implementation of controls throughout the data life cycle – including, but not limited to, data capture, processing, retention, preparation and submission of reports – to ensure that the data submitted are reliable. Reporting entities are also expected to review the ongoing effectiveness and appropriateness of these controls and any assumptions used in the preparation of reports throughout time.

5. Although there is a strong need for all of the data to be accurate, the ***agencies*** expect reporting entities to use the priority categorisation, with the benchmarks, as an indicator of where to focus data quality management practices and to understand the accuracy of the data required by the ***agencies***. This prioritisation has been incorporated into data quality benchmarks set out in Attachment A, which are calibrated according to the priority of the ***data item***, the size of the reporting entity and the type of ***data item***.

Benchmarks based on data priority

6. The priority ranking of the ***data items*** provides an indication of the relative importance of the accuracy of these ***data items*** to the ***agencies***, as primary users of the data. There are three categories of priority: '***standard***', '***high***' and '***very high***' ***priority*** items. The priority rankings of ***data items*** are set out in http://www.apra.gov.au/adi/PrudentialFramework/Documents/Priority%20listing%20for%20data%20items%20in%20the%20EFS%20collection.XLS.

7. Quantitative benchmarks apply to some ***data items***, as indicated in the tables in Attachment A. The quantitative benchmarks indicate the size of misreported ***data items*** that may impact the use of the data by the ***agencies*** and thus would be considered a ***reporting error***.

8. For other ***data items*** reporting entities are expected to exercise their judgement, taking into account the quantitative benchmarks and relative priority rankings, when determining what constitutes a ***reporting error***.

Benchmarks based on size of entity

9. The benchmarks differ according to entity size, to proportionately account for the impact of ***reporting errors*** on data quality in the ***EFS collection***.

10. Benchmarks for ***large institutions*** recognise that ***reporting errors*** by a single entity are more likely to impact industry aggregates due to their size. These benchmarks also serve to identify ***reporting errors*** relevant to the internal consistency of the entity's series.

11. Benchmarks for other reporting entities are aimed at identifying ***reporting errors*** relevant to the internal consistency of the entity's series and ***reporting errors*** that could affect the industry aggregate results if occurring across several entities simultaneously.

Notification

12. In the event of reporting errors, the ***agencies*** expect that reporting entities would notify ***APRA***. Depending on the size of the ***reporting error*** and potential impact on the ***agencies***' use of the data, ***APRA*** (in consultation with the ***agencies***) may require the data to be resubmitted.

13. ***APRA*** and the ***agencies*** also expect that, in the event of ***reporting errors***, a reporting entity would review its data quality processes and controls, including escalating knowledge of frequent or large ***reporting errors***.

Application of CPG 235 to the EFS collection and use of the data quality benchmarks

14. Good practice would be for a reporting entity to have regard to the guidance on managing data risk set out in CPG 235 when considering how to manage data quality, and, in particular, to the provisions relating to:

    a) Taking a structured and principles-based approach — data risk management is to be part of a systematic and formalised approach (paragraph 20 of CPG 235). As a foundation for managing data risk, CPG 235 envisages that an entity would assess data quality to ensure it is acceptable for the intended purpose of the data (paragraph 22(e)). The ***agencies*** expect that the data quality benchmarks will assist a reporting entity in understanding the size of a ***reporting error*** that may affect the ***agencies***' use of the data.

    b) Risk appetite and controls — Under CPG 235, ***APRA*** expects that data risk should be considered and appropriate controls implemented at each stage of the data life-cycle (paragraph 33), and be aligned to the entity's risk appetite (paragraphs 14-15). The ***agencies*** expect that a reporting entity would consider the data quality benchmarks when setting risk appetite for data quality and in the design, implementation and assessment of controls to manage EFS data quality.

    c) Data validation — CPG 235 considers data validation to be a key control for ensuring that data meets quality requirements and is assessed against fitness for use (paragraphs 51-52). For example, it would be prudent to have validation controls that manage the timeliness of data (CPG 235 definition: the degree to which data is up-to-date). The ***agencies*** expect a reporting entity to use the data quality benchmarks as part of data validation design, throughout the data's life-cycle.

    d) Monitoring and managing data issues — the ***agencies*** expect that the data quality benchmarks would be considered in monitoring and managing data issues relating to the ***EFS collection***. For example, where a data issue results in EFS data falling outside the data quality benchmarks, this would be a signal for an entity to consider an adjustment of controls. This could involve establishing a targeted data improvement program in consultation with the ***agencies*** that specifies target metrics, timeframes for resolution and associated action plans for closing any data quality gaps identified (paragraph 26). The ***agencies*** also expect the data quality benchmarks to be considered in the development of quality metrics (paragraph 64) relating to the ***EFS collection***, in order to report on the effectiveness of data risk management practices through time and to inform ongoing data improvement work.

    e) Assurance — CPG 235 provides guidance on a data risk management assurance program, including regular assurance that data quality is appropriate and data risk management is effective (paragraph 66). The ***agencies*** expect that the data quality benchmarks would be considered as part of this assurance program. For example, in considering whether the data risk management in place is consistent with the data quality benchmarks. The prioritisation of data items may also be useful for an entity to consider in setting multi-year assurance programs (paragraphs 67 and 68). The ***agencies*** consider that, to maintain the data quality over time, good practice is to conduct periodic themed deep-dive reviews of data and processes for a given set of forms or concepts.

Engagement with the agencies

15. Along with the sound data management risk practices set out in this guide, the ***agencies*** are of the view that data quality for the ***EFS collection*** will be improved through continued regular engagement between the ***agencies*** and reporting entities. As part of this approach, the ***agencies***, from time to time, may engage with reporting entities in a variety of ways including, but not limited to:

    a) *Discussion of reports on assurance processes* — ***APRA*** on behalf of the ***agencies*** may request a copy of documentation of findings from assurance processes, the recommendations given and actions taken based on those recommendations, to assist in further enhancing the standard of EFS reporting and to engage in dialogue on issues that may be impacting data quality.

    b) *Discussion of proxies and assumptions used* — ***APRA*** on behalf of the ***agencies*** may seek to engage with reporting entities to better understand the data being provided and the use and nature of proxies or assumptions used.

    c) *Peer workshops* — The findings of assurance processes and other initiatives (appropriately de-identified) may form the basis for peer workshops, which will serve as an opportunity for reporting entities, ***APRA*** and the ***agencies*** to discuss concerns and to highlight best practice. The workshops will provide an opportunity for:

        i) entities to outline areas of the instructions and guidance that are unclear or inadequate, and to discuss other reporting-related issues; and

        ii) developing practical solutions to issues and problems through discussions.

Attachment A – Data quality benchmarks

1. The data quality benchmarks are provided in Tables 1a and 1b and Tables 2a and 2b below.

Tables 1a and 1b – Benchmarks for data expressed as a dollar value, count or proportion

2. Tables 1a and 1b below set out the data quality benchmarks for identifying ***reporting errors*** for stock and flow ***data items*** expressed as a dollar value, count or proportion (e.g. term, tenor).

3. A reporting error is identified by reference to:

   a) for ***stock*** and ***flow data items*** reported as a dollar value—the difference between the reported amount and the correct amount expressed as a percentage of that series (for that institution) and/or as an absolute value.

   b) for ***stock*** and ***flow data items*** reported as a count or proportion (e.g. term, tenor)—the difference between the reported amount and the correct amount expressed as a percentage of that series (for that institution).

4. A misreported ***data item*** expressed as a dollar value, count or proportion that exceeds the benchmarks in Tables 1a or 1b constitutes a ***reporting error***.

5. For ***data items*** with a value close to or at zero, the *Application of benchmarks to series with zero value or near-zero value* section of this Attachment contains further guidance on determining ***reporting errors***.

6. For ***flow data items***, the *Application of benchmarks to volatile flow data* section of this Attachment contains further guidance on determining ***reporting errors***.

Table 1a: Benchmarks for identifying *reporting errors* for a *large institution*

Data item type	Priority	As percentage of series (%)	As absolute dollar value ($ million)
Stock	Very High	0.50	2,000
	High	5.00	
	Standard	Judgement	
Flow	Very High	5.00	250
	High	10.00	
	Standard	Judgement	

Table 1b: Benchmarks for identifying reporting errors for a reporting entity that is not a *large institution*

Data item type	Priority	As percentage of institutional series (%)	As absolute dollar value ($ million)
Stock	Very High	2.00	500
	High	10.00	
	Standard	Judgement	
Flow	Very High	10.00	100
	High	20.00	
	Standard	Judgement	

Example 1

7. A reporting entity that is not a large institution is using the benchmarks in Table 1b to assess whether misreporting on a high priority stock data item reported as a dollar value constitutes a ***reporting error***.

   The following scenarios indicate whether the misreported ***data item*** constitutes a reporting error:

   a) Case A: $50 million and representing 8 per cent of the value of that ***data item***. This would fall within agency expectations for data quality—and is not therefore a ***reporting error***—as the difference between the reported amount and correct amount is below the percentage (10 per cent) and below the maximum absolute dollar value ($500 million) benchmarks.

b) Case B: $50 million and representing 12 per cent of the value of that ***data item***. This would fall outside agency expectations for data quality—and is therefore a ***reporting error***—as the difference between the reported amount and correct amount is above the percentage (10 per cent) benchmark.

c) Case C: $550 million and representing 8 per cent of the value of that ***data item***. This would fall outside agency expectations for data quality—and is therefore a ***reporting error***—as the difference between the reported amount and correct amount is above the maximum absolute dollar value ($500 million) benchmark.

Example 2

8. A reporting entity that is not a large institution is using the benchmarks in Table 1b to assess whether misreporting on a high priority flow data item reported as a count constitutes a reporting error.

The following scenarios indicate whether the misreported data item constitutes a reporting error:

a) Case A: 25 per cent of the figure for that ***data item***. This would fall outside agency expectations for data quality—and is therefore a ***reporting error***—as the difference between the reported amount and correct amount is above the percentage (20 per cent) benchmark.

b) Case B: 15 per cent of the figure for that ***data item***. This would fall within agency expectations for data quality—and is not therefore a ***reporting error***—as the difference between the reported amount and correct amount is below the percentage (20 per cent) benchmark

Tables 2a and 2b – Benchmarks for data expressed as a rate

9. Tables 2a and 2b below provide data quality benchmarks expressed in basis points for ***data items*** reported as a rate (e.g. interest rates, margins, cost/value of funds, benchmark rate).

10. A misreported ***data item*** expressed as a rate that exceeds the benchmarks in Tables 2a or 2b constitutes a ***reporting error***.

Table 2a: Benchmarks for identifying *reporting errors* for a *large institution*

Priority	In basis points
Very high	5
Standard	15

Table 2b: Benchmarks for identifying *reporting errors* for a reporting entity that is not a *large institution*

Priority	In basis points
Very high	10
Standard	20

Example 3

11. A reporting entity that is not a ***large institution*** is using the benchmarks in Table 2b to assess whether misreporting on a ***standard priority data item*** reported as an interest rate constitutes a ***reporting error***.

    The following scenarios indicate whether the misreported data item constitutes a ***reporting error:***

    a) Case A: 25 basis points. This would fall outside agency expectations for data quality—and is therefore a ***reporting error***—as the difference between the reported amount and correct amount is above the 20 basis point benchmark.

    b) Case B: 15 basis points. This would fall within agency expectations for data quality—and is therefore not a ***reporting error***—as the difference between the reported amount and correct amount is below the 20 basis point benchmark.

Change of calculation methodology

12. For cost/value of funds, margin and benchmark rate data, the ***agencies*** do not expect that changes to a reporting entity's internal calculation methodology would be classified as a ***reporting error***. The ***agencies*** do, however, expect that changes to internal calculation methodologies expected to have a material impact on the data reported would be discussed with ***APRA*** and the ***agencies***. As part of this discussion, the ***agencies*** would expect the reporting entity to be able to provide a quantitative estimate of the impact of this methodological change on the EFS data; however, the ***agencies*** understood that a comprehensive impact assessment is unlikely to be available for all items affected by the methodological change.

13. When a reporting entity becomes aware of a change to that institution's internal calculation methodology for transfer pricing that is expected to have a material impact on the cost/value of funds, margin and/or benchmark rate data reported then the reporting entity is requested to contact ***APRA*** with a quantitative estimate of the impact of this methodological change and the date it will become effective. Where the change is large, the ***agencies*** may seek further information from the reporting entity.

Application of judgement to misreported data items

14. It would be good practice for a reporting entity to have policies and procedures in place that outline how they will apply judgement when determining whether a ***standard priority data item*** expressed as a dollar value, count or proportion is within or outside the agency

expectations for data quality (e.g. if it constitutes a ***reporting error***). The ***agencies*** expect that these policies and procedures would cover ***data items*** on two-dimensional, multi-dimensional and trade-level reporting tables.

Application of benchmarks to series with zero value or near-zero value

15. Where a ***data item*** is at or very close to zero, percentage benchmarks are unlikely to be helpful for determining whether a ***reporting error*** is outside agency expectations for data quality. As a general rule, misreporting of less than $25 million for a large institution and less than $10 million for a reporting entity that is not a ***large institution*** would not be considered as being outside the ***agencies***' expectations for data quality irrespective of the benchmarks in Table 1a and Table 1b.

Application of benchmarks to volatile flow data

16. To assess the magnitude of a ***reporting error*** for a ***flow data item*** that naturally exhibits significant period-to-period volatility, it may be appropriate to consider the difference between the data item reported and the 3- or 6-month average of the series that is required to be reported.

Use of proxy methodologies

17. RPG 701.0 guides reporting entities on the use of a proxy methodology for selected ***data items***. The guidance on the use of proxy methodologies for these series recognises the operational challenges in reporting certain EFS data.

18. Where RPG 701.0 allows use of a proxy methodology, the data quality benchmarks apply to misreporting determined by reference to the appropriately calculated proxy measure. That is, the benchmarks are applied to misreporting measured as the difference between the reported ***data item*** and that calculated using the appropriate methodology for the proxy measure.

19. For example, to report the categorisation of housing loans by state a reporting entity is using the permitted proxy methodology of allocating on the basis of the location of collateral rather than the standard treatment of allocating on the basis of the location of the property for which the funds were used. This reporting entity discovers that some ***data items*** have been misreported due to an error in the allocation of housing loans to states on the basis of the location of the collateral. The size of the ***reporting error*** would be assessed by comparing the reported data item (or items) to those calculated using the correct methodology for allocating housing loans to states on the basis of the location of collateral. For these items calculated using a permitted proxy methodology, the ***reporting error*** would not be assessed as the difference between the reported ***data item*** (or items) and that (those) calculated using the correct methodology for allocating housing loans to states on the basis of the standard treatment (the location of the property for which the funds were used).

20. Refer to RPG 701.0 for the selected ***data items*** that can be subjected to proxy methodology.

APRA