



EUROPEAN CENTRAL BANK

BANKING SUPERVISION

Data Quality at the ECB: trends 2014 -2017

Luis Suarez Tumi

Principal Economist Statistician

Banking Supervision Data Division

Supervisory Statistics Seminar

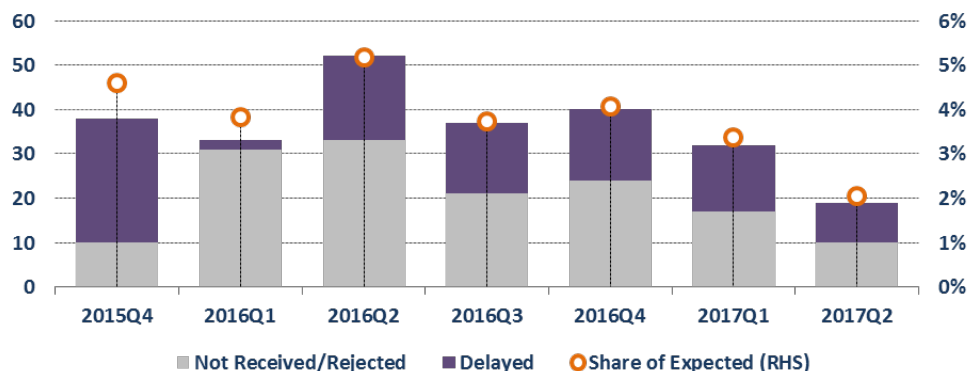
Frankfurt - 15 November 2017

European Central Bank, Frankfurt

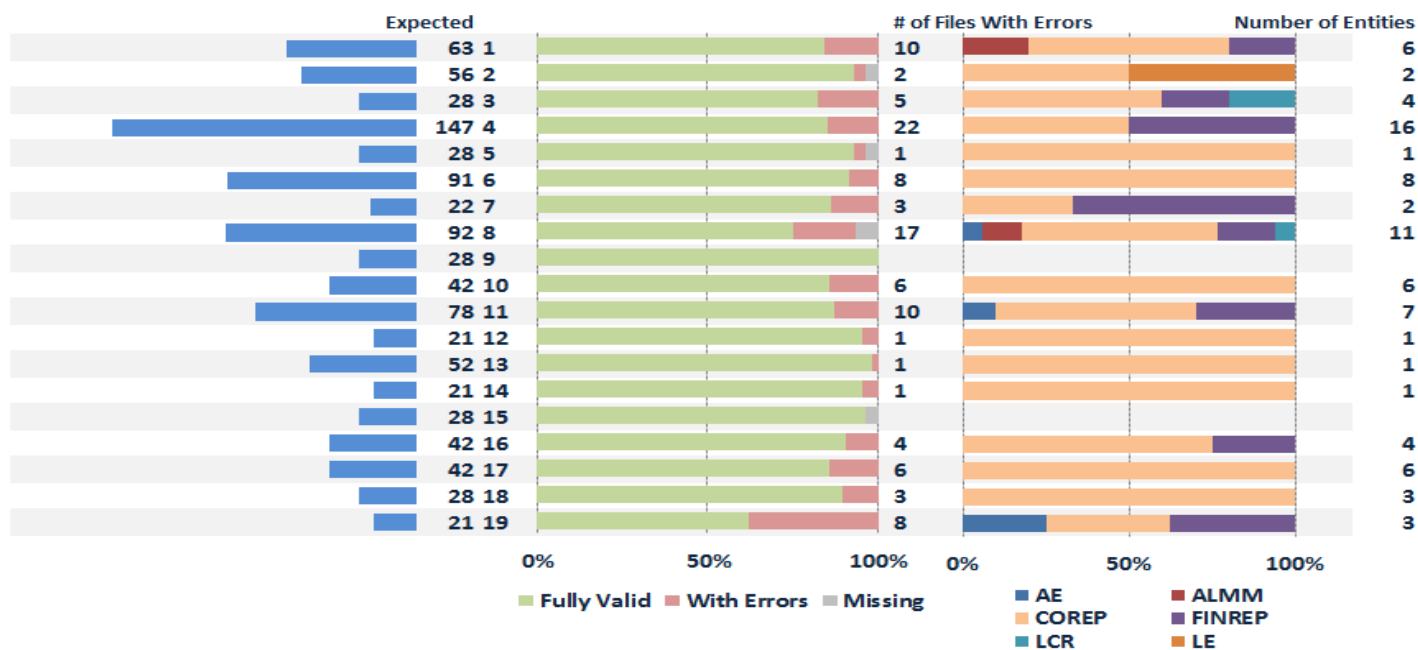
Hard Checks: Punctuality



Not received/rejected and delayed modules in 2014 - 17



Status of the submissions in Q2 2017

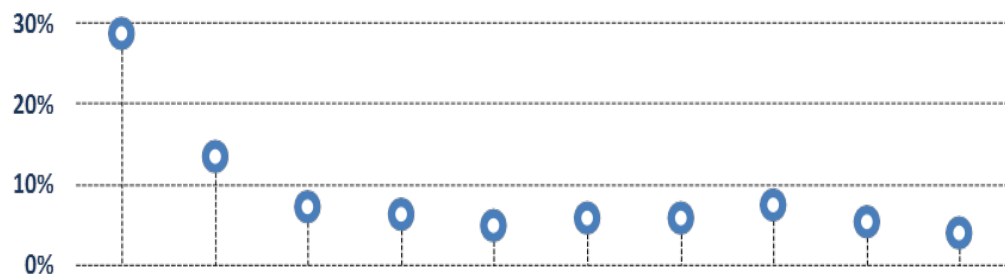


Hard Checks: Accuracy and consistency (1/4)

Examples metrics

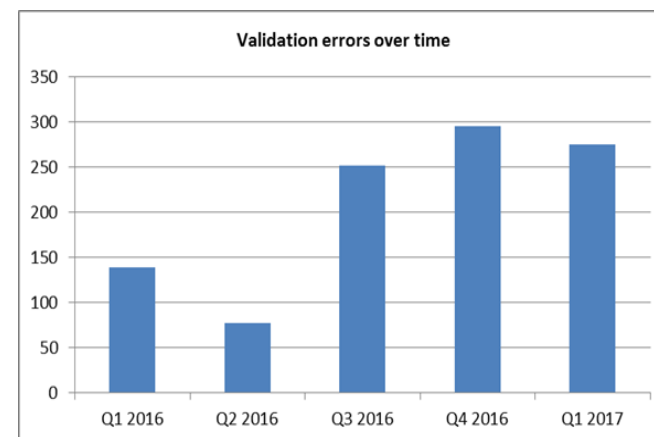
- Number of failing validation rules
- Number of validation errors

Percentage of VRs failed over applicable VRs



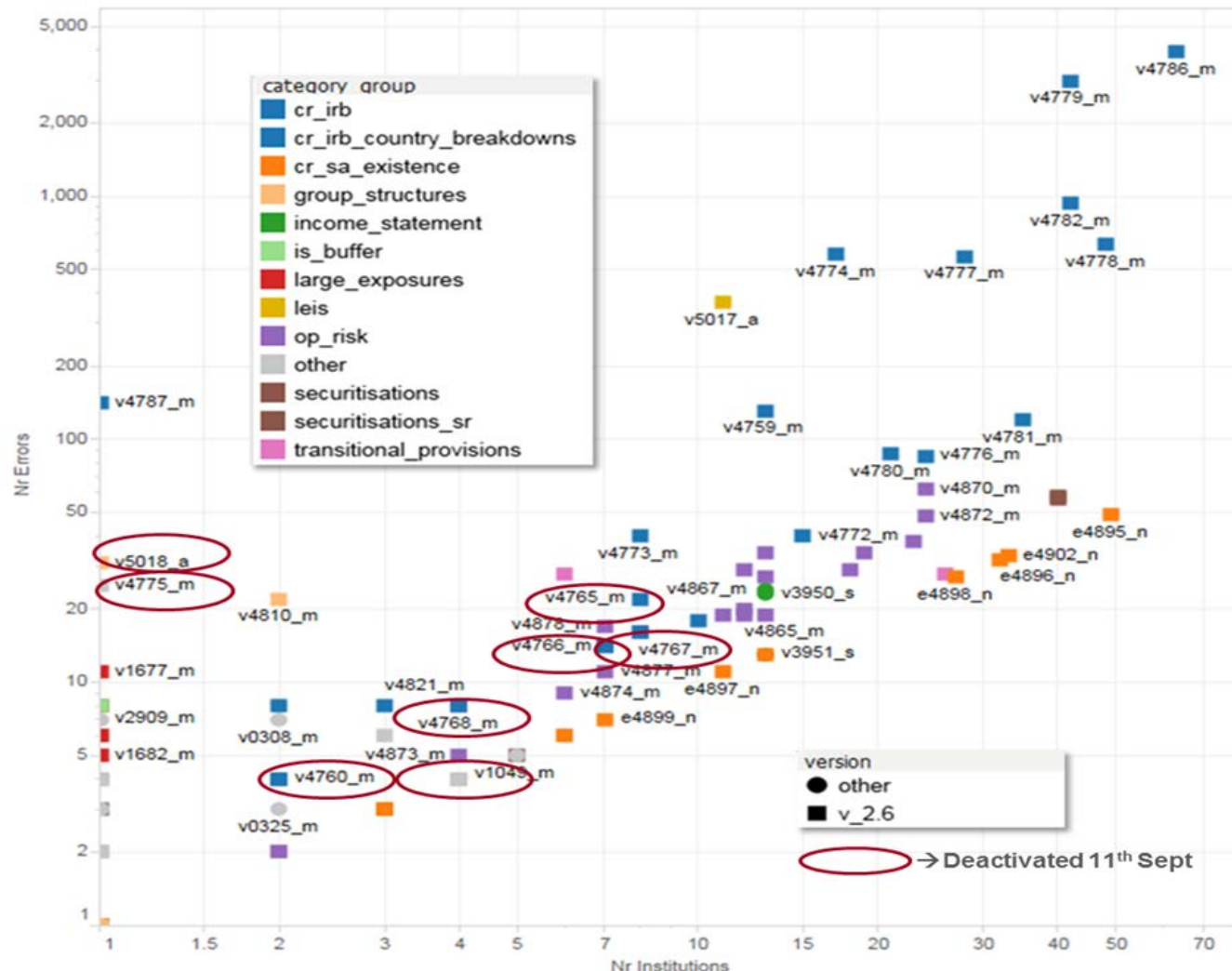
	Q4 2014	Q1 2015	Q2 2015	Q3 2015	Q4 2015	Q1 2016	Q2 2016	Q3 2016	Q4 2016	Q1 2017
Failing VRs	485	219	127	118	110	107	124	157	136	88
Applicable VRs	1,686	1,652	1,801	1,897	2,262	1,892	2,134	2,108	2,640	2,245

Number of validation errors



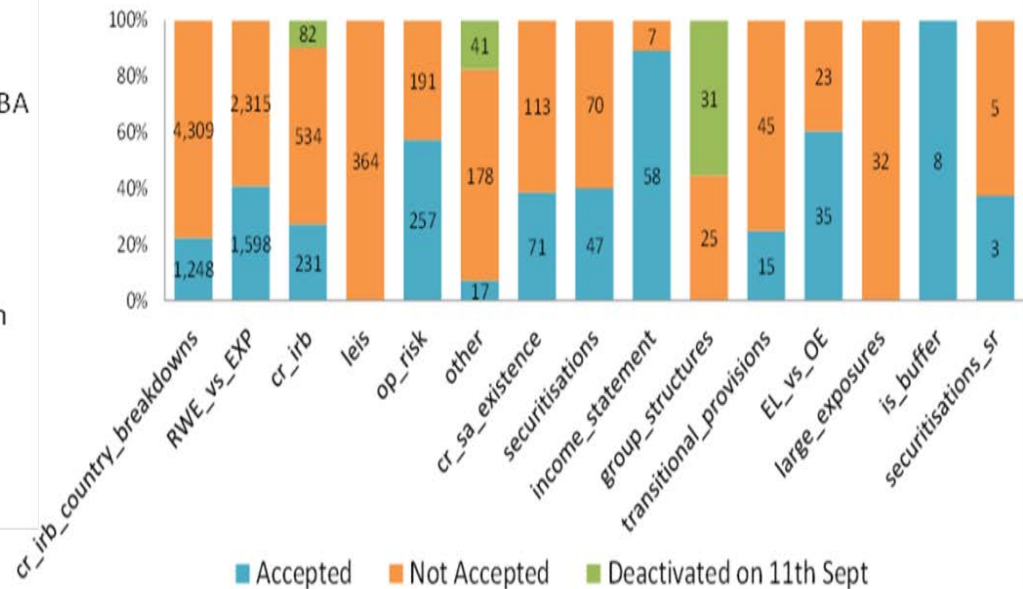
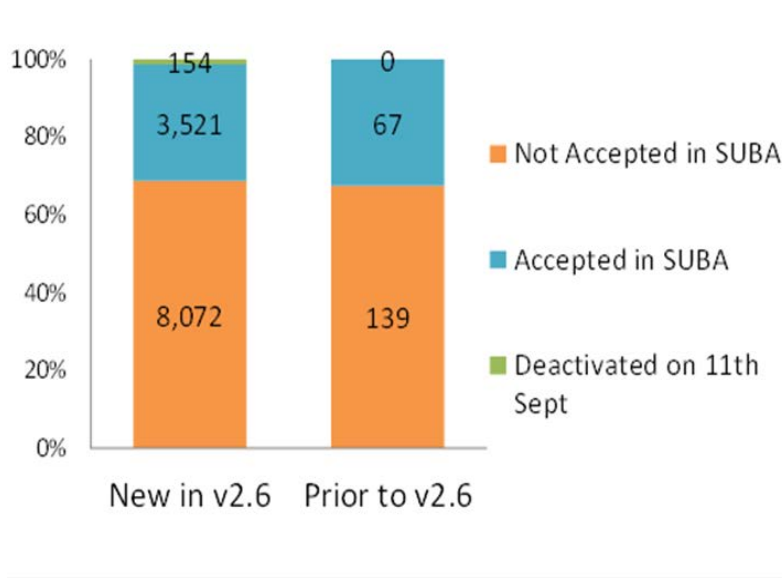
Hard Checks: Accuracy and consistency (2/4)

VRs by number of institutions affected and number of errors triggered



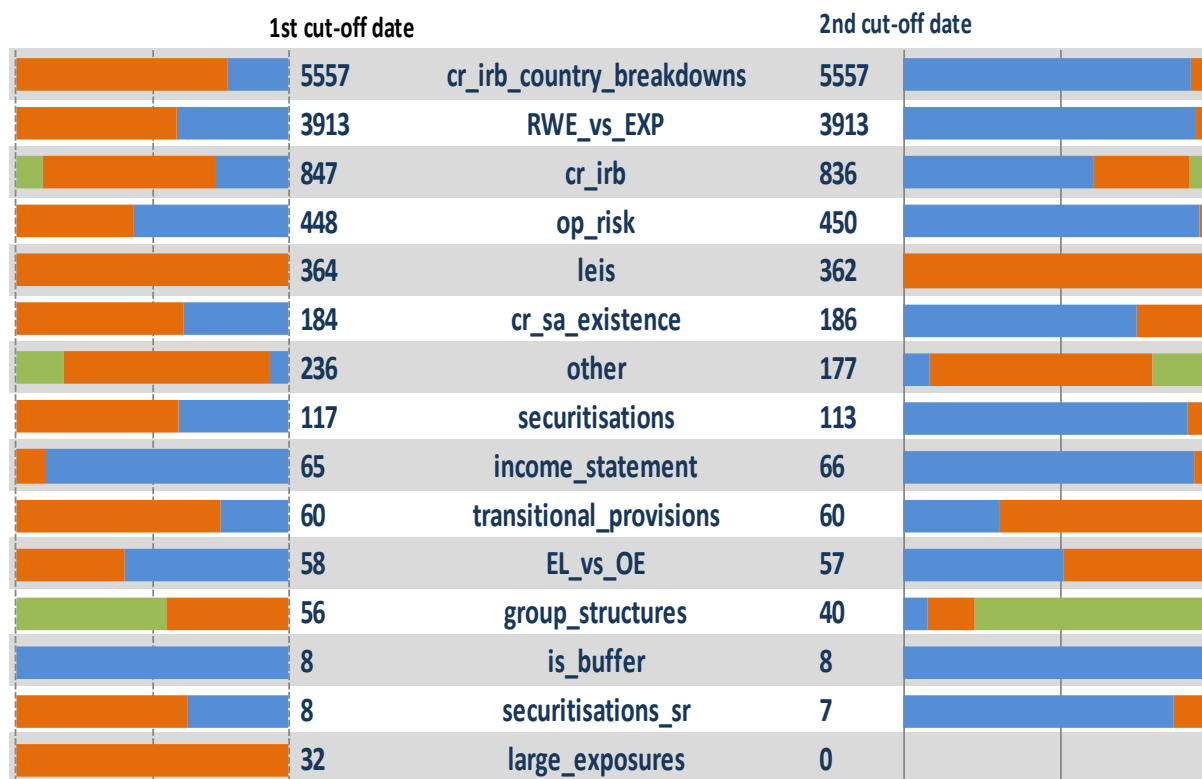
Hard Checks: Accuracy and consistency (3/4)

Q2 2017 Number of errors at ECB Remittance date Errors by VR category



Hard Checks: Accuracy and consistency (4/4)

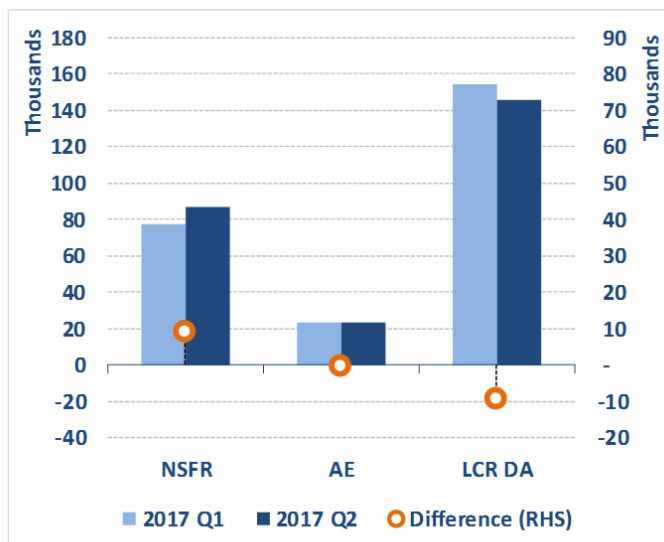
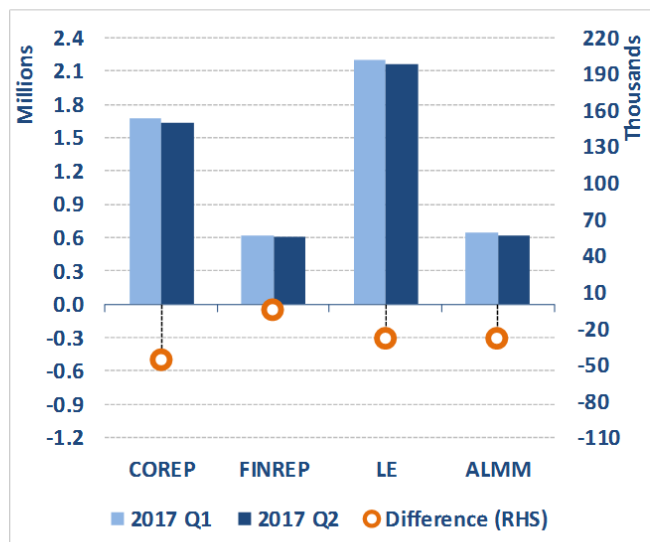
Q2 2017: Evolution validation errors



- Accepted
- Not Accepted
- Deact on 11th Sept

Soft Checks: Stability

Number of data points reported in Q1 2017 and Q2 2017



Business reasons

- Reaching the thresholds for reporting geographical breakdowns
- New financial instruments in the balance sheet*
- Disposal of a subsidiary (affects the number of countries reported)*

Soft Checks: Completeness

Completeness rates by report

Module	Quarter						
	Q4 2015	Q1 2016	Q2 2016	Q3 2016	Q4 2016	Q1 2017	Q2 2017
COREP	88%	90%	93%	94%	89%	94%	95%
FINREP	85%	79%	85%	86%	88%	87%	94%
AE	81%	78%	81%	82%	83%	78%	81%
LCR	78%	79%	NA	91%	94%	94%	92%
NSFR	86%	87%	89%	89%	90%	89%	87%
ALMM					68%	67%	67%
Total Average	83%	83%	85%	88%	89%	85%	89%

Business Model	Quarter	
	Q4 2016	Q1 2017
Corporate/Wholesale lender	85%	88%
Custodian and AM	78%	78%
Diversified lender	85%	89%
G-SIB	97%	98%
G-SIB universal	97%	98%
Not classified	84%	85%
Retail lender	87%	87%
Sectoral lender	80%	80%
Small domestic lender	81%	85%
Universal bank	95%	94%

The data used in the table above comes from a set of pre-defined data points that are considered essential by supervisors to complete key supervisory tasks and should be reported by all institutions independent of their size, business model or country of origin. However, because of differences due to business models making some data points redundant for that entity, achieving 100% is not possible and completion rates over 80% are considered as satisfactory.

Plausibility

- Outlier analysis:

Outlying unit of observations are flagged and explanations are requested to the institutions via the NCAs

We look at values with:

- extremely high (or extremely negative) **growth rates**.
- extremely high (or extremely negative) **levels**.

Examples Answers on Plausibility

Good: about a variation of XX% in Deposits:

“The decrease is mainly due to Counterparty A -7.06bln , Counterparty B -3.5bln, and Counterparty C -580mln”.

Medium: about variation of YY% in Total Risk Exposure amount:

“Integration of XXXX Lease Services (+13.9bn) o/w Credit risk (+12.2bn), operational risk (1.6bn), Market and FX risk (+0.2bn)”

Bad: about a variation of ZZ% on risk weighted exposures amounts for credit , counterparty credit and dilution risks and free deliveries

“Sale of Hungarian branch”

Resubmission expected: about a variation in financial assets held for trading-debt securities, XX% from other financial corporations to non-financial corporations:

Due to data quality analysis , an amount of XXX millions was reclassified from other financial corporations to non financial corporations

Resubmissions (an example)

Total number of resubmissions in Q2 2017

By Module			
Module	# of expected modules	# of resubmitted modules	# of institutions resubmitting
AE	134	15	15
ALM	133	35	29
COREP	134	82	58
FINREP	135	50	45
LCR	133	13	11
LE	134	48	43
NSFR	133	17	16
Total	936	260	105

By Data Points	
# of resubmitted data points	# of institutions resubmitting
962	11
22,188	3
19,914	42
9,086	39
184	3
13,758	27
908	8
67,000	83

DQIs introduced for SREP 2017

- ✓ JSTs' SREP Element 2 assessment in sub-category "Risk Infrastructure, Data & Reporting"



IMAS Screenshots

Risk Assessment Indicators

Display indicators for the past **quarters**

Data Point	Name	2017 Q1	2016 Q4	2016 Q3
DQ11	Overall data quality indicator (Unit)	2.00	2.00	2.00
DQ12	Punctuality – data quality indicator (Unit)	1.00	1.00	1.00
DQ13	Completeness – data quality indicator (Unit)	2.00	2.00	2.00
DQ14	Accuracy – data quality indicator (Unit)	2.00	2.00	2.00

- ✓ MSD's SREP horizontal analyses

- ✓ Supervisory Dialogues with banks

4.3.2 SREP 2017 - Element 2 - IG&RM - Data quality

- 25 JSTs with data quality issues in several areas revealed by the improved set of tools for JSTs.
- JSTs are now equipped with a set of tools to monitor, assess and improve the data quality.
- These tools allow us to identify data quality issues from different perspectives.
- 25 institutions face data quality issues in more than one area:
 - Data quality score Q4 2016
 - ISSUE: stress test 2017
 - Stress test 2016
 - ISMP / ISMP 2016 / 2017
 - BCBS 239 Thematic Review
 - CRM/Liquidity Monitoring Exercise
 - Other (see TR99)

2.5.1 Example of Zoom: Risk and data infrastructure

Risk Infrastructure, Data and Reporting

Data quality strengths and weaknesses

Findings of the BCBS 239 TR which are used for all banks are displayed at the bottom right JST.

- To recap bank-specific issues encountered with respect to data quality: e.g. BCBS 239 but also CRM liquidity monitoring exercise, On-site inspections, TRM, RGA, Stress test, etc.
- To recap the overall data quality indicator (DQI) for (example bank) disclosed in MAS, under Element 2 → Combined Element 2 → Display Key Risk Indicators

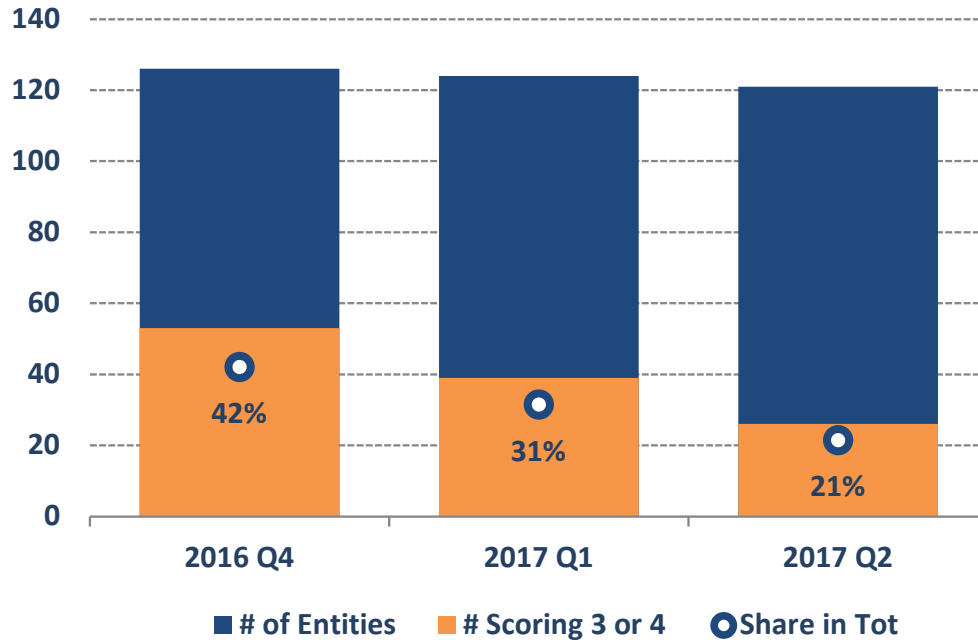
TRM/16	Data quality indicator* (Q4 2016)	SSM average
X		2.5

BCBS 239 Internal Audit Findings

- Large macroprudential risk in size and in numbers
- Several weaknesses in the data compilation processes
- IT systems do not support reconciliation of key data independently other reports
- No full automation of data compilation process
- Implementation of a number of the process and maturity of the system
- Limited data down capabilities on single data level

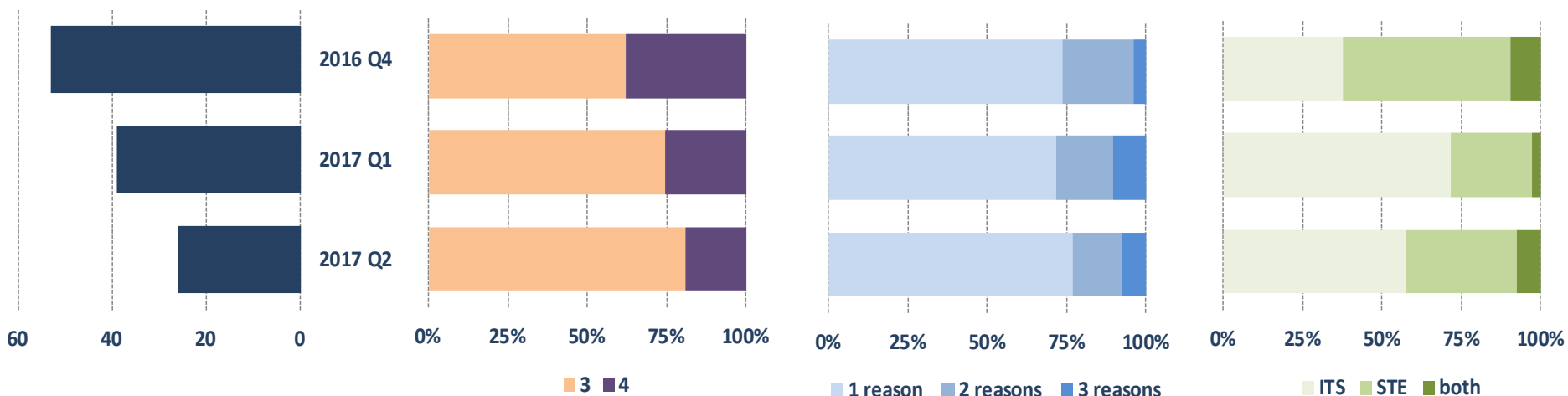
* Quality of data (data completeness) reflects aspects such as: detectability, completeness and accuracy. JST 2016 Item 1 (good) to 4 (serious concerns). www.bankingsupervision.europa.eu

Data Quality Indicators: Overall numbers

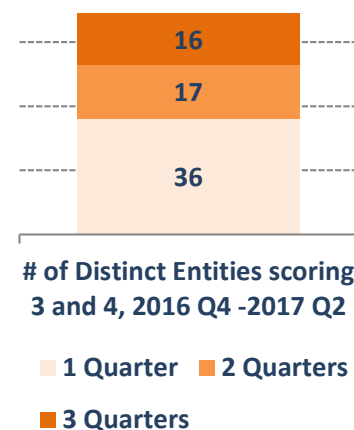


- Number of entities under direct supervision change
- Decreasing number of entities scoring 3 or 4 – absolute and relative

Data Quality Indicators: Overall Scoring 3 & 4

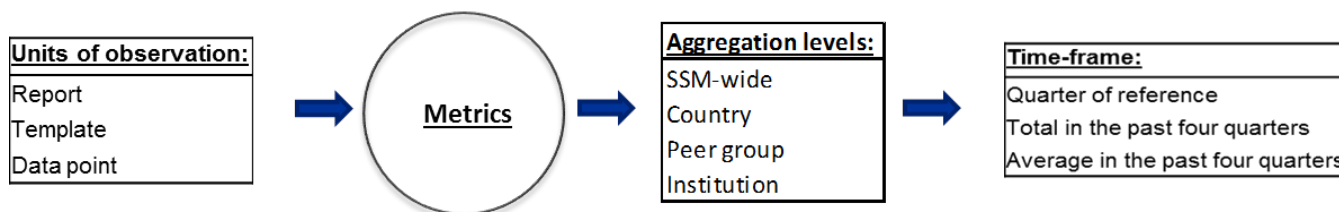


- Decreasing number of entities scoring 3 or 4 – absolute and relative
- Decreasing number of entities scoring 4
- 1 out of 4, that scored 3 or 4, scored it for more than 1 reason.
- Issues rarely are related to both reporting frameworks
- 16 out of 26 entities in Q2 2017, scored 3 or 4 in both previous quarters



Conclusion: Monitoring data quality

During the last year of activity the DG-S/SUP Data Quality Team reached a well established level of data quality issues' identification and evaluation. This was ultimately achieved by the introduction of the Data Quality Framework and the Data Quality Findings Database (shared with SSM in a quarterly basis).



How can we exploit their full potential?

CLEAR ACTION POINTS

IF NOT ADDRESSED: ESCALATION PROCESS

REFERENC_	REMITTAN_	RECEPTI_	MODULE_ID	CONS_LEV_	TEMPLATE_	Dimension	FORMULA	Rule_Type	XML_ELM_	RULE_CODE	DPT_STAT_
31DEC2016.00.		27FEB2017.12.	COREP_Con			Plausibility					Vali
31DEC2016.00.		27FEB2017.12.	COREP_Con			Plausibility					Vali
31DEC2016.00.		27FEB2017.11.	COREP_NSFR_			Plausibility					Vali
30JUN2016.00.	25AUG2016.00.		COREP_Con	CONS	C17.00.a	AccuracyAndConsistency	IF {$f(030) = f(04)$	EGDQ		EGDQ_22	
30JUN2016.00.	25AUG2016.00.		COREP_Con	CONS	C17.00.a	AccuracyAndConsistency	{$f(030) = f(040)$	EGDQ		EGDQ_23	
30JUN2016.00.	25AUG2016.00.		COREP_Con	CONS	C22.00	AccuracyAndConsistency	IF {marksIcon.	EGDQ		EGDQ_29	
30JUN2016.00.	28JUL2016.00.		COREP_LCR_	CONS	C53.00.aC53.	AccuracyAndConsistency	{$0.75 = IT.CS.$	EGDQ		EGDQ_34	
30JUN2016.00.	25AUG2016.00.		COREP_Con	CONS	C06.01	AccuracyAndConsistency	{<math>c06.01 <= 1.00.</math>	EGDQ		EGDQ_44	
30JUN2016.00.			COREP_Con	CONS	C_17.00.a	AccuracyAndConsistency	sum({$f(020) = f.$	OpRisk		RESULT_C06b	Rev
30JUN2016.00.			COREP_Con	CONS	C_02.00.C_16.	AccuracyAndConsistency	{$C02.00.$	OpRisk		RESULT_C30	EBU
30SEP2016.00.	25NOV2016.00.	23NOV2016.12.	COREP_Con	CONS	C_22.00	AccuracyAndConsistency	IF {marksIcon.	EGDQ		EGDQ_29	
30SEP2016.00.	25NOV2016.00.	23NOV2016.12.	COREP_Con	CONS	C_05.01	AccuracyAndConsistency	{<math>f(050) <= 1.00.</math>	EGDQ		EGDQ_44	
30SEP2016.00.	25NOV2016.00.	21NOV2016.16.	COREP_LE_C_	CONS	C_27.00C_28.	AccuracyAndConsistency	{<math>f(040) <= 1.00.</math>	EGDQ		EGDQ_56	
30SEP2016.00.	25NOV2016.00.	21NOV2016.16.	COREP_LE_C_	CONS	C_27.00C_28.	AccuracyAndConsistency	{<math>c050) <= 1.00.</math>	EGDQ		EGDQ_57	
30SEP2016.00.	25NOV2016.00.	21NOV2016.16.	COREP_LE_C_	CONS	C_27.00C_28.	AccuracyAndConsistency	{<math>c060) <= 1.00.</math>	EGDQ		EGDQ_58	
30SEP2016.00.	25NOV2016.00.	21NOV2016.16.	COREP_LE_C_	CONS	C_27.00	AccuracyAndConsistency	{<math>c070) <= 1.00.</math>	EGDQ		EGDQ_59	
30SEP2016.00.	25NOV2016.00.	21NOV2016.16.	COREP_LE_C_	CONS	C_27.00	AccuracyAndConsistency	{<math>c080) <= 1.00.</math>	EGDQ		EGDQ_61b	
30SEP2016.00.	25NOV2016.00.	23NOV2016.12.	COREP_Con	CONS	C_08.01.a	AccuracyAndConsistency	{$f(010).c010) = f.$	EGDQ		EGDQ_77	
31DEC2016.00.	27FEB2017.00.	27FEB2017.11.	COREP_Con	CONS	C_17.00.a	AccuracyAndConsistency	IF {$f(030) = f(04)$	EGDQ		EGDQ_22	
31DEC2016.00.	27FEB2017.00.	27FEB2017.11.	COREP_Con	CONS	C_05.01	AccuracyAndConsistency	{<math>f(050) <= 1.00.</math>	EGDQ		EGDQ_44	
31DEC2016.00.	27FEB2017.00.	23FEB2017.11.	COREP_LE_C_	CONS	C_27.00C_	AccuracyAndConsistency	{<math>c050) <= 1.00.</math>	EGDQ		EGDQ_57	
31DEC2016.00.	27FEB2017.00.	23FEB2017.11.	COREP_LE_C_	CONS	C_27.00C_	AccuracyAndConsistency	{<math>c060) <= 1.00.</math>	EGDQ		EGDQ_58	
31DEC2016.00.	27FEB2017.00.	23FEB2017.11.	COREP_LE_C_	CONS	C_28.00C_	AccuracyAndConsistency	{<math>C_28.00C_ <= 1.00.</math>	EGDQ		EGDQ_59	
31DEC2016.00.	27FEB2017.00.	23FEB2017.11.	COREP_Con	CONS	C_08.01.a	AccuracyAndConsistency	{$f(010).c010) = f.$	EGDQ		EGDQ_77	
31DEC2016.00.	27FEB2017.00.	24FEB2017.12.	FINREP_Con_J_			AccuracyAndConsistency		XBRL	eba_v3079_m	Awaiting correc. v30	
31DEC2016.00.	27FEB2017.00.	24FEB2017.12.	FINREP_Con_J_			AccuracyAndConsistency		XBRL	eba_v3080_m	Awaiting correc. v30	
31DEC2016.00.		24FEB2017.12.	FINREP_Con_J_			Plausibility					Vali

Thank you: Questions or observations