REPORTING PRACTICE GUIDE

RPG 702.0 ABS/RBA Data Quality for the EFS Collection

August 2017

Contents

About this guide

Reporting practice guides (RPGs) provide guidance on **APRA**'s view of sound practice in particular areas. RPGs frequently discuss legal requirements from legislation, regulations or APRA's prudential and reporting standards, but do not themselves create enforceable requirements.

This RPG provides guidance on managing data quality when reporting under the economic and financial statistics (EFS) data collection. Through this guide, **APRA**, and the **ABS** and the **RBA** (the agencies) seek to explain the purpose of the **EFS collection** and in turn the need for the submitted data to be of high data quality.

This guide should be read in conjunction with:

- the **EFS collection**, including *Reporting Standard ARS 701.0 ABS/RBA Definitions for the EFS Collection* (ARS 701.0), which contains definitions of the **reporting concepts** underlying the data to be reported;
- *Reporting Practice Guide RPG 701.0 ABS/RBA Reporting Concepts for the EFS Collection* (RPG 701.0), for further detail on the **reporting concepts** defined in ARS 701.0 and other relevant reporting standards; and
- *Prudential Practice Guide CPG 235 Managing Data Risk* (CPG 235).

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.

Purpose of the EFS collection

Data collected under the **EFS collection** is primarily used by the **ABS** and **RBA** for analysis, publication and policy making purposes.

The **ABS** will utilise 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** will utilise 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.

There is a strong need for the data submitted under the **EFS collection** to be relevant, timely and accurate, in order for the information to be suitable for the multiple purposes outlined above.

CPG 235 Managing Data Risk

In each **EFS reporting standard** the 'Quality control' provisions require EFS reporting entities to have in place "systems, processes and controls" that "are to assure the completeness and reliability of the information provided".

EFS reporting entities that are **ADIs** will have regard to and consider the guidance on managing data risk set out in CPG 235, in order to satisfy the above requirement.

EFS reporting entities that are **RFCs**, whilst not **APRA**-regulated institutions, should similarly consider the sound practices set out in CPG 235, in order to satisfy the above requirement.

Managing data quality – Benchmarks

The agencies expect EFS reporting entities to place high importance on the design and implementation of controls to manage and assess the quality of data being submitted in the **EFS collection**.

In order to assist EFS reporting entities in considering the controls that may be required, the agencies have provided a set of 'data quality benchmarks'. These benchmarks provide a guide on how accurate the agencies expect the data to be in order for the data to be suitable for the purposes outlined above. The data quality benchmarks are provided in attachment A to this guide. Attachment A also clarifies **errors**, i.e. when a **data item** is considered to be outside the agencies' expectations of data quality, in the context of the benchmarks and how this aligns to treatments suggested in RPG 701.0 in the use of proxies.

The agencies are of the view that it would be good practice for all EFS reporting entities to consider these data quality benchmarks when designing and implementing controls that will manage risks relating to data quality under the **EFS collection**. Further guidance on how

these benchmarks might be considered is included in the *Application of CPG 235 to the EFS collection and use of the data quality benchmarks* section of this guide.

Benchmarks based on priority

A key factor applied in setting the benchmarks is whether potential **errors** have a critical impact on the use of data by the agencies. The agencies have thus established two categories of priority for this, "standard" and "very high" data items. The agencies expect EFS reporting entities to utilise the priority categorisation, with the benchmarks, as an indicator of where to focus data quality management practices.

Attachment B to this guide lists the **very high priority** data items in each EFS reporting standard.

Benchmarks based on size of entity

The agencies have set benchmarks based on entity size, to proportionately account for the impact of **errors** on data quality in the **EFS collection**. Benchmarks for **large institutions** are aimed at identifying **errors** that could affect industry aggregates, for example growth rates in key series. Also, such benchmarks 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 **errors** relevant to the internal consistency of the entity's series.

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

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

Good practice would be for an EFS reporting entity to have particular regard to the following CPG 235 concepts when considering how to manage data quality in the EFS collection, and in using the benchmarks in data risk management practices:

- Structured and Principles-based approach - data risk management is to be part of a systematic and formalised approach (para. 20). 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 (para. 22e). The agencies expect that the data quality benchmarks will assist an EFS reporting entity in understanding quality levels that need to be achieved before providing data under the **EFS collection**.
- 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 (para. 33), and be aligned to the entity's risk appetite (para. 14-15). The agencies expect that an EFS reporting entity would consider the data quality benchmarks when setting risk appetite around data quality and in the design, implementation and assessment of controls to manage EFS data quality.

- 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 (para. 51-52). The agencies expect an EFS reporting entity to use the data quality benchmarks as part of data validation design, throughout the data's life-cycle. 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).
- Monitor and manage 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 to consider an adjustment of controls. The agencies also expect the data quality benchmarks to be considered in the development of quality metrics relating to the *EFS collection*, in order to report on the effectiveness of controls in place.
- 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 (para. 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 appropriately utilising the data quality benchmarks. The prioritisation of data items may also be useful for an EFS reporting entity to consider in setting multi-year assurance programs (paras. 67-68). The agencies view 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

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 EFS reporting entities. As part of this approach, the agencies, from time to time, may engage with EFS reporting entities in a variety of ways including, but not limited to:

- *Discussion of reports on assurance processes* - 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.
- *Discussion of proxies and assumptions used* – the agencies may seek to engage with EFS reporting entities to better understand the data being provided and the use and nature of proxies or assumptions used.
- *Peer workshops* - The findings of assurance processes and other initiatives (appropriately de-identified) will form the basis for peer workshops, which will serve as an opportunity for EFS reporting entities 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.

Glossary

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

ABS	Australian Bureau of Statistics established under the *Australian Bureau of Statistics Act 1975*.
APRA	Australian Prudential Regulation Authority established under the *Australian Prudential Regulation Authority Act 1998*.
data item	Refers to the information required to be entered in a specific cell of a form; the intersection of one or more reporting concepts along a given dimension.
EFS collection	Comprises the ***EFS reporting standards*** and data collected under the ***EFS reporting standards***.
error	Refers to a difference between the data reported to ***APRA*** and the data required to be reported to ***APRA*** as set out in the relevant reporting standards and guidance documents. An ***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.
flow	Refers to a ***data item*** with a reporting basis of 'during' the reporting period, as specified in the instructions for the relevant reporting standard.
large institution	Refers to an ***ADI*** or ***RFC*** with greater than or equal to $200 billion in total assets measured on a ***domestic books*** basis.
reporting concept	Refers to a theoretical construct – typically economic or statistical in nature – that is identified in the instructions and/or definitions. Frequently these constructs will be used across more than one ***data item***.
standard priority	Refers to a ***data item*** in an ***EFS reporting standard*** that is not ***a very high priority data item***.
stock	Refers to 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	Refers to a ***data item*** listed in Attachment B of this reporting practice guide.

Attachment A – Data quality benchmarks

The data quality benchmarks are provided in table 1 and table 2 below.

Table 1 – Benchmarks for data expressed as a dollar value, count or proportion

Table 1 below provides data quality benchmarks expressed as absolute dollar values and as a percentage of the institution-level series, for **stock** and **flow data items** that are reported as a dollar value, a count or a proportion of a year (e.g. term, tenor).

A **data item** expressed as a dollar value, count or proportion is considered to be outside agency expectations of data quality if:

- An **error** in that **data item** exceeds the benchmarks in Table 1 specified in (1a) and (1b); or
- An **erro**r in that **data item** exceeds the benchmarks in Table 1 specified in (2).

For data items reported as a dollar value, refer to the percentage and absolute dollar benchmarks. For data items reported as a count or as a proportion, refer only to the percentage benchmarks.

Table 1

Data item type	Priority	Benchmarks for errors that are within agency expectations: as an absolute dollar value (1a)	Benchmarks for errors that are outside agency expectations for: a large institution: as percentage of institution series (1b)	a large institution: as an absolute dollar value (2)	an ADI or RFC that is not a large institution: as percentage of institution series (1b)	an ADI or RFC that is not a large institution: as an absolute dollar value (2)
Stock	Very high	$25m	0.50%	$2,000m	2.00%	$500m
Stock	Standard	$25m	5.00%	$2,000m	10.00%	$500m
Flow	Very high	$10m	5.00%	$250m	10.00%	$100m
Flow	Standard	$10m	10.00%	$250m	20.00%	$100m

Example 1

An **ADI** or **RFC** that is not a **large institution** is using the benchmarks in Table 1 to assess the adequacy of their processes and controls on a **standard priority stock data item** reported as a

dollar value, and is considering whether it is likely to produce an **error** that is within or outside agency expectations.

The potential for **error** on that **data item** is identified as likely to be:

- Case A: Around $50 million and representing around 8 per cent of the value of that **data item**. This would fall *within* agency expectations for data quality as the **error** is:

  1. above the minimum absolute dollar value benchmark of $25 million, but
  2. below the percentage (10%) and maximum absolute dollar value ($500 million) benchmarks.

- Case B: Around $50 million and representing around 12 per cent of the value of that **data item**. This would fall *outside* agency expectations for data quality as the **error** is:

  1. above the minimum absolute dollar value benchmark of $25 million, and
  2. above the percentage (10%) benchmark.

- Case C: Around $20 million and representing around 12 per cent of the value of that **data item**. This would fall *within* agency expectations for data quality as the **error** is:

  1. below the minimum absolute dollar value benchmark of $25 million.

- Case D: Around $550 million and representing not more than 8 per cent of the value of that **data item**. This would fall *outside* agency expectations for data quality as the **error** is:

  1. above the maximum absolute dollar value ($500 million) benchmark.

Example 2

An **ADI** or **RFC** that is not a **large institution** is using the benchmarks in Table 1 to assess the adequacy of their processes and controls on a **standard priority flow data item** reported as a count, and is considering whether it is likely to produce an **error** that is within or outside agency expectations.

The potential for error on the **data item** is identified as likely to be:

- Case A: Around 25 per cent of the figure for that **data item**. This would fall *outside* agency expectations for data quality as the **error** is above the percentage (20%) benchmark.
- Case B: Around 15 per cent of the figure for that **data item**. This would fall *within* agency expectations for data quality as the **error** is below the percentage (20%) benchmark.

Table 2 – Benchmarks for data expressed as a rate

Table 2 below provides 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*).

A *data item* expressed in basis points is considered to be outside agency expectations of data quality if an *error* in that rate data exceeds the benchmarks in Table 2.

Table 2

Priority	Benchmarks for errors that are outside agency expectations for:	
	a large institution	an ADI or RFC that is not a large institution
Very high	5 bps	10 bps
Standard	15 bps	20 bps

Example 3

An *ADI* or *RFC* that is not a *large institution* is using the benchmarks in Table 2 to assess the adequacy of their processes and controls on a *standard priority data item* reported as an *interest rate*, and is considering whether it is likely to produce an *error* that is within or outside agency expectations.

This *error* is identified as being:

- 25 basis points. This would fall *outside* agency expectations for data quality as the *error* is above the 20 basis point benchmark.
- 15 basis points. This would fall *within* agency expectations for data quality as the *error* is below the 20 basis point benchmark.

What constitutes an error

Change of calculation methodology

For *cost/value of funds*, *margin* and *benchmark rate* data, the agencies do not expect that changes to an EFS reporting entity's internal calculation methodology would be classified as an *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 the agencies, either as part of a regular discussion or on an *ad hoc* basis. As part of this discussion, the agencies would expect the reporting institution 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.

Use of proxy methodologies

RPG 701.0 guides EFS 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.

Where RPG 701.0 allows use of a proxy methodology, the data quality benchmarks apply to errors determined by reference to the appropriately calculated proxy measure. That is, the benchmarks are applied to calculation *errors* in the use of the proxy against the intended methodology, not between the proxy and the underlying data.

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

Attachment B – Series with 'Very High' priority

The table below will be updated to include attribute codes once the reporting forms are available in D2A.

Form	Item(s)	Attribute code(s)
ARF 720.A	Item 3, column 1	
	Item 4, column 1	
	Item 6, all columns	
	Items 11 and 11.1, column 1	
	Items 12 and 12.1, column 1	
	Items 13, 13.1 and 13.1.1, column 1	
	Item 14, all columns	
	Item 16, all columns	
	Items 19 and 19.1, column 1	
	Items 20, 20.1 and 20.2, column 1	
	Item 21, column 1	
ARF 720.0B	Item 3, all columns	
	Item 4, all columns	
	Item 6, all columns	
	Item 11, all columns	
	Item 12, all columns	
	Items 13, 13.1 and 13.1.1, all columns	
	Item 14, all columns	
	Item 16, all columns	
	Item 19, all columns	
	Items 20 and 20.1, all columns	
	Item 21, all columns	

ARF 720.1A	Items 1.1 and 1.1.1, columns 1, 4 and 5
	Items 1.1.1.1, 1.1.1.1.1, 1.1.1.1.1.1, 1.1.1.1.1.2, 1.1.1.1.2, 1.1.1.3, 1.1.1.3.2, 1.1.1.3.3, 1.1.1.5 and 1.1.2, column 1
	Item 2.1, all columns
	Item 3.1, columns 1 and 3
	Item 4.1, all columns
	Item 5.1, all columns
ARF 720.1B	Items 1.1 and 1.1.1, all columns
	Items 1.1.1.1, 1.1.1.1.1, 1.1.1.1.1.1, 1.1.1.1.1.2, 1.1.1.1.2, 1.1.1.3, 1.1.1.3.2, 1.1.1.3.3, 1.1.1.5 and 1.1.2, column 1
	Item 2.1, all columns
	Item 3.1, all columns
	Item 4.1, all columns
ARF 720.2A	Items 1 and 1.1, columns 1, 3, 5 and 7
	Items 1.1.1, 1.1.3, 1.1.3.2, 1.1.3.3, 1.1.5 and 1.2, column 7
ARF 720.2B	Items 1 and 1.1, all columns
	Items 1.1.1, 1.1.3, 1.1.3.2, 1.1.3.3, 1.1.5 and 1.2, column 4
ARF 720.3	Items 1.1 and 1.1.5, column 1
	Item 2.1, column 1
	Items 3.1 and 3.1.1, all columns
	Item 4.1, column 1
ARF 720.4	Items 1, 1.1 and 1.2, column 1
	Items 2, 2.1 and 2.4, column 1
ARF 720.6	Item 1, columns 1 and 2
	Item 2, columns 1 and 3
	Item 3, columns 1 and 3
	Item 4, column 1

ARF 730.0	Item 1.1, column 1
	Item 3.1, column 1
	Item 4.1, column 1
	Item 10.1, column 1
ARF 741.0	Item 2.1, all columns
	Item 3.1, all columns
	Item 4.1, all columns
ARF 742.0A	Items 1.1, 1.1.1, 1.1.2 and 1.1.3, columns 2, 3 and 4
	Items 2.1, 2.1.2, 2.1.3 and 2.1.4, columns 2 and 3
	Items 7.1 and 7.13, column 1
	Item 8.1, all columns
ARF 742.0B	Items 1.1, 1.1.1, 1.1.2 and 1.1.3, columns 2, 3 and 4
	Items 2.1, 2.1.2, 2.1.3 and 2.1.4, columns 2 and 3
	Items 7.1 and 7.13, column 1
	Item 8.1, all columns
ARF 743.0	Items 1.1, 1.1.1 and 1.1.2, column 5
	Items 2.1, 2.1.1 and 2.1.2, column 5
	Items 4.1, 4.3 and 4.3.1, all columns
	Items 5.1 and 5.14, all columns
	Items 6.1 and 6.3, all columns
ARF 744.0A	Items 1.1 and 1.2, columns 2, 3, 4 and 5
	Item 2.1, columns 2, 3, 7 and 8
ARF 744.0B	Items 1.1 and 1.2, columns 2, 3, 4 and 5
	Item 2.1, columns 2, 3, 7 and 8
ARF 746.0A	Item 1.1, columns 2 and 3
	Item 2.1, columns 2 and 3

ARF 746.0B	Item 1.1, columns 2 and 3
	Item 2.1, columns 2 and 3
ARF 747.0A	Items 1.1, 1.1.1, 1.1.2 and 1.1.3, columns 4 and 5
	Items 2.1 and 2.1.1, columns 2 and 3
	Items 3.1, 3.1.1, 3.2 and 3.2.1, columns 4 and 5
	Item 4.1, columns 4 and 5
	Item 5.1, columns 4 and 5
ARF 747.0B	Items 1.1, 1.1.1, 1.1.2 and 1.1.3, columns 4 and 5
	Items 2.1 and 2.1.1, columns 2 and 3
	Items 3.1, 3.1.1, 3.2 and 3.2.1, columns 4 and 5
	Item 4.1, columns 4 and 5
	Item 5.1, columns 4 and 5
ARF 748.0A	Item 1.3, columns 4 and 5
	Item 2.3, columns 4 and 5
ARF 748.0B	Item 1.3, columns 4 and 5
	Item 2.3, columns 4 and 5