Freie Universität Berlin
Fachbereich Wirtschaftswissenschaft

Data Quality Assessments of the System of National Accounts at the International Level

Inaugural-Dissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft Dr. rer. pol.
des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin

vorgelegt von Dipl.-Kfm. Patrick Müller
aus
Berlin

Oktober 2013

Tag der Disputation: 10. Februar 2014

Dekan: Prof. Dr. Klaus Ruhnke

Erstgutachter: Prof. Irwin Collier, PhD.

Zweitgutachter: Prof. em. Dr. Hans-J. Lenz

Tag der Disputation: 10. Februar 2014

Acknowledgements

Many have contributed to the success of this dissertation project by providing the framework to write and do research. I wish to thank all those, who have participated, supported my efforts, and encouraged me and thus became stakeholders.

Thanks to those who have provided information and answered questions. These are Tomas Duchon and Silke Stapel from Eurostat; Olga Laveda, Rene Piche, Manik Shrestha, Mick Silver, and Kim Zieschang from the IMF; and David Barraclough, David Brackfield, and Catherine Larosa-Elakim from OECD. A special thanks to Herman Smith from the UN for his comments and advice on my work. Also special thanks for comprehensive discussions to Andreas Lorenz from Deutsche Bundesbank (German Central Bank) and to his colleague Andreas Dietrich. In particular also to Norbert Räth and Stefan Hauf from Deutsches Statistisches Bundesamt (German Federal Statistical Office) for comments and advice. Thanks to the kind members of staff and the librarians for providing documents and statistics for research and answering questions.

I wish to thank Boyko Amarov from the Freie Universität Berlin for his introduction to “R” and setting up a world map to present the results.

Moreover, I thank my supervisors Irwin Collier and Hans-J. Lenz, who encouraged me in my work, met to discuss contents, contributions and progress and of my study, willingly offered to read through the draft, requested revisions, and gave comments.

Further, I thank all my friends and my family and also the church family. Special thanks also to my personal mentor Siggi Kind, his encouragement and counsel.

S.D.G.

Table of Contents

List of Tables

List of Definitions

Regarding definitions of SNA terms like capital account, gross domestic product, net lending, etc. refer to “1993 System of National Accounts” (United Nations 1993) and “1993 System of National Accounts – Glossary” (OECD 2000) also available in “National Accounts Questionnaire - Supporting Booklet” (UNSD 2009a), respectively “2008 System of National Accounts” (United Nations 2009f), which also includes a glossary of main terms.

List of Figures

List of Abbreviations

ABS	Australian Bureau of Statistics
ANA	Annual National Accounts
BEA	Bureau of Economic Analysis
BIS	Bank for International Settlements
BIP	Bruttoinlandsprodukt
BOP	Balance of Payments
CARICOM	Caribbean community
CB	Central Bank
cf.	compare
CIS	Commonwealth of Independent States
DatQAM	Data Quality Assessment Methods and Tools
DEF.	Definition
DQ	Data Quality
DQAF	Data Quality Assessment Frameworks
EC	European Commission
ECA	Economic Commission for Africa
ECB	European Central Bank
ECE	Economic Commission for Europe
ECLAC	Economic Commission for Latin America and the Caribbean
ECOFIN	Economic and Financial (Council)
EFC	Economic and Financial Committee
e.g.	for example
EMU	European Monetary Union
ESA	European System of National Accounts
ESCAP	Economic and Social Commission for Asia and the Pacific
ESCWA	Economic and Social Commission for Western Asia
ESS	European Statistical System
et al	and others

etc.	and so forth
EU	European Union
EU27	Country group of the EU27
Eurostat	Statistical Office of the European Communities
FAO	Food and Agriculture Organization
FSB	Financial Stability Board
G20	Country group of the G20
GNDI	Gross National Disposable Income
GDP	Gross Domestic Product
GNI	Gross National Income
i.e.	that is
IMF	International Monetary Fund
IO	International Organization
IQ	Information Quality
ISIC	International Standard Industrial Classification of All Economic Activities
ISIC Rev. 4	International Standard Industrial Classification of All Economic Activities, Revision 4
ISO	International Standards Organisation
ISWGNA	Intersecretariat Working Group on National Accounts
LDC	Least Developed Countries
LLDC	Landlocked Developing Countries
M49	List of all countries, areas, and territories
MADT	Main Aggregates and Detailed Tables
MAR	Mean Absolute Revision
MDG	Millennium Development Goal
MEI	Main Economic Indicators
MR	Mean Revision
MRDS	Minimum Requirement [or Required] Data Set
MS	Milestone
NA	National Accounts
N/A	Not Applicable

NAQ	National Accounts Questionnaire
NIAP	National Income and Product Accounts
NPISH	Non-Profit Institutions Serving Households
NSI	National Statistical Institute
NSO	National Statistical Office
OECD	Organization for Economic Cooperation and Development
OECD.Stat	OECD Statistical Data Warehouse
Org.	Organizations
PDF	Portable Document Format
PEEI	Principal European Economic Indicators
QNA	Quarterly National Accounts
QoQ	Quarter-on-Quarter
RMAR	Relative Mean Absolute Revision
Rosstat	Russian Federal State Statistics Service
SADC	Southern African Development Community
SAR	Special Administrative Region
SIDC	Small Island Developing Countries
SNA	System of National Accounts
UN	United Nations
UNdata	UN data portal
UNECE	United Nations Economic Commission for Europe
UN-NAQ	United Nations National Accounts Questionnaire
UNSD	United Nations Statistics Division
WEO	World Economic Outlook
YB	Yearbook
YoY	Year-on-Year

List of Variables

1.1, 1.2	NAQ tables 1.1 and 1.2 of GDP by expenditure tables in current and constant prices, respectively
1.3	NAQ table 1.3 Relations among product, income, saving, and net lending aggregates
2.1, 2.2	NAQ tables 2.1 and 2.2 of value added by industry tables in current and constant prices, respectively
4.1, …, 4.7	NAQ tables 4.1 to 4.7 of institutional sector accounts
4.8, 4.9	NAQ tables 4.8 and 4.9 of institutional sector accounts for unconsolidated aggregated institutional sector data (sectors: financial and non-financial, and households and NPISH, respectively)
A	Availability referring to a data object
A(1.2,SUR)	Availability of NAQ table "1.2" for Suriname
A(1.3,SUR)	Availability of NAQ table "1.3" for Suriname
A(GDP,c)	Availability referring to the existence of the data object "GDP" for a country $c \in C$
$A(GDP_{constant},SUR)$	Availability of "GDP" in "constant price data" for Suriname
$A(GDP_{current},SUR)$	Availability of "GDP" in "current price data" for Suriname
A(net lending,SUR)	Availability of "NL" (current price data) for Suriname
$A(T_{x.y},SUR)$	Availability of NAQ table "$T_{x.y}$" for Suriname
A(x,c)	Availability referring to the existence of data object $x \in X$ for a country $c \in C$
$A_c(x)$	Availability referring to the existence of data object $x \in X$ for a fixed country $c \in C$
$A_c^{-1}(1)$	Preimage set of the set of data objects $x \in X$, that are *available* for a fixed country $c \in C$
$A_{GDP}(c)$	Refer to A(GDP,c)
ANA	Data object $ANA \in \mathit{ANA}$ of ANA data reported to the UN (by a specific country c), i.e. "ANA" = $x \in X$ (the set of data objects)
ANA	Set of ANA data (NA indicators and corresponding reference years (t_{RY})), given by the National Accounts Questionnaire (NAQ)
ANAK	Data object $ANAK \in \mathit{ANA}$ of "known" ANA data, i.e. the most recent reference year t_{RY} in the current yearbook (for a specific country c) was already published in a previous YB
$A_{net\_lending}^{-1}(1)$	Preimage set of countries that have the indicator "net lending" available

$A_{SUR}^{-1}(1)$	Preimage set of assessed data objects that are available for Suriname
$A_x(c)$	Availability referring to the existence of a fixed data object $x \in X$ for a country $c \in C$
$A_x^{-1}(1)$	Preimage set of the set of countries $c \in C$ for which a fixed data object $x \in X$ is available
c	Country
C	Set of countries
C_{1993}	Set of countries that fulfills 1993 MRDS
$C_{1993(P)}$	Set of countries that fulfills at least partial 1993 MRDS, i.e. missing at most one NAQ table to fulfill the 1993 MRDS criteria
C_{1r}	Set C_{1r} of countries that fulfills "relaxed" milestone level 1, i.e. L_{1r}
C_{1s}	Set C_{1s} of countries that fulfills "strict" milestone level 1, i.e. L_{1s}
C_{2008}	Set of countries that fulfills 2008 MRDS
$C_{2008(P)}$	Set of countries that fulfills at least partial 2008 MRDS, i.e. missing at most one NAQ table to fulfill the 2008 MRDS criteria
C_{2r}	Set C_{2r} of countries that fulfills "relaxed" milestone level 2, i.e. L_{2r}
C_{2s}	Set C_{2s} of countries that fulfills "strict" milestone level 2, i.e. L_{2s}
$C_{4.6+4.7\vee4.9}$	Set $C_{4.6+4.7\vee4.9}$ of countries that fulfills $RS_{4.6+4.7\vee4.9}$
C_A	Set of countries that are "annual reports" of new ANA data
C_{All}	Set of all countries (referring to countries, areas, and territories) for which official NA data are produced, according to the country list of the MDG regional classification
$C_{Developed}$	Set C of countries of the "*Developed Regions*"
C_{FT}	Set of countries that are "first time reports" of ANA data
$C^{f}_{x.y}$	Set of countries that produce NAQ table $T^{f}_{x.y} \in T^*_{2008}$
C_{IS}	Set of countries that have all required NAQ tables $T^{nl}_{x.y} \in T_{IS}$, referring to institutional sectors until NL
$C_{IS(P)}$	Set of countries that produce at least all but one of the required set of NAQ tables $T^{nl}_{x.y} \in T_{IS}$, referring to institutional sectors until NL
C_K	Set of countries with "known" ANA data in the YB (the latest available reference year t_{RY} refers to a reference year that is at most five years back from the YB publication year (i.e. $t_{RY} \geq t\text{-}5 < t\text{-}1$))

C_{K2}	Subset ($C_{K2} \subseteq C_K$) of countries with "known" ANA data that have been already published by a previous YB, for which the most current available reference year t_{RY} of ANA data refers to the reference year two years before the publication year t_{PY} of the YB (i.e. t_{RY} = t-2)
CMR	Country code (ALPHA3) for Cameroon
C_{NA}	Set of countries that are "non-annual reports" of ANA data
$C^{nl}_{x.y}$	Set of countries that produce NAQ table $T^{nl}_{x.y} \in T^*_{2008}$
C_{OT}	Set of countries that are "onetime reports" of ANA data
C_R	Subset of countries c $\in$ C, that reported a set of ANA data for a new reference year t_{RY}(c), previously not published in the Yearbook (YB). $C_R \subseteq C_{YB} \subseteq C$
$C_{R(All)}$	Set of "MDG list" countries that reported data, i.e. set $C_{ALL} \cap C_R$ ($C_{R(All)} \supset C_{R(UN)}$)
$C_{R(UN)}$	Set of UN Member States that reported ANA data, i.e. set $C_{UN} \cap C_R$
C_{R2}	Subset ($C_{R2} \subseteq C_R$) of countries that reported a new reference year t_{RY} of ANA data, for which the most current available reference year of ANA data t_{RY} is at most two years back from the publication year t_{PY} of the YB (i.e. $t_{RY} \geq$ t-2)
C^{TLR}_n	Set of countries with time lag of data publication of value n $\in$ $\boldsymbol{N}$ in years
C^T_n	Set of countries with timeliness T(c) of value n $\in$ $\boldsymbol{N}$ in years
$C^T_{YB,n}$	Set of countries with timeliness of value n $\in$ $\boldsymbol{N}$ in years (by UN timeliness assessment methodology)
CUB	Country code (ALPHA3) for Cuba
C_{UN}	Set of UN Member States
C_{YB}	Set of countries for which ANA data are published in the YB (the latest available reference year t_{RY} is at most five years back from the YB publication year (i.e. $t_{RY} \geq$ t-5)), $C_{YB} = C_R \cup C_K$
C_{YB2}	Subset ($C_{YB2} \subseteq C_{YB}$) of countries published in the YB, for which the most current available reference year t_{RY} of ANA data is at most two years back from the publication year t_{PY} of the present YB (i.e. $t_{RY} \geq$ t-2), $C_{YB2} = C_{R2} \cup C_{K2}$
$d_{1993}(s,\tilde{s})$	Distance between n-tuple s inside the superset S_{1993} (s $\in$ S_{1993}) and the subset of relevant n-tuples given by $MRDS_{1993}$ ($\tilde{s} \in MRDS_{1993}$)
$d_{2008}(s,\tilde{s})$	Distance between n-tuple s inside the superset S_{2008} (s $\in$ S_{2008}) and the subset of relevant n-tuples given by $MRDS_{2008}$ ($\tilde{s} \in MRDS_{2008}$)

$d_{IS}(s,\tilde{s})$	Distance between n-tuple s inside the superset S_{IS} ($s \in S_{IS}$) and the subset of relevant n-tuples given by IS ($\tilde{s} \in IS$)
DEU	Country code (ALPHA3) for Germany
DZA	Country code (ALPHA3) for Algeria
EGY	Country code (ALPHA3) for Egypt
$\Delta^{*}_{i,r}$	Refer to $R_{r,i}$, i.e. the "Revision" if intermediate revisions are observed
Δ^{*}_{r}	Refer to Δ^{*t}_{r}, the index t can be dropped, since the different vintages of θ are of the same reference period
$\Delta^{*t}_{0,r}$	Refer to R^{t}_{r}, i.e. the measure "Revision"
Δ^{*t}_{r}	Refer to $\Delta^{*t}_{0,r}$; $\Delta^{*t}_{r} = \Delta^{*t}_{0,r}$
$\Delta_{0,r}$	Refer to $\Delta^{t}_{0,r}$, the index t can be dropped, since the different vintages of θ are of the same reference period
$\Delta_{i,r}$	Absolute change of any (earlier) released numerical value (θ_i, with i = 0, 1, …, 99) to a subsequent released numerical value (θ_r, with i < r)
Δ_{r}	Refer to $\Delta_{0,r}$
$\Delta^{t}_{0,r}$	Absolute change of the initial (first) released numerical value (θ_0) to the subsequent released numerical value(s) (θ_r, with r = 1, 2, … 100)
Δ^{t}_{r}	Refer to $\Delta^{t}_{0,r}$
ε	Absolute deviation of the observed numerical value (θ) to the fully correct (true) numerical value (θ^{*}) for the evaluated indicator
ε_r	Absolute deviation, refer to ε, dependent on different observations of θ, i.e. θ_r, with r = 0, …, 100
$\mathbf{F}_{1.3\vee 4.1}(c)$	Image vector consisting of the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, of NAQ tables 1.3 or 4.1, i.e. where $T_{1.3} \in T_{1993}$ and $T_{4.1} \in T_{1993}$
$\mathbf{F}_{1.3\vee 4.1(nl)}(c)$	Image vector comprising the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, of NAQ tables 1.3 or 4.1 until NL, i.e. $T^{nl}_{1.3}$ and $T^{nl}_{4.1}$, where $T^{nl}_{x.y} \in T_{2008}$
$\mathbf{F}_{1993}$	Vectorial mapping indicating which of the NAQ tables of T_{1993} are available in $c \in C$ (that is produced by c) and which are not
$\mathbf{F}_{1993}(c)$	Image vector consisting of the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, where $T_{x.y} \in T_{1993}$

$\mathbf{F}_{1993}{}^{-1}(\mathrm{MRDS}_{1993})$ Preimage set of countries that fulfill 1993 MRDS criteria, i.e. of MRDS_{1993} under F_{1993}

$\mathbf{F}_{1993}{}^{-1}(\mathrm{MRDS}_{1993(p)})$ Preimage set of countries that fulfill partial 1993 MRDS, i.e. of $\mathrm{MRDS}_{1993(p)}$ under F_{1993}

$\mathbf{F}_{1r}$ Vectorial mapping indicating which of the NAQ tables of T_{1r} are available in $c \in C$ (that is produced by c) and which are not

$\mathbf{F}_{1r}(c)$ Image vector consisting of the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, where $T_{x.y} \in T_{1r}$

$\mathbf{F}_{1r}{}^{-1}(L_{1r})$ Preimage set of countries that fulfill "relaxed" milestone level 1, i.e. set L_{1r}, i.e. of L_{1r} under $\mathbf{F}_{1r}$

$\mathbf{F}_{1s}$ Vectorial mapping indicating which of the NAQ tables of T_{1s} are available in $c \in C$ (that is produced by c) and which are not

$\mathbf{F}_{1s}(c)$ Image vector consisting of the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, where $T_{x.y} \in T_{1s}$

$\mathbf{F}_{1s}{}^{-1}(L_{1s})$ Preimage set of countries that fulfill "strict" milestone level 1, i.e. set L_{1s}, i.e. of L_{1s} under $\mathbf{F}_{1s}$

$\mathbf{F}_{2008}$ Vectorial mapping indicating which of the NAQ tables of T_{2008} are available in $c \in C$ (that is produced by c) and which are not

$\mathbf{F}_{2008}(c)$ Image vector consisting of the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, where $T_{x.y} \in T_{2008}$, respectively $T^{nl}_{x.y} \in T_{2008}$

$\mathbf{F}_{2008}{}^{-1}(\mathrm{MRDS}_{2008})$ Preimage set of MRDS_{2008} under F_{2008}, i.e. of countries that fulfill 2008 MRDS criteria

$\mathbf{F}_{2008}{}^{-1}(\mathrm{MRDS}_{2008(p)})$ Preimage set of countries that fulfill partial 2008 MRDS, i.e. of $\mathrm{MRDS}_{2008(p)}$ under F_{2008}

$\mathbf{F}_{2r}$ Vectorial mapping indicating which of the NAQ tables of T_{2r} are available in $c \in C$ (that is produced by c) and which are not

$\mathbf{F}_{2r}(c)$ Image vector consisting of the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, where $T_{x.y} \in T_{2r}$

$\mathbf{F}_{2r}{}^{-1}(L_{2r})$ Preimage set of countries that fulfill "relaxed" milestone level 2, i.e. set L_{2r}, i.e. of L_{2r} under $\mathbf{F}_{2r}$

$\mathbf{F}_{2s}$ Vectorial mapping indicating which of the NAQ tables of T_{2s} are available in $c \in C$ (that is produced by c) and which are not

$\mathbf{F}_{2s}(c)$ Image vector consisting of the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, where $T_{x.y} \in T_{2s}$, $T^{nl}_{x.y} \in T_{2s}$, and $T^{f}_{x.y} \in T_{2s}$

$\mathbf{F}_{2s}{}^{-1}(L_{2s})$ Preimage set of countries that fulfill "strict" milestone level 2, i.e. set L_{2s}, i.e. of L_{2s} under $\mathbf{F}_{2s}$

$\mathbf{F}_{4.3+4.4\vee4.8}(c)$	Image vector comprising the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, of NAQ tables 4.3, 4.4, and 4.8 until NL, i.e. $T^{nl}_{4.3}$, $T^{nl}_{4.4}$, and $T^{nl}_{4.8}$, where $T^{nl}_{x.y} \in T_{2008}$
$\mathbf{F}_{4.6+4.7\vee4.9}(c)$	Image vector comprising the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, of NAQ tables 4.6, 4.7, and 4.9 until NL, i.e. $T^{nl}_{4.6}$, $T^{nl}_{4.7}$, and $T^{nl}_{4.9}$, where $T^{nl}_{x.y} \in T_{2008}$
$F^{f}_{x.y}(c)$	Function indicating whether or not NAQ table $T^{f}_{x.y}$ is available for country $c \in C$
${F^{f}_{x.y}}^{-1}(1)$	Preimage set of countries for which NAQ table $T^{f}_{x.y}$ is available, i.e. of 1 under $F^{f}_{x.y}$
$\mathbf{F}_{L}$	Vectorial mapping indicating which of the NAQ tables of milestone level L are available in $c \in C$ (that is produced by c) and which are not
$\mathbf{F}_{IS}(c)$	Image vector comprising the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, of NAQ tables $T^{nl}_{x.y} \in T_{IS}$, referring to institutional sectors until NL
$\mathbf{F}_{IS}^{-1}(IS_{(P)})$	Preimage set of countries fulfilling partial *IS*, i.e. that produce at least all but one of the required set of NAQ tables, i.e. of $IS_{(P)}$ under $\mathbf{F}_{IS}$
$F^{nl}_{x.y}(c)$	Function indicating whether or not NAQ table $T^{nl}_{x.y}$ is available for country $c \in C$
${F^{nl}_{x.y}}^{-1}(1)$	Preimage set of countries for which NAQ table $T^{nl}_{x.y}$ is available, i.e. of 1 under $F^{nl}_{x.y}$
$F_{x.y}(c)$	Function indicating whether or not NAQ table $T_{x.y}$ respectively $T^{nl}_{x.y}$ is available for country $c \in C$
$F_{x.y}^{-1}(1)$	Preimage set of countries for which NAQ table $T_{x.y}$ is available, i.e. of 1 under $F_{x.y}$
GDP	Data object GDP (Gross Domestic Product) data, i.e. "GDP" = $x \in X$ (the set of data objects)
$GDP_{(y,q)}$	GDP value for quarter (y,q), if t = (y,q) refers to the reference period of year y and quarter q
$GDP_{(y-1,q)}$	GDP value for the same quarter as $GDP_{(y,q)}$, one year earlier (y-1,q)
GNDI	Data object GNDI (Gross National Disposable Income) data, i.e. "GNDI" = $x \in X$ (the set of data objects)
GNI	Data object GNI (Gross National Income) data, i.e. "GNI" = $x \in X$ (the set of data objects)
GRD	Country code (ALPHA3) for Grenada
g^{t}_{month}	YoY growth rate of monthly data for reference period t = (y,m) with year y and month m

$h^{nl}_{x.y}$ — Share (fraction) of countries that have NAQ table $T^{nl}_{x.y}$ available, in relation to the total number of countries C, with $0 \leq h^{nl}_{x.y} \leq 1$

$h^{f}_{x.y}$ — Share of countries that have NAQ table $T^{f}_{x.y}$ available, in relation to the total number of countries C, with $0 \leq h^{f}_{x.y} \leq 1$

h_{1993} — Share of countries that fullfill the 1993 MRDS in relation to the total number of countries C, with $0 \leq h_{1993} \leq 1$

$h_{1993(p)}$ — Share of countries that fullfill partial 1993 MRDS in relation to the total number of countries C, with $0 \leq h_{1993(p)} \leq 1$

h_{1r} — Share of countries that fullfill "relaxed" milestone level 1, i.e. $L_{1r,}$ in relation to the total number of countries C, with $0 \leq h_{1r} \leq 1$

h_{1s} — Share of countries that fullfill "strict" milestone level 1, i.e. $L_{1s,}$ in relation to the total number of countries C, with $0 \leq h_{1s} \leq 1$

h_{2008} — Share of countries that fullfill the 2008 MRDS in relation to the total number of countries C, with $0 \leq h_{2008} \leq 1$

$h_{2008(p)}$ — Share of countries that fullfill partial 2008 MRDS in relation to the total number of countries C, with $0 \leq h_{2008(p)} \leq 1$

h_{2r} — Share of countries that fullfill "relaxed" milestone level 2, i.e. $L_{2r,}$ in relation to the total number of countries C, with $0 \leq h_{2r} \leq 1$

h_{2s} — Share of countries that fullfill "strict" milestone level 2, i.e. $L_{2s,}$ in relation to the total number of countries C, with $0 \leq h_{2s} \leq 1$

$h_{4.6+4.7\vee4.9}$ — Share of countries that fullfill $RS_{4.6+4.7\vee4.9}$, in relation to the total number of countries C, with $0 \leq h_{4.6+4.7\vee4.9} \leq 1$

h_{IS} — Share of countries that produce the set IS available, in relation to the total number of countries C, with $0 \leq h_{IS} \leq 1$

h_n — Share of countries regarding the set C^{T}_{n} of countries with a particular timeliness of value $n \in \boldsymbol{N}$

$h_{net\_lending}$ — Share (h) of countries for which "net lending" is available

h_x — Share (h) of countries for which the data object $x \in X$ is available

$h_{x.y}$ — Share of countries that have NAQ table $T_{x.y}$ available, in relation to the total number of countries C, with $0 \leq h_{x.y} \leq 1$

h_{xna} — Share (h) of countries $c \in C$ where the data object $x \in R \subseteq X$ of the relevant set of objects is missing

$h_{YB,n}$	Share of countries regarding the set $C^{T}_{YB,n}$ of countries with a particular timeliness of value n $\in$ $\boldsymbol{N}$ (by UN timeliness methodology)
$I_{1993}(c)$	Indicator function whether or not a country c $\in$ C is member of the set C_{1993} of countries that fulfills 1993 MRDS
$I_{1993(p)}(c)$	Indicator function whether or not a country c $\in$ C is member of the set $C_{1993(p)}$ of countries that fulfills partial 1993 MRDS
$I_{1993(p)}^{-1}(1)$	Preimage set of countries that are members of $C_{1993(p)}$, i.e. of 1 under $I_{1993(p)}$
$I_{1993}^{-1}(1)$	Preimage set of countries that that are members of C_{1993}, i.e. of 1 under I_{1993}
$I_{1r}(c)$	Indicator function whether or not a country c $\in$ C is member of the set C_{1r} of countries that fulfills "relaxed" milestone level 1, i.e. L_{1r}
$I_{1r}^{-1}(1)$	Preimage set of countries that that are members of C_{1r}, i.e. of 1 under I_{1r}
$I_{1s}(c)$	Indicator function whether or not a country c $\in$ C is member of the set C_{1s} of countries that fulfills "strict" milestone level 1, i.e. L_{1s}
$I_{1s}^{-1}(1)$	Preimage set of countries that that are members of C_{1s}, i.e. of 1 under I_{1s}
$I_{2008}(c)$	Indicator function whether or not a country c $\in$ C is member of the set C_{2008} of countries that fulfills 2008 MRDS
$I_{2008(p)}(c)$	Indicator function whether or not a country c $\in$ C is member of the set $C_{2008(p)}$ of countries that fulfills partial 2008 MRDS
$I_{2008(p)}^{-1}(1)$	Preimage set of countries that are members of $C_{2008(p)}$, i.e. of 1 under $I_{2008(p)}$
$I_{2008}^{-1}(1)$	Preimage set of countries that are members of C_{2008}, i.e. of 1 under I_{2008}
$I_{2r}(c)$	Indicator function whether or not a country c $\in$ C is member of the set C_{2r} of countries that fulfills "relaxed" milestone level 2, i.e. L_{2r}
$I_{2r}^{-1}(1)$	Preimage set of countries that that are members of C_{2r}, i.e. of 1 under I_{2r}
$I_{2s}(c)$	Indicator function whether or not a country c $\in$ C is member of the set C_{2s} of countries that fulfills "strict" milestone level 2, i.e. L_{2s}
$I_{2s}^{-1}(1)$	Preimage set of countries that that are members of C_{2s}, i.e. of 1 under I_{2s}
$I_{4.6+4.7\vee 4.9}(c)$	Indicator function whether or not a country c $\in$ C is member of the set $C_{4.6+4.7\vee 4.9}$ of countries that fulfills $RS_{4.6+4.7\vee 4.9}$
$I_{4.6+4.7\vee 4.9}^{-1}(1)$	Preimage set of countries that are members of $C_{4.6+4.7\vee 4.9}$, i.e. of 1 under $I_{4.6+4.7\vee 4.9}$
$I^{f}_{x.y}$	Indicator function regarding whether or not a country c $\in$ C is member of the set $C^{f}_{x.y}$ of countries for which NAQ table $T^{f}_{x.y}$ is available

$I^{f}_{x.y}{}^{-1}(1)$	Preimage set of countries for which NAQ table $T^{f}_{x.y}$ is available, i.e. of 1 under $I^{f}_{x.y}$
$I_{IS(P)}{}^{-1}(1)$	Preimage set of countries for which all at least all but one required NAQ tables $T^{nl}_{x.y} \in T_{IS}$ are available, i.e. of 1 under $I_{IS(p)}$
$I_{IS}{}^{-1}(1)$	Preimage set of countries for which all required NAQ tables $T^{nl}_{x.y} \in T_{IS}$ (institutional sectors until NL) are available, i.e. of 1 under I_{IS}
$I^{nl}_{x.y}$	Indicator function regarding whether or not a country c $\in C$ is member of the set $C^{nl}_{x.y}$ of countries for which NAQ table $T^{nl}_{x.y}$ is available
$I^{nl}_{x.y}{}^{-1}(1)$	Preimage set of countries for which NAQ table $T^{nl}_{x.y}$ is available, i.e. of 1 under $I^{nl}_{x.y}$
IS	Relevant subset of state vectors, i.e. where s $\in S_{IS}$ i.e. where s indicates the existence of all of the required NAQ tables $T^{nl}_{x.y} \in T_{IS}$
$IS_{(P)}$	set of relevant n-tuples that differ by not more than one missing NAQ table $T_{x.y}$ indicated by $t_{x.y}$ from the set of n-tuples given by *IS*
I^{T}_{n}	Indicator function regarding whether or not a specific country c is element of the preimage set C^{T}_{n} for a specific timeliness in years n $\in N$
$I_{x.y}$	Indicator function regarding whether or not a country c $\in C$ is member of the set $C_{x.y}$ of countries for which NAQ table $T_{x.y}$ is available
$I_{x.y}{}^{-1}(1)$	Preimage set of countries for which NAQ table $T_{x.y}$ is available, i.e. of 1 under $I_{x.y}$
L_{1r}	Subset of state vectors fulfilling “relaxed” milestone level 1 criteria
L_{1s}	Subset of state vectors fulfilling “strict” milestone level 1 criteria
L_{2r}	Subset of state vectors fulfilling “relaxed” milestone level 2 criteria
L_{2s}	Subset of state vectors fulfilling “strict” milestone level 2 criteria
LAO	Country code (ALPHA3) for Lao People’s Democratic Republic
LBY	Country code (ALPHA3) for Arab Jamahiriya
$M_{(y,m)}$	Value of a monthly time series for month (y,m), if t = (y,m) refers to the reference period of year y and month m
$M_{(y-1,m)}$	Value of a monthly time series for the same month as $M_{(y,m)}$, , one year earlier (y-1,m)
MAR	Country code (ALPHA3) for Morocco
$MAR^{t}_{0,r}$	Mean absolute revision (in percentage points) from vintage 0 to r (with “r” referring in the empirical study to specific later revision time points)

MAR_{r}^{t}	Refer to $\mathrm{MAR}_{0,r}^{t}$
$MRDS_{1993}$	Relevant subset of state vectors of 1993 MRDS
$MRDS_{1993(P)}$	Set of relevant n-tuples s that fulfill at least partial 1993 MRDS, i.e. the n-tuple $s \in MRDS_{1993(P)}$ only differ by the non-existence of maximal one needed NAQ table of the n-tuple $s \in MRDS_{1993}$
$MRDS_{2008}$	Relevant subset of state vectors of 2008 MRDS
$MRDS_{2008(P)}$	Set of relevant n-tuples s that fulfill at least partial 2008 MRDS, i.e. the n-tuple $s \in MRDS_{2008(P)}$ only differ by the non-existence of maximal one needed NAQ table of the n-tuple $s \in MRDS_{2008}$
$\mathrm{MR}_{0,r}^{t}$	Mean revision (in percentage points) from vintage 0 to vintage r (with "r" referring in the empirical study to specific later revision time points)
MR_{r}^{t}	Refer to $\mathrm{MR}_{0,r}^{t}$
M_t	Value of a monthly time series in month t
M_{t-12}	Value of a monthly time series for month t of the previous year
n	Lengths of the n-tuple (of availability measure)
n	$\|RQ\|$ = n, the number of observed reference quarters for a fixed revision interval [0,r], from the first (initial) value (vintage 0) to the later at value (vintage r) (of revision measures)
n	Value $n \in \boldsymbol{N}$ and $n \in \boldsymbol{N}_0$, respectively in years of "timeliness" and "time lag"
$\boldsymbol{N}$	Set of natural numbers, i.e. {1, 2, …}
$\boldsymbol{N}_0$	Set of natural numbers including 0, i.e. {0, 1, 2, …}
NAQ	Set of NA indicators requested by the NAQ
NL	Data object NL (Net lending) data, i.e. "NL" = $x \in X$ (the set of data objects)
Ω	Total set Ω of considered objects (i.e. whole)
P	Subset $P \subseteq \Omega$ of considered objects
$\Pi_{1993}(s)$	Projective mapping for a given state vector $s \in S_{2008}$ that gives $\tilde{s} \in S_{1993}$
q	Reference period of quarter q
QNA	Data object QNA data, i.e. "QNA" = $x \in X$ (the set of data objects)
Q_t	Value of a quarterly time series in quarter t
Q_{t-1}	Value of a quarterly time series for the previous quarter of t

Q_{t-4}	Value of a quarterly time series for quarter t of the previous year
r	Cardinality $r < 2^n$, of subset of relevant state vectors within $S = \{0,1\}^n$
$\bar{R}$	Mean revision, refer to $MR^t_{0,r}$
R	Subset $R \subseteq X$ of relevant data objects
$\boldsymbol{R}^+$	Set of real numbers, not including 0
R^t_r	Revision, i.e. directional size of revision (i.e. the change considering the sign) from the initial vintage 0 to vintage r, with r = 1, 2, … 100 (with "r" referring in the empirical study to specific later revision time points)
$RMAR_{McK\&A}$	Relative Mean Absolute Revision of McKenzie and Adam (2007)
$RMAR_{Our}$	Relative Mean Absolute Revision of our study
$RMAR^t_{0,r}$	Relative mean absolute revision (in percentage points) from vintage 0 to vintage r (with "r" referring in the empirical study to specific later revision time points)
$RMAR^t_r$	Refer to $RMAR^t_{0,r}$
ROM	Country code (ALPHA3) for Romania
RQ	Set of reference periods t with data for a fixed revision interval [0,r], from the first (initial) value (vintage 0) to the later value (vintage r) of the same reference period (t)
$R_{r,i}$	Intermediate revisions $R_{r,i}$ between an earlier released value θ_i and a later released value θ_r
$RS_{1.3\vee4.1}$	Relevant set of state vectors, i.e. where $s \in RS_{1.3\vee4.1}$, i.e. where s indicates the existence of at least one of the NAQ tables 1.3 or 4.1
$RS_{1.3\vee4.1(nl)}$	Relevant set of state vectors, i.e. where $s \in RS_{1.3\vee4.1(nl)}$, i.e. where s indicates the existence of at least one of NAQ tables 1.3 or 4.1 until NL
$RS_{4.3+4.4\vee4.8}$	Relevant set of state vectors, i.e. where $s \in RS_{4.3+4.4\vee4.8}$, i.e. where s indicates the existence of both NAQ tables 4.3 and 4.4 until NL or NAQ table 4.8 until NL
$RS_{4.6+4.7\vee4.9}$	Relevant set of state vectors, i.e. where $s \in RS_{4.6+4.7\vee4.9}$, i.e. where s indicates the existence of both NAQ tables 4.6 and 4.7 until NL or NAQ table 4.9 until NL
$\tilde{s}$	State vectors, i.e. n-tuples, of superset S
s	State vectors, i.e. n-tuples, of superset S
S	Superset of state vectors, i.e. n-tuples
$S_{1.3\vee4.1}$	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of at least one of NAQ tables 1.3 or 4.1, i.e. of $T_{1.3} \in T_{1993}$ and $T_{4.1} \in T_{1993}$

$S_{1.3\vee 4.1(\mathrm{nl})}$	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of NAQ tables 1.3 and 4.1 until NL, i.e. of $\mathrm{T}_{1.3}^{\mathrm{nl}} \in T_{2008}$ and $\mathrm{T}_{4.1}^{\mathrm{nl}} \in T_{2008}$
S_{1993}	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $\mathrm{T}_{\mathrm{x.y}} \in T_{1993}$
$S_{1\mathrm{r}}$	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $\mathrm{T}_{\mathrm{x.y}} \in T_{1\mathrm{r}}$
$S_{1\mathrm{s}}$	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $\mathrm{T}_{\mathrm{x.y}} \in T_{1\mathrm{s}}$
S_{2008}	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $\mathrm{T}_{\mathrm{x.y}}$ respectively $\mathrm{T}_{\mathrm{x.y}}^{\mathrm{nl}} \in T_{2008}$
$S_{2\mathrm{r}}$	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $\mathrm{T}_{\mathrm{x.y}} \in T_{2\mathrm{r}}$
$S_{2\mathrm{s}}$	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $\mathrm{T}_{\mathrm{x.y}} \in T_{2\mathrm{s}}$, $\mathrm{T}_{\mathrm{x.y}}^{\mathrm{nl}} \in T_{2\mathrm{s}}$, and $\mathrm{T}_{\mathrm{x.y}}^{\mathrm{f}} \in T_{2\mathrm{s}}$
$S_{4.3+4.4\vee 4.8}$	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, NAQ tables 4.3, 4.4, and 4.8 until NL, i.e. of $\mathrm{T}_{4.3}^{\mathrm{nl}} \in T_{2008}$, $\mathrm{T}_{4.4}^{\mathrm{nl}} \in T_{2008}$, and $\mathrm{T}_{4.8}^{\mathrm{nl}} \in T_{2008}$
$S_{4.6+4.7\vee 4.9}$	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, NAQ tables 4.6, 4.7, and 4.9 until NL, i.e. of $\mathrm{T}_{4.6}^{\mathrm{nl}} \in T_{2008}$, $\mathrm{T}_{4.7}^{\mathrm{nl}} \in T_{2008}$, and $\mathrm{T}_{4.9}^{\mathrm{nl}} \in T_{2008}$
S_{IS}	Superset of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $\mathrm{T}_{\mathrm{x.y}}^{\mathrm{nl}} \in T_{\mathrm{IS}}$
t	Time point referring to a calendar year, quarter, or generally a date
t_0	Current year regarding the time point of publication of the YB
t_0+12M	12 month later vintage, referred to as "t_0+1Y", i.e. one year later vintage after first publication of data for a given reference period
t_0+1Y	One year later vintage after first publication of data for a given reference period
t_0+24M	24 month later vintage after first publication of data for a given reference period, referred to as "t_0+2Y"
t_0+2Y	Two years later vintage after first publication of data for a given reference period
t_0+36M	36 month later vintage after first publication of data for a given reference period, referred to as "t_0+3Y"

t_0+3Y	Three years later vintage after first publication of data for a given reference period
t_0+5M	Five month later vintage, i.e. sub-annual vintage, after first publication of data for a given reference period
t_0+end	the latest available vintage of data for a given reference period, given it is at least 48 month after vintage t_0
t_E	End time point of the reference period (i.e. end of the described event)
$t^{f}_{x.y}$	Indicator variable, regarding the existence and non-existence, respectively, of each NAQ table $T^{f}_{x.y}$
$t^{nl}_{x.y}$	Indicator variable, regarding the existence and non-existence, respectively, of each NAQ table $T^{nl}_{x.y}$
t_P	Time point of the publication of data (related to t_E)
$t_{P(0)}$(x,c)	Time point of publication of the initial vintage of data object x $\in X$, for country c $\in C$
$t_{P(r)}$(x,c)	Time point of publication of the later vintage of data object x $\in X$, for country c $\in C$
t_{PY}	Publication year
t_{PY}(c)	Publication year of ANA data publication at the international level (i.e. at the UN) of country c $\in C$, referring to t_{PY}(ANA,c)
t_{PYK}	Publication year of the first time release of “known” data
t_{RY}	Reference year
t_{TR}	Time point of data *reporting* (transmission) to the IO
$t_{x.y}$	Indicator variable, regarding the existence and non-existence, respectively, of each NAQ table $T_{x.y}$ respectively $T^{nl}_{x.y}$
T(ANA,c)	Timeliness of ANA publication at the international level (i.e. at the UN) of country c $\in C_R$, abbreviated by T(c)
T(ANAK,c)	Timeliness of “known” ANA data (“ANAK”) in years n $\in \boldsymbol{N}$ of country c $\in C_K$, at time point (year t_{PYK}) when published by UN for the first time
T(c)	Refer to T(ANA,c); the symbol “ANA” is dropped for convenience, since we study ANA data
T(v,c)	Timeliness of publication of NAQ indicator (i.e. variable v) at the at the UN of country c $\in C_R$
T(x,c)	Timeliness of publication of data object x $\in X$ of country c $\in C$
T(x.y,c)	Timeliness of publication of NAQ table x.y $\in X.Y$ at the UN of country c $\in C_R$

T^*_{1993}	Subset of T_{1993}, of individual NAQ tables assessed for their availability
T^*_{2008}	Subset of T_{2008}, of individual NAQ tables assessed for their availability
$\mathrm{T}_{1.1,}\ \mathrm{T}_{1.2,}\ \mathrm{T}_{1.3}$	NAQ table 1.1, NAQ table 1.2, respectively NAQ table 1.3
T_{1993}	Set of potential NAQ tables of 1993 MRDS
$T_{1\mathrm{r}}$	Set of potential NAQ tables of "relaxed" milestone level 1
$T_{1\mathrm{s}}$	Set of potential NAQ tables of "strict" milestone level 1
$\mathrm{T}_{2.1,}\ \mathrm{T}_{2.2,}\ \mathrm{T}_{2.3}$	NAQ table 2.1, NAQ table 2.2, respectively NAQ table 2.3
T_{2008}	Set of potential NAQ tables of 2008 MRDS
$T_{2\mathrm{r}}$	Set of potential NAQ tables of "relaxed" milestone level 2
$T_{2\mathrm{s}}$	Set of potential NAQ tables of "strict" milestone level 2
$\mathrm{T}_{4.1,}\ \mathrm{T}_{4.2}$	NAQ table 4.1, respectively NAQ table 4.1
$\mathrm{T}^{\mathrm{f}}_{4.2}$	Financial account, symbolized by "f", of NAQ table 4.2
$\mathrm{T}^{\mathrm{f}}_{\mathrm{x.y}}$	Financial account, symbolized by "f", of NAQ table x.y, $\mathrm{T}_{\mathrm{x.y}} \in T_{\mathrm{x.y}}$
T_{IS}	Set of NAQ tables 4.1, …, 4.7, and 4.9 referring to institutional sectors
$\mathrm{T}^{\mathrm{nl}}_{1.3}$	NAQ table 1.3 existent until "net lending" indicator of capital account
$\mathrm{T}^{\mathrm{nl}}_{4.1}, \ldots, \mathrm{T}^{\mathrm{nl}}_{4.9}$	NAQ tables 4.1, …, 4.9 existent until "net lending" indicator of capital account
$\mathrm{T}^{\mathrm{nl}}_{\mathrm{x.y}}$	NAQ table $\mathrm{T}^{\mathrm{nl}}_{\mathrm{x.y}}$, where nl indicates the NAQ table needs to be existent until the "net lending" indicator of the capital account
$\mathrm{T_x(c)}$	Timeliness of a fixed data object x $\in X$ of country c $\in C$
$\mathrm{T}_{\mathrm{x.y}}$	NAQ table x.y, $\mathrm{T}_{\mathrm{x.y}} \in T_{\mathrm{x.y}}$
$\mathrm{T_{YB}(c)}$	Measure for the kind of timeliness assessed by the UN methodology of ANA data publication of country c $\in C$, referring to $\mathrm{T_{YB}(ANA,c)}$
THA	Country code (ALPHA3) for Thailand
TLR	Time Lag of Reported ANA data, referring to a time period
TLR(ANA,c)	Time Lag of Reported ANA at the time point of their publication at the UN of country c $\in C_{\mathrm{R}}$
TLR(c)	refer to TLR(ANA,c); the symbol "ANA" is dropped, since we study ANA data
TM(x,c)	Maturity of a singular data object x $\in X$ at time point t_{Pr} of publication of the later vintage of data for a country c $\in C$

TO_{1993}	Observation period (TO) for the availability assessment of the NAQ tables of the 1993 MRDS
TO_{2008}	Observation period (TO) regarding NAQ tables of 2008 MRDS
TUN	Country code (ALPHA3) for Tunisia
θ	Observed numerical value
θ^*	Fully correct (true) numerical value
θ_r	Different observations of θ, i.e. θ_r, with r = 0, …, 100
θ_0^t	First published value of θ for reference period t
θ_0^t	Values of the indicator θ at vintages 0 of the same reference period (t)
θ_r^t	Value of θ of the reference period t at time point of a later vintage r
θ_r^t	Values of the indicator θ at later released vintages r of the same reference period (t)
v	Data object "NA indicator" (variable $v \in V$ [the set of NA indicators]), i.e. "v" = $x \in X$ (the set of data objects)
x	Data object $x \in X$
x.y	Data object NAQ table $x.y \in X.Y$ (the index set of NAQ table numbers), i.e. "x.y" = $x \in X$ (the set of data objects)
x^t	YoY rates of change in percent for quarterly GDP
x_0^t	Value of quarterly GDP in constant price data YoY growth rate in percent, at the time point of the first vintage 0, for reference period t
x_{QoQ}^t	QoQ rates of change in percent for quarterly GDP
x_r^t	Value of quarterly GDP in constant price data YoY growth rate in percent, at time point of any vintage r other than the first release, for reference period t
X	Set of data objects
$X.Y$	Set of data objects (NAQ tables) identified by the index set of NAQ table numbers
$X.Y_{1993}$	Index set regarding NAQ tables $T_{x.y} \in T_{1993}$
$X.Y_{2008}$	Index set regarding NAQ tables $T_{x.y} \in T_{2008}$, respectively $T_{x.y}^{nl} \in T_{2008}$
$X.Y_{IS}$	Index set to the NAQ tables $T_{x.y}^{nl} \in T_{IS}$
y	Reference period of year y

I. Introduction

In this part we provide the motivation, followed by an executive summary of our study, and a review of previous works considered relevant for our study.

1 Motivation

The *SNA framework* provides a structured, "coherent, consistent and integrated" set of National Accounts (NA) data[1] (United Nations 1993, 1)[2]. This data set provides *information* for monitoring economic development and effects of economic policies.

For the evaluation of *Data Quality (DQ)* of official statistics, the International Organizations (IOs), i.e. primarily Eurostat, IMF and OECD (and some national agencies) have set up "*quality frameworks*". They are used by IOs to assess either *quality* of their own data (e.g. OECD self assessment) or of data produced by national agencies (e.g. IMF assessment of member countries). The OECD and UN carry out assessments of the DQ dimensions *reliability, timeliness*, and *availability* of official NA data. These assessments are based on the national data sets reported to these IOs. The *information* about the DQ of NA data of member countries is also used by international (and regional) *organizations* "to identify countries and regions which need technical assistance" (cf. United Nations 2001c, 8). Thus particular interest is paid to countries with poor data quality.

Producers and users of NA data are interested in the *quality* of NA statistics published by national agencies like National Statistical Institutes (NSIs) or Central Banks (CBs). The international community of users relies on data published through IOs, making the NA data available on behalf of member countries. This is particularly important in case national methodologies differ from the SNA standard. For data published by the UN via the NA Yearbook (YB), such differences are footnoted and explained. Data might even be computed by a country according to required standards especially for reporting to the IO (for example the US NIPA consistent with 1993 SNA for OECD).

---

[1] For definitions of SNA terms like capital account, gross domestic product, net lending, etc. refer to United Nations (1993), OECD (2000), UNSD (2009a), respectively United Nations (2009f).

[2] In the following the text will read "1993 SNA" instead of the reference to "United Nations (1993)", ditto for the 2008 SNA, i.e. the text will read "2008 SNA" instead of "United Nations (2009f)".

Countries often publish, i.e. produce, ANA data with significant *time lag* (particularly subsets of detailed data). *Time lags* of ANA data publication at the international level hamper cross-country comparison and economic analysis of countries and regions by researchers and IOs. An assessment of the *timeliness* of "reported" Annual National Accounts (ANA) data was done by the UN in 2008 (cf. UNSD 2008). This study provides only limited information regarding the assessment methodology. Moreover, the considered range of years and the clustering of assessed regions are insufficient for a detailed analysis of timeliness of NA data publication by regions.

The *availability* of ANA data is important for analysis of countries and cross-country comparison. Assessments of the *availability* of ANA data are regularly carried out by the UN. Again, the subdivision of considered country groups is insufficient. The reports following the 2008 study (UNSD 2008) provide even less detailed results. They are neither sufficient to inform data users about the data available at the UN, nor to identify country groups with weak NA data production capacity. They assess the availability of the data set that is currently considered the minimum requirement given the 1993 SNA ("1993 MRD"). While the extended minimum requirement given the 2008 MRDS ("2008 MRDS") is agreed on, it has not yet been assessed by the UN.

Reliability of the first disseminated GDP figure is important, since it is used for decision making and policy making. Later *revisions* of the first released figure may alter these decisions. Besides NSOs analyzing their own data, studies including cross-country comparisons have been carried out considering a number of OECD countries. However, these studies consider at most 20 countries, while we assess 34 countries. Moreover, previous cross-country comparisons consider a time period of analysis that includes changes between SNA versions (from 1993 SNA to ESA1995), i.e. data are computed using different SNA methodologies. This affects the results of the studies. Further, the type of growth rates analyzed is not suited for international comparison, for some of the utilized country data it should not be used to analyze revisions.

Our objectives are to assess the *timeliness*, *availability*, and *reliability* of NA data published at the international level. Moreover, we improve existing assessments by using more appropriate DQ measures, and we improve the methodologies of the UN assessments of *timeliness*, and *availability*, and of the OECD assessment of reliability.

2 Executive Summary

In the following we provide a summary of our analysis and suggested improvements of DQ assessments of *timeliness*, *availability*, and *reliability*. The questions included in this chapter are answered by our studies of the three DQ dimensions.

Timeliness

Our *assessment* of *timeliness* is based on ANA data reported by national agencies and published by the UN. The UN has the mandate to globally collect and disseminate official NA data (cf. United Nations 2009b, v). The data are published by the UN via the NA Yearbook (YB) and via the UN data portal "UNdata"[3]. We consider data published in 2009 via the 2008 YB (United Nations 2009a, b, c, d).

1) Will our assessment of *timeliness* of ANA publication find the data quality of countries to be much better than found by the UN assessment (UNSD 2008)?

We re-formulate the *definition* of the DQ dimension "Timeliness" (cf. DEF. 5, p. 18), to increase its precision and give the corresponding *formula*, since a *formula* is not provided by the DQ frameworks for official statistics. To assess *timeliness* of ANA data publication, we define two *DQ measures*. We present a verbal *definition* and a *formula* for the DQ measure of "timeliness" of ANA data publication at the international level by UN (cf. DEF. 32, p. 117). Further, we give a verbal *definition* and a *formula* for the "time lag" of ANA data published by UN (cf. DEF. 33, p. 122). Our definitions of these DQ measures cover only those countries that report a *new* reference year of ANA data to the UN. Thus, we assess the actual *timeliness* of ANA data at the time point of their *first* publication by the UN.

2) Does the UN study use an appropriate DQ *measure* to assess the *timeliness* of "reported" (cf. UNSD 2008, 7) ANA data?

The 2008 UN study (UNSD 2008) does not define the DQ measures of *timeliness* of "reporting" of ANA data. We derive a *definition* and a *formula* for this DQ measure (cf. DEF. 34, p. 138). Further, we describe the UN's assessment methodology. Our analysis finds that the used *methodology* provides misleading information about the

---

[3] The official country data are published in National Accounts: Main Aggregates and Detailed Tables and via UN's data dissemination platform "UNdata" (UNSD 2013c) online at http://data.un.org/.

actual status of *timeliness*. The UN assessment does not evaluate the *timeliness* of publication of data, when first published at the international level, but how *up-to-date* the ANA data are that are published in the YB. Thus, the results interfere with re-publications of pre-existing, i.e. “known”, data from previous reporting rounds. These “known” data have potentially been re-published for several years in previous YBs.

The results of the UN study (UNSD 2008) of *timeliness* of “*reported*” ANA data indicate a poor performance of countries. We expect significantly better results by our *improved methodology* using an appropriate DQ measure for *timeliness*.

3) Are the currently presented *subdivisions* of *regions* sufficient to facilitate a meaningful analysis?

4) What is the *range* of *timeliness* of the considered countries and what the worst *timeliness* of published data (given by value n $\in$ N in years), i.e. what is the longest *time lag*, indicating the poorest capacity to produce timely ANA data?

Deficiencies of the UN timeliness study are observed regarding both, the *clustering of country groups* and the computed *values* of *timeliness* in years. Only a subset of *occurring values* in years of *timeliness* is analyzed by the UN study. Furthermore, the UN study considers only a limited *set of countries* (e.g. the UN Member States) though data for additional countries are published by the UN.

5) How about the *timeliness* of countries, which are *not* UN Member States?

6) What is the difference in performance between *developed* and *developing regions*, respectively between different *country groups*?

7) Which *regions* might benefit from *capacity building* and *technical assistance*?

All the above questions are answered by our analysis and suggested improvements of the *assessment methodology* and our *empirical assessment*, respectively. Our improvements of the DQ measures and the clustering of country groups eliminate the major deficiencies of the 2008 UN study. Our empirical assessment is carried out using our *suggested methodology*. We present results for all occurring values n $\in \boldsymbol{N}$ in years of timeliness. Our assessment covers additional *regions* and subdivisions of *regions* and *country groups*. We also assess all suggested *regions*, based on the set of all countries (referring to countries, areas and territories) which includes 32 additional non-UN countries. Further, we make assessments of special country groups (like least developed countries), regional and international organizations (like OEDC), and important country groups (like EU27).

The *results* of our assessment are *compared* to the previous UN assessment (UNSD 2008). To facilitate better *comparison* of the differences between the results by the UN methodology and our improved methodology, we calculate the results by the UN assessment methodology using the same data set as in our new assessment.

The assessments of *timeliness* in our study demonstrate the successful application of our *improved methodology*. We can *summarize* that we define a DQ measure of *timeliness* of publication, as opposed to merely calculating how "up-to-date" data are. Due to our improvements, we calculate more meaningful results. Our study reveals a much stronger *timeliness* of ANA data publication, than the UN study (UNSD 2008).

Users of international data *benefit* from our assessment, as it informs better about the status of *timeliness* of internationally published NA statistics. Our assessment allows *researchers* to collect information regarding the time point (in years) when they can expect that a specific reference year of ANA data becomes available at the UN. Thus our assessments are an information source which *users* of NA can refer to for their work. We can identify regions and country groups with particularly strong and poor *timeliness*. Our results facilitate IOs to identify country groups that potentially benefit from aid and technical assistance to strengthen the NA capacity of countries.

Availability

The DQ dimension "Availability" is not part of the DQ frameworks of Eurostat, IMF, and OECD. We give the *definition* and *formula* for *availability* (cf. DEF. 7, p. 22). We assess the *availability* of ANA data in countries against internationally agreed data sets. We *define* our DQ measures, i.e. of the currently used 1993 "Minimum Requirement Data Set" (MRDS), and the 2008 MRDS (cf. DEF. 25, p. 59 respectively DEF. 26, p. 64). Further, "milestone levels" 1 and 2 are used to assess the progress of countries' regarding the production of SNA concepts at different levels. We assess milestone levels 1 and 2 according to the currently "relaxed" criteria of required data (cf. DEF. 30, p. 105 respectively DEF. 31, p. 109), and according to the original "strict" criteria (cf. DEF. 28, p. 98 respectively DEF. 29, p. 101). First time results regarding the fulfillment of "strict" milestone levels 1 and 2 are presented together with our 2008 MRDS assessment.

We closely follow the assessment methodology of the UN. We assess availability based on the data available in the 2008 YB and 2009 YB (United Nations 2010a). The clustering of analyzed country groups is the same as in our timeliness assessment.

1) Does the UN use appropriate DQ measures to assess ANA data availability?

The *shortcomings* of the *UN methodology* are particularly "relaxed" definitions of required data and the efforts necessary to obtain information regarding the existence of data for a country. Moreover, the status of data *availability* is assessed only of the current 1993 MRDS. Data availability of 2008 MRDS has not yet been studied.

The *methodology* of the *availability* assessment of 1993 MRDS is revised and that of 2008 MRDS is designed. Our 1993 and 2008 MRDS assessments are based on fixed time frames. The required data need to be available for at least one reference year in the time period of observation. The UN approach requires information about the reported NA data of multiple reporting rounds of the National Accounts Questionnaire[4] (NAQ). For the 1993 MRDS assessment, the condition to count NAQ tables as available is sufficient. For the 2008 MRDS, the condition is changed to fulfill the criteria for the NAQ tables requiring data in addition to the 1993 MRDS.

2) Will our 2008 MRDS assessment reveal as good data availability as for the 1993 MRDS, or will we find previously not detected or unexpected data gaps?

3) Will the change to the original "strict" criteria of the milestone levels 1 and 2 (with the 2008 MRDS) reveal data gaps that were previously covered up?

Our *availability assessment* shows which NAQ tables, corresponding to NA data sets like *GDP approaches* and different *SNA concepts*) are produced by a country. The unavailability of data in the *availability assessment* identifies a *data gap*.

The 1993 MRDS assessments by the UN indicate rather good data *availability* for the UN Member States. The results of our 1993 MRDS assessment are compared to previous 1993 MRDS assessments. These turn out to be nearly identical. Large differences are revealed between the 1993 and 2008 MRDS, as well as for the "relaxed" and "strict" milestone levels.

---

[4] Cf. 1993 National Accounts Questionnaire, UNSD (2012). Sample questionnaire online as download: http://unstats.un.org/unsd/nationalaccount/docs/SampleQuestionnaire.xls. (Last updated: 28-11-2012). Last accessed: 16-05-2013.

Our study provides new *insights* regarding *availability* of ANA data by regions and country groups. Our assessment allows *researchers* to collect information about *availability* of data of their interest. Regions and country groups can be identified that potentially benefit from technical assistance to *strengthen* NA *capacity* of countries.

Reliability

We re-formulate the *definition*, to increase its precision, and give a *formula* of the DQ dimension "Reliability" (cf. DEF. 11, p. 29). We use the established DQ measures "Mean Absolute Revision" (MAR) and "Relative Mean Absolute Revision" (RMAR) to assess reliability of first published data (cf. DEF. 19, p. 46 and DEF. 20, p. 47).

We make assessments of the needs of revision of first published GDP figures for 34 countries for which sufficient *revision* data are made available by the OECD. Results are presented for each country and for the average of the 30 OECD and four non-OECD G20 countries (separately). We provide first time assessments for 14 countries.

1) Are the analyzed time periods of previous studies suited for sensitive analysis?
2) Is the type of growth rate analyzed by previous studies appropriate?
3) Are the frequency and number of analyzed revision time points sufficient?

We analyze the assessment methodology of revision studies based on the OECD's Main Economic Indicators (MEI) Original Release Data and Revisions Database (cf. Di Fonzo 2005a). Two major deficiencies can be identified. Firstly, the results are based on time series that cover multiple revision regimes (with possibly different revision needs). Secondly, the revision intervals are calculated across different SNA series versions (which may affect the needs of revision due to changed concept, methods, or source data). This applies particularly in case of changes from 1968 SNA to ESA95 or 1993 SNA.

To facilitate appropriate cross-country comparability, we analyze Year-on-Year (YoY) growth rates (DEF. 21, p. 48) of Quarterly National Accounts (QNA) statistics. Previous studies use QoQ (Quarter-on-Quarter) growth rates (DEF. 23, p. 49).

The outcome of a *revision study* is considerably affected by the observed time period and the observed revision interval. Our *reliability* assessment provides data users with information about *needs of revision* of the first released GDP YoY growth rates compared to later *revised* values, i.e. at later vintages (cf. DEF. 13, p. 31) at revision time points five to 48 months after the first published vintage. Only data produced by the same SNA version (SNA 1968, SNA 1993, or ESA 95) are analyzed. Thus changes to the first published values due to *revisions* across SNA versions are eliminated. The studied time periods are cleansed from "estimated" values, i.e. real "estimates" (cf. DEF. 6, p. 20) that are not the result of a regular SNA production process. Those would affect results. Further, we state and eliminate shortcomings of the definition of the "latest" vintage (cf. DEF. 17, p. 41) of previous studies.

4) What is the reliability of the data of the 14 countries that have not been assessed until now, but for which OECD actually disseminates data?
5) Will our improved methodology show that reliability of first released GDP is better for the OECD member countries than found by previous studies?

Our cross-country comparison identifies countries that perform better or worse than their peer group. Depending whether the assessment is based on the *relative* or *absolute* size of *revisions*, i.e. RMAR respectively MAR, the rating of a country in the cross-country comparison might change or even reverse.

The results of our study can be used by researchers as resource for information about the reliability of countries' first released GDP data they utilize. They allow national agencies to identify their performance compared to other countries. A comparison of our results with those of a previous study reveals that our results indicate much lower *needs of revision*, and thus indicate a higher *reliability* of first released GDP data, than formerly found.

3 Previous Work

The *definition of "quality"* has changed over the past. *Definitions* include the "totality of features and characteristics of an entity that bears on its ability to satisfy stated and implied needs" (ISO 8402-1986 Quality-Vocabulary), simply the "fitness for use" (Juran 1974), the "conformance to requirement" (Crosby 1979), the "degree of excellence" (Concise Oxford Dictionary 1979), or the "fitness for use given a purpose" (Borowski and Lenz 2008).

Quality frameworks for (official) statistics were developed by several international organizations and by individual countries. In 2003 the *IMF* developed the "*Data Quality Assessment Framework (DQAF)"* as a generic framework (IMF 2003a) and as specialized frameworks for national accounts statistics, consumer price index, producer price index, government finance statistics, monetary statistics, balance of payments statistics, and external debt statistics. *Eurostat* developed the "*Definition of quality in statistics"*, the guidelines *"Standard Quality Report"* and a corresponding *"Glossary of Quality Terms"* (cf. Eurostat 2003a, 2003b, and 2003c, respectively). *OECD* developed the "*Quality Framework and Guidelines for OECD Statistical Activities"* (cf. OECD 2003[5]).

For an introduction to DQAF, its description and uses see: IMF's Data Quality Assessment Framework (Weisman, Balyozov and Venter 2010). Laliberté, Grünewald and Probst (2004) made a detailed analysis of *differences* between the Eurostat and the IMF *frameworks*. A *comparison* of all three frameworks, which includes the OECD's approach, can be found in United Nations (2009a, 123ff.).

Other organizations started to build their own *quality frameworks* around the existing ones, e.g. FAO (2004) and (2005). Also some *countries* developed their own *quality guidelines*, thus providing a producers perspective on *quality frameworks*, e.g. Statistics Canada (1998 and 2002) and Statistics Finland (2007).

Valuable *guidelines* like Eurostat's (2009a) "ESS Handbook for Quality Reports" provide explanations, examples and discussions building upon the 2003 quality standards. The "Handbook on Data Quality Assessment Methods and Tools" (DatQAM) released by Eurostat, provides guidance for National Statistical Institutes (NSIs) to increase data quality (Ehling and Körner 2007).

The "Data and Metadata Reporting and Presentation Handbook" was prepared by OECD (2007a). It is a reference manual compiled from the international frameworks that offers practical, substantial, and comprehensive guidance and recommendations.

The OECD developed a comprehensive *glossary of statistical terms*, containing more than 6000 terms (cf. OECD 2007b). The UNECE compiled a "Glossary of Terms on Statistical Data Editing" (UNECE 2000) which is also made available on "K-Base"[6], the knowledge base of UNECE Data Editing Group, and by the OECD glossary.

[5] The latest version was released January 2012 (Version 2011/1) cf. OECD (2012a), STD/QFS(2011)1.

[6] Online at: http://www1.unece.org/stat/platform/display/kbase/Glossary.

Disciplines *other than statistics* usually consider a different set of *quality dimensions*, or the *dimensions* are defined differently. Batini and Scannapieco (1998) provide a substantial introduction to issues of *data quality* from the perspective of *computer science*. Scannapieco and Catarci (2002) compare six proposals of data *quality* dimensions, each with multiple elements.

Naumann (2002) made a contribution to *Information Quality (IQ)* for data available in the *World Wide Web* and information retrieval**.** He carefully developed the definitions for the quality dimensions. The discourse on data *quality* in such systems was continued by Scannapieco and Pernici (2003).

Karr, Sanil and Banks (2002) consider a similar range of DQ dimensions as the international statistical quality frameworks, yet they relate DQ to *database systems developments* and *Total Quality Management* (TQM).

In *economic statistics*, *DQ assessments* of the *availability* and *timeliness* of *high frequency data* were carried out on a global level by the UNSD (2009b). Three international seminars covered *data quality* aspects of *high frequency data, short term statistics* and other *economic statistics***,** including their *timeliness*, were held between 2009 and 2010 under the auspices of the Statistics Canada, Statistics Netherlands, Russian Federal State Statistics Service (Rosstat), United Nations, and Eurostat (cf. UNSD 2009c and 2010)[7].

Assessments of *availability* of *balance sheets* and *accumulation accounts* compiled within the framework of the SNA, of the G20 and other advanced economies[8] were done by IMF in 2010. The results are presented in two papers (IMF 2011a, 2011b).

A *comprehensive assessment* of the *availability* of Annual National Accounts data at the UN, based on data published for its Member States in 2006, has been published in the SNA News and Notes, Issue 25/26, May 2008, pp. 6-9. This assessment includes the "conceptual implementation" of the 1993 SNA and the "timeliness of reporting".

---

[7] For further documents by NSOs see: http://www.unstats.un.org/unsd/nationalaccount/workshops.

[8] Other advanced economies (advanced economies in the WEO classification not included in the G20): Austria, Belgium, Cyprus, Czech Republic, Denmark, Finland, Greece, Hong Kong SAR, Iceland, Ireland, Israel, Luxembourg, Malta, Netherlands, New Zealand, Norway, Portugal, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Taiwan Province of China.

The *transmission schedule* of the European Union (cf. European Parliament and Council of the European Union 2007) for NA data indicates the expected *transmission time points* of ANA and QNA data for different data objects (i.e. different *SNA concepts* like main aggregates of GDP or financial accounts by institutional sectors). Thus this release schedule indicates the "*expected timeliness*" of NA data production by countries regarding the specified NA data sets.

The "Status Report of Information Requirements in the European Monetary Union" provides the *legal dates* and *target dates* of data *transmission* regarding the set of QNA data (four NA indicators) included within the set of Principal European Economic Indicators (PEEI). Additionally the *timeliness* of publication by Eurostat of the European aggregate for these four QNA indicators is indicated. Further indicated are the individual *transmission time points* of EU Member States and EU Area Member States to Eurostat (cf. EFC and ECOFIN Council 2010).

Ahmad, Bournot, and Koechlin (2004) made a comparative analysis of seven OECD countries (the G7 countries) regarding the reliability of the first released GDP growth rates. Also included in their study is an overview regarding the timeliness of the country's growth rate publication at the OECD via the monthly Main Economic Indicators (MEI) publication. Considered is the timeliness of data publication at the IO (i.e. their "filtered" [Di Fonzo 2005a] timeliness of publication at the international level) for each individual quarter and for the annual data. Di Fonzo (2005b) assessed the average timeliness of publication of the first released quarterly GDP growth rate by the OECD for 18 OECD Member States. For the same set of 18 countries this assessment was repeated by McKenzie (2006a). Nearly at the same time Tosetto and Lequiller (2006) compiled a release calendar for these OECD countries for the publication their first quarterly GDP growth rate by 1st to 4th quarter in the past 10 years of MEI publications.

The importance of *improving* the *availability* and *timeliness* of NA data is recurrently pointed out by experts and governments (for example: FSB and IMF 2009, IMF 2012).

The significance of compiling *revision histories* of vintages of data at different points of maturity (so called "real-time datasets") was re-iterated after the 2008 financial crises by contributions on their importance and applications for example Croushore (2011) and Fernandez, Koening, and Nikolsko-Rzhevskyy (2011). Further, Eurostat

(2009b) dedicated a Selected Readings issue to *revisions analysis*, featuring the summaries of 64 papers on this topic.

The OECD's "MEI Original Release Data and Revisions Database"[9, 10] was established only in 2006. A set of two sequent *revision spreadsheets*[11] to analyze *revision data* in a two-step process, which we use to analyze the revisions data published via the OECD database, was developed and explained by Di Fonzo (2005a).

How to *interpret* the results of *revision analysis* and use this information to improve data at the national level was illustrated (generalized or by case studies) by Mazzi (2001), McKenzie and Gamba (2008), and Hoven (2008).

Seven *studies* were prepared that *analyzed* reliability of first released GDP estimates of OECD countries utilizing the tools (pre-programmed Excel spreadsheets for revision analysis) by Di Fonzo (2005a) together with the data published by the OECD MEI monthly editions (cf. Di Fonzo 2005b; Tosetto and Lequiller 2006, McKenzie 2006a and 2006b, McKenzie and Adam 2007, Tosetto and Brackfield 2009, and Neumayr 2010). McKenzie (2006c) also carried out a study of short term statistics including the Index of Industrial Production (IIP) and Retail trade volume. ECB made a study similar to McKenzie (2006c) utilizing own revision data (Branchi et al. 2007).

Some countries inform about their revision policies and the results of *revision studies* e.g. Australia, Canada, Germany and the US. Results of nationally produced revision studies are typically not comparable since different revision measures or assessment methodologies are employed.

International and regional organizations have set up initiatives and training programs to *strengthen* the NA *capacity* of countries. These provide the framework to assist countries in developing statistical and institutional *capacities* and to produce NA data according to the methodology of 2008 SNA (for further details see: United Nations 2010b and 2011).

---

[9] Cf. OECD (2012b), Main Economic Indicators (MEI) Revisions Database, 2007.
Online at: http://stats.oecd.org/Index.aspx?querytype=view&queryname=206#.

[10] Databases used for NA data are also referred to as data warehouses (cf. Jarke, Lenzerini, Vassiliou and Vassiliadis 1999, 2-4). Examples of such a data warehouses are OECD.Stat and UNdata.

[11] Available as download for monthly and quarterly data via
http://www.oecd.org/std/automatedprogramstoperformrevisionsanalysis.htm, last accessed 28-10-2013.

II. Definitions of Data Quality Dimensions and Measures

This part, introduces the existing *quality assessment frameworks* of *IMF*, *Eurostat* and *OECD*. The chief contributions are the *definitions* of the *quality dimensions* and further of *quality measures* used for assessing the country data regarding the DQ dimensions *timeliness*, *availability*, and *reliability*. *Quality frameworks* usually define *DQ dimensions* and *measures* without providing formulas. Own definitions are developed where existing ones are missing and existing definitions are revised to improve the wordings and formulas. Further, each chapter points to different definitions and interpretations of the quality *dimension* and *measure* in existing *DQ frameworks* and in different disciplines (i.e. areas of application).

1 Selected Data Quality Frameworks in Official Statistics

Three *DQ frameworks* are typically referred to; that of Eurostat, IMF, and OECD.

The *DQ framework* of **Eurostat**, defined and described in Eurostat (2003a, 2003b, and 2003c), consists of six *dimensions* (cf. Table 1, p. 14). It focuses on quantitative (cf. DEF. 2, p. 16) information about the statistics and is used to assess the ability of a Member State to comply with the quality guidelines of Eurostat. Countries can also use the Eurostat *framework* for self assessments.

The *Data Quality Assessment Framework (DQAF)* of the **IMF** (2003a), as well as the corresponding specialized *DQAF* version for National Accounts (IMF 2003b), consist of five *dimensions*. Zieschang (2005) explains, how the *DQAF* relates to NA and other economic statistics as well as to the different IMF data dissemination standards. The *DQAF* for NA is used for country assessments in the data module of the IMF Reports on the Observant of Standards and Codes (ROSC). National institutions (like NSOs and central banks) can use the IMF *framework* for self assessment. The IMF's and Eurostat's data quality *frameworks* are complementary, meaning they inform about different aspects of the DQ dimensions. The *IMF framework* is concerned with qualitative (cf. DEF. 1, p. 16) information regarding the DQ. Considered are the governance of the statistical system, the application of quality assurance activities, the employed country practices against best practices, and the country's adherence to internationally agreed methodologies at the stage of data collection, -production, and -dissemination. The *Eurostat framework*, on the other hand, yields quantitative information by computing measures from the statistical data (e.g. the timeliness of data released by a Member State). Where the *DQ dimensions* of the Eurostat and IMF *frameworks* are similar to each other, for example in case of "accuracy" (Eurostat) and of "accuracy and reliability" (IMF) (cf. Table 1), the used measures target the quantitative aspect (Eurostat) respectively the qualitative aspect (IMF). For example Eurostat targets quantitative results like the mean revision, and the IMF qualitative information like whether or not revision studies are performed to assess the quality of statistical outputs (cf. DEF. 3, p. 17).

The following reference table indicates the *DQ dimensions* of the three *DQ frameworks*. The *DQ dimensions* are in alphabetic order to facilitate easy comparison between the three *DQ frameworks*. Note that in some cases a single *DQ dimension* is constructed from two components (e.g. "Accessibility and Clarity" together form one *DQ dimension* in the Eurostat *framework*).

Table 1 Quality Dimensions in the Quality Frameworks of IMF, Eurostat, and OECD

		Eurostat	**IMF**	**OECD**
Data Quality Dimension	1	Accessibility and Clarity	Accessibility	Accessibility
	2	Accuracy	Accuracy and Reliability	Accuracy
	3	Coherence		Coherence
	4	Comparability		
	5			Credibility
	6		Integrity	
	7			Interpretability
	8		Methodological Soundness	
	9	Relevance		Relevance
	10		Serviceability	
	11	Timeliness and Punctuality		Timeliness

Source: Own illustration compiled from IMF (2003a), Eurostat (2003a), and OECD (2003)
Tables are produced using Microsoft Excel and Word

Because of the differences regarding the application at country level or organizational level, and the differences regarding the type of DQ assessment (i.e. qualitative or quantitative DQ aspects), a direct comparison of the *DQ dimensions* considered by the three *DQ frameworks* is not possible. According to Laliberté, Grünewald, and Probst (2004), the major difference between the Eurostat and IMF *framework* is that **Eurostat** is particularly assessing the statistical data as such, while the **IMF** *framework* is more universal, considering the quality of the governance of statistical systems, the processes of data production, and data dissemination process.

The **OECD** *framework* (cf. OECD 2003[12]) consists of seven *dimensions*. Covered are both types of DQ aspects. OECD primarily uses the *DQ framework* to assess the quality of internal data sets and of data sets being disseminated by the OECD to the public. It is tailored to OECD activities, yet it is also suited for use by IOs in general.

The major difference of the *OECD framework*, compared to Eurostat and IMF is that it applies to the IO itself, i.e. the OECD is the subject of assessment. Meaning, at the OECD, the *DQ framework* is used to assess own statistical activities as a disseminator of statistics. Contrary, Eurostat and IMF use the DQ *frameworks* to assess the data produced by the countries. Besides the difference in application, the OECD's DQ dimensions are similar to those of the two other *DQ frameworks* while covering qualitative as well as quantitative measures (cf. OECD 2003). For example OECD covers qualitative information regarding the types of user groups interested in their data and the way they request data, and quantitative information about the number of sold copies of publications within the last 12 month.

Nationally developed *DQ frameworks* cover similar dimensions and quality measures as the international ones, or are building on them. Vice versa, existing national *DQ frameworks*, practices, and principles were a base for the internationally developed *DQ frameworks* (cf. OECD 2003, 5; Weisman, Balyozov, and Venter 2010, 2).

Figure 1 illustrates the *frameworks* of Eurostat, IMF, and OECD in reference to the *type* of assessment they target (qualitative or quantitative DQ measures) and the *subject* of application i.e. for IOs or assessments at the country level.

[12] For the latest version (QFS 2011/1) see: OECD (2012a).

Type \ *Subject*	**International Organization**	**Country**
Qualitative assessment	OECD	IMF
Quantitative assessment	OECD	Eurostat

Figure 1 Quality Frameworks of Eurostat, IMF, and OECD by Type of Assessment (Qualitative or Quantitative) and Subject of Application (IO or Country)

Source: Own illustration

Figures are produced using Microsoft Excel and Word

DEF. 1: Qualitative Quality Measures
Measures that provide meta-information about factors that influence the DQ of the actual statistics at the stage of data collection, production, and dissemination. For example: Qualitative quality measures “consist of results from user satisfaction surveys, descriptions of methods used for calculating standard errors and sources of error” (see ONS 2013), or information whether or not “statistical policies and practices are guided by professional principles”, e.g. if “the terms and conditions under which statistics are collected, processed, and disseminated are available to the public” (see: IMF 2003a, 1).

Source: Own definition, with examples by ONS (2013) and IMF (2003a, 1)

DEF. 2: Quantitative Quality Measures
Measures computed from the actual statistics that yield a quantitative value to inform about the DQ of the statistics. “Examples of quantitative quality measures are: standard error, imputation rates, non-response rates, editing rate, proxy response rate, time lag between data collection and data release” (see: ONS 2013).

Source: Own definition, with example by ONS (2013)

DEF. 3: Statistical Outputs
Statistical outputs are the statistics derived from (primary/secondary) source data (including censuses, surveys, and administrative sources) and intermediate results (i.e. data prior to the balancing process). These data are the end-product of the computation process. Example: the published national accounts indicators are statistical outputs, derived from e.g. sales data from business surveys, surveys of capital expenditure by businesses, and administrative sources such as central and local government reports. See also: DEF. 4: Source Data.

Source: Own definition

DEF. 4: Source Data
Underlying statistics and statistical outputs (survey results, administrative source data, price and production indices, etc.) used to produce NA output statistics.

Source: own definition

2 Data Quality Dimensions

The three most important DQ *dimensions* for NA statistics at the international level (always referring to the statistical outputs) are defined in this chapter. The DQ *measures* used to assess these DQ *dimensions* are defined in the next chapter.

2.1 Timeliness

DEF. 5: Timeliness

Timeliness (T) of a data object $x \in X$ (the set of data objects) published for a country $c \in C$, is the length of the time interval given by the time point of data publication (t_P) that makes it available to the public and the end time point of the reference period (t_E) of the event or phenomenon it describes. Formally expressed:

$$X := \text{set of data objects}$$
$$C := \text{set of countries}$$

$$\mathrm{T} : X \times C \rightarrow \boldsymbol{R}^{+}$$
$$(\mathrm{x,c}) \in X \times C \mapsto \mathrm{T(x,c)} = \mathrm{t_P(x,c)} - \mathrm{t_E(x,c)}, \qquad \text{(II.1)}$$

where

T(x,c): the timeliness of a singular data object $x \in X$, published for a country $c \in C$,

$t_P(x,c)$: the time point of data publication of $x \in X$, for country $c \in C$,

$t_E(x,c)$: the end time point of the reference period of $x \in X$, for $c \in C$.

If a published data object $x \in X$ consists of more than one data item, the timeliness T of the data object x is measured by the most recent data item within x, i.e. by the data item referring to the most recent reference period $t_E(x,c)$ within the data object $x \in X$.

Source: own definition based on Eurostat (2003a)

The above given definition is based on the well established definition of *timeliness* by Eurostat (2003a, 3), which reads: “Timeliness of information reflects the length of time between its availability and the event or phenomenon it describes”. The Eurostat definition lacks the provision of a formula to calculate the defined measure. Further, the precision of the Eurostat definition is improved with our revised definition by relating *timeliness* to the time points of data “publication” (making the data publicly available) and the “end time point” of the described event or phenomenon.

The ‘*publication time point*’ and the ‘*end time point of the event or phenomenon*’ given by the time points t_P and t_E can for instance refer to the release of the results of an election, or to the publication of the annual report regarding the last business year. Thus the event or phenomenon itself may refer to a point in time (an Election Day) or a time period (a calendar year). In any case, the described event or phenomenon is specified regarding the time point or time period, e.g. GDP for the reference year 2010, 2011, etc.

In the *DQ frameworks* for statistics of Eurostat and OECD, *timeliness* refers to the time length that has passed, following the described phenomenon, until the release (publication) of the corresponding information or data. Eurostat considers “Timeliness and Punctuality” as one *DQ dimension*. *Punctuality* is assessed by Eurostat and OECD. In both cases it relates to the existence of a previously announced time point at which data should be released respectively reported. It is assessed whether or not the announced time point or time frame is exceeded by the actual day of data release or reporting (cf. Eurostat 2003a, 3 and OECD 2003, 9).

Also included in the *DQ dimension* “Timeliness and Punctuality” of the Eurostat framework is the *time-related DQ measure* “data freshness” (cf. Eurostat 2003a, 12). The IMF framework includes the *DQ measure* “periodicity” in addition to “*timeliness*” in order to inform about the *DQ dimension* “serviceability” (cf. Laliberté, Grünewald, and Probst 2004, 10) (for definitions see the quoted references). Further definitions of *timeliness* and of other *time-related quality dimensions* used in computer science can be found in Batini and Scannapieco (1998, 41).

For statistics that are subsequently revised (i.e. re-calculated) until final data are computed (like in case of NA), the *timeliness* can be measured for different releases e.g. for the first release of preliminary data (initial estimate) (cf. DEF. 6 Estimate) or for the release of final data (final estimate).

We consider the *timeliness* of a data object $x \in X$. The data object x can consist of one or more data items which constitute the data object.

Figure 3, illustrates the relation between “Data Object” and “Data Item” following Unified Modeling Language (UML) 2.1 notation. This relation is a composition (composite aggregation) where “Data Object” is the composite object (i.e. aggregate or whole) and “Data Item” refers to the composed objects (i.e. parts). For example, the data object “GDP by expenditures” (i.e. the SNA table for GDP by expenditure components) consist of multiple data items. These are the GDP indicator and the NA indicators needed to calculate GDP. On the other hand, the data object “Exports of goods” (i.e. without disaggregation) only consists of this single data item.

We are concerned with assessing the *timeliness* of a country’s ANA data at the time point when these data are made publicly available for the first time at the UN level through the NA Yearbook (YB). Thus our considered data object “ANA” consists of the set of ANA indicators that are reported by an individual country to the UN for publication. (For the DQ measures used to assess the *timeliness* dimension refer to part II, chapter 3.3, p. 117ff.)

DEF. 6: Estimate
“In the strict sense, an estimate is the particular value yielded by an estimator in a given set of circumstances” (see: Dodge 2006, 136). In NA, estimates refer to the statistical outputs at the end of the data computation process. (The first release of NA data is also referred to as the initial estimate, subsequent releases as subsequent estimates, and the final data as the final estimate.)

Source: Marriott (1990, 223) and own definition

Sample Application 1: Timeliness

Germany, QNA

Reference Period	Time Point of Publication
4. quarter 2012	February 2013 (released 22-02-2013)*
* Cf. Statistisches Bundesamt (2013).	

Germany (DEU)[13] published the detailed QNA data of the 4^{th} quarter of 2012 (ending 31. December 2012) in late February 2013. Thus the timeliness of these QNA data at time point of their first publication is 2 month. Or expressed in days this is 53 days.

Referring to the formula $T(x,c) = t_P(x,c) - t_E(x,c)$ (c.f. DEF. 5), without specification of the granularity of the time period, this equates for the QNA data of Germany to

$$T(QNA, DEU) = t_P(QNA, DEU) - t_E(QNA, DEU)$$

$$= \text{February, 2013} - \text{December, 2012} = 2 \text{ month,}$$

Respectively if expressed in days we get as formula

$$T(QNA, DEU) = 22\text{-}02\text{-}2013 - 31\text{-}12\text{-}2012 = 53 \text{ days.}$$

[13] Country abbreviation corresponding to ISO ALPHA-3 code. “The Standard Country or Area Codes for Statistical Use (Series M, No.49/Rev 4), generally referred to as ‘M49’, presents a list of names of countries or areas used as a common reference for statistical processing purposes by the United Nations Statistics Division” (UNSD 2006, 1). For updates see “Countries or areas, codes and abbreviations” (UNSD 2013a) available online at: http://unstats.un.org/unsd/methods/m49/m49alpha.htm (status: 11-02-2013), last accessed 11-06-2013.

2.2 Availability

DEF. 7: Availability
Availability (A(x,c)) refers to the existence of data object x $\in$ *X* - for example a single NA indicator, NA table, and the like - for a country c $\in$ *C* (evaluated based on published data for c). The data object in question can either be available or not available for an evaluated country c. (The latter case, "not available", indicates a data gap.) Formally expressed:

$$X := \text{set of data objects}$$
$$C := \text{set of countries}$$
$$\mathrm{A} : X \times C \to \{0,1\} \quad \text{with}$$
$$(\mathrm{x},\mathrm{c}) \in X \times C \mapsto \mathrm{A}(\mathrm{x},\mathrm{c}) = \begin{cases} 1 \text{ if data object x} \in X \text{ is available for c} \in C, \\ 0 \text{ else.} \end{cases} \qquad \text{(II.2)}$$

Note: the determination of availability might be subject to an availability condition regarding the data object - meaning the precision of the description of the data object for which the availability is assessed, i.e. attributes that define the data object.

Source: own definition

Examples for a condition regarding the availability of a NAQ table are the availability of some variables for the NAQ table, or a specific variable within the NAQ table, or the availability of all variables for the NAQ table.

"Availability" is not part of the DQ frameworks by Eurostat, IMF, and OECD. *Availability* has to do with the existence of data (cf. Van Nederpelt 2009, 57). Conversely Naumann (2002, 33), where it is used synonymously for "accessibility, technical reliability, retrievability, performability" and thus relates to reliable and dependable access to data, or the possibility to retrieve Web data that are in principle *available* in a specific data source [e.g. database or Web page]). In our study, *"availability"* relates to official country data (cf. DEF. 9) published through IOs.

For the DQ dimension *"Relevance"* "the rate of non-available statistics is an indicator of the degree to which completeness has been achieved concerning a specific product"

(Eurostat 2003b, 4). The meaning of *completeness* (cf. DEF. 10) is generally interpreted similarly, but varies according to the subject and context. For an overview of definitions of *completeness* see Batini and Scannapieco (1998, 41). Eurostat is concerned with the number of indicators or *completed* (answered) queries within a questionnaire or survey etc., where the set of requested indicators is determined by the questionnaire or survey. Contrary to Eurostat, the *rate* of non-available statistics is not relevant for the *availability* assessment presented in this study, where *"Availability"* is a DQ dimension targeting the existence and non-existence of data objects.

Availability of NA data is assessed for defined sets of NA data which should be produced by each country. These NA data sets, which provide information about different economic concepts, are either available or not. Hence, *availability* refers to the data that are existent in a country, i.e. produced by the country.

The *availability* assessment is based upon the set of official country data published by the UN, thus it is subject to the data that are previously reported by the country to the UN. For the data *availability* assessments by the UN it was internationally agreed that "basing the assessment on data actually available at the international level was preferred as more objective" (cf. United Nations 1999, 3) than other ways to assess the data production of countries. Consequently official NA data not being published for a country by the UN are considered not being produced by the country, assuming that a country would report the requested data if they were produced. Equally, a country for which no NA data are published by the UN is considered not producing NA data. Thus, in this study *non-availability* of an assessed data object (i.e. a NA indicator etc.) implies a "*data gap*" (cf. DEF. 8).

DEF. 8: Data Gap (Associated with Availability)
The non-existence of at least one data object $x \in R$ of the subset $R \subseteq X$ of relevant data objects indicates a data gap as an event for $c \in C$. A data gap occurs if $$\exists\, x \in R : A_c(x) = A(x,c) = 0 \qquad \text{II.3}$$ i.e. we have $A_c(x) = 0$, for at least one $x \in R \subseteq X$.

Source: own definition

Note: If for example: R = {GDP} is a singular set, i.e. a data item ($x \in R \Leftrightarrow x$ = GDP), the data gap for $c \in C$ occurs, if $A_c(GDP) = A(GDP,c) = 0$.

The purpose of the *availability* assessments in line with the UN mandate is "to identify countries (or country groups) that need the particular attention of the international community" (cf. United Nations 2000, 8). The current UN assessments, as well as the assessments in this study present the results for the *available* data. This provides important information about the depth of *produced* data and the *statistical capacities* of countries, and further allows making conclusions regarding the *data gaps*. It is intuitive that the *data gap*, as defined above, is most interesting for the aforementioned purpose of pointing out those countries that need particular attention.

However, *non-availability* of data at the international level does not necessarily imply *non-availability* (i.e. a *data gap*) at the national level. For instance, *non-availability* at the international level might result from underreporting of existing data, a different NA compilation system in the country that yields different NA indicators than those requested by the IO, exclusion of produced data from the disseminated data set (by the IO, or already by the NSO) due to significant quality issues or restrictions due to confidentiality of information and privacy protection (e.g. anonymity of micro data).

To assess the *availability* of a fixed data object $x \in X$ for the set C of countries, we consider (based on DEF. 7) specially the function

$$A_x : C \to \{0,1\} \tag{II.4}$$

$$c \mapsto A_x(c) := A(x,c) = \begin{cases} 1 \text{ if the data object } x \in X \text{ is available for } c \in C\text{, and} \\ 0 \text{ else.} \end{cases}$$

For example, if the data object in question is GDP, then

$$A_{GDP} = A(GDP, .\,) : C \to \{0,1\} \tag{II.5}$$

$$c \mapsto A_{GDP}(c) := A(GDP,c) = \begin{cases} 1 \text{ if GDP is available for } c \in C\text{, and} \\ 0 \text{ else.} \end{cases}$$

Likewise for a fixed country $c \in C$ and a set X of data objects, we consider

$$A_c : X \rightarrow \{0,1\} \tag{II.6}$$

$$x \mapsto A_c(x) := A(x,c) = \begin{cases} 1 \text{ if the country } c \in C \text{ has data object } x \in X, \\ 0 \text{ else,} \end{cases}$$

to describe which data in c are *available* or not.

Then the set of countries $c \in C$ for which a fixed data object $x \in X$ is available is given by the preimage set

$$A_x^{-1}(1) := \{c \in C \mid A_x(c) = A(x,c) = 1\}. \tag{II.7}$$

The set of data objects $x \in X$, that are *available* for a fixed country $c \in C$, is represented as the preimage set

$$A_c^{-1}(1) := \{x \in X \mid A_c(x) = A(x,c) = 1\}. \tag{II.8}$$

For the set C of considered countries the share (h) of countries for which the data object $x \in X$ is *available* is delineated by

$$h_x := \frac{|A_x^{-1}(1)|}{|C|} \tag{II.9}$$

Likewise, the set of countries $c \in C$ where a *data gap* occurred for the fixed data object $x \in R \subseteq X$ are represented as members of the preimage

$$A_x^{-1}(0) := \{ c \in C \mid Ax(c) = A(x,c) = 0\} \tag{II.10}$$

and the share (h) of countries $c \in C$ where the data object $x \in R \subseteq X$ of the relevant set of objects is *missing* is delineated by

$$h_{xna} := \frac{|A_x^{-1}(0)|}{|C|} = 1 - \frac{|A_x^{-1}(1)|}{|C|} \tag{II.11}$$

The assessment used to inform about the dimension *Availability* is of a type that provides a qualitative measure for the DQ of the country's NA data. Assessed is the *availability* of data sets of particular interest, specifically the NA tables of the 1993 and 2008 MRDS, and the achievement of the milestone levels 1 and 2. (For the DQ measures used to assess the *availability* dimension of NA refer to chapter 3.2, p. 52ff.)

DEF. 9: Official Country Data
Statistics that are published by the national government agency mandated to disseminate these data as "official data".

Source: own definition

National statistics may be *produced* and *disseminated* by other (government) agencies that are not mandated to *disseminate* the official country data. For instance, NA data may be *produced* by the NSO (typically also the agency mandated to *disseminate* the official country data), the Central Bank (CB), a government department or ministry, or an institute for economic research. Sometimes different agencies *produce* the different data sets of the NA, yet only one agency is mandated to *disseminate* the NA data set as official data. In rare cases the data set *disseminated* as the official country data are even *produced* as part of a technical assistance project by an IO or an external consultant, e.g. for Timor-Leste (cf. United Nations 2010a).

DEF. 10: Completeness

Extent (given by the share) of the present set (i.e. fraction, part) with respect to the total set (i.e. whole). Completeness (C) is calculated as:

$$\mathrm{C} = \frac{|P|}{|\Omega|}, \text{ for any subset } P \subseteq \Omega \qquad (0 \leq \mathrm{C} \leq 1), \tag{II.12}$$

where

$|P|$: cardinality of the subset $P \subseteq \Omega$, and

$|\Omega|$: cardinality of the total set Ω of considered objects (i.e. whole).

Source: own definition

Sample Application 2: Availability

Specific NA Indicator

Suriname

Available NAQ Tables*	Missing NAQ Tables*
1.1 GDP by expenditure, current prices 1.3 Relations among income, saving and net lending aggregates 2.1 GDP by expenditure, current prices 2.2. GDP by expenditure, constant prices Specifically: 1.3 is not available until NA indicator "net lending".	1.2 GDP by expenditure, constant prices 2.3 Value added components by industry 4.1 - 4.9 Institutional sector accounts
* For the NAQ tables numbers and corresponding NAQ table names of these NA data sets refer to Table 3 Minimum Requirement Data Sets for the 1993 SNA and 2008 SNA on p. 56.	

Thus regarding the availability of a *specific NA indicator* for Suriname (SUR), we see that for example the GDP indicator is available in both current and constant prices. It is a component of the NAQ tables 1.1, 1.3, and 2.1 in case of current price data GDP, and the constant price data GDP is a component of NAQ table 2.2. On the other hand, the indicator net lending is not available, though NAQ table 1.3 is available, yet not until the net lending indicator.

Referring to the formulas (cf. DEF. 7 and DEF. 8), for the mentioned NA indicators, this equates to

$A(GDP_{current},SUR) = 1,$

$A(GDP_{constant},SUR) = 1,$

$A(\text{net lending},SUR) = 0$

Thus, the last equation is indicating a data gap regarding the net lending indicator.

For the availability of an NAQ table instead of a specific indicator this equates to e.g.

$A(1.3,SUR) = 1,$

$A(1.2,SUR) = 0.$

(continued on next page)

The preimage set of data objects, for example of NAQ tables ($T_{x.y}$), that are available for Suriname consists of NAQ tables 1.1, 1.3, 2.1, and 2.2.

Referring to the formula (II.8) this equates to

$$A_{SUR}^{-1}(1) = \{ T_{x.y} \in \mathcal{T}_{x.y} \,|= A(T_{x.y},SUR) = 1\} = \{ T_{1.1}, T_{1.3}, T_{2.1}, T_{2.2} \}.$$

Fixed Data Object

Regarding the set of countries that have a *fixed data object* available we now we look at the UN Member States in the region Northern Africa and the data object net lending.

Northern Africa Countries	Availability of Net Lending
Algeria	Yes
Egypt	Yes
Libyan Arab Jamahiriya	No
Morocco	Yes
Tunisia	Yes

Thus Libya is the only country (c) in that region that is not a member of the preimage set of countries that have the indicator “net lending” available.

Referring to the formula (II.7) this equates to

$$A_{net\_lending}^{-1}(1) := \{c \in \mathit{Northern\ Africa} \,|= A(net\_lending,c) = 1\}$$

$$= \{ \text{Algeria, Egypt, Morocco, Tunisia} \}$$

This corresponds to a share of 0,8 (i.e. 80 percent) of the countries that have net lending available in the Northern Africa region.

Referring to the formulas this is given by the cardinality of the members of the preimage set of countries with the data object “net lending” of the capital account over the cardinality of the countries in the region (II.9) which equates to

$$h_{net\_lending} := \frac{\left|A_{net\_lending}^{-1}(1)\right|}{\left|Northern\_Africa\right|} = \frac{4}{5} = 0{,}8$$

2.3 Reliability

DEF. 11: Reliability (of National Accounts)
The absolute change (Δ_r^t) of the initial (first) released numerical value (θ_0) to the subsequent released (more mature) numerical value(s) (θ_r, with r = 1, 2, … 100[14]) of the same reference period (t) for vintages 0,r; the absolute change being marked by the closeness of the values of vintage 0 and r, i.e. the insignificance of the absolute change[15]. Reliability is calculated as: $$\Delta_r^t = \Delta_{0,r}^t := \lvert\theta_r^t - \theta_0^t\rvert, \qquad \text{(II.13)}$$ where θ_0^t and θ_r^t : the values of the indicator θ (e.g. GDP, or GDP growth rate) at vintages 0 respectively r of the same reference period (t), i.e. at time point of the initial release (first vintage) and any later vintage (i.e. at a later time point, for example a revised release five month later). If $\lvert\theta_r^t - \theta_0^t\rvert = 0$, than θ_0^t is a perfectly reliable representative for θ_r^t . If $\theta_{0(t)}$ is a preliminary value, yet $\theta_{0(t)}$ is not subsequently revised to incorporate new information, reliability cannot be assessed by $\lvert\theta_r^t - \theta_0^t\rvert$. Since the different vintages of θ are of the same reference period, the index t can be dropped in the formula. Thus we have $\Delta_r = \Delta_{0,r} = \lvert\theta_r - \theta_0\rvert$. [The reliability of any (earlier) released numerical value (θ_i, with i = 0, 1, …, 99) to a subsequent released numerical value (θ_r) can be calculated as $\Delta_{i,r} := \lvert\theta_r - \theta_i\rvert$, with i < r.] See also: DEF. 13: Vintage, DEF. 14: Maturity, DEF. 4: Source Data, DEF. 16: Final Data, DEF. 17: Latest Vintage.

Source: own definition

[14] Setting r = 100 equates for example to four releases of QNA data per year over 25 years and one quarter with continuous revisions of all previously releases (back data) with every release of new QNA data. Thus, for example in case of seasonally adjusted data, this means 100 revised vintages of θ.

[15] The meaning of reliability depends on the precise definition of Δ_r e.g. the fixing of an upper boundary.

The above definition for *reliability* of NA is based on the regularly quoted definition by Carson and Laliberté (2002, 4): "The closeness of the initial estimated value(s) to the subsequent estimated values." The authors add: "Assessing reliability involves comparing estimates over time".

Typically, *reliability* is evaluated based on an average measure computed from multiple observations (at least 20 [cf. Ahmad, Bournot, and Koechlin 2004]) regarding the difference between the first and a subsequent released value. The later released value refers to a data vintage (cf. DEF. 13) of higher maturity (cf. DEF. 14). For the *reliability* dimension we consider the absolute change, whereas the "*Revision*" and "*Mean Revision*" (cf. DEF. 15, p. 37 and DEF. 18, p. 45) examine the change taking into consideration the sign, i.e. whether it is positive or negative. (For the DQ measures used to assess the *reliability* of NA data refer to chapter 3.1.2, p. 44ff.)

DEF. 12: Accuracy (of National Accounts)
The absolute deviation (ε) of the observed numerical value (θ) to the fully correct (true) numerical value (θ^*) for the evaluated indicator of the same reference period (t)[16]; with the deviation being marked by the closeness of the observed to the fully correct value, i.e. the insignificance of the deviation[17]. Accuracy is calculated as: $$\varepsilon^t = \lvert\theta^t - \theta^{t*}\rvert \qquad \text{(II.14)}$$ θ^{t*} is unknown (a notional value) in national accounts, therefore ε is a theoretical value that cannot be assessed in reality. Theory: if $\lvert\theta^t - \theta^{t*}\rvert = 0$, than θ^t is a perfectly accurate representative for θ^{t*}. Practice: Since the fully correct, true, value (θ^{t*}) is unknown for the SNA indicator (e.g. GDP), the accuracy cannot be assessed by $\lvert\theta^t - \theta^{t*}\rvert$. [For different observations of θ^t we define $\varepsilon_r^t = \lvert\theta_r^t - \theta^{t*}\rvert$ with r = 0, …, 100.]

Source: own definition

[16] The index t can be dropped, since the different vintages of θ are of the same reverence period.

[17] The meaning of accuracy depends on the precise definition of ε for example, a fixed upper boundary.

The above definition for *accuracy* is a more formal and detailed formulation of the usually quoted definition for *accuracy* of NA: "The closeness of the estimated value to the (unknown) true value that the statistic is intended to measure." by Carson and Laliberté (2002, 4). The authors continue this definition stating that: "In practical terms, there is no single aggregate or overall measure of accuracy."

In other words, *accuracy* can be defined as the proximity of an estimated value to the true (i.e. exact or correct) value that was intended to be measured, yet in NA this value is usually an unknown, notional value. And, on the other hand, *reliability* can be defined as the proximity of a preliminary estimate (typically the first published value) to a subsequent revised estimate of the same reference period [compare to definitions in Aspden (1990), ABS (2003), and Ahmad, Bournot, and Koechlin (2004)].

DEF. 13: Vintage
Vintage refers to the 'harvest' of data at a specific time point. It is "a given release of data at a given point in time" (see: Brown, Buccellato, Chamberlin, Dey-Chowdhury, and Youll 2009, 44).

Source: own combined with Brown, Buccellato, Chamberlin, Dey-Chowdhury, and Youll (2009, 44)

Carson and Laliberté (2002) define *accuracy* and *reliability* for NA and BOP statistics in context of the dimension *"*Accuracy and Reliability*"* of the IMF's DQAF. In the Eurostat and OECD frameworks, the analysis of revisions to initial released data, i.e. the assessment of the *"reliability"* of data is part of the DQ dimension *"*Accuracy*"*. "Accuracy in the general statistical sense denotes the closeness of computations or estimates to the exact or true values […]" (see: Dodge 2006, 4). Corresponding to "Accuracy", the definition of "Error" is "the difference between an occurring value and its 'true' or 'expected' value" (see: Marriott 1990, 69); respectively "the difference between an estimated value and the true value […]" (see: Dodge 2006, 134). "The larger the error, the lower the accuracy." (see: OECD 2007b, 263).

Note that the definitions of "Accuracy" and "Reliability" give a specification regarding the targeted size of the difference. This refers to the smallness of difference, i.e. its insignificance, if the data is marked by *accuracy* or *reliability*, respectively. Vice versa, data marked by *inaccuracy* and *unreliability* has a significant difference between the two observations. That does mean, the smaller the size of the difference the higher the quality of the data. The meaning of *accuracy* and *reliability* depend on the specific definition of the difference [18] and the margins for which data are considered *accurate* or *inaccurate*, and *reliable* or *unreliable*. Moreover, using a general term like "difference", "discrepancy", or "distance" changes the meaning of the definition, since using "difference" would result in the definition of the "*Error*".

Regarding the quality dimensions "*Accuracy*" and "*Reliability*" in NA, Fixler and Grimm (2005, 8) note that in Statistics, "the term 'accuracy' refers to the total measurement error, which in the NIPAs [National Income and Product Accounts][19] is never observed" [similar also noted by Young (1993)].

Sometimes the terms *accuracy* and *reliability* are not clearly distinguished or are used ambiguously. Naumann (2002, 30) found that the terms "data quality, error rate, correctness, reliability, integrity, [and] precision" were used as synonyms for *accuracy* by various authors.

In areas other than Statistics, *reliability* is used as meaning something different or it might belong to a completely different quality dimension providing information about the credibility or trustworthiness of data and a data provider (the data producing agency) (cf. Batini and Scannapieco 1998, 38.35.217). Others, like Karr, Sanil and Banks (2002, 8) refer to *reliability* as the application of appropriate statistical methods.

For NA statistics the Office for National Statistics (ONS) notes: "For the preliminary estimates of GDP, the key quality issues would seem to be how early, and how accurately, its value can be assessed with some certainty. […] This approach to measuring quality puts information on revisions at the centre of the quality assessment" (see: Brown, Buccellato, Chamberlin, Dey-Chowdhury, and Youll 2009, 47).

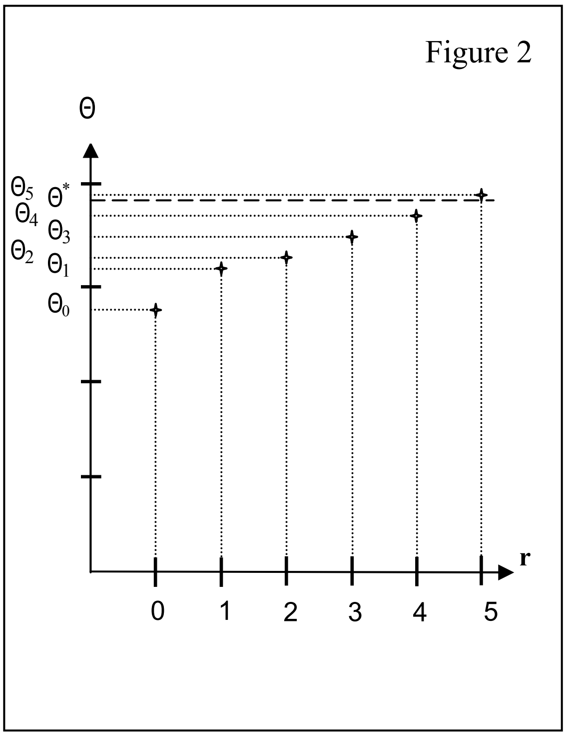
IOs do not carry out DQ assessment of the *accuracy* of NA, therefore this study focuses on the assessment of the *reliability* of NA. At the national level NSOs can investigate inconsistency within a NA data set and across other data sets (e.g. in comparison to BOP statistics) and expert knowledge of the data compilers (subjective

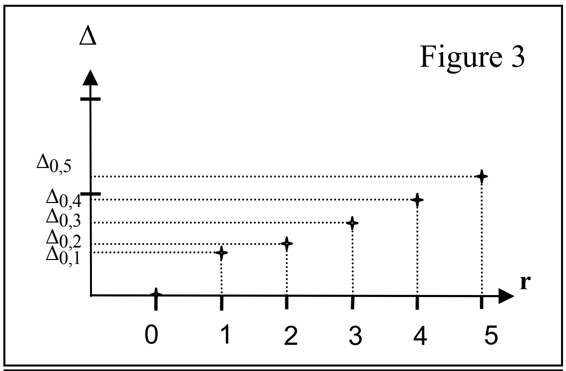
[18] Van Nederpelt (2009 ,23) defines "Accuracy of data" as "the degree to which an estimate of these data are close to the exact values of these data".

[19] Added by the author.

evaluation) to obtain indications (rather than precise measures) about the *accuracy* of the NA statistics, in addition to the information from revision studies for the *reliability* of data (cf. Carson and Laliberté 2002).

It should be noted that there is (usually) interdependence between NA becoming more *reliable* and more *accurate*, claimed e.g. by Bier and Ahnert (2001), Carson and Laliberté (2002), ABS (2003), Eurostat (2003b). This is because early estimates of NA statistics undergo routine revisions to incorporate new, more complete source data (cf. DEF. 4). For example Australian Bureau of Statistics (ABS) notes regarding their data: "In practice, at least in the Australian national accounts, accuracy and reliability tend to be interwoven and reinforcing. Ideally, as the estimate for a particular period passes through a sequence of revisions the size of the revisions gets smaller (the statistics become more reliable) and the estimate moves closer to the true value (the statistics become more accurate)" (ABS 2003, 18). (For further explanation refer to chapter 3.1.1 Definition of Revisions and Vintages, p. 37.)





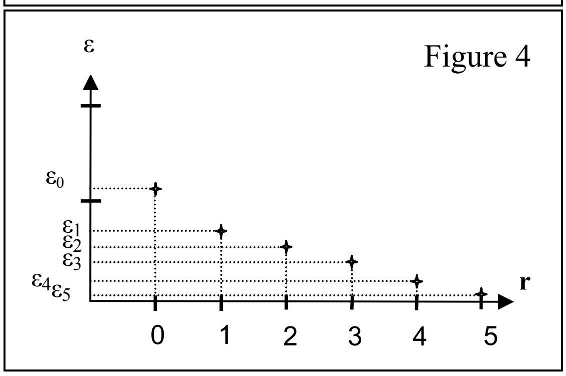


Figure 2 Θ-Diagram **Figure 3 Δ-Diagram** **Figure 4 ε-Diagram**

Source: own illustrations

Illustrated in Figure 2 are the initial release θ_0 of data for any NA indicator, e.g. GDP, and the subsequent releases of θ_r (with r = 1, …, 5) and the theoretical value for θ^*.

The *revised* releases (vintages) of θ_r are above the initial released vintage θ_0 and closer to the theoretical "true" value of θ. Figure 3 presents the *reliability* of the initial released value for θ_0 , i.e. the changes $\Delta_{1,r}$ from θ_0 to the subsequent releases of θ. Figure 4 presents the *accuracy* of the observed values for θ_r (with r = 0, …, 5) to the *theoretical* value θ*, i.e. the deviation ε_i (i = 1, …, 5) of θ_r (with r = 0, …, 5) to θ*. Since θ* is merely a theoretical value, ε is not computed in reality for NA.

DEF. 14: Maturity
Maturity refers to the 'age' of the statistical output and the underlying source data from which a given vintage of data is computed. It is the time period (TM) that has passed since the first (initial) release of the data until the release of the later vintage. Maturity is used as an attribute for a vintage of data of a specific event or phenomenon at different points in time. It is relevant only in context of data revisions where output statistics and source data are revised to incorporate new and more complete information that have become available over time. Maturity is calculated as: $\mathrm{TM} : X \times C \to \boldsymbol{R}^+$ $(x,c) \in X \times C \mapsto \mathrm{TM}(x,c) = t_{P(r)}(x,c) - t_{P(0)}(x,c),$ (II.15) where TM(x,c) : the maturity of a singular data object $x \in X$ at time point t_{Pr} of publication of the later vintage of data for a country $c \in C$, $t_{P(0)}(x,c)$: the time point of publication of the initial vintage of data object $x \in X$, for country $c \in C$, $t_{P(r)}(x,c)$: the time point of publication of the later vintage of data object $x \in X$, for country $c \in C$.

Source: own definition

Sample Application 3: Reliability(and Accuracy)

Netherlands

First and subsequent vintages of GDP Year-on-Year growth rate, 1. quarter 2007

Release time points / r, Θ_r	First estimate in t_0	t_0+5M	t_0+12M	t_0+24M	t_0+36M	t_0+end
Vintage (r)	0	1	2	3	4	5
Value (Θ_r)	2,8	3,2	3,3	3,5	3,7	3,9

t_0+5M: five month later vintage
t_0+12M: 12 month later vintage, etc.
t_0+end: the latest available vintage, given it is at least 48 month later than t_0.
Note that r referrers to specific release time points given in the first row.

We see that the first released Year-on-Year (YoY) growth rate (cf. DEF. 22, p. 48) is revised by +0,4 percentage points in the vintage observed five month later. The later observed vintages one, two, and three years later show a revision to the first released data of +0,5 , +0,7 , and +0,9 percentage points respectively. The latest observation, being a vintage observed at least four years after the first released data, shows a revision of +1,1 percentage points. For the reliability measure only the absolute change to the initial released value is of interest, e.g. 0,9 percentage points for the reliability of the first estimate to the three years later estimate.

Referring to the formula for the *absolute change* Δ_r (cf. DEF. 11) this equates to

Absolute Change Δ_r

r	θ_0	θ_r	Δ_r
1	2,8	3,2	0,4
2	2,8	3,3	0,5
3	2,8	3,5	0,7
4	2,8	3,7	0,9
5	2,8	3,9	1,1

The meaning of that absolute change depends on the definition of e.g. an upper boundary of the change which is considered reliable. Compared to the average of observed absolute changes to the initial released GDP YoY growth rates of the Netherlands (cf. Table 50 on p. 265) the provided example shows higher changes. For example the absolute change to the three years later observed value is exactly 0,2 percentage points higher than the average absolute change this time interval.

(continued on next page)

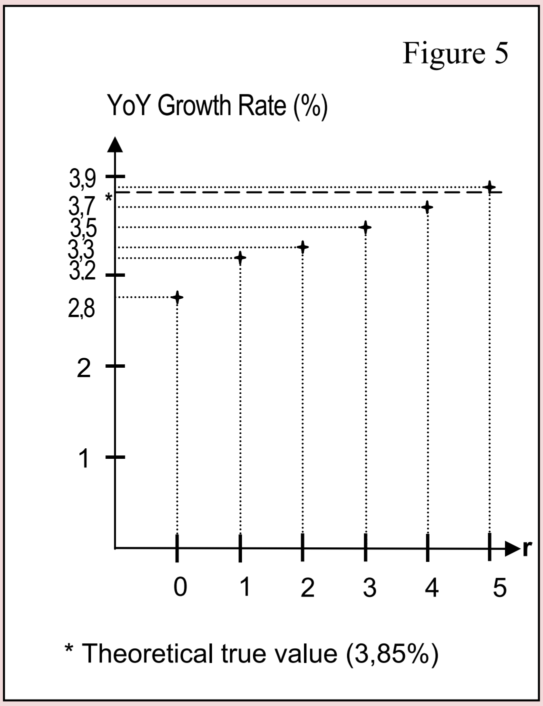
Accuracy cannot be determined since the true value is theoretical. However, we can calculate an example under the assumption that the true GDP YoY growth rate for the 1. quarter of 2007 is 3,85 percent. Thus we have a absolute deviation of the first released value to the true value of 1,05 percentage points. The accuracy of the three years later estimate is 0,15 percentage points.

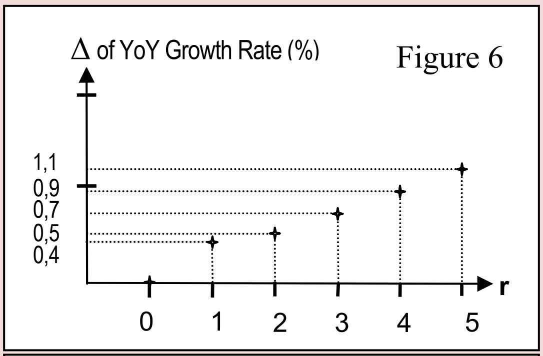
Referring to the formula for the *absolute deviation* ε_r to the theoretical 'true' value θ^* (cf. DEF. 12) this equates to

Absolute Deviation ε_r

r	θ_r	θ^*	ε_r
0	2,8	3,85	1,05
1	3,2	3,85	0,75
2	3,3	3,85	0,55
3	3,5	3,85	0,35
4	3,7	3,85	0,15
5	3,9	3,85	0,05

The reliability of the first value to the subsequent released values and the accuracy of these values to the true (theoretical) value are illustrated in the following Figure 5, Figure 6 and Figure 7.





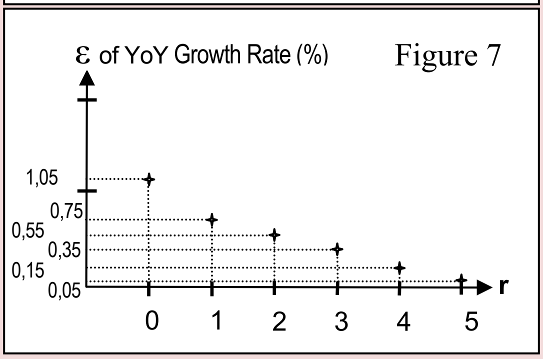


Figure 5 YoY Growth Rate Diagram

Figure 6 Δ of YoY Growth Rate Diagram

Figure 7 ε of YoY Growth Rate Diagram

3 Data Quality Measures

In this chapter, we define the DQ measures we use to assess the three DQ dimensions.

3.1 Reliability Measures

In the following we define and discuss measures for *reliability* of NA statistics.

3.1.1 Definition of Revisions and Vintages

Revision (R) is calculated as follows[20].

DEF. 15: Revision (R)
Is the size of change (revision) from the first (initial) value to the later revised value (level of the estimate or growth rate) of the same reference periods (t). The Revision (R) considers the sign and thus indicates whether the change to the later value is positive or negative. $$R_r^t = \Delta_r^{*t} = \Delta_{0,r}^{*t} := \theta_r^t - \theta_0^t, \qquad \text{(II.16)}$$ where R_r^t : the directional size of revision (i.e. the change Δ_r^{*t} considering the sign) from the initial vintage 0 to vintage r, with r = 1, 2, … 100. Where vintages 0,r refer to the same reference period t, and θ_0^t respectively θ_r^t the values referencing the same reference period t. θ_0^t : the first published value of θ for reference period t, θ_r^t : the value of θ for the same reference period t at time point of a later vintage r (i.e. republication at a later time point).   (continued on next page)

[20] Hitherto existing writing: "$R_t = L_t - E_t$, where L is the change or percent change in the later vintage quarterly or annual estimates, and E is the change or percent change in the earlier vintage estimates." (cf. Fixler, Greenaway-McGrevy, Grimm 2011).

Since the different vintages of θ are of the same reference period, the index t can be dropped in the formula. Thus we have $\Delta^*_r = \theta_r - \theta_0$.

Referring to the *reliability* dimension (DEF. 11) we have $\Delta_{i,r} = | \Delta^*_{i,r} |$ for $0 \leq 1 < r$.

[R can be calculated for any two values of different maturity, i.e. the intermediate revisions $R_{r,i}$ between an earlier released value θ_i and a later released value θ_r is: $R_{r,i} = \Delta^*_{i,r} = \theta_r - \theta_i$, with $i < r$.

Ditto for all following revision measures for the reliability study.]

Source: Author's verbal definition; formula: Fixler, Greenaway-McGrevy, Grimm (2011) modified by the author.

It is intuitive that the size of *revisions* of NA data are calculated for the different vintages of indicators of the same reference period, therefore (as done in the following) the index t for the reference period is often omitted in the formulas for the *revision* measures (cf. Statistisches Bundesamt 2011).

Since R provides the direction (i.e. it considers the sign) besides the size of the change between two vintages of different maturity, it is important to deduct the earlier vintage form the later vintage, not the other way around. Fixler and Grimm (2002, 9) define *revisions* as: "the changes from an earlier vintage of estimates to a later vintage".

According to Carson and Laliberté (2002, 4), in NA, "reliability refers to assessing revisions". The size of revision (R) to the initial released estimate of a given reference period is the basis for the measures used in *revision* studies. These studies evaluate the average *reliability* of first published values based on the observations for multiple reference periods. To provide meaningful results the studies should cover at least 20 observation periods (cf. Ahmad, Bournot, and Koechlin 2004).

Sample Application 4: Revision

Netherlands

Changes to first estimate of GDP Year-on-Year growth rate for 3. quarter 2006

Release time points / r, Θ_r	First estimate in t_0	t_0+5M	t_0+12M	t_0+24M	t_0+36M	t_0+end
Vintage (r)	0	1	2	3	4	5
Value (Θ_r)	2,9	3,0	2,8	3,1	3,1	3,3

Looking at the revisions, i.e. the change (considering the sign) to the first released Year-on-Year (YoY) growth rate (cf. DEF. 22, p. 48) we see that the first value of 2,9 percent is revised by +0,1 percentage points by the time of the five month later vintage (see value 3,0 below vintage 1). By the time point of the one year later vintage the first estimate is revised downwards by -0,1 percentage points. The two and three years later vintages show a revision to the first released data of +0,2 in both cases. The latest observation, being a vintage observed at least four years after the first released data, shows a revision of +0,4 percentage points.

Referring to the formula for the *revision* R_r^t (cf. DEF. 15) this equates to

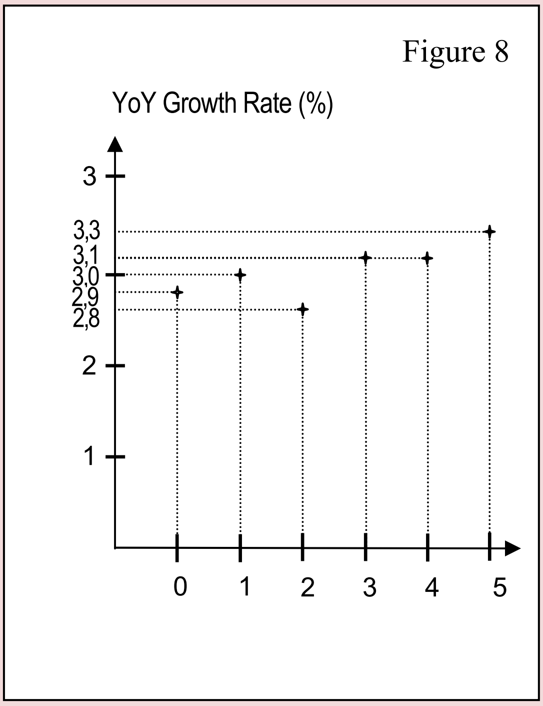
Revision R_r^t

r	θ_r	θ_0	R_r^{2006Q3}
1	3,0	2,9	0,1
2	2,8	2,9	-0,1
3	3,1	2,9	0,2
4	3,1	2,9	0,2
5	3,3	2,9	0,4

The meaning of the revision depends on the definition of a fix point for comparison e.g. a lower and upper boundary.

The YoY growth rates and revisions-diagram are given in Figure 8 and Figure 9.

(continued on next page)



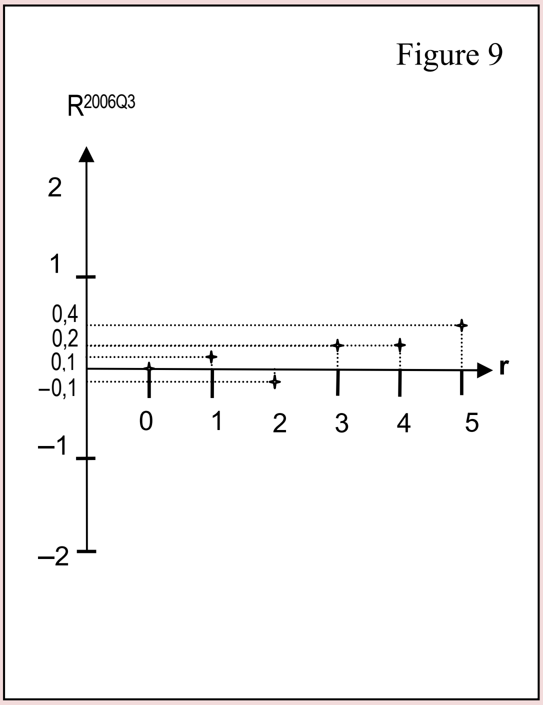


Figure 8 YoY Growth Rate-Diagram

Figure 9 R-Diagram

3.1.1.1 Final Data and the Latest Vintage

The following defines the "*final*" and "*latest*" data. These are typically the center of interest. Their definition in context of a reliability study is particularly important since the definitions of these terms often differ between revision studies. Further, these definitions influence the results of the revision studies.

DEF. 16: Final Data
The result of the last scheduled revision of preliminary estimates to include new source data. They are usually computed from all needed, i.e. complete, source data. The time point of the last scheduled revision is based on the country's regular revision cycle. See also: DEF. 10: Completeness See also: DEF. 3: Statistical Outputs

Source: own definition

DEF. 17: Latest Vintage
The "latest" vintage is the latest available data with at least four years maturity since the first release of a given reference period.

Source: own definition

Final data (also referred to as *final* vintage or *final* estimate) are typically calculated based on complete source data. Hence, the *final* data are not scheduled for further revisions. The initial released estimate (θ_0) may be revised several times until all required source data are available to compute the *final* revision of the later vintage (θ_r).

Sometimes it is argued, that NA data can never be marked as "*final*", since they could always be revised again, for example Di Fonzo (2005a). Revisions of "*final*" data occur for example to correct mistakes or with the introduction of new concepts and methods and data sources. However, backward revisions of "*final*" data in conjunction with a comprehensive revision are for the purpose of building coherent time series. Conversely, revisions of preliminary vintages of NA data require further revisions since they are still based on preliminary (incomplete) source data.

"The *final* NA data do represent the best estimates of the true values" (cf. Denton and Oksanen 1973, 446), i.e. the *final* GDP is the best representative for the "*true*" unknown GDP. This can be based on the argument that "the differences between the early release and most current estimates are largely news" (i.e. new information is incorporated to compute the revised GDP, thus the current estimate contains more information about the "*true*" unobserved GDP) (cf. Fixler and Nalewaik 2004, 4 and 2010, 4). For the (i) kinds of source data used to compute NA statistics (QNA and ANA), and analysis regarding the (ii) kinds of source data used to compute the NA statistics at vintages of different maturity (for GDP or a more comprehensive NA data set), and the (iii) shares of kinds of source data in the computed GDP at vintages of different maturity, refer for example to Grimm and Weadock (2006), Holdren and Grimm (2008), or Landefeld, Seskin, and Fraumeni (2008).

Even if all possible sources were considered the (*final*) data might be inaccurate, because the sources were inaccurate, thus the computation of accurate NA statistics was impossible (cf. Di Fonzo 2005b). Considering the lack of complete information at the time of computation of the preliminary data and the completion of source data as estimates mature, revisions also refer to the tradeoff between accuracy and timeliness.

The definition of the **"latest" vintage** varies across the revision studies of different intuitions and researchers. Sometimes it simply refers to the *latest available* vintage for the reference period. However in that case it might return data of different maturity. For example, in the existing revision studies some of the *latest available* vintages refer to data at the time point of the first quarterly revision, e.g. Branchi et al. (2007) and Deutsche Bundesbank (2011). Yet again, if the reference period is a decade ago the actual "*final*" data might have been subject to several comprehensive revisions, e.g. McKenzie and Adams (2007). The "*latest*" vintage might also be conditioned to a vintage with a minimum maturity, e.g. at least three years after initial release, e.g. McKenzie and Adams (2007). The revision cycle could also be limited to only consider the vintages until data are "*final*", thus the "*latest*" vintage would not refer to revisions of back data, meaning a re-calculation of back years after introduction of a new methodology, e.g. Deutsche Bundesbank (2011).

3.1.1.2 Classes of Revisions

The terminologies to identify different kinds of revisions differ between studies. Therefore it is necessary to define relevant classes in the context of this study (see Table 2 for the definitions of classes). There is broad agreement regarding the term "*current*" (regular) *revisions*, which is used to refer to the inclusion of more complete source data for preliminary estimates.

Beside the "*current*" (regular) *revisions* to the preliminary estimates, *comprehensive revisions* are carried out about every 5 to 10 years. *Comprehensive revisions* cover the changes in concepts (definitions and classifications), of methods (e.g. change between seasonally and non-seasonally adjustment), and of data (e.g. inclusion of new data sources). Thus *comprehensive revisions* occur due to methodological changes (concepts and methods) or revisions of the kind and scope of utilized data. Consequently all kinds of *comprehensive revisions* are changing the benchmark of the computed NA data.

For our revision study we are not concerned about the reason for the *comprehensive revision*. Yet, relevant is whether the revision has a huge impact on the levels of data and growth rates from which the reliability measures are computed. Hence, this study suggests distinguishing between three *classes of revisions*: *regular revisions*, and "*minor*" and "*major*" *comprehensive revisions*, as defined in the following table.

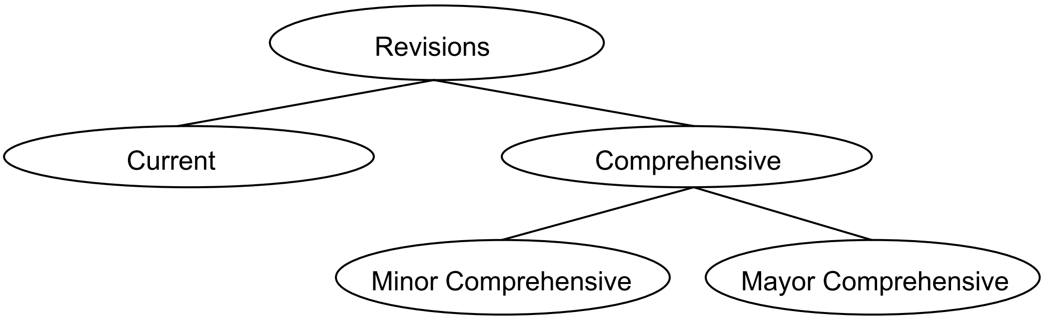


Figure 10 Classes of Revisions
Source: own illustration

Table 2 Classes of Revisions

Class of Revision	**Definition**
Current revision (alternatively: regular or routine)	The revisions to initial released data due to incorporation of more complete source data over the course of the regular revision cycle. This excludes any revisions due to changes in standards, methods or concepts.
(Minor) Comprehensive revision	Revisions to a time series of data that occur on a regular basis. They render adaptations and (minor) changes in standards, methods, concepts and data sources (like changes regarding seasonal adjustment), yet not major methodological changes or introduction of substantial new data sources.
"Major" comprehensive revision	The introduction of a new SNA version, e.g. change from 1968 to 1993 SNA; substantial changes to concepts and methods or inclusion of new data sources that have significant impact on the growth rates and data levels[21].

Source: own compilation

[21] For example the major revision of the Chinese time series in 2005.

3.1.2 **Selected Measures of Revision Analysis**

We study the *reliability* of the first published value of the quarterly GDP[22] in constant price data YoY growth rate (%). Revision analysis uses different *revision measures* to evaluate the observed individual revisions. These *measures* provide information about the *average reliability* of the first published vintage, compared to a (revised) vintage published at a later time point, i.e. a vintage of higher maturity.

"Of particular interest are basic summary statistics which give information on the normal expectations for revision." (McKenzie and Gamba 2008, 3). The typically published measures by NSOs or IOs in revision studies are the: *Mean Revision* (MR), *Mean Absolute Revision (*MAR), and *Relative Mean Absolute Revision* (RMAR), while the latter is not as frequently used.

This subchapter defines the MR, MAR, and RMAR for the first published value of the quarterly GDP in constant price data YoY growth rate (%). While the verbal parts of the definitions of these measures are carefully crafted by the author, the specialty of these definitions is a more mathematical correct version of the formulas compared to the previous ones.

Further measures, that can be used as alternative, additional, or supplementary measures for analyzing revisions are explained in McKenzie and Gamba (2008).

[22] Partially seasonally adjusted data.

The *Mean Revision* (MR), also referred to as $\bar{R}$, calculated as follows[23]:

DEF. 18: Mean Revision (MR)
The average of the change (i.e. of the revisions R_r^t), for a fixed revision interval [0,r], from the first (initial) value (vintage 0) to the later revised value (vintage r) of the same reference period (t), observed for n such periods. The MR_r^t considers the sign of the revisions and thus indicates whether the revision to the first published vintage until the publication of vintage r is an average positive or negative. $$MR_r^t = \frac{1}{n} \sum_{t \in RQ} \left(x_r^t - x_0^t \right), \qquad (II.17)$$ with $MR_r^t = MR_{0,r}^t$: the mean revision (in percentage points) from vintage 0 to vintage r. The vintages 0, r referring to the same reference period t (quarter), and x_r^t respectively x_0^t are the values referencing the same t (with "r" referring in the empirical study to specific later revision time points), RQ : the set of reference periods t and n = $\|RQ\|$ the number of these periods.[24] x_r^t: the value of the quarterly GDP in constant price data YoY growth rate in percent, at time point of any vintage other than the first release, for reference period t, x_0^t: the value at the time point of the first release (for the reference period t), and   n : the number of observations (different reference periods).

Source: Formula by Fixler, Greenaway-McGrevy, Grimm (2011) modified by the author, with own verbal definition.

[23] Hitherto existing writing: $MR = \sum_t R_t / n$, t=1,..,n (cf. Fixler and Grimm 2011).

[24] For example RQ = {2006Q3, 2006Q4, …, 2012Q1} does mean n = 23.

The MR indicates whether the revisions are an average positive or negative and the size of it. Sometimes it is referred to as "bias" (cf. Statistisches Bundesamt 2013). However, this term should be avoided, since no bias exists, "when the mean revision is not significantly different from zero" (Di Fonzo 2005a, 10). Ahmad, Bournot, and Koechlin (2004, 16) note the "full assessment of whether bias does exist is best made by each statistical office, since they will have full access to the nature of revisions".

The *Mean Absolute Revision (*MAR) for multiple reference years is calculated as follows[25]:

DEF. 19: Mean Absolute Revision (MAR)
The mean of the absolute size of revisions of the same revision interval (0 to r) of multiple reference periods (t). The MAR does not consider the sign of the revisions and thus indicates the average absolute size of revision (average absolute change) until publication of vintage r, compared to the first published value. $$\mathrm{MAR}_r^t = \frac{1}{n}\sum_{t \in RQ}\left|x_r^t - x_0^t\right|, \qquad \text{(II.18)}$$ with $\mathrm{MAR}_r^t = \mathrm{MAR}_{0,r}^t$: the mean absolute revision (in percentage points) from vintage 0 to vintage r. The vintages referring to the same reference period t (quarter), and x_r^t respectively x_0^t are the values referencing the same t. x_r^t respectively x_0^t, *RQ*, and n : see DEF. 18.

Source: Fixler, Greenaway-McGrevy, Grimm (2011) modified by the author; own verbal definition.

The MAR eliminates the offsetting effect of positive and negative revisions by not considering the sign of the individual revisions. Thus it provides the average absolute deviation of the revised value from the first value (for the observed reference periods).

[25] Hitherto existing writing: $MAR = \sum_t |R_t| / n,\ t = 1, \ldots, n$ (cf. Fixler and Grimm 2011).

The *Relative Mean Absolute Revision* (RMAR) is calculated as follows[26]:

DEF. 20: Relative Mean Absolute Revision (RMAR)
Relates the sum of the absolute size of revisions of the same revision interval (0 to r) of multiple reference periods (t) to the absolute sum of the values of the vintage (r) in period (t). $$\mathrm{RMAR}_r^t = \frac{\sum_{t \in RQ} \left| x_r^t - x_0^t \right|}{\sum_{t \in RQ} \left| x_r^t \right|}, \qquad \text{(II.19)}$$ with $\mathrm{RMAR}_r^t := \mathrm{RMAR}_{0,r}^t$: the relative mean absolute revision (in percentage points) from vintage 0 to vintage r. The vintages referring to the same reference period t (quarter), and x_r^t respectively x_0^t are the values referencing the same t, x_r^t respectively x_0^t, *RQ* and n: see DEF. 18.

Source: Tosetto and Brackfield (2009) modified by the author; own verbal definition.

The value of RMAR_r^t gives the average absolute change of x_0^t to x_r^t across different reference periods t, in relation to the average value of the absolute value of the revised values x_r^t observed across the same reference periods. The RMAR measure is not only relevant for national reliability studies, but particularly valuable for cross-country comparisons, as it provides a measure that is adjusted for the individual growth rates of the NA data of a country. Thus it takes into account differences in the growth rates between countries in seasons of high growth and those in seasons of low growth.

[26] Hitherto existing writing: $RMAR = \frac{\sum_{t=1}^{n} |R_t|}{\sum_{t=1}^{n} |L_t|}$ (cf. Tosetto and Brackfield 2009); for R_t and L_t see footnote 20, p. 41.

DEF. 21 Year-on-Year (YoY) Growth Rates
Year-on-year growth rates are rates of change expressed over the corresponding period (month or quarter in relation to the frequency of the data) of the previous year. Such rates are expressed as $(M_t/M_{t-12}) - 1$ or $(Q_t/Q_{t-4}) - 1$. M_t denotes the value of a monthly time series in month t and Q_t the value of a quarterly time series in quarter t.
[Year-on-Year growth rates are also] referred to as year-over-year growth rates, year-to-year growth rate, rate of change from the previous year, or 12-month rate of change.

Source: OECD (2007a, 50.158)

We consider YoY rates of change in percent for quarterly GDP in constant price data. Corresponding to DEF. 21 we define our growth rate of interest, while modifying the hitherto exiting writing (though intuitive) to a stricter mathematical notation:

DEF. 22 Quarterly GDP Year-on-Year (YoY) Growth Rates (%)

If t = (y,q) refers to the reference period of year y and quarter q, then we get for the YoY rates of change in percent for quarterly GDP [27]

$$x^t = \left(\frac{GDP_{(y,q)}}{GDP_{(y-1,q)}} - 1 \right) \times 100,$$

where

x^t : the YoY rates of change in percent for quarterly GDP, and

$GDP_{(y,q)}$ refers to the GDP value for quarter (y,q) and $GDP_{(y-1,q)}$ refers to the GDP value for the same quarter one year earlier (y-1,q) (thus for example for t = 2001Q3 we get: y = 2000, q = 3 and the same quarter one year earlier given by y = 1999, q =3).

Source: OECD (2007a, 50.158) modified by the author and adapted to quarterly GDP levels.

[27] Analogous for any monthly time series of data for reference period t = (y,m) with month m we get the YoY growth rates: $g^t_{month} = \frac{M_{(y,m)}}{M_{(y-1,m)}} - 1$ (for example t = 2011M10 does mean y = 2011, m = 10).

On the other hand, the Quarter-on-Quarter (QoQ) growth rates are also often used when analyzing revisions data. QoQ growth rates are also referred to as "Quarter-on-previous-quarter growth rates", and "Quarter-to-quarter (Period-to-period) growth rates, Quarter-over-quarter growth rates, 1-quarter growth rates, or Rate of change on the previous quarter", and are defined as "rates of change expressed with respect to the previous quarter […] expressed as $(Q_t / Q_{t-1}) - 1$." (cf. OECD 2007a, 50.151 for all three quotes). Q_{t-1} refers to the value for the previous quarter of t.

Matching our definition for quarterly GDP YoY growth rates in percent (cf. DEF. 22) the strict mathematical notation for quarterly GDP QoQ growth rates in percent is calculated as follows:

DEF. 23 Quarterly GDP Quarter-on-Quarter (QoQ) Growth Rates (%)

If t = (y,q) refers to the reference period of year y and quarter q, then we get for the QoQ rates of change in percent for quarterly GDP

$$x_{QoQ}^{t} = \left(\frac{GDP_{(y,q)}}{GDP_{(y,q-1)}} - 1 \right) \times 100,$$

where

x_{QoQ}^{t} : the QoQ rates of change in percent for quarterly GDP, and

$GDP_{(y,q)}$ refers to the GDP value for quarter (y,q) and $GDP_{(y,q-1)}$ refers to the GDP value for the previous quarter of the same year (y,q-1) (thus for example for t = 2001Q3 we get: y = 2000, q = 3 and the previous quarter of the same year given by y = 2000, q =2).

Source: OECD (2007a, 50.151) modified by the author and adapted to quarterly GDP levels.

Sample Application 5: Mean Revision (MR), Mean Absolute Revision (MAR), and Relative Mean Absolute Revision (RMAR)

Netherlands

GDP Year-on-Year growth rate (x_r^t) for 2006Q1 - 2007Q1 at Different Release Time Points (vintages "r") and Corresponding Revision Measures

Release time points of x_r^t / Reference Period (t)	First estimate to	t_0+5M	t_0+12M	t_0+24M	t_0+36M	t_0+end
Memorial item: vintage (r)	0	1	2	3	4	5
2006Q1	2,1	2,3	2,6	2,9	3,3	3,3
2006Q2	2,6	3,0	3,2	3,7	3,7	3,7
2006Q3	2,9	3,0	2,8	3,1	3,1	3,3
2006Q4	2,9	2,9	2,9	3,5	3,5	3,4
2007Q1	2,8	3,2	3,3	3,5	3,7	3,9
Revision Measures						
MR		0,22	0,30	0,68	0,76	0,86
MAR		0,22	0,34	0,68	0,76	0,86
RMAR		0,08	0,11	0,20	0,22	0,24

t_0+5M: five month later vintage, i.e. sub-annual vintage.
t_0+12M: 12 month later vintage, in the following referred to as "t_0+1Y", i.e. one year later vintage, etc.

Looking at the revisions to first released YoY growth rates of the QNA data of five consecutive quarters we can observe the MR, MAR and RMAR to the first estimate with respect to the later observed vintages. For the first to the fiver month later vintage we see that the MR is 0,22, the MAR is 0,22, and the RMAR is 0,076.

Referring to the formulas for MR, MAR, and RMAR (cf. DEF. 18 to DEF. 20) this equates for the revision from the release of the first vintage at time point t_0 to the release of the later vintage at time point t_0+5M to

$$MR_{t_0+5M} = \frac{1}{5}(2,3-2,1)+(3,0-2,6)+(3,0-2,9)+(2,9-2,9)+(3,2-2,8) = 0,22,$$

$$MAR_{t_0+5M} = \frac{1}{5}\left(|2,3-2,1|+|3,0-2,6|+|3,0-2,9|+|2,9-2,9|+|3,2-2,8|\right) = 0,22,$$

$$RMAR_{t_0+5M} = \frac{|2,3-2,1|+|3,0-2,6|+|3,0-2,9|+|2,9-2,9|+|3,2-2,8|}{|2,3+3,0+3,0+2,9+2,8|} = 0,08.$$

(continued on next page)

For the revisions to the first estimate for the later observed vintages, analogous.

Since we observe for the one year later vintage of 2006Q3 a negative revision to the initial released estimate, the MR for the vintages with maturity t_0+12M is lower than the corresponding MAR. For all other vintages these two revision measures yield the same values.

3.2 Availability Measures

This chapter presents the availability[28] measures and corresponding formulas to assess the data availability for countries, regions and other country groups.

Of foremost importance among the availability measures are the sets of NAQ tables defined as the "Minimum Requirement Data Set (MRDS)" for the 1993 SNA and 2008 SNA versions, and the "Milestone Levels", particularly levels 1 and 2. The latter are typically included in the data availability assessments of the UN.

Provided in the following, is an overview of the availability measures defined in this chapter. The information "n" following the availability measure refers to the superset "$\{0,1\}^n$" of n-tuples of lengths n, called state vectors s. The state vectors identify the existence and non-existence of the NAQ tables[29] considered for the fulfillment of the availability measures. Thus there are 2^n possible state vectors. The subset of relevant state vectors within $\{0,1\}^n$ is characterized by the cardinality $r < 2^n$; the value for r is provided, too. [30]

1. The Minimum Requirement Data Set (MRDS):
    a. 2008 MRDS: n = 14, r = 15.
    b. 1993 MRDS: n = 8, r = 3.
2. At least the partial fulfillment of the MRDS (i.e. at least all but one of the MRDS criteria are fulfilled), formally delineated by a distance measure to the previously delineated n-tuples for the 1993 and 2008 MRDS:
    a. Substantial part of the 2008 MRDS: n = 14, r = 164.
    b. Substantial part of the 1993 MRDS: n = 8, r = 6.

---

[28] Availability refers to the nationally produced data, measured by the data transmitted by a national agency to the UN and made available through publication.

[29] The NAQ table number x.y (as per Table 3) is used in the following within the text to refer to the NAQ tables (within formulas the object $T_{x.y}$ and the variable $t_{x.y}$ are used).

[30] Generally r is small compared to $\{0,1\}^n$.

3. Availability of individual NAQ tables relevant for the particular MRDS:

    a. 2008 MRDS NAQ tables: n = 1, r = 1.
    (For availability criteria of NAQ tables[31] 1.1, 1.2, 2.1, 2.2, 2.3, required by both MRDS versions see b. Availability criteria of NAQ tables 1.3, 4.y, y = 1, …, 9 until the Net Lending (NL) indicator of the capital account. In the following NL refers to net lending of the capital account.[32])

    b. 1993 MRDS: n = 1, r = 1. (NAQ tables 1.1, 1.2, 2.1, 2.2, 2.3, 4.2, without requirement of a specific NA indicator. Regarding NAQ tables 1.3 and 4.1, without requirement of a specific NA indicator see: 4.)

4. Alternative NAQ table combinations:

    a. NAQ table 1.3 or 4.1 of 2008 MRDS (i.e. required until NL): n = 2, r = 3.

    b. NAQ table 1.3 or 4.1 of 1993 MRDS (i.e. no requirement of a specific NA indicator): n = 2, r = 3.

    c. NAQ table 4.3 and 4.4 or at least 4.8 for 2008 MRDS (i.e. required until NL): n = 3, r = 5.

    d. NAQ table 4.6 and 4.7 or at least 4.9 for 2008 MRDS (i.e. required until NL): n = 3, r = 5.

5. Institutional Sectors of the 2008 MRDS:

    a. All requirement institutional sectors (NAQ tables 4.y, y = 1, …, 7, 4.9, all required until NL): n = 8, r = 5.

    b. At least a substantial part of the institutional sectors (i.e. at least all but one of the needed criteria are fulfilled), formally delineated by a distance measure to the previously delineated n-tuples for all needed institutional sectors): n = 8, r = 33.

[31] Regarding the NAQ table numbers and corresponding NAQ table names see: Table 3, p. 49.

[32] NAQ table 4.8 is considered in addition to the defined requirement NAQ tables.

6. Milestone Levels 1 and 2 according to the original “strict” and later “relaxed interpretation of needed NAQ tables:
    a. Milestone Level 1 (strict): n = 4, r = 1.
    b. Milestone Level 2 (strict) (with NAQ tables 4.1, 4.2 until NL and the financial accounts for 4.2): n = 8, r = 3.
    c. Milestone Level 1 (relaxed): n = 4, r = 9.
    d. Milestone Level 2 (relaxed) (for NAQ tables 4.1, 4.2 without requirement of a specific NA indicator): n = 6, r = 27.

The availability of data for a country is assessed based on the data published in the YB. Thus it is based on official statistics that are produced by the mandated national agency (e.g. NSO) and subsequently transmitted by the country to the UN.

C is the set of assessed countries[33] (e.g. the UN Member States, OECD countries, a set by economic classification like “transition economies”, a special classification like “Least Developed Countries” (LDC), an economic group like the “G20”, or a region).

A country $c \in C$ that produces for example the minimum requirement of NA data for the 1993 MRDS fulfills it. The set of countries that fulfills a specific availability measure, here the 1993 MRDS is denominated as $C_{1993} \subseteq C$. For the sets of countries producing another availability measure e.g. the 2008 MRDS, analogous, i.e. $C_{2008} \subseteq C$.

The availability[34] study in part III, chapter 2 uses the above measures, i.e. for the 1993 and 2008 MRDS etc. to assess the availability of NA data for a country $c \in C$. The results for the specific availability measures are presented in aggregated form for the sets of countries that fulfill these measures e.g. by C_{1993}.

These sets (i.e. C_{1993} etc.) are presented for different country sets C (country clusters), e.g. for regions or by membership in organizations. The results are presented in Table 34 to Table 39 (pp. 214 - 221) of the 1993 MRDS assessment and in Table 40 to Table 45 (pp. 223 - 230) of the 2008 MRDS assessment, respectively.

[33] The set of countries varies, e.g.: all UN countries, all countries including the 32 non-UN countries, regions, and organizations (for all assessed sets of countries see regional aggregates in part III, chapter 2).

[34] Availability refers to the nationally produced data, measured by the data transmitted by a national agency to the UN and made available through publication.

3.2.1 **Minimum Required Data Set**

DEF. 24: Minimum Requirement Data Set (General Definition)
The 1993 and 2008 Minimum Requirement Data Sets (MRDS) of NA that should be produced by each country are given by sets of National Accounts Questionnaire (NAQ) tables as fixed in United Nations (2001b and 2011). Note: the 1993 MRDS is a part of the 2008 MRDS. See also: DEF. 25 for the 1993 MRDS and DEF. 26 for the 2008 MRDS.

Source: own definition based on United Nations (2001b and 2011).

In this study, we refer to the set of NAQ tables that make up the MRDS in case of the 1993, respectively, 2008 SNA as the 1993 MRDS respectively 2008 MRDS. Table 3 gives the required NAQ tables for the 1993 and 2008 MRDS[35] (by number and name). The required sets of NAQ tables are internationally agreed as published in the reports of the Intersecretariat Working Group on National Accounts (ISWGNA) (cf. United Nations 2001b, 10.11 and 2011, 13, respectively).

The considered NAQ tables reflect the data sets of different SNA concepts. The data set that is requested by each NAQ table is identified by the NAQ table name. The tables of NA data in the SNA standard correspond to the NAQ tables (i.e. the spreadsheets of requested NA indicators in the questionnaire).

The NAQ tables given in Table 3 are components of the 1993 and 2008 MRDS. Yet, not all of the NAQ tables given in Table 3 are required to fulfill the criteria of the corresponding MRDS (see footnotes for possible combinations of NAQ tables and specific required SNA indicators). That does mean that combinations (i.e. the relevant subsets) of these NAQ tables are sufficient to fulfill the MRDS criteria. The latter are the subsets of NAQ tables that need to be existent in any $c \in C$. For the formal definition of the 1993 and 2008 MRDS refer to DEF. 25, p. 59 respectively DEF. 26, p. 64.

[35] A subset of 8 and 14 of these NAQ tables is considered for the 1993 and 2008 MRDS, respectively.

Table 3 Minimum Requirement Data Sets for the 1993 SNA and 2008 SNA

NAQ Table No	SNA Table Name	1993 MRDS	2008 MRDS
1.1	Expenditures of the GDP[/a] in current prices	☒	☒ 1
1.2	Expenditures of the GDP[/a] in constant prices	☒	☒ 1
1.3	Relations among product, income, saving, and net lending aggregates	☒ 2/ 3	☒ 2/ 4/ 5
2.1	Value added and GDP[/b] in current prices by industry	☒	☒ 1
2.2	Value added and GDP[/b] in constant prices by industry	☒	☒ 1
2.3[/c]	Value added components by industry[/d] in current prices	☒	☒
4.1	Total economy institutional sector accounts	☒ 2/ 3	☒ 2/ 4/ 5
4.2	Rest of the world institutional sector accounts	☒ 3	☒ 4/ 5
4.3	Non-financial corporations institutional sector accounts	☐	☒ 5
4.4	Financial corporations institutional sector accounts	☐	☒ 5
4.5	General government institutional sector accounts	☐	☒ 5
4.6	Households institutional sector accounts	☐	☒ 5/ 6
4.7	Non-profit institutions serving households sector accounts	☐	☒ 5/ 6
4.9	Households and non-profit institutions serving households sector accounts	☐	☒ 5/ 6

/a In the NAQ renamed to: GDP by expenditure.
/b In the NAQ renamed to: Value added by industry.
/c Covers the industries by ISIC categories (requesting output, value added, income components, and fixed assets).
/d In the NAQ renamed to: Output, value added and fixed assets by industry.

1 Quarterly accounts are required in addition to annual accounts for tables 1.1 and 1.2, or 2.1 and 2.2.
2 Required is table 1.3 or 4.1.
3 The definition included the requirement of this table until net lending aggregate (of the capital account), in practice this additional requirement was dropped i.e. only the table was required, no specific indicator of it (cf. UNSD 2008).
4 Quarterly accounts are required in addition to annual accounts.
5 Required until net lending aggregate (of the capital account).
6 Required are tables 4.6 and 4.7, or the institutional sector accounts of table 4.9 that presents tables 4.6 and 4.7 together.

Source: Own illustration compiled from the SNA data set tables in United Nations (2001b and 2011)

To fulfill the 1993 MRDS criteria, countries are not required to compute both of the NAQ tables 4.1 and 1.3, since both tables cover the same main income indicators. Further, though the 1993 MRDS as per United Nations (2001b, 10.11) requested the availability of data for the NAQ tables 1.3, 4.1, and 4.2 until the indicator NL, this was in practice not measured by the 1993 MRDS assessment.

In practice the 1993 MRDS requirements were redefined to not request any specific indicator. This is obvious from statements in current 1993 MRDS assessments. These read that countries that fulfilled the 1993 MRDS reported ("at least")[36] "parts of the integrated economic accounts for the total economy and the rest of the world" (UNDS 2008, 9). Since the accounts for the total economy and the rest of the world refer to the NAQ tables 4.1 and 4.2, respectively, the latter does mean that (at least) parts of NAQ tables 4.1 or 4.2 are reported. The redefinition of the requirements applies likewise to NAQ table 1.3 since it covers the same information as table 4.1. However, with the quoted description no specific SNA indicator is identified regarding the detail of the reported NAQ table. Since NAQ tables 1.3, as well as 4.1 to 4.9 cover the integrated accounts starting form GDP and it components until the "net lending" balance of the capital account respectively of the financial account, the possible depth of reporting can differ substantially. A country may for example report indicators for all accounts or merely the basic data (referring to GDP until GNI). The redefined requirements make the availability of any indicator for these NAQ tables sufficient without need for checking the availability of a specific SNA indicator. The redefinition is reflected by the formal delineation of the 1993 MRDS in the following subchapter.

The 2008 MRDS extends the 1993 MRDS requirements by inclusion of NAQ tables for all institutional sectors and the specification of a particular SNA indicator for the institutional sectors. Regarding the additional NAQ tables for intuitional sectors (NAQ tables 4.1 to 4.7), countries need to produce either NAQ tables 4.6 and 4.7 or alternatively NAQ table 4.9 with the unconsolidated aggregated data of the first two tables[37]. Further, the NAQ tables 1.3, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, and 4.9 need to

[36] Added in United Nations (2011, 22).

[37] Providing unconsolidated data in NAQ table 4.9 does mean that the positions of transactions and flows between the households (4.7) and NPISH institutional sector accounts (4.8) don't cancel each other out.

On the other hand: "Consolidation involves the elimination of those transactions or debtor or creditor relationships that occur between two institutional units belonging to the same institutional sector or subsector" (2008 SNA, paragraph 22.14); it is "a method of presenting statistics for a set of units as if they constituted a single unit" (2008 SNA, paragraph 22.79).

be available until the SNA indicator "net lending" of the capital account in order to count as available for the 2008 MRDS assessment. It is important to specify this indicator in the formula as "net lending of the capital account" (NL). The "net lending" indicator of the financial account is derived separately from the financial indicators. The requirement "until net lending" (United Nations 2011, 25) implies the required existence of the sequence of accounts subsequently derived from GDP until the closing balance, i.e. the final SNA indicator, of the capital account.

To determine the production of data until "net lending" for the institutional sector accounts, the existence of the "net lending" indicator of the capital account is sufficient. (For a detailed explanation of the sufficiency of any indicator given the required NAQ tables of the 1993 MRDS and of the specific indicator NL for the 2008 MRDS refer to part III, chapter 2.2.2 Availability Conditions of NAQ Tables, p. 203.)

Note that the currently used UN 1993 National Accounts Questionnaire (1993 NAQ) only covers annual data, not quarterly data. The 2008 MRDS also requires QNA data for some of the NAQ tables. The 2008 SNA introduced additional indicators[38], thus a future 2008 NAQ will have a different set of variables. Besides these differences the 1993 NAQ covers the requirement NAQ tables[39] for both MRDS. All variables (NA indicators) to each NAQ table can be obtained from the currently used 1993 NAQ.

The 1993 and 2008 MRDS and the corresponding sets $MRDS_{1993}$ and $MRDS_{2008}$ are formally delineated in the following sections 3.2.1.1 and 3.2.1.2.

[38] The NA indicators differs at times for the 2008 SNA since some items (e.g. financial instruments) are renamed or presented in more detail. For the NA indicators of the 2008 SNA refer to United Nations (2009f).

[39] In the Excel-version of the NAQ, as well as in some other documents, the "dot" in the NAQ table number x.y is replaced by "0", thus reading e.g. "102" instead of "1.2"

3.2.1.1 1993 Minimum Requirement Data Set

DEF. 25: 1993 Minimum Requirement Data Set

The 1993 MRDS reflects the set of possible combination of NAQ tables $T_{x.y}$ as objects, considered the minimum requirement in case of the 1993 SNA as per United Nations (2001b), it is given by the existence (requirement) of all of the NAQ tables 1.1, 1.2, 2.1, 2.2, 2.3, 4.2 and at least one of the NAQ tables 4.1 or 1.3.

The 1993 MRDS is formally delineated as follows:

Let the set of potential NAQ tables be:

$$\mathcal{T}_{1993} := \{T_{1.1}, T_{1.2}, T_{1.3}, T_{2.1}, T_{2.2}, T_{2.3}, T_{4.1}, T_{4.2}\}.$$

Only the existence of a NAQ table is of interest, indicated by the value of the indicator variable $t_{x.y}$. That is

$$t_{x.y} = \begin{cases} 1 & \text{if NAQ table } T_{x.y} \text{ exists (is a requirement)} \\ 0 & \text{else.} \end{cases}$$

Now we define the superset $\mathcal{S}_{1993}$ of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $T_{x.y} \in \mathcal{T}_{1993}$:

$$\mathcal{S}_{1993} := \{s = (t_{1.1}, t_{1.2}, t_{1.3}, t_{2.1}, t_{2.2}, t_{2.3}, t_{4.1}, t_{4.2}) \mid t_{x.y} \in \{0,1\}\} = \{0,1\}^8.$$

As relevant subset of state vectors we define the set

$$\text{MRDS}_{1993} := \{ s \in \mathcal{S}_{1993} \mid t_{1.1} * t_{1.2} * t_{1.3} * t_{2.1} * t_{2.2} * t_{2.3} * t_{4.2} = 1 \vee t_{1.1} * t_{1.2} * t_{2.1} * t_{2.2} * t_{2.3} * t_{4.1} * t_{4.2} = 1\}.$$ [40, 41]

Source: own definition based SNA data sets table in United Nations (2001b, 10.11).

For example: state vector s = (0,0,0,0,0,0,0,0) $\in \mathcal{S}_{1993}$ indicates the non-existence of any tables in $\mathcal{T}_{1993}$, s = (1,1,1,1,1,1,1,1) $\in \mathcal{S}_{1993}$ the existence of all tables in $\mathcal{T}_{1993}$, and

[40] That is all factors $t_{x.y}$ in these products are =1. That does mean, s $\in \boldsymbol{S}_{1993}$ signals that all of the components $t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}, t_{2.3}, t_{4.2}$ and at least one of the components $t_{1.3}$ and $t_{4.1}$ have to be 1.

[41] Where "∨" symbolizes the "logical or".

s = (1,0,0,1,1,0,1,1) ∈ S_{1993} the existence of the tables $T_{1.1}$, $T_{2.1}$, $T_{2.2}$, $T_{4.1}$, $T_{4.2}$ and the missing of $T_{1.2}$, $T_{1.3}$, $T_{2.3}$.

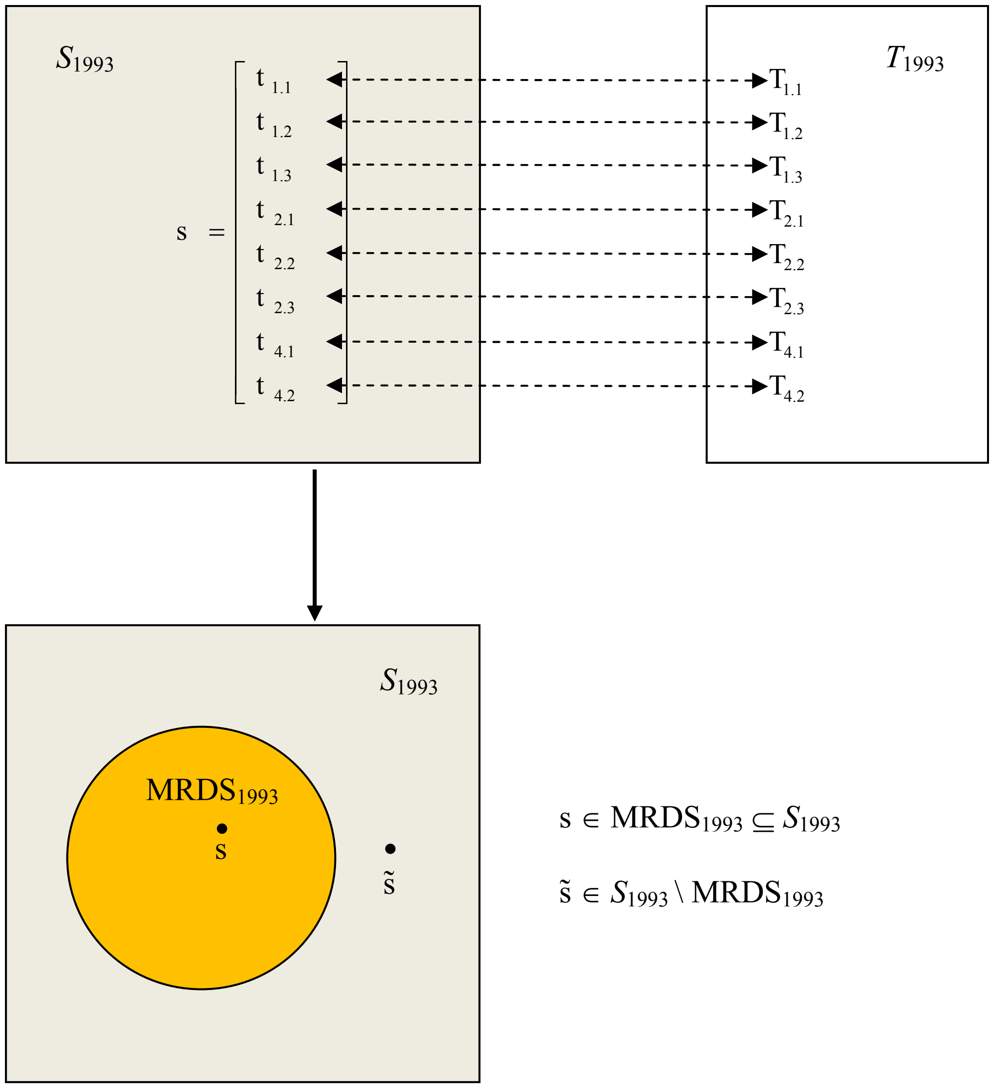


Figure 11 Relation between Set of Tables T_{1993} and State Vector s, and between Set of State Vectors S_{1993} and Set $MRDS_{1993}$

Figure 11 illustrates exemplarily the relations between T_{1993} - the set of NAQ tables that are elements of the state vector s - and the set S_{1993} - the superset of state vectors s. Further, the set $MRDS_{1993}$ represents the set of state vectors s ∈ $MRDS_{1993} \subseteq S_{1993}$ that fulfill the 1993 MRDS criteria. The state vector s̃ ∈ $S_{1993} \setminus MRDS_{1993}$ does not fulfill the minimum requirement criteria.

The relations illustrated above apply analogous for the NAQ table sets T, supersets of state vectors S, and subsets of relevant state vectors defined in the following.

1993 MRDS Fulfillment by the Available Country Data

To describe whether or not a country $c \in C$ fulfills the 1993 MRDS criteria we define the mapping $\mathbf{F}_{1993} : C \rightarrow S_{1993}$ from the domain C of considered countries to the codomain $S_{1993} = \{0,1\}^8$ with

$$c \in C \mapsto \mathbf{F}_{1993}(c) = (t_{1.1}, t_{1.2}, t_{1.3}, t_{2.1}, t_{2.2}, t_{2.3}, t_{4.1}, t_{4.2}) \qquad \text{(II.20)}$$

and

$$t_{x.y} = F_{x.y}(c) = \begin{cases} 1, \text{ if NAQ table } T_{x.y} \text{ is available}^{42} \text{ for } c \in C, \text{ and} \\ 0 \text{ else.} \end{cases}$$

The range of $\mathbf{F}_{1993}$ is the set $\mathbf{F}_{1993}(c) = \{ \mathbf{F}_{1993}(c) \mid c \in C \}$. With other words, the vectorial mapping $\mathbf{F}_{1993}(c)$ indicates which of the NAQ tables of T_{1993} are available in c (that is produced by c) and which are not.

Thus, $\mathbf{F}_{1993}(c) = s \in \text{MRDS}_{1993}$ indicates for a specific country $c \in C$ that it fulfills the criteria of 1993 MRDS. Then the countries meeting the 1993 MRDS criteria are represented as the members of the preimage set of $\text{MRDS}_{1993} \subset S_{1993}$, i.e.

$$C_{1993} := \mathbf{F}_{1993}^{-1}(\text{MRDS}_{1993}) = \{c \in C \mid \mathbf{F}_{1993}(c) \in \text{MRDS}_{1993}\}. \qquad \text{(II.21)}$$

Further we define the indicator function:

$$I_{1993} : C \rightarrow \{0,1\} \text{ with} \qquad \text{(II.22)}$$

$$I_{1993}(c) = \begin{cases} 1 \text{ if } c \in C_{1993} \\ 0 \text{ else.} \end{cases}$$

Thus we have

$$C_{1993} = I_{1993}^{-1}(1). \qquad \text{(II.23)}$$

As the result of interest for the assessment by regions and country groups, we calculate the number of countries within a region or country group that fulfill the

[42] Availability refers to the nationally produced data, measured by the data transmitted by a national agency to the UN and made available through publication.

criteria of the 1993MRDS. Thus we are interested in the cardinality of C_{1993}, which is the sum of I_{1993}(c) over all c ∈ C:

$$\left|C_{1993}\right| = \sum_{c \in C} \mathrm{I}_{1993}(\mathrm{c}) . \tag{II.24}$$

Corresponding to the cardinality of C_{1993}, the share h_{1993} of countries that fullfill the 1993 MRDS in relation to the total number of countries C is calucualted by

$$\mathrm{h}_{1993} := \frac{\left|C_{1993}\right|}{\left|C\right|}, \qquad \text{(with } 0 \leq \mathrm{h}_{1993} \leq 1\text{).} \tag{II.25}$$

There are only three n-tuples (state vectors s) which comprise the set $MRDS_{1993}$. In Table 4 the state vectors s_1 , s_2 , s_3 are the row vectors of the availability matrix. Table 4 reflects the combinations of produced NAQ tables $T_{x.y}$ of a country c ∈ C fulfilling the 1993 MRDS criteria. The availability (meaning production i.e. existence respectively non-existence)[42] of each NAQ table $T_{x.y}$ is indicated by $F_{x.y}$(c).

Table 4 Availability Matrix of $T_{x.y} \in T_{1993}$ for the Cases where F_{1993}(c) Fulfills the 1993 MRDS

$\mathbf{F}_{1993}$(c) = s reflecting the NAQ table availability indicated by $t_{x.y} = F_{x.y}(c)$ for $T_{x.y}$ for each of the possible cases for s ∈ $MRDS_{1993}$ for any country c ∈ C.								
$t_{x.y}$ / $\mathbf{F}_{1993}$(c)	$t_{1.1}$	$t_{1.2}$	$t_{1.3}$	$t_{2.1}$	$t_{2.2}$	$t_{2.3}$	$t_{4.1}$	$t_{4.2}$
s_1	1	1	1	1	1	1	1	1
s_2	1	1	1	1	1	1	0	1
s_3	1	1	0	1	1	1	1	1
$t_{x.y}$: indicator variable for the existence and non-existence, respectively, of $T_{x.y}$. $\mathbf{F}_{1993}$(c): the image of c ∈ C under the map $\mathbf{F}_{1993} : C \rightarrow S_{1993}$ indicating for cases s_1, …, s_3 which NAQ tables are available or not.								

Source: own compilation

The following examples show the results of the function $\mathbf{F}_{1993}$(c) and I_{1993}(c) under different scenarios of data availability for a specific country c ∈ C. The first three scenarios correspond to the row vectors s_i (i = 1, 2, 3) in Table 4, these render the exhaustive enumeration of the cases for which the function I_{1993}(c) = 1. Further the next three exemplary scenarios (4 to 6) cover the complete cases where the data availability of a country $c \in C$ does not fulfill the 1993 MRDS criteria. (For the other availability measures only an example and counter-example are given.)

Example: The fulfillment of the 1993 MRDS by a specific country $c \in C$ is assessed. The first three scenarios fulfill the set of NAQ tables required by the 1993 MRDS.

Scenario 1: All NAQ tables that count towards the fulfillment of the 1993 MRDS are available for the assessed country c.
$\mathbf{F}_{1993}(c) = (1,1,1,1,1,1,1,1) = s_1 \in MRDS_{1993}$, that is $I_{1993}(c) = 1$.

Scenario 2: Only NAQ table 4.1 is missing, but 1.3 is available.
$\mathbf{F}_{1993}(c) = (1,1,1,1,1,1,0,1) = s_2 \in MRDS_{1993}$, that is $I_{1993}(c) = 1$.

Scenario 3: Only NAQ table 1.3 is missing, but 4.1 is available.
$\mathbf{F}_{1993}(c) = (1,1,0,1,1,1,1,1) = s_3 \in MRDS_{1993}$, that is $I_{1993}(c) = 1$.

For all other cases, the availability (production) of the required set of NAQ tables is not fulfilled, i.e. for any other n-tuple $\mathbf{F}_{1993}(c) \in S_{1993}$ identifying the NAQ table availability is $I_{1993}(c) = 0$. This is illustrated by the following exemplary scenarios.

Scenario 4: Both NAQ tables 1.3 and 4.1 are missing.
$\mathbf{F}_{1993}(c) = (1,1,0,1,1,1,0,1) = s \notin MRDS_{1993}$, that is $I_{1993}(c) = 0$.

Scenario 5: NAQ table 1.1 is missing.
$\mathbf{F}_{1993}(c) = (0,1,1,1,1,1,1,1) = s \notin MRDS_{1993}$, that is $I_{1993}(c) = 0$.

Scenario 6 (similar to scenario 5): Any one or more than one of the NAQ tables 1.1, 1.2, 2.1, 2.2, 2.3, or 4.2 are missing.
For $c \in C$ with:

$\mathbf{F}_{1993}(c) = (t_{1.1}, t_{1.2}, t_{1.3}, t_{2.1}, t_{2.2}, t_{2.3}, t_{4.1}, t_{4.2}) \notin MRDS_{1993}$ because $t_{x.y} = F_{x.y} = 0$ for at least one of the indicated indices 1.1, 1.2, 2.1, 2.2, 2.3, or 4.2 and therefore
$F_{1.1}(c) * F_{1.2}(c) * F_{2.1}(c) * F_{2.2}(c) * F_{2.3}(c) * F_{4.2}(c) = 0$.

Therefore $I_{1993}(c) = 0$.

In such case the non-fulfillment of the 1993 MRDS does not depend on the non-existence of $T_{1.3}$ and $T_{4.1}$ indicated by $t_{1.3}$ and $t_{4.1}$, respectively.

3.2.1.2 2008 Minimum Requirement Data Set

DEF. 26: 2008 Minimum Requirement Data Set

The 2008 MRDS reflects the set of possible combination of NAQ tables $T_{x.y}$ as objects considered the minimum requirement in case of the 2008 SNA as per United Nations (2011), it is given by the existence (requirement) of all of the NAQ tables 1.1, 1.2, 2.1, 2.2, 2.3, 4.2 4.3, 4.4, 4.5, further at least one of the tables 1.3 or 4.1, and either both tables 4.6 and 4.7 or table 4.9. The NAQ tables 1.3, 4.1 to 4.7, and 4.9 need to be available until the NA indicator "net lending" of the capital account.

The 2008 MRDS is formally delineated as follows:

Let the set of potential NAQ tables be:

$T_{2008} := \{T_{1.1}, T_{1.2}, T_{1.3}^{nl}, T_{2.1}, T_{2.2}, T_{2.3}, T_{4.1}^{nl}, T_{4.2}^{nl}, T_{4.3}^{nl}, T_{4.4}^{nl}, T_{4.5}^{nl}, T_{4.6}^{nl}, T_{4.7}^{nl}, T_{4.9}^{nl}\}$, where nl indicates the NAQ table needs to be existent until the "net lending" indicator.

The existence of a NAQ table is indicated by the value of $t_{x.y}$. That is

$$t_{x.y} = \begin{cases} 1 \text{ if NAQ table } T_{x.y} \text{ respectively } T_{x.y}^{nl} \text{ exists (is a requirement)} \\ 0 \text{ else.} \end{cases}$$

Now we define the superset S_{2008} of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $T_{x.y}$ respectively $T_{x.y}^{nl} \in T_{2008}$:

$$S_{2008} := \{s = (t_{1.1}, t_{1.2}, t_{1.3}, t_{2.1}, t_{2.2}, t_{2.3}, t_{4.1}, t_{4.2}, t_{4.3}, t_{4.4}, t_{4.5}, t_{4.6}, t_{4.7}, t_{4.9}) \mid t_{x.y} \in \{0,1\}\} = \{0,1\}^{14}.$$

As relevant subset of state vectors we define the set

$$\begin{aligned} MRDS_{2008} := \{ s \in S_{2008} \mid\ & t_{1.1} \cdot t_{1.2} \cdot t_{1.3} \cdot t_{2.1} \cdot t_{2.2} \cdot t_{2.3} \cdot t_{4.2} \cdot t_{4.3} \cdot t_{4.4} \cdot t_{4.5} \cdot t_{4.6} \cdot t_{4.7} = 1 \\ \vee\ & t_{1.1} \cdot t_{1.2} \cdot t_{1.3} \cdot t_{2.1} \cdot t_{2.2} \cdot t_{2.3} \cdot t_{4.2} \cdot t_{4.3} \cdot t_{4.4} \cdot t_{4.5} \cdot t_{4.9} = 1 \\ \vee\ & t_{1.1} \cdot t_{1.2} \cdot t_{2.1} \cdot t_{2.2} \cdot t_{2.3} \cdot t_{4.1} \cdot t_{4.2} \cdot t_{4.3} \cdot t_{4.4} \cdot t_{4.5} \cdot t_{4.6} \cdot t_{4.7} = 1 \\ \vee\ & t_{1.1} \cdot t_{1.2} \cdot t_{2.1} \cdot t_{2.2} \cdot t_{2.3} \cdot t_{4.1} \cdot t_{4.2} \cdot t_{4.3} \cdot t_{4.4} \cdot t_{4.5} \cdot t_{4.9} = 1\}. \end{aligned}$$ [43]

Source: own definition based on SNA data sets table in United Nations (2011, 25).

[43] That is all factors $t_{x.y}$ in these products are =1. That does mean $s \in \boldsymbol{S}_{2008}$ signals that all components $t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}, t_{2.3}, t_{4.2}, t_{4.3}, t_{4.4}, t_{4.5}$, at least one of $t_{1.3}$ and $t_{4.1}$, and both $t_{4.6}$ and $t_{4.7}$ or $t_{4.9}$ have to be 1.

For example: s = (0,0,0,0,0,0,0,0,0,0,0,0,0,0) $\in S_{2008}$ indicates the non-existence of any tables in T_{2008}, s = (1,1,1,1,1,1,1,1,1,1,1,1,1,1) $\in S_{2008}$ the existence of all tables in T_{2008}, and s = (1,0,1,1,0,0,1,1,0,0,1,1,0,0) $\in S_{2008}$ the existence of the tables $T_{1.1}, T^{nl}_{1.3}, T_{2.1}, T^{nl}_{4.1}, T^{nl}_{4.2}, T^{nl}_{4.5}, T^{nl}_{4.6}$ and the missing of $T_{1.2}, T_{2.2}, T_{2.3}, T^{nl}_{4.3}, T^{nl}_{4.4}, T^{nl}_{4.7}, T^{nl}_{4.9}$ (with the existence of $T_{x.y}$ until the net lending indicator of the capital account symbolized by $T^{nl}_{x.y}$).

Building upon the superset S_{2008} a projective mapping $\Pi_{1993} : S_{2008} \rightarrow S_{1993}$ could have been defined, which for a given s = $(t_{1.1}, t_{1.2}, ..., t_{4.2}, ..., t_{4.7}, t_{4.9}) \in S_{2008}$ gives a $\tilde{s} = \Pi_{1993}(s) := (t_{1.1}, t_{1.2}, ..., t_{4.2}) \in S_{1993}$. That does mean that $\tilde{s} = \Pi_{1993}(s)$ is the partial n-tuple of s that is relevant for the 1993 MRDS.

The problem of this approach is the inconsistence resulting from the requirements regarding the completeness of the individual tables $T_{x.y}$, which are relevant for both, the 1993 and 2008 MRDS[44]. For example in case of the 2008 MRDS we have $T^{nl}_{1.3}$, i.e. the NAQ table is required until the closing balance "net lending" of the capital account, while this is not required in case of the 1993 MRDS.

Moreover, using two individual sets, meaning $S_{1993} = \{0,1\}^8$ and $S_{2008} = \{0,1\}^{14}$, is superior to basing both MRDS measures on the superset S_{2008}. This is because using individual sets embraces the possibility of extensions to future versions of the MRDS (e.g. the requirement of financial accounts for all institutional sectors [NAQ tables 4.1 to 4.9]). An advantage, for mathematical purposes, of basing both MRDS measure on a common superset would be the direct comparability of the n-tuples which indicate the NAQ table availability for a country c $\in C$. However, we observe the results by the preimage sets of countries that fulfill the availability measure of interest.

[44] The 2008 MRDS requires the existence of $T_{x.y}$ until the net lending indicator of the capital account, while in case of the 1993 MRDS the mere availability of any data for the NAQ table $T_{x.y}$ is sufficient.

2008 MRDS Fulfillment by the Available Country Data

To describe whether or not a country $c \in C$ fulfills the 2008 MRDS criteria we define the mapping $\mathbf{F}_{2008} : C \rightarrow S_{2008}$ with $S_{2008} = \{0,1\}^{14}$,

$$c \in C \mapsto \mathbf{F}_{2008}(c) = \left(t_{1.1}, t_{1.2}, t_{1.3}, t_{2.1}, t_{2.2}, t_{2.3}, t_{4.1}, t_{4.2}, t_{4.3}, t_{4.4}, t_{4.5}, t_{4.6}, t_{4.7}, t_{4.9}\right) \quad \text{(II.26)}$$

and

$$t_{x.y} = F_{x.y}(c) = \begin{cases} 1 \text{ if } T_{x.y} \text{ respectively } T^{nl}_{x.y} \text{ is available}^{45} \text{ for } c \in C\text{, and} \\ 0 \text{ else.} \end{cases}$$

$F_{x.y}$ are the individual component functions of the vectorial mapping $\mathbf{F}_{2008}$. In particular the image vector $\mathbf{F}_{2008}(c)$ consists of the components $F_{x.y}(c) = t_{x.y} \in \{0,1\}$, where $T_{x.y} \in T_{2008}$, respectively $T^{nl}_{x.y} \in T_{2008}$.

With other words, $\mathbf{F}_{2008}$ indicates which of the tables of T_{2008} are available in c and which are not. Thus $\mathbf{F}_{2008}(c) = s \in MRDS_{2008}$ indicates for a specific country $c \in C$ that it fulfills the criteria of the 2008 MRDS. For the NAQ table $T^{nl}_{x.y}$ of T_{2008} the "nl" indicates the required availability until the "net lending" indicator of the capital account. Basic data for $T_{x.y}$ might be produced, yet not until the defined NA indicator. Thus, this is an extension of the 1993 MRDS requirements.

Then the countries meeting the 2008 MRDS criteria are given by members of the preimage set of $MRDS_{2008} \subset S_{2008}$, i.e.

$$C_{2008} := \mathbf{F}_{2008}{}^{-1}(MRDS_{2008}) = \{c \in C \mid \mathbf{F}_{2008}(c) \in MRDS_{2008}\}. \quad \text{(II.27)}$$

Further we define the indicator function:

$$I_{2008} : C \rightarrow \{0,1\} \text{ with } \quad I_{2008}(c) = \begin{cases} 1 \text{ if } c \in C_{2008} \\ 0 \text{ else.} \end{cases} \quad \text{(II.28)}$$

Thus we have

$$C_{2008} = I_{2008}{}^{-1}(1). \quad \text{(II.29)}$$

[45] Availability refers to the nationally produced data, measured by the data transmitted by a national agency to the UN and made available through publication.

We are interested in the number of countries that fulfill the requirement criteria of the 2008MRDS, i.e. the cardinality of C_{2008}, which is the sum of $\mathrm{I}_{2008}(\mathrm{c})$ over all $\mathrm{c} \in C$:

$$|C_{2008}| = \sum_{\mathrm{c} \in C} \mathrm{I}_{2008}(\mathrm{c}) . \tag{II.30}$$

The share h_{2008} is calculated analougously to Formula II.25, by the cardinality of C_{2008} in relation to the total number of countries C.

Table 5 Availability Matrix of $T_{x.y} \in T_{2008}$ for the Cases where $F_{2008}(c)$ Fulfills the 2008 MRDS

$\mathbf{F}_{2008}(\mathrm{c})$ = s reflecting the NAQ table availability indicated by $t_{x.y} = F_{x.y}(c)$ for $T_{x.y}$ respectively $T^{nl}_{x.y}$ for each of the possible cases for s $\in$ MRDS_{2008} for any country $c \in C$.

$\mathbf{F}_{2008}(c)$ \ $t_{x.y}$	$t_{1.1}$	$t_{1.2}$	$t_{1.3}$	$t_{2.1}$	$t_{2.2}$	$t_{2.3}$	$t_{4.1}$	$t_{4.2}$	$t_{4.3}$	$t_{4.4}$	$t_{4.5}$	$t_{4.6}$	$t_{4.7}$	$t_{4.9}$
s_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
s_2	1	1	1	1	1	1	1	1	1	1	1	1	1	**0**
s_3	1	1	1	1	1	1	1	1	1	1	1	1	**0**	1
s_4	1	1	1	1	1	1	1	1	1	1	1	**0**	1	1
s_5	1	1	1	1	1	1	1	1	1	1	1	**0**	**0**	1
s_6	1	1	1	1	1	1	**0**	1	1	1	1	1	1	1
s_7	1	1	1	1	1	1	**0**	1	1	1	1	1	1	**0**
s_8	1	1	1	1	1	1	**0**	1	1	1	1	1	**0**	1
s_9	1	1	1	1	1	1	**0**	1	1	1	1	**0**	1	1
s_{10}	1	1	1	1	1	1	**0**	1	1	1	1	**0**	**0**	1
s_{11}	1	1	0	1	1	1	1	1	1	1	1	1	1	1
s_{12}	1	1	0	1	1	1	1	1	1	1	1	1	1	**0**
s_{13}	1	1	0	1	1	1	1	1	1	1	1	1	**0**	1
s_{14}	1	1	0	1	1	1	1	1	1	1	1	**0**	1	1
s_{15}	1	1	0	1	1	1	1	1	1	1	1	**0**	**0**	1

$t_{x.y}$: indicator variable for the existence and non-existence, respectively, of $T_{x.y}$ respectively $T^{nl}_{x.y}$.

$\mathbf{F}_{2008}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{2008} : C \rightarrow S_{2008}$ indicating for cases $s_1, \ldots, s_{15}$ which tables are available or not. (Legend also applicable to all following availability matrices.)

Source: own compilation

Table 5 reflects the combinations of produced NAQ tables $T_{x.y}$ respectively $T^{nl}_{x.y}$ of a country $c \in C$ for each possible $\mathbf{F}_{2008}(c) = s \in MRDS_{2008}$. The availability respectively missing of each NAQ table $T_{x.y}$ respectively $T^{nl}_{x.y}$ in $c \in C$ is indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{2008}(c)$ for $T_{x.y}$ respectively $T^{nl}_{x.y}$. (For easier visibility of the differences between the row vectors the "0" are printed in bold font.)

Example: The fulfillment of the 2008 MRDS criteria is assessed analogous to the 1993 MRDS measure. Only one example scenario for the fulfillment and two for the non-fulfillment are provided for the 2008 MRDS measure, since a comprehensive example is provided for the 1993 MRDS (cf. chapter 3.2.1.1, p. 59). The scenarios show the results of the function $\mathbf{F}_{2008}(c)$ and $I_{2008}(c)$ for a specific country $c \in C$.

Scenario 1: Both NAQ tables 4.6 and 4.7 are missing, but 4.9 is available.

$\mathbf{F}_{2008}(c) = (1,1,1,1,1,1,1,1,1,1,1,0,0,1) = s_5 \in MRDS_{2008}$, that is $I_{2008}(c) = 1$.

Scenario 2: NAQ table 4.6, 4.7, and 4.9 are missing.

$\mathbf{F}_{2008}(c) = (1,1,1,1,1,1,1,1,1,1,1,0,0,0) = s \notin MRDS_{2008}$, that is $I_{2008}(c) = 0$.

Scenario 3: Here any one or more than one of the NAQ tables 1.1, 1.2, 2.1, 2.2, 2.3, 4.2, 4.3, 4.4, or 4.5 are missing.
For $c \in C$ with:

$\mathbf{F}_{2008}(c) = \left(t_{1.1},t_{1.2},t_{1.3},t_{2.1},t_{2.2},t_{2.3},t_{4.1},t_{4.2},t_{4.3},t_{4.4},t_{4.5},t_{4.6},t_{4.7},t_{4.9}\right) \notin MRDS_{2008}$ because $t_{x.y} = F_{x.y} = 0$ for at least one of the indicated indices 1.1, 1.2, 2.1, 2.2, 2.3, 4.2, 4.3, 4.4, or 4.5, and therefore
$F_{1.1}(c) * F_{1.2}(c) * F_{2.1}(c) * F_{2.2}(c) * F_{2.3}(c) * F_{4.2}(c) * F_{4.3}(c) * F_{4.4}(c) * F_{4.5}(c) = 0$.

Therefore $I_{2008}(c) = 0$.

In such case the non-fulfillment of the 2008 MRDS does not depend on the non-availability of $T^{nl}_{1.3}, T^{nl}_{4.1}, T^{nl}_{4.6}, T^{nl}_{4.7}$ and $T^{nl}_{4.9}$ indicated by $t_{1.3}, t_{4.1}, t_{4.6}, t_{4.7}$ and $t_{4.9}$, respectively.

3.2.1.3 Partial 1993 and 2008 Minimum Requirement Data Set

In addition to the previously defined availability measures the NA data availability assessment also considers the cases where a country fulfills at least a substantial part of the 1993 and 2008 MRDS, i.e. at least all but one of the required criteria.

As previously stated, 1993 MRDS is a part of 2008 MRDS. Figure 12 illustrates the preimage sets of countries that fulfill 1993 and 2008 MRDS or at least a substantial part. The formal expression of the measures is provided in the subsequent sections. The color palette follows the categories of "heat", meaning the intensity of the colors indicate the intensity of the requirement criteria regarding the availability of NAQ tables. Hence, regarding fulfillment of these requirement criteria, the color intensity indicates the increased ability in NA production of the members of the country sets.

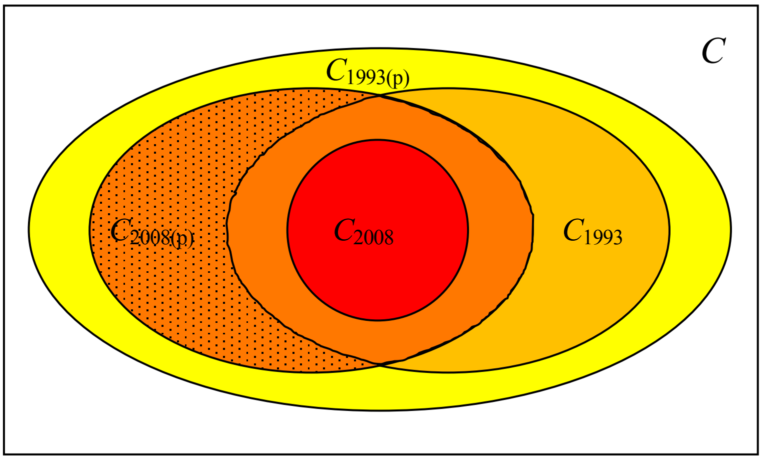


Figure 12 Country Sets Relating to the Fulfillment of MRDS Requirements

Source: own illustration

Explanation of the colored sets in Figure 12.

	<brick red>	$C_{2008} = \mathbf{F}_{2008}^{-1}(MRDS_{2008})$
	<orange-red, dots>	$C_{2008(P)} \setminus C_{1993}$, set of countries missing exactly one criterion of the 1993 MRDS (yet fulfilling all additional requirements of the 2008 MRDS and thus falling into the set $C_{2008(P)}$.
	<orange-red>	$(C_{2008(P)} \cap C_{1993}) \setminus C_{2008}$, countries fulfilling the 1993 MRDS criterion, but missing exactly one criterion of the 2008 MRDS.
	<orange-yellow>	$C_{1993} \setminus C_{2008(P)}$, set of countries fulfilling all 1993 MRDS criteria but missing at least two of the additional 2008 MRDS criteria.
	<yellow>	$C_{1993(P)} \setminus (C_{2008(P)} \cup C_{1993})$, set of countries missing exactly one criterion of the 1993 MRDS and further at least two of the additional 2008 MRDS criteria.

$C_{1993(P)}$ and $C_{2008(P)}$: set of countries that fulfills at least partial 1993 and 2008 MRDS, respectively, i.e. missing at most one NAQ table to fulfill the criteria.

Thus we have $C_{2008} \subseteq C_{1993} \subseteq C_{1993(P)}$. (II.31)

And $C_{2008} \subseteq C_{2008(P)} \subseteq C_{1993(P)}$. (II.32)

And, in a strict mathematical sense, $C_{2008(P)} \not\subseteq C_{1993}$. (II.33)

With other words, a country that fulfills the 2008 MRDS requirements already fulfills the lower requirements of the 1993 MRDS. Fulfilling the 2008 MRDS also includes the measure indicating the fulfillment of both, a substantial part of the 2008 MRDS and a substantial part of the 1993 MRDS.

The last formula does mean that a country that produces all but one of the NAQ tables required to fulfill the 2008 MRDS does not necessarily produce all NAQ tables required by the 1993 MRDS. This is because the missing NAQ table might belong to the basic NA data covered by the 1993 MRDS (for example one of the NAQ tables 1.2 or 2.2 for constant price data), while the advanced NAQ tables i.e. in particular the institutional sectors (4.1 to 4.7 and 4.9) required by the 2008 MRDS are produced.

For the fulfillment of at least a substantial part of the 2008 MRDS we define the distance d_{2008}, to measure the deviation of n-tuple s inside the superset S_{2008} to the subset of relevant n-tuples given by $MRDS_{2008}$. For $MRDS_{1993}$ (the relevant subset) and n-tuple s inside the superset $S_{1993,}$ in the same way by distance d_{1993}.

Let $MRDS_{2008(P)}$ be the set of relevant n-tuples s that fulfill at least a substantial part of the 2008 MRDS, i.e. the n-tuple $s \in MRDS_{2008(P)}$ only differ by the non-existence of maximal one needed NAQ table of the n-tuple $s \in MRDS_{2008.}$ Analogous for $MRDS_{1993(P)}$ and $MRDS_{1993}$.

Now, let $X.Y_{2008}$ be the index set regarding the NAQ tables $T_{x.y} \in T_{2008}$ respectively $T^{nl}_{x.y} \in T_{2008}$.

$$X.Y_{2008} = \{ x.y \mid T_{x.y} \in T_{2008} \vee T^{nl}_{x.y} \in T_{2008} \} = \{1.1, 1.2, 1.3, 2.1, 2.2, 2.3, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.9\} \quad \text{(II.34)}$$

and let $X.Y_{1993,}$ defined as

$$X.Y_{1993} = \{ \text{x.y} \mid \text{T}_{\text{x.y}} \in T_{1993} \} = \{ 1.1 , 1.2 , 1.3 , 2.1, 2.2 , 2.3 , 4.1 , 4.2 \} \qquad \text{(II.35)}$$

be the index set regarding the NAQ tables $\text{T}_{\text{x.y}} \in T_{1993.}$

Then the distance between two n-tuples $s = (t_{1.1},..,t_{4.9})$, $\tilde{s} = (\tilde{t}_{1.1},..,\tilde{t}_{4.9}) \in S_{2008}$ is given by:

$$\text{d}_{2008}(s,\tilde{s}) = \sum_{\text{x.y} \in X.Y_{2008}} \left| \text{t}_{\text{x.y}} - \tilde{\text{t}}_{\text{x.y}} \right| \qquad \text{(II.36)}$$

respectively for $s = (t_{1.1},..,t_{4.2})$, $\tilde{s} = (\tilde{t}_{1.1},..,\tilde{t}_{4.2}) \in S_{1993}$ by

$$\text{d}_{1993}(s,\tilde{s}) = \sum_{\text{x.y} \in X.Y_{1993}} \left| \text{t}_{\text{x.y}} - \tilde{\text{t}}_{\text{x.y}} \right| \qquad \text{(II.37)}$$

with d_{2008} and d_{1993} applied to the 2008 and 1993 MRDS, respectively.

Consequently, if for $s \in S_{2008} \setminus \text{MRDS}_{2008}$ there exists a state vector $\tilde{s} \in \text{MRDS}_{2008}$ with $\text{d}_{2008}(s,\tilde{s}) = 1$, it does mean that in this case exactly one of the requirements is missing to fulfill the MRDS_{2008} criteria. With other words, a country $c \in C$ fulfills at least a substantial part of the 2008 MRDS by its available (produced) set $s = \mathbf{F}_{2008}(c)$ of NAQ tables $\text{T}_{\text{x.y}} \in T_{2008}$ respectively $\text{T}^{\text{nl}}_{\text{x.y}} \in T_{2008}$, if there exists a $\tilde{s} \in \text{MRDS}_{2008}$ with $\text{d}_{2008}(s,\tilde{s}) \leq 1$.

Analogous: If for $s \in S_{1993} \setminus \text{MRDS}_{1993}$ there exists a state vector $\tilde{s} \in \text{MRDS}_{1993}$ with $\text{d}_{1993}(s,\tilde{s}) = 1$, it does mean that in this case exactly one of the requirements is missing to fulfill the MRDS_{1993} criteria. With other words, a country $c \in C$ fulfills at least a substantial part of the 1993 MRDS by its available (produced) set $s = \mathbf{F}_{1993}(c)$ of NAQ tables $\text{T}_{\text{x.y}} \in T_{1993}$, if there exists a $\tilde{s} \in \text{MRDS}_{1993}$ with $\text{d}_{1993}(s,\tilde{s}) \leq 1$.

Example: For a state vector $s = (1,0,1,1,1,1,1,1,1,1,1,1,1,0) \in S_{2008} \setminus \text{MRDS}_{2008}$ - that is $s \notin \text{MRDS}_{2008}$ - and $\tilde{s} = (1,1,1,1,1,1,1,1,1,1,1,1,1,0) \in \text{MRDS}_{2008}$ the distance is:

$$\text{d}_{2008}(s,\tilde{s}) = \sum_{\text{x.y} \in X.Y_{2008}} \left| \text{t}_{\text{x.y}} - \tilde{\text{t}}_{\text{x.y}} \right| = 1.$$

This does mean that for $s \in S_{2008} \setminus \mathrm{MRDS}_{2008}$ exactly one requirement is missing to fulfill the MRDS_{2008} criteria. Thus a country $c \in C$ with $\mathbf{F}_{2008}(c) = s = (1,0,1,1,1,1,1,1,1,1,1,1,0)$ fulfilled a substantial part of the 2008 MRDS (the here missing requirement is NAQ table $T_{1.2}$).

On the other hand, for state vector $s = (1,1,1,1,1,1,0,0,1,1,1,0,0,0) \in S_{2008} \setminus \mathrm{MRDS}_{2008}$ there exists no state vector $\tilde{s} \in \mathrm{MRDS}_{2008}$ with $d_{2008}(s,\tilde{s}) = 1$. Within MRDS_{2008} the state vector with the minimum distance to s is $\tilde{s} = (1,1,1,1,1,1,0,1,1,1,1,0,0,1)$, giving $d_{2008}(s,\tilde{s}) = 2$. Therefore, for any $c \in C$ with $\mathbf{F}_{2008}(c) = s$, already two of the requirements to fulfill the MRDS_{1993} criteria (here NAQ tables $T_{4.2}$ and $T_{4.9}$) are missing.

Via $d_{2008}(\,.\,,\,.\,)$, we get once more the description of the set $C_{2008(p)}$. Thus we observe

$$\mathrm{MRDS}_{2008(p)} = \mathrm{MRDS}_{2008} \cup \{s \in S_{2008} \setminus \mathrm{MRDS}_{2008} \mid \exists\, \tilde{s} \in \mathrm{MRDS}_{2008} : d_{2008}(s,\tilde{s}) = 1\} = \mathrm{MRDS}_{2008(p)} = \{\tilde{s} \in \mathrm{MRDS}_{2008} : d(s,\tilde{s}) \leq 1\}. \quad \text{(II.38)}$$

By this view we get

$$C_{2008(p)} := \mathbf{F}_{2008}^{-1}(\mathrm{MRDS}_{2008(p)}) = C_{2008} \cup \{c \in C \setminus C_{2008} \mid \exists\, s \in \mathrm{MRDS}_{2008} : d_{2008}(\mathbf{F}_{2008}(c),s) = 1\}. \quad \text{(II.39)}$$

Further we can define the indicator function:

$$I_{2008(p)} : C \rightarrow \{0,1\} \text{ with } I_{2008(p)}(c) = \begin{cases} 1 \text{ if } c \in C_{2008(p)} \\ 0 \text{ else.} \end{cases} \quad \text{(II.40)}$$

Thus we have

$$C_{2008(p)} = I_{2008(p)}^{-1}(1). \quad \text{(II.41)}$$

As the result of interest for the assessment by regions and country groups, we calculate the number of countries within a region or country group that fulfill the substantial part of the requirement criteria of the 2008MRDS. Thus we are interested in the cardinality of $C_{2008(p)}$ that is the sum of $I_{2008(p)}(c)$ over all $c \in C$:

$$\left|C_{2008\ (p)}\right| = \sum_{c \in C} I_{2008(p)}(c)\,. \quad \text{(II.42)}$$

The share $h_{2008(p)}$ of countries that fullfill a substantial part of the 2008 MRDS in relation to the total number of countries *C* is calucualted by

$$h_{2008(p)} := \frac{|C_{2008(p)}|}{|C|}, \qquad \text{(with } 0 \leq h_{2008(p)} \leq 1\text{)} \qquad \text{(II.43)}$$

In analogous way (by replacing the index 2008 with 1993) we get similar formulas for the 1993 MRDS instead of the 2008 MRDS.

Sample Application 6: Minimum Requirement Data Set Measures

Example 1: Cameroon

Available NAQ Tables	**Missing NAQ Tables**
1.1, 1.2, 1.3, 2.1, 2.2, 2.3, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6 and 4.7. Specifically: 1.3 and 4.1 to 4.7 available until the NA indicator “net lending” of the capital account.	4.8[*/ **], 4.9[***]

[*] NAQ table 4.8 is not part of the two MRDS measures, however it is here considered with the NAQ table availability of purpose of comprehensiveness and for use in later sample applications.

[**] NAQ table 4.8 can be produced in addition or in lieu, i.e. as substitute for the combined existence of NAQ tables 4.3 and 4.4.

[***] NAQ table 4.9 can be produced in addition or in lieu, i.e. as substitute for the combined existence of NAQ tables 4.6 and 4.7.

From the available NAQ tables, we see that Cameroon (CMR) fulfills the 1993 MRDS and also the 2008 MRDS (as defined in DEF. 25, p. 59, respectively DEF. 26, p. 64) and necessarily also the partial MRDS measures (cf. chapter 3.2.1.3) i.e. a subset missing exactly one of the required MRDS criteria.

Referring to the formulas (II.20 ff. and II.26 ff.), of the 1993 MRDS we get $\mathbf{F}_{1993}$(CMR) = (1,1,1,1,1,1,1,1), respectively regarding the 2008 MRDS we get $\mathbf{F}_{2008}$(CMR) = (1,1,1,1,1,1,1,1,1,1,1,1,1,0).

(continued on next page)

Defined by $MRDS_{1993}$ and $MRDS_{2008}$ are the relevant sets of state vectors (within the superset of potential state vectors) that meet the requirement criteria of the MRDS measures.

Thus we have $\mathbf{F}_{1993}(CMR) \in MRDS_{1993}$. Consequently Cameroon is member of the preimage set C_{1993} of the set of countries that fulfill the 1993 MRDS criteria; for the indicator function signaling that a country fulfills this criteria we get $I_{1993}(CMR) = 1$. We also have $\mathbf{F}_{2008}(CMR) \in MRDS_{2008}$. Consequently Cameroon is member of the preimage set C_{2008} and $I_{2008}(CMR) = 1$.

Regarding the partial fulfillment of the MRDS measures (cf. Formula II.37), we get for the distance between the n-tuple s = $\mathbf{F}_{1993}(CMR)$ and state vector s $\in MRDS_{1993}$: $d_{1993}(s,s) = 0 \leq 1$, thus Cameroon is also member of the preimage set $C_{1993(p)}$, further $I_{1993(p)}(CMR) = 1$.
Analogue for the partial fulfillment of the 2008 MRDS measure evaluated by distance d_{2008} (cf. Formula II.36).

Example 2: Thailand

Available NAQ Tables	Missing NAQ Tables
1.1, 1.2, 1.3, 2.1, 2.2, 2.3. Specifically: 1.3 until "net lending" of the capital account.	4.1 - 4.7 and 4.9.

Thus Thailand (THA) does not fulfill the 1993 or 2008 MRDS. The 2008 MRDS is not fulfilled since the integrated accounts (4.1 to 4.7 and 4.9) until net lending are missing. The 1993 MRDS is not fulfilled since 4.2 is missing.

However, Thailand is member of the set of countries fulfilling a substantial part of the 1993 MRDS, since here only one criterion (table 4.2) is missing. Because NAQ tables 1.3 and 4.1 can be substituted for each other, the missing of 4.1 is compensated by the existence of 1.3.

(continued on next page)

Referring to the formulas (see Example 1), for the 1993 MRDS this equates to the vectorial mapping $\mathbf{F}_{1993}$(THA) = (1,1,1,1,1,1,0,0) $\notin$ $MRDS_{1993}$ and in case of the 2008 MRDS to $\mathbf{F}_{2008}$(THA) = (1,1,1,1,1,1,0,0,0,0,0,0,0,0) $\notin$ $MRDS_{2008}$. Consequently Thailand is not member of the preimage set C_{1993} respectively C_{2008}, further I_{1993}(THA) = 0 and I_{2008}(THA) = 0.

Regarding the fulfillment of a substantial part of the 2008 MRDS, we get for the distance between s = $\mathbf{F}_{2008}$(THA) and all state vectors $\tilde{s} \in MRDS_{2008}$ the distance: $d_{2008}(s,\tilde{s}) \geq 5 > 1$, thus Thailand is not in the set $C_{2008(p)}$, furthermore $I_{2008(p)}$(THA) = 0.

For s = $\mathbf{F}_{1993}$(THA) and state vector $\tilde{s}$ = (1,1,1,1,1,1,0,1) $\in MRDS_{1993}$ we get the distance $d_{1993}(s,\tilde{s}) = 1$, thus Thailand is member of the set $C_{1993(p)}$, further $I_{1993(p)}$(THA) = 1.

Exemplarily for the regional analysis, we calculate the numbers of countries in South-Eastern Asia that fulfill the criteria of the MRDS and partial MRDS measures.

The South-Eastern Asia region consists of eleven UN Member States, of which four countries fulfill a substantial part of the 1993 MRDS and two fulfill the 1993 MRDS. None of these countries fulfills the 2008 MRDS or a substantial part of it.

Referring to the formulas the cardinality of the country sets that fulfill the MRDS measures is:

Cardinality of Country Sets of South-Eastern Asia that Fulfill the MRDS Measures

C / **t**	$\boldsymbol{C_t}$**(p)**	$\boldsymbol{C_t}$
1993	4	2
2008	0	0

For example for $C_t(p)$ with t = 1993 the formula is

$$\left|C_{1993\ (p)}\right| = \sum_{c \in South-Eastern\ Asia} I_{1993(p)}\,(c) = 4.$$

(continued on next page)

Corresponding to the cardinalities of $C_{1993(p)}$, C_{1993}, $C_{2008(p)}$, C_{2008} the shares (h) in relation to the total number of countries in the South-Eastern Asia region that fullfill a particular MRDS measure is

Share (h) of Countries of South-Eastern Asia that Fulfill the MRDS Measures

t \ h	$h_t(p)$	h_t
1993	0,36	0,18
2008	0	0

For example for $h_t(p)$ with t = 1993 the formula is

$$h_{1993(p)} = \frac{\left|C_{1993\ (p)}\right|}{\left|South-Eastern\ \ Asia\right|} = \frac{4}{11} \qquad = 0{,}36.$$

3.2.2 Sets of Substitutional NAQ Tables

The 1993 and 2008 MRDS allow certain NAQ tables to be replaced by other NAQ tables in order to fulfill the MRDS criteria (since these NAQ tables cover similar or identical NA data). We assess the fulfillment of the criteria of such NAQ table sets. These sets are delineated in this subchapter. (An example for such substitutional NAQ tables is specified by the requirement of at least one of the NAQ tables 1.3 or 4.1, that does mean that NAQ table 1.3 can be replaced by NAQ table 4.1 and vice versa.)

The sets of substitutional NAQ tables are delineated similarly to the MRDS measures, i.e. they are given by the relevant state vectors that fulfill the requirement criteria. To each of the four sets of substitutional NAQ tables (see 1. to 4.), we define in the following the supersets S of s (n-tuples) indicating the existence respectively non-existence, of each NAQ table. Corresponding to the superset, we further define the set of relevant state vectors $RS \subset S$. Moreover, we define the vectorial maps **F**(c) for the assessment of the NAQ table availability for a country c $\in C$.

The referenced NAQ tables $T_{x.y}$ and $T^{nl}_{x.y}$ are elements of the NAQ table sets T_{1993} and T_{2008}, respectively. Like in the 2008 MRDS, the requirement until NL of the NAQ table is symbolized by $T_{x.y}^{nl}$. Like in the 1993 MRDS, for the NAQ tables $T_{x.y}$ without symbol no specific NA indicator is needed. The requirement and non-requirement (existence and non-existence) for $T_{x.y}$ is indicated by the indicator variable $t_{x.y}$.

1. Regarding the existence of at least one of the NAQ tables 1.3 or 4.1 ($T_{x.y} \in T_{1993}$), which are substitutional for the fulfillment of the 1993 MRDS, we define the superset $S_{1.3\vee4.1}$ and as the set of relevant state vectors $RS_{1.3\vee4.1}$:

$$S_{1.3\vee4.1} := \{s = (t_{1.3,} t_{4.1}) \mid t_{x.y} \in \{0,1\}\} = \{0,1\}^2. \qquad \text{(II.44)}$$

$$RS_{1.3\vee4.1} := \{s \in S_{1.3\vee4.1} \mid t_{1.3} = 1 \vee t_{4.1} = 1\}. \qquad \text{(II.45)}$$

Thus, $RS_{1.3\vee4.1}$ is the set of state vectors of those possible cases where $s \in RS_{1.3\vee4.1}$, i.e. where s indicates the existence of at least one of the substitutional NAQ tables.

$RS_{1.3\vee4.1} = \{s_1, s_2, s_3\}$ with s_i (i = 1, …, 3) is fulfilled in case of the NAQ table availability as given by the row vectors of the availability matrix in Table 6.

Further, we define the vectorial mapping $\mathbf{F}_{1.3\vee4.1} : C \rightarrow S_{1.3\vee4.1}$ with $S_{1.3\vee4.1} = \{0,1\}^2$,

$$c \in C \mapsto \mathbf{F}_{1.3\vee4.1}(c) = (t_{1.3,} t_{4.1}). \qquad \text{(II.46)}$$

$\mathbf{F}_{1.3\vee4.1}(c) = s$ is reflecting the NAQ table availability (production) for $c \in C$ indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{1.3\vee4.1}(c)$ for $T_{x.y.}$

Table 6 Availability Matrix of $T_{x.y}$ for the Cases that Fulfill the Existence of at Least one of the Substitutional NAQ Tables 1.3 or 4.1

$t_{x.y}$ / $\mathbf{F}_{1.3\vee4.1}(c)$	$t_{1.3}$	$t_{4.1}$
s_1	1	1
s_2	1	0
s_3	0	1
$\mathbf{F}_{1.3\vee4.1}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{1.3\vee4.1} : C \rightarrow S_{1.3\vee4.1}$		

Source: own compilation

2. Analogue to 1., yet for the 2008 MRDS regarding the existence of NAQ table 1.3 or 4.1 until NL ($T^{nl}_{x.y} \in T_{2008}$), we define the superset of all state vectors $S_{1.3\vee4.1(nl)}$ and $RS_{1.3\vee4.1(nl)}$ as the set of relevant state vectors:

$$S_{1.3\vee4.1(nl)} := \{s = (t_{1.3}, t_{4.1}) \mid t_{x.y} \in \{0,1\}\} = \{0,1\}^2. \qquad \text{(II.47)}$$

$$RS_{1.3\vee4.1(nl)} := \{s \in S_{1.3\vee4.1(nl)} \mid t_{1.3} = 1 \vee t_{4.1} = 1\}. \qquad \text{(II.48)}$$

$RS_{1.3\vee4.1(nl)} = \{s_1, s_2, s_3\}$ with s_i (i = 1, …, 3) is fulfilled in case of the NAQ table availability as given by the row vectors of the availability matrix in Table 7.

We define the vectorial mapping $\mathbf{F}_{1.3\vee4.1(nl)} : C \rightarrow S_{1.3\vee4.1(nl)}$ with $S_{1.3\vee4.1(nl)} = \{0,1\}^2$,

$$c \in C \mapsto \mathbf{F}_{1.3\vee4.1(nl)}(c) = (t_{1.3}, t_{4.1}). \qquad \text{(II.49)}$$

$\mathbf{F}_{1.3\vee4.1(nl)}(c) = s$ is reflecting the NAQ table availability (production) for $c \in C$ indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{1.3\vee4.1(nl)}(c)$ for $T^{nl}_{x.y}$.

Table 7 Availability Matrix of $T_{x.y}$ that Fulfill the Existence of at Least One of the Substitutional NAQ Table 1.3∨4.1 until Net Lending

$t_{x.y}$ / $\mathbf{F}_{1.3\vee4.1(nl)}(c)$	$t_{1.3}$	$t_{4.1}$
s_1	1	1
s_2	1	0
s_3	0	1
$\mathbf{F}_{1.3\vee4.1(nl)}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{1.3\vee4.1(nl)} : C \rightarrow S_{1.3\vee4.1(nl)}$		

Source: own compilation

3. Regarding the existence of both NAQ tables 4.3 and 4.4 until NL or NAQ table 4.8 until NL ($T^{nl}_{x.y} \in T_{2008}$), i.e. the possibility to substitute the missing of both NAQ tables 4.3 and 4.4, by the existence of NAQ table 4.8 to fulfill the 2008 MRDS criteria, we define the superset $S_{4.3+4.4\vee4.8}$ and the set of relevant state vectors $RS_{4.3+4.4\vee4.8}$:

$$S_{4.3+4.4\vee4.8} := \{s = (t_{4.3}, t_{4.4}, t_{4.8}) \mid t_{x.y} \in \{0,1\}\} = \{0,1\}^3. \qquad \text{(II.50)}$$

$$RS_{4.3+4.4\vee4.8} := \{s \in S_{4.3+4.4\vee4.8} \mid t_{4.3} * t_{4.4} = 1 \vee t_{4.8} = 1\}. \qquad \text{(II.51)}$$

$RS_{4.3+4.4\vee4.8} = \{s_1 , \ldots , s_5\}$ with s_i (i = 1, …, 5) is fulfilled in case of the NAQ table availability as given by the row vectors of the availability matrix in Table 8.

We define the vectorial mapping $\mathbf{F}_{4.3+4.4\vee4.8} : C \to S_{4.3+4.4\vee4.8}$ with $S_{4.3+4.4\vee4.8} = \{0,1\}^3$,

$$c \in C \mapsto \mathbf{F}_{4.3+4.4\vee4.8}(c) = (t_{4.3}, t_{4.4}, t_{4.8}). \qquad \text{(II.52)}$$

$\mathbf{F}_{4.3+4.4\vee4.8}(c) = s$ is reflecting the NAQ table availability (production) for $c \in C$ indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{4.3+4.4\vee4.8}(c)$ for $T^{nl}_{x.y}$.

Table 8 Availability Matrix of $T_{x.y}$ that Fulfill the Criteria of the Substitutional NAQ Tables Set 4.4+4.6 $\vee$ 4.8

$t_{x.y}$ / $\mathbf{F}_{4.3+4.4\vee4.8}(c)$	$t_{4.3}$	$t_{4.4}$	$t_{4.8}$
s_1	1	1	1
s_2	1	1	0
s_3	1	0	1
s_4	0	1	1
s_5	0	0	1
$\mathbf{F}_{4.3+4.4\vee4.8}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{4.3+4.4\vee4.8} : C \to S_{4.3+4.4\vee4.8}$			

Source: own compilation

4. Similar to 3., regarding the existence of both NAQ tables 4.6 and 4.7 until NL or NAQ table 4.9 until NL ($T^{nl}_{x.y} \in T_{2008}$), we define the superset $S_{4.6+4.7\vee4.9}$, and $RS_{4.6+4.7\vee4.9}$ as the set of relevant state vectors:

$$S_{4.6+4.7\vee4.9} := \{s = (t_{4.6}, t_{4.7}, t_{4.9}) \mid t_{x.y} \in \{0,1\}\} = \{0,1\}^3. \qquad \text{(II.53)}$$

$$RS_{4.6+4.7\vee4.9} := \{s \in S_{4.6+4.7\vee4.9} \mid t_{4.6} * t_{4.7} = 1 \vee t_{4.9} = 1\}. \qquad \text{(II.54)}$$

$RS_{4.6+4.7\vee 4.9}$ = {s_1 , … , s_5} with s_i (i = 1, …, 5) is fulfilled in case of the NAQ table availability as given by the row vectors of the availability matrix in Table 9.

We define the vectorial mapping $\mathbf{F}_{4.6+4.7\vee 4.9} : C \rightarrow S_{4.6+4.7\vee 4.9}$ with $S_{4.6+4.7\vee 4.9} = \{0,1\}^3$,

$$c \in C \mapsto \mathbf{F}_{4.6+4.7\vee 4.9}(c) = (t_{4.6,}\, t_{4.7,}\, t_{4.9}). \qquad \text{(II.55)}$$

$\mathbf{F}_{4.6+4.7\vee 4.9}(c) = s$ is reflecting the NAQ table availability (production) for $c \in C$ indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{4.6+4.7\vee 4.9}(c)$ for $T^{nl}_{x.y}$.

Table 9 Availability Matrix of $T_{x.y}$ that Fulfill the Criteria of the Substitutional NAQ Table Set 4.6+4.7$_\vee$4.9

$t_{x.y}$ / $\mathbf{F}_{4.6+4.7\vee 4.9}(c)$	$t_{4.6}$	$t_{4.7}$	$t_{4.9}$
s_1	1	1	1
s_2	1	1	0
s_3	1	0	1
s_4	0	1	1
s_5	0	0	1
$\mathbf{F}_{4.6+4.7\vee 4.9}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{4.6+4.7\vee 4.9} : C \rightarrow S_{4.6+4.7\vee 4.9}$			

Source: own compilation

The availability of the needed NAQ table combinations that fulfill the requirements of these sets of substitutional NAQ tables is assessed in the availability study presented in part III chapter 2.3, p. 212 ff. The assessment regarding the countries that fulfill these criteria is analogous to that of the MRDS measures; this applies also to the assessments of the members of the preimage sets of the substitutional NAQ table sets, and to the corresponding indicator functions. An exemplary assessment for one substitutional NAQ table set is provided in the following.

Referring to 1. – 4., where we defined the sets of substitutional NAQ tables, we now define exemplary the preimage sets of countries that fulfill these sets, the indicator functions indicating that a country $c \in C$ is member of that preimage set, and further the calculation of the cardinality of that preimage set.

The availability of the required NAQ tables, i.e. the fulfillment of the substitutional NAQ table set, is assessed analogous to those of $MRDS_{1993}$ and $MRDS_{2008}$. The assessment is done *exemplarily* for the last (4.) of the above defined sets concerning the existence of both NAQ tables 4.6 and 4.7 or the substitutional NAQ table 4.9:

The map $\mathbf{F}_{4.6+4.7\vee 4.9} : C \rightarrow S_{4.6+4.7\vee 4.9}$ with $S_{4.6+4.7\vee 4.9} = \{0,1\}^3$,

$$c \in C \mapsto \mathbf{F}_{4.6+4.7\vee 4.9}(c) = (t_{4.6}, t_{4.7}, t_{4.9}). \tag{II.55}$$

and

$$t_{x.y} = F_{x.y}(c) = \begin{cases} 1 \text{ if NAQ table } T^{nl}_{x.y} \text{ is available}^{34} \text{ for } c \in C \text{ and} \\ 0 \text{ else,} \end{cases}$$

indicates which of the tables $T^{nl}_{4.6}$, $T^{nl}_{4.7}$, $T^{nl}_{4.9}$ of T_{2008} are available in $c \in C$, (that is produced by c) and which are not.

Then the countries producing both of the NAQ tables 4.6 and 4.7 or at least table 4.9 are represented as the members of the preimage set

$$C_{4.6+4.7\vee 4.9} := \mathbf{F}_{4.6+4.7\vee 4.9}{}^{-1}(RS_{4.6+4.7\vee 4.9}) = \{c \in C \mid \mathbf{F}_{4.6+4.7\vee 4.9}(c) \in RS_{4.6+4.7\vee 4.9}\}. \tag{II.56}$$

Further we define the indicator function:

$$I_{4.6+4.7\vee 4.9} : C \rightarrow \{0,1\} \quad \text{with } I_{4.6+4.7\vee 4.9}(c) = \begin{cases} 1 \text{ if } c \in C_{4.6+4.7\vee 4.9} \\ 0 \text{ else.} \end{cases} \tag{II.57}$$

The indicator function indicates for each country $c \in C$ whether or not it is element of the preimage set.

Thus we have

$$C_{4.6+4.7\vee 4.9} = I_{4.6+4.7\vee 4.9}{}^{-1}(1). \tag{II.58}$$

As the result of interest for the assessment by regions and country groups, we calculate the number of countries within a region or country group for which the assessed NAQ table combination is available. Thus we are interested in the cardinality of $C_{4.6+4.7\vee 4.9}$, that is the sum of $I_{4.6+4.7\vee 4.9}(c)$ over all $c \in C$:

$$\left|C_{4.6+4.7\vee 4.9}\right| = \sum_{c\in C} \mathrm{I}_{4.6+4.7\vee 4.9}(c). \tag{II.59}$$

Corresponding to the cardinality of $C_{4.6+4.7\vee4.9}$, the share $h_{4.6+4.7\vee4.9}$ of countries that have the NAQ table combination available relative to the total number of countries C:

$$h_{4.6+4.7\vee4.9} := \frac{\left|C_{4.6+4.7\vee4.9}\right|}{\left|C\right|}, \qquad \text{(with } 0 \le h_{4.6+4.7\vee4.9} \le 1\text{).} \tag{II.60}$$

The availabilty of the other sets of substitutional NAQ tables, given by Table 6 to Table 8 are assessed for country c $\in C$ by the vectorial mappings $\mathbf{F}_{1.3\vee4.1}(c)$, $\mathbf{F}_{1.3\vee4.1(nl)}(c)$, and $\mathbf{F}_{4.3+4.4\vee4.8}(c)$, analogous as done above with $\mathbf{F}_{4.6+4.7\vee4.9}$.

Sample Application 7: Alternative NAQ Table Combinations

Example 1: Cameroon

We build upon the data availability of Cameroon and Thailand as given in Sample Application 6. Now, the availability of sets of substitutional NAQ tables is assessed, that does mean the existence of one NAQ table compensates the missing of another.

We consider the availability of the substitutional NAQ tables 1.3 or 4.1

1. given the existence of any NA indicator within these NAQ tables, and
2. given the existence of the indicator NL (specifying the completeness),

further we consider the availability of

3. NAQ table 4.8 or of both the NAQ tables 4.3 and 4.4 (all until NL), and of
4. NAQ table 4.9, or of both 4.6 and 4.7 (all until NL).

Thus, Cameroon is included in the preimage sets of countries that have the sets of substitutional NAQ table 1.3 or 4.1 available (until NL) i.e. referring to 1. and 2., further of 3., since both NAQ tables 4.3 and 4.4 are available, and also of 4., since both NAQ tables 4.6 and 4.7 are available (all until NL).

Referring to the formulas, the sets of relevant state vectors for the sets of substitutional NAQ tables describes above (1. to 4.) are defined by II.45, II.48, II.51, and II.54. Further, the availability of the NAQ tables equates to the following mappings (corresponding to Formulas II.46, II.49, II.52, II.55):

$\mathbf{F}_{1.3\vee4.1}(\text{CMR}) = (1,1)$,
$\mathbf{F}_{1.3\vee4.1(\text{nl})}(\text{CMR}) = (1,1)$),
$\mathbf{F}_{4.3+4.4\vee4.8}(\text{CMR}) = (1,1,0)$, and
$\mathbf{F}_{4.6+4.7\vee4.9}(\text{CMR}) = (1,1,0)$,

Availability of Substitutional NAQ Tables by Cameroon

$\boldsymbol{RS}$	$\mathbf{F}_{\boldsymbol{RS}} \in \boldsymbol{RS}$
$RS_{1.3\vee4.1}$	✓
$RS_{1.3\vee4.1(\text{nl})}$	✓
$RS_{4.3+4.4\vee4.8}$	✓
$RS_{4.6+4.7\vee4.9}$	✓

(continued on next page)

where the relevant set of state vectors fulfilling the criteria of the substitutional NAQ tables sets is defined by $RS_{1.3\vee 4.1}$, $RS_{1.3\vee 4.1(nl)}$, $RS_{4.3+4.4\vee 4.8}$, and $RS_{4.6+4.7\vee 4.9}$.

Consequently Cameroon is member of the sets $C_{1.3\vee 4.1}$, $C_{1.3\vee 4.1(nl)}$, $C_{4.3+4.4\vee 4.8}$, and $C_{4.6+4.7\vee 4.9}$. Further, the indicator functions are $\mathrm{I}_{1.3\vee 4.1}(\mathrm{CMR}) = 1$, $\mathrm{I}_{1.3\vee 4.1(nl)}(\mathrm{CMR}) = 1$, $\mathrm{I}_{4.3+4.4\vee 4.8}(\mathrm{CMR}) = 1$, and $\mathrm{I}_{4.6+4.7\vee 4.9}(\mathrm{CMR}) = 1$.

Example 2: Thailand

Since NAQ table 1.3 is available until NL, this compensates for the missing of NAQ table 4.1, therefore Thailand is included in the preimage sets of countries that have the set of substitutional NAQ tables 1.3 or 4.1 available given cases 1. and 2 above.

Thailand is not included in the preimage sets of countries for cases 3. and 4.

Referring to the formulas (see references in Example 1), the availability of the NAQ tables equates to the mappings

$\mathbf{F}_{1.3\vee 4.1}(\mathrm{THA}) = (1,1)$,
$\mathbf{F}_{1.3\vee 4.1(nl)}(\mathrm{THA}) = (1,0)$,
$\mathbf{F}_{4.3+4.4\vee 4.8}(\mathrm{THA}) = (0,0,0)$, and
$\mathbf{F}_{4.6+4.7\vee 4.9}(\mathrm{THA}) = (0,0,0)$.

Availability of Substitutional NAQ Tables by Thailand

RS	$\mathbf{F}_{RS} \in RS$
$RS_{1.3\vee 4.1}$	✓
$RS_{1.3\vee 4.1(nl)}$	✓
$RS_{4.3+4.4\vee 4.8}$	×
$RS_{4.6+4.7\vee 4.9}$	×

Consequently Thailand is member of the sets $C_{1.3\vee 4.1}$ and $C_{1.3\vee 4.1(nl)}$, but not of $C_{4.3+4.4\vee 4.8}$ and $C_{4.6+4.7\vee 4.9}$. The indicator functions are $\mathrm{I}_{1.3\vee 4.1}(\mathrm{THA}) = 1$, $\mathrm{I}_{1.3\vee 4.1(nl)}(\mathrm{THA}) = 1$, $\mathrm{I}_{4.3+4.4\vee 4.8}(\mathrm{THA}) = 0$, and $\mathrm{I}_{4.6+4.7\vee 4.9}(\mathrm{THA}) = 0$.

3.2.3 **Availability of Individual NAQ Tables**

The availability of individual NAQ tables in a country $c \in C$ is assessed for the NAQ tables of the 2008 and 1993 MRDS. These assessments are similar to those of the sets $MRDS_{2008} \subset S_{2008}$ and $MRDS_{1993} \subset S_{1993}$. This section formally delineates the measures for the assessment regarding the availability of individual NAQ tables.

Assessing the availability regarding the individual NAQ tables needed to fulfill the MRDS measures (similarly applicable to the availability of substitutional NAQ tables) allows us to identify which of the required tables and corresponding NA concepts are well available and which ones are less frequently available, i.e. marking where the bottle-necks are. Thus capacity building and technical assistance programs by IOs can focus on these weak points.

The assessment results regarding the availability of the individual NAQ tables are presented together with the results tables of the 1993 and 2008 MRDS assessments in part III, chapter 2 in Table 34 to Table 39 (pp. 214 - 221) of the 1993 MRDS assessment and in Table 40 to Table 45 (pp. 223 - 230) of the 2008 MRDS assessment.

The individual NAQ tables of the 1993 MRDS that are assessed for their availability are given by the set

$$T^*_{1993} = \{T_{1.1}, T_{1.2}, T_{2.1}, T_{2.2}, T_{2.3}\} = T_{1993} \setminus . \{T_{1.3}, T_{4.1}\},$$

and in case of the 2008 MRDS by the set

$$T^*_{2008} = \{T^{nl}_{4.1}, T^{nl}_{4.2}, T^{f}_{4.2}, T^{nl}_{4.3}, \ldots, T^{nl}_{4.8}, T^{nl}_{4.9}\} = (T_{2008} \setminus T^*_{1993}) \cup \{T^{nl}_{4.8}, T^{f}_{4.2}\}.$$

Note: The NAQ table $T^{nl}_{4.8}$ provides the unconsolidated aggregated data of NAQ tables $T^{nl}_{4.3}$ and $T^{nl}_{4.4}$. Further, $T^{f}_{4.2}$ refers to the *financial account*, symbolized by "f", of NAQ table 4.2. The availability of the *financial account* is evaluated based on the SNA indicator "*financial derivatives*". Identical to the 1993 and 2008 MRDS the existence of any SNA indicator within the NAQ table $T_{x.y}$ is sufficient if not otherwise indicated. Further, the requirement until the *net lending* indicator of the capital account (NL) is symbolized by $T^{nl}_{x.y}$ for the NAQ table. The latter is evaluated based on the existence of the NL indicator. (Regarding the condition to determine the existence of a required

NAQ table [by defined detail], and its sufficiency, refer to part III, chapter 2.2.2 Availability Conditions of NAQ Tables, p. 203.)

Now, let $C^{nl}_{x.y}$ be the set of countries that produce the NAQ table $T^{nl}_{x.y} \in T^*_{2008}$, respectively $C^{f}_{x.y}$ for $T^{f}_{x.y} \in T^*_{2008}$ and $C_{x.y}$ for $T_{x.y} \in T^*_{1993}$.

The existence and non-existence of each NAQ table $T^{nl}_{x.y}$ respectively $T^{f}_{x.y}$ and $T_{x.y}$ is indicated by the value of the indicator variable $t^{nl}_{x.y}$ respectively $t^{f}_{x.y}$ and that is

$$t^{nl}_{x.y} = \begin{cases} 1 \text{ if NAQ table } T^{nl}_{x.y} \text{ exists} \\ 0 \text{ else.} \end{cases}$$

For $t^{f}_{x.y}$ and $T^{f}_{x.y}$ respectively $t_{x.y}$ and $T_{x.y}$, analogous.

Than we define $F^{nl}_{x.y} : C \rightarrow \{0,1\}$ for every $T^{nl}_{x.y} \in T^*_{2008}$ (i.e. each NAQ table individually) by

$$c \in C \mapsto t^{nl}_{x.y} = F^{nl}_{x.y}(c) = \begin{cases} 1, \text{ if NAQ table } T^{nl}_{x.y} \text{ is available for } c \in C, \\ 0 \text{ else.} \end{cases} \tag{II.61}$$

Analogous we define $F^{f}_{x.y} : C \rightarrow \{0,1\}$ for every $T^{f}_{x.y} \in T^*_{2008}$ by

$$c \in C \mapsto t^{f}_{x.y} = F^{f}_{x.y}(c) = \begin{cases} 1, \text{ if NAQ table } T^{f}_{x.y} \text{ is available for } c \in C, \\ 0 \text{ else,} \end{cases} \tag{II.62}$$

and $F_{x.y} : C \rightarrow \{0,1\}$ for every $T_{x.y} \in T^*_{1993}$ by

$$c \in C \mapsto t_{x.y} = F_{x.y}(c) = \begin{cases} 1, \text{ if NAQ table } T_{x.y} \text{ is available for } c \in C, \\ 0 \text{ else.} \end{cases} \tag{II.63}$$

With other words, $F^{nl}_{x.y}(c)$ indicates for an individual NAQ table $T^{nl}_{x.y}$, respectively $F^{f}_{x.y}$ for $T^{f}_{x.y}$ and $F_{x.y}$ for $T_{x.y}$, whether or not it is available in c (that is produced by c).

Then the countries for which $T^{nl}_{x.y}$ respectively $T^{f}_{x.y}$ or $T_{x.y}$ is available are represented as the members of the preimage set

$C^{nl}_{x.y} := F^{nl\ -1}_{x.y}(1) = \{c \in C \mid F^{nl}_{x.y}(c) = 1\}$, for all $T^{nl}_{x.y} \in T^*_{2008}$. (II.64)

For $C^{f}_{x.y} := F^{f\ -1}_{x.y}(1)$ and

for $C_{x.y} := F_{x.y}^{-1}(1)$, analogous.

Further we define the indicator function:

$I^{nl}_{x.y} : C \rightarrow \{0,1\}$, for all $T^{nl}_{x.y} \in T^*_{2008}$, with (II.65)

$$I^{nl}_{x.y}(c) = \begin{cases} 1 \text{ if } c \in C^{nl}_{x.y}, \\ 0 \text{ else.} \end{cases}$$

For $I^{f}_{x.y} : C \rightarrow \{0,1\}$, for all $T^{f}_{x.y} \in T^*_{2008}$ and

for $I_{x.y} : C \rightarrow \{0,1\}$, for all $T_{x.y} \in T^*_{1993}$, analogous.

Thus, we have

$C^{nl}_{x.y} = I^{nl\ -1}_{x.y}(1)$, for all $T^{nl}_{x.y} \in T^*_{2008}$. (II.66)

Analogously, $C^{f}_{x.y} = I^{f\ -1}_{x.y}(1)$, for all $T^{f}_{x.y} \in T^*_{2008}$ and

for $C_{x.y} = I_{x.y}^{-1}(1)$, for all $T_{x.y} \in T^*_{1993}$.

Thus the sum of $I^{nl}_{x.y}(c)$ over all $c \in C$ is equal to the cardinality of $C^{nl}_{x.y}$, that is the number of countries that have the NAQ table $T^{nl}_{x.y} \in T_{2008}$ available. The total number of countries within a region or country group is of interest for indentifying the country clusters with good or particularly poor capacity to produce the corresponding NA data. Note that for the availability of individual NAQ tables, the result of the indicator function, $I^{nl}_{x.y}$, is that of $F^{nl}_{x.y}$, while for the other availability measures we use solely the indicator function to calculate the number of countries within the corresponding preimage set.

As formula this is:

$$\left|C^{nl}_{x.y}\right| = \left|I^{nl\ -1}_{x.y}(1)\right| = \left|F^{nl\ -1}_{x.y}(1)\right| = \sum_{c \in C} F^{nl}_{x.y}(c) = \sum_{c \in C} I^{nl}_{x.y}(c). \quad \text{(II.67)}$$

For $I^{f}_{x.y}(c)$ and $T^{f}_{x.y} \in \mathcal{T}_{2008}$ and for $I_{x.y}(c)$ and $T_{x.y}$, we get analogous results.

Corresponding to the cardinality of $C^{nl}_{x.y}$, the share $h^{nl}_{x.y}$ of countries that have the NAQ table available, in relation to the total number of countries C:

$$h^{nl}_{x.y} := \frac{\left|C^{nl}_{x.y}\right|}{C}, \qquad \text{(with } 0 \le h^{nl}_{x.y} \le 1\text{).} \qquad \text{(II.68)}$$

For the shares $h^{f}_{x.y}$ and $h_{x.y}$, corresponding to the cardinality of $C^{f}_{x.y}$ respectively $C_{x.y}$, analogously.

Sample Application 8: Individual NAQ Tables

Example 1: Cameroon

Building upon the preceding Sample Application 6 the availability of individual NAQ tables is assessed, namely of tables 1.1, 1.2, 2.1, 2.2, 2.3, further of tables 4.1 to 4.9 until NL, and of table 4.2 regarding the financial account. All but tables 4.8 and 4.9 are available, i.e. until NL where specified and also the financial account of table 4.2.

Thus Cameroon is included in the set of countries that has, for example the individual NAQ table 1.1 available, ditto for 1.2 - 4.7, and it is not included in the preimage sets regarding the NAQ tables 4.8 - 4.9 and for the financial account of 4.2.

Referring to the formulas (II.61 ff.), the availability of the NAQ tables is given by $F_{1.y}(CMR) = 1$ for y = 1, 2, $F_{2.y}(CMR) = 1$ for y = 1, 2, 3, $F^{nl}_{4.y}(CMR) = 1$ for y = 1, …, 7, $F^{nl}_{4.y}(CMR) = 0$ for y = 8, 9, and $F^{f}_{4.2}(CMR) = 0$.

Thus, Cameroon is member of the sets $C_{1.1}$, $C_{1.2}$, $C_{2.1}$, $C_{2.2}$, $C_{2.3}$, and $C^{nl}_{4.1}$ to $C^{nl}_{4.7}$ and not of $C^{nl}_{4.8}$ to $C^{nl}_{4.9}$ and $C^{f}_{4.2}$. The indicator functions are $I_{1.y}(CMR) = 1$ for y = 1, 2, $I_{2.y}(CMR) = 1$ for y = 1, 2, 3, $I^{nl}_{4.y}(CMR) = 1$ for y = 1, …, 7, 9, and $I^{f}_{4.y}(CMR) = 0$.

Example 2: Thailand

Thailand is included in the sets of countries that have the individual NAQ tables 1.1 to 2.3, and is not included in the preimage sets of countries for the individual NAQ tables 4.1 to 4.9 and for the financial account of table 4.2.

Referring to the formulas (see references in Example 1), the availability of the NAQ tables is given by $F_{1.y}(THA) = 1$ for y = 1, 2, $F_{2.y}(THA) = 1$ for y = 1, 2, 3, and by $F^{nl}_{4.y}(THA) = 0$ for y = 1, …, 9 and $F^{f}_{4.2}(THA) = 0$.

Consequently Thailand is member of the preimage sets $C_{1.1}$, $C_{1.2}$, $C_{2.1}$, $C_{2.2}$, and $C_{2.3}$, and not of $C^{nl}_{4.1}$ to $C^{nl}_{4.9}$ and $C^{f}_{4.2}$. Further, the indicator functions $I_{1.y}(THA) = 1$ for y = 1, 2, $I_{2.y}(THA) = 1$ for y = 1, 2, 3, and $I^{nl}_{4.y}(THA) = 0$ for y = 1, …, 9 and $I^{f}_{4.y}(THA) = 0$.

3.2.4 **Institutional Sectors Accounts of the 2008 MRDS**

The 2008 MRDS chiefly extends the 1993 MRDS requirements by the institutional sectors (i.e. NAQ tables 4.1 - 4.5, further 4.6 and 4.7 or 4.9). Therefore, we separately assess the set of countries that produce all and at least all but one of the required institutional sectors until NL (represented by the corresponding NAQ tables until NL).

Required Institutional Sectors of the 2008 MRDS

The considered set of NAQ tables is a subset of the 2008 MRDS tables. This assessment informs us whether the non-fulfillment of the 2008 MRDS criteria is due to the missing of any one or more of the institutional sectors until NL, or due to the missing of any one or more of the NAQ tables already required by the 1993 MRDS. The latter would be the case, if the number of countries fulfilling the 2008 MRDS is smaller than the number of countries with availability of the required set of institutional sectors until NL.

The availability (existence) of the set of institutional sectors until NL that are required to fulfill the 2008 MRDS criteria (cf. DEF. 26, p. 64) is formally delineated as follows:

Let the set of NAQ tables referencing institutional sector tables be denominated by

$$T_{\mathrm{IS}} := \{ \mathrm{T}^{\mathrm{nl}}_{4.1}, \mathrm{T}^{\mathrm{nl}}_{4.2}, \mathrm{T}^{\mathrm{nl}}_{4.3}, \mathrm{T}^{\mathrm{nl}}_{4.4}, \mathrm{T}^{\mathrm{nl}}_{4.5}, \mathrm{T}^{\mathrm{nl}}_{4.6}, \mathrm{T}^{\mathrm{nl}}_{4.7}, \mathrm{T}^{\mathrm{nl}}_{4.9} \}. \qquad \text{(II.69)}$$

Thus, $T_{\mathrm{IS}} \subset T_{2008.}$

We define the superset S_{IS} of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $\mathrm{T}^{\mathrm{nl}}_{\mathrm{x.y}} \in T_{\mathrm{IS}}$:

$$S_{\mathrm{IS}} := \{ \mathrm{s} = (\mathrm{t}_{4.1}, \mathrm{t}_{4.2}, \mathrm{t}_{4.3}, \mathrm{t}_{4.4}, \mathrm{t}_{4.5}, \mathrm{t}_{4.6}, \mathrm{t}_{4.7}, \mathrm{t}_{4.9}) \mid \mathrm{t}_{\mathrm{x.y}} \in \{0,1\} \} = \{0,1\}^8. \qquad \text{(II.70)}$$

As relevant subset of state vectors we define the set $IS \subset S_{\mathrm{IS}}$:

$$IS := \{ \mathrm{s} \in S_{\mathrm{IS}} \mid \mathrm{t}_{4.1} \cdot \mathrm{t}_{4.2} \cdot \mathrm{t}_{4.3} \cdot \mathrm{t}_{4.4} \cdot \mathrm{t}_{4.5} \cdot \mathrm{t}_{4.6} \cdot \mathrm{t}_{4.7} = 1$$
$$\vee\ \mathrm{t}_{4.1} \cdot \mathrm{t}_{4.2} \cdot \mathrm{t}_{4.3} \cdot \mathrm{t}_{4.4} \cdot \mathrm{t}_{4.5} \cdot \mathrm{t}_{4.9} = 1 \}. \qquad \text{(II.71)}$$

To describe whether or not a country $c \in C$ produces the required set of institutional sectors until NL we define the mapping $\mathbf{F}_{IS} : C \to S_{IS}$, with $S_{IS} = \{0,1\}^8$

$$c \in C \mapsto \mathbf{F}_{IS}(c) = \left(t_{4.1}, t_{4.2}, t_{4.3}, t_{4.4}, t_{4.5}, t_{4.6}, t_{4.7}, t_{4.9}\right) \quad \text{(II.72)}$$

and

$$t_{x.y} = F_{x.y}(c) = \begin{cases} 1 \text{ if NAQ table } T^{nl}_{x.y} \text{ is available}^{46} \text{ for } c \in C\text{, and} \\ 0 \text{ else.} \end{cases}$$

With other words, $\mathbf{F}_{IS}$ indicates which of the tables of T_{IS} are available in c, (that is produced by c) and which are not.

Then the countries that produce all required institutional sectors until NL are represented as the members of the preimage set of $IS \subset S_{IS}$, i.e. by

$$C_{IS} := \mathbf{F}_{IS}^{-1}(IS) = \{c \in C \mid \mathbf{F}_{IS}(c) \in IS\}. \quad \text{(II.73)}$$

Further we define the indicator function

$$I_{IS} : C \to \{0,1\} \text{ with } \quad I_{IS}(c) = \begin{cases} 1 \text{ if } c \in C_{IS} \\ 0 \text{ else.} \end{cases} \quad \text{(II.74)}$$

Thus we have

$$C_{IS} = I_{IS}^{-1}(1). \quad \text{(II.75)}$$

As the result of interest for the assessment by regions and country groups, we calculate the number of countries within a region or country group for which the required institutional sector accounts until NL are available. Thus we are interested in the cardinality of C_{IS}, that is the sum of $I_{IS}(c)$ over all $c \in C$:

$$|C_{IS}| = \sum_{c \in C} I_{IS}(c). \quad \text{(II.76)}$$

[46] Availability refers to the nationally produced data, measured by the data transmitted by a national agency to the UN and made available through publication.

Corresponding to the cardinality of C_{IS}, the share h_{IS} of countries that produce the set *IS*, i.e. all required institutional sectors until NL in relation to the total number of countries *C*:

$$h_{IS} := \frac{|C_{IS}|}{C}, \qquad \text{(with } 0 \leq h_{IS} \leq 1\text{).} \qquad \text{(II.77)}$$

$\mathbf{F}_{IS}(c) = s \in IS$ indicates for a specific country $c \in C$ that the set of required NAQ tables until NL is available.

The following table presents the availability matrix for each possible $\mathbf{F}_{IS}(c) \in IS$ and corresponding $T^{nl}_{x.y}$ indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{IS}(c)$ for $T^{nl}_{x.y}$. (For easier visibility of the differences between the row vectors, "0" are printed in bold font).

For s_i with i = 1, …, 5, the row vectors of the availability matrix of Table 10 are similar to the first five rows of Table 5, p. 67 for the fulfillment of the 2008 MRDS, yet without the "first six" columns of that table (referring to NAQ tables 1.1. to 2.3).

Table 10 Availability Matrix of $T_{x.y} \in T_{IS}$ for the Cases where $F_{IS}(c)$ Fulfills the Set *IS*

$t_{x.y}$ / $\mathbf{F}_{IS}(c)$	$t_{4.1}$	$t_{4.2}$	$t_{4.3}$	$t_{4.4}$	$t_{4.5}$	$t_{4.6}$	$t_{4.7}$	$t_{4.9}$
s_1	1	1	1	1	1	1	1	1
s_2	1	1	1	1	1	1	1	**0**
s_3	1	1	1	1	1	1	**0**	1
s_4	1	1	1	1	1	**0**	1	1
s_5	1	1	1	1	1	**0**	**0**	1
$\mathbf{F}_{IS}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{IS} : C \rightarrow S_{IS}$								

Source: own compilation

Example: The fulfillment of the *IS* criteria is assessed analogous to the 2008 MRDS measure. For a comprehensive example refer to chapter 3.2.1.1, p. 59. Here, only one example for the fulfillment and the non-fulfillment of *IS* are provided. The scenarios show the results of the function $\mathbf{F}_{IS}(c)$ and $I_{IS}(c)$ for a specific country $c \in C$.

Scenario 1: Both NAQ tables 4.6 and 4.7 are missing, but 4.9 is available (corresponding to row vector s_5 in the above availability matrix).

$\mathbf{F}_{IS}(c) = (1,1,1,1,1,0,0,1) = s_5 \in IS$, that is $I_{IS}(c) = 1$.

Scenario 2: NAQ table 4.1 is missing.

$\mathbf{F}_{IS}(c) = (0,1,1,1,1,1,1,1) = s \notin IS$, that is $I_{IS}(c) = 0$.

Partial Availability of Institutional Sectors of the 2008 MRDS

Similar to the case of the 1993 and 2008 MRDS, the partial availability of the required set of institutional sectors until NL is assessed. This measure helps indentifying those countries missing one of the requirement institutional sectors, yet have very advanced NA production capacities. Here, we consider the set of countries that produce at least a substantial part of the required set of NAQ tables 4.1 to 4.5 along with either both of 4.6 and 4.7 or at least 4.9. This does mean that in case of a partial fulfillment, one of the NAQ tables 4.1 to 4.5, or all of 4.6, 4.7 and 4.9 are allowed to be missing. The latter case, i.e. the missing of three NAQ tables is the non-fulfillment of only one requirement criteria, since NAQ table 4.9 is the substitutional NAQ table for the combined availability of both tables 4.6 and 4.7.

Likewise to the partial fulfillment of the MRDS measures, a distance term is defined to measure the distance to the n-tuples given by the set *IS*. By this distance we get the set of n-tuples (indicating the NAQ table availability) differing exactly one criterion from the requirement criteria to fulfill the set *IS*.

Let the set of countries producing at least the substantial part of institutional sectors be denominated $C_{IS(P)}$.

$$\text{Thus } C_{IS} \subseteq C_{IS(P)}. \qquad \text{(II.78)}$$

Let $IS_{(P)}$ be the set of relevant n-tuples that differ by not more than one missing NAQ table $T_{x.y}$ indicated by $t_{x.y}$ from the set of n-tuples given by *IS*.

Let $X.Y_{IS}$ be the index set to the NAQ tables $T^{nl}_{x.y} \in T_{IS}$.

$$X.Y_{IS} = \{ \text{x.y} \mid T^{nl}_{x.y} \in T_{IS} \} \qquad \text{(II.79)}$$

$$= \{4.1\,, 4.2\,, 4.3\,, 4.4\,, 4.5\,, 46\,, 4.7\,, 4.9\}.$$

Than the distance is measured by d_{IS}:

$$d_{IS}(s,\tilde{s}) = \sum_{x.y \in X.Y_{IS}} \left| t_{x.y} - \tilde{t}_{x.y} \right|. \tag{II.80}$$

Consequently, if for a s $\notin$ *IS* there exists a $\tilde{s} \in$ *IS* with $d(s,\tilde{s}) = 1$, than this does mean that exactly one of the criteria is missing to fulfill the availability measure *IS*. With other words, if for any country c $\in$ *C* the available set of NAQ tables $T^{nl}_{x.y} \in T_{IS}$ produces the mapping $\mathbf{F}_{IS}(c) = s$ and there exists a state vector $\tilde{s} \in$ *IS* with $d(s,\tilde{s}) \leq 1$, than this country fulfills at least a substantial part of the required institutional sectors accounts. In that case at maximum one of the NAQ tables 4.1 to 4.5 is missing, or up to all of the NAQ tables 4.6, 4.7 and 4.9 (for all NAQ tables referring to the availability until NL).

Then the countries fulfilling the $IS_{(P)}$ criteria are represented as the members of the preimage set of $IS_{(P)} \subset S_{IS}$, i.e. by

$$C_{IS(P)} := \mathbf{F}_{IS(P)}^{-1}(IS) = \{ c \in C \mid \mathbf{F}_{IS}(c) \mid d(s,\tilde{s}) \leq 1 \in IS \}. \tag{II.81}$$

Thus, $\mathbf{F}_{IS(P)}(c) = s \in IS_{(P)}$ indicates for a specific country c $\in$ *C* that it produces at least all but one of the required institutional sectors until NL (with other words at least a substantial part of *IS*).

Further we define the indicator function

$$I_{IS(P)} : C \rightarrow \{0,1\} \text{ with } \quad I_{IS(P)}(c) = \begin{cases} 1 \text{ if } c \in C_{IS(P)} \\ 0 \text{ else.} \end{cases} \tag{II.82}$$

Thus we have

$$C_{IS(P)} = I_{IS(P)}^{-1}(1). \tag{II.83}$$

The result of interest, the cardinality of C_{IS}, that is the sum of $I_{IS(P)}(c)$ over all c $\in$ *C* is

$$\left| C_{IS(P)} \right| = \sum_{c \in C} I_{IS(P)}(c). \tag{II.84}$$

The share $h_{IS(P)}$ is calculated analougously to Formula II.25, by the cardinality of $C_{IS(P)}$ in relation to the total number of countries *C*.

Sample Application 9: Institutional Sector Accounts

Example 1: Cameroon

Building upon the data availability in Sample Application 6, the availability of the required institutional sector accounts until NL (required to fulfill the 2008 MRDS) is assessed. The required NAQ tables are 4.1 to 4.5, and both 4.6 and 4.7 or at least 4.9 (as defined in II.71). Also assessed is the availability of the substantial part of the required set, i.e. the subset where exactly one of the required criteria is allowed to be missing (the partial availability is defined via the distance in II.80).

For Cameroon all of the NAQ tables 4.1 to 4.7 are available until NL and 4.9 is missing. As NAQ table 4.9 is a substitute for the existence of both NAQ tables 4.7 and 4.7, the missing of NAQ table 4.9 is compensated for by the combined availability of the other two. Thus Cameroon is included in the sets of countries that have the required set of institutional sector accounts available, and necessarily also at least the substantial part of them.

Referring to the formulas, for the required institutional sector accounts this equates to the image vector (II.72) $\mathbf{F}_{IS}(CMR) = (1,1,1,1,1,1,1,0) \in IS$. Consequently Cameroon is in the preimage sets C_{IS}, and the indicator function $I_{IS}(CMR) = 1$.

Regarding the fulfillment of a substantial part of *IS*, we get as distance between the n-tuple $s = \mathbf{F}_{IS}(CMR)$ and state vector $s \in IS$ the formula $d_{IS}(s,s) = 0 \leq 1$, thus Cameroon is also member of the preimage set $C_{IS(p)}$, further $I_{IS(p)} = 1$.

Example 2: Thailand

Since none of the NAQ tables 4.1 to 4.9 are available for Thailand, it is not included in the sets of countries that have the required set, nor a substantial part of it available.

Referring to the formulas, for the required institutional sector accounts this equates to the image vector $\mathbf{F}_{IS}(THA) = (0,0,0,0,0,0,0,0) \notin IS$. Consequently Thailand is not member of the preimage sets C_{IS} and, the indicator functions $I_{IS}(THA) = 0$.

Regarding the fulfillment of a substantial part of the set *IS*, we get for $s = \mathbf{F}_{IS}(THA)$ and all $\tilde{s} \in IS$ the distance $d_{IS}(s,\tilde{s}) \geq 6 > 1$, thus Thailand is not in the preimage set $C_{IS(p)}$, further $I_{IS(p)} = 0$.

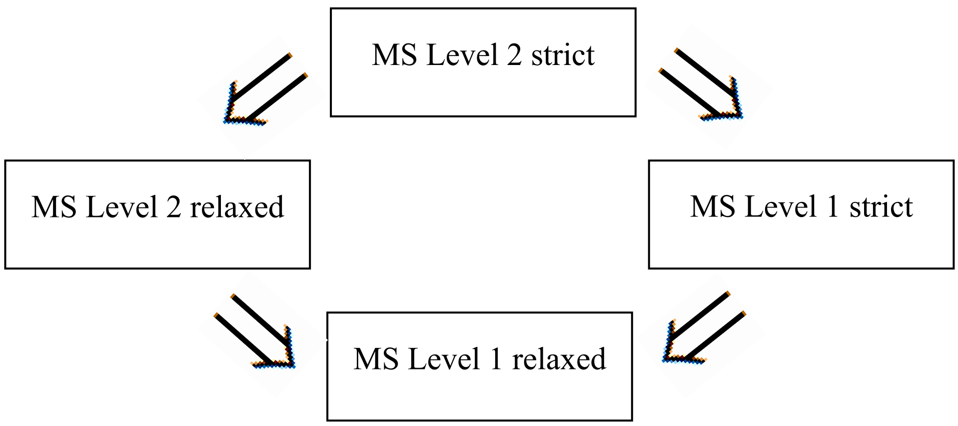
3.2.5 Milestone Levels 1 and 2 of National Accounts Production

The milestone levels 1 and 2 also assess the NA data availability in countries. The NAQ tables needed to fulfill these measures are arranged in a different way than the MRDS measures. They measure the production of SNA concepts at different levels. This subchapter defines and compares the milestone levels 1 and 2. We apply these measures in our availability assessment in part III, chapter 2 (p. 197 ff.).

DEF. 27: Milestones
"[A] set of recommendations for the phased implementation of the 1993 SNA. The so called '6 Milestones' […] to full SNA implementation" (ISWGNA 1996, 1). They are defined by: "Criteria against which progress regarding the country implementation of the System of National Accounts, 1993 (1993 SNA) could be judged. […] Higher milestone levels correspond to increasingly complex national accounting systems. Countries at level 1 are characterized by compiling gross domestic product at current and constant prices, broken down by kind of economic activity and final expenditure. Level-2 group countries add to the gross domestic product (GDP) information the main macroeconomic aggregates for the national economy and complete accounts for the rest of the world. Levels 3 to 6 then introduce institutional sector accounts, ranging from production accounts (level 3), income and capital accounts (level 4) and financial accounts (level 5) to balance sheet information (level 6)." United Nations (1999, 3). See also: DEF. 28 for milestone level 1 (strict), DEF. 29 for milestone level 2 (strict), DEF. 30 for milestone level 1 (relaxed), DEF. 31 for milestone level 2 (relaxed).

Source: United Nations (1999, 3)

There are two interpretations regarding the combination of NAQ tables that make up the milestone levels 1 and 2. The first being the original (strict) and the second the later (relaxed) interpretation of the NA data needed to reach these milestone levels. In the following the terms "strict" and "relaxed" will be used to refer to these different interpretations. Note that the "relaxed" milestone levels are only a part of the "strict" milestone levels, i.e. the “strict” interpretation requires more detailed NA data. Further, milestone level 1 is part of milestone level 2. Yet, since the “relaxed” milestone level 2 builds upon the “relaxed” milestone level 1, the “strict” milestone level 1 already is not part of the “relaxed” milestone level 2, because the “strict” interpretation already requires a combination of NAQ tables that is not needed for the “relaxed” milestone levels. The following figure illustrates the relations between the Milestone (MS) levels as logic conditions.



$\Rightarrow$: symbolizes the logical implication regarding the milestone levels.

Figure 13 Relation between Milestone Levels 2 “Strict”, 2 “Relaxed”, 1 “Strict” and 1 “Relaxed”

Source: own illustration

Thus if we look at the MS Levels as logic conditions, as illustrated by Figure 13, we get the following logic causal structure where Level 2 “strict” implies the compliance (symbolized by “⇒”) with Level 2 “relaxed” and with Level 1 “strict”. Further, through Level 2 “relaxed” and Level 1 “strict”, the Level 2 “strict” also implies the compliance with Level 1 “relaxed”, since this is directly implied by the aforementioned Levels 2 “relaxed” and Level 1 “strict”. The formal delineation of the four milestone level measures is given in the upcoming sections.

We argue for the re-revision of the “relaxed” to the “strict” interpretation of the milestone levels based on the extended NA data required to fulfill the 2008 MRDS. (For further details and comparison of results of the first milestone assessment with the “strict” interpretation and the later assessment with the “relaxed” interpretation given the 1993 SNA implementation, see Part III, chapter 2 (p. 197 ff.).

The milestone levels can be used to better evaluate the progress countries made in the process of producing different of the SNA concepts. The advantage of the milestone levels is that they focus on specific SNA concepts. Fulfilling the first two milestone levels given the "relaxed" criteria is not as demanding as fulfilling the full 1993 MRDS. They can be fulfilled by a subset of the NAQ tables of the 1993 MRDS.

3.2.5.1 "Strict" Milestone Levels

In this subsection we provide the measures given the original "strict" criteria.

"Strict" Milestone Level 1

DEF. 28: Milestone Level 1 "Strict"

The combination of NAQ tables $T_{x.y}$ as objects, that make up the original "strict" milestone level 1 is given by the existence (requirement) of all of the NAQ tables 1.1, 1.2, 2.1, and 2.2 (cf. ISWGNA 1996).

Milestone level 1 "strict" is formally delineated as follows:

Let the set of potential NAQ tables be

$$T_{1s} := \{T_{1.1}, T_{1.2}, T_{2.1}, T_{2.2}\}.$$

Regarding the NAQ tables $T_{x.y}$ the existence and non-existence (i.e. requirement and non-requirement) is indicated by the indicator variable $t_{x.y}$.

Now we define the superset S_{1s} of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $T_{x.y} \in T_{1s}$ by

$$S_{1s} := \{s = (t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}) \mid t_{x.y} \in \{0,1\}\} = \{0,1\}^4.$$

As relevant subset of state vectors fulfilling the level 1 "strict" criteria we define the set

$$L_{1s} := \{ s \in S_{1s} \mid t_{1.1} * t_{1.2} * t_{2.1} * t_{2.2} = 1\}.$$ [47]

Source: own definition based on ISWGNA (1996)

[47] That is all factors $t_{x.y}$ in these products are =1. That does mean, $s \in \boldsymbol{S}_{1s}$ signals that all of the components $t_{1.1}, t_{1.2}, t_{2.1},$ and $t_{2.2}$ have to be 1.

For example: $s = (0,0,0,0) \in S_{1s}$ indicates the non-existence of any tables in T_{1s}, $s = (1,1,1,1) \in S_{1s}$ the existence of all tables in T_{1s}, and $s = (1,0,1,0) \in S_{1s}$ the existence of the tables $T_{1.1}$ and $T_{2.1}$ and the missing of $T_{1.2}$ and $T_{2.2}$.

The above definition of T_{1s}, S_{1s}, and L_{1s} for the "strict" milestone level 1 is delineated from the description of that milestone level as used since 1996: "The basic indicators of Gross Domestic Product (GDP). Final expenditures on GDP current and constant prices, GDP by industry at current and constant prices." (ISWGNA 1996, 2).

"Strict" Milestone Level 1 Fulfillment by the Available Country Data

To describe whether or not a country $c \in C$ fulfills the "strict" milestone level 1 criteria we define the mapping $\mathbf{F}_{1s} : C \rightarrow S_{1s}$ with $S_{1s} = \{0,1\}^4$,

$$c \in C \mapsto \mathbf{F}_{1s}(c) = \left(t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}\right) \qquad \text{(II.85)}$$

and

$$t_{x.y} = F_{x.y}(c) = \begin{cases} 1 \text{ if NAQ table } T_{x.y} \text{ is available}^{48} \text{ for } c \in C, \text{ and} \\ 0 \text{ else.} \end{cases}$$

The image vector $\mathbf{F}_{1s}(c)$ indicates which of the tables of T_{1s} are available in c and which are not. $\mathbf{F}_{1s}(c) = s \in L_{1s}$ indicates for a specific country $c \in C$ that the country fulfills the "strict" milestone level 1. Then the countries meeting the "strict" milestone level 1 criteria are represented as the members of the preimage set of $L_{1s} \subset S_{1s}$, i.e. by

$$C_{1s} := \mathbf{F}_{1s}^{-1}(L_{1s}) = \{ c \in C \mid \mathbf{F}_{1s}(c) \in L_{1s} \}. \qquad \text{(II.86)}$$

Further we define the indicator function:

$$I_{1s} : C \rightarrow \{0,1\} \text{ with } \qquad I_{1s}(c) = \begin{cases} 1 \text{ if } c \in C_{1s} \\ 0 \text{ else.} \end{cases} \qquad \text{(II.87)}$$

Thus we have

$$C_{1s} = I_{1s}^{-1}(1). \qquad \text{(II.88)}$$

[48] Availability refers to the nationally produced data, measured by the data transmitted by a national agency to the UN and made available through publication.

We calculate the number of countries within a region or country group which fulfill the "strict" milestone measure 1. Thus we are interested in the cardinality of C_{1s}, that is the sum of $I_{1s}(c)$ over all $c \in C$:

$$|C_{1s}| = \sum_{c \in C} I_{1s}(c)\,. \qquad \text{(II.89)}$$

The share h_{1s} is calculated analougously to Formula II.25, by the cardinality of C_{1s} in relation to the total number of countries C, i.e. by:

$$h_{1s} = \frac{|C_{1s}|}{|C|}\,, \qquad \text{(with } 0 \le h_{1s} \le 1) \qquad \text{(II.90)}$$

The following Table 11 reflects the single combination of produced NAQ tables $T_{x.y}$, for which the country $c \in C$ fulfills the "strict" milestone level 1 criteria. The availability (meaning production i.e. existence respectively non-existence) of each NAQ table is indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{1s}(c)$ for $T_{x.y.}$

Table 11 Availability Matrix of $T_{x.y} \in T_{1s}$ for the Cases where L_{1s} Fulfills Milestone Level 1 "Strict"

$t_{x.y}$ / $\mathbf{F}_{1s}(c)$	$t_{1.1}$	$t_{1.2}$	$t_{2.1}$	$t_{2.2}$
s_1	1	1	1	1
$\mathbf{F}_{1s}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{1s} : C \rightarrow S_{1s}$ indicating for s_1 which tables are available or not.				

Source: own compilation

For all possible cases, other than $\mathbf{F}_{1s}(c) = (1,1,1,1) = s_1 \in L_{1s}$, that is $I_{1s}(c) = 1$, the availability of the required set of NAQ tables for the „strict" milestone level 1 is not fulfilled by the country, i.e. for any other n-tuple $\mathbf{F}_{1s}(c) \in S_{1s}$ is $I_{1s}(c) = 0$. This is illustrated for $c \in C$ with:

$\mathbf{F}_{1s}(c) = (t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}) \notin L_{1s}$, because of:

$t_{x.y} = F_{x.y} = 0$ for at least one of the indices 1.1, 1.2, 2.1, 2.2, and therefore

$F_{1.1}(c) * F_{1.2}(c) * F_{2.1}(c) * F_{2.2}(c) = 0$. Therefore $I_{1s}(c) = 0$.

“Strict” Milestone Level 2

DEF. 29: Milestone Level 2 “Strict”
The combination of NAQ tables $T_{x.y}$ as objects, that make up the “strict” milestone level 2 is given by the existence (requirement) of NAQ tables 1.1, 1.2, 2.1, 2.2, 4.2 and at least one of the tables 1.3 and 4.1. The NAQ tables 1.3, 4.1, and 4.2 need to be available until the NA indicator “*net lending*” of the capital account, further for NAQ table 4.2 the “*financial account*” (f) needs to be available (cf. ISWGNA 1996). Milestone level 2 “strict” is formally delineated as follows: Let the set of potential NAQ tables be $T_{2s} := \{ T_{1.1}, T_{1.2}, T^{nl}_{1.3}, T_{2.1}, T_{2.2}, T^{nl}_{4.1}, T^{nl}_{4.2}, T^{f}_{4.2} \}.$ Regarding the NAQ tables $T_{x.y,}$ $T^{nl}_{x.y}$, and $T^{f}_{x.y}$ the indicator variables $t_{x.y}$, $t^{nl}_{x.y}$, and $t^{f}_{x.y}$ indicate their existence and non-existence (i.e. requirement and non-requirement).[49] Now we define the superset S_{2s} of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $T_{x.y} \in T_{2s}$, $T^{nl}_{x.y} \in T_{2s}$, and $T^{f}_{x.y} \in T_{2s}$: $S_{2s} := \{ s = (t_{1.1}, t_{1.2}, t^{nl}_{1.3}, t_{2.1}, t_{2.2}, t^{nl}_{4.1}, t^{nl}_{4.2}, t^{f}_{4.2}) \mid t_{x.y} \in \{0,1\} \} = \{0,1\}^8.$ As relevant subset of state vectors fulfilling the level 2 “strict” criteria we define the set $L_{2s} := \{ s \in S_{2s} \mid t_{1.1} \cdot t_{1.2} \cdot t^{nl}_{1.3} \cdot t_{2.1} \cdot t_{2.2} \cdot t^{nl}_{4.2} \cdot t^{f}_{4.2} = 1$ $\vee \; t_{1.1} \cdot t_{1.2} \cdot t_{2.1} \cdot t_{2.2} \cdot t^{nl}_{4.1} \cdot t^{nl}_{4.2} \cdot t^{f}_{4.2} = 1 \}.$ [50]

Source: own definition based on ISWGNA (1996)

The above definition of T_{2s}, S_{2s}, and L_{2s} is delineated from the description of that milestone level, which reads: “Gross national income and other primary indicators.

[49] See: chapter 3.2.3.

[50] That is all factors $t_{x.y}$ in these products are =1. That does mean, $s \in \boldsymbol{S}_{2s}$ signals that all of the components $t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}, t^{nl}_{4.2}, t^{f}_{4.2}$ and at least one of $t^{nl}_{1.3}$ and $t^{nl}_{4.1}$ have to be 1.

External account of primary incomes and current transfers, capital and financial accounts for rest of the world", together with the detail, that these "countries can calculate gross national income, national disposable income, national saving and net lending/borrowing" (ISWGNA 1996, 2 and 1, respectively). Thus the "strict" milestone level 2 is an extension of the data required for the "strict" milestone level 1. Compared to the NAQ tables that comprise the 1993 MRDS we see that the "strict" milestone 2 does not require NAQ table 2.3. On the other hand the *financial account* for NAQ table 4.2 is required here, which is not part of the 1993 MRDS requirements. Furthermore, the NAQ tables 1.3, 4.1 and 4.2 are required until the "*net lending*" indicator of the capital account (NL), as is the criteria required by the 2008 MRDS.

Identical to the 1993 and 2008 MRDS the existence of any SNA indicator within the NAQ table $T_{x.y}$ is sufficient if not otherwise indicated. The requirement until NL is symbolized by $T^{nl}_{x.y}$ for the NAQ table, and the requirement of the *financial account* is symbolized by $T^{f}_{x.y}$ for the NAQ table. (Since net lending of the financial account is computed from the financial indicators, the availability of financial account data is evaluated separately, i.e. in addition to net lending of the capital account.)

Note that:

> from $t^{nl}_{x.y} = 1$ respectively $t^{f}_{x.y} = 1$ follows necessarily that $t_{x.y} = 1$ (since $t_{x.y}$ indicates the existence of any number equal or greater than one of SNA indicators within the NAQ table $T_{x.y}$).
>
> But $t_{x.y} = 1$ does not imply $t^{nl}_{x.y} = 1$ or $t^{f}_{x.y} = 1$.

The latter does mean: the values for $t^{nl}_{x.y}$ and $t^{f}_{x.y}$ are independent of each other (i.e. the one does not imply the other. A country may for example produce SNA data until *net lending* of the *capital account*, yet not the *financial account* and vice versa.)

"Strict" Milestone Level 2 Fulfillment by the Available Country Data

To describe whether or not a country $c \in C$ fulfills the "strict" milestone level 2 criteria we define the mapping $\mathbf{F}_{2s} : C \rightarrow S_{2s}$ with $S_{2s} = \{0,1\}^8$,

$$c \in C \mapsto \mathbf{F}_{2s}(c) = \left(t_{1.1}, t_{1.2}, t^{nl}_{1.3}, t_{2.1}, t_{2.2}, t^{nl}_{4.1}, t^{nl}_{4.2}, t^{f}_{4.2} \right) \qquad \text{(II.91)}$$

and

$$\begin{array}{l} t_{x.y} = F_{x.y}(c) \\ t^{nl}_{x.y} = F^{nl}_{x.y}(c) \\ t^{f}_{x.y} = F^{f}_{x.y}(c) \end{array} = \begin{cases} 1 \text{ if NAQ table } T_{x.y}\text{, respectively } T^{nl}_{x.y}\text{, respectively} \\ T^{f}_{x.y} \text{ is available for } c \in C\text{, and} \\ 0 \text{ else,} \end{cases}$$

$\mathbf{F}_{2s}(c)$ indicates which of the tables of T_{2s} are available in c and which are not.

Then the set of countries meeting the "strict" milestone level 2 criteria are represented as the preimage set of $L_{2s} \subset S_{2s}$, i.e. by

$$C_{2s} := \mathbf{F}_{2s}^{-1}(L_{2s}) = \{ c \in C \mid \mathbf{F}_{2s}(c) \in L_{2s} \}. \qquad \text{(II.92)}$$

Thus $\mathbf{F}_{2s}(c) = s \in L_{2s}$ indicates that country $c \in C$ fulfills "strict" milestone level 2.

Further we define the indicator function

$$I_{2s} : C \rightarrow \{0,1\} \quad \text{with} \qquad I_{2s}(c) = \begin{cases} 1 \text{ if } c \in C_{2s} \\ 0 \text{ else.} \end{cases} \qquad \text{(II.93)}$$

Thus we have

$$C_{2s} = I_{2s}^{-1}(1). \qquad \text{(II.94)}$$

We calculate the number of countries which fulfill the "strict" milestone measure 2. Thus, we are interested in the cardinality of C_{2s}, that is the sum of $I_{2s}(c)$ over all $c \in C$, calculated by:

$$\left| C_{2s} \right| = \sum_{c \in C} I_{2s}(c). \qquad \text{(II.95)}$$

The share h_{2s} is calculated analougously to Formula II.25, by the cardinality of C_{2s} in relation to the total number of countries C.

The following table reflects the availability matrix of produced NAQ tables $T_{x.y}$ of a country $c \in C$, for all of the three cases that fulfill the "strict" milestone level 2 criteria. The availability regarding each NAQ table is indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{2s}(c)$ for $T_{x.y}$, and by $t_{x.y} = F^{nl}_{x.y}(c)$, the component functions of $\mathbf{F}_{2s}(c)$ for $T^{nl}_{x.y}$, and by $t_{x.y} = F^{f}_{x.y}(c)$, the component function of $\mathbf{F}_{2s}(c)$ for $T^{f}_{x.y}$.

Table 12 Availability Matrix of $T_{x.y} \in T_{2s}$ for the Cases where L_{2s} Fulfills Milestone Level 2 "Strict"

$t_{x.y}$ / $\mathbf{F}_{2s}(c)$	$t_{1.1}$	$t_{1.2}$	$t^{nl}_{1.3}$	$t_{2.1}$	$t_{2.2}$	$t^{nl}_{4.1}$	$t^{nl}_{4.2}$	$t^{f}_{4.2}$
s_1	1	1	1	1	1	1	1	1
s_2	1	1	1	1	1	0	1	1
s_3	1	1	0	1	1	1	1	1
$\mathbf{F}_{2s}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{2s} : C \to S_{2s}$ indicating for $s_1, \ldots, s_3$ which tables are available or not.								

Source: own compilation

Likewise to the earlier 1993 MRDS measure, the "strict" milestone level 2 requires the availability of only one of the NAQ tables 1.3 or 4.2. Therefore, though these availability measures comprise of different sets of NAQ tables, the differences between the scenarios that determine the fulfillment of the measures are analogue.

Example: The fulfillment of the "strict" milestone level 2 criteria is assessed analogous to the 1993 MRDS measure. Here we provide twp examples, for a comprehensive example refer to chapter 3.2.1.1, p. 59ff. The scenarios show the results of the function $\mathbf{F}_{2s}(c)$ and $I_{2s}(c)$ for a specific country $c \in C$.

Scenario 1: Only NAQ table 1.3 is missing[51], but table 4.1 is available.

$\mathbf{F}_{2s}(c) = (1,1,0,1,1,1,1,1) = s_3 \in L_{2s}$, that is $I_{2s}(c) = 1$.

Scenario 2: NAQ table 1.1 is missing, thus "strict" milestone level 2 is not fulfilled.

$\mathbf{F}_{2s}(c) = (0,1,1,1,1,1,1,1) = s \notin L_{2s}$, that is $I_{2s}(c) = 0$.

[51] For NAQ tables 1.3 and 4.1, the missing of a NAQ table does mean that is isn't available until the net lending indicator of the capital account.

3.2.5.2 “Relaxed” Milestone Levels

This subsection provides the milestone measures given the original “relaxed” criteria.

“Relaxed” Milestone Level 1

DEF. 30: Milestone level 1 “Relaxed”
The combination of NAQ tables $T_{x.y}$ as objects, that make up the “relaxed” milestone level 1 is given by the existence (requirement) of at least one of the NAQ tables 1.1 and 1.2 along with at least one of the NAQ tables 2.1 and 2.2 (cf. United Nations 2001a, 5). Milestone level 1 “relaxed” is formally delineated as follows: Let the set of potential NAQ tables be $T_{1r} := \{ T_{1.1}, T_{1.2}, T_{2.1}, T_{2.2} \}$. Regarding the NAQ tables $T_{x.y}$ the existence and non-existence (i.e. requirement and non-requirement) is indicated by the indicator variable $t_{x.y}$. Now we define the superset S_{1r} of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $T_{x.y} \in T_{1r}$: $S_{1r} := \{s = (t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}) \mid t_{x.y} \in \{0,1\}\} = \{0,1\}^4$. As relevant subset of state vectors we define for milestone level 1 “relaxed” the set $L_{1r} := \{ s \in S_{1r} \mid t_{1.1} * t_{2.1} = 1, \vee t_{1.1} * t_{2.2} = 1, \vee t_{1.2} * t_{2.1} = 1, \vee t_{1.2} * t_{2.2} = 1\}$. [52]

Source: own definition based on ISWGNA (1996) and United Nations (2001a, 5)

The above definition of T_{1r}, S_{1r}, and L_{1r} for the “relaxed” milestone level 1 is delineated from the note regarding the “relaxed” measure for milestone level 1 in United Nations (2001a, 5) which now only requires *current* or *constant* prices data on “final expenditures on GDP” along with “GDP by industry”, i.e. only *one* of the two types of prices is required, but for both approaches for GDP computation. The 1996

[52] That is all factors $t_{x.y}$ in these products are =1. That does mean $s \in \boldsymbol{S}_{1r}$ signals that at least one of the components $t_{1.1}, t_{1.2}$ and at least one of the components $t_{2.1}$ and $t_{2.2}$ have to be 1.

definition required data in both types of prices for the two approaches for GDP computation, i.e. in current *and* constant prices (cf. ISWGNA 1996)[53].

"Relaxed" Milestone Level 1 Fulfillment by the Available Country Data

To describe whether or not a country $c \in C$ fulfills the "relaxed" milestone level 1 criteria we define the mapping $\mathbf{F}_{1r} : C \rightarrow S_{1r}$ with $S_{1r} = \{0,1\}^4$,

$$c \in C \mapsto \mathbf{F}_{1r}(c) = \left(t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}\right) \tag{II.96}$$

and

$$t_{x.y} = F_{x.y}(c) = \begin{cases} 1 \text{ if NAQ table } T_{x.y} \text{ is available for } c \in C \text{, and} \\ 0 \text{ else.} \end{cases}$$

With other words, $\mathbf{F}_{1r}(c)$ indicates which of the tables of T_{1r} are available in c (that is produced by c) and which are not.

Then the countries meeting the "relaxed" milestone level 1 criteria represented as the members of the preimage set of $L_{1r} \subset S_{1r}$, i.e. by

$$C_{1r} := \mathbf{F}_{1r}^{-1}(L_{1r}) = \{c \in C \mid \mathbf{F}_{1r}(c) \in L_{1r}\}. \tag{II.97}$$

Thus $\mathbf{F}_{2s}(c) = s \in L_{1r}$ indicates for a specific country $c \in C$ that it fulfills the "relaxed" milestone level 1 criteria.

Further we define the indicator function

$$I_{1r} : C \rightarrow \{0,1\} \quad \text{with} \quad I_{1r}(c) = \begin{cases} 1 \text{ if } c \in C_{1r} \\ 0 \text{ else.} \end{cases} \tag{II.98}$$

Thus,

$$C_{1r} = I_{1r}^{-1}(1). \tag{II.99}$$

[53] See also the note to the milestone measures in United Nations (2004, 20 and 2005, 19).

As the result of interest for the assessment by regions and country groups, we calculate the number of countries which fulfill the “relaxed” milestone level 1. Thus we are interested in the cardinality of C_{1r}, that is the sum of $I_{1r}(c)$ over all $c \in C$:

$$\left|C_{1r}\right| = \sum_{c \in C} I_{1r}(c) . \tag{II.100}$$

The share h_{1r} is calculated analougously to Formula II.25, by the cardinality of C_{1r} in relation to the total number of countries C.

The following table presents the availability matrix of produced NAQ tables $T_{x.y}$ of a country $c \in C$ for each possible $\mathbf{F}_{1r}(c) \in L_{1r}$. The availability respectively missing of each NAQ table $T_{x.y}$ in $c \in C$ is indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{1r}(c)$ for $T_{x.y}$. (The row vectors s_4 to s_6 and s_7 to s_9 have the same availability regarding tables 2.1 and 2.2 as row vectors s_1 to s_3, yet with either table 1.2 or 1.1 missing, respectively. For easier visibility of the differences between the row vectors the “0” are printed in bold font.)

Table 13 Availability Matrix of $T_{x.y} \in T_{1r}$ for the Cases where L_{1r} Fulfills Milestone Level 1 “Relaxed”

$t_{x.y}$ / $\mathbf{F}_{1r}(c)$	$t_{1.1}$	$t_{1.2}$	$t_{2.1}$	$t_{2.2}$
s_1	1	1	1	1
s_2	1	1	1	**0**
s_3	1	1	**0**	1
s_4	1	**0**	1	1
s_5	1	**0**	1	**0**
s_6	1	**0**	**0**	1
s_7	**0**	1	1	1
s_8	**0**	1	1	**0**
s_9	**0**	1	**0**	1
$\mathbf{F}_{1r}(c)$: the image of $c \in C$ under the map $\mathbf{F}_{1r} : C \rightarrow S_{1r}$ indicating for $s_1, \ldots, s_9$ which tables are available or not.				

Source: own compilation

Example: The fulfillment of the „relaxed" milestone level 1 is assessed analogous to the previous milestone measure. One example for the fulfillment and one for the non-fulfillment are provided, for a comprehensive example refer to chapter 3.2.1.1, p. 59. The exemplary scenarios show the results of the function $\mathbf{F}_{1r}(c)$ and $I_{1r}(c)$ for a specific country $c \in C$.

Scenario 1: NAQ table 1.1 is missing, but 1.2 is available, further NAQ table 2.1 is missing, but 2.2 is available (which corresponds to row vector 9).

$\mathbf{F}_{1r}(c) = (0,1,0,1) = s_9 \in L_{1r}$, that is $I_{1r}(c) = 1$.

Scenario 2: NAQ table 1.1 and 1.2 are missing.
$\mathbf{F}_{1r}(c) = (0,0,1,1) \notin L_{1r}$, that is $I_{1r}(c) = 0$.

If, like in scenario 2, both of the NAQ tables 1.1 or 1.2 are missing the fulfillment of the "relaxed" milestone level 1 does not depend on the non- existence of the NAQ tables 2.1 or 2.2, i.e. of $T_{2.1}$ or $T_{2.2}$, indicated by $t_{2.1}$ and $t_{2.2}$, respectively.

That is, for $c \in C$ with:

$\mathbf{F}_{1r}(c) = \left(t_{1.1}, t_{1.2}, t_{2.1}, t_{2.2}\right) \notin L_{1r}$ because of
$t_{x.y} = F_{x.y} = 0$ for both of the indices 1.1, 1.2 and therefore
$F_{1.1}(c) = 0$, further $F_{1.2}(c) = 0$.

Therefore,

$t_{1.1} * t_{2.1} = 0$, as well as
$t_{1.1} * t_{2.2} = 0$, as well as
$t_{1.2} * t_{2.1} = 0$ as well as
$t_{1.2} * t_{2.2} = 0$.

Therefore $I_{1r}(c) = 0$.

“Relaxed” Milestone Level 2

DEF. 31: Milestone level 2 “Relaxed”

The combination of NAQ tables $T_{x.y}$ as objects, that make up the “relaxed” milestone level 2 is given by the existence (requirement) of the “relaxed” milestone level 1 (at least one of the NAQ tables 1.1 and 1.2 along with at least one of the NAQ tables 2.1 and 2.2), along with at least one of the NAQ tables 1.3 and 4.1 (cf. ISWGNA 1996 in conjunction with UNSD 2008).

Milestone level 2 “relaxed” is formally delineated as follows:

Let the set of potential NAQ tables be

$$\mathcal{T}_{2r} := \{ T_{1.1}, T_{1.2}, T_{1.3}, T_{2.1}, T_{2.2}, T_{4.1} \}.$$

Regarding the NAQ tables $T_{x.y}$ the existence and non-existence (i.e. requirement and non-requirement) is indicated by the indicator variable $t_{x.y}$.

Now we define the superset $\mathcal{S}_{2r}$ of all state vectors (n-tuples) s, indicating the existence and non-existence, respectively, of each NAQ table $T_{x.y} \in \mathcal{T}_{2r}$:

$$\mathcal{S}_{2r} := \{ s = (t_{1.1}, t_{1.2}, t_{1.3}, t_{2.1}, t_{2.2}, t_{4.1}) \mid t_{x.y} \in \{0,1\} \} = \{0,1\}^6.$$

As relevant subset of state vectors we define for milestone level 2 “relaxed” the set

$$\begin{aligned} \mathcal{L}_{2r} := \{ s \in \mathcal{S}_{2r} \mid\ & t_{1.1} \cdot t_{1.3} \cdot t_{2.1} = 1 \ \vee\ t_{1.1} \cdot t_{1.3} \cdot t_{2.2} = 1 \\ & \vee\ t_{1.1} \cdot t_{2.1} \cdot t_{4.1} = 1 \ \vee\ t_{1.1} \cdot t_{2.2} \cdot t_{4.1} = 1 \\ & \vee\ t_{1.2} \cdot t_{1.3} \cdot t_{2.1} = 1 \ \vee\ t_{1.2} \cdot t_{1.3} \cdot t_{2.2} \cdot = 1 \\ & \vee\ t_{1.2} \cdot t_{2.1} \cdot t_{4.1} = 1 \ \vee\ t_{1.2} \cdot t_{2.2} \cdot t_{4.1} = 1 \}. \end{aligned}$$ [54]

Source: own definition based on based on ISWGNA (1996) in conjunction with UNSD (2008)

The above definition of $\mathcal{T}_{2r}$, $\mathcal{S}_{2r}$, and $\mathcal{L}_{2r}$ is delineated from the current methodology applied by the UN to assess the milestone measure 2 (which we identified as the “relaxed” version). The statement regarding the data availability of a country that

[54] That is all factors $t_{x.y}$ in these products are =1. That does mean $s \in \boldsymbol{S}_{2r}$ signals that at least one of the components $t_{1.1}$, $t_{1.2}$ and at least one of $t_{2.1}$ and $t_{2.2}$ and at least one of $t_{1.3}$ and $t_{4.1}$ have to be 1.

fulfills this milestone measure reads: “With this milestone, a reporting country provides at least value added by industry and gross domestic product by expenditures, either in current or constant prices, and gross national income.” (UNSD 2008, 9).

The “relaxed” milestone level 2 is an extension of the data requirement for the “relaxed” milestone level 1. Yet, it does not require as much additional data as in case of the “strict” milestone level 2. It is no longer necessary to evaluate the availability of the NAQ table 4.2 at all. Further, the statement relaxes the data requirements regarding the completeness of the NA indicators within the needed NAQ table. The statement identifies which NA indicator is at least available for the country; this is “*gross national income*”. “*Gross national income*” marks the first balance following GDP regarding the balances from GDP to the net lending aggregate. As the first balance, this is the minimum information that would be provided by a country within the NAQ table 1.3 “Relations among product, income, saving, and net lending aggregates”, else wise none of the data are computed for that NAQ table. Consequently, regarding the specification of the completeness of that NAQ table, no specific NA indicator needs to be identified as requirement. This applies also for NAQ table 4.1, since 1.3 and 4.1 are substitutional NAQ tables. Thus to fulfill the “relaxed” milestone level 2, the mere availability of any NA indicator for these tables is sufficient. (For a detailed explanation refer to part III, chapter 2.2.2 Availability Conditions of NAQ Tables, p. 203.)

“Relaxed” Milestone Level 2 Fulfillment by the Available Country Data

To describe whether or not a country $c \in C$ fulfills the “relaxed” milestone level 2 criteria we define the mapping $\mathbf{F}_{2r} : C \rightarrow S_{2r}$ with $S_{2r} = \{0,1\}^6$,

$$c \in C \mapsto \mathbf{F}_{2r}(c) = \left(t_{1.1}, t_{1.2}, t_{1.3}, t_{2.1}, t_{2.2}, t_{4.1}\right) \quad \text{(II.101)}$$

and

$$t_{x.y} = F_{x.y}(c) = \begin{cases} 1 \text{ if NAQ table } T_{x.y} \text{ is available for } c \in C\text{, and} \\ 0 \text{ else.} \end{cases}$$

With other words, $\mathbf{F}_{2r}(c)$ indicates which of the tables of T_{2r} are available in c (that is produced by c) and which are not.

Then the countries meeting the "relaxed" milestone level 2 criteria are represented as the members of the preimage set of $L_{2r} \subset S_{2r}$, i.e. by

$$C_{2r} := \mathbf{F}_{2r}^{-1}(L_{2r}) = \{ c \in C \mid \mathbf{F}_{2r}(c) \in L_{2r} \}. \qquad \text{(II.102)}$$

Thus $\mathbf{F}_{2r}(c) = s \in L_{2r}$ indicates for a specific country $c \in C$ that it fulfills the "relaxed" milestone level 2 criteria.

Further we define the indicator function

$$I_{2r} : C \rightarrow \{0,1\} \quad \text{with} \qquad I_{2r}(c) = \begin{cases} 1 \text{ if } c \in C_{2r} \\ 0 \text{ else.} \end{cases} \qquad \text{(II.103)}$$

Thus we have

$$C_{2r} = I_{2r}^{-1}(1). \qquad \text{(II.104)}$$

As the result of interest for the assessment by regions and country groups, we calculate the number of countries which fulfill the "relaxed" milestone measure 2. Thus we are interested in the cardinality of C_{2r}, that is the sum of $I_{2r}(c)$ over all $c \in C$:

$$|C_{2r}| = \sum_{c \in C} I_{2r}(c). \qquad \text{(II.105)}$$

The share h_{2r} is calculated analougously to Formula II.25, by the cardinality of C_{2r} in relation to the total number of countries C.

The following table presents the availability matrix for each possible $\mathbf{F}_{2r}(c) \in L_{2r}$ and $T_{x.y}$ indicated by $t_{x.y} = F_{x.y}(c)$, the component functions of $\mathbf{F}_{2r}(c)$ for $T_{x.y}$. (The row vectors s_3 to s_4 and s_5 to s_6 are similar to s_1 to s_2, yet with one of the tables 2.2 and 2.1 missing, respectively. Row vectors s_{10} to s_{18} and s_{19} to s_{27} are similar to s_1 to s_9, yet with one of the tables 1.2 and 1.1 missing, respectively. For easier visibility of the differences between the row vectors the "0" are printed in bold font.)

Table 14 Availability Matrix of $T_{x.y} \in T_{2r}$ for the Cases where L_{2r} Fulfills Milestone Level 2 "Relaxed"

$t_{x.y}$ / $\mathbf{F}_{2r}(c)$	$t_{1.1}$	$t_{1.2}$	$t_{1.3}$	$t_{2.1}$	$t_{2.2}$	$t_{4.1}$
s_1	1	1	1	1	1	1
s_2	1	1	1	1	1	**0**
s_3	1	1	1	1	**0**	1
s_4	1	1	1	1	**0**	**0**
s_5	1	1	1	**0**	1	1
s_6	1	1	1	**0**	1	**0**
s_7	1	1	**0**	1	1	1
s_8	1	1	**0**	1	**0**	1
s_9	1	1	**0**	**0**	1	1
s_{10}	1	**0**	1	1	1	1
s_{11}	1	**0**	1	1	1	**0**
s_{12}	1	**0**	1	1	**0**	1
s_{13}	1	**0**	1	1	**0**	**0**
s_{14}	1	**0**	1	**0**	1	1
s_{15}	1	**0**	1	**0**	1	**0**
s_{16}	1	**0**	**0**	1	1	1
s_{17}	1	**0**	**0**	1	**0**	1
s_{18}	1	**0**	**0**	**0**	1	1
s_{19}	**0**	1	1	1	1	1
s_{20}	**0**	1	1	1	1	**0**
s_{21}	**0**	1	1	1	**0**	1
s_{22}	**0**	1	1	1	**0**	**0**
s_{23}	**0**	1	1	**0**	1	1
s_{24}	**0**	1	1	**0**	1	**0**
s_{25}	**0**	1	**0**	1	1	1
s_{26}	**0**	1	**0**	1	**0**	1
s_{27}	**0**	1	**0**	**0**	1	1

$\mathbf{F}_{2r}(c)$: the image of $c \in C$ under the map
$\mathbf{F}_{2r} : C \rightarrow S_{2r}$ indicating for $s_1, \ldots, s_{27}$ which tables are available or not.

Source: own compilation

Example: The fulfillment of the "relaxed" milestone level 2 criteria is exemplarily assessed for one scenario where a country $c \in C$ fulfills the criteria and one where it does not. The fulfillment is assessed analogous to previous measures. For a comprehensive example refer to chapter 3.2.1.1, p. 59. The scenarios show the results of the function $\mathbf{F}_{2r}(c)$ and $I_{2r}(c)$ for a specific country $c \in C$.

Scenario 1: NAQ table 2.2 is missing, yet 2.1 is available, and 4.1 is missing, but table 1.3 is available (which corresponds to row vector s_4).

$\mathbf{F}_{2r}(c) = (1,1,1,1,0,0) = s_4 \in L_{2r}$, that is $I_{2r}(c) = 1$.

Scenario 2: NAQ table 2.1 and 2.2 are missing.

$\mathbf{F}_{2r}(c) = (1,1,1,0,0,1) \notin L_{2r}$, that is $I_{2r}(c) = 0$.

Here, the non-fulfillment of "relaxed" milestone level 2 does not depend on the non-existence of $T_{1.1}$ and $T_{1.2}$, or $T_{1.3}$ and $T_{4.1}$, indicated by $t_{1.1}$ and $t_{1.2}$, and $t_{1.3}$ and $t_{4.1}$.

3.2.5.3 Relations among Country Sets that Fulfilled the Milestone Levels

In the following, we illustrated the relations between the country sets that fulfill the four milestone levels that we defined in the previous sections (cf. Figure 13, p. 97).

Based on our milestone level definitions, we can now describe the relations between the preimage sets of C_{2s}, C_{2r}, C_{1s}, and C_{1r}, i.e. of the sets of countries that fulfill the requirements of the corresponding milestone levels.

The sets presented in Figure 14 are defined via their preimage sets, i.e.

$$C_{2s} := \mathbf{F}_{2s}^{-1}(L_{2s}),\ C_{2r} := \mathbf{F}_{2r}^{-1}(L_{2r}),\ C_{1s} := \mathbf{F}_{1s}^{-1}(L_{1s}), \text{ and } C_{1r} := \mathbf{F}_{1rs}^{-1}(L_{1r}).$$

As relations between these country sets we have:

$$C_{2s} \subseteq C_{1s} \subseteq C_{1r}. \qquad \text{(II.106)}$$

$$\text{And } C_{2s} \subseteq C_{2r} \subseteq C_{1r}. \qquad \text{(II.107)}$$

Note: In a strict mathematical sense, there is no subset relation between C_{2r} and C_{1s}.

With other words, a country that fulfills "strict" milestone level 2 already fulfills the lower requirements of "strict" milestone level 1 and also those of "relaxed" milestone level 2. It is therefore a member of the set of countries that fulfill all three measures.

On the other hand, fulfilling the "strict" milestone level 2 includes the fulfillment of the "relaxed" milestone level 1 via either of two aforementioned milestone measures, i.e. levels 1 "strict" or 2 "relaxed". Hence, a country that fulfills the "strict" milestone level 2 or either one of levels 1 "strict" or 2 "relaxed" is also member of the preimage set of countries that fulfilled the "relaxed" milestone level 1.

Yet note that the requirement criteria of the “relaxed” milestone level 2 are fulfilled without the necessity of the availability of all the NAQ tables 1.1, 1.2, 2.1 and 2.2. Consequently, there is no mathematical subset relation between the preimage set of the countries that fulfilled the “relaxed” milestone level 2 and those that fulfilled the “strict” milestone level 1, which requires the existence of all four NAQ tables. The set of countries that fulfills the “relaxed” milestone level 2 and the “strict” milestone level 1 is given by the set $C_{2r} \cap C_{1s}$. In the latter case, the set of NAQ tables required to fulfill the “relaxed” milestone level 2 are available for a country, while particularly all of the NAQ tables 1.1, 1.2, 2.1 and 2.2 are available (which corresponds to the row vectors 1, 2, and 7 of Table 14, p. 112).

Illustrated by the following figure are the sets C_{2s}, C_{2r}, C_{1s}, and C_{1r}.

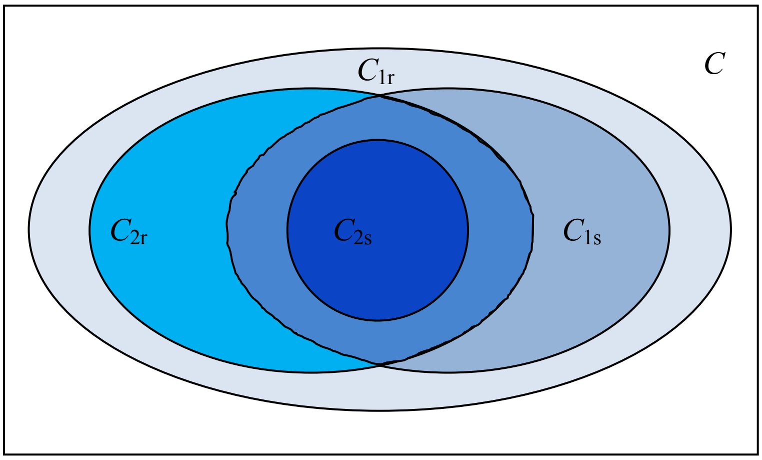


Figure 14 Country Sets Relating to the Fulfillment of Milestone Levels

Source: own illustration

Explanation of the coloring for the sets in Figure 12 (from dark blue to light blue).

- $C_{2s} := \mathbf{F}_{2s}^{-1}(L_{2s})$ <dark blue>, countries fulfilling the “strict” milestone level 2.
- $(C_{2r} \cap C_{1s}) \setminus C_{2008}$, countries fulfilling the “relaxed” milestone level 2 and “strict” milestone level 1 criterion, but not the “strict” milestone level 2.
- $C_{2r} \setminus C_{1s}$, set of countries fulfilling the “relaxed” milestone level 2, missing at least one criterion of the “strict” milestone level 1.
- $C_{1s} \setminus C_{2r}$, set of countries fulfilling all “strict” milestone level 1criteria but missing the additional “relaxed” milestone level 2 criteria.
- $C_{1r} \setminus (C_{2r} \cup C_{1s})$ <light blue>, set of countries fulfilling the “relaxed” milestone level 1 criteria, but missing at least one criterion of the “strict” milestone level 1 and at least one of the “relaxed” milestone level 2 criteria.

Sample Application 10: Milestone Measures

Example 1: Cameroon

Building upon the data availability in Sample Application 6, we consider the availability of the required NAQ tables regarding the fulfillment of the "strict" milestone levels 1 and 2, and the "relaxed" milestone levels 1 and 2.

Cameroon fulfills the "strict" milestone levels 1 and 2 (as defined in DEF. 28, p. 98, respectively DEF. 29, p. 101). Necessarily it also fulfills the less demanding "relaxed" milestone levels 1 and 2 (DEF. 30, p. 105, respectively DEF. 31, p. 109).

Referring to the formulas (DEF. 28 to DEF. 31), defined by $L_{1\text{s}}$, $L_{2\text{s}}$, $L_{1\text{r}}$, and $L_{2\text{r}}$ are the relevant sets of state vectors (within the superset of potential state vectors) that meet the requirement criteria of the different milestone level measures. Further, the vectorial mappings relating to formulas II.85, II.91, II.96, II.101 equate to:

$\mathbf{F}_{1\text{s}}(\text{CMR}) = (1,1,1,1)$ for milestone level 1 "strict",

$\mathbf{F}_{2\text{s}}(\text{CMR}) = (1,1,1,1,1,1,1,1)$ for milestone level 2 "strict",

$\mathbf{F}_{1\text{r}}(\text{CMR}) = (1,1,1,1)$ for milestone level 1 "relaxed",

$\mathbf{F}_{2\text{r}}(\text{CMR}) = (1,1,1,1,1,1)$ for milestone level 2 "relaxed".

Fulfillment of the Milestone Measures by Cameroon

$\boldsymbol{L}$	$\mathbf{F}_{\boldsymbol{L}} \in \boldsymbol{L}$
$L_{1\text{s}}$	✓
$L_{2\text{s}}$	✓
$L_{1\text{r}}$	✓
$L_{2\text{r}}$	✓

Thus, because $\mathbf{F}_{1\text{s}}(\text{CMR}) \in L_{1\text{s}}$, Cameroon is member of the set $C_{1\text{s}}$ of countries that fulfilled the criteria of "strict" milestone level 1 and the indicator function signaling the membership in $C_{1\text{s}}$, $\text{I}_{1\text{s}}(\text{CMR}) = 1$. Analogous, Cameroon is member of the sets $C_{2\text{s}}$, $C_{1\text{r}}$, $C_{2\text{r}}$, and $\text{I}_{2\text{s}}(\text{CMR}) = 1$, respectively $\text{I}_{1\text{r}}(\text{CMR}) = 1$, and $\text{I}_{2\text{r}}(\text{CMR}) = 1$.

That does mean Cameroon is included in all four sets of countries, it fulfills the four milestone level measures, and the indicator functions indicate the membership.

(continued on next page)

Example 2: Thailand

Thailand fulfills the milestone level 1 “strict”, necessarily also level 1 “relaxed”. Furthermore, it fulfills the “relaxed” milestone level 2 since the criteria only require the existence of either NAQ table 4.1 or 1.3, with the latter being available. While the missing of NAQ table 4.1 until NL is compensated for by the availability of NAQ table 1.3 until NL, Thailand does not fulfill the criteria for milestone level 2 “strict”, because the NAQ table 4.2 and specifically the financial accounts of NAQ table 4.2 are missing.

Referring to the formulas (see Example 1), this equates to the following vectorial mappings:

$\mathbf{F}_{1s}(\text{THA}) = (1,1,1,1);$

$\mathbf{F}_{2s}(\text{THA}) = (1,1,1,1,1,0,0,0);$

$\mathbf{F}_{1r}(\text{THA}) = (1,1,1,1);$

$\mathbf{F}_{2r}(\text{THA}) = (1,1,1,1,1,0).$

Fulfillment of the Milestone Measures by Thailand

$\boldsymbol{L}$	$\mathbf{F}_{\boldsymbol{L}} \in \boldsymbol{L}$
L_{1s}	✓
L_{2s}	×
L_{1r}	✓
L_{2r}	✓

Consequently Thailand is member of the preimage sets C_{1s}, C_{1r}, and C_{2r}, and not of C_{2s}. Further, the indicator functions $\mathrm{I}_{1s}(\text{THA}) = 1$, $\mathrm{I}_{1r}(\text{THA}) = 1$, $\mathrm{I}_{2r}(\text{THA}) = 1$, but $\mathrm{I}_{2s}(\text{THA}) = 0$.

3.3 Timeliness Measures

In this chapter we provide the *DQ measures* for the *"filtered" timelines* of ANA data publication by the IO, and for the *time lag of reported ANA data*. Further, we define the formulas that assign (*map*) the country data to *timeliness* and *time lag* values, and the formulas for the *sets of countries* with a specific *timeliness* of value $n \in \boldsymbol{N}$ in years.

3.3.1 **Timeliness of Data Publication**

DEF. 32: Timeliness of Annual National Accounts Publication (T)
The length of the time period (T) measured in years between the end time point of the reference year (t_{RY}) for which ANA data are reported by the NSO to the IO, and the time point of the first time availability of these ANA data at the international level (t_{PY}) (i.e. their first time publication in the YB). Formally expressed: C_R := subset ($C_R \subseteq C$) of countries $c \in C$, that reported a set of ANA indicators for a new reference year $t_{RY}(c)$, previously not published in the Yearbook (YB). *ANA* := set of ANA data (NA indicators and corresponding reference years (t_{RY})) $\mathrm{T} : ANA \times C_R \longrightarrow \boldsymbol{N}$ (II.108) $(\mathrm{ANA,c}) \in ANA \times C_R \mapsto \mathrm{T(ANA,c)} = t_{PY}(\mathrm{ANA,c}) - t_{RY}(\mathrm{ANA,c})$ [55], where $\mathrm{T(ANA,c)}$ [55]: the timeliness of ANA data in years $n \in \boldsymbol{N}$ referring to the most recent reference year within the present set of ANA data of country $c \in C_R$, when published by the UN for the first time, $t_{PY}(\mathrm{ANA,c})$[55]: the year when the YB volume is published (with a new reference year of ANA data for country $c \in C_R \subseteq C_{YB}$), and $t_{RY}(\mathrm{ANAc})$ [55]: the most recent reference year within the set of published ANA data for country $c \in C_R \subseteq C_{YB}$, with $t_{PY}(c) > t_{RY}(c)$.

Source: own definition

[55] T(ANA,c) corresponds to T(x,c) (see: DEF. 5) with data object x = "ANA". The symbol "ANA" is dropped from now on for convenience, since we study ANA data. For $t_{PY}(\mathrm{ANA,c})$ and $t_{RY}(\mathrm{ANA,c})$, ditto.

The data object "ANA" of reported NA indicators and corresponding reference years, provides at least one new reference year of ANA data. It typically includes revisions of previous reference years of "known" data, too. Of interest for our *DQ assessment* is the *timeliness* of ANA data publication given the time point of the first release of a new reference year of ANA data. This means, not of interest is the *timeliness* of "known" data, i.e. ANA data for a reference that has previously been published via the YB. Also not of interest is the *timeliness* of a revised reference year of "known" data. Therefore, we do not consider data of a reference year that has been published in a previous YB for a given country (even in case the statistics were revised).

The year $t_{PY} - 1$ is the expected most recent reference year in the YB published[56] in year t_{PY}. Obviously, the year $t_{PY} - 1$ is the most recent reference year for which ANA data can be produced, thus this is the minimum *timeliness* of ANA data publication.

Example 1: The YB covering ANA data for country $c \in C_R$, that reported a new reference year of ANA data, is published in year 2010, i.e. $t_{PY}(c) = 2010$. Further, the most recent available reference year in the YB for country $c \in C_R$ is the past reference year, i.e. $t_{RY}(c) = 2009$. We get as the *timeliness* of the new published ANA data for country $c \in C_R$

$$T(c) = t_{PY}(c) - t_{RY}(c) = 2010 - 2009 = 1 \text{ (year)}.$$

Example 2: The YB is published in year 2010, thus $t_{PY}(c) = 2010$, and the most recent reference year in the YB for country $c \in C_R$ is 2007, thus $t_{RY}(c) = 2007$. Then we get

$$T(c) = 2010 - 2007 = 3 \text{ (years)}.$$

Rather than saying that the data for a country $c \in C_R$ that reported new ANA data has the *timeliness* T(c) of $n \in \boldsymbol{N}$ years, we can say that the most recent reference year (t_{RY}) of published ANA equals the publication year (t_{PY}) of the data minus n years. Thus instead of indicting the *value* of the *timeliness* by year(s) $n = t_{PY}(c) - t_{RY}(c)$ we express the *timeliness* of the most recent available reference year (t_{RY}) in relation to the publication year (t_{PY}) of the YB, i.e. by

$$t_{RY}(c) = t_{PY}(c) - n. \tag{II.109}$$

[56] Data reported to the UN after the YB publication (ca. September) are deferred to the next year's YB.

Our DQ measure *timeliness of ANA* data refers to a time period in years. Further, the time point t_{RY} for the reference year, as well as time point t_{PY} for the year of the YB publication, refer to calendar years. Since the YB is an annual publication, any publication year t_{PY} corresponds to a specific calendar year. To emphasize the *timeliness* in years of value $n \in \boldsymbol{N}$ of the published ANA data, rather than identifying the actual reference year t_{RY}, we abbreviate the above notation of Formula II.109 by writing

$$t_{RY}(c) = t - n. \qquad \text{(II.110)}$$

Thus we write that the ANA data of country $c \in C_R$ refer to the reference year "t-n", where $t := t_{PY}(c)$, and n indicates the time period in year(s) $n \in \boldsymbol{N}$, marking the *timeliness* of the ANA data in relation to the year $t = t_{PY}(c)$.

For our *timeliness* measure (cf. DEF. 32), the availability of country data at the international level is the subject of interest, i.e. their publication by an IO. Thus, not of interest is the data publication at the national level, i.e. by a national institution. The type of DQ *measure*, concerning country data publication at the international level, was introduced by Di Fonzo (2005b) as *"filtered" timeliness*. We use the term "*IO (filtered) timeliness*", or in short "*IO timeliness*" to refer to the *timeliness* of data publication for a country $c \in C_R$ that reported new data to an IO. The country data are produced by a national institution (e.g. NSO) and subsequently published by the IO. Likewise, "*NSO timeliness*" refers to the data publication at the national level.

The time point of data publication at the *international* level is typically later than the time point of publication at the *national* level. The *IO timeliness* is subject to the time required until the national data are submitted to the IO as well as the time needed for data processing by the IO, until the data dissemination through publications or online.

Due to the specifics of the publication process, the *IO (filtered) timeliness* of individual countries may include "extra waiting time", until the time point of publication of multiple countries at once. For example, in case of the UN with an annual YB publication, the data for all countries are published at once, say in September (considering the available ANA data for any country $c \in C$ for the past 12 reference years [t_{RY} = t-1, …, t-12]). The data of the individual countries have been successively received and processed over the previous months.

The data published in the YB of a country c $\in$ C_R that reported a new set of ANA data for the present YB, are expected to refer to the reference year t_{RY} = t-1, i.e. the benchmark reference period for ANA production by countries. Countries reporting a new set of ANA data with the most reference year t_{RY} referring to a reference year earlier than the year t-1, have a *time lag* (cf. DEF. 33 Time Lag of Reported Annual National Accounts Data (TLR), p. 122).

Measurement of Timeliness

Based on our definition of timeliness T(c) (cf. DEF. 32), the set of countries with a specific *timeliness* value n $\in$ $\boldsymbol{N}$ in years is the preimage set

$$C_n^T := T^{-1}(n) = \{ c \in C_R \mid T(c) = n \}. \qquad \text{(II.111)}$$

Further we define the indicator function regarding whether or not a specific country c is element of the preimage set C_n^T for a specific *timeliness* value given by n $\in$ $\boldsymbol{N}$ in years:

$$I_n^T : C_R \rightarrow \{0,1\} \text{ with } \qquad I_n^T(c) = \begin{cases} 1 \text{ if } c \in C_n^T \\ 0 \text{ else.} \end{cases} \qquad \text{(II.112)}$$

$$\text{Thus, } C_n^T = {I_n^T}^{-1}(1). \qquad \text{(II.113)}$$

For the regional analysis and the analysis of special country groups (e.g. Northern Africa or OECD countries), we calculate the number of countries in C_R with a particular *timeliness* of value n $\in$ $\boldsymbol{N}$ in years. Thus we are interested in the cardinality of C_n^T that is the sum of $I_n^T(c)$ over all c $\in$ C_R:

$$\left|C_n^T\right| = \sum_{c \in C_R} I_n^T(c). \qquad \text{(II.114)}$$

For the set C_R of considered countries, i.e. that reported new data, the share (h) of countries with a particular *timeliness* of value n $\in$ $\boldsymbol{N}$ in years is given by

$$h_n := \frac{\left|C_n^T\right|}{\left|C_R\right|} = \frac{\left|T^{-1}(n)\right|}{\left|C_R\right|}, \qquad \text{(with } 0 \le h_n \le 1\text{).} \qquad \text{(II.115)}$$

For most regions, the set C of countries that assemble a region (e.g. Northern Africa) or a country group (e.g. OECD) is larger than the set C_R of countries that actually reported a new reference year of data. Using the set $C_R \subseteq C$ as basis set to calculate the share h_n (cf. Formula II.115), implies that the total of shares h_n over all occurring values $n \in \boldsymbol{N}$ of *timeliness* in years is one, i.e. $\sum_{n=1}^{5} h_n = 1$ for n = 1, …, 5. Using the set C instead of set C_R as basis set to calculate the share h_n, in case of $C_R \subset C$, yields lower results for each *timeliness* value $n \in \boldsymbol{N}$ in years. The cumulative shares h_n remains below one. This is because the set C includes all those countries, which did not report new ANA data (i.e. the data published in the YB are "known") and it also includes all those countries that do not even produce NA.

The sets of countries $C_n^T = T^{-1}(n)$ with a specific *timeliness* T(c) = n = 1, 2, 3, 4 are illustrated in Figure 15. We use the color code "heat" to indicate the superior and inferior *timeliness* of ANA publication, where "brick red" indicates the most intense heat level, referring to the top-quality.

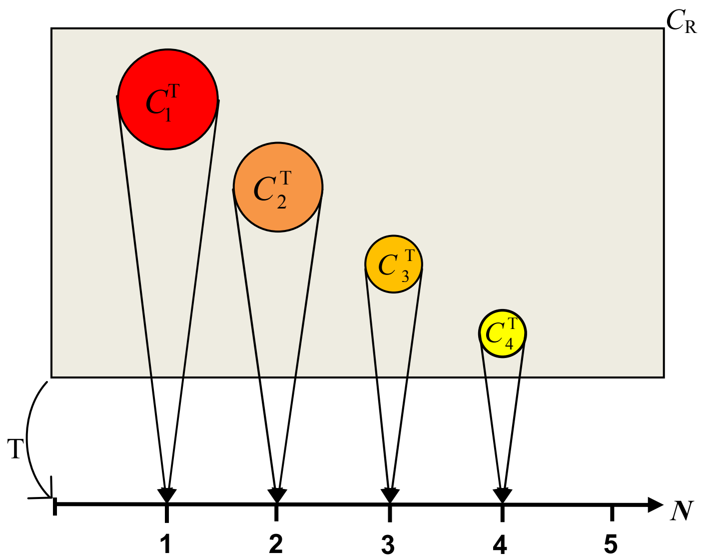


Figure 15 Preimage Sets of Countries by Timeliness (T) of value n ∈ *N* in Years

Source: own illustration

3.3.2 **Time Lag of Reported Data**

DEF. 33: Time Lag of Reported Annual National Accounts Data (TLR)
The delay by the NSO of a country $c \in C_R$ in reporting the ANA data to the IO, is given by the time period (TLR) measured in years $n \in \boldsymbol{N}$ between the expected year (t*) of reported ANA data ($t^* = t_{PY} - 1$) and the reported year of ANA data (t_{RY}). Formally expressed: $\text{TLR} : ANA \times C_R \rightarrow \boldsymbol{N}_0$ (II.116) $(\text{ANA},c) \in ANA \times C_R \mapsto \text{TLR}(c) = t^* - t_{RY}(c) = t_{PY} - t_{RY}(c) - 1$ [57], where TLR(c) [57]: the time lag of the reported ANA data (i.e. the delay in reporting data for publication to the IO) for country $c \in C_R$ that reported new ANA data, t* : the expected year of reported ANA data by the NSO ($t^* = t_{PY} - 1$, i.e. the past reference period), and [t_{RY} and t_{PY} : see: DEF. 32].

Source: own definition

Instead of providing results for the *timeliness* of data, a *timeliness assessment* might communicate the *time lag* of reported data by countries (cf. UNSD 2008). The *timeliness* (T) and the *time lag of reported ANA data* (TLR) are obviously related to each other. Since the most recent set of ANA data is expected to be produced for the past reference year (t_{RY} = t-1), there is no *time lag* (TLR(c) = 0) for a country $c \in C_R$ that reported t-1 data (i.e. T(c) = 1). Therefore an actual *time lag*, i.e. TLR > 0, occurs only for a country $c \in C_R$ that reported new data referring to the reference year before the previous year, i.e. for $t_{RY} \leq$ t-2. In case of ANA data referring to the reference year

[57] $\text{TLR}(c) = \text{TLR}_{ANA}(c) = \text{TLR}(\text{ANA},c)$. The symbol "ANA" is dropped from now on for convenience.

t-2 (i.e. T(c) = 2), the *time lag* is one year (i.e. TLR(c) = 1). Thus the value of the *time lag* in years is always one less than the value of the *timeliness* in years.

Example: The expected reference year of reported ANA (t*) is 2009, for new ANA data published in the YB in year (t_{PY}) 2010. Let the most recent reference year for country c $\in C_R$ in the YB be the reference year t_{RY} = 2007. Thus

TLR(ANA,c) $= t^* - t_{RY}(c)$, respectively, (cf. DEF. 33)

TLR(ANA,c) $= 2009 - 2007 = 2 = t_{PY} - t_{RY}(c) - 1 = 2010 - 2007 - 1 = 2$.

Thus the *time lag* of the ANA data of country c is two years.

The *time lag* can also be derived from the *timelines* as per DEF. 32 by deducting one (i.e. TLR(c) = T(c) – 1). For example, the data of a country with *timeliness* of three years has a *time lag* of two years, meaning for the timeliness T(c) = 3 (years), we get the *time lag* TLR(c) = 3 – 1 = 2 (years).

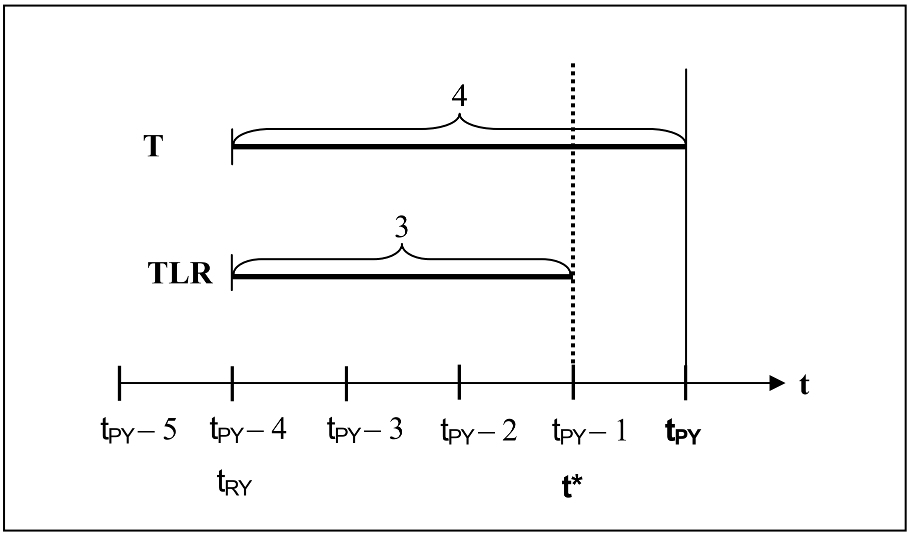


Figure 16 Timeliness (T) and Time Lag (TLR) of Annual National Accounts

Source: own illustration

The measure for the *time lag* is related to the time-related measure "*punctuality*". *Punctuality* can only be evaluated against a previously announced target publication date. This is done by examining whether or not the data were released prior to the targeted publication date or within the targeted timeframe. If the data are released after the target date, the length of time between the targeted release date and the actual release date can be evaluated (i.e. the delay) (cf. Eurostat 2003a and OECD 2003).

The measures for *timeliness* and the *time lag* of the ANA data at the time of their publication at the IO, i.e. at the *international* level, do not provide information regarding the *timeliness* of data production and publication at the *national* level. Nor do they provide information regarding the *time lag* of the data at time point t_{TR} of data *reporting* (transmission) to the IO. Thus these measures are less meaningful for assessing the actual NA compilation capacities. For example countries that report data comparably early are delayed by the IO's publication procedure until the annual YB is published; on the other hand countries that missed the publication deadline for the YB are deferred to the next year's YB publication.

However, due to the time point of the UN's YB publication late into the current year, the results do also provide valuable information for the identification of countries and country clusters with particularly poor NA data *production capacities* as regard to the *timeliness* of the *produced* data. This applies especially to countries with a *timeliness* of more than two years, i.e. a *time lag* of more than one year. (More meaningful for an assessment of the NA compilation capacity of a country would be the *timeliness* respectively the *time lag* at the time point of reporting of the national data to the IO, i.e. the time point when the data are received by the IO. However, most useful would be the *timeliness* at the national level, i.e. at time point of first time publication by the national institution, for example by the NSO or by the central bank.)

The *timeliness study* is further relevant for scholars interested in research of NA at the international level e.g. in international comparisons of countries. It provides information about the *timeliness* of the *published* ANA data by *regions* and *country groups*. Thus we can for example determine, until which publication year of a YB we have to defer a study of ANA data about the economic development within the first five years following the latest financial crisis. That does mean we can conclude the publication year t_{PY}, when we expect to have sufficient data available for specific regions like Northern Africa or groups like the OECD.

Measurement of Time Lag

Building upon our definition of *time lag* TLR(c) (cf. DEF. 33), the set of countries with a specific *time lag* of value $n \in \boldsymbol{N}$ in years is the preimage set

$$C_{n}^{TLR} := TLR^{-1}(n) = \{c \in C_{R} \mid TLR(c) \in n\}. \qquad (II.117)$$

Further we define the indicator functions regarding whether or not a specific country c is element of the preimage set C_n^{TLR} for a specific *time lag* of value n $\in$ $\boldsymbol{N}$ in years:

$$I_n^{TLR} : C_R \rightarrow \{0,1\} \text{ with } \qquad I_n^{TLR}(c) = \begin{cases} 1 \text{ if } c \in C_n^{TLR} \\ 0 \text{ else.} \end{cases} \tag{II.118}$$

$$\text{Thus, } C_n^{TLR} = I_n^{TLR\,-1}(1). \tag{II.119}$$

The preimage set C_n^T of countries with a specific *timeliness* of value n $\in$ $\boldsymbol{N}$ is equal to the preimage set C_n^{TLR} for TLR(c) = T(c) – 1. Thus for example the countries that are the members of the preimage set C_2^{TLR} of countries with a *time lag* of two years are identical to those countries that are the members of the preimage set C_3^T of countries with the *timeliness* of three years. Consequently, the preimage sets C_0^{TLR} to C_3^{TLR} in Figure 17 correspond exactly to the preimage sets C_1^T to C_4^T illustrated in Figure 15, p. 121.

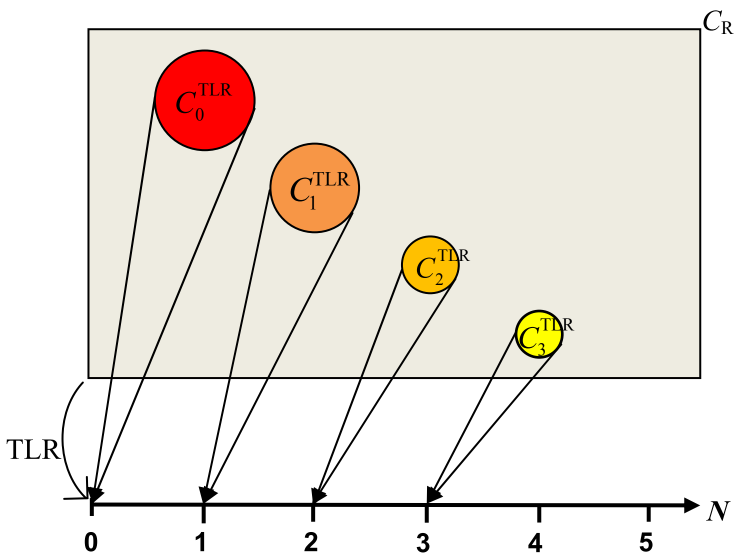


Figure 17 Preimage Sets of Countries by Time Lag (TLR) of value n $\in$ $\boldsymbol{N}$ in Years

Source: own illustration

Sample Application 11: Timeliness of Annual National Accounts

Country (Northern Africa Region) $c \in$ *Northern Africa*	Published Reference Period t_{RY}	Time Point of Publication t_{PY}	Timeliness T	Time Lag TLR
Algeria	2003	Previous YB	N/A	N/A
Egypt	2006	09-2009*	3 (years)	2 (years)
Libyan Arab Jamahiriya	2007	Previous YB	N/A	N/A
Morocco	2007	09-2009*	2 (years)	1 (year)
Tunisia	2007	09-2009*	2 (years)	1 (year)

* 2008 Yearbook as PDF publication.
N/A: not applicable

Algeria and also **Libyan Arab Jamahiriya** did not report new data for the 2008 YB publication (published in 2009). Therefore the ANA data published in the 2008 YB refers to "known" data, i.e. data that were already published in a previously produced YB. Consequently the *timeliness* of these ANA data at time point of their first publication at the international level cannot be assessed using the 2008 YB (since the results would merely inform how *up-to-date* the published YB is with regard to the total countries and corresponding data covered by the assessed YB).

Example 1: Egypt (EGY) reported data for reference year 2006 - meaning 2006 is the most recent reference year for which an ANA indicator is available within the dataset reported to the UN. Thus the *timelines* of the ANA data is *three years*.

Example 2: For Morocco (MAR) and Tunisia (TUN), the most recent reference year is 2007, thus the *timeliness* of the ANA data of these countries is *two years*.

The reference year 2008 is the expected reference year for which ANA should be produced and published for any country during the fiscal year 2009. Thus the *time lag* of the data of Egypt is two years. For Morocco and Tunisia it is one year.

Regarding the *formulas* for *timeliness* of ANA data, T(ANA,c) (c.f. DEF. 32), and *time lag* of reported ANA data, TLR(ANA,c) (c.f. DEF. 33), we drop the symbol "ANA" in the following for convenience.

(continued on next page)

Timeliness and Time Lag in Years

The formula for *timeliness* in years, calculated via the publication year of the YB and the most recent reference year of ANA data, i.e. $T(c) = t_{PY}(c) - t_{RY}(c)$ equates for our sample countries Egypt, Morocco, and Tunisia, to

$$T(EGY) = 2009 - 2006 = 3 \text{ (years)},$$
$$T(MAR) = 2009 - 2007 = 2 \text{ (years)},$$
$$T(TUN) = 2009 - 2007 = 2 \text{ (years)}.$$

For the *time lag* of reported ANA data, calculated via the expected reference year (i.e. the publication year of the YB less one year) and the reported reference year, i.e. by $TLR(c) = t^* - t_{RY}(c) = t_{PY} - t_{RY}(c) - 1$, this equates to

$$TLR(EGY) = 2008 - 2006 = 2009 - 2006 - 1 = 2 \text{ (years)},$$
$$TLR(MAR) = 2008 - 2007 = 2009 - 2007 - 1 = 1 \text{ (year)},$$
$$TLR(TUN) = 2008 - 2007 = 2009 - 2007 - 1 = 1 \text{ (year)},$$

This is illustrated by the following Figure 18 for the country Egypt (EGY).

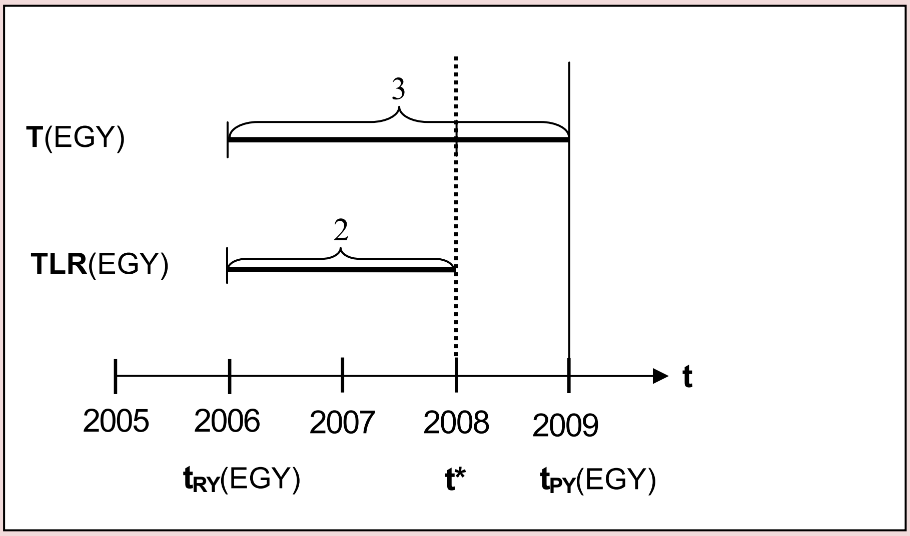


Figure 18 Timeliness (T) and Time Lag (TLR) of Annual National Accounts of Egypt

Source: own illustration

(continued on next page)

Set of Countries with a Specific Timeliness

Regarding the *set of countries* with a specific *timeliness* of value n $\in \boldsymbol{N}$ in years, we now look at the UN Member States in the region Northern Africa and the *timeliness* of two years. Morocco and Tunisia are in the set of countries with the *timeliness* value of two years. Referring to the formulas (cf. II.111) this equates to

$$C_2^{\mathrm{T}} = \mathrm{T}^{-1}(2) = \{\, \mathrm{c} \in \textit{Northern Africa}_{\mathrm{R}} \mid \mathrm{T(c)} = 2 \,\} = \{\, \text{Morocco, Tunisia}\}.$$

Further, of the region Northern Africa, Egypt is the single one member country of the set of countries with the *timelines* value three years. As formula, we have

$$C_3^{\mathrm{T}} = \mathrm{T}^{-1}(3) = \{\, \mathrm{c} \in \textit{Northern Africa}_{\mathrm{R}} \mid \mathrm{T(c)} = 3 \,\} = \{\, \text{Egypt} \,\}.$$

For the *set of countries* with a specific *time lag* of value n $\in \boldsymbol{N}$ in years we get the same sets of countries as for the *timeliness* with value n + 1 (i.e. T(c) = TLR(c) + 1). We get these sets of countries via the preimage set, too. Thus, as formula (cf. II.117) we have

$$C_1^{\mathrm{TLR}} = \mathrm{TLR}^{-1}(1) = \{\mathrm{c} \in \textit{Northern Africa}_{\mathrm{R}} \mid \mathrm{TLR(c)} \in 1\} = \{\text{Morocco, Tunisia}\},$$

$$C_2^{\mathrm{TLR}} = \mathrm{TLR}^{-1}(2) = \{\, \mathrm{c} \in \textit{Northern Africa}_{\mathrm{R}} \mid \mathrm{TLR(c)} \in 2\} = \{\, \text{Egypt} \,\}$$

The membership of a country within the *set of countries* of a specific *timeliness* or *time lag* value is indicated by the indicator function (cf. Formula II.112 and II.118, respectively). Thus we have

$\mathrm{I}_2^{\mathrm{T}}(\mathrm{EGY}) = 0;$ $\quad \mathrm{I}_1^{\mathrm{TLR}}(\mathrm{EGY}) = 0,$

$\mathrm{I}_2^{\mathrm{T}}(\mathrm{MAR}) = 1;$ $\quad \mathrm{I}_1^{\mathrm{TLR}}(\mathrm{MAR}) = 1,$

$\mathrm{I}_2^{\mathrm{T}}(\mathrm{TUN}) = 1;$ $\quad \mathrm{I}_1^{\mathrm{TLR}}(\mathrm{TUN}) = 1$, and correspondingly,

$\mathrm{I}_3^{\mathrm{T}}(\mathrm{EGY}) = 1;$ $\quad \mathrm{I}_2^{\mathrm{TLR}}(\mathrm{EGY}) = 1,$

$\mathrm{I}_3^{\mathrm{T}}(\mathrm{MAR}) = 0;$ $\quad \mathrm{I}_2^{\mathrm{TLR}}(\mathrm{MAR}) = 0,$

$\mathrm{I}_3^{\mathrm{T}}(\mathrm{TUN}) = 0;$ $\quad \mathrm{I}_2^{\mathrm{TLR}}(\mathrm{TUN}) = 0.$

(continued on next page)

Regional Analysis: Count of Countries with a Specific Timeliness

For the regional analysis, we calculate the numbers of countries in Northern Africa with a specific *timeliness* of value n $\in$ ***N*** in years. In this example we have two countries with *timeliness* of two years and one country with *timeliness* of one year.

There are no countries with any other *timeliness* value for Northern Africa. When calculating the *cardinality* of C_n^T i.e. the sum of $I_n^T(c)$ over all c $\in$ C_R we get

$$\left|Northern\_Africa_n^T\right| = \left|\sum_{c \in Northern\_Africa_R} I_n^T(c)\right|.$$

Count of Northern African Countries with Timeliness n in Years

n	Timeliness (T)
1	0
2	2
3	1
4	0

For the set C_R of countries that reported a new reference year of ANA data the *share* (h) of countries with a particular *timeliness* value n $\in$ ***N*** in years is given by

$$h_n := \frac{\left|T^{-1}(n)\right|}{\left|Northern\_Africa_R\right|}.$$

Share (h) of Northern African Countries that Reported New Data with Timeliness n in Years

n	h_n
1	0
2	2/3
3	1/3
4	0

Thus we have 0% countries where the timeliness is one year, 66% where the timeliness is two years, 33% where the timeliness is 3 years, and 0% where the timeliness is 4 years.

III. Data Quality Assessments at the International Level

This part covers the data quality assessments. Each chapter *analyzes* first the *deficiencies* of the *current* NA data quality (DQ) assessments and suggests *improvements*. We use our *improved* methodology for our subsequent *assessments*.

1 Timeliness Study

This chapter provides an assessment of the *timeliness* of Annual National Accounts (ANA) reported by countries to the UN and subsequently published in the NA Yearbook (YB). The currently employed *methodology* of the UN for assessing the *timeliness* of published NA data is *reviewed* and *improved* to provide better and more meaningful results. We employ the DQ measures *Timeliness of ANA* (T) and *Time Lag of Reported data* (TLR) for our timeliness assessment. We defined these *measures* in DEF. 32, p. 117, respectively DEF. 33, p. 122. In contrast to the UN, we improve particularly the clustering of assessed country groups and compute the full range of timeliness of ANA publication.

It has to be considered that our *timeliness assessment* reflects the "IOs" *timeliness* of data published at the *international* level. The "NSOs" *timeliness* of data available at the *national* level might be better. - For example a country might have failed to report data or to do so on time. - Yet, even if ANA data become available more *timely* at the *national* level (and thus could be utilized for example by government for policy making etc.), they are not available to the *international* community of users of data at the UN.

1.1 Introduction

Not all countries $c \in C_{YB}$ for which ANA data are published in the YB compute their ANA data for the past reference year. The most recent available reference year $t_{RY}(c)$ of ANA data for a country $c \in C_{YB}$ in the YB should be the reference year preceding the publication year t_{PY} of the present YB (i.e. $t_{RY}(c) = t_{PY}(c) - 1$; in the following we write: $t_{RY} = t\text{-}1$ [where $t := t_{PY}$)][58]). Many countries $c \in C_{YB}$ in the YB that currently

[58] $t_{PY} = t_{PY}(c)$ and $t_{RY} = t_{RY}(c)$, for the publication year t_{PY} and reference year t_{RY} of ANA data, we drop the symbol for the variable "c" indicating the country.

have data available (i.e. produce data)[59] have a *time lag* (TLR(c)) for ANA production. Thus, the ANA data for the past reference year (t_{RY} = t-1), which is the benchmark reference year, become available in the country and at the international level with possibly several years *time lag*. (For example, the ANA data for the current reference year will be produced in three years from now.)

Our *timeliness assessment* identifies how many countries $c \in C_R \subseteq C_{YB}$ reported new ANA data to the UN, moreover we identify the most current reference year t_{RY} of ANA for the countries $c \in C_R$ that reported a new reference year of ANA. In other words, we assess *timeliness* T(c), respectively *time lag* TLR(c), of ANA published via the YB of those countries $c \in C_R$ that reported a new set of ANA data.

We distinguish the following sets of countries:

> C_K: the set of countries with "known" ANA data, i.e. where the most recent available ANA data have been already published by a previous YB (the latest available reference year t_{RY} refers to a reference year that is at most five years back from the YB publication year (i.e. $t_{RY} \geq$ t-5 < t-1)).[60]
>
> C_R: the set of countries that reported a new reference year t_{RY} of ANA data.
>
> C_{YB}: the set of countries for which ANA data are published in the YB (the latest available reference year t_{RY} at most five years back from the YB publication year (i.e. $t_{RY} \geq$ t-5)), $C_{YB} = C_R \cup C_K$.

The current UN assessment of *timeliness* is marked by several *deficiencies*. The *methodology* employed by the UN does not take into account whether or not the NA data published for a country $c \in C_{YB}$ included in the YB were actually published for

[59] NA data not being published by the UN are considered not being produced. It is internationally agreed that "basing the assessment on data actually available at the international level was preferred as more objective" (cf. United Nations 1999, 3).

[60] We consider the countries $c \in C / C_{YB}$ (i.e. the set $C / (C_R \cup C_K)$) as currently not having ANA data available, i.e. as currently not producing ANA data. The set C / C_{YB} of countries currently not having ANA data available comprises those countries for which official ANA data were never reported and those countries for which the most current reference year t_{RY} of ANA data is at least six years back from the publication year t_{PY} of the present YB (i.e. $t_{RY} \leq$ t-6).

the first time by the present YB. This means, the UN methodology does not distinguish between the set $C_K = C_{YB} / C_R$ of countries with "known" data, included in the YB, for which no new reference year t_{RY} of ANA data had been reported, and the set C_R of countries that provided new ANA data. For some of the countries in the YB, the present data have been already published by a previous YB (possibly for several years). With other words, these ANA data are "*known*" and such country is member of the subset $C_K \subseteq C_{YB}$ of countries with "known" data.

The UN methodology also needs improvement of the subdivision of assessed *regional aggregates*. They are not sufficiently detailed for a sensitive analysis of differences between individual *regions* or of country *groups* (like G20 countries and other advanced economies). Furthermore, the assessed *range of timeliness* (i.e. $n \in \boldsymbol{N}$ in year(s) with n = 1, 2 and the cumulative it) is insufficient. The UN study (UNSD 2008) only displays results for those cases, where the most current available reference year t_{RY} is not more than two years back from the publication year t_{PY} of the YB (i.e. $t_{RY} \leq$ t-2). Thus only a subset of the occurring years $n \in \boldsymbol{N}$ of *timeliness* are computed respectively presented. With other words, the UN *timeliness assessment* only considers the subset $C_{YB2} \subseteq C_{YB}$ of countries in the YB, which we define as follows:

> Let C_{YB2} be the subset ($C_{YB2} \subseteq C_{YB}$) of countries published in the YB, for which the most current available reference year t_{RY} of ANA data is at most two years back from the publication year t_{PY} of the present YB (i.e. $t_{RY} \leq$ t-2).

Further we distinguish the two basis country sets C_{All}, referring to our grand total of considered countries, and $C_{UN} \subset C_{All}$, which are defined as follows:

> C_{All}: the set of countries in the MDG country list (224 countries[61]).[62.]
>
> C_{UN}: the set of UN Member States (192 countries[63])

---

[61] There are now 225 "MDG list" countries in the set $C_{All} := C$, but at the time of the analyses it was 224.

[62] The set of countries considered by the MDG country list (UNSD 2003) is based on the Standard Country or Area Codes for Statistical Use (Series M, No.49/Rev 4) (UNSD 2013a). It is available as download online at: http://mdgs.un.org/unsd/mdg/Resources/Static/Data/Regional%20groupings.doc and http://mdgs.un.org/unsd/mdg/Resources/Static/Data/MDGRegionCodes_200611.xls, subject to updates per (UNSD 2013b) available online at: http://unstats.un.org/unsd/methods/m49/m49regin.htm (status: 11-02-2013), last accessed 11-06-2013.

In part II, chapter 3.3, p. 117ff., we *improved* the definition of the *timeliness measures*. In the following subchapter, we *improve* the methodology of the *timeliness* assessment in terms of further (all occurring) years $n \in \boldsymbol{N}$ for these *timeliness measures*, and in terms of the *granularity* of assessed country clusters. Thus our present assessment *updates* the previous UN assessment and provides better information regarding the *timeliness* of ANA publication for different *country clusters*.

The *improved* assessment *methodology* takes into account solely the set C_R of those countries that provided new ANA data. That does mean, not considered for the evaluation of the *timeliness* of NA data publication at the international level, are the ANA data of those countries $c \in C_K \subseteq C_{YB}$ that are merely republished by the YB. Therefore, the results of our assessment do not interfere with "known" (i.e. "old" respectively "preexisting" data), as is the case in current UN *timeliness assessments*.

The *improved granularity* of the analyzed *country groups* covers developed-, transition- and developing economies, with further *regional* disaggregation. Moreover, regarding these *groups* by economic classification and regions, we consider the set C_{UN} of UN Member States and also the set $C_{All} \supset C_{UN}$ of all countries (and areas and territories) including the non-UN members as the basis set. Further, we analyze *groups* defined by membership in organizations like OECD, special country *groups* like LDCs, and the UN Regional Commissions (i.e. the UN's regional organizations).

The YB *volume* used to assess the *timeliness* of NA publication is the 2008 Yearbook (United Nations 2009b, c, d, and e). This volume was produced in September 2009 (i.e. in publication year $t_{PY} = 2009$) with data for the last 12 reference years, covering 1997 to 2008. The most current reference year requested for ANA reporting (i.e. the benchmark) given the publication year $t_{PY} = 2009$ is the reference year $t_{RY} = 2008$. The optimum *timeliness* for a country's ANA is T(c) = 1, i.e. the reporting of 2008 reference year data for the 2008 YB, published in 2009. A country that reports data for the reference year t-1 in the publication year t has therefore no *lag*, i.e. TLR(c) = 0.

The NA statistics in the YB are official country data. They are obtained by the United Nations Statistics Division (UNSD) directly from a national statistical agency, i.e. NSO, CB, etc., or received through an international or regional organization, i.e. by

[63] There are now 193 UN members, but at the time of the analyses it was 192.

“the Organisation for Economic Co-operation and Development (OECD), the United Nations Economic Commission for Europe (ECE) and the Caribbean Community (CARICOM)”, “on behalf of their constituents” (United Nations 2009b, vi and vii). The ANA statistics published in the YB are subsequently released via the data dissemination platform “UNdata” (UNSD 2013c).

The world map (Figure 19, p. 135) indicates the results for the *timeliness* of ANA publication T(c) in years n $\in \boldsymbol{N}$ (1 to 4). The result 1, …, 4 years refers to the most current reference year t_{RY} for the country c $\in C_R$ that reported a new reference year of ANA, in relation to the year t_{PY} of the ANA data publication (t_{PY} = 2009).[64] The country coloring follows the color code “heat”, where the best timeliness (T(c) = 1 year) is indicated by “brick red”, i.e. the “top-heat”. T(c) = 4 (“yellow”), does mean that the *timeliness* of the most recent reference year, of the new ANA data for a country c $\in C_R$ is four years. With other words the country’s most current ANA data were reported for the reference year (t_{RY}) dating four years back from their first time availability in the analyzed YB.

Those countries in the world map marked as *timeliness* ‘n/a (“known” data, ≤ 5 years)’ refer to the set $C_K = C_{YB} / C_R$ of countries included in the YB with “known” data (published in a previously released YB). [These are mainly countries that replied to the previous YB and further two that replied to the two year’s earlier YB (the latter are Greenland and Gabon, both onetime reporters of ANA data).]

Countries left blank in the world map are those considered as having no data available. This refers to the set C / C_{YB} of those countries not included in the set $C_{YB} = C_R \cup C_K$. With other words, exclude are those countries indeed not included in the YB and those countries considered as currently having no data available since the most current reference year t_{RY} of ANA data is at least six years back from the publication year t_{PY} of the YB (i.e. $t_{RY} \leq$ t-6). [Further, one country c $\in C_R$ (Timor-Leste) which reported new data, indicating a timeliness of six years (T(c) = 6) of ANA publication, is left blank for technical purpose.[65]]

[64] Countries, areas or territories included in the ANA data of another (“main”) country are displayed (in the world map, only) by the same color as the “main” country, although no individual data are reported.

[65] We conclude the data were produced at an earlier point in time by external assistance and currently no NA data are produced by the national authorities (cf. chapter 1.3.1, p. 186).

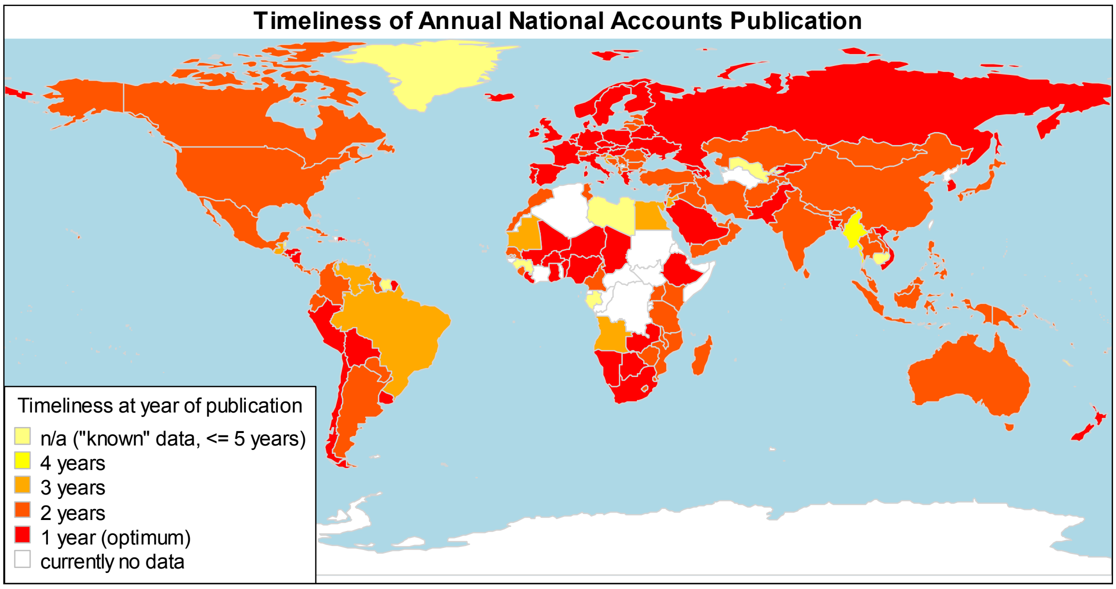


Figure 19 World Map of Timeliness of Annual National Accounts Publication

Source: own illustration. (Based on 2008 YB, published in 2009.)

World maps are produced using R

The world map shows that overall *timeliness* T(c) of ANA is at a level where nearly all countries report annual data for the previous year (t-1) or at least the year t-2, thus within a maximum *lag* of one year to the optimal reference year.

As explained in part II, chapter 3.3, we also describe the *timeliness* by referring to the reported reference year t_{RY} in relation to the publication year t_{PY}, i.e. we say that the reference year t_{RY} published in t refers to t-n (again, with t as abbreviation for t_{PY}). We prefer the notation t-n, i.e. "t-1" etc. over using the description "previous year data" etc. Either of these expressions, put the latest published reference year t_{RY} in relation to the publication year t_{PY} and emphasizes the NA compilation capacities of countries regarding the *timeliness* T(c) of ANA data publication respectively production. That is, a country with data being published for the year t-1, has the capacity to produce ANA for the previous reference year. On the other hand, a country which reported t-4 data as most recent available reference year t_{RY} of ANA has a weak NA production capacity. In the latter case the published data have a *time lag* (TLR(c) = 3) of three years to the expected reference year (that is the previous reference year, i.e. $t_{RY} = t-1$).

We assume that, given no changes, a country $c \in C_R$ that reported new data for the current YB will do so in the coming year, too, and with the *same timeliness* of the

ANA data reported for publishing. Regarding a country $c \in C_{YB} / C_R$ that did not report new data, the future NA production and reporting ca not be concluded with certainty. For example the non-reporting by a country might be due to changed NA production capacities resulting from political instability in a country.

Our study reveals a much better *timeliness* of NA data than suggested by the previous UN assessment (cf. UNDS 2008). We find that the NA data *reported* by UN Member States does not have more than a one year *lag* in over 90% of the cases. In contrast, the UN indicates that less than 2/3 of the *UN Member States* have such *timeliness*. This is primarily, because the UN assessment does not consider that a significant number of the UN Member States did not report a new set of ANA data or do not even compile NA.

Table 15 Comparison of Timeliness of UN Member States by UN Methodology and by Our Improved Methodology

Percentage of countries for which the latest reference year of ANA data does not have more than one year time lag (i.e. with reference year $t_{RY} \geq$ t-2)	
UN Methodology[/a]	**Improved Methodology**[/b]
64%	92%
/a Considering set C_{YB} of countries in the YB, percentage of total number of UN Member States C_{UN}. /b Considering set C_R of countries that reported new data, percentage also based on set C_R.	

Source: UNSD (2008, 7) and own compilation. (Based on 2006 YB, published in 2007 (UNSD) respectively 2008 YB, published in 2009 (own).)

The UNSD (2008) assessment uses a *different methodology* than our timeliness assessment. The following subchapters explain the *deficiencies* of the *UN methodology* and describe our *improved methodology*. The methodological analysis and perspectives for *further improvements* is followed by the *results* of our empirical *timeliness assessment*.

1.2 Improvement of the Current UN Assessment of Timeliness

The aim of our *improvements* is to provide

i. better DQ measures for the *timeliness* of ANA data publication at the international level,
ii. full range, i.e. each occurring year $n \in \boldsymbol{N}$, of *timeliness* T(c), and
iii. finer subdivision of the assessed *regions* and *country groups*.

This facilitates better analysis for researchers and provides detailed information that could be used to identify country clusters that could benefit from assistance by international and regional organizations to improve the NA capacities of the national institutions (e.g. of NSOs and CBs).

1.2.1 Methodology

This section provides the details for the *deficiencies* and *improvements* of the current *UN assessment methodology*.

The *UN's timeliness assessment*, as well as our *improved* approach, both calculate the *timeliness* based on the data value that refers to the most current reference year within the reported data set. Thus *both approaches* are subject to the limitation that some of the detailed ANA data of the set of NA indicators requested by the NAQ, or even the majority of ANA data, might have an additional *time lag*. Say for example a country $c \in C_R \subseteq C_{YB}$ reported new ANA data with the *timeliness* T(c) = 1, i.e. data referring to the previous reference year. However, this country might have data with such *timeliness* available only for GDP by expenditure and the corresponding components. Yet, the data for GDP by industry and the corresponding components might become available with additional *time lag*, e.g. with a *timeliness* of 2 years (i.e. T(c) = 2).

The major difference is that the *UN methodology* actually does not assess the *timeliness* of published ANA data. It merely identifies the most current reference year t_{RY} of data in the YB, even if these data are "known" data, i.e. reported by the country $c \in C_K$ with "known" data during a previous YB reporting round. Thus, such an assessment only indicates how *up-to-date* the ANA data are for the set C_{YB} of countries included in the YB.

Deficiency of the Current UN Assessment Methodology

We start with a definition and formula ($T_{YB}(c)$) for the calculation of the "timeliness of reporting" (UNSD 2008, 7) *measured* by the UN, because these are not included in the UN study. This allows a comparison between our definition of *timeliness* with the UN's definition. The UN methodology, considers a different set of countries in the YB (also covering a subset of countries $c \in C_K$ with "known" data). Therefore, the kind of *timeliness* assessed by the UN, alters our definition for the actual *timeliness* (cf. DEF. 32) as follows:

DEF. 34: Timeliness of Annual National Accounts by UN Methodology
The length of the time period (T_{YB}) measured in years between the end time point of the latest available reference year (t_{RY}) of ANA data for any country included in the Yearbook (YB), and the time point of publication (publication year t_{PY}) of the present YB. Formally expressed: C_{YB2}: subset ($C_{YB2} \subseteq C_{YB}$) of countries published in the YB, for which the most current available reference year t_{RY} of ANA data is at most two years back from the publication year t_{PY} of the present YB (i.e. $t_{RY} \leq$ t-2). $T_{YB} : ANA \times C_{YB2} \rightarrow \boldsymbol{N}$ (1.1) $(ANA,c) \in ANA \times C_{YB2} \mapsto T_{YB}(ANA,c) = t_{PY}(ANA,c) - t_{RY}(ANA,c)$[66], where $T_{YB}(ANA,c)$[66]: the timeliness of ANA data in years $n \in \boldsymbol{N}$ referring to the most recent reference year within the present set of ANA data of country $c \in C_{YB2}$ included in the present YB, $t_{PY}(ANA,c)$[66]: the year when the present YB volume is published, and $t_{RY}(ANA,c)$[66]: the most recent reference year within the set of published ANA data for country $c \in C_{YB2}$, with $t_{PY}(c) > t_{RY}(c)$.

Source: own definition based on UNSD (2008)

[66] $T_{YB}(c) = T_{YB}(ANA,c)$. The symbol for data object "ANA" is dropped from now on for convenience. For $t_{PY}(c) = t_{PY}(ANA,c)$ and $t_{RY}(c) = t_{RY}(ANA,c)$, ditto.

Then the set of countries with a specific *timeliness* value $n \in \boldsymbol{N}$ in years is the preimage set

$$C^{T}_{YB,n} := T^{-1}_{YB}(n) = \{ c \in C_{YB} \mid T_{YB}(c) = n \}. \tag{1.2}$$

Further we define the indicator function regarding whether or not a specific country c is element of the preimage set $C^{T}_{YB,n}$ for a specific *timeliness* given by value $n \in \boldsymbol{N}$ in years:

$$I^{T}_{YB,n} : C_R \rightarrow \{0,1\} \text{ with } \qquad I^{T}_{YB,n}(c) = \begin{cases} 1 \text{ if } c \in C^{T}_{YB,n} \\ 0 \text{ else.} \end{cases} \tag{1.3}$$

$$\text{Thus, } C^{T}_{YB,n} = {I^{T}_{YB,n}}^{-1}(1). \tag{1.4}$$

The cardinality of $C^{T}_{YB,n}$ that is the sum of $I^{T}_{YB,n}(c)$ over all $c \in C_{YB}$, is

$$\left| C^{T}_{YB,n} \right| = \left| \sum_{c \in C_{YB}} I^{T}_{YB,n}(c) \right|. \tag{1.5}$$

Using the *UN assessment methodology*, explained in this subsection, the *UN results* refer to the "Latest year of available data reported by countries as percentage of total number of countries" (cf. UNSD 2008, 7). Assessed is the availability of ANA data in the YB where the most recent reference year (t_{RY}) refers to t-1 or t-2 (and the sum of these two groups, i.e. where $t_{RY} \geq$ t-2). That does mean considered is only the subset $C_{YB2} \subseteq C_{YB}$ of countries for which the most current available reference year t_{RY} of ANA data is at most two years back from the publication year t_{PY} of the present YB. Regarding the subset $C_{YB2} = C_{R2} \cup C_{K2}$ we define the following subsets of countries:

C_{K2}: the subset ($C_{K2} \subseteq C_K$) of countries with "known" ANA, data that have been already published by a previous YB, for which the most current available reference year t_{RY} of ANA data refers to the reference year two years before the publication year t_{PY} of the present YB (i.e. t_{RY} = t-2).

C_{R2}: the subset ($C_{R2} \subseteq C_R$) of the countries that reported a new reference year t_{RY} of ANA data, for which the most current available reference year t_{RY} of ANA data is at most two years back from the publication year t_{PY} of the present YB (i.e. $t_{RY} \leq$ t-2).

The methodology to derive these results has the following *deficiencies*:

1) Not the "timeliness of reporting" (UNSD 2008, 7) is assessed for the countries in the YB (i.e. for the set $c \in C_{YB}$ of countries in the YB). Instead, the latest available reference year (t_{RY}) of a country in the YB is identified. Yet, this is only equivalent to the *timeliness of data publication* T(c) for the subset of the countries $c \in C_R$ ($C_R \subseteq C_{YB}$) that reported new data during the current YB reporting round.

2) The results are only *produced* for the *reference years* t_{RY} = t-1 and t_{RY} = t-2 (and the cumulative of the two).

3) Considered is only the subset $C_{YB2} = C_{R2} \cup C_{K2}$ of countries for which the most current available reference year of ANA data t_{RY} is at most two years back from the publication year t_{PY} of the YB. This does mean, the UN methodology includes the set C_{K2} comprising all those countries $c \in C_K$ with "known" data, that did last report new ANA data with a *timeliness* of one year for the *previously* published YB. Further, those countries $c \in C_R$ that reported new ANA data, yet with a timeliness of more than two years (i.e. t_{RY} > t-2) are not included in the UN assessment. The latter could be simply changed to cover all countries we consider currently having data available, i.e. for which the most recent available reference year of ANA is at most five years back for the present YB publication year (i.e. $t_{RY} \leq$ t-5). Therefore the considered set of countries in DEF. 34 (C_{YB2}) merely needs to be changed to the set C_{YB} of countries in the YB (for which, as per definition, the latest available reference year t_{RY} is at most five years back from the YB publication year). That is

$$T_{YB} : ANA \times C_{YB} \longrightarrow \boldsymbol{N} \qquad (1.6)$$

$$(\mathrm{ANA},c) \in ANA \times C_{YB} \mapsto T_{YB}(\mathrm{ANA},c) = t_{PY}(\mathrm{ANA},c) - t_{RY}(\mathrm{ANA},c)^{66},$$

with $t_{PY}(c) > t_{RY}(c)$.

4) The further a reference year t_{RY} is past the publication year t_{PY} (e.g. $t_{RY} = t-4$) for a country $c \in C_{YB}$ in the YB, the less meaningful is this information. This is because the set C_{YB} of countries in the YB includes the set C_K with "known" data, which have been published by a previous YB. For a country $c \in C_{YB}$, a reference year in the distance past (e.g. 4 years back, i.e. $t_{RY} = t-4$) could result either from new reported data of a country $c \in C_R$ with a long *time lag* for the data production. Or the back reference year might result from the republication of "known" data. The data are possibly republished since several years for a country $c \in C_K$ with "known" data, which might currently even not compile NA data. Further, at the time point of the year t_{PY} of first publication of the ANA data, the country $c \in C_{YB} / C_R$ might had have a small or no time lag TLR(c) of reporting. The problem of the UN timeliness assessment is that the *UN results mingle* the *reasons* for the lack of up-to-date data.

5) Only the UN Member States ($C_{UN} \subset C_{All}$) are reflected in the *UN assessment*. Yet, the UN collects data from all countries (i.e. referring to countries, areas, and territories) that produce official NA data (cf. United Nations 2009b, v). Precisely, the set C_{All} of countries in the MDG country list[67] is considered by the UN for the ANA data collection.

   Thus the countries $c \in C_{YB} / C_{UN}$, that are included in the YB yet not being a UN Member State, are not reflected by the existing *UN assessment*.

6) The UN results do not provide the *absolute numbers* of countries (i.e. the cardinality $\left|C^{T}_{YB,n}\right|$) with a particular timeliness of year $n \in \boldsymbol{N}$, but just the *percentages* regarding the set $C^{T}_{YB,n}$. Moreover, these percentages relate to the total number of countries ($|C|$) comprising the corresponding region (respectively country group) [e.g. the UN Member States or a region like Northern Africa[68]]).

[67] The countries included in the MDG (Millennium Development Goals) country list, cf. UNSD (2003).

[68] For the considered sets of countries (C) see the regional aggregates in chapter 1.2.3, p. 159.

The share $h_{YB,n}$ for any subset $C^{T}_{YB,n} \subseteq C$ of countries with a particular *timeliness* value $n \in \boldsymbol{N}$ in years is calculated by

$$h_{YB,n} = \frac{\left|C^{T}_{YB,n}\right|}{|C|} = \frac{\left|T^{-1}_{YB}(n)\right|}{|C|} \qquad (0 \leq h_{YB,n} \leq 1). \qquad (1.7)$$

The cardinality of countries is important, as it provides an absolute measure. Opposite, the derived shares depend on the defined corresponding total quantity, i.e. the basis set. This basis set can for example refer to the total number of the set C of countries within a region; or the total number of the set of those countries that have at one point reported data and are still included in the YB; or the total number of the set C_R of those countries that reported data during the reporting period for the present YB [where $C_R \subseteq C_{YB} \subseteq C$]).

For regions where not all countries $c \in C$ comprising a region, reported new data (i.e. $C_R \subset C$), the results suggests a *too small share* (h) of countries that are able to report timely data. This is because the shares are not calculated based on the set C_R of countries that actually report new data, but on the total number of the set C of countries assembling a region. Thus, the sum of all occurring years $n \in \boldsymbol{N}$ of timelines $\sum_{n=1}^{5} h_{YB,n} < 1$ remains below one. With other words, the total of the individual percentages regarding the sets $C^{T}_{YB,n}$ of countries with a particular timeliness, do not add up to 100% of the corresponding set C of countries of the assessed region. Thus, the *share* or *percentage* based on a regions total (i.e. $|C|$) leads to misinterpretation of the presented results. (Moreover, of the set C_R of countries that report new data, all countries $c \in C_R$ actually might have reported new data with the optimum timeliness, i.e. for the previous reference year (t-1).

The presentation of *shares* or *percentages* might particularly lead to misinterpretations if, for the *timeliness measure* $T_{YB}(c) = n$, not all occurring *timeliness* values in years $n \in \boldsymbol{N}$ are presented by the results of a study.

Specifically in case of the UN assessment, which considers solely the timeliness values for years n = 1, 2 we get $\sum_{n=1}^{2} h_{YB,n} < \sum_{n=1}^{5} h_{YB,n}$ (in case of occurring values n = 1, …, 5 of years of timeliness).

Furthermore, the countries $c \in C / C^{T}_{YB,n}$ that are not captured by the covered years of value n of *timeliness* $T_{YB}(c)$ might either have a *different timeliness* (n > 2) or do not report NA data (i.e. refer to members of the set C / C_{YB} considered not producing data). No information at all is provided by the *UN study* (UNSD 2008) regarding these (excluded) countries that are members of the set $C / (C^{T}_{YB,1} \cup C^{T}_{YB,2})$.

Table 16 Share of Countries with Timeliness of one Year given Different Basic Sets

Share (h) corresponding to a set of 64 UN countries with ANA data for the reference year t_{RY} = t-1 (i.e. C^{T}_{n} for n =1; C^{T}_{1} = 64) given different basic sets (C, C_{YB}, and C_R) of countries for the considered region of UN Member States.

Basic set*	C	C_{YB}**	C_R***
Total countries	192	178	135
Share (h) of countries with t_{RY} = t-1 data	0,33 ($0 \le h < 1$)	0,36 ($0 \le h < 1$)	0,47 ($0 \le h \le 1$)

Note: The share of countries with ANA data for reference year t_{RY} = t-1 is below 1 in case of set C_{UN} of all countries of the considered region and also in case of the set C_{YB} of all countries included in the YB. This is because not all UN countries $c \in C_{UN} \supseteq C_{YB}$ produce ANA data, nor have all countries $c \in C_{YB}$ that are included in the YB reported a new set of ANA data.

* Basic set for the assessed region of UN Member States (i.e. C = *UN Member States* =: C_{UN}).

** Referring to those UN Member States included in the YB, i.e. the set $C_{UN} \cap C_{YB}$.

*** Referring to those UN Member States that reported new ANA data, i.e. the set $C_{UN} \cap C_R$.

Source: own compilation

Beside these specific *deficiencies* of the *UN assessment methodology*, the following *limitations* apply in general to *assessments* based on internationally available country data, collected by IOs from national agencies:

1) The availability of data at the IO is subject to reporting of the data by the national agency.

2) The publication of the most recently reported data of a country in the IOs' publications [in case of the UN the YB] is subject to the cutoff date for inclusion of new data into the upcoming publication of the IO.[69]

3) The most recent reference year (in case of annual data) included in the set of data reported by a country [$c \in C_R$ that reported a new reference year t_{RY} of ANA data] is subject to the transmission time point, when the country decides to report their data. In best case countries report the new set of [ANA] data immediately after computation of a new reference year, in worst case a new set of data might be produced by the country just after having reported a data set to the IO.

4) Less than the data set requested by the IO [in case of the UN, the set *ANA* of ANA indicators given by the NAQ[70]] might be reported by countries due to underreporting. That means that only a subset of the requested data is reported by the country, though the total set of data are produced by the country.

   It is particularly difficult to track and re-request missing data, without any indication that these data are indeed produced by the country. For example, in case an entire subset of requested data [like an entire NAQ tables] is missing, respectively was not reported, or if a country did not report revisions of previously reported data [which in case of ANA are typically computed for several back periods]. Such failures to report existing data might lead to wrong ratings of countries in data availability assessments or in studies that assess the reliability of first released data.

[69] Data being reported too late are published with the following publication (i.e. in case of the UN the one year later YB), or potentially get replaced by more recent data being reported by the country prior to the following publication. Thus the (filtered) timeliness of data publication might vary between countries, although the data are produced with the same or similar timeliness by the NSOs.

[70] The 1993 National Accounts Questionnaire (1993 NAQ) includes more than 1300 variables.

5) Data actually reported by a country to the IO, might not have been published by the IO (all data, or a partial data set).[71]

Improved Methodology

This subsection describes in short our suggested improved *methodology*. The following sections provide further details concerning the *granularity* of the presented assessment results. Granularity relates to the computed values $n \in \boldsymbol{N}$ for the *years* of *timeliness* and the composition of *regional aggregates*, i.e. the country groups for which the *timeliness* is assessed of the each country $c \in C$ that is member of the considered country group.

1) We describe *timeliness*, T(c) = n, by the most current reference year t_{RY} *relative to the publication year* t_{PY}, i.e. as $t_{RY} = t_{PY} - n$ respectively $t_{RY} = t\text{-}n$. Thus for a country $c \in C_R$ with the timeliness T(c) = 4, we say that data for the reporting year t_{RY} = t-4 are published, or simply that t-4 data are published. (This is equivalent to a *time lag* of the data of 3 years, i.e. TLR(c) = 3.)

2) Only data reported during the current YB reporting round are considered, i.e. only those countries $c \in C_R$ that *provided new NA data* for the present YB. Thus previously published data, i.e. those countries ($c \in C_K = C_{YB} / C_R$) being republished with "known" data, which would wrongly indicate larger *time lags* are excluded. This *approach* ensures that the real *time lag* TLR(c) respectively the real *timeliness* T(c) of the ANA data published by the UN is captured.

3) All occurring values $n \in \boldsymbol{N}$ of timeliness T(c) = n are computed (see the following section).

4) The results are displayed in *absolute numbers* as well as in *percentages* regarding the set C_n^T (cf. Formula II.111) of countries with a specific timeliness T(c) = n (i.e. with an individual value $n \in \boldsymbol{N}$ of timeliness in years, or as the cumulative of several such years in addition to the individual years).

[71] For example due to reporting an incomplete data set or errors within the reported data set.

5) These *percentages* are computed for the set C_n^T of those countries with the timeliness value n $\in$ ***N*** of year(s), relative to the set C_R of countries *that reported new data*, i.e. calculated via the share

$$h_n = \frac{|C_n^T|}{|C_R|} \times 100 \text{ for any subset } C_n^T \subseteq C_R \ (0 \le h_n \le 1) \text{ (cf. Formula II.115).}$$

The set $C_R \subseteq C_{YB} \subseteq C$ of those countries that have reported data (for a specific region or country group, e.g. the *UN Member States*) for the current YB is the appropriate basis for calculating the percentage or share. (In contrast, the share based on the set *C* of the total number of countries comprising a region is not appropriate.)

The subset C_R is better than the total set of countries *C* (e.g. the total of the *UN Member States*), because

- not all countries c $\in C_{YB}$ in the YB (which we consider countries currently producing ANA data) produce ANA data annually,
- others might have missed a reporting round, and again
- the countries c $\in C / C_{YB}$ do not compile any NA (i.e. never produced or considered as currently not producing ANA data).

6) Our assessment includes *results for non-UN countries*, too.[72] Thus, as the first assessment, we consider only the member countries of the UN that actually reported new ANA data, i.e. the set $C_{R(UN)} = C_{UN} \cap C_R$ (with subdivision by each assessed region). Then we assess the set $C_{R(All)} = C_{All} \cap C_R$ of all countries that actually reported new ANA data (i.e. $C_{R(All)} \supset C_{R(UN)}$). We assess once more all regions (with subdivisions), since the additional countries c $\in C_{All} / C_{UN}$ that are not UN members are distributed over different regions. [For the considered *regional aggregates* see section 1.2.3 of this chapter. For the *results* of the *assessment* for the country set C_{UN} considering only UN Member States and the set C_{All} considering all countries including the non-UN member countries, see Table 25, p. 177 respectively Table 26, p. 180.]

---

[72] Since “the Statistics Division of the United Nations should publish regularly the most recent available data on national accounts for as many countries and areas as possible” (United Nations 2009b, v).

As an *outcome* of our changed approach - i.e. particularly the changed calculation of the share h_n of the set C_n^T of countries that report data with a specific *timeliness* in years n $\in$ $\boldsymbol{N}$, and the exclusion of countries c $\in$ C_{YB} / C_R that did not report data for the current YB - we expect significantly different results to the UNSD (2008) assessment. For our *timeliness assessment* see the chapter 1.3, p. 175 ff. A *comparison* with the UNSD (2008) study, and a *re-assessment* employing the UN methodology based on the data set we used four our present study, can be found in chapter 1.3.4, p. 184 ff.

1.2.2 **Range of Timeliness**

The deficiencies of the measurement of *timeliness* in the UN assessment methodology were identified in the previous section. This section refers to the *granularity* of computed *values of timeliness* n $\in$ $\boldsymbol{N}$ in years.

Deficiency of the Currently Composed Range of Timeliness

The *results* presentation in the *UN assessment* regarding the computed values of timeliness in years is marked by the following *deficiency*:

1) Only *two* years of *timeliness values* n $\in$ $\boldsymbol{N}$ in years, “t-1” and “t-2”, are considered, i.e. countries with ANA data for the previous year (optimal, *no lag*) and those countries with *one year lag*. Yet, not computed by the UN assessment are the *results* for these countries with a different *timeliness value*, i.e. *longer time lags* (TLR(c) > 1 the equivalent of *timeliness* T(c) > 2). Thus, only a *subset* of the occurring years n $\in$ $\boldsymbol{N}$ of *timeliness* are assessed.

2) Information about the *range* of n $\in$ $\boldsymbol{N}$, i.e. about the occurring values of years n $\in$ $\boldsymbol{N}$ of timeliness (other than 1 and 2 years), and the *number* of countries being affected is not provided.

3) The additionally displayed group of countries, respectively the additional column for the results regarding the percentage of countries “reporting with lag of not more than one year” (cf. UNSD 2008, 7) is merely the sum of the columns for years “t-1” and “t-2”. Thus this does not provide additional insight.

Improved Range of Timeliness

Our improved *granularity* of the *results* regarding the years of timeliness is as follows:

1) *All occurring values* n $\in$ $\boldsymbol{N}$ of timeliness in years, i.e. the complete range of occurring *values*, of those countries c $\in$ C_{R} that reported data for the new YB is calculated.

2) The *percentages* are computed based on the number of those countries c $\in$ C_{R} that reported new data for the considered region (i.e. based on the cardinality of the set C_{R}). Therefore, the cumulative *percentages* of all individual sets of countries $C_{\mathrm{n}}^{\mathrm{T}}$ with a specific timeliness in year(s) n $\in$ $\boldsymbol{N}$ adds up to 100% for the assessed region. That does mean, concerning the percentage of the share $\mathrm{h_n} = \frac{\left|C_{\mathrm{n}}^{\mathrm{T}}\right|}{\left|C_{\mathrm{R}}\right|}$ (cf. Formula II.115), we know that $\sum_{\mathrm{n}=1}^{5} \mathrm{h_n} = 1$ for n = 1, …, 5. Thus, in case the results table of a given timeliness study does not display[73] all occurring values of timeliness, the reader can conclude that the corresponding *percentage* of countries that is missing falls outside the displayed values of years. With other words, the missing share $\mathrm{h_n}$ (for a single value n $\in$ $\boldsymbol{N}$ in year(s) or the cumulative of multiple values n $\in$ $\boldsymbol{N}$ in years) cannot result from the set $C_{\mathrm{YB}} / C_{\mathrm{R}}$ of countries that did not report data or from the set C / C_{YB} of countries that do not produce NA.

For an *example* of the improved granularity of computed years of *timeliness* in the *results tables*, presenting the outcome of our assessment, see Figure 20, p. 156 at the end of the following section.

1.2.3 Clustering of Country Groups

This section refers to the *clustering* of composed *geographical* regions, sub-regions, and other country groups. Described are the *deficiencies* of the assessments by the UN. The subsequent presented *clustering*, allows identifying *country groups* that should be in the focus of technical assistance and capacity building programs more easily.

[73] In our assessment, the table presenting the results for timeliness by regions and years of timeliness contains one exception. The timeliness of a single country isn't included in the table, but in the text. This refers to a country which reported "new" ANA data for a single table with a time lag of five years.

Deficiency of the Currently Composed Country Groups

The computed *regions* of the timeliness assessment (and the availability assessment) are merely nine *country groups,* based on the 192 UN Member States. These are given in the following table.

Table 17 Clustering of Regions of the UN Assessment

By economic classification	**Count**
Developed economies	37
Transition economies	19[74]
Developing economies	136
Total	192
Of developing economies, classification by regions	
Africa	53
Caribbean and Latin America	33
Western Asia	14[75]
Eastern, South-eastern, and Southern Asia	24
Oceania	12
Total	136

The *deficiency* of this regional subdivision is:

1) The *regional subdivision* of *developing economies* is disaggregated into five major regions according to the *MDG regional classification*[76]. Nevertheless, the presented *regional breakdown* is still not granular enough to provide detailed views on *individual regions* concerning the timeliness of their ANA

[74] There are now 17 transition economies, but at the time of the analyses it was 19. For recent changes of regional classifications see UNSD (2013b), the excerpt from the United Nations publication, Standard Country or Area Codes for Statistical Use (Series M, No.49/Rev 4), available online at: http://unstats.un.org/unsd/methods/m49/m49regin.htm (status: 11-02-2013), last accessed 11-06-2013).

[75] There are now 15 countries in the Western Asia region, but at the time of the analyses it was 14.

[76] Based on the Standard Country or Area Codes for Statistical Use (Series M, No.49/Rev 4). The MDG regional classification is included in the MDG country list (cf. footnote 67).

data (ditto for availability of ANA data). The *MDG regional classification* includes further compositions of subdivisions of regions. For example the region "Eastern, South-eastern, and Southern Asia" could be subdivided into its three individual sub-regions.

1) The UN assessment only considers the *UN Member States,* i.e. the set C_{UN}. Yet, the UN collects data for the set $C_{\mathrm{All}} \supset C_{\mathrm{UN}}$ of all countries which produce official ANA data (cf. United Nations 2009b, v).

The results of a more granular *regional assessment* could be used to better identify *regions* that might benefit from technical assistance and capacity building initiatives. For such *regions*, the responsible *regional organization* could provide this assistance.

Improved Subdivisions of Country Groups

The *regions* for our timeliness assessment (and for availability) cover all *regions* and *country groups* already considered by the UN assessment as above. The *MDG regional classification*[76] includes additional *regions* and *country groups* which we include, too. These are *subdivision* of the *transition economies* and of the *developing regions*, and further the *official MDG country groups*. We also consider some selected *organizations* and *economic groups*.

1) Concerning the *regional subdivision* we suggested the computation of additional geographical *regions*. Regarding the breakdown by **economic classification** we display the two subgroups for the *transition economies*. Furthermore, the **developing regions** (by the same main regions as in the UN assessment) have additional sub-regions. These additional sub-regions are given in the following table; numbers refer to the count of UN Member States.

Table 18 Additional Sub-Regions of Transition Economies and Developing Regions

Region	Sub-Region	Count
Transition economies	South-eastern Europe, and Commonwealth of Independent States (CIS)	6 11
Total		17[77]
Africa	Northern, and Sub-Sahara	5 48
Total		53
Caribbean and Latin America	Caribbean, and Latin America	13 20
Total		33
Eastern, Southeastern, and Southern Asia	Eastern, South-eastern, and Southern Asia	4 9 11
Total		24
Western Asia	-	15
Total		15[78]
Oceania	-	12
Total		12

For an *example* of the improved *regional subdivisions* of the major regions, see Figure 20, p. 156 (at the end of the following section) presenting a sample *results tables* of our timeliness assessment.

2) In addition to the assessment for the 192 UN Member States, a second one includes the set C_{All} / C_{UN} of **32 countries, areas and territories** considered by the set of countries in the MDG country list[76], that are not UN members. We use the same *clustering* as above for the assessment considering the set C_{All} of "*all countries*" (i.e. countries, areas, and territories) by adding the

[77] One transition economy is now counted to the developed and another one to the developing countries.

[78] One former transition economy of the Western Asia region is now counted a developing country.

additional countries c ∈ C_{All} / C_{UN} to the corresponding *regions*. We also refer to the additional 32 countries as "other countries from M49 list"[79].

3) The groups of the official **Millennium Development Goals** (**MDG countries**) are presented separately.

Table 19 Millennium Development Goals Countries

Millennium Development Goals (MDG) countries	**Count**
Landlocked Developing Countries (LLDCs),	31
Small Island Developing States (SIDS), and	51
Least Developed Countries (LDCs).	50

4) Selected **country groups** are presented separately, i.e. **international organizations** and **economic groups**. These are in given in the following table:

Table 20 International Organizations and Economic Groups

International organizations and **economic groups**	**Count**
Caribbean Community (CARICOM),	20
Organization for Economic Cooperation and Development (OECD),	30
Southern African Development Community (SADC),	15
EU27 countries,	27
G20 countries, and the	19
non-G20 advanced economies (also referred to as "other advanced economies").	23

5) The **UN Regional Commissions** are presented, see Table 21. The countries represented by the *UN Regional Commissions* are subdivided by their

[79] The country list in the Standard Country or Area Codes for Statistical Use (Series M, No.49/Rev 4) includes nine additional countries or areas that are not members of the set C_{All} of "MDG list" countries.

geographical regions and *economic classification* (using the clustering of the regional and economic classification given above).

Table 21 UN Regional Commissions

UN Regional Commissions	Count
Economic Commission for Africa (ECA)	52
Economic Commission for Europe (ECE)	56
Economic and Social Commission for Asia and the Pacific (ESCAP)	49
Economic and Social Commission for Western Asia (ESCWA)	14
Economic Commission for Latin America and the Caribbean (ECLAC)	44

Our assessment with the above outlined granularity is conducted in *three parts*[80]:

i. The *first part* shows the results for the economic classifications and regions considering only the set C_{UN} of *UN Member States* (192 countries). (The results for this part of the *timeliness assessment* are presented by Table 25 Timeliness for UN Member States by Most Recent Reference Year, p. 177.)

ii. The *second part* considers the set C_{All} of *all countries, areas and territories* (i.e. 32 countries in addition to the set C_{UN}). This table also *includes* the three special *MDG country groups*. (The results for this part of the *timeliness assessment* are presented by Table 26 Timeliness for All Countries, Areas and Territories by Most Recent Reference Year, p. 180.)

iii. The *third part* provides the *country groups* according to membership in *organizations* and *economic groups*. Further, it also includes the *UN Regional Commissions* and the corresponding subdivisions. (The results for this part are

[80] Thanks to Jillian Campbell from the United Nations, for setting up an Excel workbook, that aggregates country results by regions and country groups for analysis according to the MDG regional classification, UN Regional Commissions, and other country groups (other than some of the analyzed organizations and economic groups, i.e. OECD, EU27, G20, non-G20 advanced economies).

presented by Table 27 Timeliness for Membership in Other Organizations/ Groups by Most Recent Fiscal Year, p. 183.)

The last part of our *timeliness assessment* (Table 27) is particularly interesting for analyzing *country groups* like the G20 and OECD countries, the other two parts (Table 25 and Table 26) are particularly interesting for studying *specific regions*.

1.2.4 **Description of the Results Tables**

In this section we describe the *format* of the *results table* of our *timeliness assessment*.

For the description we consider the *results table* for part one of our *timeliness assessment* that considers the set C_{UN} of *UN Member States* (i.e. Table 25, p. 177). The *general structure* of the *results tables* is identical in all three parts of our *timeliness assessment*. We refer to the formulas and symbols we defined in part II of this study in order to explain the context and results presented by the different *sections* of the *results table*. Figure 20 (p. 156) illustrates Table 25 with references in brackets [the numbers] used to explain the corresponding *sections*.

The table *columns* (refer to **[1]** in the figure) begin with the region respectively country group (refer to chapter 1.2.3) indicating the considered set C of countries, that we assess for the *timeliness* of its represented member countries $c \in C$. (The set C is for example the set of "*UN Member States*", or the region "*Northern Africa*"). The *second column* **[2]** is the total number of countries that assemble the set C (i.e. the cardinality $|C|$) of each considered region (respectively country group). – For example in case of the set "*UN Member States*" this is $|C| = |C_{\mathrm{UN}}| = 192$ countries, and for the set "*Northern Africa*" this is $|C| = |C_{\mathrm{Northern\_Africa}}| = 5$ countries. - *Column three* **[3]** gives first the number of countries of the set C_{R} that reported new data for the 2008 YB (i.e. the cardinality $|C_{\mathrm{R}}|$), followed by the corresponding percentage relative to the region's total (i.e. the cardinality $|C|$). For the set C of countries that assemble a considered region (respectively country group), we calculate the share h_{R} of the set C_{R} of countries that reported new ANA data by

$$h_{\mathrm{R}} := \frac{|C_{\mathrm{R}}|}{|C|} \text{ for any subset } C_{\mathrm{R}} \subseteq C \qquad (0 \le h_{\mathrm{R}} \le 1). \tag{1.8}$$

The following *columns* refer to the most recent available reference year[81] t_{RY} **[4]** (for the countries $c \in C_R$ that reported new data). The four reference years are 2005 to 2008 (t_{RY} = 2005, …, 2008 equating to t_{RY} = t-4, …, t-1)[82]. The values in the *rows* below each reference year (**[6]**, **[7])** are: First **[6]** the number of countries to each reference year, i.e. the cardinality $\left|C_n^T\right|$ (cf. Formula II.114) of the set C_n^T (cf. Formula II.111) of countries with the timeliness T(c) of value $n \in \boldsymbol{N}$ in years [for n = 4, … 1]. This is followed by **[7]** the percentage of countries with a specific timeliness (calculated via the share $h_n = \frac{\left|C_n^T\right|}{\left|C_R\right|}$, refer to Formula II.115; i.e. the percentages are based on the cardinality of the set C_R of countries that reported new data, indicated in column three **[3]**. The *last column* **[5]** gives the cumulative of the reference years 2007 and 2008 ($t_{RY} \geq$ t-2), i.e. referring to those countries with a *time lag* of not more than one year. These values are the cardinality $\left|C_1^T\right| + \left|C_2^T\right|$ of the set C_n^T of countries with the timeliness T(c) = n for n = 1, 2 and the shares $h_1 + h_2$ as percentage. Differences between individual percentage values for 2007 and 2008 and their cumulative percentage value are due to rounding.

[Note that the share h_R is based on the cardinality of the set *C* of countries comprising the considered region (or country group). In case of the UN Member States this are 192 countries. However, for five of these countries, the country data are included in the ANA data of another ("main") country. These five countries do not report individual data. Therefore they are not considered members of the set C_R of countries that reported new data, even if the "main" country is member of the set C_R. Thus for the region "*UN Member States*" the share h_R cannot reach one ($0 \leq h_R < 1$). Likewise for the other affected regions as follows: Developing Regions (all five countries), Caribbean (two countries), Latin America (one country), Sub-Sahara Africa (two countries). Moreover, in case of the assessment of the 224 countries considered by the MDG country list, one non-UN member country is added. Thus the affected regions

[81] In the results table we use the term "fiscal year", accentuating that the reference year might cover a calendar year differently than the Western Calendar Year (i.e. with a different start or end date).

[82] For the purpose of readability, the reference yeas referring to t-5, t-6 are not included in the tables. A note that one country (Timor-Leste) apparently reported "new" data for t-6 is provided in the text.

are: All (M49) countries, areas and territories (all six countries), Developing Regions (all six countries), Northern Africa (one country).]

Reporting round for 2008 fiscal year data (Assessment of 2008 Yearbook data as of September 2009)	[2] Total number of countries for region	Of total: Replies for reporting round for 2008 fiscal year data	Of replies: Most recent fiscal year available [where 2005 ≡ T(c) = 4, …, 2008 ≡ T(c) = 1] 2005 $\lvert C_4^T \rvert$ (h_4)	2006 $\lvert C_3^T \rvert$ (h_3)	[4] 2007 $\lvert C_2^T \rvert$ (h_2)	2008 $\lvert C_1^T \rvert$ (h_1)	Total 2007+2008 [5] $\sum_{n=1}^{2} \lvert C_n^T \rvert$ (h_1+h_2)
UN Member States	**192**	**158 (82%)**	**2 (1%)**	**10 (6%)**	**79 (50%)**	**66 (42%)**	**145 (92%)**
Developed Regions [1]	**38**	**37 (97%)**	**0 (0%)**	**2 (5%)**	**13 (35%)**	**22 (59%)**	**35 (95%)**
[8] **Transition Countries**	**17**	15 ([illegible]) [3]	**0 (0%)**	1 (7%) [6]	**7 (47%)**	[7] 7 (47%)	**14 (93%)**
Transition countries of South-eastern Europe	6	6 (100%)	0 (0%)	1 (17%)	5 (83%)	0 (0%)	5 (83%)
Commonwealth of Independent States (CIS)	11	9 (82%)	0 (0%)	0 (0%)	2 (22%)	7 (78%)	9 (100%)
Developing Regions	**137**	**106 (77%)**	**2 (2%)**	**7 (7%)**	**59 (56%)**	**37 (35%)**	**96 (91%)**
Africa	53	34 (64%)	0 (0%)	3 (9%)	16 (47%)	15 (44%)	31 (91%)
Northern Africa	5	3 (60%)	0 (0%)	1 (33%)	2 (67%)	0 (0%)	2 (67%)
Sub-Saharan Africa	48	31 (65%)	0 (0%)	2 (6%)	14 (45%)	15 (48%)	29 (94%)
Latin America and the Caribbean	33	27 (82%)	0 (0%)	3 (11%)	15 (56%)	9 (33%)	24 (89%)
Caribbean	13	9 (69%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
Latin America	20	18 (90%)	0 (0%)	3 (17%)	9 (50%)	6 (33%)	15 (83%)
Asia	39	37 (95%)	2 (5%)	1 (3%)	23 (62%)	10 (27%)	33 (89%)
Eastern, Southern, and South-eastern Asia	24	22 (92%)	1 (5%)	0 (0%)	14 (64%)	6 (27%)	20 (91%)
Eastern Asia	4	3 (75%)	0 (0%)	0 (0%)	2 (67%)	1 (33%)	3 (100%)
Southern Asia	9	9 (100%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
South-eastern Asia	11	10 (91%)	1 (10%)	0 (0%)	6 (60%)	2 (20%)	8 (80%)
Western Asia	15	15 (100%)	1 (7%)	1 (7%)	9 (60%)	4 (27%)	13 (87%)
Oceania	12	8 (67%)	0 (0%)	0 (0%)	5 (63%)	3 (38%)	8 (100%)

Figure 20 Structure of the Results Table (Example)

Source: own illustration

The following provides two **examples** how to read the information provided with the *results table* for the *assessment* of *timeliness of ANA data publication* at the UN.

Example 1: The *row* "Developed Regions" (refer to **[8]** in the first *column* **[1]**) indicates that we consider the set *C* of countries of the "*Developed Regions*" (i.e. here the set C = *Developed Regions* = $C_{\text{Developed}}$)[83]. The *information* provided by the following *column* **[2]** is the cardinality of the set *C* of countries, thus a total of 38 countries belong to this region. The next *column* **[3]** indicates that 37 ($|C_R|$) out of 38 countries ($|C|$), i.e. 97%, have reported new data that was published in the 2008 YB (published in 2009 (t_{PY})). The following *columns* (2005, 2006, 2007, 2008) **[4]** indicate the most recent reference year (t_{RY}) that was reported. The *column* "2007" indicates that: 13 countries (refer to **[6]**), i.e. 35% (refer to **[7]**) of the set C_R of countries that reported new data, have as most recent ANA data, data for the reference year t_{RY} = 2007. Thus we say that these 13 countries have the *timeliness* "t-2" years,

[83] Regarding the considered set of countries *C*, the specific country group or region is given by the name of set *C*. We drop the additional index to *C*, since the variables and symbols are defined referring to *C*.

which is a *time lag* of one year. The last *column* **[5]** “Total of 2007 and 2008” indicates that 35 developed countries (95%) have reported new data, with not more than one year *time lag*, i.e. either for the reference year t_{RY} = 2007 or t_{RY} = 2008.

Example 2: Referring to the set of “*UN Member States*” (*top row* of the country groups in the first *column* **[1]**) and the columns associated to **[4]**, we observe that: 1% (2 countries), out of the 158 countries (column **[3]**) assembling the set C_R of countries that reported new ANA data, did report new data for the reference year t_{RY} = 2005 (we say “t-4”). Further, 6% (10 countries) reported ANA data for the reference year t_{RY} = t-3 (2006); 50% (79 countries) for the reference year t-2 (2007); and 42% (66 countries) reported t-1 (2008) data. Additionally, one country apparently reported “new” ANA data for a reference year which is not included in the columns of the results table. The data refer to the reference year t-6 (2003).[84]

The cumulative of these sets of countries with different reference years of the reported ANA data is 100% (i.e. 158 countries that reported new data).

Plotting these *percentages* for the *timeliness* T(c) = n, respectively by t_{RY} = t-n of the set of *UN Member States* leads to the following figure.

Cumulative Percentages of Timeliness Results for UN Member States

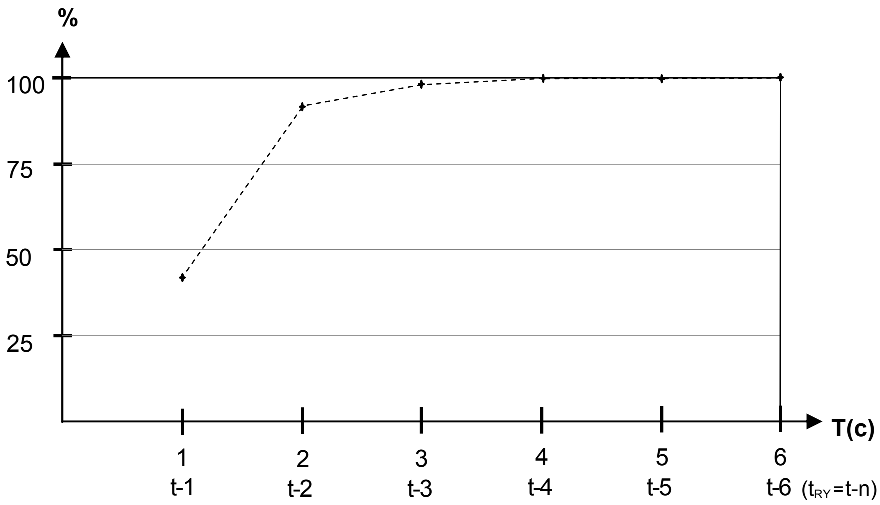


Figure 21 Timeliness, T(c) = n, as Cumulative Percentages of UN Member States (Considered is *C = UN Member States* of Figure 20)

Source: own illustration

[84] Refer for further explanation to the note in the timeliness assessment, chapter 1.3.1, p. 191ff.

1.2.5 Perspectives for Further Improvement of the Assessment Methodology

In the following subsections we outline perspectives for further improvements of our *assessment methodology*. We consider additional *country groups* for the assessed *regional aggregates* and different *data objects* for which we can assess the timeliness of ANA data publication.

Particularly, for those countries $c \in C_{YB} / C_R$ that are included in the YB, yet did not report new data for the present YB, we assess the *timeliness* via the *last time point* when a new reference year of ANA data was published in the YB. Further, we consider the identification of *annual reporters*, and regarding the *non-annual reporters* the distinction of *onetime* and *first time reporters* of ANA data.

Sample Applications, using real country data, are provided for the assessment of timeliness of ANA publication for the *set* C_K of countries with *"known" data*. Sample Applications are also provided for the assessment of timeliness of *individual* NAQ tables and of *individual* NA indicators within the set of published ANA data.

1.2.5.1 Different Subdivisions of Country Groups

With our subdivision of *regions* we follow the *MDG regional classification*[76] of countries. We could use *further country groups*, e.g. subdivisions of Africa or Latin America as per the "Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings" (UNSD 2013b)[85]. The *additional regions* could be employed for the *timeliness assessment* (as well as for the *availability assessment*).

We can consider the *additional subdivisions* of regions given in Table 22, below. The numbers refer to the count of 192 UN Member States in the corresponding sub-region.

[85] Excerpted from the United Nations publication, Standard Country or Area Codes for Statistical Use, Revision 4 (United Nations publication, Sales No. 98.XVII.9), available online at: http://unstats.un.org/unsd/methods/m49/m49regin.htm (status: 11-02-2013), last accessed 11-06-2013.

Table 22 Additional Sub-Regions of Africa, Europe, and Latin America

Region	Sub-Region	Count
Africa	Eastern Africa	17
	Middle Africa	10
	Northern Africa	5
	Southern Africa	5
	Western Africa	16
Total		53
Europe	Eastern Europe	10
	Northern Europe	10
	Southern Europe	14
	Western Europe	9
Total		43
Latin America	Central America	8
	South America	12
Total		20

We could further include the region *Caucasus and Central Asia*, which are currently not included in our subdivision. This is due to the subdivision of developing countries of Asia in our assessment. We include the countries of the *Caucasus and Central Asia* region in the set of Member States of the CIS (Commonwealth of Independent States) that are from Asia. Thus, these countries are classified as transition economies and consequently are not included again in the set of developing regions.

We could also consider further *regional organizations*, like the *Secretariat of the Pacific Community* (SPC). Such regional organizations could potentially foster the NA capacities of countries in their responsibility.

Moreover, apart from the regional classification or membership in organizations, we could consider *country groups* like the *top 50 countries* regarding the value of GDP or GNI in US dollars, or the group that makes up *90% of world GDP* or world GNI.

1.2.5.2 Timeliness of Non-Reporters

DEF. 35: Timeliness of Countries in the Yearbook with "Known" ANA Data

The length of the time period (T) measured in years between the end time point of the most recent reference year (t_{RY}) within the set of ANA data of a country published in the present YB with "known" data, and the publication year (t_{PYK}) of the past YB that made these data publicly available for the first time. Formally expressed:

ANAK := data object ANAK $\in$ *ANA* of "known" ANA data, i.e. the most recent reference year t_{RY} of ANA data in the current yearbook (for a specific country c) was already published in a previous YB.

C_K := set of countries with "known" ANA data in the YB (the latest available reference year t_{RY} refers to a reference year that is at most five years back from the YB publication year (i.e. $t_{RY} \geq$ t-5 < t-1)).

$$T : ANA \times C_K \longrightarrow \boldsymbol{N} \tag{1.9}$$

$$(\text{ANAK},c) \in ANA \times C_K \mapsto T(\text{ANAK},c) = t_{PYK}(\text{ANAK},c) - t_{RY}(\text{ANAK},c),$$

where

T(ANAK,c): timeliness of "known" ANA data ("ANAK") in years n $\in \boldsymbol{N}$ of country c $\in$ *C*; at the time point (year t_{PYK}) when published by the UN for the first time,

t_{PYK}(ANAK,c): year when the YB was published, that included the currently most recent reference year t_{RY}(c) of "known" ANA data for c $\in C_K$ for the first time. (Being the publication year t_{PY}(ANA,c) of that YB, when country c $\in C_K$ was last time member of the set C_R of countries that reported a new reference year of ANA data.)

t_{RY}(ANAK,c): the most recent reference year within the set of "known" ANA data for country c $\in C_K \subseteq C_{YB}$, with t_{PYK}(c) > t_{RY}(c).

Source: own definition

Our *assessment* of *timeliness* of ANA data publication, in subchapter 1.3, p. 175, includes only those countries $c \in C_R$ that reported new data during the latest YB reporting round. This could be extended to assessing the *timeliness* $T_K(c)$ of the additional countries $c \in C_K$ with "known" data, currently not considered by our assessment.

The set C_K of countries that are included in the YB with "known" data comprise of 16 countries (i.e. the cardinality $|C_K| = 16$). The *timeliness* T(ANAK,c) of the countries $c \in C_K$ with "known" data gives an indication regarding the ANA data production capacity of these countries. Yet, the production capacity, indicated by the *timeliness* T(ANAK,c) of countries with "known" data might have changed since the last time point of data reporting to the YB (published in year t_{PYK}).

Changing the set of *considered countries*, from the set C_R of countries that reported a new reference year t_{RY} of ANA data, to the set C_K of countries in the YB with "known" data, alters our definition of *timeliness* T(c) = T(ANA,c) (cf. DEF. 32).

We do not calculate the *timeliness* T(ANAK,c) as the difference between the most recent available reference year t_{RY} of ANA data publication year of the present YB. This would yield the same results as the *timeliness measure* of the UN methodology. Rather we calculate the *timeliness* T(ANAK,c) for the data object "known" ANA data ("ANAK") given the time point (publication year t_{PYK}) when the YB was published which made the presently most current reference year t_{RY} of ANA data available for the first time. Thus we actually calculate the *timeliness* T(c) of the ANA data publication for the time point when the country reported the present data as a new reference year of ANA data. With other words, we carry the latest actual *timeliness* T(c) of value $n \in \boldsymbol{N}$ in years forward in case no new data are reported.

The countries $c \in C_K = (C_{YB} / C_R)$ with "known" ANA data, i.e. those countries that did not report data for the present YB cannot be annual reporters of ANA data. Thus, when considering these countries in the *timeliness assessment*, also the *frequency* of reporting of countries could be identified by the assessment and provided as supplementary information. Further, an extended *timeliness study* could differentiate between *annual* reporters and the *timeliness* of these countries, and *non-annual* reporters and the *timeliness* of those countries.

To identify *annual* and *non-annual* reporters we consider the reporting of new data to the four most recent YBs, produced in publication year $t_{PY} = t_0 - 3$, $t_0 - 2$, $t_0 - 1$, t_0. Where t_0 refers to the current year regarding the time point of publication of the YB. For example identify *annual* and *non-annual* reporters given the most recent YB is published in 2009 we consider YBs of the publication years t_{PY} = 2006, …, 2009.

Table 23 Reporting to the Four Most Recent Yearbooks of "Annual Reporters"

Reporting to Yearbooks by publication year $t_{PY} = t_0-3, t_0-2, t_0-1, t_0$ of annual reporters ($c \in C_A$).				
t_{PY} / Country	**t_0 –3**	**t_0 –2**	**t_0 –1**	**t_0**
$c \in C_A$	☑	☑	☑	☑
		☑	☑	☑
			☑	☑

Source: own definition

We can define a country as an "*annual reporter*", i.e. a country $c \in C_A \subseteq C_R \subseteq C_{YB}$, if it reported a new reference year t_{RY} of ANA data at least for the current YB and for the previous YB and that has moreover no gap in reporting, in case ANA data were reported for earlier YBs (considering the reporting of data starting from the YB published in $t_{PY} = t_0 - 3$).

Thus, only a country $c \in C_R$ that reported new ANA data to the current YB can be a member of the set C_A of countries that are "*annual reporters*". [Consequently a country $c \in C_K$ with "known" data in the YB cannot be a member of the set C_A of "*annual reporters*", but is a "non-annual reporter".] However, a country $c \in C_R$ that reported new ANA data is not necessarily member of the set C_A of countries that are "*annual reporters*". With other words, it is a prerequisite to be member of the set C_R of countries that reported new ANA data to the current YB, in order to be member of the set C_A of countries that are "*annual reporters*". However, being a member of the set C_R of countries that reported new data does not imply membership in the set C_A of countries that are "*annual reporters*".

All countries c $\in C_{YB} / C_A$ that are included in the YB (i.e. set $C_R \cup C_K$), yet not members of the set C_A of countries that are "*annual reporters*", qualify as "n*on-annual reporters*", i.e. countries that are members of the set C_{NA}.

We may further distinguish within the set C_{NA} of *non-annual* reporters, those countries, which are "*onetime*" reporters. A "*onetime*" reporter of ANA data is a country c $\in C_{OT} \subseteq C_{NA} \subseteq C_{YB}$ that reported new ANA data for exactly *one* of the three YBs published in year $t_{PY} = t_0 - 3$, $t_0 - 2$, $t_0 - 1$. Thus not included in this set of considered YBs is the present YB, published in publication year $t_{PY} = t_0$.

Further, a country that only reported new ANA data for the *most current* YB, produced in publication year $t_{PY} = t_0$, is considered "*first-time*" reporters, i.e. a country c $\in C_{FT} \subseteq C_{NA}$.

Thus, the set $C_{NA} / (C_{OT} \cup C_{FT})$ indicates the set of the "*true*" non-annual reporters, as it excludes the *onetime* and *first-time* reporters of ANA data to the YB.

The distinction of these sets by reporting pattern helps users and IOs to identify potential countries with particularly poor NA data production capacity, e.g. *onetime* reporters. Further, it helps to anticipate the potential next transmission year of ANA data by a country, e.g. in case of a country that reports new data every *second* year. And it also provides information about the potential timeliness of the ANA data at time point of their publication for those countries that do not report new data annually.

The following table provides the reporting respectively non-reporting of new ANA data to the four most recent YBs ($t_{PY} = t_0 - 3$, $t_0 - 2$, $t_0 - 1$, t_0) for "*non-annual*" reports (c $\in C_{NA}$). The table is arranged by countries c $\in C_{FT}$ that are "*first-time*" reporters, countries c $\in C_{OT}$ that are "*onetime*" reporters, and countries c $\in C_{NA} / (C_{OT} \cup C_{FT})$ that are "*true*" non-annual reporters.

Table 24 Reporting respectively Non-Reporting of Non-Annual Reporters to the Four Most Recent Yearbooks by "First-Time", "Onetime", and "True" Non-Annual Reporters

Non-annual reporters $c \in C_{NA}$, divided by "first-time", "onetime", and "true" non-annual reporting to the four most recent Yearbooks (YB) by publication year $t_{PY} = t_0-3, t_0-2, t_0-1, t_0$.				
t_{PY} / Country	$\mathbf{t_0-3}$	$\mathbf{t_0-2}$	$\mathbf{t_0-1}$	$\mathbf{t_0}$
$\mathbf{c \in C_{FT}}$				☑
$\mathbf{c \in C_{OT}}$			☑	
		☑		
	☑			
$\mathbf{c \in C_{NA} / (C_{OT} \cup C_{FT})}$	☑		☑	☑
	☑	☑		☑
		☑		☑
	☑			☑
	☑	☑	☑	
		☑	☑	
	☑		☑	
	☑	☑		

Source: own definition

As already mentioned, these possible extensions are not considered in our *timeliness assessment*. Such extensions would require additional information about the reasons why new ANA data were not reported by some countries (i.e. $c \in C_{YB} / C_R$) included in the YB. For example, a reason a country $c \in C_{YB} / C_R$ in the YB did not report new data might be, that due to *current circumstances* (e.g. politic or economic) the NSO is not in a position to compile official NA data. Another reason might be that the *general statistical capacity* of the country is such, that the NA data can only be produced every second year.

It is also possible that the ANA data of a country have been produced during a project or technical assistance by *external experts*. In this case, or in case a country has *ceased* to produce ANA data, the country should be excluded form an assessment that aims to inform about the *timeliness* of ANA publication, since currently no data are produced for the country.

The following *sample application* provides an *example* for the assessment of the *timeliness* of ANA publication for countries which last reported a new reference year of ANA data for a previous YB. With other words, we assess the timeliness T(ANAK,c) of ANA data publication for the set C_K of countries with "known" data, (given the time point when the currently published ANA data were reported by the country and subsequently published via the YB for the first time as new ANA data).

Sample Application 12: Timeliness of Annual National Accounts in Case of "Known" Data

Country (Northern Africa Region)* **$c \in$ *Northern Africa***	**Published Reference Period** **t_{RY}**	**Time Point of Publication** **t_{PY}**	**Timeliness** **T(ANAK,c)**	**Time Lag** **TLR(ANAK,c)**
Algeria	2003	2005	2 (years)	1 (year)
Libyan Arab Jamahiriya	2007	2008	1 (year)	0 (years)

* Those countries $c \in C_K$ for which the latest available reference year was published in a previous YB.

Now we consider those countries of Northern Africa, which we did not assess in Sample Application 11, p. 126 regarding their timeliness T(c) of ANA publication, since their available ANA data were already published in a previous YB. Algeria (DZA) and Libyan Arab Jamahiriya (LBY) did not report new data for the 2008 YB (published in 2009). Thus their available ANA data refer to "known" data.

Example 1: The *most recent* available ANA data for *Algeria* refer to the reference year 2003. These data were *first* made publicly available in the YB published in year 2005. The reference year 2004 was the *expected reference year* for which ANA should have been produced and made available in the publication year 2005. Thus back then, when *Algeria* reported this ANA data set with the new reference year 2003, the ANA had a *timeliness* of *two* years, i.e. a *time lag* of *one* year.

Example 2: Similarly for *Libyan Arab Jamahiriya*, the ANA data for the reference year 2007 were *first* published by the YB that was published in 2008. Therefore, the actual *timeliness* of publication, at the time point when these data were first published, was *one* year, i.e. *no time lag*.

Referring to the formula for the *timeliness* of ANA of the set C_K with "known" data (c.f. DEF. 35, Formula 1.9), calculated via the publication year of that YB which first published these data and the most recent available reference year of ANA data for a country by $c \in C_K \mapsto T(ANAK,c) = t_{PYK}(c) - t_{RY}(c)$ this equates to

$$T(ANAK,DZA) = 2005 - 2003 = 2 \text{ (years)},$$
$$T(ANAK,LBY) = 2008 - 2007 = 1 \text{ (year)}.$$

1.2.5.3 Timeliness of SNA Concepts

Timeliness T(x,c) (cf. DEF. 5, p. 18) can be assessed for different *data objects* x $\in X$ (the set of data objects). Our timeliness assessment informs about the *timeliness* regarding the overall data set of *ANA* requested by the NAQ. That does mean, our *timeliness measure* T(c), which is our abbreviation for T(ANA,c) (cf. DEF. 32), has the *data object* x = ANA. [This *data object* (with multiple data items) refers to the full range of NA indicators (variables) within the NAQ by reference years (t_{RY}). Thus the *data object* "ANA" is synonymous for "NAQ", because the set of ANA data reported and subsequently published for a country is limited to the set of ANA data requested by the NAQ.]

Instead of considering the whole set of *ANA* data, the *timeliness* of ANA publication of countries could be assessed separately for each *individual subset* of published ANA data (i.e. according to different *SNA concepts* like for example: GDP by expenditures and GDP by value added of industries, or Gross National Income (GDP) and Gross National Disposable Income (GNDI)). To do so we need to assess the *timeliness* T(x,c) of a *data object* x $\in X$ published for a country c $\in C$, where our *data object* x $\in X$ refers to an *individual* set of NA data (i.e. a specific *SNA concept*).

We define these *timeliness measures* that assess the *timeliness* of *individual NAQ tables* (representing different *SNA concepts*) in the following. At the end of this chapter we provide an *exemplary assessment* that employs these measures.

The NAQ tables of the NAQ cover different *SNA concepts* which may become available (are produced and subsequently published by countries) at *different time points*. Hence, we can assess the *timeliness* of individual *NAQ tables* within the set of published *NAQ tables*, i.e. we assess the *timeliness* T(x,c) where the *data object* x $\in X$ refers to a single "*NAQ table*". We identify an individual *NAQ table* by its two-digit *NAQ table number* x.y $\in X.Y$:= {1.1, 1.2, 1.3, 2.1, 2.2, 2.3, 3.1, 3.2, 4.1, …, 4.9, 5.1}.

DEF. 36: Timeliness of Individual NAQ Tables

The Timeliness (T) of an individual NAQ table x.y $\in X.Y$ (the index set of NAQ table numbers regarding the set of NAQ tables) is the length of the time period measured in years n $\in \boldsymbol{N}$ between the end time point of the reference year (t_{RY}) for which the ANA data for the individual NAQ table with table number x.y are published in the YB, and the time point of their first time publication (year t_{PY})[86] in the YB. Formally expressed:

$X.Y$:= set of data objects (NAQ tables) identified by the index set of NAQ table numbers

C_R := subset ($C_R \subseteq C$) of countries c $\in C$, that reported a set of ANA data for a new reference year $t_{RY}(c)$, previously not published in the Yearbook (YB).

$$\mathrm{T} : X.Y \times C_R \rightarrow \boldsymbol{N} \tag{1.10}$$

$$(x.y,c) \in X.Y \times C_R \mapsto \mathrm{T}(x.y,c) = t_{PY}(x.y,c) - t_{RY}(x.y,c)$$

$$\text{with } t_{PY}(x.y,c) > t_{RY}(x.y,c),$$

where

T(x.y,c): the timeliness of ANA data publication in years n $\in \boldsymbol{N}$ of the individual NAQ table with table number x.y of country c $\in C_R$,

$t_{PY}(x.y,c)$: the year when the YB volume is published (that makes available a new reference year for NAQ table x.y for country c $\in C_R$), and

$t_{RY}(x.y,c)$: the most recent reference year (within the set of published data) for NAQ table x.y for country c $\in C_R \subseteq C_{YB}$ (of data being made available by the YB published in year t_{PY}).

Source: own definition

Each *NAQ table* consists of a set of data items, i.e. the product of the covered NA indicators and corresponding reference years (t_{RY}). Since a *NAQ table* consist of more than one data item, its *timeliness* is measured by the most recent available reference year within the set of published NA indicators (cf. last paragraph in DEF. 5, p. 18).

[86] Referring to the YB making the data available for the first time.

Some NAQ tables cover *multiple SNA concepts* (NAQ tables 1.3, 2.3, 4.1, …, 4.9) which again may become available with different *timeliness*. We can identify *NA indicators* (variables) within these NAQ tables that are representatives of different *SNA concepts*. Thus, we can assess the *timeliness* T(x,c) = T(v,c), where our *data object* v $\in$ *ANA* refers to an *individual NA indicator*. We define the timeliness of ANA data publication for an individual NA indicator v as follows:

DEF. 37: Timeliness of Individual NA Indicators
The Timeliness (T) of an individual NAQ indicator (variable) v $\in ANA$ (the set of NA indicators) is the length of the time period measured in years n $\in \boldsymbol{N}$ between the end time point of the reference year (t_{RY}) for which the ANA indicator is published in the YB, and the time point of its first time publication (year t_{PY}) in the YB. Formally expressed: ANA := set of ANA indicators C_R := set of countries that reported a new reference year t_{RY} of ANA data $T : ANA \times C_R \rightarrow \boldsymbol{N}$ (1.11) $(v,c) \in ANA \times C_R \mapsto T(v,c) = t_{PY}(v,c) - t_{RY}(v,c)$ with $t_{PY}(v,c) > t_{RY}(v,c)$, where T(v,c): the timeliness in years n $\in \boldsymbol{N}$ of ANA data publication regarding the individual NA indicator (v), $t_{PY}(v,c)$: the year when the YB volume is published (that makes available a new reference year of ANA for NA indicator v for c $\in C_R$), and $t_{RY}(v,c)$: the most recent reference year (within the set of published data) for NA indicator v for country c $\in C_R \subseteq C_{YB}$ (of data being made available by the YB published in year t_{PY}).

Source: own definition

For example, NAQ table 1.3 Relations among product, income, saving, and net lending aggregates covers the *main aggregates* (i.e. main NA indicators) beginning with the NA indicators GDP to the first balance GNI, then from GNI to GNDI, to saving, and to Net Lending/net borrowing (NL).

Thus, for NAQ table 1.3, we could assess the *timeliness* for the *NA indicator* "GNI" by T(GNI,c) = T(v,c) (i.e. the *data object* is GNI). The result for T(GNI,c) indicates the *timeliness* regarding the availability of all the NA indicators until the balance "GNI" (Thus we know the timeliness of all components needed for the computation of GNI from GDP, i.e. of the NA indicators representing the SNA concept of GNI). We could likewise assess the *timeliness* for *each following balance*, until the final balance of the capital account, with the closing *NA indicator* "NL", calculated by the *timeliness measure* T(NL,c) (i.e. the *data object* is NL).

The latter extension is not considered in our *timeliness assessment*, as the reporting and releasing of NA data at the UN is currently limited to once a year. Thus, due to the annual reporting, all NA data produced within the 12 month cycle are reported together. (With other words the ANA data of all different SNA concepts become available at the UN via the YB at the same time and consequently have the same timeliness of ANA data publication). Further, the *aim* of our present study is to determine the *timeliness* of produced (published) ANA data *in general*. It is of primary interest whether the first ANA main aggregates of the past reference period become available for a country $c \in C$ with no *time lag*, i.e. are produced for the reference year $t_{RY} = t\text{-}1$, or with additional *time lag*, particularly if this *lag* is more than one year, i.e. referring to a reference year $t_{RY} < t\text{-}2$.

However, the *timeliness* of publication of different *NAQ tables* or even of specific *NA indicators* might be of interest for researchers in academia and other organizations, who depend on particular data sets (like the set of constant price data) or NA indicators (like saving).

The following provides an example for the determination of the *timeliness* T(x.y,c) of *individual NAQ tables* x.y published for a country $c \in C_R$. Moreover, the sample application provides an example or the determination of the *timeliness* T(ANA,c) of ANA data publication for the data set "ANA" requested by the NAQ (i.e. the general timeliness for the set of ANA data produced by a country). The latter discriminates between the cases of publication of one and (as usual) multiple NAQ tables x.y for the country $c \in C_R$ in the YB.

Sample Application 13: Timeliness of Individual National Accounts Tables

Country	ANA data by NAQ Table No.	Published Reference Period t_{RY}	Time Point of Publication t_{PY}	Timeliness by NAQ Table T(x.y, c)	Timeliness of the ANA data of the country T(ANA,c)
Grenada	1.1 2.1, 2.2	2007 2008	09-2009*	2 years 1 year	1 year
Lao People's Democratic Republic	2.1	2007	09-2009*	2 years	2 years

* Publication of the 2008 Yearbook as PDF version.

Example 1: Let a country (as in case of Lao People's Democratic Republic (LAO)) report new ANA data for the 2008 YB published in year 2009. Yet, it only reports new data for a *single NAQ table*. Here this is NAQ table 2.1 for GDP by *production* in current prices, covering the ANA data for value added by industries and GDP. The data are published up to the reference year 2007.

Thus, the *timeliness* of publication of ANA data of the individual NAQ table 2.1 is two years. Consequently, since only one NAQ table is published, the *timeliness* of ANA publication of this country is also two years.

Example 2: Let a country (as in case of Grenada (GRD)) report new ANA data subsequently published in the 2008 YB (in publication year 2009). It provides ANA data for *three NAQ tables*. We can now assess the *timeliness* of the published *ANA data set*, and for each *individual NAQ table* (representing different SNA concepts).

The *timeliness* of a *data object* consisting of more than one data item is based on the most recent data item within this data object (cf. annotation in DEF. 5, p. 18).

Regarding Grenada, within NAQ table 1.1 (GDP by *expenditures* in current prices), all reported NA indicators are available until the reference year 2007. Within NAQ tables 2.1 and 2.2 (GDP by *production* in current respectively constant prices), the NA indicators for NAQ tables 2.1 and 2.2 are available until the reference year 2008.

(continued on next page)

Thus, NAQ table 1.1 has the *timeliness* of two years. Further NAQ tables 2.1 and 2.2 both have a *timeliness* of one year. That is: the ANA data of the NA concept *GDP by expenditures* (in current price data) become available in Grenada with an additional year of *time lag*. While the ANA data for the NA concept *GDP by production* become available with *no time lag*, for both the current and constant price data.

The *timeliness* of the *ANA data* publication T(c) is also one year, because the reference year 2008 is the most current reference year within the set of reported ANA data.

Referring to the formula (cf. DEF. 36) regarding *individual NAQ tables* in case of Grenada (GRD), this equates for the *timeliness* $T(x.y, c)$ for the NAQ tables with number x.y = 1.1, 2.1, 2.2 to

$$T(1.1,GRD) = t_{PY}(1.1,GRD) - t_{RY}(1.1,GRD) = 2009 - 2007 = 2 \text{ (years)},$$
$$T(2.1,GRD) = t_{PY}(2.1,GRD) - t_{RY}(2.1,GRD) = 2009 - 2008 = 1 \text{ (year)},$$
$$T(2.2,GRD) = t_{PY}(2.2,GRD) - t_{RY}(2.2,GRD) = 2009 - 2008 = 1 \text{ (year)}.$$

The recent reference year t_{RY} = 2008 is the most recent within the set of ANA data. Regarding the formula of *timeliness* of ANA $T(ANA,c) = t_{PY}(ANA,c) - t_{RY}(ANA,c)$ (cf. DEF. 32), this equates to

$$T(ANA,GRD) = t_{PY}(ANA,GRD) - t_{RY}(ANA,GRD) = 2009 - 2008 = 1 \text{ (year)},$$

Regarding the earlier case of Lao People's Democratic Republic (LAO) for the timeliness of *individual NAQ tables* $T(x.y, c)$ we get for NAQ table 2.1

$$T(2.1,LAO) = t_{PY}(2.1,LAO) - t_{RY}(2.1,LAO) = 2009 - 2007 = 2 \text{ (years)}.$$

Consequently, for the *timeliness* of ANA publication $T(ANA,c)$ we get

$$T(ANA,LAO) = t_{PY}(ANA,LAO) - t_{RY}(ANA,LAO) = 2009 - 2008 = 2 \text{ (years)}.$$

Sample Application 14: Timeliness of Individual NA Indicators (Variables) within National Accounts Tables

Country	ANA data by NA Indicator	Published Reference Period t_{RY}	Time Point of Publication t_{PY}	Timeliness by NA Indicator (i.e. of SNA concept) T(v, c)
Cuba	GDP GNI GNDI Saving NL	2007 2007 2007 2007 2007	09-2009*	2 years
Romania	GDP GNI GNDI Saving NL	2007 2007 2006 2006 2006	09-2009*	2 years 2 years 3 years 3 years 3 years

* Publication of the 2008 Yearbook as PDF version.

Example 1: Let a country (like Cuba) report new ANA data for NAQ table 1.3 (Relations among product, income, saving, and net lending aggregates), i.e. covering the *main aggregates* GDP, GNI, GNDI, Saving, and NL (and the components of these *NA indicators*). These five main aggregates (variables "v") indentify different *SNA concepts*. For all five SNA concepts, ANA data are published for the reference year 2007. Consequently, since the data are published in the publication year 2009, all five *SNA concepts* are published with a *timeliness* of two years.

Example 2: Another country (as in case of Romania) has different *timeliness* for the production of these SNA concepts (i.e. the SNA data represented by the five balances covered within the NAQ table 1.3). We observe that ANA data for the *main aggregates* GDP and GNI are published for the reference year 2007. The ANA data for the *main aggregates* GNDI, Saving, and NL are only published for the reference year 2006, thus with one year time lag to the aforementioned *NA indicators*.

Thus, GDP and GNI are produced with the *timeliness* of two years, while the other *SNA concepts* (GNDI, Saving, and NL) are published with *timeliness* three years.

(continued on next page)

Referring to the formula (cf. DEF. 37) for the *timeliness* of individual *NA indicators*, $(v,c) \in V \times C_R \mapsto T(v,c)$, for example for the *NA indicator* GDP this equates to

$$T(\mathrm{GDP},c) = t_{PY}(\mathrm{GDP},c) - t_{RY}(\mathrm{GDP},c).$$

Regarding our two sample countries, we assess a set of selected *NA indicators* given by *NAI* = {GDP, GNI, GNDI, Saving, and NL}. In case of Cuba (CUB), where all five main aggregates are published with the same *timeliness*, T(v,c) equates to

$$T(v, \mathrm{CUB}) = 2009 - 2007 = 2 \text{ (years)}, \quad \text{for all } v \in NAI.$$

In case of Romania (ROM), where the five main aggregates are published in two sets of different *timeliness*, we get

$$T(\mathrm{GDP},\mathrm{ROM}) = T(\mathrm{GNI},\mathrm{ROM}) = 2009 - 2007 = 2 \text{ (years), and}$$
$$T(\mathrm{GNDI},\mathrm{ROM}) = T(\mathrm{Saving},\mathrm{ROM}) = T(\mathrm{NL},\mathrm{ROM}) = 2009 - 2006 = 3 \text{ (years).}$$

1.3 Assessment of Timeliness

In this subchapter, we present the updated and improved results for the *assessment* of *timeliness* corresponding to our two *timeliness measures* (T(c) respectively TLR(c)), which we defined and delineated in part II, chapter 3.3, p. 117 ff. We employ our improved methodology, explained in detail in the previous subchapter.

Our *assessment* extends over three parts as explained in chapter 1.2.3, p. 148ff. First we present the *assessment* of *timeliness* for the 192 *UN Member States*, i.e. the set C_{UN}. Then, for all 224 countries, i.e. the set C_{ALL} (referring to the additional 32 countries $c \in C_{\mathrm{ALL}} / C_{\mathrm{UN}}$ in the country list of the MDG regional classification that are *non-UN member countries*). The second section also covers the special country groups defined with the Millennium Development Goals (MDG), i.e. *special interest groups* (like LDCs). In the third section we assess, *economic groups* (e.g. G20), *other international organizations* (e.g. OECD), and *UN regional organizations* (e.g. ECA).

Subsequent to our *timeliness assessments* we provide two *comparisons*, first of our results to those of the assessment by the UNSD (2008), and furthermore to results we computed by the UN methodology using the same source data as for our study.

Ultimately we will see, that in contrast to the UN *assessment* (UNSD 2008), our *assessment* indicates that the *timeliness* T(c) of ANA data reported to and subsequently published by the UN is better than suggested by the results of the UN study (see chapter 1.3.4, p. 184ff. at the end of this subchapter, with the *comparison* to the results of the UN assessment).

1.3.1 **UN Member States**

The findings of our *assessment* (cf. Table 25, p. 177) allow the *conclusion* that overall the *timeliness* T(c) of published ANA data for the *UN Member States* is "good". Precisely, that does mean we observe for the considered set $C_{\mathrm{UN}} \cap C_{\mathrm{R}}$ of *UN Member States* (C_{UN}) that actually reported new data (C_{R}), that for *over 90%* the most recent reference year t_{RY} of ANA is either t-1 or t-2 (i.e. the reference year 2008 or 2007). It is a surprising finding that merely a very small share of the countries $c \in C_{\mathrm{R}}$ that reported new data for the 2008 YB published in year $t_{\mathrm{PY}} = 2009$, did so with more than one year *time lag* TLR(c) behind the benchmark reference year $t_{\mathrm{RY}} = 2008$.

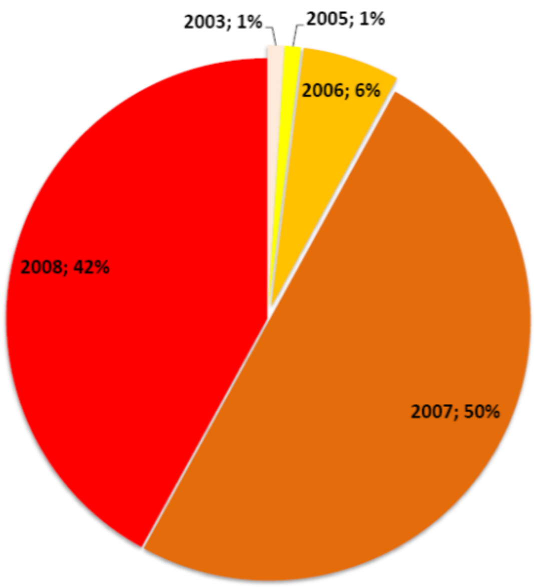


Figure 22 indicates for the *UN Member States*, the *timeliness* T(c) of published ANA data by reference year t_{RY}, where 2008 refers to t_{RY} = t-1 data. It is evident that the ANA of nearly all countries doesn't have a *time lag* TLR (c) of more than one year. For less than 10% of the countries the new set of published ANA data refers to a reference year $t_{RY} \leq$ t-3.

Figure 22 UN Member States Timeliness of Published Data by Reference Year

Source: own illustration. (Based on 2008 YB.)

The most *delayed* data is published for the reference year t_{RY} = 2003 (t-6), i.e. with a *time lag* of 5 years, yet for a single *UN Member State*. This country is Timor-Leste[87]. Further, only 2 countries (1%)[88] reported data for 2005, these are Myanmar and Qatar. [In case of Timor-Leste the *timeliness* of 5 years of the reported ANA is likely due to failure of the country to report the particular data together with the set of data, reported two years earlier, for the 2006 YB. The particular ANA data reported, are merely the constant price data for GDP by expenditures of NAQ table 1.2, corresponding to the previously published current price data (NAQ table 1.1). There is no indication that the constant price data were indeed produced with an additional *time lag* to the current price data. The country notes to the NA data reveal that the data were produced by the "UN Transitional Administration in East Timor […], with the assistance of the Boston Institute for Developing Economies, Massachusetts" (cf. United Nations 2009e, 575). Thus, it can be concluded that all NA data were

[87] This data shouldn't really be counted for the assessment of the timeliness of ANA publication and for conclusions regarding the NA production capacity, since the NA data were not produced by a national agency. Further, as we assume that the reported constant price data were accidentally reported to the UN with delay, this also sheds light on the issue of underreporting of data.

[88] Percentages, i.e. shares h_n in %, are based on the cardinality of countries C_R that reported new data.

produced at the same time during a onetime technical assistance mission. Since then, no further NA data were published.]

Table 25 Timeliness for UN Member States by Most Recent Reference Year

Reporting round for 2008 fiscal year data (Assessment of 2008 Yearbook data as of September 2009)	Total number of countries for region	Of total: Replies for reporting round for 2008 fiscal year data	Of replies: Most recent fiscal year available [where 2005 ≡ T(c) = 4, …, 2008 ≡ T(c) = 1] 2005 $\lvert C_4^T \rvert$ (h_4)	2006 $\lvert C_3^T \rvert$ (h_3)	2007 $\lvert C_2^T \rvert$ (h_2)	2008 $\lvert C_1^T \rvert$ (h_1)	Total 2007+2008 $\sum_{n=1}^{2} \lvert C_n^T \rvert$ (h_1+h_2)
UN Member States	**192**	**158 (82%)**	**2 (1%)**	**10 (6%)**	**79 (50%)**	**66 (42%)**	**145 (92%)**
Developed Regions	**38**	**37 (97%)**	**0 (0%)**	**2 (5%)**	**13 (35%)**	**22 (59%)**	**35 (95%)**
Transition Countries	**17**	**15 (88%)**	**0 (0%)**	**1 (7%)**	**7 (47%)**	**7 (47%)**	**14 (93%)**
Transition countries of South-eastern Europe	6	6 (100%)	0 (0%)	1 (17%)	5 (83%)	0 (0%)	5 (83%)
Commonwealth of Independent States (CIS)	11	9 (82%)	0 (0%)	0 (0%)	2 (22%)	7 (78%)	9 (100%)
Developing Regions	**137**	**106 (77%)**	**2 (2%)**	**7 (7%)**	**59 (56%)**	**37 (35%)**	**96 (91%)**
Africa	53	34 (64%)	0 (0%)	3 (9%)	16 (47%)	15 (44%)	31 (91%)
Northern Africa	5	3 (60%)	0 (0%)	1 (33%)	2 (67%)	0 (0%)	2 (67%)
Sub-Saharan Africa	48	31 (65%)	0 (0%)	2 (6%)	14 (45%)	15 (48%)	29 (94%)
Latin America and the Caribbean	33	27 (82%)	0 (0%)	3 (11%)	15 (56%)	9 (33%)	24 (89%)
Caribbean	13	9 (69%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
Latin America	20	18 (90%)	0 (0%)	3 (17%)	9 (50%)	6 (33%)	15 (83%)
Asia	39	37 (95%)	2 (5%)	1 (3%)	23 (62%)	10 (27%)	33 (89%)
Eastern, Southern, and South-eastern Asia	24	22 (92%)	1 (5%)	0 (0%)	14 (64%)	6 (27%)	20 (91%)
Eastern Asia	4	3 (75%)	0 (0%)	0 (0%)	2 (67%)	1 (33%)	3 (100%)
Southern Asia	9	9 (100%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
South-eastern Asia	11	10 (91%)	1 (10%)	0 (0%)	6 (60%)	2 (20%)	8 (80%)
Western Asia	15	15 (100%)	1 (7%)	1 (7%)	9 (60%)	4 (27%)	13 (87%)
Oceania	12	8 (67%)	0 (0%)	0 (0%)	5 (63%)	3 (38%)	8 (100%)

Source: own compilation

Referring to Table 25 with the detailed results for *timeliness* T(c) = n of ANA publication by **regions**: For the **developed regions** and for the **transition countries**, the cumulative percentage of countries with optimal timeliness T(c) = 1 and second best T(c) = 2, is above 90% ($\sum_{n=1}^{2} \mathrm{h}_\mathrm{n} > 0{,}9$ for the share $\mathrm{h}_\mathrm{n} = \frac{\lvert C_\mathrm{n}^\mathrm{T} \rvert}{\lvert C_\mathrm{R} \rvert}$, cf. Formula II.115. It is particularly remarkable that even in the *developing* regions nearly all (91%) of the countries report data with that *timeliness*, i.e. for either the benchmark reference year t_{RY} = t-1 (2008) or at least for the preceding reference year t_{RY} = t-2 (2007).

Yet, of these three economic groups, the *developed* regions is the single group for which the majority of countries (though still only 59%) reported data for 2008. For the *transition* economies, merely half of the countries (47% i.e. 7 countries) made the 2008 benchmark. Finally, of the *developing* regions only 37 ($\lvert C_1^\mathrm{T} \rvert$) out of the 106 countries ($\lvert C_\mathrm{R} \rvert$) that reported new data, i.e. just above one third (35%), reported data for 2008.

The **worse timeliness** among the *developed* regions is observed for two countries that have a *time lag* of three years; further one country did not report any new data. Among the *transition* economies two countries did not report new data, one has a *time lag* of three years and nearly half the countries, i.e. seven countries (47%) have a *time lag* of two years.

The *developing* regions reported new data for only about three quarters of the countries (77%). Thus the other countries either do not even produce NA data or do not do so on an annual basis, or the countries' NA data were not captured for other reasons (e.g. the contact to the NSO could not be established). However, the majority of *developing* countries (56% i.e. 59 countries) reported NA for the reference year t-2.

Among the **developing regions**, the region **Africa** has the **best** *timeliness*. Nearly half of *Sub-Sahara Africa* (48%) reported data for the benchmark reference year 2008. However, from the *Northern Africa* region, none of the countries reported data for the benchmark reference year t-1 (though these countries even produce a substantial set of ANA data (see our data availability assessments in chapter 2). Only three countries reported new data, two (66%) of them for the reference year t-2 and one (33%) for t-3.

Of **Latin America**, the **Caribbean**, **Asia**, and **Oceania** about one third of the countries reported data with the optimum timeliness (t-1). Interestingly, of *Oceania* and the *Caribbean*, all countries that reported new ANA data have data available for the reference years t-1 or t-2; yet only about *two thirds* (67% respectively 69%) of the countries in these regions reported new data. (*Oceania* and the *Caribbean* are also among the three regions with the least capability to produce the full 1993 Minimum Requirement Data Set (MRDS) of NA (see our availability assessment in chapter 2). In *Latin America* and *Western Asia* (which are particularly strong in producing the 1993 MRDS), 90% respectively 100%, reported new data. However, of *Latin America,* 3 out of 18 countries (nearly 20%) reported data with rather poor *timeliness*, precisely with a *time lag* of two years. For *Western Asia* one country (7%) reported data for the reference year t-3 and another one for t-4 (that is the worst *time lag*, with exception of Timor-Leste, only matched by one more country from *South-eastern Asia*.)

1.3.2 All Countries, Areas, and Territories

Considering *all countries*, *areas* and *territories* (set C_{All}), the *overall* result is not much different than for the UN Member States (set C_{UN}), only. However, with the inclusion of the *additional 32 countries* the results for the *timeliness* of ANA publication deteriorate, because none of the *additional countries* provided data of the benchmark reference year 2008 and one even provided t-4 data, i.e. has a *time lag* of three years.

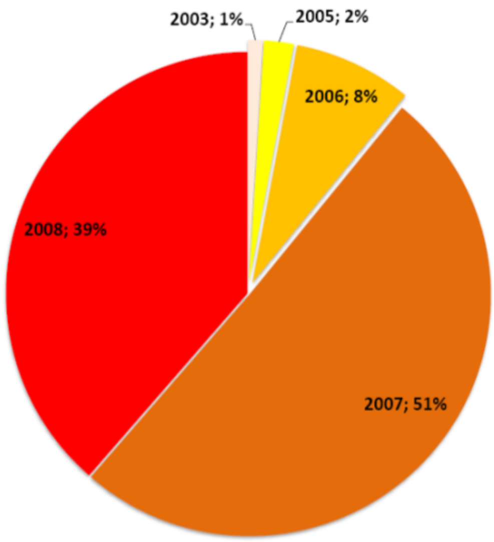


Figure 23 All (M49) Countries, Areas, and Territories Timeliness of Published Data by Reference Year

Source: own illustration. (Based on 2008 YB.)

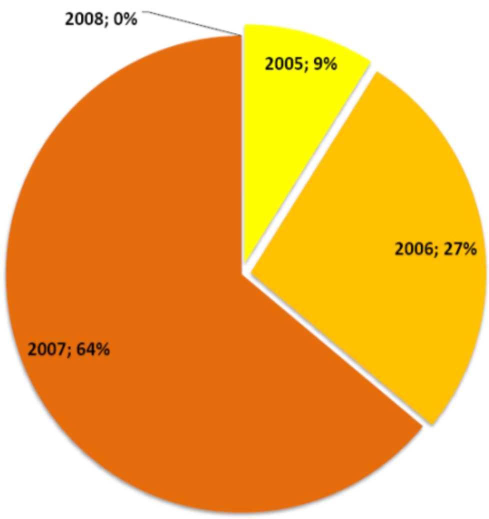


Figure 24 Non-UN Countries (from M49 List) Timeliness of Published Data by Reference Year

Source: own illustration. (Based on 2008 YB.)

Figure 23 and Figure 24 illustrate the *timeliness* of ANA publication for the set C_{All} of *all countries*, *areas* and *territories* (left) respectively for only the set C_{All} / C_{UN} of the additional 32 *non-UN member countries* (right). Figure 24 shows that the performance of the **non-UN member countries** is substantially below the global average. Merely 11 out of the additional 32 countries (34%) *reported* new data for the YB publication. Of those 11 countries ($|C_R|$) that reported new data, the majority of seven countries (64%) provided data for the reference year t-2. Yet, in case of three countries (27%), the data refers to the reference year t-3, and for one country (almost 10%) the most recent available data refer to the reference year t-4.

Table 26 Timeliness for All Countries, Areas and Territories by Most Recent Reference Year

Reporting round for 2008 fiscal year data (Assessment of 2008 Yearbook data as of September 2009)	Total number of countries for region	Of total: Replies for reporting round for 2008 fiscal year data	Of replies: Most recent fiscal year available [where 2005 ≡ T(c) = 4, …, 2008 ≡ T(c) = 1] 2005 $\lvert C_4^T\rvert$ (h_4)	2006 $\lvert C_3^T\rvert$ (h_3)	2007 $\lvert C_2^T\rvert$ (h_2)	2008 $\lvert C_1^T\rvert$ (h_1)	Total 2007+2008 $\sum_{n=1}^{2}\lvert C_n^T\rvert$ (h_1+h_2)
All (M49) Countries, Areas, and Territories	**224**	**169 (75%)**	**3 (2%)**	**13 (8%)**	**86 (51%)**	**66 (39%)**	**152 (90%)**
UN Member States	192	158 (82%)	2 (1%)	10 (6%)	79 (50%)	66 (42%)	145 (92%)
Other Countries from M49 list	32	11 (34%)	1 (9%)	3 (27%)	7 (64%)	0 (0%)	7 (64%)
Developed Regions	**43**	**38 (88%)**	**0 (0%)**	**2 (5%)**	**14 (37%)**	**22 (58%)**	**36 (95%)**
Transition Countries	**17**	**15 (88%)**	**0 (0%)**	**1 (7%)**	**7 (47%)**	**7 (47%)**	**14 (93%)**
Transition countries of South-eastern Europe	6	6 (100%)	0 (0%)	1 (17%)	5 (83%)	0 (0%)	5 (83%)
Commonwalth of Independent States (CIS)	11	9 (82%)	0 (0%)	0 (0%)	2 (22%)	7 (78%)	9 (100%)
Developing Regions	**164**	**116 (71%)**	**3 (3%)**	**10 (9%)**	**65 (56%)**	**37 (32%)**	**102 (88%)**
Africa	56	34 (61%)	0 (0%)	3 (9%)	16 (47%)	15 (44%)	31 (91%)
Northern Africa	6	3 (50%)	0 (0%)	1 (33%)	2 (67%)	0 (0%)	2 (67%)
Sub-Saharan Africa	50	31 (62%)	0 (0%)	2 (6%)	14 (45%)	15 (48%)	29 (94%)
Latin America and the Caribbean	46	32 (70%)	0 (0%)	5 (16%)	18 (56%)	9 (28%)	27 (84%)
Caribbean	24	14 (58%)	0 (0%)	2 (14%)	9 (64%)	3 (21%)	12 (86%)
Latin America	22	18 (82%)	0 (0%)	3 (17%)	9 (50%)	6 (33%)	15 (83%)
Asia	42	40 (95%)	2 (5%)	2 (5%)	25 (63%)	10 (25%)	35 (88%)
Eastern, Southern, and South-eastern Asia	26	24 (92%)	1 (4%)	0 (0%)	16 (67%)	6 (25%)	22 (92%)
Eastern Asia	6	5 (83%)	0 (0%)	0 (0%)	4 (80%)	1 (20%)	5 (100%)
Southern Asia	9	9 (100%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
South-eastern Asia	11	10 (91%)	1 (10%)	0 (0%)	6 (60%)	2 (20%)	8 (80%)
Western Asia	16	16 (100%)	1 (6%)	2 (13%)	9 (56%)	4 (25%)	13 (81%)
Oceania	20	10 (50%)	1 (10%)	0 (0%)	6 (60%)	3 (30%)	9 (90%)
Official MDG Groupings							
Landlocked developing countries (LLDCs)	31	26 (84%)	0 (0%)	0 (0%)	14 (54%)	12 (46%)	26 (100%)
Small island developing States (SIDS)	51	30 (59%)	1 (3%)	2 (7%)	16 (53%)	10 (33%)	26 (87%)
Least developed countries (LDCs)	50	33 (66%)	1 (3%)	2 (6%)	17 (52%)	12 (36%)	29 (88%)

Source: own compilation

Comparing the results in Table 25, p. 177 based on set C_{UN} of UN Member States with those in Table 26 we note: There are five *additional countries* in the **developed** regions. Yet, only one of them reported new data. These data have a one year *time lag*. There are no additional countries for the **transition** economies, thus no changes for this group. For the **developing** regions we have an additional 27 non-UN member countries. Three of them are in *Africa,* of which none reported new data. The *Caribbean* has 11 additional countries of which five reported new data. The timeliness is below that of the UN Member States of that region, with data referring to the reference years 2006 (t-3) and 2007 (t-2) (two and three countries, respectively). The *Latin America* region has two more non-UN member countries, of which none reported new data. In the *Asia* region, all three additional countries reported new data, one for 2006 and the others for 2007. The other eight countries are of the *Oceania* region, only two of them reported new data, one for 2005 (t-3) and one for 2007.

The assessment by countries of the **Millennium Development Goal (MDG) groups** in Table 26 shows that these groups do not perform below average of the *developing* regions. The *Landlocked Developing Countries* (LLDC) surprisingly perform even above global average. (For this group we observe similar results in our data availability assessment in chapter 2.) Even all three of the official *MDG groupings* perform better than the transition economies of South-eastern Europe.

1.3.3 **Other Organizations and Economic Groups**

This section provides the *assessment of the timeliness* of ANA for the major regional and international **organizations** and for important **economic groups**. A comparison of the results reveals remarkable and sometimes substantial differences between the examined *groups*.

Figure 25 presents for each group the *percentage* regarding the set C_R of countries that reported new data in relation to the group's total set C of member countries (referring to share $h_R := \frac{|C_R|}{|C|}$ for any subset $C_R \subseteq C$, cf. Formula 1.8, p. 154).

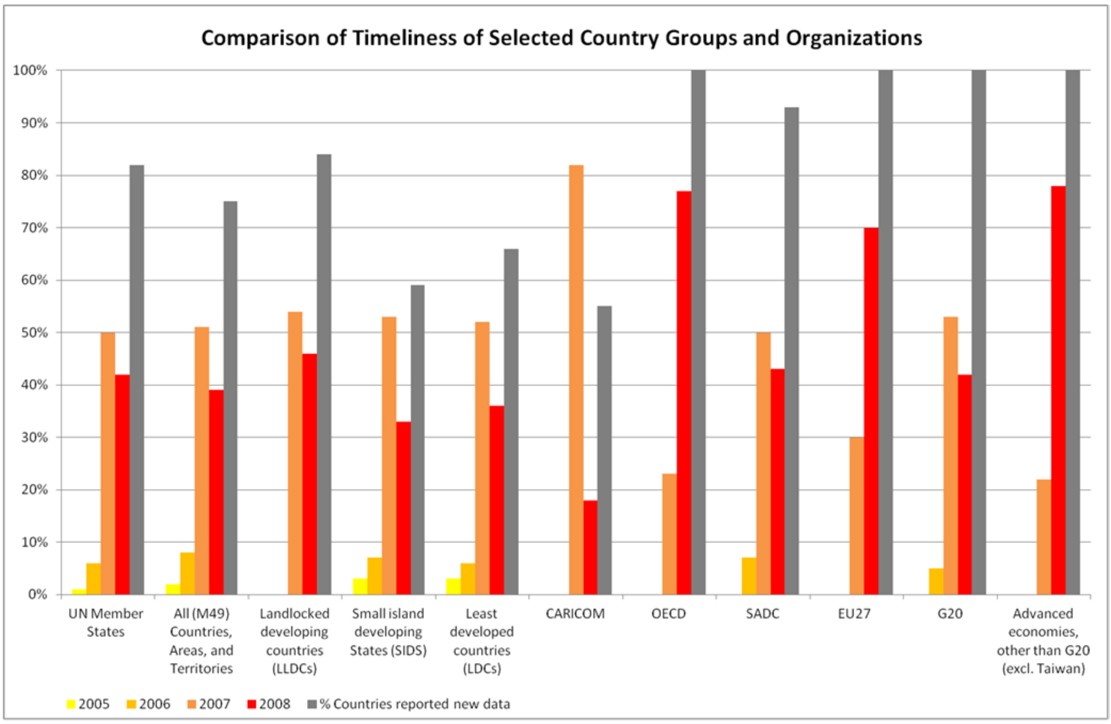


Figure 25 Comparison of Timeliness of Selected Country Groups and Organizations, Percent of Countries

Source: own illustration. (Based on 2008 YB.)

The percentage of countries that reported new ANA data is indicated by the grey bars. Where the reporting of new data to a group is less than 100%, the rest of the countries $c \in C / C_R$ that did not report new data, either do not produce data or do not report data annually.

Of the set C_R of countries that reported new data, the above figure presents the percentages (referring to the share h_n) for the sets of countries C_n^T (cf. Formula II.111 with a particular *timeliness* in years $n \in \boldsymbol{N}$. The percentages are given for the four reference year t_{RY} = 2005 to 2008, i.e. for t_{RY} = t-4 (yellow bars) to t_{RY} = t-1 (brick red bars). With this "yellow" to "brick red" *coloring* of the bars, that indicate the different years of timeliness, we continue the use of the "heat" *color code*. It indicates the superiority of performance regarding the *timeliness* of published ANA data by the intensity of the *color*, with "brick red" marking the top heat.

There are **significant differences** between the organizations *CARICOM*, *OECD*, and *SADC* and also between the groups of the *EU27*, *G20*, and *other advanced economies*. Regarding the *CARICOM* and *OECD*, the distribution of countries reporting t-1 data t-2 data are basically reversed.

The *CARICOM* has a particularly **bad** *timeliness* and low reporting of new data. The *EU27* group, *OECD*, and the *non-G20 advanced economies* (see row "Advanced, other than G20" in Table 27), are those with the **best** *timeliness*.

The *G20* countries perform below the *timeliness* of the other advanced (non-G20) economies. The *G20* and *SADC* countries indicate similar patterns of *timeliness* at a substantially lower performance than the previous groups, with the *timeliness* pattern being closer to that of the global average.

The *OECD* and *non-G20 advanced economies* slightly outperform the *EU27* group. Of the earlier groups, 77% and 78%, respectively, of the countries reported ANA data for the reference year (2008). For the EU27 this is still 70%. No country in these groups reported ANA data with a *time lag* of more than one year.

The *G20* group also performs below the average of the *developed regions*. Ten out of the 19 countries (53%) reported data with one year *time lag* and one country (5%) has a *time lag* of even two years to the benchmark reference year (i.e. reported the reference year t_{RY} = t-3). Only 8 out of 19 *G20* countries (42%) reported data for 2008, which is even below the average of the *transition economies*.

Table 27 Timeliness for Membership in Other Organizations/ Groups by Most Recent Fiscal Year

Reporting round for 2008 fiscal year data (Assessment of 2008 Yearbook data as of September 2009)	Total number of countries for region	Of total: Replies for reporting round for 2008 fiscal year data	Of replies: Most recent fiscal year available [where 2005 ≡ T(c) = 4, …, 2008 ≡ T(c) = 1]				
			2005	2006	2007	2008	Total 2007+2008
			$\lvert C_4^T \rvert$ (h_4)	$\lvert C_3^T \rvert$ (h_3)	$\lvert C_2^T \rvert$ (h_2)	$\lvert C_1^T \rvert$ (h_1)	$\sum_{n=1}^{2} \lvert C_n^T \rvert$ (h_1+h_2)
Other Organizations/Economic Groupings							
CARICOM	20	11 (55%)	0 (0%)	0 (0%)	9 (82%)	2 (18%)	11 (100%)
OECD	30	30 (100%)	0 (0%)	0 (0%)	7 (23%)	23 (77%)	30 (100%)
SADC	15	14 (93%)	0 (0%)	1 (7%)	7 (50%)	6 (43%)	13 (93%)
EU27	27	27 (100%)	0 (0%)	0 (0%)	8 (30%)	19 (70%)	27 (100%)
G20	19	19 (100%)	0 (0%)	1 (5%)	10 (53%)	8 (42%)	18 (95%)
Advanced, other than G-20 (excl. Taiwan Prov. of China)	23	23 (100%)	0 (0%)	0 (0%)	5 (22%)	18 (78%)	23 (100%)
UN Regions							
ECA	52	34 (65%)	0 (0%)	3 (9%)	16 (47%)	15 (44%)	31 (91%)
Northern Africa	5	3 (60%)	0 (0%)	1 (33%)	2 (67%)	0 (0%)	2 (67%)
Sub-Saharan Africa	47	31 (66%)	0 (0%)	2 (6%)	14 (45%)	15 (48%)	29 (94%)
ECE	56	52 (93%)	0 (0%)	3 (6%)	20 (38%)	29 (56%)	49 (94%)
Developed regions	40	39 (98%)	0 (0%)	3 (8%)	16 (41%)	20 (51%)	36 (92%)
of which: South-eastern Europe transition countries	*6*	*6 (100%)*	*0 (0%)*	*1 (17%)*	*5 (83%)*	*0 (0%)*	*5 (83%)*
Commonwealth of Independent States (CIS)	11	9 (82%)	0 (0%)	0 (0%)	2 (22%)	7 (78%)	9 (100%)
Developing regions	5	4 (80%)	0 (0%)	0 (0%)	2 (50%)	2 (50%)	4 (100%)
ESCAP	49	41 (84%)	1 (2%)	0 (0%)	24 (59%)	15 (37%)	39 (95%)
Developed regions	3	3 (100%)	0 (0%)	0 (0%)	2 (67%)	1 (33%)	3 (100%)
Commonwealth of Independent States (CIS)	8	6 (75%)	0 (0%)	0 (0%)	2 (33%)	4 (67%)	6 (100%)
Developing regions	38	32 (84%)	1 (3%)	0 (0%)	20 (63%)	10 (31%)	30 (94%)
Asia	26	24 (92%)	1 (4%)	0 (0%)	15 (63%)	7 (29%)	22 (92%)
Eastern Asia	4	3 (75%)	0 (0%)	0 (0%)	2 (67%)	1 (33%)	3 (100%)
Southern Asia	9	9 (100%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
South-eastern Asia	11	10 (91%)	1 (10%)	0 (0%)	6 (60%)	2 (20%)	8 (80%)
Western Asia	2	2 (100%)	0 (0%)	0 (0%)	1 (50%)	1 (50%)	2 (100%)
Oceania	12	8 (67%)	0 (0%)	0 (0%)	5 (63%)	3 (38%)	8 (100%)
ESCWA	14	13 (93%)	1 (8%)	3 (23%)	7 (54%)	2 (15%)	9 (69%)
ECLAC	44	38 (86%)	0 (0%)	3 (8%)	18 (47%)	17 (45%)	35 (92%)
Developed regions	10	10 (100%)	0 (0%)	0 (0%)	3 (30%)	7 (70%)	10 (100%)
Developing regions	34	28 (82%)	0 (0%)	3 (11%)	15 (54%)	10 (36%)	25 (89%)
Caribbean	13	9 (69%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
Latin America	20	18 (90%)	0 (0%)	3 (17%)	9 (50%)	6 (33%)	15 (83%)
Other (Asia)	1	1 (100%)	0 (0%)	0 (0%)	0 (0%)	1 (100%	1 (100%)

Source: own compilation

Table 28 depicts as example the group of G20 countries and advanced economies that are not members of the G20 (cf. Table 27, above), showing that G20 countries have a performance below the non-G20 advanced economies.

Table 28 Timeliness of Group of G20 Countries and of Advanced Economies Other than G20

Set of countries that reported new data with a timeliness T(c) = 1 year (i.e. no time lag).	
G20 countries	**Advanced economies other than G20**
No. (percentage)	No. (percentage)
8 (42%)	18 (78%)

Source: own compilation. (Based on 2008 YB, Published in 2009.)

Our assessment by **UN Regional Commissions**, i.e. by countries located within the Commission's responsibility, shows that the *ESCWA* region has particularly long *time lags* in reporting. Since the assessment results indicate that more than 30% of the *ESCWA* countries reported data with a *time lag* of more than one year, these countries might benefit from statistical capacity programs that foster the *timely production* of source statistics and NA data. (Such detailed *assessment* according to groups of countries within the responsibility of *UN Regional Commissions* can be used effectively to identify clusters of countries that could be served though the UN by technical assistance. This applies particularly for our NA data availability study in chapter 2.)

1.3.4 Comparison with Results of the Previous UN Assessment

The results of the **UN timeliness study** (cf. UNSD 2008) and of **our study** are not directly comparable. The two studies were carried out based on YBs published in different years. The *UN timeliness study* refers to the country data published in the 2006 YB. *Our study* is based on the two years later 2008 YB, published in 2009. Therefore, the new results interfere with changes in the timeliness of ANA data production at the country level.

Further, besides a different methodology, the *UN study* only provides results as percentages. These percentages refer to a different basic set than that, which we used for the computation of percentages in our timeliness assessment. This is explained in chapter 1.2.1, p. 137ff., concerning the deficiencies of the current UN assessment methodology. To compare the timeliness results by actual numbers of countries with a particular timeliness in year(s) $n \in \boldsymbol{N}$ of *timeliness*, the cardinalities of the sets of countries need to be calculated from the percentages presented in the UN study. Due to rounded percentages in the UN study, this yields an approximate number for the set of countries with a particular timeliness.

Table 29 presents results of the timeliness study (UNSD 2008) for the total of UN Member States and their subdivision by economic classification, based on the UN methodology (i.e. for $T_{YB}(c)$, cf. DEF. 34, p. 138). Further it presents the results of the timeliness assessment for these regions by our improved methodology (i.e. for T(c), cf. DEF. 32, p. 117). Presented are the percentage results of the set of countries with the optimum timeliness (i.e. reporting of data for the benchmark reference year t_{RY} = t-1).

Table 29 Timeliness of UN Member States by Economic Classification, by UN and Our Approach

Set of UN Member States with a timeliness of one year (optimum) as percentage.				
Region / Approach	**UN (total)**	**Developed**	**Transition**	**Developing**
$T_{YB}(c)$ (UN Approach)	33%	62%	53%	23%
T(c) (our Approach)	42%	59%	47%	35%

Source: UNSD (2008, 7) for UN approach based on 2006 YB, published in 2007, own compilation based on 2008 YB, published in 2009.

The **UN's assessment** indicates that only 33% of the *UN Member States* reported annual NA data for the reference year t_{RY} = t-1 (cf. Formula 1.7, p. 142 for the share $h_{YB,n}$ of countries with a particular timeliness of value n ∈ ***N*** in year(s)). By their economic status: 62% of the *developed* countries, 53% of the *transition* economies, and merely 23% of the *developing* countries reported data with no *time lag*. Regarding the low percentage of the developing regions, it should be noted that only 83 of the 137 countries of this region replied to the NAQ for the 2006 YB (cf. UNSD 2008, 7).

Our **improved methodology** reveals that the share h_n (cf. Formula II.115, p. 120) of the set C_n^T of countries with a particular *timeliness* T(c) in years n ∈ *N*, based on the actual number of the set C_R of countries that reported data, is substantially better. In contrast to the UN assessment, not 33%, but actually 42% of the *UN Member States* report data for the reference year t_{RY} = t-1. By their economic status, the difference of the results is particularly high for the *developing* regions, now showing 35% instead of 23% reporting data with no *time lag*. Nearly all countries of the *developed* regions and the *transition* countries reported new data. Thus, the change in percentage is primarily due to a lower number of countries that reported data for the benchmark reference year. The change for the *developing* region results from a higher number of countries that reported new ANA data for that region (107 countries) and primarily from the change of the set of countries considered for the assessment, to the set C_R of countries that reported data, which is also used as the basis set for the percentages.

To increase the *comparability* of the results yielded by these two *methodologies*, we compute the results for the *timeliness* $T_{YB}(c)$ of ANA according to the *UN methodology*, based on data published in the 2008 YB. In Table 30, p. 188 we present these results for the UN Member States and the regional aggregates introduced in chapter 1.2.3, p. 148ff, i.e. the country clusters with detailed subdivisions of regions. Thus, we *eliminate* the *differences* between the UNSD (2008) study and our present study, which result from changes due to improvements of the country's data reporting.

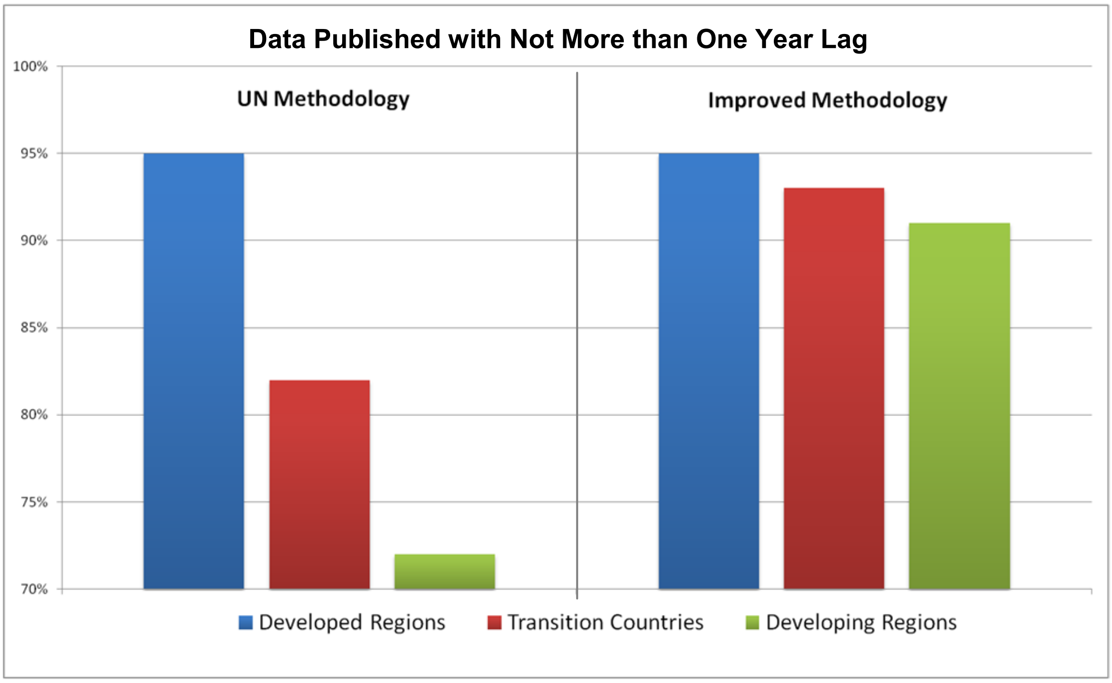


Figure 26 Data Published with Not More than One Year Lag by UN Methodology vs. Improved Methodology, Percent of Countries

Source: own illustration. (Both assessments based on 2008 YB.)

Figure 26, illustrates the comparison of the results for *timeliness* of ANA assessed by the current *UN methodology* and by our *improved methodology*, using data published in year t_{PY} = 2009 (i.e. in the 2008 YB). Displayed is the percentage of countries with *not more than one year time lag* for the country groups by *economic classification* according to the two different *methodologies*. The results by the *UN methodology* indicate a much lower percentage of countries in case of the *transition* economies and the particularly for the *developing* regions.

The absolute number (i.e. the cardinality $|C_1^T|+|C_2^T|$) of countries with the *timeliness* $T(c) \leq 2$ years (cf. the detailed results in Table 25 Timeliness for UN Member States by Most Recent Reference Year, p. 177) is in fact lower in case of our *methodology*. This is due to exclusion of those countries ($c \in C_{YB} / C_R$) in the YB that did not reported new data. The percentages (of $h_1 + h_2$) are higher by *our improved methodology*, since we define and calculate *timeliness* only for the set C_R of those countries that actually reported a new reference year of ANA data.

Only for the UN Member States that belong to the *developed* regions, the percentage is similar by both methodologies. This is because almost the full number of UN countries of the *developed* regions annually report new ANA data with the timeliness $T(c) \leq 2$ years.

Of the UN Member States, 31 *developing* countries did not report new data for the 2008 YB. Further, 21 out of the 31 countries (more than 2/3 of them) actually do not produce data (i.e. they have currently no ANA data available, that does mean the latest available reference year t_{RY} is more than five years back from the YB publication year (i.e. $t_{RY} < t\text{-}5$), which applies to all those countries that never reported ANA data or at least not for four consecutive reporting rounds). Therefore the percentage of the number of countries with a particular timeliness, based on the total number of countries C assembling the *developed* region is lower in case of the UN methodology for assessing timeliness, although the UN methodology indicates for three additional countries the timeliness $T_{YB}(c)$ = 2 years (i.e. of countries with "known" data).

With our *methodology* for the assessment of timeliness, we take into account that the countries $c \in C_{YB} / C_R$ in the YB, that did not reported new data, should not be assessed regarding the timeliness of ANA data publication. Moreover, the countries c ∈ C / C_{YB} that we consider as currently not producing ANA (i.e. that currently have no ANA data available) should not be included in the basis set, used to compute the percentage of countries with a particular timeliness. Therefore, our *improved methodology* reveals that the actual *timeliness* of ANA data in the *developing* regions is much better than suggested by the UN methodology. The results with our *improved methodology* are almost **20%p** (percentage points =: %p) **up** for the UN Member States in the *developing* region, compared to the results by the *UN methodology*.

Similarly to the developing regions, the results for *transition* economies are **up** by **10%p** if employing our *improved methodology*. Since there are no countries c ∈ C_K with "known", i.e. previously published data, for the *transition* economies, we know that the change in percentage points is solely the result of changing the basic set to the set C_R of countries that reported new data, instead of using the total set C of countries assembling a region.

Table 30 Timeliness by UN Methodology for UN Member States by Most Recent Reference Year

Reporting round for 2008 fiscal year data (Assessment of 2008 Yearbook data as of September 2009) by UN assessment methodology	Total number of countries for region	Of total: Replies for reporting round for 2008 fiscal year data	Of replies: Most recent fiscal year available [where 2005 ≡ T(c) = 4, …, 2008 ≡ T(c) = 1]				
			2005 $\lvert C_4^T \rvert$ (h4)	2006 $\lvert C_3^T \rvert$ (h3)	2007 $\lvert C_2^T \rvert$ (h2)	2008 $\lvert C_1^T \rvert$ (h1)	Total 2007+2008 $\sum_{n=1}^{2} \lvert C_n^T \rvert$ (h1+h2)
UN Member States	**192**	**158 (82%)**	**3 (2%)**	**16 (8%)**	**83 (43%)**	**66 (34%)**	**149 (78%)**
Developed Regions	**38**	**37 (97%)**	**0 (0%)**	**2 (5%)**	**14 (37%)**	**22 (58%)**	**36 (95%)**
Transition Countries	**17**	**15 (88%)**	**0 (0%)**	**1 (6%)**	**7 (41%)**	**7 (41%)**	**14 (82%)**
Transition countries of South-eastern Europe	6	6 (100%)	0 (0%)	1 (17%)	5 (83%)	0 (0%)	5 (83%)
Commonwalth of Independent States (CIS)	11	9 (82%)	0 (0%)	0 (0%)	2 (18%)	7 (64%)	9 (82%)
Developing Regions	**137**	**106 (77%)**	**3 (2%)**	**13 (9%)**	**62 (45%)**	**37 (27%)**	**99 (72%)**
Africa	53	34 (64%)	1 (2%)	5 (9%)	17 (32%)	15 (28%)	32 (60%)
Northern Africa	5	3 (60%)	0 (0%)	1 (20%)	3 (60%)	0 (0%)	3 (60%)
Sub-Saharan Africa	48	31 (65%)	1 (2%)	4 (8%)	14 (29%)	15 (31%)	29 (60%)
Latin America and the Caribbean	33	27 (82%)	0 (0%)	6 (18%)	16 (48%)	9 (27%)	25 (76%)
Caribbean	13	9 (69%)	0 (0%)	1 (8%)	7 (54%)	3 (23%)	10 (77%)
Latin America	20	18 (90%)	0 (0%)	5 (25%)	9 (45%)	6 (30%)	15 (75%)
Asia	39	37 (95%)	2 (5%)	1 (3%)	24 (62%)	10 (26%)	34 (87%)
Eastern, Southern, and South-eastern Asia	24	22 (92%)	1 (4%)	0 (0%)	15 (63%)	6 (25%)	21 (88%)
Eastern Asia	4	3 (75%)	0 (0%)	0 (0%)	2 (50%)	1 (25%)	3 (75%)
Southern Asia	9	9 (100%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
South-eastern Asia	11	10 (91%)	1 (9%)	0 (0%)	7 (64%)	2 (18%)	9 (82%)
Western Asia	15	15 (100%)	1 (7%)	1 (7%)	9 (60%)	4 (27%)	13 (87%)
Oceania	12	8 (67%)	0 (0%)	1 (8%)	5 (42%)	3 (25%)	8 (67%)

Source: own compilation

Table 30 gives the results for $T_{YB}(c)$ using the *UN methodology* for the assessment of the "timeliness of reporting" by, the "latest [reference][89] year of available data

[89] Added by the author.

reported by countries" (cf. UNSD 2008, 7 [table heading and column heading in the table presenting the results in the UNSD study]).

We observe for example that, using the *UN methodology*, the reference year 2006 is the most recent reference year for 16 of the *UN Member States* ($\left|C^{T}_{YB,3}\right|$ = 16). In our *improved* assessment there are only 10 countries for this reference year ($|C^{T}_{3}|$ = 10). Thus there are six extra countries c $\in C_K = C_{YB} / C_R$ with "*known*" data in the YB. For these extra countries included in Table 30 it is not clear whether the *actual timeliness* of the ANA data at the time point of data reporting (i.e. at time point of the publication year t_{PY} when the data of these countries were first made publicly available at the international level) had the *timeliness* T(c) = n of 1 or 2 years. Yet, using the *UN methodology* it appears as if these countries have the *timeliness* T(c) = 3 years, i.e. a *time lag* (TLR(c)) of 2 years.

This also illustrates, that the information about the *timeliness* of the ANA is less meaningful in case of the *UN methodology*, the further the indicated reference year t_{RY} is in the past. This is, because a country indicating a poor timeliness might indeed have a poor capacity to produce timely NA data, or it might not produce NA data annually, or it might have ceased to produce ANA data.

In the following paragraphs we emphasize some individual regions in order to pinpoint to *differences* between the *UN methodology* and our *improved methodology*.

In *Africa* one country indicates data for the reference year t_{RY} = 2005, i.e. t-4 by the *UN methodology*, while *our assessment* reveals that *none* of the countries c $\in C_R$ that reported new ANA has such poor *timeliness*. Particularly the percentages $h_1 + h_2$ referring to the number of countries that reported data with a *time lag* of not more than one year (TLR(c) $\leq$ 1, i.e. T(c) $\leq$ 2) is **31%p higher**, from 60% by the *UN methodology*, to 91% with our *improved methodology*.

Also *Latin America* and the *Caribbean* are **up** by **13%p**, from 76% to 89% of countries that reported new ANA data with a timeliness T(c) ≤ 2 of two years or better.

The *differences* are especially evident for the *Caribbean* and *Oceania*. These regions formerly indicated 77% respectively 67% of countries with a *timeliness* of the most recent available reference year t_{RY} referring to either t-1 or t-2. Now, both regions reveal that **100%** of the set C_R of countries that reported new data have this *timeliness*.

Table 31 Timeliness by UN Methodology for All Countries, Areas and Territories by Most Recent Reference Year

Reporting round for 2008 fiscal year data (Assessment of 2008 Yearbook data as of September 2009) by UN assessment methodology	Total number of countries for region	Of total: Replies for reporting round for 2008 fiscal year data	Of replies: Most recent fiscal year available [where 2005 ≡ T(c) = 4, …, 2008 ≡ T(c) = 1]				
			2005	2006	2007	2008	Total 2007+2008
			$\lvert C_4^T \rvert$ (h_4)	$\lvert C_3^T \rvert$ (h_3)	$\lvert C_2^T \rvert$ (h_2)	$\lvert C_1^T \rvert$ (h_1)	$\sum_{n=1}^{2} \lvert C_n^T \rvert$ (h_1+h_2)
All (M49) Countries, Areas, and Territories	**224**	**169 (75%)**	**5 (2%)**	**20 (9%)**	**90 (40%)**	**66 (29%)**	**156 (70%)**
UN Member States	192	158 (82%)	3 (2%)	16 (8%)	83 (43%)	66 (34%)	149 (78%)
Other Countries from M49 list	32	11 (34%)	2 (6%)	4 (13%)	7 (22%)	0 (0%)	7 (22%)
Developed Regions	**43**	**38 (88%)**	**1 (2%)**	**2 (5%)**	**15 (35%)**	**22 (51%)**	**37 (86%)**
Transition Countries	**17**	**15 (88%)**	**0 (0%)**	**1 (6%)**	**7 (41%)**	**7 (41%)**	**14 (82%)**
Transition countries of South-eastern Europe	6	6 (100%)	0 (0%)	1 (17%)	5 (83%)	0 (0%)	5 (83%)
Commonwalth of Independent States (CIS)	11	9 (82%)	0 (0%)	0 (0%)	2 (18%)	7 (64%)	9 (82%)
Developing Regions	**164**	**116 (71%)**	**4 (2%)**	**17 (10%)**	**68 (41%)**	**37 (23%)**	**105 (64%)**
Africa	56	34 (61%)	1 (2%)	5 (9%)	17 (30%)	15 (27%)	32 (57%)
Northern Africa	6	3 (50%)	0 (0%)	1 (17%)	3 (50%)	0 (0%)	3 (50%)
Sub-Saharan Africa	50	31 (62%)	1 (2%)	4 (8%)	14 (28%)	15 (30%)	29 (58%)
Latin America and the Caribbean	46	32 (70%)	0 (0%)	9 (20%)	19 (41%)	9 (20%)	28 (61%)
Caribbean	24	14 (58%)	0 (0%)	4 (17%)	10 (42%)	3 (13%)	13 (54%)
Latin America	22	18 (82%)	0 (0%)	5 (23%)	9 (41%)	6 (27%)	15 (68%)
Asia	42	40 (95%)	2 (5%)	2 (5%)	26 (62%)	10 (24%)	36 (86%)
Eastern, Southern, and South-eastern Asia	26	24 (92%)	1 (4%)	0 (0%)	17 (65%)	6 (23%)	23 (88%)
Eastern Asia	6	5 (83%)	0 (0%)	0 (0%)	4 (67%)	1 (17%)	5 (83%)
Southern Asia	9	9 (100%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
South-eastern Asia	11	10 (91%)	1 (9%)	0 (0%)	7 (64%)	2 (18%)	9 (82%)
Western Asia	16	16 (100%)	1 (6%)	2 (13%)	9 (56%)	4 (25%)	13 (81%)
Oceania	20	10 (50%)	1 (5%)	1 (5%)	6 (30%)	3 (15%)	9 (45%)
Official MDG Groupings							
Landlocked developing countries (LLDCs)	31	26 (84%)	0 (0%)	0 (0%)	14 (45%)	12 (39%)	26 (84%)
Small island developing States (SIDS)	51	30 (59%)	1 (2%)	8 (16%)	17 (33%)	10 (20%)	27 (53%)
Least developed countries (LDCs)	50	33 (66%)	1 (2%)	4 (8%)	18 (36%)	12 (24%)	30 (60%)

Source: own compilation

Referring to Table 31 "Timeliness by UN Methodology for All Countries, Areas and Territories by Most Recent Reference Year", we observe results similar to the above assessment, where we consider only the *UN Member States*. the overall result is **20%p** (percentage points =: %p) **up** For the now 224 countries with our *improved methodology* regarding the percentages of countries with not more than one year *lag* (cf. Table 26, p. 180). Considering only the *32 additional countries*, the percentage went up from 22% to 64%, i.e. **up** by **46%p**. This is because only one third of the *32 additional countries* actually reported new ANA data. Most of the additional countries belong to the *developing* region. Thus, the cumulative of countries with data for the reference year t-1 and t-2 by the *UN methodology* indicate a low percentage for this region of only 64% of the countries. Again, this is almost **90%** by *our methodology*.

Of the *MDG country groups*, particularly the *LDCs* and *SIDS* groups now reveal a much better *timeliness* of their member countries. This is because only 2/3 of these countries, and even less in case of the SIDS, reported new ANA data; however, nearly **90%** of them reported ANA data with a *timeliness* T(c) ≤ 2 years.

Table 32 Timeliness by UN Methodology for Membership in Other Organizations/ Groups by Most Recent Reference Year

Reporting round for 2008 fiscal year data (Assessment of 2008 Yearbook data as of September 2009) by UN assessment methodology	Total number of countries for region	Of total: Replies for reporting round for 2008 fiscal year data	Of replies: Most recent fiscal year available [where 2005 ≡ T(c) = 4, …, 2008 ≡ T(c) = 1]				
			2005 $\lvert C_4^T \rvert$ (h_4)	2006 $\lvert C_3^T \rvert$ (h_3)	2007 $\lvert C_2^T \rvert$ (h_2)	2008 $\lvert C_1^T \rvert$ (h_1)	Total 2007+2008 $\sum_{n=1}^{2} \lvert C_n^T \rvert$ (h_1+h_2)
Other Organizations/Economic Groupings							
CARICOM	20	11 (55%)	0 (0%)	4 (20%)	10 (50%)	2 (10%)	12 (60%)
OECD	30	30 (100%)	0 (0%)	0 (0%)	7 (23%)	23 (77%)	30 (100%)
SADC	15	14 (93%)	0 (0%)	1 (7%)	7 (47%)	6 (40%)	13 (87%)
EU27	27	27 (100%)	0 (0%)	0 (0%)	8 (30%)	19 (70%)	27 (100%)
G20	19	19 (100%)	0 (0%)	1 (5%)	10 (53%)	8 (42%)	18 (95%)
Advanced other than G-20 (excl. Taiwan Prov. of China	23	23 (100%)	0 (0%)	0 (0%)	5 (22%)	18 (78%)	23 (100%)
UN Regions							
ECA	52	34 (65%)	1 (2%)	5 (10%)	17 (33%)	15 (29%)	32 (62%)
Northern Africa	5	3 (60%)	0 (0%)	1 (20%)	3 (60%)	0 (0%)	3 (60%)
Sub-Saharan Africa	47	31 (66%)	1 (2%)	4 (9%)	14 (30%)	15 (32%)	29 (62%)
ECE	56	52 (93%)	0 (0%)	3 (5%)	21 (38%)	29 (52%)	50 (89%)
Developed regions	40	39 (98%)	0 (0%)	3 (8%)	17 (43%)	20 (50%)	37 (93%)
of which: South-eastern Europe transition countries	*6*	*6 (100%)*	*0 (0%)*	*1 (17%)*	*5 (83%)*	*0 (0%)*	*5 (83%)*
Commonwealth of Independent States (CIS)	11	9 (82%)	0 (0%)	0 (0%)	2 (18%)	7 (64%)	9 (82%)
Developing regions	5	4 (80%)	0 (0%)	0 (0%)	2 (40%)	2 (40%)	4 (80%)
ESCAP	49	41 (84%)	1 (2%)	1 (2%)	25 (51%)	15 (31%)	40 (82%)
Developed regions	3	3 (100%)	0 (0%)	0 (0%)	2 (67%)	1 (33%)	3 (100%)
Commonwealth of Independent States (CIS)	8	6 (75%)	0 (0%)	0 (0%)	2 (25%)	4 (50%)	6 (75%)
Developing regions	38	32 (84%)	1 (3%)	1 (3%)	21 (55%)	10 (26%)	31 (82%)
Asia	26	24 (92%)	1 (4%)	0 (0%)	16 (62%)	7 (27%)	23 (88%)
Eastern Asia	4	3 (75%)	0 (0%)	0 (0%)	2 (50%)	1 (25%)	3 (75%)
Southern Asia	9	9 (100%)	0 (0%)	0 (0%)	6 (67%)	3 (33%)	9 (100%)
South-eastern Asia	11	10 (91%)	1 (9%)	0 (0%)	7 (64%)	2 (18%)	9 (82%)
Western Asia	2	2 (100%)	0 (0%)	0 (0%)	1 (50%)	1 (50%)	2 (100%)
Oceania	12	8 (67%)	0 (0%)	1 (8%)	5 (42%)	3 (25%)	8 (67%)
ESCWA	14	13 (93%)	1 (7%)	3 (21%)	7 (50%)	2 (14%)	9 (64%)
ECLAC	44	38 (86%)	0 (0%)	6 (14%)	19 (43%)	17 (39%)	36 (82%)
Developed regions	10	10 (100%)	0 (0%)	0 (0%)	3 (30%)	7 (70%)	10 (100%)
Developing regions	34	28 (82%)	0 (0%)	6 (18%)	16 (47%)	10 (29%)	26 (76%)
Caribbean	13	9 (69%)	0 (0%)	1 (8%)	7 (54%)	3 (23%)	10 (77%)
Latin America	20	18 (90%)	0 (0%)	5 (25%)	9 (45%)	6 (30%)	15 (75%)
Other (Asia)	1	1 (100%)	0 (0%)	0 (0%)	0 (0%)	1 (100%	1 (100%)

Source: own compilation

Among the *country groups*, *organizations*, and *UN regional commissions* assessed in Table 32, we notice a substantial difference especially for *CARICOM*, if using the *UN methodology*. Only 11 out of the 20 *CARICOM* member countries reported new data, yet the *UN methodology* provides results for 16 countries. Four countries (of the set C_{YB} / C_R) that did not report new ANA are indicated in the above table, showing our results according to the UN methodology, as having the reference year t_{RY} = 2006 as most recent reference year, further one country indicates the reference year t_{RY} = 2007. The cumulative percentage of countries with not more than one year *lag* is indicated as 60%, while it is actually **100%** of the set C_R of countries that reported new data, as revealed by our *improved assessment* (cf. Table 27, p. 183).

This repeats similarly for the *UN regional commission ECLAC*, for which 13 of its 44 member countries also belong to the *CARICOM* region. The *largest difference* can be observed for the *ECA* where the result by the *UN methodology* is 62% regarding the percentage of countries with not more than one year *lag* (i.e. $\sum_{n=1}^{2} h_{YB,n} = 0{,}62$). This result by the *UN methodology* is nearly **30%p lower** than by our *improved methodology* (i.e. $\sum_{n=1}^{2} h_{n} = 0{,}91$).

Table 33 Comparison of Timeliness by UN Methodology ($T_{YB}(c)$) and Improved Methodology ($T(c)$)

Results by region/ country group for $T_{YB}(c) \leq 2$ and $T(c) \leq 2$, i.e. where the most recent reference year refers to the year $t_{RY} \geq t-2$, and the corresponding percentages $\sum_{n=1}^{2} h_{YB,n}$ x 100 (based on the regions total of countries), respectively $\sum_{n=1}^{2} h_{n}$ x 100 (based on the number of countries in a region that reported new data).

Region / Country group \ **Methodology**	**UN ($T_{YB}(c)$)**	**Our Improved ($T(c)$)**
	Count (percentage)	Count (percentage)
UN Member States	149 (78%)	145 (92%)
All (M49) Countries	156 (70%)	152 (90%)
Landlocked Developing Countries (LLDCs)	26 (84%)	26 (100%)
Small Island Developing States (SIDS)	27 (53%)	26 (87%)
Least Developed Countries (LDCs)	30 (60%)	29 (88%)
CARICOM	12 (60%)	11 (100%)
ECA	32 (62%)	31 (91%)
ESAP	40 (82%)	39 (95%)
ECLAC	36 (82%)	35 (92%)

Source: own compilation. (Both assessments based on 2008 YB, published in 2009.)

Table 33 depicts our findings for a selection of assessed regions and country groups. Our *improved methodology* reveals that the actual *timeliness* of publication is much *better* than found by the *UN assessment methodology*. For example we find by our approach (cf. column "Improved Methodology") for the group of UN Member States, that *92%* of the countries c $\in C_R$ that report new ANA data have as *timeliness* of publication T(c) $\leq$ 2 years, i.e. the most recent reference year refers to the year $t_{RY} \geq$ t-2. Contrary, when we calculate the results by the UN approach, we find that apparently *only 74%* of the countries assembling the set C_{UN} of UN Member States have ANA data for the reference year $t_{RY} \geq$ t-2. Here the results of the *UN approach* reflect both, the calculation of the percentages relative to the total number of countries that assemble a region (i.e. cardinality $|C|$), rather than calculated relative to the cardinality of the set C_R of those countries that have reported new data, and the consideration of the countries c $\in C_K$ with "known" data (four additional countries).

1.4 Conclusion

The *UN assessment methodology*, assessing "timeliness of reporting" (UNSD 2008, 7), is *improved* mainly by excluding those countries that last reported ANA data for a previously published YB, i.e. the elimination of the set C_K of countries with "known" data. The *timeliness* of ANA data publication at the international level through the UN needs to be determined based on that *publication year* t_{PY}, which made these data available in the YB for the *first time*. Therefore the set $C_K = C_{YB} / C_R$ of countries that last reported new ANA data for a previous YB needs to be excluded from the total set C_{YB} of countries in the present YB. Consequently, only those countries c $\in C_R \subseteq C_{YB}$ that *reported new data* are considered. Moreover, a country c $\in C_{YB} / C_R$ that is currently included in the YB, without having reported new data, might have ceased to produce NA data. [Regarding the definition of the country sets C_K, C_R, and C_{YB} refer to chapter 1.1, p. 130 ff. of this study.]

Our *improved methodology* truly reflects the *timeliness* T(c) of the country's ANA data, when published for the *first time* at the UN. Contrary, the later *republication* of these ANA data in any subsequent published YB does not express the *timeliness* of publication (or production or reporting of ANA data), but describes how "*up to date*" the ANA data are (cf. chapter 1.2.1, p. 137ff. on the deficiency of the current UN assessment methodology). Thus, in contrast to our timeliness assessment, the *UN* *methodology* actually assesses for any present YB how *up-to-date* the YB issue is concerning the included ANA data. This deficiency results from considering countries for the UN timeliness assessment (UNSD 2008), regardless whether or not the published country data are actually already “known”.

Since we based our results for *timeliness* solely on the set C_R of countries that indeed report *new data*, we also use this set of countries for the calculation of the true *share* h_n (cf. Formula II.115, p. 120) of the set C^T_n (cf. Formula II.111, p. 120) of countries with a timeliness of value $n \in \boldsymbol{N}$ in years. Contrary in the UN study, the presented *percentages*, for the set $C^T_{YB,n}$ (cf. Formula 1.2, p. 139) of countries with a particular timeliness of value $n \in \boldsymbol{N}$ in years, are based on the total number of countries (i.e. the cardinality $|C|$) *assembling* the corresponding region or country group. Thus, the UN study even *includes* countries in the basic set, which ceased to or never produced ANA data. Consequently the *percentages* presented by the UN study are too low and do not add up to 100%, for those assessed regions or country groups, that includes countries $c \in C / C_{YB}$ which currently do not have ANA data available.

In *our timeliness assessment*, we calculate *all occurring years* $n \in \boldsymbol{N}$ of *timeliness* T(c) (by individual years and the cumulative of the reference years t-1 and t-2) (cf. chapter 1.2.2). Opposite, the *UN study* only calculates results for the *two most recent reference years* and the cumulative of these two (i.e. $t_{RY} = t-1$, $t_{RY} = t-2$, and $t_{RY} \geq t-2$). Thus, the UN study only presents results for the countries $c \in C_{YB2} \subseteq C_{YB}$ published in the YB, for which the most current available reference year t_{RY} of ANA data is at most *two years back* from the publication year t_{PY} of the present YB. Consequently not all *occurring years* $n \in \boldsymbol{N}$ of timeliness are covered for the set C_{YB} of countries included in the YB. Neither is the timeliness of *all countries* $c \in C_R \subseteq C_{YB}$ that reported a *new* reference year t_{RY} of ANA data assessed, nor does the UN study consider the countries $c \in C_K / C_{K2}$ with “known” ANA data, for which the most current available reference year t_{RY} of ANA data refers to a reference year earlier than *two years* prior to the publication year t_{PY} of the YB (i.e. $t_{RY} \leq t-3$). [Regarding the definition of the country sets C_{K2} and C_{YB2} refer to chapter 1.1, p. 130 ff. of this study.]

Based on the UN *timeliness assessment*, we delineate a definition and a formula for the UN’s *timeliness measure* which we label $T_{YB}(c)$ (cf. DEF. 34, p.138). It is calculated by the most recent reference year t_{RY} within the present set of ANA data for a country $c \in C_{YB2}$ published by the present YB with the available ANA data referring to either of the two most recent reference years (i.e. $t_{RY} \geq$ t-2), in relation to the publication year t_{PY} of the present YB.

In order to assess *all occurring years* $n \in \boldsymbol{N}$ of timeliness $T_{YB}(c)$ for the set C_{YB} of countries included in the YB, we merely need to change the considered set of countries in the formula for the UN timeliness measure $T_{YB}(c)$ to the set C_{YB} of countries in the YB (for which the latest available reference year t_{RY} is at most five years back from the YB publication year t_{PY}).

Specifically, for the set C_1^T of countries with a *timeliness* of one year, i.e. that produced t-1 data, we know that they have a strong ANA data *production* capacity, as these data are *produced* within the first nine month of the new reference year. On the other hand, especially for the sets C_n^T with $n > 2$, of countries with a *timeliness* of ANA publication of more than two year, we know that these have a poor ANA data *production* capacity. In the latter case the countries lack the statistical capacity to *produce* ANA data even within 2 years after the end of the reference year. Thus the results of our study provide sufficient detail and information. Thus the *timeliness assessment* based on internationally available ANA data in the YB of the UN provides the needed detail for the assessment of the ANA data *production* capacity.

The problem of considering the set C_K of countries with "known" data is particularly obvious in case of ANA data published with a *time lag* of more than one year, i.e. where the *timeliness* $T_{YB}(c) \geq 2$ years. Since our *improved approach* provides the actual *timeliness* of ANA data publication for a country $c \in C_R$ that reported *new* ANA data, there is no uncertainly regarding the number of countries (cardinality $\left|C_n^T\right|$) that have a particular *timeliness in years* $n \in \boldsymbol{N}$. Since the *UN methodology* considers also those countries $c \in C_K = C_{YB} / C_R$ with "known" data already published by a previous YB, it is not certain how many of the countries (cardinality $\left|C_{YB,n}^T\right|$) indeed have a *particular timeliness* (other than the countries with data for t-1, as the data for these countries are definitely published for the first time).

The *timeliness* T(c) of ANA data needs improvement for all countries $c \in C_R \subseteq C_{YB}$ for which the ANA are currently not published for the reference year t_{RY} = t-1. For these countries, the NA production capacity needs to be strengthened to produce ANA data more *timely*. Further, the NA production capacity of those countries c $\in C_{YB} / C_R$ that produce, or at least produced, ANA data in the recent past, but did not report new data, i.e. that do not report data annually or have ceased to produce data, need to be improved. Those countries c $\in C / C_{YB}$ not included in the YB and thus considered countries that do not produce ANA data (cf. chapter 1.1, footnote 59, p. 131), need to develop a basic NA production capacity.

Furthermore, regarding the set $C_K = C_{YB} / C_R$ of *countries* included in the YB, yet not having reported new data, we could assess the *timeliness* of ANA publication of a country c $\in C_K$, based on the publication year $t_{PYK}(c)$ when these data were published internationally for the first time in a YB (cf. chapter 1.2.5.2). This approach would provide information regarding the (former) timeliness of ANA data publication at the last time point of new ANA data reporting by the country. Still, a country that did not report new ANA data might not be able to provide ANA data annually or might have ceased to produce ANA.

Moreover, we could assess the *timeliness* of ANA publication of different *SNA concepts* (like current and constant price data, GDP by expenditure, or production, or the integrated accounts), i.e. by individual *NAQ tables* or even for individual *NA indicators* (like GDP, GNI, GNDI, Saving or NL) representing specific *SNA concepts* (cf. chapter 1.2.5.3). However, the ANA data are currently reported and subsequently published once a year. Therefore, all ANA data produced since the previously reported data set are reported together. Thus, differences could only be detected if the ANA data of individual *SNA concepts* are computed with such different *timeliness* that different reference years are published in the YB.

2 Availability Study

In this chapter we analyze and improve the current UN assessment methodology for NA data availability of the NAQ tables of the present valid 1993 MRDS (cf. DEF. 25, p. 59). Subsequently, we suggest a corresponding methodology to assess the ANA data availability of the 2008 MRDS (cf. DEF. 26, p. 64). We make empirical studies of the available ANA data in countries, based on the reported official ANA statistics. We use our improved methodology for our assessment of the 1993 MRDS. Our study assesses the availability of ANA data of the new 2008 MRDS for the first time. Together with the 1993 and 2008 MRDS, we assess the NAQ table availability of the milestone levels 1 and 2 according to their "relaxed" criteria (cf. DEF. 30, p. 105, and DEF. 31, p. 109), respectively, according to their "strict" criteria (cf. DEF. 28, p. 98 and DEF. 29, p. 101).

2.1 Introduction

Assessments of ANA data availability for the 1993 MRDS have been carried out by the UN since 2001.[90] The UN assessments are prepared at the request of the Statistical Commission (United Nations 2001a). These assessments focused on the major changes in ANA data availability compared to the previous year and were done for the NAQ of tables of the 1993 MRDS. Among the previous assessments, a comparably more comprehensive study by the UN was last published in the SNA News and Notes, Issue 25/26, May 2008 (UNSD 2008), following a significant increase of ANA data availability.

It is important to note, that the data availability at the international level, i.e. the amount of data released by an IO is not necessarily the same as the availability of data at the national level. We consider the availability of ANA data at the international level based on the publication of national statistics by the UN in the annual YB.

Our improved assessment serves research at the national and international level as a source for information about the ANA data availability. It serves IOs to identify country groupings by 1) regions, 2) regional UN commissions, and 3) membership in

[90] The results for the first two assessments can be found in the 2004 report of the ISWGNA (cf. United Nations 2004, 15 ff.).

other organizations that are in need of NA capacity building. The data gap can be identified with the availability assessment as the difference between the total number of countries of a group and the number (and percentage) of countries that have the assessed data set available.

The UN has reported significant improvements in ANA data availability. More than 90% of the UN Member States report data, above 40% of the countries report the complete 1993 MRDS, and little less than 60% report the substantial part (i.e. at least all but one of the required NAQ tables) of the 1993 MRDS (cf. UNSD 2008 and United Nations 2011, 2012, 2013).

Yet, is ANA data availability indeed as good as suggested by these statements? The UN assessments are not based on the extended set of required NAQ tables of the 2008 MRDS, which is the new benchmark of the minimum set of ANA data that NSOs should produce.

The following map shows the depth of ANA data reporting by individual countries regarding the NAQ tables required to fulfill the 1993 and 2008 MRDS criteria.

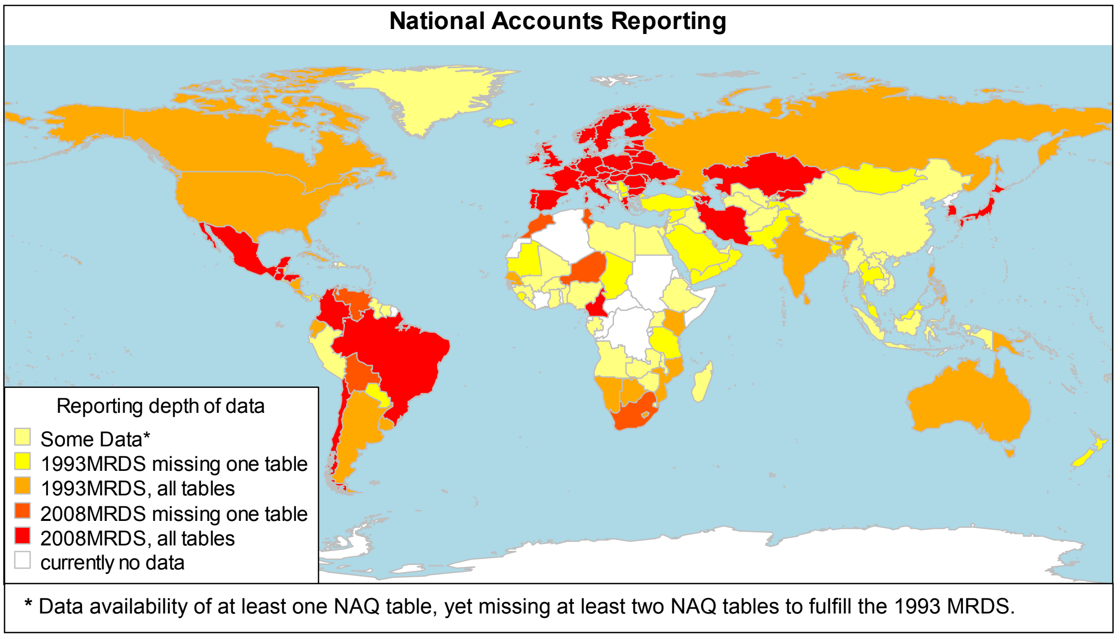


Figure 27 World Map of Annual National Accounts Data Availability

Source: own compilation. (Based on 2008 YB and 2009 YB)

The world map (Figure 27, above) shows that the most detailed data are available in Europe, for the Commonwealth of Independent States and in America. North America, i.e. Canada and the US, interestingly show a reporting depth of the 1993 MRDS only, rather than for the 2008 MRDS. Similarly Australia and the Russian Federation, two major countries in the world and also big advanced economies like Argentina and India fall short of the 2008 MRDS benchmark. This surprising outcome results from data gaps in the reporting of NAQ tables 4.1, …, 4.7, and 4.9, i.e. the "institutional sectors accounts", until NL of the capital account. Particularly the required data for NAQ tables 4.3, … 4.7 the individual "domestic sectors" (other than NAQ table 4.5 "government sector") and NAQ table 4.2 "external sector" are often missing. Regions that show the most significant data gaps (i.e. countries that only report some basic economic statistics or even no data) are foremost African and Asian countries (with some exceptions where even well advanced data are available) and the countries in the Caribbean and Oceania regions.

As main finding from the assessment, our study reveals much lower actual data availability than indicated by previous assessments. We provide updated and refined results for the 1993 MRDS and first time results for the 2008 MRDS.

2.2 Improvement of 1993 MRDS Assessment and Design of 2008 MRDS Assessment

In this subchapter the UN methodology with its deficiencies and our suggested improvements are explained, and our suggested methodology for the 2008 MRDS assessment.

The utilized source data for the 1993 MRDS assessment is the data published in the 2008 YB. The 2008 YB was released in September 2009 (first as PDF publication, also as print publication in 2010); the statistics were also released online via the data portal of the United Nations "UNdata" (UNSD 2013c)[91]. The assessment for the institutional sectors (NAQ tables 4.y, with y = 1, …, 9) for the 2008 MRDS is based on the 2009 YB, which was published in 2010 as PDF publication (United Nations 2010a) and the data were released via UNdata in November 2010[92].

The countries used the SNA standards, 1993 SNA and 1968 SNA, to produce the ANA statistics. All NAQ tables required by the 2008 MRDS can be compiled using any of the SNA standards (though small differences exist regarding individual NA

[91] Cf. http://data.un.org/.

[92] The most recent release is as of 2 November 2011 for data of the 2010 Yearbook.

indicators). Thus the presented assessments are not assessment of the implementation of a specific SNA version by NAQ tables of the MRDS, but are assessment of the capacity of countries regarding the production of specific NAQ tables (i.e. the corresponding SNA concepts). Thus, the assessment indeed shows the available of specific sets of ANA data by NAQ tables.

2.2.1 **Time Period of Assessment**

The following subsections explain the time frame considered by the current assessment and the revised time frames. The suggested improvements ensure a more current time frame for the data that is considered for the assessment. Further, two individual time periods are suggested for the NAQ tables of the 1993 MRDS and the institutional sectors (NAQ tables 4.y with y = 1, …, 9) assessed with the 2008 MRDS.

Time Period of the UN Approach

The availability of data is evaluated based on the data reported during the last five NAQ reporting rounds. Data for the individual NAQ tables of the 1993 MRDS needed to be present for at least one year in one of the NAQs that were reported over the last five NAQ reporting rounds. This part of the methodology is also explained in United Nations (2004, 15ff. and 2011, 21ff.).

Any time lag of the reported data is irrelevant for the assessment. Each NAQ table of the NAQ covers 12 reference years of data. The reported ANA data can refer to any of the 12 reference years covered by the NAQ in order to be considered for the assessment. Thus the current methodology covers 17 reference years for which data could have been reported.

In some cases the data reported by the NAQ might merely provide updates to previously reported data. The NAQ might also be used to report ANA data for back periods. This might happen e.g. in case of failure to report the requested historic periods earlier. In either case data are considered for the availability assessment.

The advantage of this methodology is that it captures the data reported by a country during any of the five recent reporting rounds. Thus, even if a country was unable to report new data during the past four reporting rounds, and additionally produces ANA data with substantial time-lag of three or four years, these data would still be captured.

Obviously, such approach might lead to misleading results about the actual availability of ANA data for current years.

The UN methodology does not allow executing the assessment directly based on the available ANA statistics in the database used by the UN to store the received country statistics. Moreover, the UN methodology requires information which cannot be obtained from the database. The information to each country about the time points of reporting over the last five NAQ reporting rounds and the specific NAQ tables that were reported with each of the NAQs need to be maintained and analyzed. The analysis regarding NAQ tables for which data were reported cannot be limited to the latest reported NAQ of a country, since the coverage of data may differ between reported NAQs.

Revised Time Period for the 1993 and 2008 MRDS

The observation period during which data need to be present for at least one reference year is changed to a fixed time frame.

The revised observation period (TO) for the availability assessment of the NAQ tables of the 1993 MRDS is set to the past five reference years of ANA data, regarding the current year t:

$$TO_{1993} = \{t-5, \ldots, t-1\}. \tag{2.1}$$

Example: $TO_{1993} = \{2004, \ldots, 2008\}$ for ANA data published by the UN in t = 2009.

The suggested observation period of the institutional sector accounts until NL, i.e. NAQ tables 4.1, … , 4.9, assessed with the 2008 MRDS is set to the past ten reference years starting form t:

$$TO_{2008} = \{t-10, \ldots, t-1\}. \tag{2.2}$$

Example: $TO_{2008} = \{2000, \ldots, 2009\}$ for the 2009 YB, and data published in t = 2010.

Using a fixed time period does not require additional information about the actual set of data reported by countries during each of the last five NAQ reporting rounds. The major advantage of the new approach is the possibility to change the data source to obtain all the information needed to assess the data availability. Selecting a fixed time frame, the assessment can be executed directly from the available NA statistics in the

UN database. This procedure could also be used to easily execute assessments for the data availability of different past periods. The results of our 1993 MRDS assessment for the individual countries are run directly on the UN database (data availability as per the 2008 YB). The availability of data of NAQ tables 4.1, …, 4.9 until NL and the corresponding 2008 MRDS assessment are based on country data retrieved from the UN data portal "UNdata" (data availability as per the 2009 YB).

Using the past five reference years as observation period, places the focuses on the actual availability of current data. This observation period is chosen for the 1993 MRDS, since it assesses the availability of basic economic data including total economy aggregates in case of the NAQ tables 1.3, 2.3, 4.1, and "exports and imports" in NAQ table 4.2. Countries usually produce these ANA data within one year after the reference period.

The five year period also captures those cases, where the production process for the data of the 1993 MRDS takes several years. Even data that are reported with a time lag of four years after the reference period are captured. As our assessment of timeliness shows (see part III, chapter 1.3), there are not any UN Member States that reported data with a lag of more than three years (other than one country that reported data that were previously missed out). The short period is also sufficient to capture countries that missed the most recent reporting round, e.g. where countries do not have the capacity to produce NA annually but only every other year. Five years seems to be optimal, since it captures the data if the latest available year has a substantial lag, while excluding those cases where data were not reported during the recent NAQ reporting rounds. Hence, the five year period correctly indicate a data gap in the latter case.

A much longer time period of ten years is chosen for NAQ tables 4.1, …, 4.9. The 2008 MRDS criteria require the production regarding NAQ tables 4.1, …, 4.7, and 4.9 until the net lending indicator of the capital account (NL). We apply this requirement also to NAQ table 4.8 which we consider in our 2008 MRDS assessment. The reason for the longer time period is that even in developed economies some of the institutional sector accounts (NAQ tables 4.y with y = 1, …, 9) only become available 36 month after the reference period. Thus, developing countries, particularly if they need more than one year to produce basic economic data, likely need substantially more time to produce these detailed statistics. Adding three years to the five year observation period seems to be the minimum for NAQ tables 4.1, …, 4.9 until NL. Countries, which still have to build the statistical capacity to produce these detailed data, might encounter additional delays. A shortcoming of this long time period is that it might merely capture old country data. This can happen, if a country discontinuous the production of NAQ tables 4.1, …, 4.9 until NL or ceased to report data for a long time (e.g. in case of adverse political circumstances).

Based on the above explanations, our chosen time periods seem to be reasonable for the assessment of the availability of the 1993 and 2008 MRDS. It should be noted here, that until 2004 the methodology used by the UN to assess the 1993 MRDS and the milestone levels considered a six year period, starting with the reference year t-2, for which at least one year of data needed to be available, cf. United Nations (2000) for the milestone assessment and United Nations (2001a) for the 1993 MRDS assessment.

2.2.2 **Availability Conditions of NAQ Tables**

In the following we explain the deficiency of the currently used condition to determine whether or not a NAQ table is counted as available under the UN methodology. Then we summarize the suggested conditions for our study, which we defined and explained in part II, chapter 3.2.

Deficiency of the Current NAQ Table Availability Condition

Currently, in the data availability assessments of the UN, a NAQ table of the *1993 MRDS* is *counted as available* if any NA indicator of this NAQ table is available[93]. This condition is sufficient in case of the basic ANA data sets covered by the 1993 MRDS (cf. DEF. 25, p. 59).

Data reported for any of the approaches to compute GDP (NAQ tables 1.1, 1.2, 2.1, and 2.2) would typically be available with full detail of the components. It might be that components are reported in aggregated form, e.g. the data for several industries are aggregated and reported as one figure (referring to NAQ tables 2.1, 2.2, and 2.3).

[93] In case of the availability assessment in 2004 “a set of most important items was determined. If at least 50 per cent of these critical items had been submitted to the United Nations for a given table for at least one fiscal year in the observation period, then the table was counted as ‘available’.”(cf. United Nations 2004,16).

The NAQ table 2.3 covers “output”, “gross value added” and “operating surplus”. Thus, NAQ table 2.3 covers NA indicators of the “production” side and of the “income” side. The current method does not allow concluding whether the indicators of the one or the other side are available. Usually the whole NAQ table is available, however it might be in exceptional cases that a country only prepares one of these data sets. Only the availability of NA data for the “total economy” is assessed, not by individual industries. Further covered are the NA indicators for “gross capital formation” and “employment” by industry. These NA indicators are isolated from the other NA indicators of this NAQ table; hence their availability cannot be concluded.

In case of NAQ tables 1.3, 4.1 and 4.2 it can only be concluded that NA data are available until the first balance for the sequence of balances covered by these NAQ tables. The UN methodology is inadequate for the assessment of the availability of the NAQ tables 1.3, 4.1, and 4.2 until “net lending” of the capital account (NL). These NAQ tables start with basic data (“GDP”, respectively “output” and “intermediate consumption” in case of NAQ tables 3.1 respectively 4.1, and “export” and “import” components of the “external balance of goods and services” in case of NAQ table 4.2). Yet, they also cover the more detailed NA indicators until NL of the capital account, that are not available in many countries. Consequently, the mere availability of any data for these NAQ tables only indicates that at least the basic data are available. Therefore, the UN concludes regarding NAQ tables 4.1 and 4.2 that (“at least”)[94] “parts of the integrated economic accounts for the total economy and the rest of the world” (UNDS 2008, 9) are available. Given NAQ table 1.3, we can specify the basic data availability to “at least the data until GNI”, because this is the first balance in the sequence calculated based on GDP. Otherwise no data is reported for this NAQ table.

Suggested Availability Conditions of NAQ Tables

In the following, we summarize the availability conditions given in part II, chapter 3.2, where we formalize and explain them with our definition of the availably measures.

For the NAQ tables of the 1993 MRDS we use the current UN methodology. Thus any of the NA indicators covered by a specific NAQ table needs to be available for at

[94] Added in United Nations (2011, 22).

least one reference year of the observed time period (TO_{1993}). With other words, the cardinality of reported data of the NAQ table needs to be greater than zero.

Using the improved time period and the current availability condition, gives evidence that in one case (Aruba), the data for GDP are available for NAQ table 1.2 “GDP by expenditure in constant prices”, yet not the components (cf. United Nations 2009b).

The 2008 MRDS, includes the institutional sector accounts (NAQ tables 4.1, …, 4.9). The NAQ table is required until NL to count as available. The existence of the NL indicator ensures the reporting of the complete institutional sector accounts starting from GDP[95]. Therefore, the existence of the NL indicator is the condition in our assessment to count NAQ table 4.1, …, 4.9 as available. The NL indicator needs to be existent for at least one reference year of our observed time period (TO_{2008}).

In addition to the NAQ tables of the 2008 MRDS, we assess the availability of the “financial account” of NAQ table 4.2. We measure this by the availability of the “financial derivatives”. This NA indicator is chosen over the “net lending” indicator of the financial account, since it also implies the existence of the detailed financial indicators. Contrary, the main aggregate “net lending” of the financial account might be published without the detailed financial indicators. The “financial derivatives” indicator of NAQ table 4.2 needs to be existent for at least one reference year of our observed time period (TO_{2008}) to count the financial account as available.

2.2.3 Clustering of Country Groups

This section briefly describes the deficiencies of the current subdivisions of country groups of the UN assessments and our improved subdivisions. The clustering of the country groups is exactly the same as for the assessment of timeliness, where they are described in detail (cf. 1.2.3 Clustering of Country Groups, p. 148ff.).

Deficiency of Currently Composed Country Groups

The currently considered country groups provide only limited information to facilitate detailed analysis about the data availability for specific regions, smaller country

[95] GDP is also covered by the production account. Since the net lending indicator is based on the GDP of the income account, only the availability of data starting with this account can be ensured. Yet, even if a country doesn’t report the production account, the institutional sector accounts count as available.

clusters and for the member countries of organizations (e.g. OECD). Information for more detailed subdivisions of country groups would be useful for IOs to provide tailored technical assistance and capacity building aid to those country groups.

Improved Subdivisions of Country Groups

We provide further subdivisions by regions of interest and smaller subsets of these regions. We also cover country groups that are of particular importance or frequently monitored (e.g. LDCs). The assessment by regional and international organizations includes the assessment of UN Member States by UN Regional Commissions.

Moreover, we calculate results for our suggested subdivisions by regional first for the set of countries that comprises only of the UN Member States; then again for the set of countries that includes also the non-UN members for which ANA data are requested (32 additional countries).

Refer to Figure 28 at the end of the following section for an example of the presentation of the improved subdivision of regional aggregates in the results tables. We present for each assessed country group the absolute number and percentage of countries that have data available for the assessed NAQ table or availability measure.

2.2.4 Assessed NAQ Tables, MRDS Measures, and Milestone Levels

This section describes for which NAQ tables data availability is measured. We list first the NAQ tables and availability measures (like the 1993 MRDS) considered by the UN, then those of our availability assessment. The references in brackets refer to the chapters in part II, where we define the availability measure and describe the measurement of the country set that fulfills it.

Current NAQ Tables, 1993 MRDS, and “Relaxed” Milestone Levels

The current 1993 MRDS assessments by the UN cover the following individual NAQ tables and sets of NAQ tables (compare UNSD 2008):

- NAQ tables 1.1, 1.2, 1.3, 2.2 and 2.3, individually (cf. chapter 3.2.3, p. 85),
- NAQ table 1.3 or 4.1, i.e. substitutional (cf. chapter 3.2.2, p. 76),
- NAQ table 4.2, (cf. chapter 3.2.3, p. 85),

- 1993 MRDS (cf. chapter 3.2.1.1, p. 59),
- Partial 1993 MRDS (i.e. availability of at least all but one of the NAQ tables required to fulfill the 1993 MRDS criteria) (cf. 3.2.1.3, p. 69),
- Milestone level 1 and 2 ("relaxed" criteria) (cf. 3.2.5.2, p. 105).

The UN study provides only the absolute numbers of countries that have data available for the assessed availability measure. Thus, not included are the percentage or shares of the set of countries that has the data for the assessed availability measure, in relation to the total number of countries corresponding to each region. The only exceptions are the milestone levels which include percentages. Providing the percentages to each assessed region and NAQ table is a matter of user friendliness of the data presentation.

The 1993 MRDS assessment does not cover any of the additional NAQ tables required by the 2008 MRDS.

Suggested NAQ Tables, 2008 MRDS, and "Strict" Milestone Levels

We assess the availability of the individual NAQ tables that need to be available to fulfill the 1993 or 2008 MRDS criteria, the 1993 and 2008 MRDS, and the milestone levels 1 and 2. Since we suggest the re-revision of the criteria to fulfill the milestone measures, we assess their availability according to both the "relaxed" and the "strict" criteria. We provide the absolute number and percentage of countries that have data available for the assessed availability measure for each country group.

The availability measures of the 1993 MRDS remain unchanged, since the needed NAQ tables to fulfill the 1993 MRDS are already covered by the current UN methodology. A breakdown into an individual assessment of NAQ tables 1.3 and 4.1 is not necessary and the other NAQ tables are already assessed individually.

In the 2008 MRDS assessment we cover only the availability measures, which are not already covered in the 1993 MRDS assessment. These are:

- NAQ table 1.3 or 4.1 until NL, substitutional (cf. chapter 3.2.2, p. 76),
- NAQ tables 4.1, …, 4.9 until NL, individually (cf. chapter 3.2.3, p. 85),
- NAQ table 4.2 financial account (cf. chapter 3.2.3, p. 85),

- NAQ table 4.3 and 4.4, or 4.8, substitutional (cf. chapter 3.2.3, p. 85),
- NAQ table 4.6 and 4.7, or 4.9, substitutional (cf. chapter 3.2.3, p. 85),
- 2008 MRDS (cf. chapter 3.2.1.2, p. 64),
- Partial 2008 MRDS (i.e. availability of at least all but one of the NAQ tables required to fulfill the 2008 MRDS criteria) (cf. 3.2.1.3, p. 69),
- Subset of all of the NAQ tables 4.1, …, 4.7, and 4.8 (i.e. the institutional sector accounts) required by the 2008 MRDS criteria (cf. 3.2.4, p. 90),
- Subset of at least all but one (i.e. partial availability) of the NAQ tables 4.1, …, 4.7, and 4.8 required by the 2008 MRDS criteria (cf. 3.2.4, p. 90),
- Milestone level 1 and 2 (“strict” criteria) (cf. 3.2.5.1, p. 98).

The assessment of the availability of individual NAQ tables and sets of NAQ tables allows insight into the NA data production according to different SNA concepts and indicates “bottlenecks”, i.e. data gaps, in the NA data production.

2.2.5 **Description of our Results Tables**

The following paragraphs describe the *format* of the *results table* of our *availability* assessment. The 1993 MRDS and 2008 MRDS assessment are shown in two separate assessments. The tables presenting the results for the 1993 and 2008 MRDS assessment consists of two parts due to the length of the Excel spreadsheets. The general structure of the results tables is identical for both, the 1993 and 2008 MRDS assessment and all considered country groups.

- Part a) shows the individual NAQ tables and substitutional NAQ tables (see Figure 28, p. 210).
- Part b) (see Figure 29, p. 210) shows the results for the data availability of the
    - o 1993 respectively 2008 MRDS, including the partial fulfillment,
    - o milestone levels 1 and 2 (“relaxed” criteria in case of the 1993 MRDS assessment and “strict” criteria in case of the 2008 MRDS assessment),
    - o Subset of all of NAQ tables 4.1, …, 4.7, and 4.8 until NL, including the partial availability, shown in part b) of the 2008 MRDS assessment.

For the description of part a) of the results tables, we refer to Table 34 (p. 214) of our *availability assessment*. This table shows the availability of the individual NAQ tables and substitutional NAQ tables regarding the 1993 MRDS assessment and the country groups composed of the set C_{UN} of *UN Member States*. We continue with part b) of the results table by referring to Table 41 (p. 224) of our 2008 MRDS assessment.

To explain the context and data presented by the different *sections* of the *results table*, we refer to the formulas and symbols defined in part II, chapter 3.2 of our study. Figure 28 and Figure 29 (p. 210) illustrates Table 34 respectively Table 41, with references in brackets [the numbers] used to explain the corresponding *sections*.

Similar to our assessment of timeliness, the columns of the results table begin with the region respectively country group (refer to [1] in the figure below). It indicates the considered set C of countries that we assess regarding the data availability of its represented member countries $c \in C$. (The set C is for example the set of "*UN Member States*", or the region "*Northern Africa*"). The second column [2] is the total number of countries (i.e. the cardinality $|C|$) for each region respectively country group. Columns one and two are included in part a) and b) of the results table.

For "part a)" of the results table, the following columns [3] specify the individual NAQ tables of the 1993 MRDS (cf. part II, chapter 3.2.3) and combinations of NAQ tables (cf. part II, chapter 3.2.2). In "part b)" of the results table, the third and fourth column [4] shows the results for the countries that were able to report all, and all but one, of the required institutional sectors (cf. part II, chapter 3.2.4). The next two columns [5] present the results for the countries that were able to fulfill the MRDS criteria (here regarding 2008 MRDS) and those that reported at least all but one of the required NAQ tables to fulfill the MRDS criteria (cf. part II, chapter 3.2.1). In case of the 2008 MRDS assessment we include a column [6] for the results regarding the availability of the "financial accounts" for NAQ table 4.2 (cf. part II, chapter 3.2.3). The last two columns [7] provide information regarding the number of countries that fulfilled milestone level 1 or higher, and milestone level 2 (cf. part II, chapter 3.2.5). Regarding the latter, we present the results for the "relaxed" criteria with the 1993 MRDS and for the "strict" criteria with the 2008 MRDS.

Data available between 2004-2008	Total number of countries [2] $\lvert C \rvert$	Tables of the 1993 SNA Minimum Requirement Data Set (MRDS) 1.1 $\lvert C_{1.1} \rvert$ ($h_{1.1}$)	1.2 $\lvert C_{1.2} \rvert$ ($h_{1.2}$)	2.1 $\lvert C_{2.1} \rvert$ ($h_{2.1}$)	2.2 [3] $\lvert C_{2.2} \rvert$ ($h_{2.2}$)	2.3 $\lvert C_{2.3} \rvert$ ($h_{2.3}$)	1.3/4.1 $\lvert C_{1.3 \vee 4.1} \rvert$ ($h_{1.3 \vee 4.1}$)	4.2 $\lvert C_{4.2} \rvert$ ($h_{4.2}$)
UN Member States	**192**	**158 (82%)**	**129 (67%)**	**166 (86%)**	**159 (83%)**	**119 (62%)**	**134 (70%)**	**89 (46%)**
Developed Regions [1]	**38**	**37 (97%)**	**36 (95%)**	**36 (95%)**	**35 (92%)**	**36 (95%)**	**37 (97%)**	**32 (84%)**
Transition Countries [8]	**17**	16 (94%)	13 ([illegible])	17 (100%	**16 (94%)**	**13 (76%)**	**12 (71%)**	**11 (65%)**
Transition countries of South-eastern Europe	6	6 (100%)	4 (67%)	6 (100%	6 (100%)	5 (83%)	3 (50%)	2 (33%)
Commonwalth of Independent States (CIS)	11	10 ([illegible]) [9]	9 (82%) [10]	[illegible]	10 (91%)	8 (73%)	9 (82%)	9 (82%)
Developing Regions	**137**	05 (77%)	80 (58%)	113 (82%	**108 (79%)**	**70 (51%)**	**85 (62%)**	**46 (34%)**
Africa	53	32 (60%)	27 (51%)	37 (70%)	34 (64%)	20 (38%)	23 (43%)	12 (23%)
Northern Africa	5	3 (60%)	2 (40%)	4 (80%)	3 (60%)	3 (60%)	4 (80%)	2 (40%)
Sub-Saharan Africa	48	29 (60%)	25 (52%)	33 (69%)	31 (65%)	17 (35%)	19 (40%)	10 (21%)
Latin America and the Caribbean	33	31 (94%)	23 (70%)	31 (94%)	31 (94%)	24 (73%)	26 (79%)	16 (48%)
Caribbean	13	11 (85%)	5 (38%)	11 (85%)	11 (85%)	6 (46%)	8 (62%)	2 (15%)
Latin America	20	20 (100%)	18 (90%)	20 (100%)	20 (100%)	18 (90%)	18 (90%)	14 (70%)
Asia	39	36 (92%)	27 (69%)	36 (92%)	34 (87%)	24 (62%)	29 (74%)	14 (36%)
Eastern, Southern, and South-eastern Asia	24	21 (88%)	18 (75%)	21 (88%)	20 (83%)	10 (42%)	16 (67%)	9 (38%)
Eastern Asia	4	3 (75%)	1 (25%)	3 (75%)	2 (50%)	2 (50%)	3 (75%)	2 (50%)
Southern Asia	9	9 (100%)	8 (89%)	8 (89%)	9 (100%)	5 (56%)	8 (89%)	4 (44%)
South-eastern Asia	11	9 (82%)	9 (82%)	10 (91%)	9 (82%)	3 (27%)	5 (45%)	3 (27%)
Western Asia	15	15 (100%)	9 (60%)	15 (100%)	14 (93%)	14 (93%)	13 (87%)	5 (33%)
Oceania	12	6 (50%)	3 (25%)	9 (75%)	9 (75%)	2 (17%)	7 (58%)	4 (33%)

Figure 28 Structure of the Results Table, Part a) for 1993 MRDS (Example)

Source: own illustration. (Data based on 2008 YB, published in 2009.)

Data available between 2000-2009 for institutional sectors, 2004-2008 for other data	[2] Total number of countries $\lvert C \rvert$	2008 MRDS sectors: All tables $\lvert C_{IS} \rvert$ (h_{IS})	2008 MRDS sectors: All but one tables $\lvert C_{IS(p)} \rvert$ ($h_{IS(p)}$)	2008 MRDS: All tables $\lvert C_{2008} \rvert$ (h_{2008})	2008 MRDS: All but one tables $\lvert C_{2008(p)} \rvert$ ($h_{2008(p)}$)	Fin. acc. of NAQ Table 4.2 $\lvert C^f_{4.2} \rvert$ ($h^f_{4.2}$)	Strict milestone level: 1 or higher [11] $\lvert C_{1s} \rvert$ (h_{1s})	Strict milestone level: 2 $\lvert C_{2s} \rvert$ (h_{2s})
UN Member States	**192**	**47 (24%)**	**56 (29%)**	**45 (23%)**	**52 (27%)**	**43 (22%)**	[illegible]	**40 (21%)**
Developed Regions [1]	**38**	**27 (71%)**	**27 (71%)**	**27 (71%)**	**27 (71%)**	**26 (68%)**	**34 (89%)**	**26 (68%)**
Transition Countries [8]	**17**	[illegible] (47%)	[illegible] (53%)	[illegible] (47%)	[illegible] (47%	1 (6%)	13 (76%)	[illegible] (6%)
Transition countries of South-eastern Europe	6	1 (17%)	2 (33%)	1 (17%)	1 (17%	[6] [illegible]	4 (67%)	0 (0%)
Commonwalth of Independent States (CIS)	11	7 (64%) [4]	7 (64%)	7 (64%) [5]	7 (64%	1 (9%)	9 (82%) [7]	1 (9%)
Developing Regions	**137**	12 (9%)	20 (15%)	10 (7%)	17 (12%	16 (12%)	79 (58%)	13 (9%)
Africa	53	3 (6%)	8 (15%)	1 (2%)	5 (9%)	3 (6%)	27 (51%)	2 (4%)
Northern Africa	5	1 (20%)	3 (60%)	0 (0%)	2 (40%)	3 (60%)	2 (40%)	2 (40%)
Sub-Saharan Africa	48	2 (4%)	5 (10%)	1 (2%)	3 (6%)	0 (0%)	25 (52%)	0 (0%)
Latin America and the Caribbean	33	6 (18%)	7 (21%)	6 (18%)	8 (24%)	6 (18%)	23 (70%)	6 (18%)
Caribbean	13	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (38%)	0 (0%)
Latin America	20	6 (30%)	7 (35%)	6 (30%)	8 (40%)	6 (30%)	18 (90%)	6 (30%)
Asia	39	3 (8%)	5 (13%)	3 (8%)	4 (10%)	7 (18%)	26 (67%)	5 (13%)
Eastern, Southern, and South-eastern Asia	24	2 (8%)	3 (13%)	2 (8%)	2 (8%)	5 (21%)	17 (71%)	4 (17%)
Eastern Asia	4	1 (25%)	2 (50%)	1 (25%)	1 (25%)	2 (50%)	1 (25%)	1 (25%)
Southern Asia	9	1 (11%)	1 (11%)	1 (11%)	1 (11%)	2 (22%)	7 (78%)	2 (22%)
South-eastern Asia	11	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (9%)	9 (82%)	1 (9%)
Western Asia	15	1 (7%)	2 (13%)	1 (7%)	2 (13%)	2 (13%)	9 (60%)	1 (7%)
Oceania	12	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (25%)	0 (0%)

Figure 29 Structure of the Results Table, Part b) for 2008 MRDS (Example)

Source: own illustration. (Data based on 2009 YB, published in 2010.)

In the following we provide two **examples** how to read the information in the *results table* of the assessment of *availability* of ANA data.

Example 1: We consider the data in row "Developed Regions" (refer to [8] in the first *column* [1] regarding the country groups). In the next column we see [2] that a total of 38 countries belong to that region. Referring to the individual NAQ tables assessed in part a) [3], the column for NAQ table "1.2" indicates that 36 developed countries [9] have reported data for this NAQ table, this is 95% of the developed countries [10]. In part b) of the results table, the column “Strict milestone level, 1 or higher” [11] indicates that 34 developed countries (89%) have the data required for milestone 1 according to the “strict” milestone level 1 criteria.

Example 2: Again for the row “Developed Regions” (refer to [8]) in part b) of the results table, we observe that 27 countries, i.e. 71% fulfill the 2008 MRDS criteria [5] (see “All 2008 MRDS tables”). Consequently the set of these countries also fulfills at least the partial MRDS (see “All but one 2008 MRDS tables”) and also the subset of required institutional sectors [4], i.e. of NAQ tables 4.1, …, 4.7, and 4.9 until NL. Data for the “financial account” of NAQ table 4.2 is available for 26 out of the 38 countries in the developed region (68%).

2.2.6 Perspectives for Further Improvement of the Assessment Methodology

Further possible improvements include different compositions of regions, by including for example further subdivisions, or additional organizations. For our suggestions of different subdivisions of country groups refer to chapter 1.2.5.1, p. 158 in our timeliness assessment.

Regarding the assessed availability measures, our availability assessment can be easily extended to assess the availability of the “financial accounts” of all of the institutional sectors, i.e. the other NAQ tables 4.1 and 4.3, …, 4.9.

Regarding the availability condition of the NAQ tables we can refine our conditions further by testing the availability of all main indicators of a given NAQ table. For example for NAQ table 1.1 “GDP by expenditure in current prices” this are all components to calculate GDP. However, additional conditions are not pragmatic and might lead to the unintended result that a NAQ table is not available. For example regarding NAQ table 2.1 “Value added by industry in current prices” (i.e. the production approach), the industries may or may not be reported individually since countries may aggregate single industries.

2.3 Assessment of 1993 MRDS and 2008 MRDS

In this subchapter the results of our availability assessment are presented. First, we present the results of the 1993 MRDS. In the next section follow the results for the 2008 MRDS.

The data availability assessments consist of three parts considering different sets of countries, identical to those of the timeliness assessment (cf. chapter 1.2.3, p. 148). The three parts cover:

i. UN Member States.

ii. All countries, areas, and territories, for which data are requested by the UN, considering 32 additional countries from the MDG country list in addition to the UN Member States. This includes the MDG groups.

iii. Organizations and economic groups. This results table also includes the UN Regional Commissions.

2.3.1 **1993 MRDS Assessment**

This section provides the 1993 MRDS assessment. We start with the UN member countries.

2.3.1.1 UN Member States

Our data availability assessment shows that only 77 out of the 192 UN Member States[96], i.e. 40%, of the countries fulfill the 1993 MRDS criteria. Significantly more countries, yet still only 105 countries, i.e. 55% of the UN Member States, are able to compile at least all but one of the NAQ tables of the 1993 MRDS.

Figure 30 illustrates that another 34% of the UN Member States report at least some NA data (i.e. they report at least one NAQ table, yet missing at least two NAQ tables of the 1993 MRDS criteria). Thus, nearly 90% of the UN Member States report NA data, of which the majority fulfills at least all but one of the 1993 MRDS criteria.

[96] There are now 193 UN members, but at the time of the analyses it was 192.

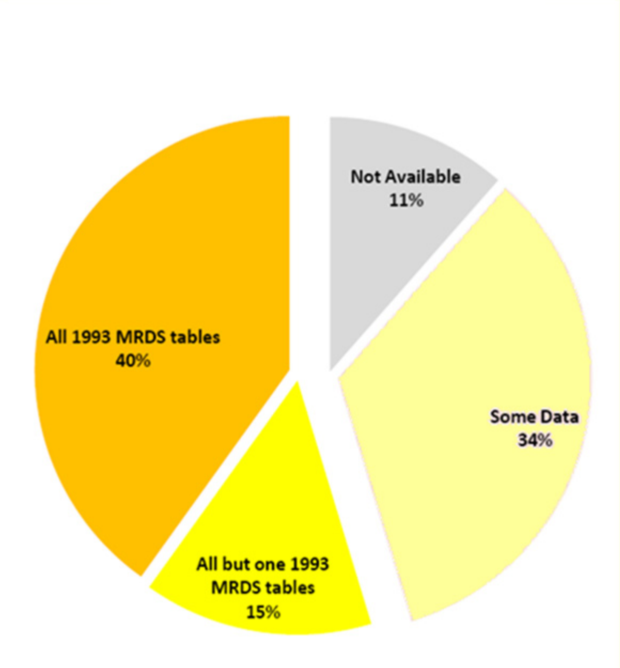


Figure 30 UN Member States Data Availability of 1993 MRDS

Source: own illustration. (Based on 2008 YB.)

The NAQ table that can be compiled by most member states is NAQ table 2.1 "GDP by industry in current prices". These data are available for 86%, i.e. 166 countries (cf. Table 34, p. 214). The availability of corresponding constant price data, i.e. NAQ table 2.2, is not significantly lower, being 83%, i.e. 159 countries. The availability for NAQ table 1.1 "GDP by the expenditure" approach in current prices is still above 80%, precisely 82%, i.e. 158 countries. However, the data in constant prices, i.e. NAQ table 1.2, are only available for 67%, i.e. 129 countries. Also 70%, i.e. 134 UN Member States can provide at least "Gross National Income" (GNI). This can be verified by the fact that they compile data for NAQ tables 1.3 or 4.1 (cf. col. eight "1.3/4.1"). The NAQ table 2.3, which provides "value added" and "income" components corresponding to NAQ table 2.1 for the total economy and individual industries can only be compiled by 62%, i.e. 119 countries. The NAQ table with the least availability is the NAQ table 4.2 "rest of the world" produced by only 46%, i.e. 89 of the 192 UN Member States.

Table 34 1993 MRDS for UN Members, Part a: Availability of Individual Required NAQ Tables

Data available between 2004-2008	Total number of countries	Tables of the 1993 SNA Minimum Requirement Data Set (MRDS)						
		1.1	1.2	2.1	2.2	2.3	1.3/4.1	4.2
	$\lvert C \rvert$	$\lvert C_{1.1} \rvert$ ($h_{1.1}$)	$\lvert C_{1.2} \rvert$ ($h_{1.2}$)	$\lvert C_{2.1} \rvert$ ($h_{2.1}$)	$\lvert C_{2.2} \rvert$ ($h_{2.2}$)	$\lvert C_{2.3} \rvert$ ($h_{2.3}$)	$\lvert C_{1.3\vee 4.1} \rvert$ ($h_{1.3\vee 4.1}$)	$\lvert C_{4.2} \rvert$ ($h_{4.2}$)
UN Member States	**192**	**158 (82%)**	**129 (67%)**	**166 (86%)**	**159 (83%)**	**119 (62%)**	**134 (70%)**	**89 (46%)**
Developed Regions	**38**	**37 (97%)**	**36 (95%)**	**36 (95%)**	**35 (92%)**	**36 (95%)**	**37 (97%)**	**32 (84%)**
Transition Countries	**17**	**16 (94%)**	**13 (76%)**	**17 (100%)**	**16 (94%)**	**13 (76%)**	**12 (71%)**	**11 (65%)**
Transition countries of South-eastern Europe	6	6 (100%)	4 (67%)	6 (100%)	6 (100%)	5 (83%)	3 (50%)	2 (33%)
Commonwalth of Independent States (CIS)	11	10 (91%)	9 (82%)	11 (100%)	10 (91%)	8 (73%)	9 (82%)	9 (82%)
Developing Regions	**137**	**105 (77%)**	**80 (58%)**	**113 (82%)**	**108 (79%)**	**70 (51%)**	**85 (62%)**	**46 (34%)**
Africa	53	32 (60%)	27 (51%)	37 (70%)	34 (64%)	20 (38%)	23 (43%)	12 (23%)
Northern Africa	5	3 (60%)	2 (40%)	4 (80%)	3 (60%)	3 (60%)	4 (80%)	2 (40%)
Sub-Saharan Africa	48	29 (60%)	25 (52%)	33 (69%)	31 (65%)	17 (35%)	19 (40%)	10 (21%)
Latin America and the Caribbean	33	31 (94%)	23 (70%)	31 (94%)	31 (94%)	24 (73%)	26 (79%)	16 (48%)
Caribbean	13	11 (85%)	5 (38%)	11 (85%)	11 (85%)	6 (46%)	8 (62%)	2 (15%)
Latin America	20	20 (100%)	18 (90%)	20 (100%)	20 (100%)	18 (90%)	18 (90%)	14 (70%)
Asia	39	36 (92%)	27 (69%)	36 (92%)	34 (87%)	24 (62%)	29 (74%)	14 (36%)
Eastern, Southern, and South-eastern Asia	24	21 (88%)	18 (75%)	21 (88%)	20 (83%)	10 (42%)	16 (67%)	9 (38%)
Eastern Asia	4	3 (75%)	1 (25%)	3 (75%)	2 (50%)	2 (50%)	3 (75%)	2 (50%)
Southern Asia	9	9 (100%)	8 (89%)	8 (89%)	9 (100%)	5 (56%)	8 (89%)	4 (44%)
South-eastern Asia	11	9 (82%)	9 (82%)	10 (91%)	9 (82%)	3 (27%)	5 (45%)	3 (27%)
Western Asia	15	15 (100%)	9 (60%)	15 (100%)	14 (93%)	14 (93%)	13 (87%)	5 (33%)
Oceania	12	6 (50%)	3 (25%)	9 (75%)	9 (75%)	2 (17%)	7 (58%)	4 (33%)

Legend*

In columns: MRDS tables by table number, with absolute number of countries reporting data and percentage of group total.

In rows: the country groups, with aggregates and breakdowns (breakdowns indented).

*Applicable to all 1993 MRDS availability assessment tables, part a.

Source: own compilation. (Based on 2008 YB.)

Table 35 1993 MRDS for UN Members, Part b: MRDS and Milestone Level Results

Data available between 2004-2008	Total number of countries	1993 MRDS		Milestone level	
		All tables	All but one tables	1 or higher	2
	$\lvert C \rvert$	$\lvert C_{1993} \rvert$ (h_{1993})	$\lvert C_{1993(p)} \rvert$ ($h_{1993(p)}$)	$\lvert C_{1r} \rvert$ (h_{1r})	$\lvert C_{2r} \rvert$ (h_{2r})
UN Member States	**192**	**77 (40%)**	**105 (55%)**	**156 (81%)**	**129 (67%)**
Developed Regions	**38**	**31 (82%)**	**34 (89%)**	**36 (95%)**	**36 (95%)**
Transition Countries	**17**	**9 (53%)**	**12 (71%)**	**16 (94%)**	**12 (71%)**
Transition countries of South-eastern Europe	6	1 (17%)	3 (50%)	6 (100%)	3 (50%)
Commonwalth of Independent States (CIS)	11	8 (73%)	9 (82%)	10 (91%)	9 (82%)
Developing Regions	**137**	**37 (27%)**	**59 (43%)**	**104 (76%)**	**81 (59%)**
Africa	53	11 (21%)	16 (30%)	31 (58%)	21 (40%)
Northern Africa	5	1 (20%)	2 (40%)	3 (60%)	3 (60%)
Sub-Saharan Africa	48	10 (21%)	14 (29%)	28 (58%)	18 (38%)
Latin America and the Caribbean	33	15 (45%)	20 (61%)	31 (94%)	26 (79%)
Caribbean	13	2 (15%)	3 (23%)	11 (85%)	8 (62%)
Latin America	20	13 (65%)	17 (85%)	20 (100%)	18 (90%)
Asia	39	9 (23%)	21 (54%)	36 (92%)	29 (74%)
Eastern, Southern, and South-eastern Asia	24	6 (25%)	12 (50%)	21 (88%)	16 (67%)
Eastern Asia	4	1 (25%)	2 (50%)	3 (75%)	3 (75%)
Southern Asia	9	3 (33%)	6 (67%)	9 (100%)	8 (89%)
South-eastern Asia	11	2 (18%)	4 (36%)	9 (82%)	5 (45%)
Western Asia	15	3 (20%)	9 (60%)	15 (100%)	13 (87%)
Oceania	12	2 (17%)	2 (17%)	6 (50%)	5 (42%)

Legend*

In columns: the MRDS compliance and milestone level results, with absolute number of countries and percentage of group total.

In rows: the country groups, with aggregates and breakdowns (breakdowns indented).

*Applicable to all 1993 MRDS availability assessment tables, part b.

Source: own compilation. (Based on 2008 YB.)

The assessment of the fulfillment of the “relaxed” milestone level 1 criteria indicates that 81%, i.e. 156 countries, are able to provide data for “GDP by expenditure” (i.e. NAQ table 1.1 or 1.2) and “GDP by industry” (i.e. NAQ table 2.1 or 2.2). To fulfill the “relaxed” milestone level 1 criteria the data can be produced in either “current” or “constant” prices. Thus this milestone level does not provide information about the capacity to produce “constant price data”. The assessment of the (relaxed) milestone level 2 criteria, requiring at least GNI (NAQ table 1.3 or 4.1), shows that only 67%, i.e. 129 countries, can compile this data in addition to the milestone level 1 criteria.

Table 34 shows that the *developed regions* have a high data availability for all NAQ tables 1.1, …, 4.2 needed for the 1993 MRDS and “relaxed” milestone levels 1 and 2 criteria. The availability of these NAQ tables is generally around 95% of the developed regions. It is lower for NAQ table 4.2, the “rest of the world”, being 84%, i.e. 32 out of the 38 countries. The latter lowers the number of countries that fulfill the 1993 MRDS, being 82%, i.e. 31 countries. However, almost 90%, precisely 34 of the 38 developed regions reported a substantial part of the 1993 MRDS (meaning at least all but one of the NAQ tables needed to fulfill the 1993 MRDS).

The *transition countries* show top results for the availability of NAQ tables 1.1, 2.1 and 2.2. The CIS countries are high for all NAQ tables. The transition economies of South-eastern Europe show a significant lack regarding the availability of GNI (i.e. NAQ table 1.3 or 4.1) and of data for the “rest of the world” (i.e. NAQ table 4.2). Consequently, only one country fulfills the full 1993 MRDS and still only 50% fulfill a substantial part of it. Further the milestone assessments drop from 100% for milestone level 1 to 50% for milestone level 2 (“relaxed” criteria).

The *developing regions* show good results for the availability of “GDP by industry” in “current” and “constant” prices (NAQ table 2.1 respectively 2.2) as well as for “GDP by expenditure” in “current” prices (NAQ table 1.1), all ranging around 80%. Hence, “relaxed” milestone level 1 is fulfilled by nearly 76%, i.e. 104 countries. Yet, only half of the developing regions, precisely 51%, i.e. 71 countries, compile the components for “value added by industry” (NAQ table 2.3). The availability for the “rest of the world” sector (NAQ table 4.2) is particularly low with 34%, i.e. 46 countries. The worst results regarding availability of NAQ table 4.2 are 15% for the Caribbean and 23% for Africa. Concerning Latin America, and Southern- and Western Asia approx. 90% of the countries have the NAQ tables required to fulfill the "relaxed" milestone 2 criteria and 75% to 100% in regard to the "relaxed" milestone level 1 criteria. Only Africa and Oceania show a considerable data gap regarding "relaxed" milestone level 1 and 2. The data availability of "relaxed" milestone level 1 is 58% for Africa and 50% for Oceania. With regard to "relaxed" milestone level 2 it is 40% for Africa and 42% for Oceania. South-eastern Asia drops from 82% of the countries fulfilling "relaxed" milestone level 1 to 45% fulfilling "relaxed" milestone level 2. All developing regions, other than Latin America, Western Asia and Southern Asia show significant gaps for "constant price" data (i.e. NAQ tables 1.2 and 2.2.) and for GNI (i.e. NAQ tables 1.3 or 4.1).

The biggest data gaps regarding the 1993 MRDS is observed for Oceania, where merely two countries fulfill the 1993 MRDS. Also NAQ table 2.3 is only compiled by (the same) two countries. "GDP by expenditure" in "current" prices (NAQ table 1.1) is available for just half of the 12 countries. The corresponding data in "constant" prices (NAQ table 1.2) is only available for 25%, i.e. 3 countries. However, 75%, i.e. 9 countries can compile data for "GDP by industry" in "current" and "constant" prices (NAQ tables 2.1 and 2.2).

In Table 34 we observe a substantial data gap for the transition countries of South-eastern Europe and all developing regions, besides Latin America regarding NAQ table 4.2 "rest of the world" (cf. last col.). Thus we can expect that the availability of the other institutional sectors, i.e. NAQ tables 4.3, …, 4.9, which are not included in the 1993 MRDS assessment, is also limited.

2.3.1.2 All Countries, Areas, and Territories

This part of our data availability assessment considers all 224 countries, areas, and territories, according to the MDG country list. The results for the additional 32 countries are displayed in row three of Table 36 and Table 37, p. 218 (see "Other countries from M.49 list").

Due to the low data availability for the additional 32 countries, the percent of countries reporting data for each NAQ table is lower than in the previous assessment, which covered only UN Member States. The best availability is for NAQ tables 1.1 and 2.1. More than 40% of the additional 32 countries produce these NAQ tables (i.e. 14 countries regarding NAQ table 1.1 and 13 countries regarding NAQ table 2.1). The availability of NAQ tables 1.2, 2.2 and 4.2 is approx. 20%. For NAQ tables 2.3 and

1.3 it is approx. 30%. Compared to the UN Member States this is a substantially worse data availability. Only 41%, i.e. 13 countries, out of the 32 additional countries fulfill “relaxed” milestone level 1 and 28% fulfill “relaxed” milestone level 2. With the inclusion of the additional countries, the compliance with 1993 MRDS increased by only two countries in and thus drops the percent of countries fulfilling this data set to 35%, i.e. 79 out of the total 224 countries. The reporting of at least all but one of the NAQ tables required to fulfill the 1993 MRDS criteria increased by five countries in total and drops the corresponding percent to just 49%, i.e. 110 countries.

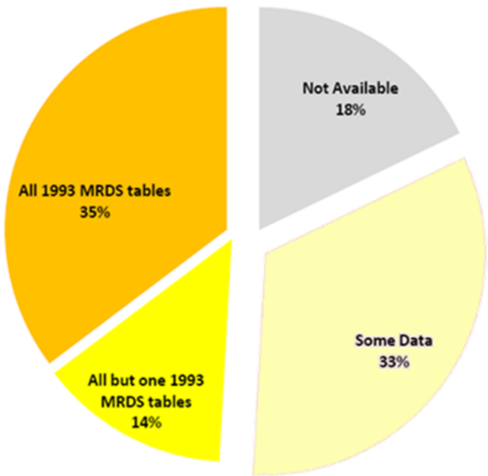


Figure 31 All (M49) Countries, Areas, and Territories Data Availability of 1993 MRDS
Source: own illustration. (Based on 2008 YB.)

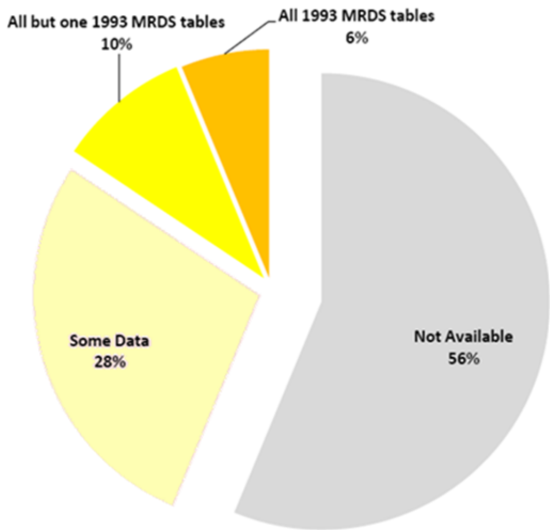


Figure 32 Non-UN Countries (from M49 list) Data Availability of 1993 MRDS
Source: own illustration. (Based on 2008 YB.)

Figure 31 and Figure 32 illustrate the shift in global data availability when considering the additional 32 countries that are not UN members. The overall data availability does not change dramatically (compare Figure 30, p. 213 and Figure 31, above). However, the share of countries with no NA data increased substantially since 56%, i.e. 18 out of the 32 non-UN member countries fall into that group. Particularly the percent of countries that fulfill 1993 MRDS falls five percentage points from 40% to 35%, leading to a cumulative total of only 49% of the 224 countries with availability of at least the substantial part of the 1993 MRDS.

Table 36 1993 MRDS for all Countries, Areas and Territories, Part a: Availability of Individual Required NAQ Tables

Data available between 2004-2008	Total number of countries $\lvert C \rvert$	Tables of the 1993 SNA Minimum Requirement Data Set (MRDS) 1.1 $\lvert C_{1.1} \rvert$ ($h_{1.1}$)	1.2 $\lvert C_{1.2} \rvert$ ($h_{1.2}$)	2.1 $\lvert C_{2.1} \rvert$ ($h_{2.1}$)	2.2 $\lvert C_{2.2} \rvert$ ($h_{2.2}$)	2.3 $\lvert C_{2.3} \rvert$ ($h_{2.3}$)	1.3/4.1 $\lvert C_{1.3 \vee 4.1} \rvert$ ($h_{1.3 \vee 4.1}$)	4.2 $\lvert C_{4.2} \rvert$ ($h_{4.2}$)
All (M49) Countries, Areas, and Territories	**224**	**172 (77%)**	**136 (61%)**	**179 (80%)**	**166 (74%)**	**129 (58%)**	**143 (64%)**	**95 (42%)**
UN Member States	192	158 (82%)	129 (67%)	166 (86%)	159 (83%)	119 (62%)	134 (70%)	89 (46%)
Other Countries from M49 list	32	14 (44%)	7 (22%)	13 (41%)	7 (22%)	10 (31%)	9 (28%)	6 (19%)
Developed Regions	**43**	**39 (91%)**	**37 (86%)**	**38 (88%)**	**36 (84%)**	**38 (88%)**	**39 (91%)**	**33 (77%)**
Transition Countries	**17**	**16 (94%)**	**13 (76%)**	**17 (100%)**	**16 (94%)**	**13 (76%)**	**12 (71%)**	**11 (65%)**
Transition countries of South-eastern Europe	6	6 (100%)	4 (67%)	6 (100%)	6 (100%)	5 (83%)	3 (50%)	2 (33%)
Commonwalth of Independent States (CIS)	11	10 (91%)	9 (82%)	11 (100%)	10 (91%)	8 (73%)	9 (82%)	9 (82%)
Developing regions	**164**	**117 (71%)**	**86 (52%)**	**124 (76%)**	**114 (70%)**	**78 (48%)**	**92 (56%)**	**51 (31%)**
Africa	56	32 (57%)	27 (48%)	37 (66%)	34 (61%)	20 (36%)	23 (41%)	12 (21%)
Northern Africa	6	3 (50%)	2 (33%)	4 (67%)	3 (50%)	3 (50%)	4 (67%)	2 (33%)
Sub-Saharan Africa	50	29 (58%)	25 (50%)	33 (66%)	31 (62%)	17 (34%)	19 (38%)	10 (20%)
Latin America and the Caribbean	46	37 (80%)	25 (54%)	37 (80%)	34 (74%)	27 (59%)	30 (65%)	18 (39%)
Caribbean	24	17 (71%)	7 (29%)	17 (71%)	14 (58%)	9 (38%)	12 (50%)	4 (17%)
Latin America	22	20 (91%)	18 (82%)	20 (91%)	20 (91%)	18 (82%)	18 (82%)	14 (64%)
Asia	42	39 (93%)	30 (71%)	39 (93%)	36 (86%)	27 (64%)	32 (76%)	17 (40%)
Eastern, Southern, and South-eastern Asia	26	23 (88%)	20 (77%)	23 (88%)	21 (81%)	12 (46%)	18 (69%)	11 (42%)
Eastern Asia	6	5 (83%)	3 (50%)	5 (83%)	3 (50%)	4 (67%)	5 (83%)	4 (67%)
Southern Asia	9	9 (100%)	8 (89%)	8 (89%)	9 (100%)	5 (56%)	8 (89%)	4 (44%)
South-eastern Asia	11	9 (82%)	9 (82%)	10 (91%)	9 (82%)	3 (27%)	5 (45%)	3 (27%)
Western Asia	16	16 (100%)	10 (63%)	16 (100%)	15 (94%)	15 (94%)	14 (88%)	6 (38%)
Oceania	20	9 (45%)	4 (20%)	11 (55%)	10 (50%)	4 (20%)	7 (35%)	4 (20%)
Official MDG Groupings								
Landlocked developing countries (LLDCs)	31	26 (84%)	21 (68%)	28 (90%)	26 (84%)	13 (42%)	18 (58%)	13 (42%)
Small island developing States (SIDS)	51	32 (63%)	16 (31%)	35 (69%)	32 (63%)	16 (31%)	27 (53%)	10 (20%)
Least developed countries (LDCs)	50	29 (58%)	25 (50%)	34 (68%)	32 (64%)	10 (20%)	18 (36%)	7 (14%)

Source: own compilation. (Based on 2008 YB.)

Table 37 1993 MRDS for all Countries, Areas and Territories, Part b: MRDS and Milestone Level Results

Data available between 2004-2008	Total number of countries $\lvert C \rvert$	1993 MRDS All tables $\lvert C_{1993} \rvert$ (h_{1993})	All but one tables $\lvert C_{1993(p)} \rvert$ ($h_{1993(p)}$)	Milestone level 1 or higher $\lvert C_{1r} \rvert$ (h_{1r})	2 $\lvert C_{2r} \rvert$ (h_{2r})
All (M49) Countries, Areas, and Territories	**224**	**79 (35%)**	**110 (49%)**	**169 (75%)**	**138 (62%)**
UN Member States	192	77 (40%)	105 (55%)	156 (81%)	129 (67%)
Other Countries from M49 list	32	2 (6%)	5 (16%)	13 (41%)	9 (28%)
Developed Regions	**43**	**31 (72%)**	**35 (81%)**	**38 (88%)**	**38 (88%)**
Transition Countries	**17**	**9 (53%)**	**12 (71%)**	**16 (94%)**	**12 (71%)**
Transition countries of South-eastern Europe	6	1 (17%)	3 (50%)	6 (100%)	3 (50%)
Commonwalth of Independent States (CIS)	11	8 (73%)	9 (82%)	10 (91%)	9 (82%)
Developing regions	**164**	**39 (24%)**	**63 (38%)**	**115 (70%)**	**88 (54%)**
Africa	56	11 (20%)	16 (29%)	31 (55%)	21 (38%)
Northern Africa	6	1 (17%)	2 (33%)	3 (50%)	3 (50%)
Sub-Saharan Africa	50	10 (20%)	14 (28%)	28 (56%)	18 (36%)
Latin America and the Caribbean	46	15 (33%)	21 (46%)	37 (80%)	30 (65%)
Caribbean	24	2 (8%)	4 (17%)	17 (71%)	12 (50%)
Latin America	22	13 (59%)	17 (77%)	20 (91%)	18 (82%)
Asia	42	11 (26%)	24 (57%)	39 (93%)	32 (76%)
Eastern, Southern, and South-eastern Asia	26	7 (27%)	14 (54%)	23 (88%)	18 (69%)
Eastern Asia	6	2 (33%)	4 (67%)	5 (83%)	5 (83%)
Southern Asia	9	3 (33%)	6 (67%)	9 (100%)	8 (89%)
South-eastern Asia	11	2 (18%)	4 (36%)	9 (82%)	5 (45%)
Western Asia	16	4 (25%)	10 (63%)	16 (100%)	14 (88%)
Oceania	20	2 (10%)	2 (10%)	8 (40%)	5 (25%)
Official MDG Groupings					
Landlocked developing countries (LLDCs)	31	9 (29%)	15 (48%)	26 (84%)	18 (58%)
Small island developing States (SIDS)	51	6 (12%)	9 (18%)	31 (61%)	24 (47%)
Least developed countries (LDCs)	50	4 (8%)	11 (22%)	29 (58%)	17 (34%)

Source: own compilation. (Based on 2008 YB.)

Referring to the economic classification of countries in Table 36 and Table 37, p. 218, we observe that none of the five additional *developed regions* provides 1993 MRDS. Only one country reports the substantial part of it, i.e. all but one of the NAQ tables to fulfill the 1993 MRDS criteria. The results for *transition countries* remain unchanged, because all are UN Member States. The *developing regions* show two additional countries that report 1993 MRDS (total percent falls to 24%, i.e. 39 countries) and four additional countries having at least all but one of the required NAQ tables (total percent falls to 38%, i.e. 63 countries). The two countries that report 1993 MRDS belong to the Western Asia and the Eastern Asia region.

The *Millennium Development Goals (MDG)* country groups show that Landlocked Developing Countries (LLDCs) are above average regarding NA data production compared to the grand total of the developing regions. Small Island Developing States (SIDS) and particularly Least Developed Countries (LDCs) are below average. However, in order to assess the economy and economic developments of small countries, not necessarily all NAQ tables that make up the 1993 MRDS need to be available. SIDS and LDCs have the worst compliance 1993 MRDS (12% and 8%, respectively, fulfill 1993 MRDS). "Relaxed" milestone level 1 is on the other hand fulfilled by 61% of the SIDS group and 58% of the LDCs. This means they can produce NA statistics for "GDP by expenditure" and "GDP by industry" in "current" or "constant prices". However, less than half of the countries in these groups produce enough data to fulfill "relaxed" milestone level 2, meaning they are not capable of providing GNI.

2.3.1.3 Other Organizations and Economic Groups

The results presented in this part of our 1993 MRDS assessment are particularly interesting as we shows the NA production and data gaps in important economic groups like OECD member countries, EU27 countries, G20 countries, and other advanced economies (i.e. advanced economies, other than G20 countries).

Regarding the UN Regional Commissions, Table 39, p. 221 shows the lowest data availability for the Economic Commission for Africa (ECA), which comprises of the same set of countries as the African region. The best countries are the ten developed regions of the Economic Commission for Latin America and the Caribbean (ECLAC), reporting the 1993 MRDS and also fulfilling "relaxed" milestone level 2.

Table 38 1993 MRDS for Other Organizations/ Groups, Part a: Availability of Individual NAQ Tables

Data available between 2004-2008	Total number of countries	Tables of the 1993 SNA Minimum Requirement Data Set (MRDS)						
		1.1	1.2	2.1	2.2	2.3	1.3/4.1	4.2
	$\lvert C\rvert$	$\lvert C_{1.1}\rvert$ ($h_{1.1}$)	$\lvert C_{1.2}\rvert$ ($h_{1.2}$)	$\lvert C_{2.1}\rvert$ ($h_{2.1}$)	$\lvert C_{2.2}\rvert$ ($h_{2.2}$)	$\lvert C_{2.3}\rvert$ ($h_{2.3}$)	$\lvert C_{1.3\vee 4.1}\rvert$ ($h_{1.3\vee 4.1}$)	$\lvert C_{4.2}\rvert$ ($h_{4.2}$)
Other Organizations/Economic Groupings								
CARICOM	20	16 (80%)	4 (20%)	16 (80%)	15 (75%)	8 (40%)	13 (65%)	2 (10%)
OECD	30	30 (100%)	30 (100%)	30 (100%)	30 (100%)	30 (100%)	30 (100%)	27 (90%)
SADC	15	12 (80%)	11 (73%)	14 (93%)	13 (87%)	8 (53%)	10 (67%)	6 (40%)
EU27	27	27 (100%)	27 (100%)	27 (100%)	26 (96%)	27 (100%)	27 (100%)	26 (96%)
G20	19	19 (100%)	18 (95%)	19 (100%)	18 (95%)	17 (89%)	18 (95%)	15 (79%)
Advanced, other than G-20 (excl. Taiwan Province of China)	23	23 (100%)	23 (100%)	23 (100%)	22 (96%)	23 (100%)	23 (100%)	20 (87%)
UN Regional Organizations								
ECA	52	32 (62%)	27 (52%)	37 (71%)	34 (65%)	20 (38%)	23 (44%)	12 (23%)
Northern Africa	5	3 (60%)	2 (40%)	4 (80%)	3 (60%)	3 (60%)	4 (80%)	2 (40%)
Sub-Saharan Africa	47	29 (62%)	25 (53%)	33 (70%)	31 (66%)	17 (36%)	19 (40%)	10 (21%)
ECE	56	53 (95%)	48 (86%)	53 (95%)	51 (91%)	48 (86%)	49 (88%)	42 (75%)
Developed regions	40	39 (98%)	36 (90%)	38 (95%)	37 (93%)	37 (93%)	36 (90%)	31 (78%)
of which: South-eastern Europe transition countries	6	6 (100%)	4 (67%)	6 (100%)	6 (100%)	5 (83%)	3 (50%)	2 (33%)
Commonwalth of Independent States (CIS)	11	10 (91%)	9 (82%)	11 (100%)	10 (91%)	8 (73%)	9 (82%)	9 (82%)
Developing regions	5	4 (80%)	3 (60%)	4 (80%)	4 (80%)	3 (60%)	4 (80%)	2 (40%)
ESCAP	49	39 (80%)	31 (63%)	43 (88%)	41 (84%)	21 (43%)	34 (69%)	21 (43%)
Developed regions	3	3 (100%)	3 (100%)	3 (100%)	3 (100%)	3 (100%)	3 (100%)	2 (67%)
Commonwalth of Independent States (CIS)	8	7 (88%)	6 (75%)	8 (100%)	7 (88%)	5 (63%)	6 (75%)	6 (75%)
Developing regions	38	29 (76%)	22 (58%)	32 (84%)	31 (82%)	13 (34%)	25 (66%)	13 (34%)
Asia	26	23 (88%)	19 (73%)	23 (88%)	22 (85%)	11 (42%)	18 (69%)	9 (35%)
Eastern Asia	4	3 (75%)	1 (25%)	3 (75%)	2 (50%)	2 (50%)	3 (75%)	2 (50%)
Southern Asia	9	9 (100%)	8 (89%)	8 (89%)	9 (100%)	5 (56%)	8 (89%)	4 (44%)
South-eastern Asia	11	9 (82%)	9 (82%)	10 (91%)	9 (82%)	3 (27%)	5 (45%)	3 (27%)
Western Asia	2	2 (100%)	1 (50%)	2 (100%)	2 (100%)	1 (50%)	2 (100%)	0 (0%)
Oceania	12	6 (50%)	3 (25%)	9 (75%)	9 (75%)	2 (17%)	7 (58%)	4 (33%)
ESCWA	14	13 (93%)	7 (50%)	13 (93%)	11 (79%)	13 (93%)	11 (79%)	5 (36%)
ECLAC	44	42 (95%)	34 (77%)	42 (95%)	42 (95%)	35 (80%)	37 (84%)	27 (61%)
Developed regions	10	10 (100%)	10 (100%)	10 (100%)	10 (100%)	10 (100%)	10 (100%)	10 (100%)
Developing regions	34	32 (94%)	24 (71%)	32 (94%)	32 (94%)	25 (74%)	27 (79%)	17 (50%)
Caribbean	13	11 (85%)	5 (38%)	11 (85%)	11 (85%)	6 (46%)	8 (62%)	2 (15%)
Latin America	20	20 (100%)	18 (90%)	20 (100%)	20 (100%)	18 (90%)	18 (90%)	14 (70%)
Other (Asia)	1	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)

Source: own compilation. (Based on 2008 YB.)

The fulfillment of 1993 MRDS is particularly low for the Caribbean Community (CARICOM). This is not surprising with the Caribbean being the region with one of the lowest NA capacities. Only one of the 20 countries produces the 1993 MRDS. The bottleneck is the unavailability of NAQ tables 2.3 and 4.2. However, the NA data needed for “relaxed” milestone level 1 are compiled by 80%, i.e. 16 countries, and for “relaxed” milestone level 2 by 65%, i.e. 13 out of the 20 countries. The EU27 has the best fulfillment of 1993 MRDS with nearly 100%, only lacking one country. All EU27 fulfill the “relaxed” milestone level 1 and 2 criteria. The results for OECD and for the other advanced economies are also very good. Nearly all countries fulfill the 1993 MRDS. Only NAQ table 4.2 is not available for three countries in both groups. Further, one country of the other advanced economies does not produce data for NAQ table 2.3. The G20 countries show a lower fulfillment of 1993 MRDS, which is below 80% for this group, i.e. only 15 out of the 19 countries. One G20 country also does not fulfill “relaxed” milestone 2, meaning that this country even has not reported GNI.

Table 39 1993 MRDS for Other Organizations/ Groups, Part b: MRDS and Milestone Level Results

Data available between 2004-2008	Total number of countries $\lvert C \rvert$	1993 MRDS: All tables $\lvert C_{1993} \rvert$ (h_{1993})	1993 MRDS: All but one tables $\lvert C_{1993(p)} \rvert$ ($h_{1993(p)}$)	Milestone level: 1 or higher $\lvert C_{1r} \rvert$ (h_{1r})	Milestone level: 2 $\lvert C_{2r} \rvert$ (h_{2r})
Other Organizations/Economic Groupings					
CARICOM	20	1 (5%)	4 (20%)	16 (80%)	13 (65%)
OECD	30	27 (90%)	30 (100%)	30 (100%)	30 (100%)
SADC	15	6 (40%)	7 (47%)	12 (80%)	9 (60%)
EU27	27	26 (96%)	26 (96%)	27 (100%)	27 (100%)
G20	19	15 (79%)	17 (89%)	19 (100%)	18 (95%)
Advanced, other than G-20 (excl. Taiwan Province of China)	23	20 (87%)	22 (96%)	23 (100%)	23 (100%)
UN Regional Organizations					
ECA	52	11 (21%)	16 (31%)	31 (60%)	21 (40%)
Northern Africa	5	1 (20%)	2 (40%)	3 (60%)	3 (60%)
Sub-Saharan Africa	47	10 (21%)	14 (30%)	28 (60%)	18 (38%)
ECE	56	39 (70%)	45 (80%)	52 (93%)	48 (86%)
Developed regions	40	29 (73%)	33 (83%)	38 (95%)	35 (88%)
of which: South-eastern Europe transition countries	6	1 (17%)	3 (50%)	6 (100%)	3 (50%)
Commonwalth of Independent States (CIS)	11	8 (73%)	9 (82%)	10 (91%)	9 (82%)
Developing regions	5	2 (40%)	3 (60%)	4 (80%)	4 (80%)
ESCAP	49	15 (31%)	24 (49%)	39 (80%)	32 (65%)
Developed regions	3	2 (67%)	3 (100%)	3 (100%)	3 (100%)
Commonwalth of Independent States (CIS)	8	5 (63%)	6 (75%)	7 (88%)	6 (75%)
Developing regions	38	8 (21%)	15 (39%)	29 (76%)	23 (61%)
Asia	26	6 (23%)	13 (50%)	23 (88%)	18 (69%)
Eastern Asia	4	1 (25%)	2 (50%)	3 (75%)	3 (75%)
Southern Asia	9	3 (33%)	6 (67%)	9 (100%)	8 (89%)
South-eastern Asia	11	2 (18%)	4 (36%)	9 (82%)	5 (45%)
Western Asia	2	0 (0%)	1 (50%)	2 (100%)	2 (100%)
Oceania	12	2 (17%)	2 (17%)	6 (50%)	5 (42%)
ESCWA	14	2 (14%)	7 (50%)	13 (93%)	11 (79%)
ECLAC	44	26 (59%)	31 (70%)	42 (95%)	37 (84%)
Developed regions	10	10 (100%)	10 (100%)	10 (100%)	10 (100%)
Developing regions	34	16 (47%)	21 (62%)	32 (94%)	27 (79%)
Caribbean	13	2 (15%)	3 (23%)	11 (85%)	8 (62%)
Latin America	20	13 (65%)	17 (85%)	20 (100%)	18 (90%)
Other (Asia)	1	1 (100%)	1 (100%)	1 (100%)	1 (100%)

Source: own compilation. (Based on 2008 YB.)

2.3.2 **2008 MRDS Assessment**

Our 2008 MRDS assessment, provided in the following, is structured like our 1993 MRDS assessment, meaning in three parts.

2.3.2.1 UN Member States

The assessment of 2008 MRDS reveals substantial data gaps that are significantly larger than identified by the 1993 MRDS assessment.

All country groups show substantial data gaps for the production of the institutional sector accounts (NAQ tables 4.1, …, 4.9) until NL. Production of these NAQ tables until NL requires advanced NA capacity. The 1993 MRDS assessment only required the availability of any data for NAQ tables 4.1 and 4.2. Some of the countries that do not fulfill the 2008 MRDS requirements might provide at least some data for NAQ tables 4.1, …, 4.9. This is evident from the higher availability of NAQ table 4.2 and NAQ tables 3.1 or 4.1 in the 1993 MRDS compared to the 2008 MRDS.

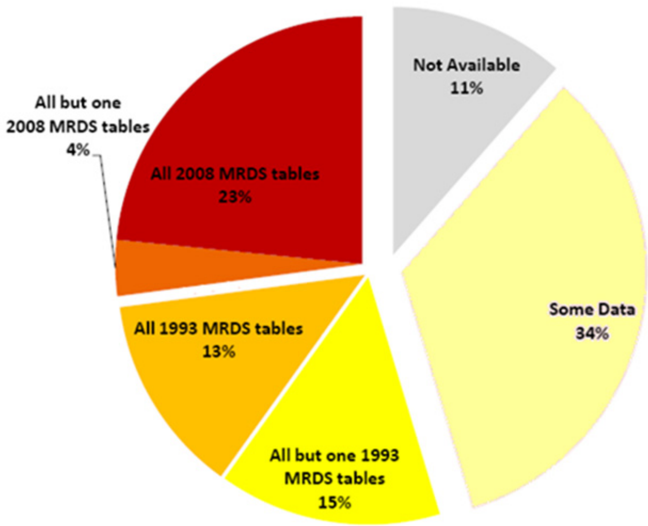


Figure 33 UN Member States Data Availability of 1993 and 2008 MRDS

Source: own illustration. (Based on 2008 YB and 2009 YB.)

Figure 33 indicates the available data for UN Member States. As already found by our 1993 MRDS assessment, 45% of the countries have either only some data, i.e. below the level of fulfilling a substantial part of the 1993 MRDS, or no data. Further we found that the other 55% fulfill 1993 MRDS or a substantial part of it. Yet, only half of these countries, i.e. 27%, fulfill 2008 MRDS (23%) or at least a substantial part of it (13%).

Referring to Table 40, p. 223, we observe for the UN Member States that most NA data are available for NAQ table 4.1 "total economy" with 43% (82 countries) and for NAQ table 4.2 "rest of the world" with 40% (77 countries). Of the institutional sectors, NAQ table 4.5 "government" has the best data availability with 37% (71 countries). Data for NAQ table 4.6 "households" and 4.7 "NPISH" show the least availability.

For the developed regions, data availability of NAQ tables 4.1, …, 4.9 until NL is still strong (cf. Table 40, p. 223, row two). Nevertheless, nearly 30%, i.e. eleven countries do not fulfill 2008 MRDS (cf. Table 41, p. 223). The reporting for CIS transition economies is at a level where two thirds, i.e. seven countries, report 2008 MRDS. Of the South-eastern Europe *transition economies* only one third, i.e. two countries, report at least a substantial part of the 2008 MRDS.

Table 40 2008 MRDS for UN Members, Part a: Availability of Individual Required NAQ Tables

Data available between 2000-2009 for institutional sectors, 2004-2008 for other data	Total number of countries $\|C\|$	Institutional secor account tables of the *2008 SNA* Minimum Requirement Data Set (MRDS) until net lending/ net borrowing										
		1.3/4.1 $\|C_{1.3\vee 4.1(nl)}\|$ $(h_{1.3\vee 4.1(nl)})$	4.1 $\|C^{nl}_{4.1}\|$ $(h^{nl}_{4.1})$	4.2 $\|C^{nl}_{4.2}\|$ $(h^{nl}_{4.2})$	4.3 $\|C^{nl}_{4.3}\|$ $(h^{nl}_{4.3})$	4.4 $\|C^{nl}_{4.4}\|$ $(h^{nl}_{4.4})$	4.8 (combined 4.3 & 4.4) $\|C^{nl}_{4.8}\|$ $(h^{nl}_{4.8})$	4.3 and 4.4, or 4.8 $\|C_{4.3+4.4\vee 4.8}\|$ $(h_{4.3+4.4\vee 4.6})$	4.5 $\|C^{nl}_{4.5}\|$ $(h^{nl}_{4.5})$	4.6 $\|C^{nl}_{4.6}\|$ $(h^{nl}_{4.6})$	4.7 $\|C^{nl}_{4.7}\|$ $(h^{nl}_{4.7})$	4.9 (combined 4.6 & 4.7) $\|C^{nl}_{4.9}\|$ $(h^{nl}_{4.9})$	4.6 and 4.7, or 4.9 $\|C_{4.6+4.7\vee 4.9}\|$ $(h_{4.6+4.7\vee 4.9})$
UN Member States	**192**	**111 (58%)**	**82 (43%)**	**77 (40%)**	**57 (30%)**	**58 (30%)**	**26 (14%)**	**60 (31%)**	**71 (37%)**	**53 (28%)**	**45 (23%)**	**27 (14%)**	**53 (28%)**
Developed Regions	**38**	**32 (84%)**	**31 (82%)**	**28 (74%)**	**27 (71%)**	**27 (71%)**	**24 (63%)**	**30 (79%)**	**34 (89%)**	**23 (61%)**	**23 (61%)**	**25 (66%)**	**31 (82%)**
Transition Countries	**17**	**12 (71%)**	**12 (71%)**	**9 (53%)**	**9 (53%)**	**9 (53%)**	**0 (0%)**	**9 (53%)**	**9 (53%)**	**9 (53%)**	**9 (53%)**	**0 (0%)**	**9 (53%)**
Transition countries of South-eastern Europe	6	3 (50%)	3 (50%)	1 (17%)	2 (33%)	2 (33%)	0 (0%)	2 (33%)	2 (33%)	2 (33%)	2 (33%)	0 (0%)	2 (33%)
Commonwealth of Independent States (CIS)	11	9 (82%)	9 (82%)	8 (73%)	7 (64%)	7 (64%)	0 (0%)	7 (64%)	7 (64%)	7 (64%)	7 (64%)	0 (0%)	7 (64%)
Developing Regions	**137**	**67 (49%)**	**39 (28%)**	**40 (29%)**	**21 (15%)**	**22 (16%)**	**2 (1%)**	**21 (15%)**	**28 (20%)**	**21 (15%)**	**13 (9%)**	**2 (1%)**	**13 (9%)**
Africa	53	21 (40%)	15 (28%)	15 (28%)	8 (15%)	8 (15%)	0 (0%)	8 (15%)	11 (21%)	9 (17%)	4 (8%)	0 (0%)	4 (8%)
Northern Africa	5	4 (80%)	3 (60%)	3 (60%)	3 (60%)	3 (60%)	0 (0%)	3 (60%)	3 (60%)	3 (60%)	1 (20%)	0 (0%)	1 (20%)
Sub-Saharan Africa	48	17 (35%)	12 (25%)	12 (25%)	5 (10%)	5 (10%)	0 (0%)	5 (10%)	8 (17%)	6 (13%)	3 (6%)	0 (0%)	3 (6%)
Latin America and the Caribbean	33	20 (61%)	13 (39%)	10 (30%)	8 (24%)	9 (27%)	1 (3%)	8 (24%)	9 (27%)	7 (21%)	6 (18%)	1 (3%)	6 (18%)
Caribbean	13	5 (38%)	0 (0%)	1 (8%)	0 (0%)	1 (8%)	0 (0%)	0 (0%)	1 (8%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Latin America	20	15 (75%)	13 (65%)	9 (45%)	8 (40%)	8 (40%)	1 (5%)	8 (40%)	8 (40%)	7 (35%)	6 (30%)	1 (5%)	6 (30%)
Asia	39	23 (59%)	10 (26%)	11 (28%)	5 (13%)	5 (13%)	1 (3%)	5 (13%)	8 (21%)	5 (13%)	3 (8%)	1 (3%)	3 (8%)
Eastern, Southern, and South-eastern Asia	24	12 (50%)	5 (21%)	6 (25%)	3 (13%)	3 (13%)	1 (4%)	3 (13%)	5 (21%)	3 (13%)	1 (4%)	1 (4%)	2 (8%)
Eastern Asia	4	2 (50%)	2 (50%)	2 (50%)	2 (50%)	2 (50%)	1 (25%)	2 (50%)	3 (75%)	1 (25%)	0 (0%)	1 (25%)	1 (25%)
Southern Asia	9	5 (56%)	2 (22%)	3 (33%)	1 (11%)	1 (11%)	0 (0%)	1 (11%)	2 (22%)	2 (22%)	1 (11%)	0 (0%)	1 (11%)
South-eastern Asia	11	5 (45%)	1 (9%)	1 (9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Western Asia	15	11 (73%)	5 (33%)	5 (33%)	2 (13%)	2 (13%)	0 (0%)	2 (13%)	3 (20%)	2 (13%)	2 (13%)	0 (0%)	1 (7%)
Oceania	12	3 (25%)	1 (8%)	4 (33%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Legend*

In columns: the MRDS tables by table number, with absolute number of countries reporting data and percentage of group total.

In rows: the country groups, with aggregates and breakdowns (breakdowns indented).

*Applicable to all 2008 MRDS availability assessment tables, part a.

Source: own compilation. (Based on 2008 YB and 2009 YB.)

Table 41 2008 MRDS for UN Members, Part b: Institutional Sector, MRDS and Milestone Level Results

Data available between 2000-2009 for institutional sectors, 2004-2008 for other data	Total number of countries $\|C\|$	2008 MRDS sectors: All tables $\|C_{IS}\|$ (h_{IS})	2008 MRDS sectors: All but one tables $\|C_{IS(p)}\|$ ($h_{IS(p)}$)	2008 MRDS: All tables $\|C_{2008}\|$ (h_{2008})	2008 MRDS: All but one tables $\|C_{2008(p)}\|$ ($h_{2008(p)}$)	Fin. acc. of NAQ Table 4.2 $\|C^{f}_{4.2}\|$ ($h^{f}_{4.2}$)	Strict milestone level: 1 or higher $\|C_{1s}\|$ (h_{1s})	Strict milestone level: 2 $\|C_{2s}\|$ (h_{2s})
UN Member States	**192**	**47 (24%)**	**56 (29%)**	**45 (23%)**	**52 (27%)**	**43 (22%)**	**126 (66%)**	**40 (21%)**
Developed Regions	**38**	**27 (71%)**	**27 (71%)**	**27 (71%)**	**27 (71%)**	**26 (68%)**	**34 (89%)**	**26 (68%)**
Transition Countries	**17**	**8 (47%)**	**9 (53%)**	**8 (47%)**	**8 (47%)**	**1 (6%)**	**13 (76%)**	**1 (6%)**
Transition countries of South-eastern Europe	6	1 (17%)	2 (33%)	1 (17%)	1 (17%)	0 (0%)	4 (67%)	0 (0%)
Commonwalth of Independent States (CIS)	11	7 (64%)	7 (64%)	7 (64%)	7 (64%)	1 (9%)	9 (82%)	1 (9%)
Developing Regions	**137**	**12 (9%)**	**20 (15%)**	**10 (7%)**	**17 (12%)**	**16 (12%)**	**79 (58%)**	**13 (9%)**
Africa	53	3 (6%)	8 (15%)	1 (2%)	5 (9%)	3 (6%)	27 (51%)	2 (4%)
Northern Africa	5	1 (20%)	3 (60%)	0 (0%)	2 (40%)	3 (60%)	2 (40%)	2 (40%)
Sub-Saharan Africa	48	2 (4%)	5 (10%)	1 (2%)	3 (6%)	0 (0%)	25 (52%)	0 (0%)
Latin America and the Caribbean	33	6 (18%)	7 (21%)	6 (18%)	8 (24%)	6 (18%)	23 (70%)	6 (18%)
Caribbean	13	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (38%)	0 (0%)
Latin America	20	6 (30%)	7 (35%)	6 (30%)	8 (40%)	6 (30%)	18 (90%)	6 (30%)
Asia	39	3 (8%)	5 (13%)	3 (8%)	4 (10%)	7 (18%)	26 (67%)	5 (13%)
Eastern, Southern, and South-eastern Asia	24	2 (8%)	3 (13%)	2 (8%)	2 (8%)	5 (21%)	17 (71%)	4 (17%)
Eastern Asia	4	1 (25%)	2 (50%)	1 (25%)	1 (25%)	2 (50%)	1 (25%)	1 (25%)
Southern Asia	9	1 (11%)	1 (11%)	1 (11%)	1 (11%)	2 (22%)	7 (78%)	2 (22%)
South-eastern Asia	11	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (9%)	9 (82%)	1 (9%)
Western Asia	15	1 (7%)	2 (13%)	1 (7%)	2 (13%)	2 (13%)	9 (60%)	1 (7%)
Oceania	12	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (25%)	0 (0%)

Legend*

In columns: the availability of institutional sector accounts, MRDS compliance, availability of financial accounts for the rest of the world sector, and milestone level results, with absolute number of countries and percentage of group total.

In rows: the country groups, with aggregates and breakdowns (breakdowns indented).

*Applicable to all 2008 MRDS availability assessment tables, part a.

Source: own compilation. (Based on 2008 YB and 2009 YB.)

The *developing regions* (refer to row six in the above results tables) indicate the lowest availability of the required data. Only ten countries fulfill the 2008 MRDS. Another ten report a substantial part of the 2008 MRDS, mainly missing the "NPISH" sector, i.e. NAQ table 4.7. Even of NAQ table 4.5 "government", which is the domestic sector with the best availability, data until NL are only available for 20%, i.e. 28 out of the 137 developing countries. For UN Member States, the fulfillment of 2008 MRDS fell to 23%, i.e. 47 countries. Thus 147 of the UN Member States do not produce the required NA data to fulfill the 2008 MRDS. Moreover, the results make it evident that countries have either very strong NA data production capacity, or only low NA data production capacity. Merely about ten to 20 countries are in an intermediate state, where they produce a substantial part, but not all data required by 2008 MRDS. In this context it needs to be pointed out that three countries (Australia, Canada, and the US) do compile their NA until NL. Yet, these countries do not report NAQ table 4.3 "non-financial" and NAQ table 4.4 "financial" sector, individually. Instead, NAQ table 4.8 is produced, i.e. the NAQ table presenting the unconsolidated aggregated data of NAQ tables 4.3 and 4.4. Thus these countries do not fulfill the 2008 MRDS criteria. Eight countries prepare NAQ table 4.9, i.e. the unconsolidated aggregated data of the individual NAQ tables 4.6 "households" and 4.7 "NPISH" instead of the individual NAQ tables, which is sufficient for the 2008 MRDS.

The strongest of the *developing regions* is Latin America, as in the 1993 MRDS assessment. The highest data gaps regarding 2008 MRDS can be observed for Oceania, the Asian region, the Caribbean and the Sub-Sahara region. The Sub-Sahara region has in particular prospects for capacity building, since already one quarter, i.e. 12 out of the 48 countries, produce NAQ tables 4.1 "total economy" and 4.2 "rest of the world" until NL. Yet, only three to eight countries have NAQ tables 4.3, … 4.7, i.e. the other domestic sectors.

The "strict" *milestone level* 1 and 2 criteria lead to fewer countries that are able to comply with these milestone levels. *Milestone level 1* is fulfilled by 66% of the countries; this is 30 countries less than for the "relaxed" milestone level criteria. This indicates that these countries cannot provide data for all of NAQ tables 1.1, … 2.2, i.e. for "GDP by expenditure" and "GDP by industry" in both "current" and "constant" prices. Regarding "strict" *milestone level 2*, particularly the requirement of "financial accounts" of NAQ table 4.2 "external sector", which is not part of the 2008 MRDS criteria, strongly lowers the fulfillment of this measure. For the developed regions, the fulfillment of *milestone level 2* dropped from 95% using the "relaxed" criteria, to 68% using the "strict" criteria. Only one of the transition countries is able to fulfill the "strict" criteria, which were previously fulfilled by 12 of the countries. For the developing regions, the fulfillment of *milestone level 2* dropped from nearly 60% by "relaxed" criteria to less than 10% by "strict" criteria.

2.3.2.2 All Countries, Areas, and Territories

The absolute number of countries reporting the 2008 MRDS, or a substantial part of it, does not change compared to the UN Member States assessment. Thus, the percent of countries fulfilling 2008 MRDS drops to 20% of the now 224 countries. Further, only four additional countries reach the "strict" milestone level 1 (all from the developing region). No additional country fulfills "strict" milestone level 2.

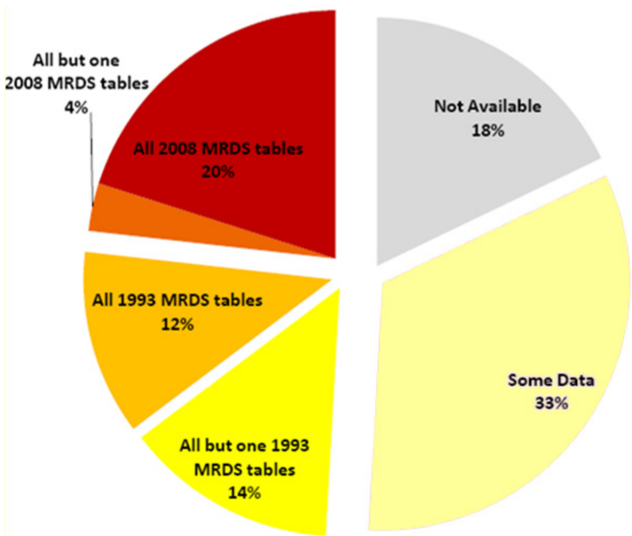


Figure 34 All (M49) Countries, Areas, and Territories
Data Availability of 1993 and 2008 MRDS
Source: own illustration. (Based on 2008 YB and 2009 YB.)

Figure 34 illustrates the severely worse data status of the 32 non-UN members on the global data availability. None of these countries reports sufficient data to fulfill more than the 1993 MRDS. The majority of them do not report any NA data (see: Figure 32, p. 217 in our 1993 MRDS assessment). Thus, the percent countries reporting only some or no data, shifts to more than 50% of the 224 countries. Of the other 49% that fulfill at least the partial 1993 MRDS, we observe that 26 % of the countries do not fulfill more than the 1993 MRDS, and 24% fulfill at least the partial 2008 MRDS.

The *MDG groups* indicate that Small Island Developing States (SIDS) and Least Developed Countries (LDCs) are well below average (cf. last three rows in Table 42, p. 227 and Table 43, p. 228). None of the latter two groups fulfill the 2008 MRDS, one LDC fulfills the substantial part of it. However, nearly 50% of the LDCs, i.e. 24 countries, meet "strict" milestone level 1 and also more than 25% of the SIDS group, i.e. 14 countries. One country of the SIDS group (2%) fulfills "strict" milestone level 2. The Landlocked Developing Countries (LLDCs) are well above average compared to the total of developing regions. Nearly 70%, i.e. 21 countries, fulfill "strict" milestone level 1, approx. one quarter, i.e. 7 countries, reports a substantial part of the 2008 MRDS, and 16%, i.e. 5 countries even fulfill 2008 MRDS.

Table 42 2008 MRDS for all Countries, Areas and Territories, Part a: Availability of Individual Required NAQ Tables

Data available between 2000-2009	Total number of countries $\lvert C \rvert$	Institutional secor account tables of the *2008 SNA* Minimum Requirement Data Set (MRDS) until net lending/ net borrowing											
		1.3/4.1 $\lvert C_{1.3\vee 4.1(nl)} \rvert$ $(h_{1.3\vee 4.1(nl)})$	4.1 $\lvert C^{nl}_{4.1} \rvert$ $(h^{nl}_{4.1})$	4.2 $\lvert C^{nl}_{4.2} \rvert$ $(h^{nl}_{4.2})$	4.3 $\lvert C^{nl}_{4.3} \rvert$ $(h^{nl}_{4.3})$	4.4 $\lvert C^{nl}_{4.4} \rvert$ $(h^{nl}_{4.4})$	4.8 (combined 4.3 & 4.4) $\lvert C^{nl}_{4.8} \rvert$ $(h^{nl}_{4.8})$	4.3 and 4.4, or 4.8 $\lvert C_{4.3+4.4\vee 4.8} \rvert$ $(h_{4.3+4.4\vee 4.6})$	4.5 $\lvert C^{nl}_{4.5} \rvert$ $(h^{nl}_{4.5})$	4.6 $\lvert C^{nl}_{4.6} \rvert$ $(h^{nl}_{4.6})$	4.7 $\lvert C^{nl}_{4.7} \rvert$ $(h^{nl}_{4.7})$	4.9 (combined 4.6 & 4.7) $\lvert C^{nl}_{4.9} \rvert$ $(h^{nl}_{4.9})$	4.6 and 4.7, or 4.9 $\lvert C_{4.6+4.7\vee 4.9} \rvert$ $(h_{4.6+4.7\vee 4.9})$
All (M49) Countries, Areas, and Territories	**224**	**116 (52%)**	**85 (38%)**	**81 (36%)**	**59 (26%)**	**60 (27%)**	**26 (12%)**	**62 (28%)**	**75 (33%)**	**54 (24%)**	**45 (20%)**	**28 (13%)**	**54 (24%)**
UN Member States	192	111 (58%)	82 (43%)	77 (40%)	57 (30%)	58 (30%)	26 (14%)	60 (31%)	71 (37%)	53 (28%)	45 (23%)	27 (14%)	53 (28%)
Other Countries from M49 list	32	5 (16%)	3 (9%)	4 (13%)	2 (6%)	2 (6%)	0 (0%)	2 (6%)	4 (13%)	1 (3%)	0 (0%)	1 (3%)	1 (3%)
Developed Regions	**43**	**33 (77%)**	**31 (72%)**	**29 (67%)**	**27 (63%)**	**27 (63%)**	**24 (56%)**	**30 (70%)**	**35 (81%)**	**23 (53%)**	**23 (53%)**	**25 (58%)**	**31 (72%)**
Transition Countries	**17**	**12 (71%)**	**12 (71%)**	**9 (53%)**	**9 (53%)**	**9 (53%)**	**0 (0%)**	**9 (53%)**	**9 (53%)**	**9 (53%)**	**9 (53%)**	**0 (0%)**	**9 (53%)**
Transition countries of South-eastern Europe	6	3 (50%)	3 (50%)	1 (17%)	2 (33%)	2 (33%)	0 (0%)	2 (33%)	2 (33%)	2 (33%)	2 (33%)	0 (0%)	2 (33%)
Commonwalth of Independent States (CIS)	11	9 (82%)	9 (82%)	8 (73%)	7 (64%)	7 (64%)	0 (0%)	7 (64%)	7 (64%)	7 (64%)	7 (64%)	0 (0%)	7 (64%)
Developing regions	**164**	**71 (43%)**	**42 (26%)**	**43 (26%)**	**23 (14%)**	**24 (15%)**	**2 (1%)**	**23 (14%)**	**31 (19%)**	**22 (13%)**	**13 (8%)**	**3 (2%)**	**14 (9%)**
Africa	56	21 (38%)	15 (27%)	15 (27%)	8 (14%)	8 (14%)	0 (0%)	8 (14%)	11 (20%)	9 (16%)	4 (7%)	0 (0%)	4 (7%)
Northern Africa	6	4 (67%)	3 (50%)	3 (50%)	3 (50%)	3 (50%)	0 (0%)	3 (50%)	3 (50%)	3 (50%)	1 (17%)	0 (0%)	1 (17%)
Sub-Saharan Africa	50	17 (34%)	12 (24%)	12 (24%)	5 (10%)	5 (10%)	0 (0%)	5 (10%)	8 (16%)	6 (12%)	3 (6%)	0 (0%)	3 (6%)
Latin America and the Caribbean	46	23 (50%)	15 (33%)	12 (26%)	10 (22%)	11 (24%)	1 (2%)	10 (22%)	11 (24%)	8 (17%)	6 (13%)	2 (4%)	7 (15%)
Caribbean	24	8 (33%)	2 (8%)	3 (13%)	2 (8%)	3 (13%)	0 (0%)	2 (8%)	3 (13%)	1 (4%)	0 (0%)	1 (4%)	1 (4%)
Latin America	22	15 (68%)	13 (59%)	9 (41%)	8 (36%)	8 (36%)	1 (5%)	8 (36%)	8 (36%)	7 (32%)	6 (27%)	1 (5%)	6 (27%)
Asia	42	24 (57%)	11 (26%)	12 (29%)	5 (12%)	5 (12%)	1 (2%)	5 (12%)	9 (21%)	5 (12%)	3 (7%)	1 (2%)	3 (7%)
Eastern, Southern, and South-eastern Asia	26	13 (50%)	6 (23%)	7 (27%)	3 (12%)	3 (12%)	1 (4%)	3 (12%)	6 (23%)	3 (12%)	1 (4%)	1 (4%)	2 (8%)
Eastern Asia	6	3 (50%)	3 (50%)	3 (50%)	2 (33%)	2 (33%)	1 (17%)	2 (33%)	4 (67%)	1 (17%)	0 (0%)	1 (17%)	1 (17%)
Southern Asia	9	5 (56%)	2 (22%)	3 (33%)	1 (11%)	1 (11%)	0 (0%)	1 (11%)	2 (22%)	2 (22%)	1 (11%)	0 (0%)	1 (11%)
South-eastern Asia	11	5 (45%)	1 (9%)	1 (9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Western Asia	16	11 (69%)	5 (31%)	5 (31%)	2 (13%)	2 (13%)	0 (0%)	2 (13%)	3 (19%)	2 (13%)	2 (13%)	0 (0%)	1 (6%)
Oceania	20	3 (15%)	1 (5%)	4 (20%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Official MDG Groupings													
Landlocked developing countries (LLDCs)	31	15 (48%)	11 (35%)	11 (35%)	8 (26%)	8 (26%)	0 (0%)	8 (26%)	10 (32%)	9 (29%)	6 (19%)	0 (0%)	6 (19%)
Small island developing States (SIDS)	51	12 (24%)	3 (6%)	8 (16%)	2 (4%)	3 (6%)	0 (0%)	2 (4%)	3 (6%)	1 (2%)	0 (0%)	1 (2%)	1 (2%)
Least developed countries (LDCs)	50	14 (28%)	5 (10%)	8 (16%)	2 (4%)	2 (4%)	0 (0%)	2 (4%)	3 (6%)	2 (4%)	0 (0%)	0 (0%)	0 (0%)

Source: own compilation. (Based on 2008 YB and 2009 YB.)

Table 43 2008 MRDS for all Countries, Areas and Territories, Part b: Institutional Sector, MRDS and Milestone Level Results

Data available between 2000-2009	Total number of countries $\lvert C\rvert$	2008 MRDS sectors: All tables $\lvert C_{IS}\rvert$ (h_{IS})	2008 MRDS sectors: All but one tables $\lvert C_{IS(p)}\rvert$ ($h_{IS(p)}$)	2008 MRDS: All tables $\lvert C_{2008}\rvert$ (h_{2008})	2008 MRDS: All but one tables $\lvert C_{2008(p)}\rvert$ ($h_{2008(p)}$)	Fin. acc. of NAQ Table 4.2 $\lvert C^{f}_{4.2}\rvert$ ($h^{f}_{4.2}$)	Strict milestone level: 1 or higher $\lvert C_{1s}\rvert$ (h_{1s})	Strict milestone level: 2 $\lvert C_{2s}\rvert$ (h_{2s})
All (M49) Countries, Areas, and Territories	**224**	**48 (21%)**	**58 (26%)**	**45 (20%)**	**52 (23%)**	**44 (20%)**	**130 (58%)**	**40 (18%)**
UN Member States	192	47 (24%)	56 (29%)	45 (23%)	52 (27%)	43 (22%)	126 (66%)	40 (21%)
Other Countries from M49 list	32	1 (3%)	2 (6%)	0 (0%)	0 (0%)	1 (3%)	4 (13%)	0 (0%)
Developed Regions	**43**	**27 (63%)**	**27 (63%)**	**27 (63%)**	**27 (63%)**	**27 (63%)**	**34 (79%)**	**26 (60%)**
Transition Countries	**17**	**8 (47%)**	**9 (53%)**	**8 (47%)**	**8 (47%)**	**1 (6%)**	**13 (76%)**	**1 (6%)**
Transition countries of South-eastern Europe	6	1 (17%)	2 (33%)	1 (17%)	1 (17%)	0 (0%)	4 (67%)	0 (0%)
Commonwealth of Independent States (CIS)	11	7 (64%)	7 (64%)	7 (64%)	7 (64%)	1 (9%)	9 (82%)	1 (9%)
Developing regions	**164**	**13 (8%)**	**22 (13%)**	**10 (6%)**	**17 (10%)**	**16 (10%)**	**83 (51%)**	**13 (8%)**
Africa	56	3 (5%)	8 (14%)	1 (2%)	5 (9%)	3 (5%)	27 (48%)	2 (4%)
Northern Africa	6	1 (17%)	3 (50%)	0 (0%)	2 (33%)	3 (50%)	2 (33%)	2 (33%)
Sub-Saharan Africa	50	2 (4%)	5 (10%)	1 (2%)	3 (6%)	0 (0%)	25 (50%)	0 (0%)
Latin America and the Caribbean	46	7 (15%)	9 (20%)	6 (13%)	8 (17%)	6 (13%)	24 (52%)	6 (13%)
Caribbean	24	1 (4%)	2 (8%)	0 (0%)	0 (0%)	0 (0%)	6 (25%)	0 (0%)
Latin America	22	6 (27%)	7 (32%)	6 (27%)	8 (36%)	6 (27%)	18 (82%)	6 (27%)
Asia	42	3 (7%)	5 (12%)	3 (7%)	4 (10%)	7 (17%)	28 (67%)	5 (12%)
Eastern, Southern, and South-eastern Asia	26	2 (8%)	3 (12%)	2 (8%)	2 (8%)	5 (19%)	18 (69%)	4 (15%)
Eastern Asia	6	1 (17%)	2 (33%)	1 (17%)	1 (17%)	2 (33%)	2 (33%)	1 (17%)
Southern Asia	9	1 (11%)	1 (11%)	1 (11%)	1 (11%)	2 (22%)	7 (78%)	2 (22%)
South-eastern Asia	11	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (9%)	9 (82%)	1 (9%)
Western Asia	16	1 (6%)	2 (13%)	1 (6%)	2 (13%)	2 (13%)	10 (63%)	1 (6%)
Oceania	20	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (20%)	0 (0%)
Official MDG Groupings								
Landlocked developing countries (LLDCs)	31	5 (16%)	8 (26%)	5 (16%)	7 (23%)	0 (0%)	21 (68%)	0 (0%)
Small island developing States (SIDS)	51	1 (2%)	2 (4%)	0 (0%)	0 (0%)	1 (2%)	14 (27%)	1 (2%)
Least developed countries (LDCs)	50	0 (0%)	2 (4%)	0 (0%)	1 (2%)	0 (0%)	24 (48%)	0 (0%)

Source: own compilation. (Based on 2008 YB and 2009 YB.)

2.3.2.3 Other Organizations and Economic Groups

Table 43, p. 228 and Table 44, p. 229 give the data availability for the other groups.

The reporting of 2008 MRDS is particularly low for CARICOM. The CARICOM member countries compile hardly any NA data until NL. Only four countries (20%) fulfill the "strict" milestone level 1.

The EU27 has the best availability of 2008 MRDS with almost 90% of the countries. However, this is two countries less than for 1993 MRDS. All but one of the EU27 countries fulfill "strict" milestone level 1. Yet, only 81% fulfill "strict" milestone level 2, which is 100% regarding the "relaxed" criteria. All 30 OECD member countries fulfill "strict" milestone level 1, yet only 80% "strict" milestone level 2. Chiefly due to unavailability of data until NL of the individual "non-financial" and "financial" sectors, i.e. NAQ tables 4.3 and 4.4, the fulfillment of 2008 MRDS is only 77%, i.e. 23 OECD countries. Australia, Canada, and the US report these sectors by NAQ table 4.8, i.e. as unconsolidated aggregated data of the financial and non-financial sectors, yet this does not fulfill the 2008 MRDS requirements.

Table 44 2008 MRDS for Other Organizations/ Groups, Part a: Availability of Individual Required NAQ Tables

Data available between 2000-2009	Total number of countries $\lvert C \rvert$	Institutional secor account tables of the *2008 SNA* Minimum Requirement Data Set (MRDS) until net lending/ net borrowing											
		1.3/4.1 $\lvert C_{1.3\vee 4.1(nl)} \rvert$ ($h_{1.3\vee 4.1(nl)}$)	4.1 $\lvert C^{nl}_{4.1} \rvert$ ($h^{nl}_{4.1}$)	4.2 $\lvert C^{nl}_{4.2} \rvert$ ($h^{nl}_{4.2}$)	4.3 $\lvert C^{nl}_{4.3} \rvert$ ($h^{nl}_{4.3}$)	4.4 $\lvert C^{nl}_{4.4} \rvert$ ($h^{nl}_{4.4}$)	4.8 (combined 4.3 & 4.4) $\lvert C^{nl}_{4.8} \rvert$ ($h^{nl}_{4.8}$)	4.3 and 4.4, or 4.8 $\lvert C_{4.3+4.4\vee 4.8} \rvert$ ($h_{4.3+4.4\vee 4.6}$)	4.5 $\lvert C^{nl}_{4.5} \rvert$ ($h^{nl}_{4.5}$)	4.6 $\lvert C^{nl}_{4.6} \rvert$ ($h^{nl}_{4.6}$)	4.7 $\lvert C^{nl}_{4.7} \rvert$ ($h^{nl}_{4.7}$)	4.9 (combined 4.6 & 4.7) $\lvert C^{nl}_{4.9} \rvert$ ($h^{nl}_{4.9}$)	4.6 and 4.7, or 4.9 $\lvert C_{4.6+4.7\vee 4.9} \rvert$ ($h_{4.6+4.7\vee 4.9}$)
Other Organizations/Economic Groupings													
CARICOM	20	6 (30%)	0 (0%)	2 (10%)	0 (0%)	1 (5%)	0 (0%)	0 (0%)	1 (5%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
OECD	30	28 (93%)	27 (90%)	24 (80%)	23 (77%)	23 (77%)	26 (87%)	26 (87%)	29 (97%)	18 (60%)	18 (60%)	27 (90%)	27 (90%)
SADC	15	6 (40%)	4 (27%)	6 (40%)	1 (7%)	1 (7%)	0 (0%)	1 (7%)	4 (27%)	2 (13%)	1 (7%)	0 (0%)	1 (7%)
EU27	27	25 (93%)	24 (89%)	25 (93%)	24 (89%)	24 (89%)	17 (63%)	24 (89%)	26 (96%)	20 (74%)	20 (74%)	17 (63%)	24 (89%)
G20	19	16 (84%)	16 (84%)	13 (68%)	10 (53%)	10 (53%)	10 (53%)	13 (68%)	14 (74%)	9 (47%)	6 (32%)	10 (53%)	11 (58%)
Advanced other than G-20 (excl. Taiw an Prov. of Chin	23	21 (91%)	18 (78%)	19 (83%)	15 (65%)	15 (65%)	14 (61%)	15 (65%)	20 (87%)	12 (52%)	13 (57%)	15 (65%)	16 (70%)
UN Regional Organizations													
ECA	52	21 (40%)	15 (29%)	15 (29%)	8 (15%)	8 (15%)	0 (0%)	8 (15%)	11 (21%)	9 (17%)	4 (8%)	0 (0%)	4 (8%)
Northern Africa	5	4 (80%)	3 (60%)	3 (60%)	3 (60%)	3 (60%)	0 (0%)	3 (60%)	3 (60%)	3 (60%)	1 (20%)	0 (0%)	1 (20%)
Sub-Saharan Africa	47	17 (36%)	12 (26%)	12 (26%)	5 (11%)	5 (11%)	0 (0%)	5 (11%)	8 (17%)	6 (13%)	3 (6%)	0 (0%)	3 (6%)
ECE	56	43 (77%)	41 (73%)	36 (64%)	34 (61%)	34 (61%)	21 (38%)	36 (64%)	40 (71%)	30 (54%)	31 (55%)	21 (38%)	36 (64%)
Developed regions	40	31 (78%)	30 (75%)	26 (65%)	27 (68%)	27 (68%)	21 (53%)	29 (73%)	32 (80%)	23 (58%)	23 (58%)	21 (53%)	29 (73%)
of which: South-eastern Europe transition countr	6	3 (50%)	3 (50%)	1 (17%)	2 (33%)	2 (33%)	0 (0%)	2 (33%)	2 (33%)	2 (33%)	2 (33%)	0 (0%)	2 (33%)
Commonwealth of Independent States (CIS)	11	9 (82%)	9 (82%)	8 (73%)	7 (64%)	7 (64%)	0 (0%)	7 (64%)	7 (64%)	7 (64%)	7 (64%)	0 (0%)	7 (64%)
Developing regions	5	3 (60%)	2 (40%)	2 (40%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (20%)	0 (0%)	1 (20%)	0 (0%)	0 (0%)
ESCAP	49	25 (51%)	16 (33%)	17 (35%)	8 (16%)	8 (16%)	3 (6%)	9 (18%)	12 (24%)	8 (16%)	6 (12%)	4 (8%)	9 (18%)
Developed regions	3	3 (100%)	3 (100%	2 (67%)	1 (33%)	1 (33%)	2 (67%)	2 (67%)	3 (100%	1 (33%)	1 (33%)	3 (100%)	3 (100%)
Commonwealth of Independent States (CIS)	8	6 (75%)	6 (75%)	5 (63%)	4 (50%)	4 (50%)	0 (0%)	4 (50%)	4 (50%)	4 (50%)	4 (50%)	0 (0%)	4 (50%)
Developing regions	38	16 (42%)	7 (18%)	10 (26%)	3 (8%)	3 (8%)	1 (3%)	3 (8%)	5 (13%)	3 (8%)	1 (3%)	1 (3%)	2 (5%)
Asia	26	13 (50%)	6 (23%)	6 (23%)	3 (12%)	3 (12%)	1 (4%)	3 (12%)	5 (19%)	3 (12%)	1 (4%)	1 (4%)	2 (8%)
Eastern Asia	4	2 (50%)	2 (50%)	2 (50%)	2 (50%)	2 (50%)	1 (25%)	2 (50%)	3 (75%)	1 (25%)	0 (0%)	1 (25%)	1 (25%)
Southern Asia	9	5 (56%)	2 (22%)	3 (33%)	1 (11%)	1 (11%)	0 (0%)	1 (11%)	2 (22%)	2 (22%)	1 (11%)	0 (0%)	1 (11%)
South-eastern Asia	11	5 (45%)	1 (9%)	1 (9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Western Asia	2	1 (50%)	1 (50%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Oceania	12	3 (25%)	1 (8%)	4 (33%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
ESCWA	14	10 (71%)	4 (29%)	4 (29%)	3 (21%)	3 (21%)	0 (0%)	3 (21%)	3 (21%)	3 (21%)	2 (14%)	0 (0%)	2 (14%)
ECLAC	44	31 (70%)	24 (55%)	19 (43%)	17 (39%)	18 (41%)	12 (27%)	19 (43%)	20 (45%)	14 (32%)	13 (30%)	12 (27%)	17 (39%)
Developed regions	10	10 (100%)	10 (100%	8 (80%)	8 (80%)	8 (80%)	10 (100%)	10 (100%)	10 (100%	7 (70%)	7 (70%)	10 (100%)	10 (100%)
Developing regions	34	21 (62%)	14 (41%)	11 (32%)	9 (26%)	10 (29%)	2 (6%)	9 (26%)	10 (29%)	7 (21%)	6 (18%)	2 (6%)	7 (21%)
Caribbean	13	5 (38%)	0 (0%)	1 (8%)	0 (0%)	1 (8%)	0 (0%)	0 (0%)	1 (8%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Latin America	20	15 (75%)	13 (65%)	9 (45%)	8 (40%)	8 (40%)	1 (5%)	8 (40%)	8 (40%)	7 (35%)	6 (30%)	1 (5%)	6 (30%)
Other (Asia)	1	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)	1 (100%)

Source: own compilation. (Based on 2008 YB and 2009 YB.)

Table 45 2008 MRDS for Other Organizations/ Groups, Part b: Institutional Sector, MRDS and Milestone Level Results

Data available between 2000-2009 for institutional sectors, 2004-2008 for other data	Total number of countries	2008 MRDS sectors		2008 MRDS		Fin. acc.	Strict milestone level	
		All tables	All but one tables	All tables	All but one tables	of NAQ Table 4.2	1 or higher	2
	$\lvert C \rvert$	$\lvert C_{IS} \rvert$ (h_{IS})	$\lvert C_{IS(p)} \rvert$ ($h_{IS(p)}$)	$\lvert C_{2008} \rvert$ (h_{2008})	$\lvert C_{2008(p)} \rvert$ ($h_{2008(p)}$)	$\lvert C^{f}_{4.2} \rvert$ ($h^{f}_{4.2}$)	$\lvert C_{1s} \rvert$ (h_{1s})	$\lvert C_{2s} \rvert$ (h_{2s})
Other Organizations/Economic Groupings								
CARICOM	20	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5%)	4 (20%)	0 (0%)
OECD	30	23 (77%)	23 (77%)	23 (77%)	23 (77%)	24 (80%)	30 (100%)	24 (80%)
SADC	15	0 (0%)	1 (7%)	0 (0%)	1 (7%)	0 (0%)	11 (73%)	0 (0%)
EU27	27	24 (89%)	24 (89%)	24 (89%)	24 (89%)	22 (81%)	26 (96%)	22 (81%)
G20	19	8 (42%)	10 (53%)	8 (42%)	9 (47%)	10 (53%)	18 (95%)	9 (47%)
Advanced other than G-20 (excl. Taiwan Prov. of Chin	23	15 (65%)	15 (65%)	15 (65%)	15 (65%)	16 (70%)	22 (96%)	16 (70%)
UN Regional Organizations								
ECA	52	3 (6%)	8 (15%)	1 (2%)	5 (10%)	3 (6%)	27 (52%)	2 (4%)
Northern Africa	5	1 (20%)	3 (60%)	0 (0%)	2 (40%)	3 (60%)	2 (40%)	2 (40%)
Sub-Saharan Africa	47	2 (4%)	5 (11%)	1 (2%)	3 (6%)	0 (0%)	25 (53%)	0 (0%)
ECE	56	33 (59%)	34 (61%)	33 (59%)	33 (59%)	25 (45%)	46 (82%)	25 (45%)
Developed regions	40	26 (65%)	27 (68%)	26 (65%)	26 (65%)	23 (58%)	34 (85%)	23 (58%)
of which: South-eastern Europe transition countr	6	1 (17%)	2 (33%)	1 (17%)	1 (17%)	0 (0%)	4 (67%)	0 (0%)
Commonwealth of Independent States (CIS)	11	7 (64%)	7 (64%)	7 (64%)	7 (64%)	1 (9%)	9 (82%)	1 (9%)
Developing regions	5	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (20%)	3 (60%)	1 (20%)
ESCAP	49	7 (14%)	8 (16%)	7 (14%)	7 (14%)	7 (14%)	30 (61%)	6 (12%)
Developed regions	3	1 (33%)	1 (33%)	1 (33%)	1 (33%)	2 (67%)	3 (100%)	2 (67%)
Commonwealth of Independent States (CIS)	8	4 (50%)	4 (50%)	4 (50%)	4 (50%)	0 (0%)	6 (75%)	0 (0%)
Developing regions	38	2 (5%)	3 (8%)	2 (5%)	2 (5%)	5 (13%)	21 (55%)	4 (11%)
Asia	26	2 (8%)	3 (12%)	2 (8%)	2 (8%)	5 (19%)	18 (69%)	4 (15%)
Eastern Asia	4	1 (25%)	2 (50%)	1 (25%)	1 (25%)	2 (50%)	1 (25%)	1 (25%)
Southern Asia	9	1 (11%)	1 (11%)	1 (11%)	1 (11%)	2 (22%)	7 (78%)	2 (22%)
South-eastern Asia	11	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (9%)	9 (82%)	1 (9%)
Western Asia	2	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (50%)	0 (0%)
Oceania	12	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (25%)	0 (0%)
ESCWA	14	2 (14%)	3 (21%)	1 (7%)	2 (14%)	2 (14%)	7 (50%)	0 (0%)
ECLAC	44	15 (34%)	16 (36%)	15 (34%)	17 (39%)	15 (34%)	34 (77%)	15 (34%)
Developed regions	10	8 (80%)	8 (80%)	8 (80%)	8 (80%)	8 (80%)	10 (100%	8 (80%)
Developing regions	34	7 (21%)	8 (24%)	7 (21%)	9 (26%)	7 (21%)	24 (71%)	7 (21%)
Caribbean	13	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (38%)	0 (0%)
Latin America	20	6 (30%)	7 (35%)	6 (30%)	8 (40%)	6 (30%)	18 (90%)	6 (30%)
Other (Asia)	1	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)	1 (100%)

Source: own compilation. (Based on 2008 YB and 2009 YB.)

The G20 countries show a large data gap. Once again, the other advanced economies provide more of the required data than the G20 group. Less than half of the G20 countries, i.e. only eight countries, fulfill the 2008 MRDS. Two thirds, i.e. 15 of the other advanced economies, fulfill the 2008 MRDS. Nearly all countries of the G20 and the other advanced economies fulfill “strict” milestone level 1 with 95% and 96%, respectively. The compliance with “strict” milestone level 2 is significantly lower, particularly due to missing “financial accounts” data of NAQ table 4.2. While still 70%, i.e. 16 of the other advanced economies fulfill “strict” milestone level 2, only 47%, i.e. nine of the 19 G20 countries are able to do so.

The SADC member countries show significantly lower data availability compared to the 1993 MRDS assessment. The number of countries fulfilling the MRDS criteria dropped from 40%, i.e. six countries regarding the 1993 MRDS to zero regarding the

2008 MRDS. One SADC member country is able to report all but one of the NAQ tables required to fulfill the 2008 MRDS.

2.3.3 Comparison with Results of the UN Assessment

The results of the UN assessment and our assessment are compared in this section. For the 1993 MRDS assessment, we follow closely the UN assessment methodology. Therefore, we expect similar results. Our 1993 MRDS assessment is based on the 2008 YB, published in 2009. There is no availability assessment based on the same data as our assessment. The UN's 1993 MRDS assessment of 2008 (UNSD 2008) is based on data in the 2006 YB, published in 2007. Thereafter, only short reports were prepared and published in the ISWGNA reports on NA. Following the UN's 2008 data availability assessment, the next data availability report (United Nations 2011) is based on the 2009 YB, published in 2010. The latest report (United Nations 2013) indicates that 58% of the UN Member States have at least all but one of the NAQ tables required to fulfill the 1993 MRDS criteria, and 44% fulfill 1993 MRDS.

Compared to the 2008 YB, the 2009 YB interferes less with changes of data availability than the 2006 YB. From the report based on the 2009 YB (United Nations 2011) we extract the information regarding the fulfillment of the 1993 MRDS criteria, and the availability of at least all but one of the NAQ tables to fulfill the 1993 MRDS criteria. In Table 46, we compare these results with our results.

Table 46 Data Availability of UN Member States by UN and Our Assessment

UN Member States as percentage of total UN Member States (192)	**All 1993 MRDS tables**	**All but one 1993 MRDS tables**
UN Assessment	41%	58%
Our Assessment	40%	55%

Source: United Nations (2011, 22) based on 2009 YB, published in 2010, and our assessment based on 2008 YB, published in 2009.

The UN report indicates that 41% of the UN Member States report all NAQ tables required by the 1993 MRDS criteria (cf. Table 46, col. two), and 58% of the countries report at least all but one of the required NAQ tables (cf. Table 46, col. three). Thus, the differences to our assessment are minimal, as expected.

The 2008 UN assessment based on the 2006 YB provides the most detailed results for comparison with our availability assessment. The 2008 availability assessment by the UN notes a "significant improvement in the availability" of official national accounts data (cf. UNSD 2008, 6). We also observe this improvement in data availability from the 2006 YB to the 2008 YB. The following Table 47 displays the improvements in data availability by NAQ tables and the results for "relaxed" milestone levels 1 and 2.

Table 47 Improvements in Data Availability from 2008 UN Assessment to Our Assessment

Data availability as absolute number of UN Member States	**NAQ table 1.1**	**NAQ table 1.2**	**NAQ table 2.1**	**NAQ table 2.2**	**NAQ table 2.3**	**NAQ table 1.3/4.1**	**NAQ table 4.2**	**Milestone level 1**	**Milestone level 2**
UN Assessment	151	128	154	149	108	131	78	148	129
Our Assessment	158	129	166	159	119	134	89	156	129

Source: UNSD (2008, Table 2, p. 8) based on 2006 YB, published in 2007, our assessment based on 2008 YB, published in 2009.

The comparison with the detailed UN assessment of 2008 indicates that the number of countries reporting data increased to our assessment. The "Number of countries that replied to UN-NAQ once between 2002-2006" (cf. UNDS 2008, Table 1, col. d, p. 7), i.e. at least once during the last five data reporting rounds, is 162 countries. In our assessment 170 countries have data available for at least one of the five most recent reference years. Thus, we can conclude that several countries successfully produced NA data for the first time since the 2008 UN assessment. Other countries improved the number of produced NAQ tables.

Between one and twelve additional countries reported data for the individual NAQ tables of the 1993 MRDS. The lowest increase is observed for NAQ table 1.2 "GDP by expenditure in constant prices", which according to our assessment is still only produced by 129 of the 192 UN Member States. The largest improvement is observed for NAQ table 2.1 "value added by industry in current prices". Also data for NAQ table 2.3 “output, value added and fixed assets by industry” and NAQ table 4.2 “rest of the world”, i.e. external sector data, are more frequently reported.

The comparison based on the “relaxed” milestone levels, which do not require all NAQ tables of the 1993 MRDS, indicates a substantial increase of eight countries for “relaxed” milestone level 1. That means, these countries can report data for “GDP by expenditure” and “GDP by production”. However, the number of countries that fulfill the “relaxed” milestone level 2 remained unchanged.

2.3.4 Comparison of 1993 MRDS and 2008 MRDS Assessments

In this section we compare our data availability assessments of the 1993 MRDS and 2008 MRDS. The question is answered whether data availability is indeed as good as suggested by the 1993 MRDS assessment.

Our comparison of the 1993 and 2008 MRDS assessments (see Figure 35 to Figure 40) indicates a drastically lower number of countries that fulfill the 2008 MRDS criteria. Thus, fewer countries are able to produce data of NAQ tables 4.1, …, 4.7, and 4.9, until the NL indicator, than of NAQ tables 1.1, …, 4.2.

In order to fulfill the availability criteria of NAQ tables 1.3, 4.1, and 4.2, defined for the 1993 MRDS, the NL indicator is not required. Thus, also those countries are assessed to have these NAQ tables available, which have any data available for these NAQ tables.

We compare the availability of NA data based on the condition that the NL indicator of the capital account is existent, respectively that any data is available for NAQ tables 1.3/4.1 and 4.2. This reveals substantial differences. The availability of NAQ tables 1.3/4.1 and 4.2 until the NL indicator, i.e. requiring advanced NA production capacity by countries, is much lower compared the availability of any data for NAQ tables 1.3/4.1 and 4.2.

1993 vs. 2008 MRDS Availability: UN Member States by Country Clusters, in Percent of Countries

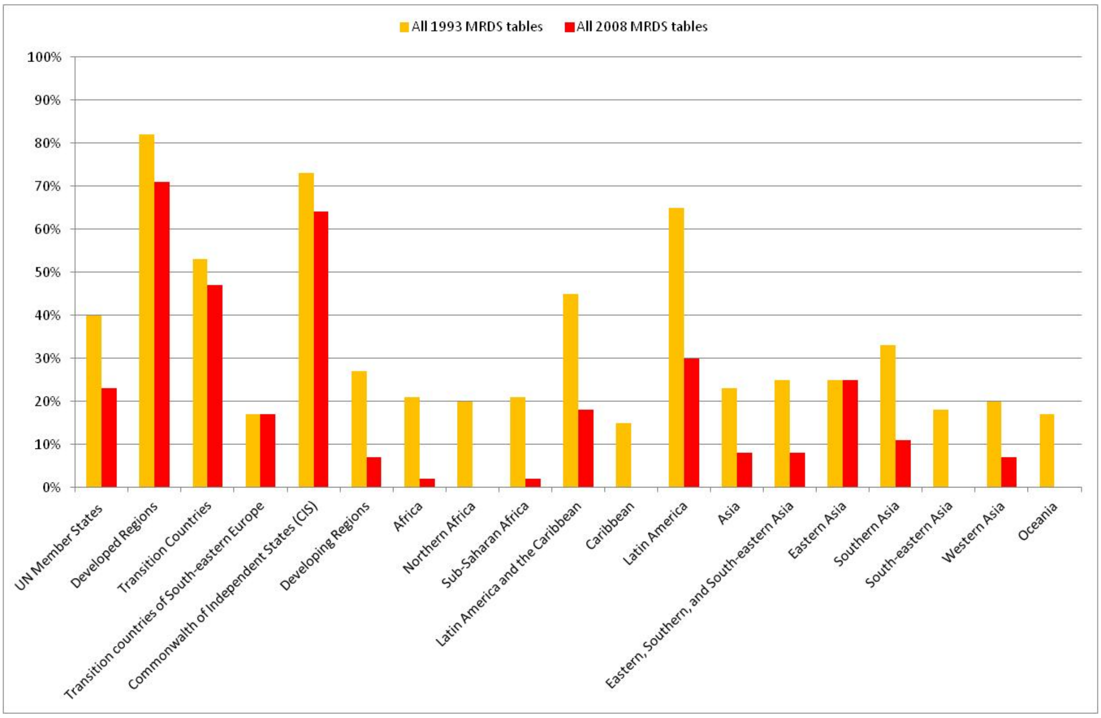


Figure 35 1993 and 2008 MRDS (UN Members), by Country Clusters, in Percent

Source: own illustration. (Based on 2008 YB and 2009 YB.)

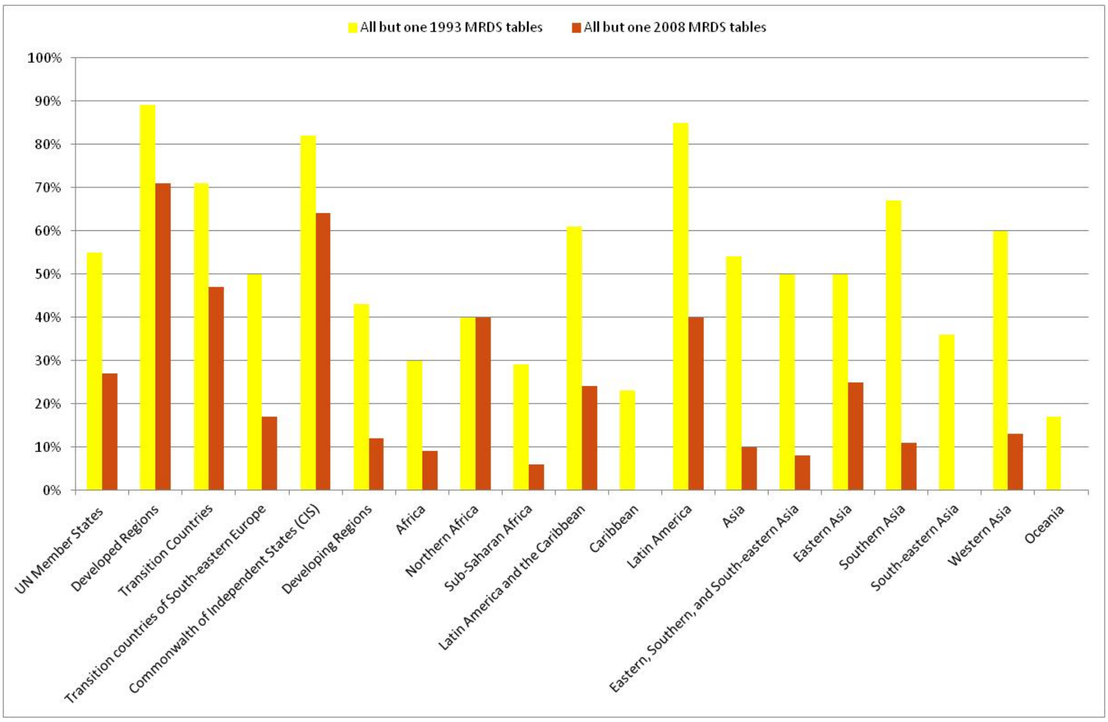


Figure 36 Partial 1993 and 2008 MRDS (UN Members), by Country Clusters, in Percent

Source: own illustration. (Based on 2008 YB and 2009 YB.)

The above two figures illustrate the actual low data availability of detailed NA data of the 2008 MRDS in the world. The first figure represents data availability by percent of countries that fulfill the 1993 MRDS vs. 2008 MRDS. The second figure shows the results regarding the availability of (at least) all but one of the NAQ tables required to fulfill these data sets. We compile these two figures for the country groups in three parts:

i) UN Member States,

ii) All countries, areas, and territories, and

iii) Organizational groups and UN regional commissions.

Figure 35 shows that the data availability fulfilling the 2008 MRDS requirements is lower than for the 1993 MRDS. Figure 35 and Figure 36 together illustrate that the number of countries being able to produce at least all but one of the NAQ tables required for the 1993 MRDS is significantly higher than the number of countries fulfilling the 2008 MRDS. Regarding the 2008 MRDS, there is not much difference between the number of countries fulfilling it and the number of countries that have at least all but one of the required NAQ tables. This indicates a large gap between countries with well developed NA capacity and those with development needs. In Figure 36 we observe for the different country clusters, that a lot of the member countries produce sufficient data to partially fulfill the 1993 MRDS, i.e. at least all but one NAQ tables needed to fulfill the 1993 MRDS criteria (for example 55% of the UN Member States). Yet, significantly fewer countries of these country clusters produce a substantial part of the 2008 MRDS (for example only 27% of the UN Member States). In Figure 35 we see that nearly all countries that fulfill a substantial part of the 2008 MRDS actually also produce all NAQ tables required by 2008 MRDS (for example by 23% of the UN Member States).

The following Figure 37 and Figure 38 show the availability of all respectively at least all but one of the NAQ tables needed to fulfill 1993 vs. 2008 MRDS for all 224 countries, i.e. including the 32 non-UN countries.

1993 vs. 2008 MRDS Availability: All Countries, Areas, and Territories by Country Clusters, in Percent of Countries

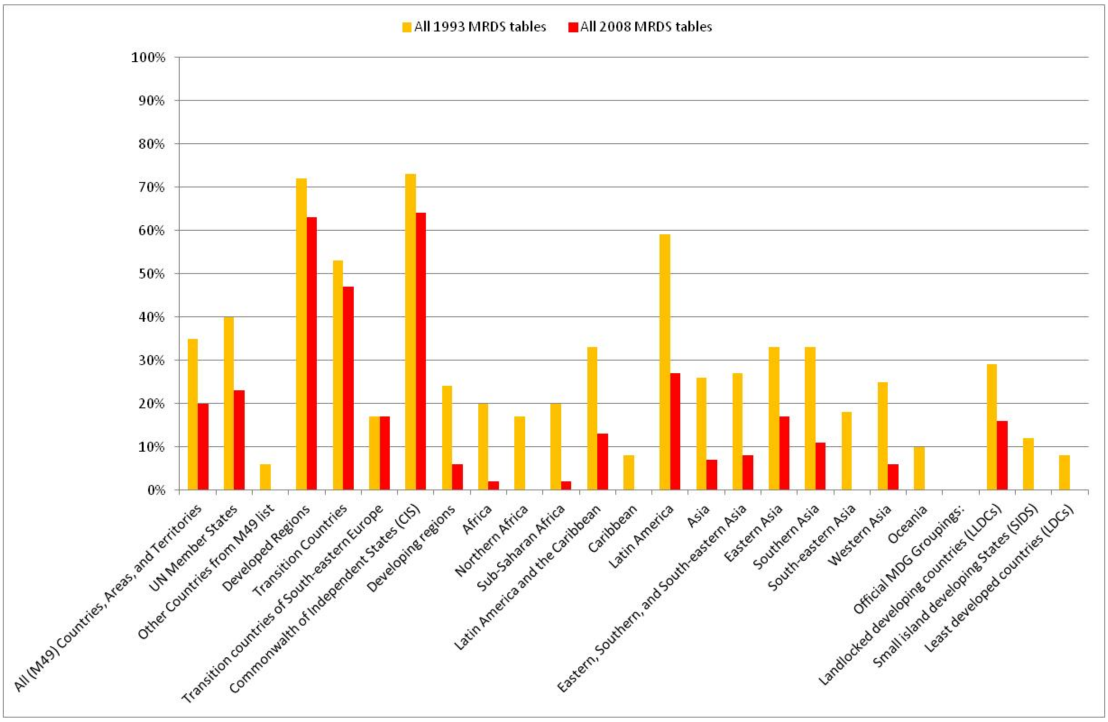


Figure 37 All NAQ Tables of 1993 and 2008 MRDS (All Countries), by Country Clusters, in Percent

Source: own illustration. (Based on 2008 YB and 2009 YB.)

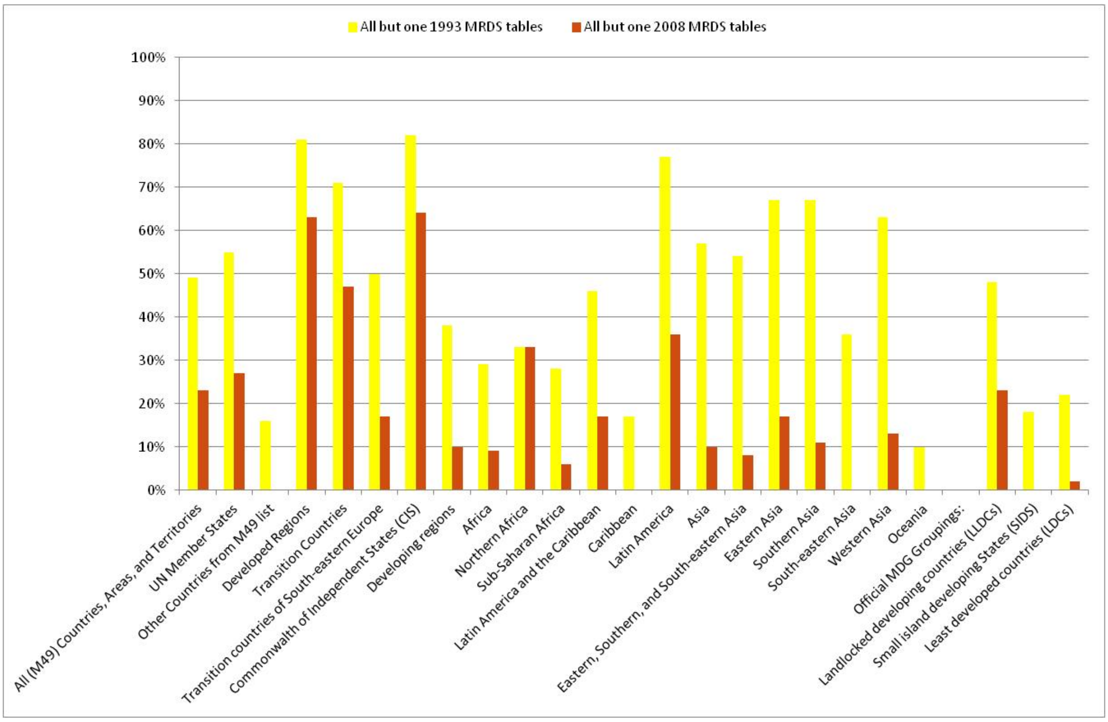


Figure 38 Partial 1993 and 2008 MRDS (All Countries), by Country Clusters, in Percent

Source: own illustration. (Based on 2008 YB and 2009 YB.)

The developed regions (cf. Figure 37, 4th pair of bars), show a surprising large gap (nine percentage points) between those countries that fulfill 1993 MRDS to those that fulfill 2008 MRDS. Since some countries only miss one NAQ table of 1993 MRDS, the gap increases regarding the percent of availability of the partial 1993 respectively 2008 MRDS (cf. Figure 38, corresponding pair of bars). Even role model countries with substantial NA capacities like Australia, Canada, and the US, do not fulfill the 2008 MRDS criteria. On the world map (cf. Figure 27, p. 198), they are illustrated as fulfilling 1993 MRDS. All three countries do not report the “financial” and “non-financial” sectors as individual sectors (i.e. NAQ tables 4.3 and 4.4., respectively). Moreover, only Australia reports the required “external” sector data, i.e. NAQ table 4.2 until NL. Most of the developed regions produce NAQ table 4.5 “government” until NL (81%) and NAQ table 4.1 “total economy” until NL (72%), but are missing some of the other institutional sectors, i.e. NAQ tables 4.y, with y = 2, 3, 4, 6, 7, 9. Thus, merely 63% of the developed regions fulfill the 2008 MRDS criteria.

The 2008 MRDS allows the production of NAQ table 4.9, i.e. the combined institutional sector of NAQ table 4.6 “households” and NAQ table 4.7 “NPISHs”, instead of the two individual institutional sectors. Would such practice for the “financial” and “non-financial” sector make a difference for the 2008 MRSD results? An analysis of each single country discloses the following results. Nine out of the 43 developed regions, i.e. 21%, produce NAQ table 4.9 for the “households” and “NPISHs” combined institutional sector instead of the individual institutional sectors. For five of these nine countries the combined NAQ table 4.9, i.e. the unconsolidated aggregation of NAQ tables 4.6 and 4.7, leads to the fulfillment of the 2008 MRDS.

For the non-financial and financial sectors, most countries that produce NAQ table 4.8, i.e. the combined institutional sector, already produce the individual institutional sectors, too. Only the three previously identified countries, Australia, Canada and the US, would be affected if NAQ tables 4.3 and 4.4 were not required individually. Australia would be reclassified from missing two NAQ tables of the individual sectors to missing none of the requirement criteria and thus fulfilling the 2008 MRDS. Canada and the US, still missing NAQ table 4.2 until NL, would be reclassified to fulfilling all but one of the 2008 MRDS criteria.

The following figures present the fulfillment of 1993 and 2008 MRDS for selected organizations and country groups, and by UN Regional Commissions.

1993 vs. 2008 MRDS Availability: Organizations and Groups by Country Clusters, in Percent of Countries

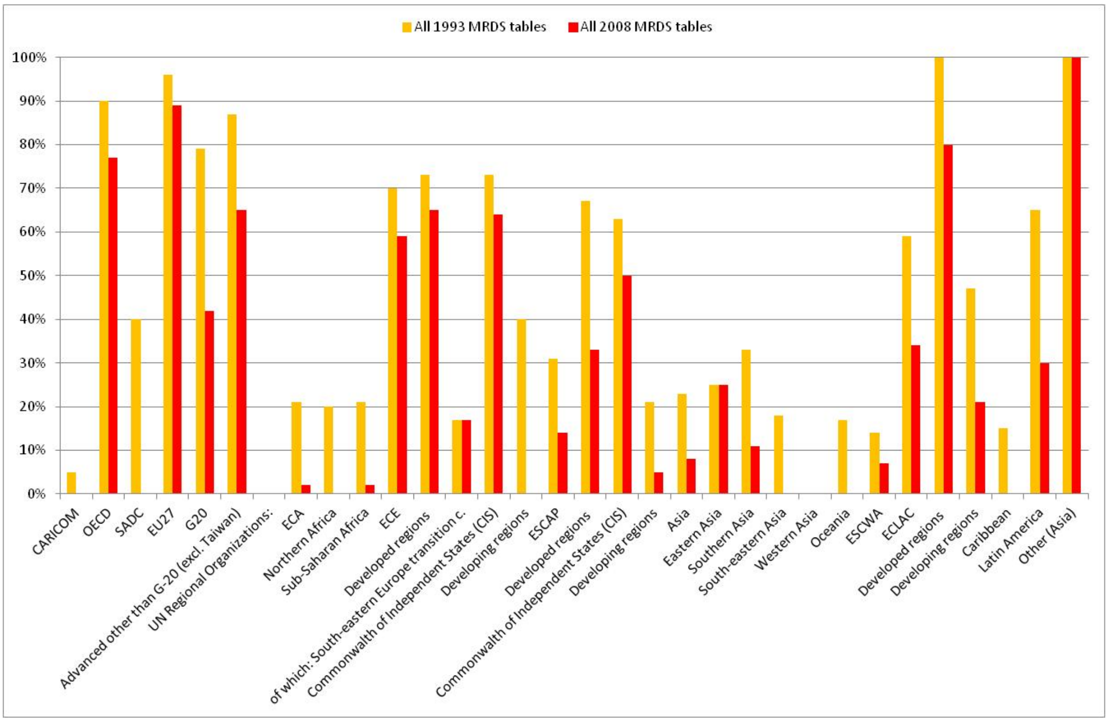


Figure 39 1993 and 2008 MRDS (Org. and Groups), by Country Clusters, in Percent

Source: own illustration. (Based on 2008 YB and 2009 YB.)

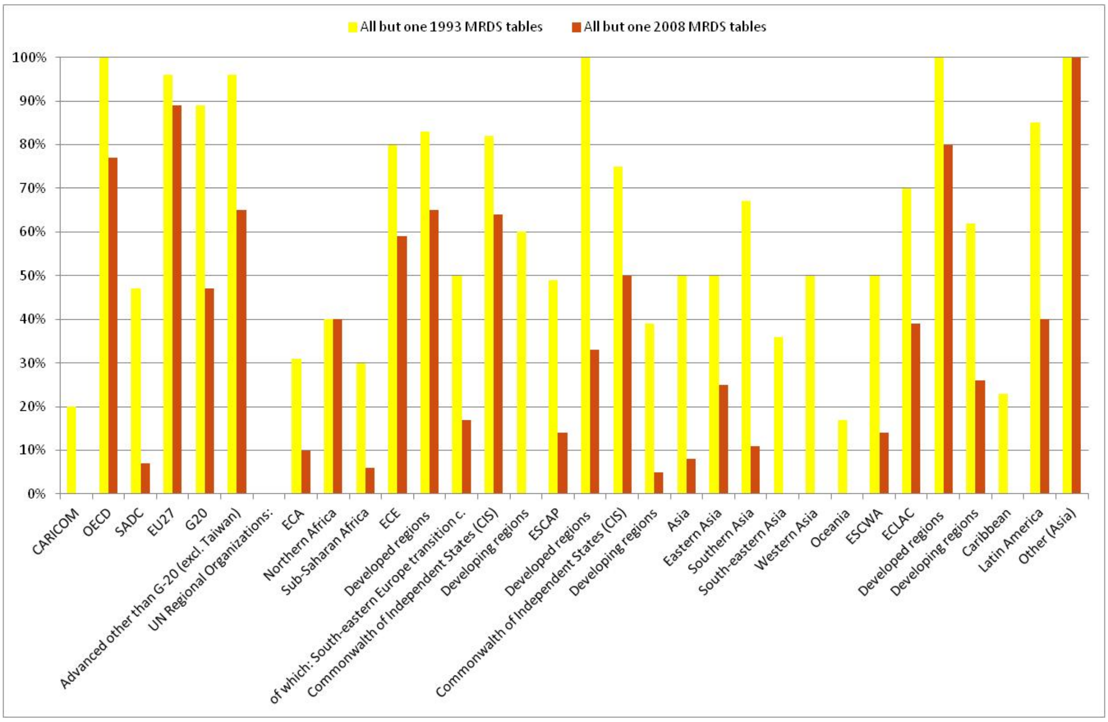


Figure 40 Partial 1993 and 2008 MRDS (Organizations and Groups), by Country Clusters, in Percent

Source: own illustration. (Based on 2008 YB and 2009 YB.)

2.4 Conclusion

The main advantage of our improved methodology of the 1993 and 2008 MRDS assessment is that it can be carried out simply based on the data available in the UN database. Other advantages like the improved clustering of country groups are already achieved with our timeliness assessment.

Compared to the UN assessments, which only assess the data availability for a few country groups against the 1993 MRDS and the "relaxed" milestone level criteria, our 2008 MRDS assessment reveals a significantly worse data availability.

Our figures and tables presenting detailed results regarding the 1993 MRDS assessment show that the NA data availability of 1993 MRDS for UN Member States is generally good for the developed regions (> 80%) and for most of the transition countries (> 50%). Yet, the data availability of 1993 MRDS for developing regions is rather weak (< 30%).

The extension of 1993 MRDS to 2008 MRDS, by inclusion of more NAQ tables and requirement of NA data until the NL indicator of NAQ tables 3.1, 4.1, …, 4.7 and 4.8 shows large data gaps regarding these NAQ tables. (Less than 25% of UN Member States fulfill 2008 MRDS.) Even regions and organizations where member countries show good results for 1993 MRDS have large data gaps regarding 2008 MRDS.

The "strict" milestone level 2, which requires the rarely available "financial account" of NAQ table 4.2 (22% of UN Member States), is primarily responsible for the decrease of countries being able to comply with the "strict" milestone level 2.

The best data availability among international and regional organizations and economic groups is achieved by the group of the EU27 countries, closely followed by the OECD and the non-G20 advanced economies.

Because of the importance of "financial accounts" and "balance sheets", these data should be produced by countries and be reported to IOs for publication. However, these data are not included in the 2008 MRDS.

Data gaps need to be addressed at the country level. The NA data produced by NSOs needs to be improved regarding the compilation of the individual institutional sectors (i.e. NAQ tables 4.1, …, 4.7) and also by more depth of the institutional sector data, i.e. until the NL indicator of the capital account.

3 Revision Analysis (Reliability Study)

This chapter investigates the current methodology used for the analysis of revisions suggested by OECD and used by researchers, NSOs and IOs. Based on the observed deficiencies, we suggest improvements to provide better information about the actual revision needs of first released GDP estimates. Our assessment of the revision needs of first published GDP estimates is presented in chapter 3.4, p. 263.

3.1 Introduction

In order to answer the question, how reliable the initial released, i.e. first estimates, of GDP are for different countries, we compute the revision measures MAR (cf. DEF. 19, p. 46) and RMAR (cf. DEF. 20, p. 47). We make an assessment of the revision needs of first published GDP estimates of 30 OECD countries and four non-OECD G20 countries.

To the best knowledge of the author, this is the largest cross-country comparison. Earlier international comparisons assessed between five and 20 countries the most (cf. Di Fonzo 2005b; Tosetto and Lequiller 2006, McKenzie 2006a and 2006b, McKenzie and Adam 2007, Tosetto and Brackfield 2009, and Neumayr 2010).

In the following the current methodology used for revision studies by OECD is checked, the limitations and deficiencies are described, and suggestions are made for improvements. The aim of our revision analysis is to provide an up-to-date picture of the revision needs of first released NA data for each country. Further, the results of the countries are compared, since revision analysis is a valuable tool for international comparison.

Compared to previous assessments our improvements are as follows:

- We only consider data prepared by the same SNA version (thus excluding revisions across different SNA versions),
- we exclude real estimated data (cf. DEF. 6, p. 20) and data gaps,
- we analyze more recent data,
- we include 14 countries that were not considered in previous international assessments,

- we use Year-on-Year (YoY) growth rates (cf. DEF. 21, p. 48), instead of Quarter-on-Quarter (QoQ) growth rates (cf. DEF. 23, p. 49), since YoY growth rates are best for international comparison (cf. Cheem 2003),
- we assess more revision intervals, i.e. revisions of the first released vintage (DEF. 13, p. 31) to a later vintage, i.e. of hither maturity (cf. DEF. 14, p. 34); as vintage at later revision time points we consider one sub-annual, three annual and the latest vintage,
- we have a higher maturity of the "latest" vintage (cf. DEF. 17, p. 41), which now refers to data released at least four years after the first publication.

Our results of reliability of the initial released GDP estimates, calculated through our improved methodology, support economic analysts, decision makers and policy makers who use first released GDP estimates.

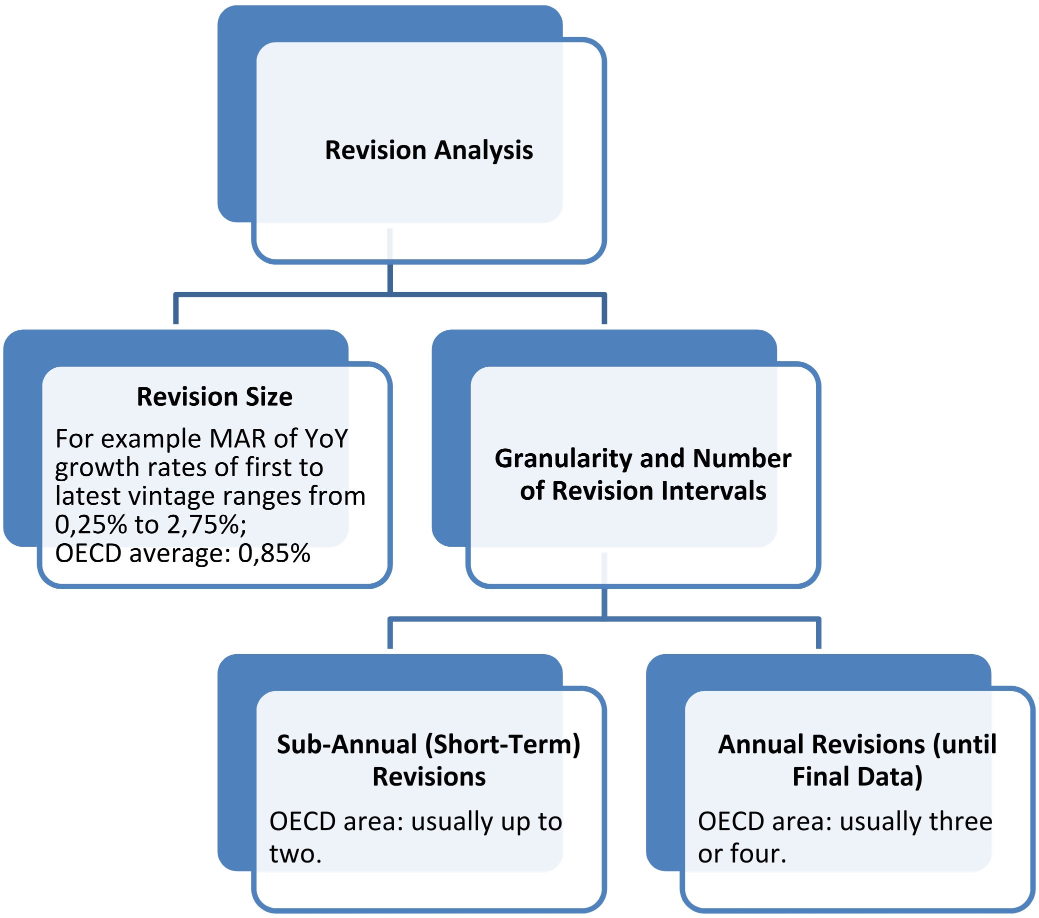


Figure 41 Revision Analysis

Source: own illustration

A main finding from our improved revision study is that the size of revisions are substantially lower than those of previous studies based on the OECD MEI database. While the reliability of first released GDP data is generally much better than before, there are large discrepancies between countries. For example the size of MAR of YoY growth rates of first to latest vintage ranges from 0,25% to 2,75%. The number of revised vintages, and the time period until data become final (cf. DEF. 16, p. 40) varies, too. For OECD countries it usually takes 3-4 years until data are declared final.

A global revision analysis for NA, similarly to the assessments of the timeliness and of the availability of ANA, based on the data reported to the UN, can unfortunately not be done. The UN database cannot store different vintages of data. Thus neither the numbers of revisions until data are final can be analyzed nor the size of revisions. For such an analysis the levels of previously reported data (via the NAQ) would need to be retained, i.e. the preliminary vintages of data would need to be maintained by the database instead of being replaced by later vintages. In computer science such retention of data is referred to as version control or revision control, which allows tracking changes of the data in the repository, see for example Microsoft Computer Dictionary-5th ed. [Microsoft Press] (2002) or TechTerms.com (2011).

The OECD Main Economic Indicators (MEI) Original Release Data and Revisions Database[97] facilitates revision analysis as it provides the required data. The analyzed time series of GDP in constant prices contains the levels of GDP to each quarter as per the monthly MEI editions. Thus it allows tracking the changes to each quarter at later editions. Further, the standards for data published by OECD facilitate cross-country comparison of the results of the assessment (cf. McKenzie 2006b).

OECD, IMF, and other IOs encourage their member countries to install "real time databases" also referred to as "revisions database" and provided guidelines how to build, maintain and use datasets suitable for revision analysis, e.g. Di Fonzo (2005b). Yet, not all countries have installed such databases.

NSOs or IOs, that already use databases to store and process NA data would not need to build such special databases separately if their database allows the retention of different vintages. Results of the complete revision analysis could be computed by functions of a database systems.

[97] Online at: http://stats.oecd.org/Index.aspx?querytype=view&queryname=206#.

3.2 Assessment Methodology and Deficiency of the OECD Approach

The following sections describe the OECD methodology of revision analysis and explain the limitations and deficiencies of this methodology of those studies that used the OECD MEI database.

Data of constant price quarterly GDP (levels), by editions of the MEI publication, are made available through the OECD MEI database. Excel spreadsheets[98] for revision studies by Di Fonzo (2005a) are utilized to calculate growth rates from the GDP levels and different measures for the revision analysis. The data for each country need to be processed manually, meaning multiple countries cannot be assessed at once, but only country by country. For description of the database and spreadsheets for revision analysis refer to Di Fonzo (2005a). For a more detailed description of measures that can be produced by these spreadsheets refer to McKenzie and Gamba (2008).

Reliability of NA data (cf. DEF. 11, p. 29) is analyzed based on the absolute change between the first estimate and the later (or final) result, e.g. Statistisches Bundesamt (2011). The OECD Excel spreadsheets for revision analysis by Di Fonzo calculate revision measures (e.g. MAR and RMAR) of the first estimates (t_0) of Quarterly National Accounts (QNA) data, compared to the revised later estimates published:

- five month after the first estimate (t_0+5M), i.e. the sub-annual revision,
- one year after first (t_0+1Y),
- two years after first (t_0+2Y),
- the latest published estimates (t_0+end) [which is not necessarily the actual "final" data].

Further calculated are the needs of revision between the following vintages:

- t_0+5M to t_0+1Y; i.e. sub-annual revision to one year after first vintage
- t_0+1Y to t_0+2Y, i.e. one year after first to two years after first vintage.[99]

[98] Available as download via http://www.oecd.org/std/automatedprogramstoperformrevisionsanalysis.htm, last accessed 28-10-2013.

[99] The index r of the revision measures e.g. MAR_r, and $RMAR_r$ identifying the later vintage, refers to the given specific revision time points. Ditto for the revision interval [i,r] between any two vintages.

Limitations of Internationally Published Data

The data of the OECD MEI database might differ from nationally published data as the NSOs compile the data according to OECD requirements. These requirements make the data more appropriate for international comparisons.

The country data in the OECD MEI database are either seasonally adjusted, which may constantly show revisions, or non-seasonally adjusted. Data of some countries even shifts between seasonally adjusted and non-seasonally adjusted data. Such shifts affect the results. Seasonally adjusted QNA data are less suited for international comparison (cf. Cheem 2003).

Shortcomings of the OECD MEI database related to the timing of published first estimates and subsequent revisions are:

- The timelines of data in the MEI database reflects the timeliness of publishing data at the international level. The timeliness of the publication of the initial estimate and subsequent revisions depends on the timeliness of reporting by the NSOs to the OECD and the time needed by the OECD for processing and disseminating the data.
- The timeliness of data publication at OECD possibly varies between countries, although the data are produced with the same or similar timeliness by NSOs.
- In case of late reporting, delays in data processing, or failure to report a vintage, the recorded values in the MEI edition may show the revision with delay or may show no revision at all if the latest reported value is re-published. Such cases affect the results of the revision analysis of the considered revision interval.
- The maturity of the first published estimate, i.e. the time between the end of the reference period and the production of the corresponding data by the NSO, is not indicated and cannot be concluded from the time of publication by the OECD. For instance the maturity of the initial estimate can be four weeks after the reference quarter as is the case for the US and six weeks in case of Germany. The data are published by OECD two months later for the US and three months later for Germany in the corresponding MEI edition.

3.2.1 **Time Period**

This section discusses the time period of analysis covered by previous studies.

The revision measures may not provide good information about the revision needs if the analyzed time period covers comprehensive revisions. (For our definition of regular and comprehensive revisions refer to part II, chapter 3.1.1, p. 37.) The reasons are that the observed data of a later vintage might be affected by a comprehensive revision and that such types of revisions may lead to substantially different revision patterns. Comprehensive revisions might also change the results of the cross-country comparison.

The *time period* for analysis is often *too long* to allow for a sensitive analysis. While outliers affect the results of a short period, the longer period has the risk of comprehensive revisions. Depending on the data availability the previous revision studies that utilized the OECD MEI database cover a minimum of five years up to a maximum of 14,5 years (58 quarters) per country.

All time periods start between 1994 and 1996 and thus prior to the introduction of ESA95 for Eurostat countries and of 1993 SNA for the other OECD countries. The introduction of a new SNA version marks a major change in methodology. Further, the underlying data used in previous studies seem to include “estimated values” and other (sometimes substantial) revisions due to the introduction of new data, concepts or methods. Thus the analyzed time periods of previous studies cover “major” revisions that could account for relative large sizes of revisions and substantially different revision patterns between the single regimes. The change in currency for all Euro-area countries also falls into the analyzed time periods (because the OECD database is prepared from data published in the monthly MEI issues, hence the data prior to the currency change were not converted to euro). This change has no impact on the analyzed growth rates and thus does not interfere with the results of the revision measures.

Typically, the results of revisions studies carried out by NSOs interfere with comprehensive revisions, too, as the observed time periods carry on across multiple methodologies and changes of source data. However, also NSOs criticized this deficiency of the studies based on OECD data. For some countries, methodological improvements that were introduced with a comprehensive revision lead to substantially smaller total revision needs. In other cases the comprehensive revisions lead to large changes of the GDP levels. For example the Netherlands commented on the study that it should distinguish between regular and comprehensive revisions, or Japan remarked that the size of revisions observed in their current regime are smaller than found by the revision studies that analyzed the full time period of available revisions data (cf. Tosetto and Lequiller 2006).

The time period used by the German Statistical Office (Statistisches Bundesamt 2011) begins with the time of ESA 95 introduction, thus excludes the major methodology change to ESA95. The recent revision study of the German Central Bank (Deutsche Bundesbank 2011), utilizing the banks own revision data, covers this major change. Yet the latter cuts off all data computed by a different methodology or source data. This leads to other deficiencies (explained in chapter 3.3.1, p. 256).

3.2.2 Vintages after a Comprehensive Revision

The vintages which interfere with a comprehensive revision are explained in this section.

Comprehensive revisions, i.e. the introduction of new concepts, methods or source data (in the following only reefed to by changes in methodology) impacts the subsequently revised vintages of multiple reference periods. New vintages, whether revisions to released (preliminary) estimates of previous reference periods or the first estimate of a new reference period, are calculated according to the new methodology.

"Provisional and final statistics are normally based on the same concepts and data collection methods" (cf. Eurostat 2003b, 17). Revision studies typically omit such breaks and calculate the total size of revision across comprehensive revisions.

As of the time point of the comprehensive revision, the later vintages do not match the methodology of the first released vintage. Depending on the lengths of the regular revision cycle to compute the "final" vintage, different numbers of reference periods with preliminary data are affected. This concerns all vintages at the end of the revision cycle for those reference periods that were started, but not finalized, under the old methodology.

Example: The annual GDP needs four years to become final, i.e. the first published annual GDP is followed by three further annual revisions. Thus if a new methodology is introduced at the time point of computation of a new annual GDP estimate, there are three reference periods with preliminary GDP estimates, which have not completed the regular revision cycle. Figure 42 illustrates the case of annual NA data and the annual revisions of GDP growth rates, i.e. at revision time points one, two and three years after the first published growth rate. The introduction of a new methodology happens in December 2012.

Vintage \ Reference year	2008	2009	2010	2011
Sep 09	**2,74**			
Sep 10	2,96	**0,62**		
Sep 11	2,33	0,92	**3,25**	
Sep 12	3,00	0,90	3,05	**1,39**
Dec 12	2,81	0,86	2,96	1,87
Sep 13				
Sep 14				
Sep 15				

Legend: Change of methodology in December 2012.
Cell color “green”: vintages compiled by the same “old” methodology.
Cell color “yellow”: data of September 2012 vintage recompiled by new methodology.
Cell color “red”: vintages of revised data are affected by the change to the new methodology.
Growth rates in bold font: first published estimate.

Figure 42 Vintages of Annual GDP Growth Rate that Interfere with Methodology Changes

Source: Own illustration

In Figure 42 the annual GDP growth rates in bold font refer to the first vintages for the reference years 2008, …, 2011. Cell color = “green” marks the vintages of data that are computed by the old methodology. For the reference years 2009, 2010, and 2011, that have not finished the regular revision cycle, the vintages until data will be finalized are marked by cell color = “red”. These vintages will be produced by the new methodology. Thus, the final vintage of reference year 2009 is impacted; of 2010 the last two vintages, and of 2011 all three annual revisions following the first vintage.

- For the reference year 2011 the first GDP figure is released in September 2012 (cf. last col., cell color = “green”, growth rate in bold). This vintage precedes the introduction of the new methodology in December 2011. In December 2011, the GDP time series is recalculated (cell color = “yellow”) utilizing the new methodology. Thus, the first estimate of annual GDP of 2011 is compiled utilizing the old methodology, while the revised annual GDP vintages and corresponding growth rates at revision time points one, two and three years later (final vintage) will be compiled by the new methodology. These revised vintages will be affected by changes of the level of GDP and changes of the GDP growth rate that result from the new methodology (color = “red”).

- For the first estimate of the 2010 annual GDP that was produced in September 2011, the first revision in September 2012 was still a “regular” revision. Yet, the revised vintages that will be produced two and three years after the initial estimate, i.e. in September 2013 and the “final” data in September 2014, are impacted by the change to the new methodology.
- Similarly to above, regarding the annual GDP of the reference year 2009, only the “final” GDP is affected by the methodology change.
- The first estimate of the 2008 annual GDP was produced in September 2009, i.e. more than three years prior to the introduction of the new methodology. Thus, all preliminary vintages and the “final” GDP are compiled by the same methodology. The changes of the first to the “final” vintage are due to regular annual revisions (cf. cell color = “green” for reference year “2008” and vintages “Sep 09, …, Sep 2012”). The December 2012 vintage that recalculates the “final” data by the methodology is affected by changes of the size of GDP and the corresponding growth rate due methodology changes.

The revision time point referring to the “latest” available vintage (given it is at least four years after the first vintage) captures the “final” data, if the regular revisions rounds end with the four years later vintage. If a comprehensive revision happens and “final” data are recalculated by the new methodology, the results of the “latest” vintage are affected by the methodology changes. The GDP growth rate corresponding to the “final” GDP of reference year 2008 is 3,00% (cf. Figure 42, reference year “2008”, vintage “Sep 12”). We observe a decrease of 0,19 percentage points to 2,18% (cf. vintage “Dec 12”)) of the growth rate based on the GDP produced by the new methodology . With other words, the “latest” available data indicates a 0,19 percentage points smaller need of revision, than the actual need of revision of the first published growth rate until “final” data.

3.2.3 Time Series Versions

This section identifies a lack of the OECD database to discriminate records computed by different concepts, methods, or source data (i.e. different NA time series versions).

The monthly MEI publication and the corresponding database do not indicate changes of the SNA version. The different series are not separated by version numbers. This would be most useful. Moreover, the MEI database does not even provide metadata to point out or explain methodological changes or changes in the source data (other than a metadata section about the changes between non-seasonal and seasonal adjustment and the presentation at annual or quarterly levels). Thus, it is not clear which data were computed by the same methodology and data sources and when a comprehensive revision took place. The time point of changes can sometimes be concluded from wide ranging changes of back-data beyond the vintages of the regular revision cycle. Yet, in cases like seasonally adjusted data, all back data may have been revised with every release of new QNA data. Therefore it is difficult to detect those revisions in the OECD MEI database that are due to a comprehensive revision.

Consequently the time period and range of considered vintages cannot easily be limited to only reflect the desired data for the analysis, i.e. to exclude all undesired comprehensive revisions. To do that, the user would need to obtain additional information (from the NSOs). This partially contradicts the benefit and purpose of the OECD MEI database, which is to provide users with data that facilitates revisions analysis and computation of comparable results for countries.

3.2.4 Revision Intervals

In this section we discuss deficiencies of the current assessment methodology related to revision intervals, i.e. the time that elapsed between the first published estimate and the later revision time points.

Effects of Comprehensive Revisions on Revision Analysis

The following describes the different types of effects on the summary revision measures from comprehensive revisions.

The size of revision observed by a revision interval across a comprehensive revision (i.e. from the first released vintage to the vintage at time point of the comprehensive revision and the subsequent vintages) reflects two kinds of revisions:

i) regular revision (i.e. more complete data, until data are "final"),

ii) change of methodology or changed source data.

In some cases comprehensive revisions may have no effects on the results of the revision analysis. This typically applies for analyzing growth rates. For example an upwards shift of the whole GDP time series by 10 %, does not change the GDP growth rate. A comprehensive revision only affects the results of revision measures of growth rates, if it changes the underlying growth rates. The size of revision due to revisions data across comprehensive revisions may be affected as follows:

- The level of the analyzed indicator (i.e. GDP) is affected by the size of the level change that results from the comprehensive revision (e.g. change regarding the definition of the boundary of GDP). Depending on the direction of the comprehensive revision, the total size of revision might be above or below the size of the regular revision.
- The growth rates of the time series may change due to the comprehensive revision. Thus the growth rate computed based on the new methodology would over- or underestimate the actual revision needs of the first published growth rate.
- Depending on the type of analysis (level, growth rates, and revision interval), the impact is likely only small and ignorable for the size of the calculated revision measure (MAR or RMAR) if the levels of the analyzed indicator, the size of revision of the initial estimate, and the growth rates of the time series did not significantly change between the two methodologies.
- In case different methodologies lead to significantly different sizes of revision of the first estimate, the average measure of the individual regimes does not provide appropriate information about the current size of revisions that can be expected.

Deficiency of the Revision Interval to the "Latest" Vintage

In this subsection we discuss the revision between the first vintage and the "latest" vintage.

The assessment of the "latest" available estimate has the deficiency that the "latest" vintage has different "maturities". The "latest" vintage might for example refer to the first vintage providing the initial GDP estimate, released only three month ago, or to the "final" GDP produced three years after the first vintage.

The time period between the "first" and the "latest" vintage can be much longer than the time period needed until data are "final" if a comprehensive revision happened, that included the revision of back data. The longer the analyzed time period, the more comprehensive revisions may fall into this period. For example if we observe a 14 year period and all back data are now revised with a comprehensive revision, then the "latest" vintage is produced 13 years after the first vintage.

Comprehensive revisions naturally only affect a limited number of preliminary estimates. The "latest" vintage often interferes with a comprehensive revision. The "final" vintages are typically back-calculated to provide longer time series of data according to the current compilation methodology and data sources.

Though back calculations of "final" data change the size of revision of the historic data, they are important to provide comparable time series, e.g. for historic analysis of growth rates. The "latest" vintage provides valuable information for example if the source data of the originally produced "final" data or the compilation process included errors. In that case the "latest" vintage would provide a correction of an earlier mistake after the regular revision cycle had passed.

In the studies based on the MEI revisions data, the "latest" available vintage (i.e. "t_0+end") typically refers to the vintage that is published at least three years after the first vintage. Yet, this vintage refers in many cases to a still preliminary GDP figure, i.e. the data are not yet "final" since they did not complete the regular revision cycle. Of older reference periods, the "latest" vintage may have been though multiple major revisions. In the latter case, the first and the final vintage of GDP may have been computed using significantly different concepts, methods, and source data.

The more comprehensive revisions the initial vintage was subject to, the higher the total impact from these revisions. Though the effects of comprehensive revisions may offset each other, the general tendency of methodology changes is to revise GDP levels upwards, cf. McKenzie (2006a).

Taking into account the outlined deficiencies and impacts of comprehensive revisions on the "latest" vintage, the value of information provided by the "latest" vintage is questionable. Though it is also included in our study, the revision interval chosen for cross-country comparison is the first to the three years later vintage. This revision interval also allows comparison of our results to those of other studies.

Observed Revision Intervals vs. National Revision Cycles

This subsection explains the deficiencies of the number and timing of vintages used for the analysis of different revision intervals. These are mainly the

- failure to capture the actual annual revisions of NA data at the country level,
- missing of a revision interval that captures the "final" data if NA are finalized by the country with the fourth annual revision, and
- missing of the short term revision in previous studies.

National agencies that undertake revision studies publish at least the total revision needs until "final" data or the "latest" vintage and usually also the annual revision intervals. Some also included the first quarterly revision. The number of analyzed revision intervals and time points of observed vintages are determined by the number and frequency of revisions during the regular revision cycle. As a different approach, some NSOs analyze the revision intervals from one preliminary vintage to the following vintage, i.e. focusing on the step-by-step revisions.

Using the OECD recommended spreadsheets for revision analysis the first revision time points refers to the vintage five month after the first vintage. This is not included by the majority of the previous studies. Yet, it is important for analyzing short term revisions. The three years later vintage has the highest precisely defined maturity in previous studies. The undefined highest maturity is the "latest" vintage (with a maturity of at least three years after the initial release.) Thus, a vintage that refers to the maturity of at least four years is missing. Such vintage is needed, as currently the "latest" vintage interferes with preliminary data (of the three years later vintage) if data are only finalized with the fourth annual revision.

Table 48 Time Points of Calculation of Size of Revision

Revision time point	**t_0+5M**	**t_0+1Y**	**t_0+2Y**	**t_0+3Y**	**t_0+end**
Description	Revision after first quarter	First annual revision.	Second annual revision	Third annual revision	"Latest" published* vintage.

t_0 = first vintage.
t_0+5M = vintage five month after first vintage.
t_0+1Y = vintage one year after first vintage.
t_0+2Y = vintage two years after first vintage.
t_0+3Y = vintage three years after first vintage.
t_0+end = the latest available vintage.
*In previous studies at least three years after first vintage; in our study at least four years after first.

Source: Own illustration

The exact time points of revisions may vary between countries and even for a specific country (in case of delays or different schedules for regular and comprehensive revisions).

With the annual revision, all quarters to a given reference year are simultaneously revised. NSOs are in the position to prepare results for the actual annual revision intervals. In the OECD recommended revision spreadsheets the observed revision time points are set exactly to 5, 12, 24, 36 month after the first publication and the latest available data. Thus, the maturity of the observed data may in fact vary between the maturity of the intended annual revision and that of the previous annual revision.

Assume the annual revision at the national level is carried out at the time point when the 4^{th} quarter has a maturity of one, two or three years. Consequently, regarding the data captured by the OECD recommended revision spreadsheets, all quarters other than the 4^{th} quarter actually still refer to the maturity of the revision previous to the intended annual revision. The QNA vintages of the 3^{rd} to 1^{st} quarter need an additional quarter of revisions for each quarter prior to the 4^{th} quarter to capture the time point of the annual revisions. However, since the first "annual revision" of the 4^{th} quarter is usually prepared by countries already six months after the first release (i.e. in September), only the 1^{st} quarter would not refer to the annual revision at that time.

The time period until the first estimate of reference period is published by OECD may vary, imagine for example variations due to delays in publishing. The later vintages captured by the revision spreadsheets depend on the time point publication the first vintage. In rare cases a delayed publication of the first vintage may affect the maturity of the observed sub-annual vintage, for example if the annual revision is carried out by the country six month after the first vintage and the corresponding data are published by OECD without any delay. In such case the revised vintage referring to the first annual revision is published by OECD five month after the delayed fist published vintage.

Figure 43 illustrates by example the deficiency of the revision spreadsheet to capture the actual annual revisions at the country level. The figure shows the third annual revision. The first annual revision is carried out at the country level in August 2009, which six month after the first release of the data for the 4^{th} quarter of the previous reference year. Moreover, this example demonstrates the deficiency to only observe the third annual vintage and a “latest” vintage with a maturity of at least three years. In case NA data are finalized with the fourth annual revision, neither vintage captures the actual “final” data.

Revision time point / Reference Quarter	Mai 08	Aug 08	Nov 08	Feb 09	Aug 09	Aug 10	Mai 11	Aug 11	Nov 11	Feb 12	Aug 12
Scheduled Annual Revisions					1st Annual revision	2nd Annual revision		3rd Annual revision			4th Annual revision
2008Q1	*Initial Q1*						t_0+3Y Q1				
2008Q2		*Initial Q2*						t_0+3Y Q2			
2008Q3			*Initial Q3*						t_0+3Y Q3		
2008Q4				*Initial Q4*						t_0+3Y Q4	

Legend: 3^{rd} annual revision at the country level in Aug 2011 of reference year 2008 data (all four quarters).
Cell color “yellow”: time point of 3^{rd} annual revision at the country level of all four reference quarters of 2008.
Cell color “red”: 36 after first vintage in revision spreadsheet, failing to capture the actual 3^{rd} annual revision.
Cell color “green”: data captured by revision spreadsheet corresponds to the actual 3^{rd} annual revision.

Figure 43 Three Years Later Vintage (t_0+3Y) versus Third Annual Revision

Source: Own illustration

Example 1: Figure 43 illustrates the case were the annual revision is carried out in August of each year that follows the end of the reference period (2008). The revision time point three years after the initial estimate (t_0+3Y) does not capture the actual 3^{rd}

annual revision of QNA data for the 1st quarter of the reference year. If the data were finalized with the 3rd annual revision, the "three years later" (t+3Y) vintage only captures "final" data for the quarters Q2, Q3 and Q4. The "three years later" vintage for the first quarter of 2008 still refers to the same maturity of the vintage that was prepared at the time point of the 2nd annual revision.

Example 2: In Germany "final" annual NA data are computed with the 4th annual revision, 44 month after the end of the reference period, i.e. regarding reference year 2008 data in August 2012 (illustrated in Figure 43, last col.). Hence, the "three years later vintage" does not reflect final data in all cases. It falls short of the vintage that captures the finalized data by 6 months for the 4th quarter, 9 months for the 3rd, 12 months for the 2nd, and 15 months for the 1st quarter of the reference year.

3.3 Improvements of the Assessment Methodology

This subchapter describes our improvements that were made to the current OECD methodology for revision analysis. Further, it explains improvements and extensions to the previous studies that utilized OECD MEI data and the corresponding revision spreadsheets.

The results of this study differ substantially from the results of previous studies. The studies cannot be directly compared, since the improved methodology of this study examines revisions of YoY growth rates of QNA data, while the other studies examine QoQ growth rates. Merely some results of the study by McKenzie and Adam (2007) are also based on YoY growth rates and therefore allow direct comparison. Thus, their study is used for comparisons.

We compute results for individual country utilizing the spreadsheets for revision analysis by Di Fonzo (2005a). Our study also provides average results for the group of the 30 OECD and the four non-OECD countries, computed from the country results. Moreover, the study updates the revision needs of 20 countries that were previously assessed in cross-country studies and includes for the first time 14 additional countries.

Through our suggested improvements we aim to facilitate sensitive statistical analysis, and thus provide better information about the actual size of revision. The main purpose is to provide useful results for individual countries and to compare the performance of countries relative to another.

3.3.1 **Time Period and Time Series Versions**

Our time period of analysis and its effect on the assessment is explained in the following. Further, this subsection gives examples how comprehensive revisions, particularly the introduction of ESA95, were handled by revision studies at the national level.

This revision study covers a more recent period than previous studies, thus it provides an up-to-date picture of the revisions of initial released data. The considered data and time periods were selected to facilitate a sensitive statistical analysis and to provide better inform about the actual revision needs than the previous studies.

The underlying data for the analysis is cleaned-up as regards to the elimination of "estimated values"; further the time periods were chosen in such way that there are no missing values and that that they only consider data that have been prepared by the same SNA version (i.e. 1993 SNA or ESA95).

With these improvements, the revision study provides:

i) better measures for the size of revision that can, in average, be expected for the observed revision intervals, and

ii) results that are more comparable across countries, thus facilitating better cross-country analysis.

It is not firmly established whether the analysis of revision to later vintages should strictly focus on a time period with the same methodology and thus only include regular (i.e. current) revisions, or whether long time series across comprehensive revisions should be considered for the analysis, e.g. Deutsche Bundesbank (2011). Branchi et al. (2007) remarks that a long revision history is needed to make definite statements about the reliability of data. NSOs and international studies often consider revision histories of 10 to 15 years or even longer.

Some NSOs calculate the size of revision resulting from the new methodology. Thus, they provide information in addition to the total size of revision that is affected by a comprehensive revision (so practiced for example by German Statistical Office).

To completely eliminate the interferences with most of the happened comprehensive revisions, a special revision study was prepared by the Deutsche Bundesbank (2011). The time period of analysis covered three different regimes, i.e. two comprehensive revisions (some minor changes in methodology were not treated as comprehensive revision). Data of a reference period produced as of the time of the comprehensive revisions were not considered. Yet, the "latest" available preliminary vintage of the old regime was used as the "final" data for the corresponding reference period. Thus, the total size of revision for the first vintage to the "final" vintage was computed based on preliminary data in these cases. While this approach ensures that the revision measures of the first to "final" vintage do not interfere with comprehensive revisions, they interfere with vintages of different maturity. The preliminary data that were used as "final" data have at times substantially smaller size of revisions than size of revision of the first to final data of the other reference periods. Therefore, the results of this approach likely underestimate the actual size of revision of first released data.

Since the OECD database does not provide information about breaks in the series versions of the presented data, i.e. comprehensive revisions, corresponding adjustments to the OECD MEI revisions data were not done. Also information about the regular revision cycles of countries is not provided, which would allow limiting the data set to the vintages until the "final" annual revision.

In this study the reference periods for countries usually cover the QNA data for 1999Q1 until 2011Q2. Thus the considered time period is of comparable lengths to the other studies. At such lengths the country data typically includes one or two comprehensive revisions.

Year 1999, with the MEI release edition of July 1999 (for the 1st quarter) was chosen as the general start of the analyzed time period, since the European countries changed to the ESA 95 methodology as of that year. This marked a major methodological change. The July 1999 MEI edition also provided back-casts according to the ESA 95 methodology until 1990. These data are not used for the analysis, as they only provide results across the major revision. The new 2008 SNA standard has not been adopted yet, other than by Australia (as of the December 2011 data release). Some comprehensive revisions were carried out by NSOs in the observed time period, however these were not across different SNA versions.

Table 49 Observation Periods

The reference periods generally refer to the quarters 1999Q1 until 2011Q2, beginning with the MEI edition of July 1999 until November 2011. Thus the changes of YoY growth rates were observed from 2000Q1 to 2011Q1 for the first to five month later revision interval (observed number of reference quarters |RQ| = 45 observations), and from 2000Q1 to 2007Q2 for the first to latest (i.e. four years or later) revision interval (|RQ| = 30 observations).
The start or begin of the reference periods and MEI issues differs for the following countries:

	First reference quarter	Begin MEI issue	Last reference quarter	End MEI issue
Standard	**1999Q1**	**6/1999**	**2011Q2**	**11/2011**
France		8/1999		
Greece	2004Q1	6/2004	2011Q1	
Hungary	2002Q4	5/2003		
Iceland	2002Q4	4/2003		
Ireland	2002Q3	1/2003		
Korea	1999Q4	4/2000		
Luxembourg	2006Q1	8/2006		
Poland	2006Q4	3/2007		
Portugal	2000Q3	3/2001		
Slovak Republic	2000Q4	4/2001		
Turkey*				
Brazil	2004Q2	10/2004	2008Q2	12/2008
India	2006Q1	6/2006	2010Q1	6/2010
Indonesia	2006Q1	7/2006	2010Q3	2/2011
South Africa	2002Q1	7/2002	2009Q2	11/2009

*No observations for MEI issues from February 2007 to April 2010 (data marked as estimated values)

Source: own illustration

Due to the exclusion of "estimated values" (i.e. real estimates that are not the results of a computation process of preliminary data and gaps in the data) the observation periods of some countries are adjusted to either start later or end earlier. For some countries the reason for the shorter time period is the unavailability of more data in the MEI revisions database, i.e. because revision data are only published since a few years. However, none of the time periods is shorter than 19 consecutive quarters. (This is only one quarter shorter than the shortest time period considered in comparable studies.)

3.3.2 **Coverage of Countries**

Our analysis considers 34 countries. This is 14 more countries than in previous studies.

The present revision study substantially extends the coverage to comparable studies which usually covered merely 18 OECD countries or even less (compare the studies of Di Fonzo 2005b, Tosetto and Lequiller 2006, McKenzie 2006a and 2006b, McKenzie and Adam 2007, and Tosetto and Brackfield 2009). The number of countries was extended to 20 countries in Neumayr (2010), with the inclusion of Czech Republic and Luxemburg.

Our study provides first results for the following countries that were not included in previous assessments: Austria, Greece, Hungary, Iceland, Ireland, Mexico, Poland, Slovak Republic, Sweden, Turkey, and the non-OECD countries, Brazil, India, Indonesia, and South Africa. (The OECD MEI database includes data for four more countries; yet, these countries could not be considered as OECD has just started to build up data for those countries in 2010 respectively 2011.)

3.3.3 **Granularity and Number of Revision Intervals**

We analyze the size of revisions of the first published vintage to the revised vintages at revision time points:

- five month later (t_0+5M),
- one year later (t_0+1Y),
- two years later (t_0+2Y),
- three years later (t_0+3Y), and
- the latest available vintage, published at least four years after the first (t_0+end).

Thus the present study includes the first quarterly revision. Further it covers three annual revisions and a vintage referring to the fourth annual revision or later.

Compared to previous studies, the "latest" vintage (t_0+end) in our study has a maturity of one additional year. This change eliminates those "latest" estimates between the third year and fourth year, which are in many cases still preliminary. This improvement ensures that the eliminated vintages do not interfere with the "latest" vintage for those countries, where the regular revision cycle ends with the 4^{th} annual revision. Since OECD countries usually take three to four years to prepare the "final" data, this is an important addition to the currently observed revision intervals. If a country only needs three years to finalize the estimates, the "final" data are already captured by the three years later vintage. Thus the revision to the initial estimates assessed at the revision time point of the three years later vintage (t_0+3) and the four years later vintage (t_0+4) would be identical (given no comprehensive revision of already "finalized" data happened.) Yet, also in our study, at least the 1st quarter still fails to capture the actual annual revision at the country level.

3.3.4 Year-on-Year Growth Rates

This section explores if revision studies should be based on QoQ growth rates or if YoY growth rates are more appropriate.

The OECD spreadsheets for revision analysis of QNA data facilitate both, revision studies of QoQ and YoY growth rates. The previous revision studies usually only analyze the QoQ growth rates. Only McKenzie and Adam (2007), present some results of YoY growth rates. Our study is based on YoY growth rates.

Both, the German Federal Statistical Office and the German Federal Reserve Bank for instance does not publish revision analyses for QoQ growth rates, but only results for revision studies based on the quarterly changes of NA compared to the previous year, i.e. based on YoY growth rates. Internally, QoQ analysis is also done by the Federal Statistical Office.

Cheem (2003) found that QoQ growth rates analyses are often carried out as supplementary information by European countries, yet YoY growth rates are most widely used. Furthermore, YoY growth rates are best suited for international comparison. QoQ growth rates need seasonally adjusted data for the analysis. But this makes it less suited for international comparison due to the applied adjustment procedures and the extent of adjustments.

The OECD MEI database provides information regarding seasonal adjustment and the presentation of annual levels of the GDP (i.e. multiplication of quarterly GDP by four). Since not all QNA data of countries are seasonally adjusted, the analysis of QoQ growth rates should not be used, but only YoY growth rates.

In the context of economic analysis Cheem (2003,9) notes that the YoY growth rate "is subject to smaller revisions" compared to QoQ or annualised QoQ growth rates, i.e. the annualised rate of change (cf. OECD 2007a, p. 140). McKenzie and Adam (2007, 5, Figure 3) include a graphical comparison of the size of RMAR for QoQ and YoY growths rates (for the revision interval of the first vintage to the five month later vintage "t_0+5M"). The displayed RMARs are higher for the QoQ growth rates than for the YoY growth rates for all 18 countries considered by their study.

3.3.5 Perspectives for Further Improvement

We consider three areas:

- a needed indication of methodology changes by the OECD in the MEI datasets,
- capturing of the NSOs' actual annual revisions with the one, two, and three years later vintage, and
- the extension of the current revision intervals to all annual vintages until finalization of the data.

Ideally "any revision analysis should distinguish" between regular revisions and revisions due to methodological changes (Branchi et al. 2007, 7). Yet even NSOs compute the need of revision of first published data utilizing data that interferes with a comprehensive revision. In such case, results should be corrected by the size of revision due to the methodology change, or countries should at least provide an indication about its size (cf. Statistisches Bundesamt 2011).

The OECD MEI database should include an indication of changes in the methodology to facilitate better analysis of the data, i.e. the database should distinguish between the different NA series versions published for the country.

Besides major comprehensive revisions due to introduction of a new SNA standard, which we excluded, any change of concepts, methods, or data sources is a (minor) comprehensive revision. Calculations of the revision needs for each individual regime separately could be used to compare the performance of different regimes. A revision study that analyzes only the performance of the current regime could provide important value added to data users, as it can inform about the average revision needs that can be expected under the current methodology. Such revision study would focus only on the NA data produced by the current regime, i.e. the first published and subsequently revised data produced by the current computation methodology. However, depending on the time point of introduction of the current methodology, the number of observations can be too limited to provide sensitive analysis. Ahmad, Bournot, and Koechlin (2004) recommend at least 20 observations.

The OECD methodology should calculate the size of revision of the data that corresponds to the actual annual revisions at the country level. Currently the OECD methodology uses the data published in the MED edition 12, 24 and 36 month after the first vintage. This does not always correspond to the annual revisions. The initial QNA estimates are usually prepared by NSOs between one and three month after the reference period. In case of the 1st quarter this is between April and June. This estimate is usually published with the next MEI edition one month later, thus between May and July. As explained in chapter 3.2.4, p. 249, the revised vintage “one year after the first” vintage would therefore refer to the data of the edition of the same month of the following year. At that time, the first annual revision is not yet produced, thus the data still refers to the latest reported quarterly revision. This applies similarly for the other annual revision vintages and the corresponding previous annual revision vintage. Consequently, the OECD methodology and revision spreadsheets could be further refined to capture the actual annual revisions.

We already improved the “latest” revision time (t_0+end) point by setting it to “at least four years after the initial estimate”. An additional revision time point could refer to the vintage at precisely “four years after the first vintage” (t_0+4Y). The advantage of this materializes if a country’s regular revision cycle is four years and a comprehensive revision happened at any time point later than 51 month (four years and one quarter) after the initial published vintage of the 4th quarter. However, if the time period covers comprehensive revisions, the revision time points likely interfere with the comprehensive revision before the precise four years later vintage (t_0+4Y).

3.4 Assessment of Reliability

In this chapter the results of our revision analysis are presented.

The main questions regarding the reliability of the first released estimates are:

1) What is the size of revision of the initial estimate at revision time point three years after the first published vintage (t_0+3Y), respectively regarding the "latest" vintage (t_0+end)?
2) Is a small size of revision of the first vintage to the first sub-annual revision (t_0+5M) a good indication that the total size of revision is low, too?
3) Are there large differences between countries?
4) Do countries that have a small Mean Absolute Revision (MAR) also have small Relative Mean Absolute Revisions (RMAR)?

To answer this question the MAR and RMAR of countries are compared. National revision analyses usually considers only the MAR to assess the revision need and further the Mean Revision (MR) (cf. DEF. 18, p. 45) to inform about a potential bias (cf. Statistisches Bundesamt 2011 or ONS 2012). RMAR, i.e. the mean absolute revision adjusted for the size of economic growth, allows cross-country comparison of revisions (cf. Ahmad, Bournot, and Koechlin 2004). To compare the individual country performances with the group average, we calculate the average for the countries in our study.

The following assessment is presented in three parts, first the results for Mean Absolute Revisions (MAR), second those for Relative Mean Absolute Revisions (RMAR), and last the results for the group average. The group averages are presented separately for the 30 OECD and the four non-OECD countries (Brazil, India, Indonesia, and South Africa). In the tables and figure for MAR and RMAR listing the countries in alphabetical order, the non-OECD countries are presented following the OECD countries. Our assessment closes with a comparison of our new results with a previous revision study by McKenzie and Adam (2007).

3.4.1 Revision by Individual Countries

In this section we present our results for the 34 countries we consider with our assessment. First we present the size of the MAR followed by the size of RMAR.

3.4.1.1 Mean Absolute Revisions

The MAR for the individual countries is presented in this subsection. Table 50, p. 265 provides the size of revision in percentage points (%p) by country and revision time points (i.e. by the time that elapsed after the publication of the first estimate). These results are subsequently illustrated in Figure 44, p. 266, in alphabetical order of countries and in Figure 45, p. 267 in ordered of size of revisions at the revision time point three years after the first vintage "t_0+3Y".

The size of MAR of the first released growth rate increases from the first revision time point "t_0+5M" to the following revision time points (cf. country results by revision time points in Table 50, p. 265 and Figure 44, p. 266).

The MAR of the first growth rate to the five month later vintage "t_0+5M"of the 30 OECD countries is an average 0,22 percentage points (cf. Table 52, p. 274, col. two). At revision time point "t_0+3Y", already yielding "final" data for some countries, the MAR of OECD countries is an average 0,68 percentage points (cf. Table 52, p. 274, col. five). Only for a few countries MAR decreased in size from an earlier to a later revision time point. For example for Belgium, MAR decreased from 0,34 percentage points at the two years after first vintage "t_0+2Y" to 0,33 percentage points at the three years after first vintage "t_0+3Y" (cf. Table 50, p. 265, col. four respectively five).

The size of MAR varies between countries. Since it is an absolute measure, the changes in size of MAR across different revision time points need to be examined for each country individually. For some countries, the revisions of the first estimate to the five month later vintage "t_0+5M" are larger, than the revision of other countries regarding the first estimate to three years later vintage"t_0+3Y". For example Japan has a MAR of 0,40 percentage points of the first to the five month later vintage "t_0+5M" (cf. Table 50, p. 265, col. one). In contrast, Germany has a MAR of 0,37 percentage points of the first to the three years after first vintage "t_0+3Y".

While some OECD countries already show revisions above 0,2 percentage points at the first revision time point (t_0+5M), all non-OECD G20 countries show rather small revisions at this revision time point. With exception of Indonesia, large revisions for non-OECD countries can be observed starting with the first annual revision (t_0+1Y).

Table 50 Mean Absolute Revisions (MAR) for OECD and G20 Countries by Revision Time Point

Country \ MAR at Revision Time Point	t_0+5M	t_0+1Y	t_0+2Y	t_0+3Y	t_0+end
AUSTRALIA	0,14	0,33	0,41	0,47	0,53
AUSTRIA	0,22	0,49	0,45	0,53	0,73
BELGIUM	0,13	0,29	0,34	0,33	0,39
CANADA	0,11	0,21	0,30	0,29	0,32
CZECH REPUBLIC	0,25	0,57	0,67	0,88	0,96
DENMARK	0,35	0,54	0,47	0,72	0,77
FINLAND	0,33	0,81	0,90	0,87	1,19
FRANCE	0,14	0,32	0,44	0,53	0,46
GERMANY	0,11	0,22	0,30	0,37	0,62
GREECE	0,22	0,46	0,70	0,64	1,00
HUNGARY	0,25	0,43	0,50	0,55	0,55
ICELAND	0,63	1,92	2,60	2,58	2,40
IRELAND	0,33	0,80	1,21	1,43	1,07
ITALY	0,10	0,18	0,22	0,30	0,51
JAPAN	0,40	0,74	0,94	0,82	0,86
KOREA	0,07	0,24	0,35	0,45	0,44
LUXEMBOURG	0,66	0,84	1,21	1,48	1,31
MEXICO	0,11	0,30	0,37	0,33	0,70
NETHERLANDS	0,20	0,34	0,49	0,70	0,79
NEW ZEALAND	0,21	0,46	0,57	0,64	0,74
NORWAY	0,24	0,46	0,67	0,91	0,90
POLAND	0,20	0,24	0,32	0,40	-
PORTUGAL	0,14	0,23	0,32	0,46	0,51
SLOVAK REPUBLIC	0,24	0,53	0,58	0,65	1,22
SPAIN	0,06	0,18	0,34	0,43	0,64
SWEDEN	0,18	0,39	0,52	0,57	0,74
SWITZERLAND	0,15	0,39	0,50	0,63	0,72
TURKEY	0,21	0,56	0,50	0,40	1,92
UNITED KINGDOM	0,16	0,22	0,30	0,36	1,00
UNITED STATES	0,17	0,32	0,63	0,63	0,57
BRAZIL	0,14	0,39	0,88	1,01	0,92
INDIA	0,11	0,50	0,87	0,77	-
INDONESIA	0,04	0,13	0,19	0,21	-
SOUTH AFRICA	0,09	0,49	0,73	0,94	1,07

Source: own compilation

It is apparent from Figure 44 (below) that for most countries the total size of revision of the YoY growth rate is less than 1 percentage point. Four countries show larger revisions "t_0+2Y" For the last published vintage "t_0+end", nine out of the 34 countries, i.e. more than 25%, indicate a total size of revision of 1 percentage point or more to the first published YoY growth rate (i.e. the first vintage at time point t_0).

The following figure illustrates MAR for each revision time point ordered by country, followed by the figure ordering the countries by size of MAR at revision time point "t_0+3Y".

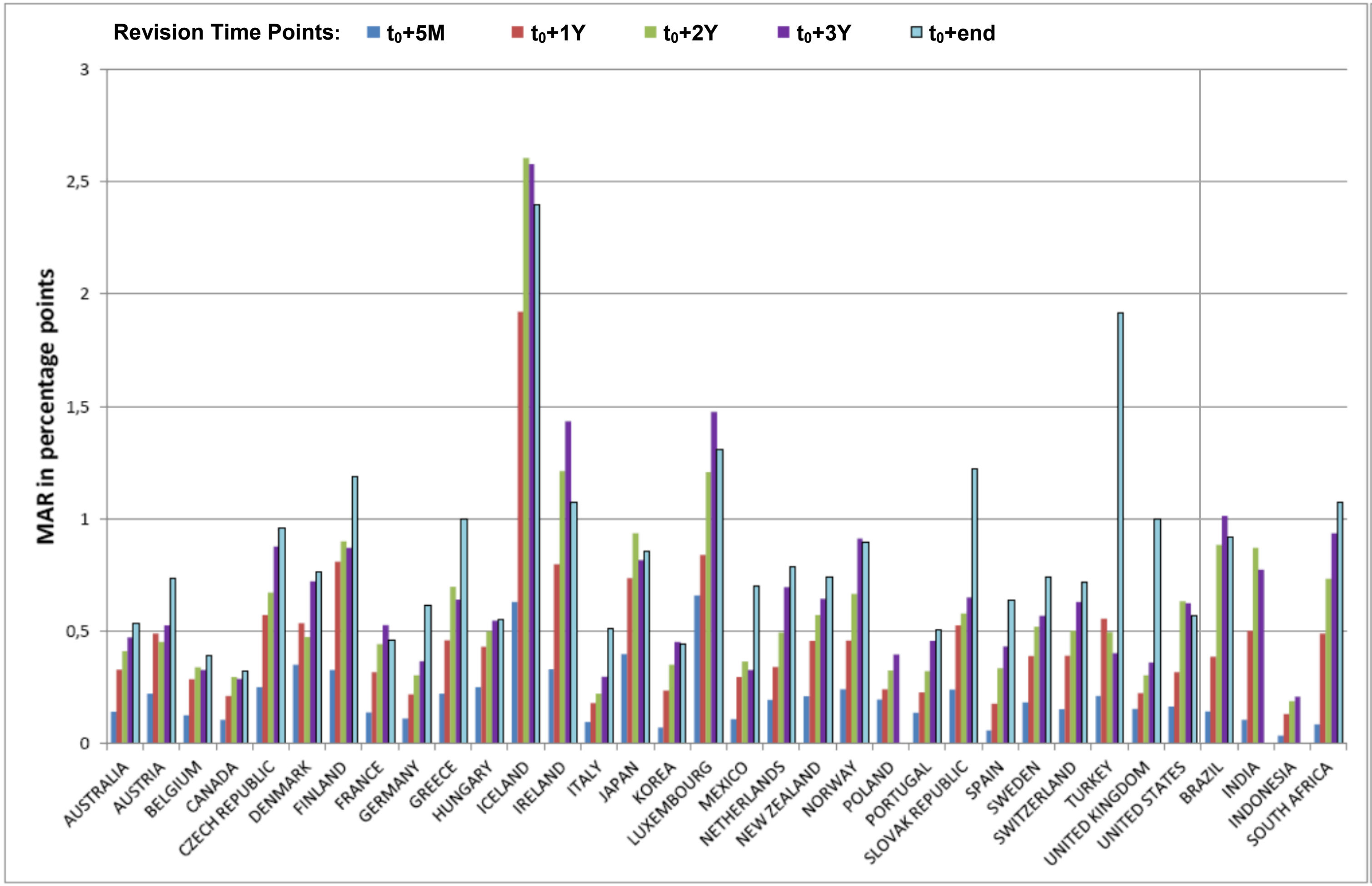


Figure 44 Mean Absolute Revisions (MARs) for OECD and G20 Countries by Revision Time Points, Countries in Alphabetical Order

Source: own compilation

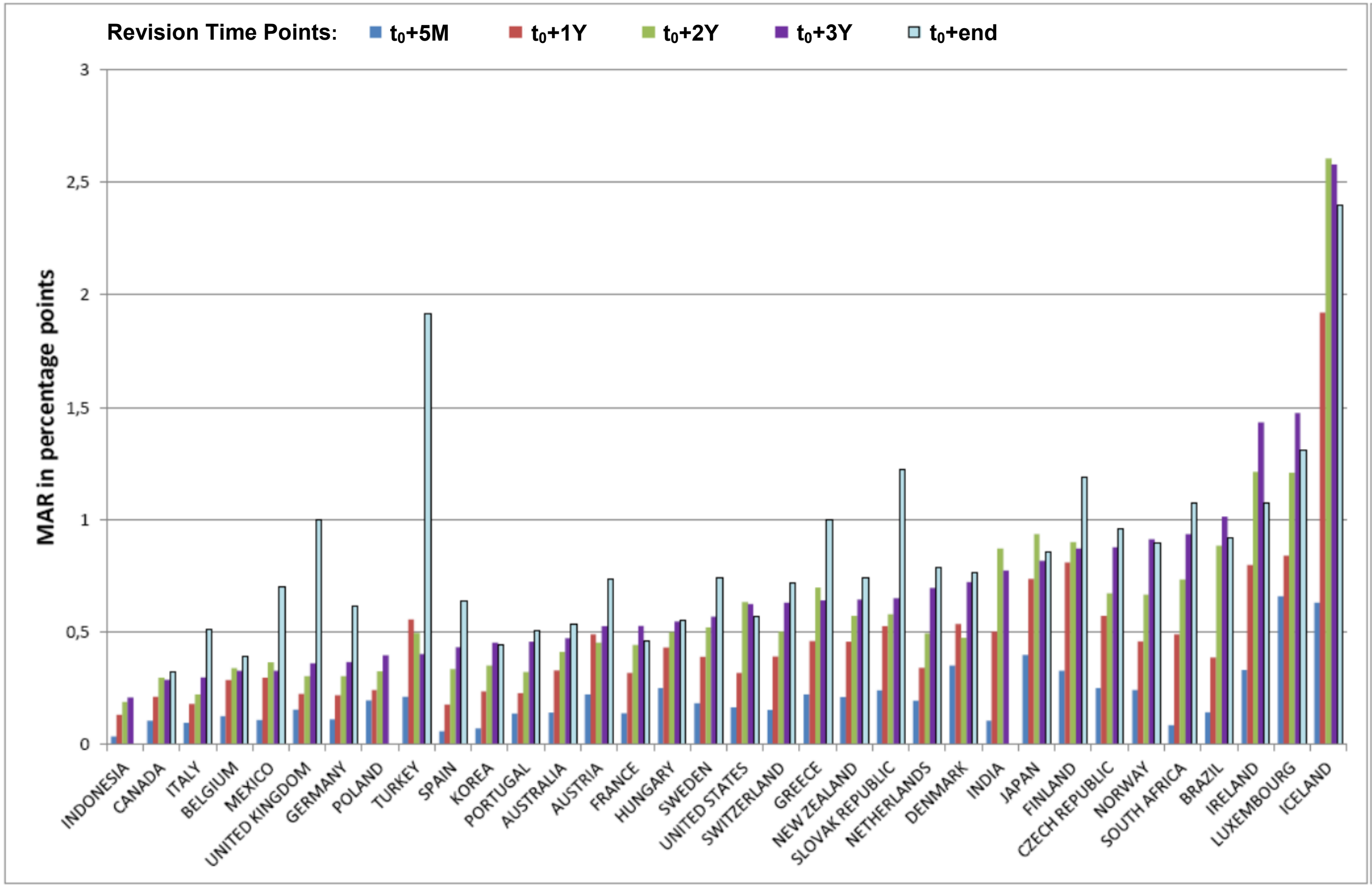


Figure 45 Mean Absolute Revision (MAR), Countries in Order of Size of MAR at Revision Time Point "t_0+3Y"

Source: own compilation

The Slovak Republic, Turkey and the United Kingdom show particularly high revisions of the latest revision time point "t_0+end" compared to the size of revisions of earlier revision time points. For these countries comprehensive revisions might have lead to changes in growth rates, which are far off the first growth rate. The "methodological changes in national accounts generally tend to slightly increase GDP" (cf. OECD 2006, 11). This can be observed for instance by the results of the two comprehensive revisions in Germany in 2005 and 2011, where the revisions across the different methodologies are 0,2 percentage points in both cases (cf. Statistisches Bundesamt 2008 and 2011).

The order of countries by size of MAR observed three years after the first estimate "t_0+3Y" (see Figure 45) better illustrates the range of size of revisions across countries. Also the performance of countries relative to another can be pointed out more easily. The non-OECD G20 countries are among those countries with the higher needs of revision of the first estimate. While South Africa and Brazil belong to the least countries, Indonesia shows the overall lowest revisions.

Among the OECD countries the large difference regarding the quality of initial released data becomes visible for example by comparing Canada, with MAR = 0,33 percentage point at revision time point three years after first vintage (t_0+3Y), and Iceland with MAR > 2,5 percentage points already at revision time point two years after first vintage (t_0+2Y).

With some exceptions the Western European countries generally show smaller revisions than countries in Eastern Europe and Scandinavia. Considering only the OECD countries outside of Europe, Canada and Mexico show relative small revisions compared to Japan, which has the highest revision needs of these countries. Interesting is also that Australia, the US and New Zealand are among those countries with medium sizes of revisions, next to Portugal and the Slovak Republic.

3.4.1.2 Relative Mean Absolute Revisions

The Relative Mean Absolute Revision (RMAR) is best suited for cross-country comparisons. Using the RMAR to assess the quality of first published estimates yields a different order of countries by size of RMAR than by size of MAR. The RMAR measure reveals that even countries of the group with rather low size of revisions in percentage points based on MAR, may substantially revise the first published growth rate relative to its size.

Only few countries show a RMAR < 10% of the first published growth rate to the three years later revision time point “t_0+3Y” (cf. Table 51, p. 270). The countries that still show a low size of revision by RMAR, already show revisions around 20% of the first to the three years later vintage. For example Germany indicates a revision of 23% of the first vintage to the three years later vintage.

The size of RMAR increases from the time point at five month after the first vintage (t_0+5M), to the time point three years after the first vintage (t_0+3Y). The first revision interval shows RMARs from 0,01 to 0,17, thus from 1% to 17% of the first published growth rate. The size of RMAR observed at revision time points one, two or three years after the first published vintage are two, three or even ten times higher. For example Switzerland shows a RMAR of 8% of the first published YoY growth rate at the revision time point five month later (t_0+5M). At the revision time point three years later (t_0+3Y), Switzerland shows a RMAR of 28% (cf. Table 51, p. 270, col. two for “t_0+5M” respectively five for “t_0+3Y”).

It is evident that the four G20 countries which are not OECD members are marked by much smaller size of revisions evaluated based on the RMAR. This is particularly the case for the sub-annual (short term) revision.

Table 51 Relative Mean Absolute Revisions (RMAR) for OECD and G20 Countries by Revision Time Point

Country \ RMAR at Revision Time Point	t_0+5M	t_0+1Y	t_0+2Y	t_0+3Y	t_0+end
AUSTRALIA	0,05	0,11	0,14	0,14	0,16
AUSTRIA	0,10	0,22	0,19	0,22	0,30
BELGIUM	0,06	0,13	0,16	0,15	0,18
CANADA	0,04	0,08	0,11	0,10	0,11
CZECH REPUBLIC	0,06	0,14	0,16	0,20	0,22
DENMARK	0,15	0,23	0,21	0,39	0,38
FINLAND	0,09	0,21	0,25	0,26	0,35
FRANCE	0,08	0,18	0,21	0,25	0,22
GERMANY	0,05	0,11	0,17	0,23	0,36
GREECE	0,06	0,14	0,21	0,17	0,25
HUNGARY	0,08	0,13	0,14	0,16	0,16
ICELAND	0,16	0,37	0,45	0,46	0,39
IRELAND	0,08	0,16	0,23	0,29	0,21
ITALY	0,06	0,11	0,15	0,23	0,33
JAPAN	0,16	0,30	0,39	0,41	0,43
KOREA	0,02	0,05	0,08	0,10	0,09
LUXEMBOURG	0,17	0,22	0,22	0,26	0,19
MEXICO	0,03	0,08	0,11	0,10	0,23
NETHERLANDS	0,10	0,16	0,23	0,32	0,37
NEW ZEALAND	0,08	0,17	0,18	0,20	0,21
NORWAY	0,13	0,23	0,31	0,42	0,43
POLAND	0,05	0,07	0,08	0,06	-
PORTUGAL	0,11	0,17	0,23	0,33	0,37
SLOVAK REPUBLIC	0,04	0,09	0,09	0,10	0,20
SPAIN	0,02	0,06	0,11	0,13	0,17
SWEDEN	0,06	0,13	0,18	0,20	0,23
SWITZERLAND	0,08	0,20	0,24	0,28	0,33
TURKEY	0,03	0,08	0,07	0,06	0,35
UNITED KINGDOM	0,06	0,09	0,12	0,14	0,33
UNITED STATES	0,06	0,11	0,24	0,26	0,22
BRAZIL	0,03	0,09	0,20	0,21	0,23
INDIA	0,01	0,06	0,09	0,08	-
INDONESIA	0,01	0,02	0,03	0,03	-
SOUTH AFRICA	0,03	0,13	0,17	0,20	0,23

Source: own compilation

From Figure 46, p. 271 it can be observed that Norway, Japan and Iceland are now the countries with the highest RMARs for the revision need of the first estimate to the latest available vintage (t_0+end). Compared to the other countries these three countries show large RMARS of the initial released estimates for all observed revision time points.

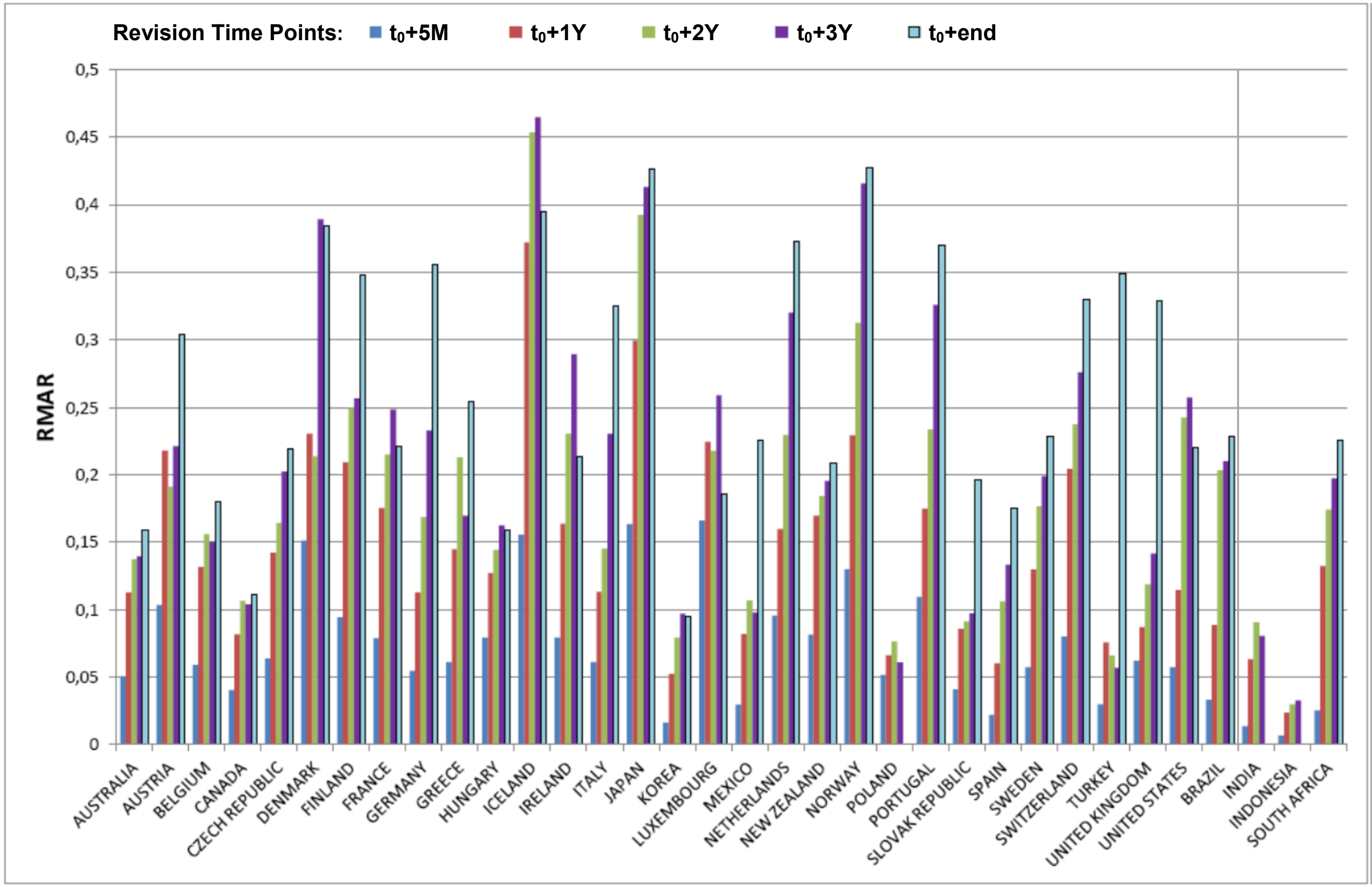


Figure 46 Relative Mean Absolute Revisions (RMARs) for OECD and G20 Countries by Revision Time Point, Countries in Alphabetical Order

Source: own compilation

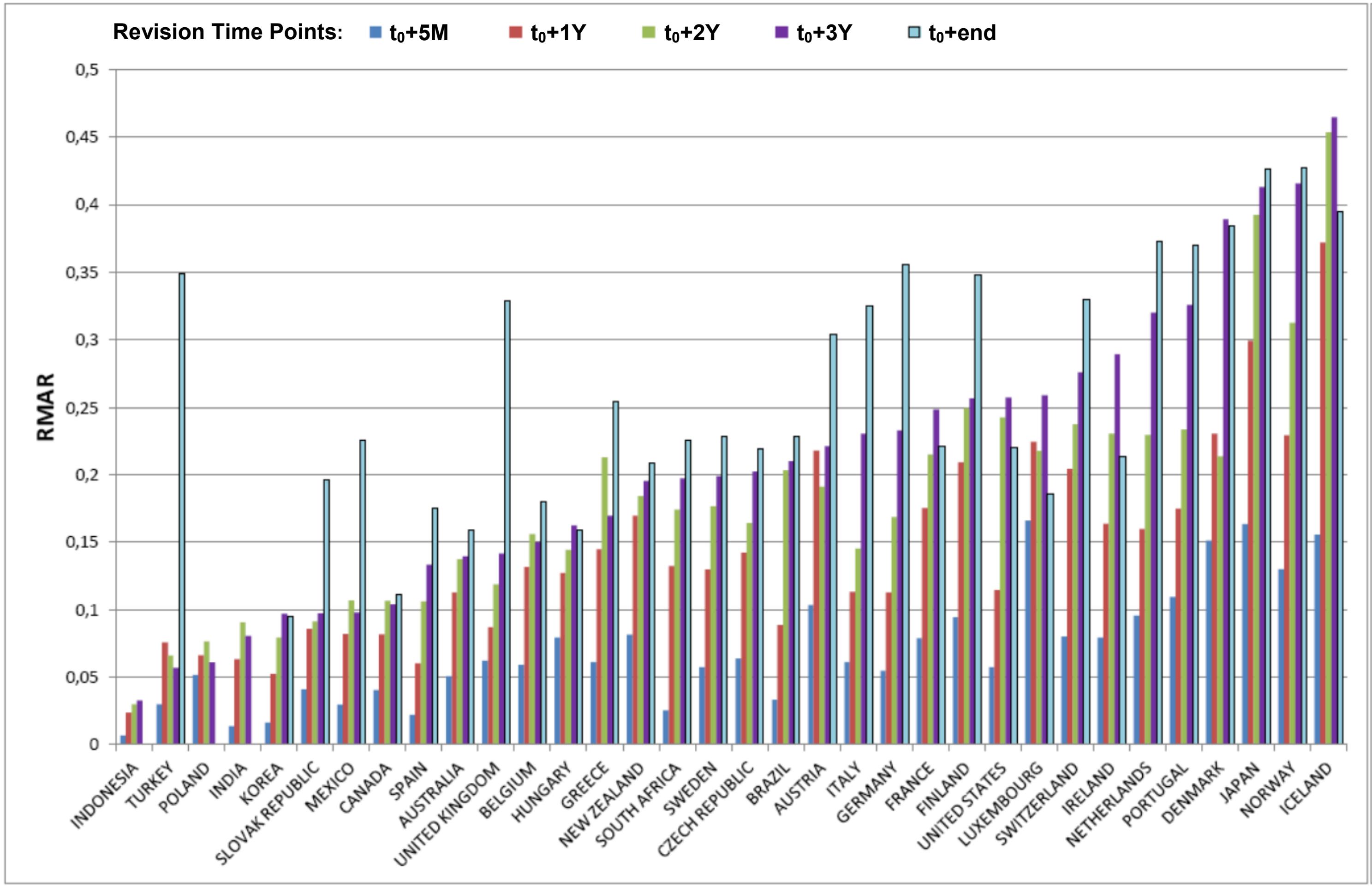


Figure 47 Relative Mean Absolute Revision (RMAR), Countries in Order of Size of RMAR at Revision Time Point "t_0+3Y"

Source: own compilation

Figure 47 presents the countries ordered by size of RMAR of the first published growth rate to the growth rate published by the three years later vintage (t_0+3Y). This leads to an order where Indonesia, Turkey, Poland, and India are among the countries with the smallest need of revision of the first published data. Iceland, Norway and Japan now belong to the group with the highest need of revisions. The country showing the smallest revisions is Indonesia, which already showed the smallest revisions for the MAR measure. The country with the absolute highest revision is Iceland, with the size of RMAR of the first published YoY growth rate to the three years later vintage (t_0+3Y) larger than 45%.

Iceland serves as an example where data users may wish to draw on later vintages of released GDP and corresponding growth rates. The short term revision "t_0+5M", i.e. the vintage five month after the first vintage already shows RMAR of 15%. The one year after the first vintage (t_0+1Y) shows RAMR above 35%. The revisions following the one year later vintage (t_0+1Y) are much closer to the "final" data (t_0+end). Thus data users who use the data for decision making may wish to use revised data as of the sub-annual revision time point or even of the first annual revision.

Considering the high reputation enjoyed by some countries, the results of our study are to an extend inverse to the expected results. Countries like Germany, France, Finland, and the US find themselves in the middle range based on RMAR at the revision time point "t_0+3Y", i.e. three years after the first published vintage. The Netherlands even show high sizes of RMAR compared to other countries. On the other hand, countries that have the smallest RMAR are not those known for applying best practice in national accounting or for providing statistical capacity building programs. Still among the countries with small RMAR, which enjoy a good reputation are e.g. Australia and Canada.

3.4.2 Average Revisions by OECD and Non-OECD Cluster

In this section, we present the arithmetic means (average) of the individual results for the set of the 30 OECD countries and the set of the four non-OECD G20 countries.

OECD Countries

Table 52 presents the average results of the OECD countries. The individual results are computed utilizing the spreadsheets for revision analysis by Di Fonzo (2005a).

Table 52 Summary Results of Revision Analysis, Average of OECD Countries

Variable \ Revision time point	t_0+5M	t_0+1Y	t_0+2Y	t_0+3Y	t_0+end
Sample (for majority)	2000Q1-2011Q1	2000Q1-2010Q2	2000Q1-2009Q2	2000Q1-2008Q2	2000Q1-2007Q2
Observed number |RQ| of reference quarters within sample with later vintage	39,9	36,8	32,9	28,9	24,9
Mean Absolute Revision (MAR)	0,22	0,47	0,60	0,68	0,85
Relative Mean Absolute Revision (RMAR)	0,08	0,15	0,19	0,22	0,27
Mean Revision (MR)	0,01	0,05	0,13	0,27	0,35
Growth rate of later vintage > growth rate of first vintage	46%	54%	56%	64%	63%
Sign of growth rate of later vintage = sign of growth rate of first vintage	99%	97%	96%	97%	95%
Note: the revision time point t_0+end only considers the results of 29 instead of the 30 countries, since Poland do not have enough revision data, i.e. because **observed number of reference quarters |RQ|** < 20.					

Source: own compilation

The average size of the sample, i.e. our analyzed reference periods, indicates a sufficiently large number of observations of reference quarters (|RQ| = n) of first vintages with data at later revision time points, i.e. a later vintage. We observe that |RQ| > 20 for all five revision time points.

Regarding MAR, the average of OECD countries indicates a need of revision of 0,47 percentage points at the revision time point "t_0+1Y", i.e. one year after the first published YoY growth rates. This is more than the revision of Germany at the three years later revision time point "t_0+3Y" (cf. Table 50, p. 265, col. five for "t_0+3Y").

Regarding RMAR, the OECD average shows that revisions are large relative to their first published growth rates. Only RMAR at the sub-annual revision time point (t_0+5M) is smaller 10%. By the time point of the three years later vintage "t_0+3Y", the initial released growth rate is revised 22% on average. Germany's RMAR is just above this average, showing 23% (cf. Table 51, p. 270, col. five for "t_0+3Y").

Regarding MR, the OECD average indicates upwards revisions that increase in size over time. This coincides with the growth rate of the later vintage being most of the times higher than the growth rate of the first vintage (cf. the row following the MR).

The vintages of later revision time points, i.e. the three years after first and subsequent vintages, are higher than the first published vintage above 50% of the time.

Referring to the last row, the revised growth rates (i.e. the later vintages) showing the same sign as the first vintage in nearly 100%. That means, if the first published growth rate is positive, so is the growth rate observed at a later revision time point. For example the growth rate published three years after the first (cf. Table 52, col. five for "t_0+3Y") is has the same sign for 97%, meaning it has only changed from positive to negative or vice versa in 3% of the observed first published growth rates.

G20 non-OECD Countries

We calculate the average of the results of the four G20 countries that are not OECD members separately. These are presented in Table 53.

Table 53 Summary Results of Revision Analysis, Average of G20 Non-OECD Countries

Revision time point / Variable	t_0+5M	t_0+1Y	t_0+2Y	t_0+3Y	t_0+end
Sample (varies)	2003/2007-2008/2011	2003/2007-2007/2010	2003/2007-2006/2009	2003/2007-2005/2008	2003/2007-2004/2007
Observed number |RQ| of reference quarters within sample with later vintage	20,5	18	13,75	10	6,75
Mean Absolute Revision (MAR)	0,09	0,38	0,67	0,73	1,00
Relative Mean Absolute Revision (RMAR)	0,02	0,08	0,12	0,13	0,23
Mean Revision (MR)	0,05	0,14	0,46	0,67	1,00
Growth rate of later vintage > growth rate of first vintage	42	56	72	89	100
Sign of growth rate of later vintage = sign of growth rate of first vintage	100%	100%	100%	100%	100%
Note: the revision time point t_0+end refers to results for Brazil and South Africa, only. This is because India and Indonesia do not have enough revision data.					

Source: own compilation

The *observation periods* vary for the non-OECD G20 countries (cf. Table 49, p. 258). The lengths of the observation periods are shorter than for OECD countries and also refer to different years for the individual countries. All countries cover recent periods.

In case of India and Indonesia the most recent available data was excluded since it is marked as estimated values.

Regarding MAR the average of the non-OECDE countries of the first estimate to the five months later vintage (t_0+5M) is well below that of the OECD countries. Also the one year later vintage (t_0+1Y) is still below OECD average. The averages for the two (t_0+2Y) and three years later vintages (t_0+3Y), are already higher than the OECD average. The MAR of the latest revision time point (t_0+end) is well above the OECD average. Results for the latest revision time point are only based on data for Brazil and South Africa, since the observation periods for India and Indonesia are too short to assess the revision time point "t_0+end".

The *RMAR* indicates needs for revision that are an average below the OECD countries for all revision time points. For the first two revision time points "t_0+5M" and "t_0+1Y" they are much lower, slowly progressing towards the same size as the OECD average for the revision time points "t_0+2Y" and "t_0+3Y".

The MR of the growth rates shows a strong tendency for growing upwards revisions. The number of revised growth rates, i.e. vintages at later revision time points, that are higher than the first published vintage, indicates more upward revisions compared to the OECD average. About 60% of the revisions of the one year later vintage (t_0+1Y) and nearly 90% at the revision time point of the three years later vintage (t_0+3Y) have a higher growth rate than the first vintage. In case of Brazil and South Africa the latest vintage (t_0+end) is higher than the first vintage in all cases.

A very positive result is that growth rates at later revision time points all have the *same sign* as the first published growth rates. Meaning the first published growth rate does not change form positive to negative or vice versa at a later time due to revisions.

3.5 Comparison of Results with a Previous Study and Summary

In the following, we first compare our results to McKenzie and Adams (2007), who consider 18 OECD countries. Then we summarize the results of our study.

Though the study of McKenzie and Adams (2007) is not the most recent, it is the only previous study that allows direct *comparison*, because it provides at least some results based on YoY growth rates. Their study present results of RMAR of the first published YoY growth rates of GDP at revision time point five month after the initial released estimate (t_0+5M). Since these results are illustrated in a figure (cf. McKenzie and Adams 2007, 5, Figure 3), we state the approximate size of RMAR in the following Table 54.

Table 54 Comparison of Size of RMAR by Our Study and McKenzie and Adam (2007)

Country / Revision time point	t_0+5M	
	RMAR of our study	RMAR of McKenzie and Adam[1]
AUSTRALIA	0,05	0,07
AUSTRIA	0,10	-
BELGIUM	0,06	0,07
CANADA	0,04	0,05
CZECH REPUBLIC	0,06	-
DENMARK	0,15	0,18
FINLAND	0,09	0,12
FRANCE	0,08	0,09
GERMANY	0,05	0,08
GREECE	0,06	-
HUNGARY	0,08	-
ICELAND	0,16	-
IRELAND	0,08	-
ITALY	0,06	0,09
JAPAN	0,16	0,2
KOREA	0,02	0,02
LUXEMBOURG	0,17	-
MEXICO	0,03	-
NETHERLANDS	0,10	0,13
NEW ZEALAND	0,08	0,12
NORWAY	0,13	0,13
POLAND	0,05	-
PORTUGAL	0,11	0,07
SLOVAK REPUBLIC	0,04	-
SPAIN	0,02	0,03
SWEDEN	0,06	-
SWITZERLAND	0,08	0,13
TURKEY	0,03	-
UNITED KINGDOM	0,06	0,11
UNITED STATES	0,06	0,04
BRAZIL	0,03	-
INDIA	0,01	-
INDONESIA	0,01	-
SOUTH AFRICA	0,03	-

[4]cf. McKenzie and Adam (2007, 5).

Source: own compilation and McKenzie and Adam (2007, 5, Figure 3)

Table 54 provides RMAR at revision time point "t_0+5M" of our study and of McKenzie and Adam (2007). A comparison of the results indicates that the size of RMARs in our study is an average 0,02 lower than previously found. The average size of RMAR in our study, for the 18 OECD countries that are also considered by McKenzie and Adam is $RMAR_{Our}$ = 0,08. The average size of RMAR for the OECD countries considered by McKenzie and Adam is $RMAR_{McK\&A}$ = 0,10. For Japan, New Zealand, and Switzerland our RMAR is even more than 0,03 lower. The biggest difference is observed for Switzerland, where the *RMAR* of our study is approximately 0,05 lower than of McKenzie and Adam. Only for Portugal and the US, our RMAR is about 0,03 higher. The smaller size of RMAR in our study are chiefly results of our improved time period of analysis, excluding major revisions across different SNA versions.

Due to the methodological improvements introduced with this study, our results provide better information and more accurate measures for the actual revision needs of the individual countries. In particular, we have a higher number of revision time points, including one sub-annual revision "t_0+5M", i.e. five month after the first published vintage and the latest vintage "t_0+end" refers to the vintage at least four years after the first vintage. Since we excluded revisions across different SNA versions, our results are not affected by changes of the growth rates due to methodological changes. Thus we obtain better results that indicate smaller sizes of the computed revision measures MAR and RMAR. This shows that the actual need for revisions of the first published data is lower than found by previous studies.

Regarding the revision time point "t_0+3Y", i.e. three years after the first published estimate of the YoY GDP growth rate for a given reference quarter, we make the following observations:

- MAR > 0,6 percentage points for nearly 50% of the 34 analyzed countries.
- MAR < 0,4 percentage points for only about 25% of the countries.
- MAR > 0,2 percentage points even for the "best" of the countries.
- MAR > 1,0 percentage points for four countries.
- Countries with the highest MAR are Ireland, Luxembourg, and Iceland. (For Iceland: MAR = 2,58 percentage points).
- The best country in the OECD domain is Canada, followed by Italy, Belgium, and Mexico. (For Canada: MAR = 0,29 percentage points.)

The RMAR indicates very large differences between the countries. Ordering the countries by size of RMAR at the revision time point "t_0+3Y" yields a different order than by size of MAR. In some cases the order by RMAR is inverse that by MAR. We make the following observations regarding the revision time point "t_0+3Y":

- RMAR = 0,22 for the OECD average. Meaning, calculating the average of the OECD countries, the first published growth rates are revised by 22%.
- The best OECD countries are Turkey and Poland (RMAR = 6%), followed by Canada, Korea, Slovak Republic, and Mexico (RMAR = 10%).
- The best of the four non-OECD countries is Indonesia (RMAR = 3%), followed by India (RMAR = 8%).
- The worst country is Iceland (RMAR = 46%), followed by Norway and Japan (RMAR = 42% and RMAR = 41%, respectively).

IV. Conclusion and Perspectives

This chapter provides a summary of our three DQ assessments. Perspectives for further improvements are summarized at the end of the chapter.

The three leading quality frameworks for official statistics issued by the IMF, Eurostat, and OECD are discussed in part II, chapter 1. Further, we define and discuss the quality dimensions and measures relevant for our study in part II, chapter 2.

Our studies of *timeliness*, *availability*, and *reliability* comprise a methodological part and an empirical assessment. The methodological part analyzes the deficiencies of the currently used assessment methodology and suggests improved concepts. Our assessments are based on our suggested concepts. We include a comparison with the assessed status of the DQ dimension by previous studies.

Our assessments inform about the status of NA data quality at the international level. Thus, we inform users about *timeliness*, *availability*, and *reliability* of data they need for their work. Moreover, we inform providers (cf. United Nations 2010b and 2011) of the initiatives and training programs that aim to strengthen NA capacity and to produce data according to 2008 SNA.

Timeliness

In part II, chapter 2.1, we provided the definition for the *DQ dimension* "*Timeliness*". We refine the wording of its verbal definition (based on Eurostat 2003a) and derive the (general) formula for the *timeliness dimension* (cf. DEF. 5: Timeliness, p. 18). Based on this definition, we define (as verbal definition and formula) two *DQ measures* (cf. part II, chapter 3.3). The first measure: *Timeliness* for *ANA data publication at the international level* T(c) (cf. DEF. 32, p. 117). Further, the corresponding measure: *Time lag*, TLR(c), of publication given the expected *timeliness* of publication (cf. DEF. 33, p. 122).

We define *timeliness* and *time lag* for *new* published ANA data, i.e. for those countries $c \in C_R$ that reported a *new* reference year of ANA data to the UN. Timeliness, T(c), is calculated by the most recent available reference year t_{RY} within the set of reported ANA data, in relation to their publication year t_{PY} by the UN, in the YB that makes the data publicly available at the international level for the first time.

Correspondingly, the time lag, TLR(c), is calculated by the most recent available reference year t_{RY} within the set of reported ANA data, in relation to the expected reference year $t^* = t_{PY}-1$, i.e. the reference year preceding the publication year t_{PY} of the YB. Thus, the relation between the *timeliness* T(c) in years $n \in \boldsymbol{N}$ and *time lag* TLR(c) in years $n \in \boldsymbol{N}$ is: TLR(c) = T(c) – 1.

In part III, chapter 1.2, we analyzed the current *assessment methodology* regarding the *timeliness assessment* of the UN (UNSD 2008), and suggested an *improved approach*. The results of *our study* are based solely on those countries $c \in C_R$ that indeed report new data. Hence, we indeed calculate the *timeliness* T(c) of the country's ANA data publication at the international level. Further, our results provide better information about the real *timeliness* of data *production* (for the country groups we consider).

Our *assessment* of *timeliness shows* that the *timeliness* of ANA data publication is at a level where 90% of all countries produce ANA data for either the past reference year (t-1) or, at least reference year t-2. Thus, the ANA are published with a maximal *delay* of one year relative to the expected reference year (i.e. t-1). In contrast, the UN assessment states that merely 64% of the UN Member States have this timeliness. Yet, the UN disregards, that a number of countries in the YB do not report ANA data annually. The data for these countries are merely republished and thus "known" data. Thus, the UN approach does not assess *timeliness* of ANA data publication, when published for the first time. Further, the UN disregards, that some countries do not even produce NA, nevertheless the presented percent values are based on all countries.

Limitations of the assessment of the DQ of country data based on data published by IOs (cf. part III, chapter 1.2.1) are mainly that the data needed for the assessment are subject to reporting (i.e. provision) by the countries. Further, the YB is an annual publication (released in September in case of the analyzed YB) and the consideration of new country data by the UN is subject to the cutoff date for inclusion of new data into the upcoming YB.

The clustering of country groups is improved (cf. part III, chapter 1.2.3) by further subdivisions of *regions* and additional *country groups*. Thus, we gain better insight into the *timeliness* of NA data production of specific *country clusters*. Besides considering those *regions* already covered by the UN assessment, we consider the set C_{All} / C_{UN} of 32 countries that are not UN Member States. Moreover, we consider those *country groups* and *regional* or *international organizations* that are of specific interest for researchers and economic analysts. These are for example the least developed countries (LDCs), leading economies like the G20 and the group of other advanced economies not within the group of G20, and organizations like the OECD.

Our three-part assessment (cf. part III, chapters 1.3.1, 1.3.2, and 1.3.3), covers

i. the UN Member States (set C_{UN}),
ii. all countries (set C_{All}), including non-UN countries, and official MDG groups,
iii. selected organizations and economic groups, and UN Regional Commissions.

We find substantial *differences* of *timeliness* between individual *country groups*. Surprisingly, the advanced economies, other than G20 show superior performance to the G20 countries. Of the advanced economies, other than G20, 78% have data with the optimal timeliness, i.e. t-1. In contrast, only 42% of the G20 countries have t-1 data. However, 95% of the G20 countries have data with a maximum delay of one year, i.e. having at least reference year t-2 data. This is the case for all advanced economies, other than G20.

Our timeliness assessments allow data users to predict when data of the reference year of interest for specific country groups will become available at the UN. *Country groups* with poor performance can be identified. To strengthen the statistical capacity of these countries the UN Regional Commission in charge, or the regional or international organization that these countries are members of, could possibly provide needed external assistance.

Availability

In part II, chapter 2.2, we define the *DQ dimension* "*Availability"* (cf. DEF. 5, p.22). Subsequently, we refine the definitions of the *DQ measures*[100] we use to assess the availability of ANA data of countries and provide corresponding formulas (cf. part II, chapter 3.2). These measures are the 1993 and 2008 MRDS (cf. DEF. 25, p. 59 respectively DEF. 26, p. 64), the milestone levels 1 and 2 (according to "strict" and "relaxed" criteria) (cf. DEF. 28, p. 98 respectively DEF. 29, p. 101 according to the "strict" criteria, and DEF. 30, p. 105 respectively DEF. 31, p. 109 by "relaxed"

[100] The measures are based on United Nations (2001a, 2001b, 2008, and 2011) and ISWGNA (1996).

criteria). Further, we assess the availability of subsets of the required NAQ tables needed to fulfill the criteria of the previous data sets.

The employed assessment methodology follows closely the UN methodology for assessing the global data availability of ANA data. The UN assessments cover the 1993 MRDS and the milestone levels 1 and 2 (according to “relaxed” criteria).

To improve the UN methodology of the availability assessment, the time period for considered data is changed to a fixed time frame. This makes the assessment independent of previously needed supplementary information about the time of reporting of data and the scope of data reported over the last five Yearbook reporting rounds. Thus, the assessment can also be conducted based directly on the NA statistics stored in the UN NA database.

This study assesses the availability of 2008 MRDS for the first time. The methodology for the 2008 MRDS assessment is based on the 2008 MRDS definition (United Nations 2010b). For the designed 2008 MRDS assessment, the condition to classify a NAQ table as “available” is revised (cf. part III, chapter 2.2). The NAQ tables 4.1, …, 4.9 for the institutional sectors now require a specific indicator “net lending/net borrowing” of the capital account. Furthermore, the requirements to fulfill milestone levels 1 and 2 are revised to reflect the original definition of the milestone levels. Corresponding formulas are developed for our assessments. The re-revision of the milestone level 1 and 2 criteria to the original “strict” criteria is necessary because of the increased number of NAQ tables and specific NA indicators needed to fulfill the 2008 MRDS criteria. Results of the milestone levels are presented for both the currently used “relaxed” criteria and the “strict” criteria. The results of our 1993 and 2008 MRDS assessment are presented by the same country groups as of our assessment of timeliness.

The *availability assessment* we undertake *indicates* the *general capacity* of a country to compile NA statistics (i.e. the capacity to produce SNA data, regardless of the SNA version used to compute the NA statistics). Our *assessments* (cf. part III, chapter 2.3.1 for 1993 MRDS and chapter 2.3.2 for 2008 MRDS) indicate that the biggest *data gaps* are in the developing countries. Particularly, Oceania and parts of Asia have significantly low levels of reported data. This confirms the results of the previous UN 1993 MRDS assessments.

The most important *finding* from the comparison of our 1993 and 2008 MRDS assessments is that there is a huge difference regarding the availability of these two data sets (cf. part III, chapter 2.3.4). The *availability* of NA data of the 2008 MRDS is much *worse* than signaled by the UN availability assessments of the 1993 MRDS (cf. part III, chapter 2.3.3).

Our assessment indicates a slope in NA data production capacity. Nearly all UN Member States of the developed regions and transition economies show a little *data gap* to 1993 MRDS. Precisely, 89% of the developed regions have at least all but one of the NAQ tables available that are needed to fulfill the 1993 MRDS criteria. For the transition countries this still applies to 71%. The 1993 MRDS criteria are still fulfilled by 82% of the developed regions, and by 53% of the transition countries. Yet, only 71% of the developed regions and 47% of the transitions economies have at least all but one of the NAQ tables available needed to fulfill 2008 MRDS. All these countries also fulfill 2008 MRDS. Thus our assessment reveals *data gaps* that were covered up in the earlier 1993 MRDS assessments. For the developing regions this is even more evident with only 43% of the countries having at least all but one of the NAQ tables needed to fulfill 1993 MRDS, 27% in case of the 1993 MRDS, 12% in case of at least all but one of the NAQ tables of 2008 MRDS, and 7% in case of 2008 MRDS.

The smallest *data gap* among the developed regions is found for the EU 27 group, where 96% fulfill the 1993 MRDS, and 89% the 2008 MRDS criteria. The *data gap* in G20 countries is much bigger than in non-G20 advanced economies. Only 79% of G20 countries and 87% of non-G20 advanced economies fulfill the 1993 MRDS. Regarding the 2008 MRDS only 42% of G20 countries fulfill the criteria, while 65% of the non-G20 advanced economies do so.

Our availability assessment informs the international user community about the availability of ANA data by NAQ tables (i.e. the different sets of SNA data) covered by the 1993 and 2008 MRDS and the “relaxed” and “strict” milestone levels 1 and 2. The *assessment* allows *identifying* clusters of countries where *capacity building* is urgently needed. Capacity building programs, like those included in the “*2008 SNA Implementation Programme”* (cf. United Nations 2010b) set out by international and regional organizations, may wish to focus on regions in the assessment that indicate *gaps* in the *data availability*. Organizations like OECD, Eurostat, etc. could utilize the assessment results for improvements of the data availability in their domain.

Reliability

In part II, chapter 2.3, we provided the definition for the *DQ dimension* "*Reliability*" of national accounts (cf. DEF. 11, p. 29) and for the definition of the *DQ dimension* "*Accuracy*" of National Accounts (cf. DEF. 12, p. 30). We refine the wording of these definitions, which are based on Carson and Laliberté (2002, 4), and define the formula. For our revision study we employ the Mean Absolute Revision (MAR) (cf. DEF. 19, p. 46) and the Relative Mean Absolute Revision (RMAR) (cf. DEF. 20, p. 47). We merely refine the definition of these two *DQ measures* (cf. part II, chapter 3.1).

In part III, chapter 3, we first analyzed the deficiencies of current *assessment methodologies*, and suggested an *improved approach*. The time period considered for our revision study is more recent and often longer than that of comparable assessments. The *most important modification* for improving the results for the revision analyses is the exclusion of major comprehensive revisions from the analyzed dataset. Those interfere with the results of previous comparable studies. These are based on the Main Economic Indicators (MEI) database of the OECD and the spreadsheet tools developed by Di Fonzo (2005a), which we slightly enhance to facilitate the analysis of other revision intervals. The analyzed data sets only include comprehensive revisions that are typically carried out every five years. All data are compiled according to ESA95/ 1993 SNA, respectively according to 1968 SNA, in case of Turkey. Thus no revisions occur across SNA series versions. The number of considered revision intervals was extended compared to previous studies to provide insight on: 1) one short term revision interval, 2) three annual intervals, and 3) an additional interval that captures the "latest" data. The "latest" revision time point is improved to refer to vintage at least four years after the first released vintage. Therefore, this interval usually captures "final" data, while earlier revision time points may more often capture data that are not already finalized. Another important change is the observation of Year-on-Year (YoY) growth rates instead of Quarter-on-Quarter (QoQ) growth rates as used in previous studies. YoY growth rates are most appropriate for international comparison (cf. Cheem 2003).

Our revision study (cf. part III, chapter 3.4) yields better and updated results for more revision intervals. The most important extension compared with other assessments of is the *empirical revision analysis* of 34 countries. First-time results in a cross-country comparison are presented for 14 countries, including four non-OECD G20 countries.

Our study covers more countries than any previous study before (previous studies based on MEI OECD data covered at most 20 countries, the first only 7 countries).

Our *revision analysis shows* significant differences between countries regarding the *reliability* of preliminary estimates (cf. part III, chapter 3.4.1). Consistent with other studies, the *revision needs* generally increase over time, i.e. they are larger at the revision time point of the three years later vintage than at an earlier revised vintage. By changing the assessment of *revisions* from an absolute to a relative size, the order of countries that produce most *reliable* or least *reliable* preliminary estimates changes, too. For example, the Mean Absolute Revision (MAR) for India is big (one of the last ranked countries), while the RMAR is small (one of the first ranks). For Portugal the order of MAR and RMAR is opposite to that of India, *showing* Portugal among the better countries by size of MAR, but among the last ranked countries by size of RMAR.

Compared to previous assessments, the size of a country's revision measure is typically smaller in our study (particularly in case of RMAR). This is expected due to the exclusion of major comprehensive revisions. Thus our approach leads to revision measures that provide more accurate information of the *reliability* of initial released YoY growth rates (cf. part III, chapter 3.5). Our study informs users of these data about the data *reliability* of countries. Our cross-country comparison allows identification of countries with strong and weak performance. Moreover, users may decide to use data produced at later revision time points instead of the first released vintage.

Perspectives

Our assessment of *timeliness* of ANA data publication can be further improved by more *subdivisions* of *regions* (e.g. for Africa). Also additional regional *organizations* can be considered like the Secretariat of the Pacific Community (SPC). (Refer to part III, chapter 1.2.5.1 with suggestions of different subdivisions of regions as a perspective for further improvements.)

Moreover, we could identify those countries that report data annually and those that *do not report* data *annually*. The *timeliness* of countries with "known" ANA data could be assessed based on the *last reporting* of data. We outline this as a perspective for improvements in part III, chapter 1.2.5.2. We could also analyze the *timeliness* of *individual data sets* (NAQ tables representing different SNA concepts), e.g. providing information on differences in the *timeliness* of current and constant price data, or of the institutional sectors. We outline this in part III, chapter 1.2.5.3.

The *availability assessment* could be *extended* to include the *financial accounts* which are collected by the UN. The NA data collected by the UN does not include the *balance sheets* by institutional sectors (i.e. NAQ tables 4.1, …, 4.9) and not *quarterly national accounts* either; otherwise the *availability* assessment could in principle be extended to include these data sets. The data *availability* of these data sets is currently of particular interest. Research on their *availability* was carried out by the IMF during 2009 to 2011 for the world's most important economies (cf. IMF 2011a and 2011b).

Regarding *revision analysis*, the OECD MEI database, which provides the "real time dataset", i.e. the "revision triangles" [101], needs to indicate any changes in the methodologies used to compute the data vintages. This would be most appropriately addressed by inclusion of series version numbers (indicating different methodologies) as metadata to the statistics. A new series version number needs to be assigned, with any change in the country's NA production process with regard to concepts, methods, or data sources.

To completely eliminate any impacts from changes in concepts, methods and data sources, besides the desired regular (current) revisions due to inclusion of more complete data, all later vintages to a first estimate (i.e. preliminary and final vintages) that are compiled by a new methodology or with changed data sources should be eliminated. However, while such practice would provide the best data to compute the *reliability* of first released estimates, it considerably limits the number of observations for the different revisions intervals.

[101] Linked to the retention of revisions data is also data availability, i.e. the availability of preliminary estimates of GDP and other NA indicators at different vintages. It is justifiably vital to retain and publish the different vintages of data. The *number of NA indicators* included in the *revisions* database for MEIs at the OECD should be extended to include all NA indicators covered by the 2008 MRDS. This would facilitate e.g. assessments of revisions of GDP in conjunction with the size of revisions of other NA indicators and analyze the NA indicators which contribute most to revision of GDP.

Bibliography

ABS (2003), A Quality Assessment of the National Accounts Information System. In: National Income, Expenditure and Product, March Quarter 2003 (cat. no. 5206.0). Australian Bureau of Statistics, Canberra. Online at: http://www.abs.gov.au/AUSSTATS/abs@.nsf/ProductsbyReleaseDate/3F5AB28F7DAF2E63CA2570A300833082?OpenDocument. Last accessed: 13-07-2012.

Ahmad, N., Bournot, S. and Koechlin, F. (2004), Revisions to Quarterly GDP Estimates: A Comparative Analysis for Seven Large OECD Countries. OECD Paper presented at OECD-ONS Workshop, Assessing and improving statistical quality – Revisions Analysis for the National Accounts, 7-8 October 2004. Online at: http://www.oecd.org/dataoecd/20/26/34350524.pdf. Last accessed: 23-05-2012.

Aspden, C. (1990), Which is the best short-term measure of Gross Domestic Product? In: Australian National Accounts: National Income and Expenditure, June Quarter 1990, (cat. no. 5206.0). Australian Bureau of Statistics, Canberra, 57-65. Online at: http://www.ausstats.abs.gov.au/ausstats/free.nsf/0/A883157CBF2E2E21CA2574DC0016C234/$File/52060_1990_JUN.pdf. Last accessed: 23-05-2012.

Batini, C. and Scannapieco, M. (1998), Data Quality. Springer, Berlin, Heidelberg, New York.

Bier, W. and Ahnert, H. (2001), Trade-Off between Timeliness and Accuracy: ECB Requirements for General Economic Statistics. (Published in Dutch in Economisch Statistische Berichten (ESB), March 15, 2001, No. 4299.) As English version online at: http://dsbb.imf.org/images/pdfs/bierahnert.pdf. Last accessed: 27-10-2012.

Borowski, E. and Lenz, H.-J. (2008), Design of a Workflow System to Improve Data Quality Using Oracle Warehouse Builder. JAQM Vol.3 (Issue 3), September 2008, 198-206.

Brown, G., Buccellato, T., Chamberlin, G., Dey-Chowdhury, S., and Youll R. (ONS) (2009), Understanding thequality of early estimates of Gross Domestic Product. Economic & Labour Market Review, Vol 3, No 12, December 2009, 43-50.

Carson C. S. and Laliberté L. (2002), Assessing accuracy and reliability: A note based on approaches used in National Accounts and Balance of Payments statistics. IMF Working Paper 02/24. IMF, Washington D.C.

Cheem, N. P. (2003), Quarterly Growth Rates. Statistics Singapore Newsletter, March 2003. Singapore Department of Statistics, Singapore. Also at: http://www.singstat.gov.sg/pubn/papers/economy/ssnmar03-pg7-10.pdf.

Concise Oxford Dictionary (1979). Oxford University Press, Oxford.

Crosby, P. B. (1979), Quality Is Free. New York, McGraw-Hill.

Croushore, D. (2011), Frontiers of Real-Time Data Analysis. Journal of Economic Literature 49 (March 2011), 72-100. Also at: https://facultystaff.richmond.edu/~dcrousho/docs/Croushore%20Real%20Time%20Frontiers%2009Dec.pdf.

Denton, F. and Oksanen, E. (1973), Data Revisions and Forecasting Accuracy: An Econometric Analysis Based on Preliminary and Revised National Accounting Estimates. Review of Income and Wealth, 1973, vol. 19, issue 4, 437-452. Also at: http://www.roiw.org/1973/437.pdf.

Deutsche Bundesbank (2011), Verlässlichkeit und Revisionsmuster ausgewählter deutscher Konjunkturindikatoren. Monatsbericht, Juli 2011. Deutsche Bundesbank, Frankfurt, 53-67.

Di Fonzo, T. (2005a), The OECD project on revisions analysis: First elements for discussion. Paper presented at the OECD STESEG Meeting, Paris, 27-28 June 2005. Also at: http://www.oecd.org/dataoecd/55/17/35010765.pdf. Corresponding revision spreadsheets for revision analysis available as download at http://www.oecd.org/std/automatedprogramstoperformrevisionsanalysis.htm, last accessed 28-10-2013.

Di Fonzo, T. (2005b), Revisions in quarterly GDP of OECD countries. STD/NAES(2005)24. Paper presented at the OECD Working Party on National Accounts, Paris, 11-14 October 2005. Also at: http://www.oecd.org/dataoecd/13/49/35440080.pdf.

Dodge, Y. (2006), The Oxford Dictionary of Statistical Terms. Oxford University Press, Oxford, New York.

Branchi, M. et al. (2007), Analysis of revisions to general economic statistics. Occasional Paper Series, No 74 / October 2007. ECB, Frankfurt.

EFC and ECOFIN Council (2010), Status Report on Information Requirements in EMU. Eurostat, Brussels. Online at: http://www.ecb.int/stats/pdf/statusreport2010.pdf?b73e705b2ed2aed0c8adddf7c55fd606. Last accessed: 07-07-2011.

Ehling, M. and Körner, T. (eds) (2007), Handbook on Data Quality Assessment Methods and Tools. Eurostat, Wiesbaden.

European Parliament and Council of the European Union (2007), Regulation (EC) No 1392/2007 of the European Parliament and the Council of 13 November 2007. In: Official Journal of the European Union L 324/1. Online at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:324:0001:0078:EN:PDF. Last accessed 06-13-2011.

Eurostat (2003a), Definition of quality in statistics, Document nr: Eurostat/A4/Quality/03/General/Definition. Eurostat, Luxembourg.

Eurostat (2003b), Standard Quality Report, Document nr: Eurostat/A4/Quality/03/General/Standard Report. Eurostat, Luxembourg.

Eurostat (2003c), Glossary of Quality Terms, Document nr: Eurostat/A4/Quality/03/General/Glossary. Eurostat, Luxembourg.

Eurostat (2009a), ESS Handbook for Quality Reports. Eurostat, Luxembourg.

Eurostat (2009b), Focus on: Revision Analysis. Selected Readings, January 2009. Online: http://epp.eurostat.ec.europa.eu/cache/ity_public/ln-sr012009/en/ln-sr012009-en.pdf. Last accessed: 03-06-2012.

FAO (2004), FAO Statistical Data Quality Framework. FAO, Rome Paper presented at Conference on Data Quality for International Organizations Committee for the Coordination of Statistical Activities Wiesbaden/Germany, 27-28 May 2004. Also at: http://faostat.fao.org/Portals/_Faostat/documents/data_quality/FAO_Data_Quality_Paper_Wiesbaden.pdf.

FAO (2005), Integrating Data Quality Indicators Into the FAO Statistical System. Quality in Statistics, Eurostat, Seventh meeting, Luxembourg, 23-24 May 2005.

Fernandez, A. Z., Koening, E. F., and Nikolsko-Rzhevskyy, A. (2011), A Real-Time Historical Database for the OECD. Federal Reserve Bank of Dallas, Working Paper No. 96. Also at: http://www.dallasfed.org/assets/documents/institute/wpapers/2011/0096.pdf.

Fixler, D. J. and Grimm, B. T. (2002), Reliability of GDP and Related NIPA Estimates. SURVEY OF CURRENT BUSINESS 85 (January 2002), 9-27; Online at: www.bea.gov/bea/pubs.htm, also: http://www.bea.gov/scb/pdf/2005/02February/0205_NIPAs.pdf, last accessed 23-04-2012.

Fixler, D. J. and Grimm, B. T. (2005), Reliability of the NIPA Estimates of U.S. Economic Activity. SURVEY OF CURRENT BUSINESS 85 (February 2005), 8-19. Online at: www.bea.gov/bea/pubs.htm, also: http://www.bea.gov/scb/pdf/2005/02February/0205_NIPAs.pdf. Last accessed: 25-03-2012.

Fixler, D. J. and Nalewaik, J. J. (2004), News, Noise and Estimates of the "True" Unobserved State of the Economy. BEA Working Paper, Working Paper 2004-08. BEA, Washington DC. Also at: http://www.bea.gov/papers/pdf/fixler_nalewaikrev.pdf.

Fixler, D. J. and Nalewaik, J. J. (2010), News, Noise and Estimates of the "True" Unobserved State of the Economy. (WP2010-04 updated WP2004-08). BEA Working Paper 2010-04. BEA, Washington DC. Also at: http://www.bea.gov/papers/pdf/fixler_nalewaikrev2.pdf.

Fixler, J. D., Greenaway-McGrevy, R. and Grimm, B. T. (2011), Revisions to GDP, GDI, and Their Major Components. SURVEY OF CURRENT BUSINESS 91 (July 2011), 8-19. Online at: www.bea.gov/bea/pubs.htm, also: http://www.bea.gov/scb/pdf/2011/07%20July/0711_revisions.pdf. Last accessed: 06-06-2012.

FSB and IMF (2009), The Financial Crisis and Information Gaps: Report to the G-20 Finance Ministers and Central Bank Governors. IMF, Washington D.C. Also at: http://www.financialstabilityboard.org/publications/r_091107e.pdf and http://www.imf.org/external/np/g20/pdf/102909.pdf.

Grimm, B. T. and Weadock, T. L. (2006), Gross Domestic Product: Revisions and Source Data. SURVEY OF CURRENT BUSINESS 86 (February): 11–15. Available online at: http://www.bea.gov/scb/index.htm, also: http://www.bea.gov/scb/pdf/2006/02february/0206_gdp.pdf, last accessed 06-04-2013.

Holdren, A. E. and Grimm, B. T. (2008), Gross Domestic Product: Revisions and Source Data. SURVEY OF CURRENT BUSINESS 88 (December 2008), 14–20. Available online at: http://www.bea.gov/scb/index.htm, also: http://www.bea.gov/scb/pdf/2008/12%20December/1208_gdi.pdf, last accessed 06-04-2013.

Hoven (2008), Using results from revisions analysis to improve compilation methods: a case study on revisions of Dutch estimates of GDP volume growth. Contribution to the joint OECD/Eurostat Task Force on "Performing Revisions Analysis for Sub-Annual Economic Statistics". Online at: http://www.oecd.org/dataoecd/44/41/40309550.pdf. Last accessed: 04-06-2012.

IMF (2003a), Data Quality Assessment Framework (DQAF). IMF, Washington D.C. Online at: http://dsbb.imf.org/images/pdfs/dqrs_Genframework.pdf. Last accessed: 10-06-2012. Also: http://dsbb.imf.org/Applications/web/dqrs/dqrsdqaf/ with reference to the different DQAF frameworks.

IMF (2003b), Data Quality Assessment Framework (DQAF) for National Accounts Statistics. IMF, Washington D.C. Also at: http://dsbb.imf.org/images/pdfs/dqrs_nag.pdf.

IMF (2011a), A Status on the Availability of Sectoral Balance Sheets and Accumulation Accounts in G-20 Economies. Conference "Strengthening Sectoral Position and Flow Data in the Macroeconomic Accounts". Jointly organized by the IMF and the OECD February 28 - March 2, 2011. IMF, Washington D.C.

IMF (2011b), A Status on the Availability of Sectoral Balance Sheets and Accumulation Accounts in Advanced Economies not Represented by Membership in G-20 Economies. Conference "Strengthening Sectoral Position and Flow Data in the Macroeconomic Accounts". Jointly organized by the IMF and the OECD February 28-March 2, 2011. IMF, Washington D.C.

IMF (2012), IMF Occasional Paper: April 2011-February 2012. TFFS 12/05. IMF, Washington D.C. Also at: http://www.tffs.org/pdf/meet/2012/imf0312.pdf.

ISO 8402-1986 Quality-Vocabulary (1986). International Standards Organisation (ISO)

ISWGNA (1996), Milestones of the SNA Implementation. SNA News and Notes, Issue 4, July 1996, 1-2. Online at: http://unstats.un.org/unsd/nationalaccount/sna/nn4_en.pdf. Last accessed: 08-06-2012.

Jarke M., Lenzerini, M., Vassiliou, Y. and Vassiliadis P. (1999), Fundamentals of Data Warehouses. Springer Verlag, Berlin, New York.

Juran, J. M. (Ed.) (1974), Quality Control Handbook, 3rd ed. McGraw-Hill, New York.

Karr, A. F., Sanil, A. P. and Banks D. L. (2005), Data Quality: A Statistical Perspective. National Institute of Statistical Sciences, Technical Report Number 151. NISS, Research Triangle Park. Also at: http://www.niss.org/sites/default/files/pdfs/technicalreports/tr151.pdf.

Landefeld, S., Seskin, E., and Fraumeni, B. (2008), Taking the Pulse of the Economy: Measuring GDP. Journal of Economic Perspectives, Volume 22, Number 2, Spring, Pages, 193–216.

Laliberté, L., Grünewald, W. and Probst, L. (2004), Data Quality: A Comparison of IMF's Data Quality Assessment Framework (DQAF) and Eurostat's quality definition. IMF, Eurostat. Online at: dsbb.imf.org/vgn/images/pdfs/dqaf_eurostat.pdf. Last accessed: 24.03.2012.

Marriott, F. H. C. (1990), A dictionary of statistical terms, Fifth edition. Longman Scientific & Technical, New York.

Mazzi, G. L. (2001), European short-term statistics: the state of the art. Eurostat A6, 11th CEIES seminar. European Communities, Luxembourg. Also at: http://www.oecd.org/dataoecd/22/58/30044960.pdf.

McKenzie, R. (2006a), Undertaking Revisions and Real-Time Data Analysis using the OECD Main Economic Indicators Original Release Data and Revisions. OECD Statistics Working Paper STD/DOC(2006)2, OECD, Paris. Also at: http://www.oecd.org/officialdocuments/displaydocumentpdf/?cote=std/doc(2006)2&doclanguage=en.

McKenzie, R. (2006b). Performing Revisions and Real-Time Data Analysis. OECD Statistics Brief, Issue 12, November 2006, Paris. Also at: http://www.oecd.org/dataoecd/46/48/37669085.pdf.

McKenzie, R. and Adam, Z. (2007) Revisions in Quarterly GDP of OECD Countries: An Update. Working Party of National Accounts, October 2007, OECD, Paris.

McKenzie, R. and Gamba, M. (2008). Interpreting the results of Revision Analyses: Recommended Summary Statistics. Contribution to the OECD/Eurostat Task Force on "Performing Revisions Analysis for Sub-Annual Economic Statistics". Online at: http://www.oecd.org/dataoecd/47/18/40315546.pdf. Last accessed: 25-04-2012.

Naumann, F. (2002), Quality-Driven Query Answering for Integrated Information Systems. Lecture notes in computer science (LNCS), 2261, Springer-Verlag, Berlin, Heidelberg.

Van Nederpelt, Ir. W. M. EMEA (2009), Checklist Quality of Statistical Output. Statistics Netherlands, Den Haag. Also available online: http://www.cbs.nl/NR/rdonlyres/4119715F-7437-4379-9A70-90A0893F949E/0/2009ChecklistQualityofStatisticalOutput.pdf.

Neumayr, H. (2010), Revisions to GDP estimates in Luxembourg, Performing Revisions Analysis by using an OECD Toolset. Working papers du STATEC. Economie et statistiques, No. 39/2010. STATEC, Luxembourg.

OECD (2000), 1993 System of National Accounts – Glossary. OECD, Paris.

OECD (2003), Quality Framework and Guidelines for OECD Statistical Activities. STD/QFS(2003)1. OECD, Paris.

OECD (2006), Minutes of the Working Party on National Accounts. Working Party on National Accounts, STD/CSTAT/WPNA/M(2006)1, 10-12 October 2006. OECD, Paris.

OECD (2007a), Data and Metadata Reporting and Presentation Handbook. OECD, Paris.

OECD (2007b), Glossary of Statistical Terms. OECD, Paris. Available online as searchable glossary at: http://stats.oecd.org/glossary/. Last accessed 27-10-2012.

OECD (2012a), Quality Framework and Guidelines for OECD Statistical Activities. STD/QFS(2011)1. OECD, Paris.

OECD (2012b), Revisions analysis, Main Economic Indicators (database). Doi: 10.1787/data-00053-en. Respectively “MEI Original release data and revisions”: Online at: http://stats.oecd.org/Index.aspx?querytype=view&queryname=206#. Last accessed: 08-05-2012.

ONS (2012), Quarterly National Accounts, 4th quarter 2011. Office for National Statistics, Newport. Also available online via: http://www.ons.gov.uk/ons/publications/all-releases.html?definition=tcm:77-23645. Last accessed 27-10-2012.

ONS (2013), FAQ on Quality Measurement and Quality Assurance. Online: http://www.ons.gov.uk/ons/guide-method/method-quality/quality/quality-information/faq-on-qm-and-qa/index.html. Last accessed 10-01-2013.

Scannapieco, M. and Catarci, T. (2002), Data Quality under the Computer Science Perspective. Journal of "Archivi & Computer", Volume 2, 2002.

Scannapieco, M., Pernici, B. (2003), Data Quality in Web Information Systems. Journal of Data Semantics, Volume 1 (LNCS 2800). Springer-Verlag, Berlin, Heidelberg.

Statistics Canada (1998), Statistics Canada's Quality Guidelines, 3rd Edition – October 1998. Statistics Canada, Ottawa.

Statistics Canada (2002), Statistics Canada's Quality Assurance Framework. Statistics Canada, Ottawa.

Statistics Finland (2007), Quality Guidelines for Official Statistics (2nd Revised Edition). Multiprint, Helsinki 2007. Also at: http://stat.fi/meta/qg_2ed_en.pdf.

Statistisches Bundesamt (2008), Volkswirtschaftliche Gesamtrechnungen, Inlandsproduktsberechnung, Detaillierte Jahresergebnisse, 2007, Qualitätsbericht. Fachserie 18 Reihe 1.4 - 2007. Statistisches Bundesamt, Wiesbaden.

Statistisches Bundesamt (2011), Volkswirtschaftliche Gesamtrechnungen, Inlandsproduktsberechnung, Detaillierte Jahresergebnisse, 2010, Qualitätsbericht. Fachserie 18 Reihe 1.4 - 2010. Statistisches Bundesamt, Wiesbaden.

Statistisches Bundesamt (2013), Volkswirtschaftliche Gesamtrechnungen, Inlandsproduktsberechnung, Vierteljahresergebnisse, 4. Vierteljahr 2012. Fachserie 18 Reihe 1.2. Statistisches Bundesamt, Wiesbaden.

TechTerms.com, 2011, Software terms (2011). Online at: http://www.techterms.com (http://www.techterms.com/definition/version_control). Last updated: 23-08-2011. Last accessed: 27-04-2012.

Tosetto, E., and Brackfield, D. (2009), REVISIONS IN QUARTERLY GDP OF OECD COUNTRIES: AN UPDATE. International Seminar on Timeliness, Methodology and

Comparability of Rapid Estimates of Economic Trends 27 – 29 May 2009, Ottawa, Canada. Online at: http://unstats.un.org/unsd/nationalaccount/workshops/2009/ottawa/AC188-S62a.PDF. Last accessed: 04-08-2012.

Tosetto, E. and Lequiller, F. (2006), Revisions in Quarterly GDP of OECD Countries: An Update. For presentation at the OECD Working Party, Paris, 10-12 October 2006, STD/CSTAT/WPNA(2006)9. Also at: http://www.oecd.org/dataoecd/2/17/37460278.pdf.

UNECE (2000), Glossary of Terms on Statistical Data Editing. United Nations, Geneva. Also in “K-Base” online: http://www1.unece.org/stat/platform/display/kbase/Glossary. Last accessed: 16-10-2013.

United Nations (1993), System of National Accounts 1993. Commission of the European Communities-Eurostat, IMF, OECD, United Nations and World Bank (eds). United Nations, New York.

United Nations (1999), Milestone assessment of the implementation of the System of National Accounts, 1993, by member States. E/CN.3/1999/3. Statistical Commission, Thirtieth session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2000), Milestone assessment of the implementation of the System of National Accounts, 1993, by member States. E/CN.3/2000/3. Statistical Commission, Thirty-first session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2001a), Assessment of the implementation of the System of National Accounts 1993. E/CN.3/2001/8. Statistical Commission, Thirty-second session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2001b), Report of the Task Force on National Accounts. E/CN.3/2001/7. Statistical Commission, Thirty-second session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2001c), Statistical Commission Report on the thirty-second session (6-9 March 2001). E/CN.3/2001/25. Statistical Commission, Thirty-second session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2004), Report of the Intersecretariat Working Group on National Accounts. E/CN.3/2004/10. Statistical Commission, Thirty-fifth session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2005), Report of the Intersecretariat Working Group on National Accounts. E/CN.3/2005/4. 2005 Statistical Commission, Thirty sixth session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2008), International Standard Industrial Classification of All Economic Activities, Revision 4. (ISIC Rev. 4). United Nations, New York.

United Nations (2009a), International Recommendations for Distributive Trade Statistics 2008. ST/ESA/STAT/SER.M/89. United Nations, New York.

United Nations (2009b), National Accounts Main Aggregates and Detailed Tables 2008, Part I. United Nations, New York.

United Nations (2009c), National Accounts Main Aggregates and Detailed Tables 2008, Part II. United Nations, New York.

United Nations (2009d), National Accounts Main Aggregates and Detailed Tables 2008, Part III. United Nations, New York.

United Nations (2009e), National Accounts Main Aggregates and Detailed Tables 2008, Part IV. United Nations, New York.

United Nations (2009f), System of National Accounts 2008. European Commission, IMF, OECD, UN, World Bank (eds). United Nations, New York. Also at: www.unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf.

United Nations (2010a), Main Aggregates and Detailed Tables 2009, Part I - V. United Nations, New York.

United Nations (2010b), Report of the Intersecretariat Working Group on National Accounts, E/CN.3/2010/8. Statistical Commission, Forty-first session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2011), Report of the Intersecretariat Working Group on National Accounts, E/CN.3/2011/6. Statistical Commission, Forty-second session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2012), Report of the Intersecretariat Working Group on National Accounts, E/CN.3/2012/4. Statistical Commission, Forty-third session, United Nations Economic and Social Council, United Nations, New York.

United Nations (2013), Report of the Intersecretariat Working Group on National Accounts, E/CN.3/2013/4. Statistical Commission, Forty-fourth session, United Nations Economic and Social Council, United Nations, New York.

UNSD (2003), MDG regional classification. Online at: http://mdgs.un.org/unsd/mdg/Resources/Static/Data/Regional%20groupings.doc and http://mdgs.un.org/unsd/mdg/Resources/Static/Data/MDGRegionCodes_200611.xls. Last accessed 11-06-2013.

UNSD (2006), Note on definition of regions for statistical analysis. SA/2006/15. Committee for the Coordination of Statistical Activities, Eighth Session, Montreal, 4-5 September 2006. Also at: http://unstats.un.org/unsd/accsub/2006docs-8th/SA-2006-15-Grouping-UNSD.pdf. Last accessed: 16-10-2013.

UNSD (2008), Reporting of National Accounts Data to the United Nations Statistics Division. SNA News and Notes, Issue 25/26, May 2008, 6-9. Online at: http://unstats.un.org/unsd/nationalaccount/sna/nn25-26-en.pdf. Last accessed: 09-09-2011.

UNSD (2009a), National Accounts Questionnaire - Supporting Booklet. UNSD, New York. Also at: http://unstats.un.org/unsd/pubs/questionnaires/Supporting%20Booklet%202009%20-%20English.pdf.

UNSD (2009b), Global assessment on the availability, periodicity, timeliness and dissemination of high frequency indicators, Bk.4. Workshop on International Economic and Social Classifications, Bamako, Mali 25-28 January 2010. United Nations, New York. Also at: http://unstats.un.org/unsd/class/intercop/training/eca10/ac211-Bk4.PDF.

UNSD (2009c), Report Ottawa: International Seminar on Timeliness, Methodology and Comparability of Rapid Estimates of Economic Trends, 27 – 29 May 2009, Ottawa, Canada. United Nations, New York. Online at: http://unstats.un.org/unsd/nationalaccount/workshops/2009/ottawa/AC188-5.PDF. Last accessed: 07-09-2011.

UNSD (2010), Report of the meeting. International Seminar on Early Warning and Business Cycle Indicators, ESA/STAT/AC. 202/5, 14 to 16 December 2009 Scheveningen, The Netherlands. United Nations, New York. Online at: http://unstats.un.org/unsd/nationalaccount/workshops/2009/netherlands/AC202-5.PDF. Last accessed: 06-09-2011.

UNSD (2012), Sample Questionnaire (1993 National Accounts Questionnaire). UNSD, New York. Online at: http://unstats.un.org/unsd/nationalaccount/docs/SampleQuestionnaire.xls. (Last updated: 28-11-2012). Last accessed: 16-05-2013.

UNSD (2013a), Countries or areas, codes and abbreviations. Online at: http://unstats.un.org/unsd/methods/m49/m49alpha.htm. (Status: 11-02-2013.) Last accessed: 16-10-2013.

UNSD (2013b), Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings. Online at: http://unstats.un.org/unsd/methods/m49/m49regin.htm. (Status: 11-02-2013.) Last accessed: 16-10-2013.

UNSD (2013c), UNdata - A world of information (UN data dissemination platform). Online at: http://data.un.org/. Last accessed: 16-10-2013.

Weisman, E., Balyozov, Z. and Venter, L. (2010), IMF's Data Quality Assessment Framework. For a Conference on Data Quality for International Organizations, Helsinki, Finland, 6 and 7 May 2010. IMF, Washington DC.

Young, A. H. (and Hirsch, A. McCully, C., Parker, R., Weadock, T., and Ziemer, R.) (1993), Reliability and Accuracy of the Quarterly Estimates of GDP. SURVEY OF CURRENT BUSINESS (October 1993), 29-43. Online at: http://www.bea.gov/scb/index.htm, also: http://www.bea.gov/scb/pdf/1993/1093cont.pdf. Last accessed: 27-10-2012.

Zieschang, K. D. (2005), IMF Statistical Metadata Standards and Classifications: Enhancing The Effectiveness Of Statistical Capacity Development. Conference of Commonwealth Statisticians, Cape Town, South Africa, September 5-9, 2005.

Zusammenfassung (German Summary)

Dieses Kapitel fasst die Hauptpunkte und Ergebnisse der Dissertation zusammen.

Wir untersuchen die Aktualität, Verfügbarkeit und Zuverlässigkeit von Daten der Volkswirtschaftlichen Gesamtrechnung (VGR) die von Ländern an internationale Organisationen gemeldet und von diesen veröffentlicht werden.

Datenqualitätsdimensionen und Maße

Hinführend zu den Definitionen der Datenqualitätsdimensionen beschreiben wir die drei Standardrahmenwerke zur Datenqualität in der Offiziellen Statistik: das des IMF, Eurostat und der OECD. Wir definiert und diskutieren die Datenqualitätsdimensionen und Maße die für unsere Studien der Aktualität, Verfügbarkeit und Zuverlässigkeit von VGR Daten relevant sind. Wir verbessern einige der bestehenden Definitionen. Wo Definitionen und Formeln nicht bereits existieren werden eigne beigebracht, dieses betrifft insbesondere die Datenqualitätsmaße.

Die Aktualität bezieht sich auf den Zeitraum bis der Datenveröffentlichung, d.h. auf den Zeitraum in Jahren zwischen dem Referenzjahr und dem Veröffentlichungsjahr der jährlichen VGR Statistiken auf internationaler Ebene.

Die VGR Daten, die das „Minimum Requirement Data Sets“ (MRDS), d.h. den erforderlichen „Mindestdatensatz“ ausmachen sind Grundlage unsere Studie der Datenverfügbarkeit. Das ursprüngliche MRDS wurde 2001 auf der Grundlage des 1993 SNA international beschlossen (siehe: United Nations 2001b). Das MRDS wurde mit der Einführung des 2008 SNA revidiert (siehe: United Nations 2011). Wir bezeichnen diese zwei MRDS als 1999 MRDS bzw. 2008 MRDS. Das 1993 und 2008 MRDS bestehen aus NAQ Tabellen, d.h. VGR Tabellen im Fragebogen der Vereinten Nationen (VN), die für ein Land existent sein, d.h. produziert werden müssen. Es gibt verschiedene Kombinationsmöglichkeiten der NAQ Tabellen um die 1993 bzw. 2008 MRDS Kriterien zu erfüllen. Wir beschreiben das 1993 und 2008 MRDS formal. Durch unsere Formeln definieren wir auch die Bedingungen für die „Verfügbarkeit“ einer NAQ Tabelle. Im Fall des 1993 MRDS reicht bereits die Verfügbarkeit irgendeines Wertes innerhalb der NAQ Tabelle aus, d.h. es ist kein spezieller VGR Indikator erforderlich. Für die Wirtschaftssektoren ist die Bedingung des 2008 MRDS die Existenz der Kontentabellen bis zum Indikator „net lending / net borrowing“, d.h. bis zum „Finanzierungssaldo“.

Die Meilenstein Stufen definieren die Produktionskapazität unterschiedlicher VGR Konzepte durch die Länder. Die Meilensteinstufen 1 und 2 können auch anhand der NAQ Tabellen gemessen werden. Wir beschrieben diese formal gegeben der „relaxed“ d.h. der milden und gegeben der „strict“, d.h. der strengen Version der Kriterien zur Erfüllung dieser Meilenstein Stufen. Die Anwendung der strengen Kriterien erweist sich aufgrund der erweiterten Anforderungen des 2008 MRDS als notwendig.

Die OECD verfügt über eine „Revisionsdatenbank“, die je Land die BIP Werte eines Referenzquartals zu verschiedenen Veröffentlichungszeitpunkten enthält. Diese ermöglichte es den Revisionsbedarf des erstveröffentlichten BIP Datums zu späteren Veröffentlichungszeitpunkten zu analysieren. Wir berechnen für einen Beobachtungszeitraum die Mittlere Absolute Revision „Mean Absolute Revision“ (MAR) und die Relative Mittlere Absolute Revision „Relative Mean Absolute Revision“ (RMAR) zu verschiedenen Veröffentlichungszeitpunkten nach der ersten Veröffentlichung.

Datenqualitätsstudien

Unsere drei Studien der Qualität von VGR Daten bestehen jeweils aus zwei Teilen, einem methodischen und einem empirischen. Im methodologischen Teil werden bestehende Ansätze zur Messung der Qualität von VGR Daten analysiert und wir schlagen verbesserte Konzepte vor. Auf den vorgeschlagenen verbesserten Konzepten basierend werden im empirischen Teil unsere Messungen zum Status der Datenqualität durchgeführt. Wir vergleichen anschließend unsere Ergebnisse zur „Aktualität“ und „Verfügbarkeit“, sowie zur „Zuverlässigkeit“ mit Ergebnissen vorheriger Studien.

Für die Messung von „Aktualität“ und „Verfügbarkeit“ untersuchen wir die Ansätze der VN, die das Mandat hat die weltweit verfügbaren jährlichen VGR Daten zu sammeln und zu veröffentlichen. Die „Aktualität“ und „Verfügbarkeit“ jährlicher VGR Daten wird gemessen und analysiert für verschiedene Regionen und Ländergruppen. Im Fall der „Zuverlässigkeit“ untersuchen wir den Ansatz der OECD.

Wir berechnen individuelle Ergebnisse für 30 OECD Länder sowie für vier G20 Staaten, die nicht OECD Mitglieder sind.

Die Studien sind insbesondere wertvoll für Datennutzer (z.B. Analysten und Wissenschaftler), die Information über den Status der Datenqualität für ihre Forschungen benötigen. Für internationale Organisationen sind diese Studien wertvoll um schwache Regionen und Länder zu identifizieren, da solche der Fokus für die Entwicklungsarbeit in der VGR sind (siehe United Nations 2000 und 2001c).

Aktualität

Die Messung der „Aktualität" gemeldeter VGR Daten wird primär dadurch verbessert, dass nur die Länder berücksichtigt werden, die tatsächlich neue Daten gemeldet haben. Somit misst unser Ergebnis wirklich die „Aktualität" zum Publikationszeitpunkt der Daten durch die VN. Zuvor (siehe Studie: UNSD 2008) wurde hingegen bemessen wie „aktuell" die im Jahrbuch enthalten Daten sind, indem das letztverfügbare Referenzjahr eines Landes gemessen wurde, unabhängig davon ob diese Daten bereits im vorherigen Jahrbuch veröffentlicht wurden und somit lediglich bereits „bekannte" Daten nochmals veröffentlicht werden.

Wir messen für alle Länder, die neue VGR Daten an die VN geliefert haben, das letztverfügbare Kalenderjahr. Die Studie der VN berücksichtigt nur die Länder, die Daten der Vorperiode (t-1, mit t = derzeitiges Jahr) oder der Vor-vorperiode (t-2) geleifert haben. Unsere Ergebnisse liefern die absolute Anzahl an Ländern pro Referenzjahr, sowie die Prozente basierend auf der Anzahl an Ländern die neue Daten geliefert haben. Letzteres ist eine wichtige Verbesserung gegenüber der VN Studie, die nur Prozentwerte liefert. In der VN Studie basieren diese auf der Gesamtzahl an Ländern der entsprechenden Region, von denen nicht immer alle Länder neue Daten gemeldet haben und teilweise gar keine VGR Daten produzieren.

Die Studie umfasst drei Teilen mit verschiedenen Regionen:

1) UN Mitgliedsstaaten,

2) „alle Länder", d.h. die VN Mitgliedsstaaten sowie 32 nicht VN Länder, und

3) andere Organisationen und Gruppen.

Anhand unserer der „Aktualität" lassen sich Regionen und Ländergruppen identifizieren, die insbesondere von externer Hilfe zur Qualitätssteigerung profitieren würden. Ferner können Datennutzer herausfinden mit welchem Publikationsjahr des VN Jahrbuchs ihnen Daten für welche Referenzjahre für welche Regionen und Ländergruppen zur Verfügung stehen werden.

Unsere Ergebnisse zeigen, dass beinahe 90% aller Länder Daten für das vergangene Kalenderjahr oder für das Vor-vorjahr liefern (d.h. für das Refernzjahr t-1 oder t-2). Zum besseren Vergleich unserer Ergebnisse mit denen der vorherigen VN Studie berechnen wir neue Ergebnisse unter Verwendung des VN Ansatzes und der VGR Daten, die unserer Studie zugrunde liegen. Bei Verwendung des VN Ansatz ist die Anzahl der Länder die Daten für eine bestimmtes Referenzjahr höher als bei unserem Ansatz, außer für das Referenzjahr t-1. Das ist damit begründet, dass der VN Ansatz auch die Länder zählt, die keine neuen Daten gemeldet haben, d.h. deren Daten „bekannt“ sind. Aufgrund der im VN Ansatz ungünstig gewählten Grundmenge ergibt sich ein wesentlich schlechter Prozentwert an Ländern deren VGR Daten maximal ein Jahr Verzögerung aufwiesen. Unser Ansatz zeigt, dass die „Aktualität“ der VGR Daten zum Publikationszeitpunkt auf internationaler Ebene durch die VN besser ist als zuvor mit dem VN Ansatz ermittelt.

Verfügbarkeit

Der Ansatz zur Messung der Verfügbarkeit der VRG Daten, die der internationalen Nutzergemeinschaft bereit stehen, folgt eng dem der VN (vgl. United Nations 1999, 2000, 2001a, 2004, 2005, und UNSD 2008). Zur Verbesserung der Messung der Verfügbarkeit von VGR Daten wurden der derzeitige Ansatz analysiert. Eine wichtige Verbesserung ist die Revision des Untersuchungszeitraums, für den wenigstens ein Referenzjahr Daten existent sein müssen, auf eine fixe Jahreszahl. Dies ermöglicht die Durchführung der Messung allein aus den in der VN Datenbank vorhandenen Statistiken, wogegen der derzeitige UN Ansatz Information über die Datenmeldungen der vergangenen fünf Jahrbücher erfordert.

Als zusätzliche Regionen, Ländergruppen und Organisationen berücksichtigen wir die aus unserer Studie der „Aktualität“. Unsere empirische Studie misst die Verfügbarkeit der einzelnen NAQ Tabellen die potentiell für die Erfüllung der 1993 und 2008 MRDS Kriterien in Frage kommen, sowie der Verfügbarkeit von denen, die die 1993 und 2008 MRDS Kriterien erfüllen. Die Erfüllung der Kriterien des derzeitig noch gültigen 1993 MRDS sowie des zukünftig (ab 2014, siehe United Nations 2011) geltenden 2008 MRDS werde in zwei Teilstudien durchgeführt. Die Erfüllung der Meilenstein Stufe 1 bzw 2 gegeben den „milden“ Kriterien wird zusammen mit dem 1993 MRDS gemessen, sowie gegeben den „strengen“ Kriterien mit dem 2008 MRDS.

Die Verfügbarkeitsanalyse erlaubt die Beurteilung der generellen Kapazität zur Produktion von VGR Daten, das heißt für verschiedene VGR Konzepte, unabhängig vom VGR Standard (z. B. SNA 1993 oder 1968) nach dem die Daten erstellt wurden. Die Ergebnisse erlauben die Identifikation von Ländergruppen die besonders große Datenlücken aufweisen, ebenso wie von Ländergruppen, die besonders große, oder sogar komplette Datenmengen zur Verfügung stellen. So können Schwerpunkte für Entwicklungsarbeit in der VGR gesetzt werden. Datennutzer können damit in Erfahrung bringen, für welche NAQ Tabellen ihnen Daten für welche Regionen und Ländergruppen durch die VN zur Verfügung stehen.

Unser Ergebnis zeigt, dass die größten Datenlücken in Entwicklungsländern bestehen. Insbesondere Ozeanien und Teile Asiens haben eine geringe VGR Datenproduktion. Dies bestätigt Resultate früherer UN Analysen zum 1993 MRDS. Unsere Analyse der Ländergruppen zeigt, dass die kleinsten Datenlücken in den entwickelten Regionen Europas mit der EU27 Gruppe sind. Andere fortgeschrittene Wirtschaften, die nicht G20 Staaten sind, haben weniger Lücken in der Datenerfassung als Staaten der G20 Gruppe. Die wichtigste Erkenntnis aus unserer detaillierten Analyse ist, dass die tatsächliche Verfügbarkeit von VGR Daten weltweit sehr viel geringer ist, als bisherige 1993 MRDS Verfügbarkeitsstudien vermuten lassen. Insbesondere die Kontentabellen der volkswirtschaftlichen Sektoren bis zum Finanzierungssaldo, die für die Erfüllung der 2008 MRDS Kriterien notwendig sind, sind in vielen Ländern nicht verfügbar. So wird das 2008 MRDS nur von 20% der VN Mitgliedsstaaten erfüllt und von keinem der nicht VN Mitgliedsstaaten. Dem Gegenüber wird das 1993 MRDS von 40 % der VN Mitgliedsstaaten erfüllt, sowie von zwei der 32 nicht VN Mitgliedsstaaten (6%). Auch die Meilenstein Stufe 2 nach „strenger“ Auslegung der Kriterien wird von weniger Ländern erreicht als nach „milder“ Auslegung. Zum Beispiel selbst unter den G20 Staaten erfüllen nur ca. 50% der Länder die Meilensteinstufe 2 nach „strenger“ Auslegung. Für das 2008 MRDS erfüllen nur 47% der G20 Staaten die Kriterien, immerhin erfüllen 53% alle außer eines der 2008 MRDS Kriterien.

Zuverlässigkeit

Das Kapitel zur Zuverlässigkeit analysiert den Ansatz und die Mängel vergleichbarer internationaler Studien (vgl. Di Fonzo 2005b; Tosetto and Lequiller 2006, McKenzie 2006a und 2006b, McKenzie und Adam 2007, Tosetto und Brackfield 2009, und Neumayr 2010). Unsere Studie, sowie die vergleichbaren Studien zur Zuverlässigkeit erstveröffentlichter BIP Daten für mehrere Länder, basiert auf den Daten der OECD „Main Economic Indicators (MEI) Original Release Data and Revisions Database" und den Excel Tabellen für Revisionsstudien von Di Fonzo (2005a).

Die Zeitperiode für die Revisionsanalyse in unserer Studie ist aktueller und oft länger als in den vergleichbaren Studien. Die wichtigste Neuerung um die Resultate der Revisionsanalyse zu verbessern ist der Ausschluss von Revisionen über verschiedene SNA Versionen hinweg. Diese beeinflussen die Resultate vergleichbarer Studien. Die gewählten Daten beinhalten lediglich solche Revisionen die üblicherweise alle fünf Jahre durchgeführt werden. Jedoch gibt es keine Revision von einer zu einer anderen SNA Version, z.B. von 1968 SNA zu 1993 SNA/ ESA95[102]. Die Anzahl von untersuchten Revisionszeitpunkten wurde gegenüber vorherigen Studien erhöht. Wir messen einen oftmals vernachlässigten, unterjährigen Zeitpunkt fünf Monate nach erster Veröffentlichung, die Zeitpunkte eins, zwei und drei Jahre nach erster Veröffentlichung, sowie die „letzte" Veröffentlichung, welche mindestens vier Jahren nach der ersten Veröffentlichung liegt. Unser „letzer" Zeitpunkt ist ein Jahr später als der von vergleichbaren Studien. Der „letzte" Veröffentlichungszeitpunkt in unserer Studie weist normalerweise die „finalen" Daten aus. Eine weitere wichtige Änderung ist die Verwendung von Wachstumsraten im Vorjahresvergleich (YoY) anstelle im Vormonatsvergleich (QoQ). Die vergleichbaren Studien, bis auf McKenzie und Adam (2007), verwenden Wachstumsraten im Vormonatsvergleich. Die OECD Daten sind oftmals Saisonbereinigt. Diese sind jedoch ungeeignet für Revisionsanalysen von Wachstumsraten im Vormonatsvergleich (siehe Cheem 2003).

Mit unserer Studie gewinnen Datennutzer Informationen zu der Zuverlässigkeit dieser Daten. Zudem vergleichen wir die Ergebnisse der Länder. Unser Hauptbeitrag ist die Qualitätsanalyse von 34 Ländern anhand empirischer Revisionsdaten. Vergleichbare ländervergleichende Analysen betrachteten maximal 20 Länder sowie kürzere

[102] Für die Türkei sind sämtliche Daten in 1968 SNA.

Zeiträume. Die Resultate unserer Analyse zeigen für die einzelnen Länder, dass tendenziell die Größe der Revision von erstveröffentlichten Daten über die ersten drei Jahre nach erster Veröffentlichung höher wird. Zwischen den Ländern zeigen sich erhebliche Unterschiede bezüglich des Revisionsbedarfs. Diese Ergebnisse sind konsistent mit anderen Studien. Jedoch zeigt ein Vergleich mit vorherigen Studien, dass unsere Ergebnisse einen geringeren Revisionsbedarf der einzelnen Länder ausweisen und somit eine höhere Zuverlässigkeit der erstveröffentlichen BIP Daten. Dies ist im Wesentlichen auf den Ausschluss der BIP Daten vor Einführung des SNA 1993 /ESA95 zurückzuführen.

Ein Vergleich zwischen MAR und RMAR zeigt, dass sich für einige Länder die Rangfolge ändert. So ist z.B. der MAR von Indien groß (das heißt: Indien ist im hinteren Teil der Rangfolge) wogegen der RMAR klein ist (das heißt: Indien ist im vorderen Teil der Rangfolge). Hingegen bei Portugal sind die Größenordnungen von MAR und RMAR umgekehrt zu denen von Indien. Ein ausgeglichenes Verhältnis zwischen der Rangfolge der Länder gemessen am MAR und RMAR ist eher die Regel. Kanada zählt in beiden Fällen zu den führenden Ländern. Die Größe der MAR beträgt zum Revisionszeitpunkt drei Jahre nach Erstveröffentlichung 0,29 Prozentpunkte für Kanada, die Größe von RMAR zu diesem Zeitpunkt beträgt 10%. Island ist in beiden Fällen an letzter Stelle, mit einem MAR von 2,58 Prozentpunkte und RMAR von 46% zum Revisionszeitpunkt drei Jahre nach Erstveröffentlichung.