

IMF Working Paper

Information Content of DQAF Indicators—Empirical Entropy Analysis

Mico Mrkaic

IMF Working Paper

Statistics Department

Information Content of DQAF Indicators—Empirical Entropy Analysis

Prepared by Mico Mrkaic

Authorized for distribution by Mohammed El Qorchi

September 2010

Abstract

The study presents an analysis of the information content of IMF's Data Quality Assessment Framework (DQAF) indicators. There are significant differences in the quantity of information between DQAF dimensions and sub-dimensions. The most informative DQAF dimension is accessibility, followed by the prerequisites of quality and accuracy and reliability. The least informative DQAF dimensions are serviceability and assurances of integrity. The implication of these findings is that the current DQAF indicators do not maximize the amount of information that could be obtained during data ROSC missions. An additional set of assessments that would refine the existing DQAF indicators would be beneficial in maximizing the information gathered during data ROSC mission. The entropy of DQAF indicators could also be used in the construction of a cardinal index of data quality.

JEL Classification Numbers: C46, C49

Keywords: Data Quality Assessment Framework, Entropy, Information theory

Author's E-Mail Address: mmrkaic@imf.org

This Working Paper should not be reported as representing the views of the IMF.

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Contents	Page
I. Introduction	3
A. The quality of macroeconomic data.....	3
B. The Data Quality Assessment Framework.....	4
II. Measuring the Quantity of Information	5
III. The Data Set.....	7
IV. Results.....	8
A. Computing the Sample Variance of Entropies.....	8
Using the Delta Method.....	8
Small Population Variance Adjustments	9
B. Entropy of Individual DQAF Indicators	10
C. Kullback-Leibler Divergences between Datasets	12
V. The Discussion.....	13
VI. References.....	14
VII. Appendix I: Tables and Figures	15
VIII. Appendix II: Matching the Old and New DQAF Codes.....	24
IX. Appendix III: DQAF Dimensions	25

Tables

Table 1: Frequencies of DQAF indicators by data sets	8
Table 2: Average Information per Indicator in bits (average across data sets)	11
Table 3: KL divergences between statistics, averages across DQAF indicators.	12
Table 4. DQAF Entropy Analysis for National Accounts Statistics.....	21
Table 5. DQAF Entropy Analysis for Monetary and Financial Statistics	21
Table 6. DQAF Entropy Analysis for External Sector Statistics.....	22
Table 7. DQAF Entropy Analysis for Government Finance Statistics.....	22
Table 8. DQAF Entropy Analysis for Producer Price Statistics.....	23
Table 9. DQAF Entropy Analysis for Consumer Price Statistics.....	23
Table 10. Correspondence of Old and New DQAF Codes.....	24

Figures

Figure 1: Entropy for DQAF Indicators—National Accounts Statistics	15
Figure 2: Entropy for DQAF Indicators—Monetary and Financial Statistics.....	16
Figure 3: Entropy for DQAF Indicators—External Sector Statistics	17
Figure 4: Entropy for DQAF Indicators—Government Finance Statistics	18
Figure 5: Entropy for DQAF Indicators—Producer Price Index.....	19
Figure 6: Entropy for DQAF Indicators—Consumer Price Index.....	20

I. INTRODUCTION

A. The quality of macroeconomic data

Successful macroeconomic policies and high quality macroeconomic research require frequent, reliable and timely macroeconomic data. Unfortunately, macroeconomic data are often less reliable than desired, disseminated with suboptimal frequencies and/or timeliness. The final users of macroeconomic data often lack either resources and/or knowledge to assess the quality of the data that they use. If macroeconomic data were experimental, they could in principle be verified. However, macroeconomic data are in most cases not verifiable. In addition, the reliability and usefulness of macroeconomic data could potentially suffer from the use of nonstandard statistical methodologies, classifications, and even fraudulent reporting practices, where agencies or authorities deliberately disseminate incorrect information to conceal their mistaken policies or instances of mismanagement.

The collection and dissemination of macroeconomic data is a resource intensive activity, which, especially in poor and developing countries, is not allocated sufficient human and financial resources to produce adequate outputs. Poor data quality often results in errors and omissions that could be cumulative. As showcased by the recent experiences with crises economies, the conduct of monetary and fiscal policies can be adversely affected by unreliable or fraudulent macroeconomic data.

Because assessing the quality of macroeconomic data is of great importance for both the conduct of economic policies and the integrity of empirical research, methodologists have developed specialized frameworks to assess data quality. Such frameworks typically assess data quality by comparing data compilation and dissemination practices in a country against the ideal benchmark—the best internationally accepted practices. The Data Quality Assessment Framework (DQAF) is a framework for data quality assessment, developed and utilized by the IMF and other international agencies. The DQAF has been developed to provide a framework for a uniform and standardized assessment of data quality and improvements of data compiling and dissemination practices. It assesses the observance of best data compilation and dissemination practices according to five dimensions of data quality: prerequisites of quality, assurance of integrity, methodological soundness, accuracy and reliability, serviceability and accessibility.¹

While the dimensions and sub-dimensions of the DQAF appear to provide an intuitively appealing assessment framework, they have been selected on a mostly ad-hoc basis. So far, no analytical work has been done to assess the extent to which the selection of the dimensions and sub-dimensions of the DQAF is appropriate and informative. Specifically, there exist no estimates of the information content different DQAF indicators. The purpose of

¹ A more detailed description of the DQAF is given in subsection I. B.

this paper is to provide such analysis and construct an analytical framework for assessing the information content of the DQAF and other indicators of data quality.²

I use information theory to compute the entropy, a standard measure of information, for each of the dimensions and sub-dimensions of the DQAF. In addition, I compute the Kullback-Leibler divergence between different macroeconomic data sets. The analysis yields assessments of the quantity of information for each DQAF indicator, measured in bits of information, and the measure of the degree of statistical independence of DQAF indicators between macroeconomic data sets.

B. The Data Quality Assessment Framework

The DQAF identifies quality-related features of governance of statistical systems, statistical processes, and statistical products. It is rooted in the UN Fundamental Principles of Official Statistics. The DQAF provides a structure for assessing existing practices against best practices, including internationally accepted methodologies. It has proved to be valuable for at least three groups of users: (i) IMF staff using data in policy evaluation, preparing the data module of Reports on the Observance of Standards and Codes (ROSCs), and designing technical assistance; (ii) country authorities preparing self-assessments; and (iii) data users evaluating data for policy analysis, forecasts, and economic performance.

The DQAF's coverage of governance, processes, and products is organized around a set of prerequisites and five dimensions of data quality—assurances of integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility. For each dimension, the DQAF identifies 3-5 elements of good practice, and for each element, several relevant indicators.³

The DQAF is the organizing model of the data module Report on the Observance of Standards and Codes. To prepare a ROSC, at the invitation of the authorities, a team of experts spends about two weeks in dialogue along the lines of the DQAF with the country officials. To date, about 100 data ROSCs have been published.⁴

Data ROSCs rank the observance of good practice for each element of each dimension with one of the following four marks: not observed (NO), largely not observed (LNO), largely

² The first paper that underlined the key importance of data quality for macroeconomic analysis is Morgenstern (1950). Unfortunately, the awareness of this critical importance has not increased appreciably since the time of that paper.

³ The elements of good practice are listed in Appendix. IX.

⁴ The first three paragraphs in this section have been adapted from the *Data Quality Assessment Framework Factsheet*, IMF, Statistics Department © 2006

observed (LO), and observed (O). These assessments are ordinal and not cardinal. For this reason, the assessments cannot be analyzed using the standard statistical machinery, which primarily relies on the moments of random variables. For example, assigning numbers from one to four to the rankings and computing sample averages and standard deviations is intuitively appealing, but conceptually misguided. Such numbering preserves the order of rankings—however, every monotonic transformation of this numbering would also preserve the rankings. The moments would of course not be preserved.⁵

A logically consistent approach to the problem of statistical analysis of ordinal data is to use the tools from information theory to compute the quantity of information contained in different DQAF indicators. In addition, information theory can be used to compute the degree of informational independence between DQAF indicators. Entropy, the standard measure of information in information theory, can be computed for any random variable. Entropy does not depend on the values that the random variable can take; it depends, only on the underlying probabilities of different outcomes of the random variable.

The remainder of the paper is structured as follows: the second section explains the basics of the information theory and show how it is applied in this paper; the third section provides a detailed description of the data set; the fourth section presents the results of the analysis and the fifth section concludes.

II. MEASURING THE QUANTITY OF INFORMATION⁶

Shannon (1948) derived a measure of information content called the *self-information* of a message m .

$$I(m) = \log\left(\frac{1}{p(m)}\right) = -\log(p(m)) \quad (1.1)$$

Here $p(m) = \Pr(M=m)$ is the probability that message m is chosen from all possible choices in the message space M . The base of the logarithm is usually chosen to be two. In this case, the measure of information is expressed in units of bits of information.

A message that is certain to occur has an information measure of zero. A compound message of two or more mutually independent messages has a quantity of information that is the sum of the measures of information of each message individually.

⁵ The same conclusion holds for computing correlations of ordinal random variables.

⁶ This section has been adapted from the Wikipedia article on Information theory.

The *entropy* of a discrete message space M is a measure of the amount of uncertainty one has about the message that will be transmitted. It is defined as the expected self-information of a message m from that message space

$$H(M) = E[I(M)] = - \sum_{m \in M} p(m) \log(p(m)). \quad (1.2)$$

Entropy is maximized when all the messages in the message space are equally probable, that is when $p(m) = 1/M$. In this case the value of entropy is $H(M) = \log|M|$.

Function H can also be expressed in terms of the probabilities of the underlying distribution

$$H(p_1, p_2, \dots, p_k) = - \sum_{i=1}^k p_i \log(p_i), \quad p_i \geq 0 \forall i \in \{1, \dots, k\}, \quad \sum_{i=1}^k p_i = 1. \quad (1.3)$$

The *joint entropy* of two discrete random variables X and Y is defined as the entropy of the joint distribution of X and Y

$$H(X, Y) = -E[\log(p(x, y))] = - \sum_{x, y} p(x, y) \log(p(x, y)). \quad (1.4)$$

If X and Y are independent, then the joint entropy is simply the sum of their individual entropies.

Given a particular value of a random variable Y , the *conditional entropy* of X given $Y=y$ is defined as

$$H(X | y) = -E_{x|y}[\log(p(x, y))] = - \sum_{x \in X} p(x | y) \log(p(x | y)). \quad (1.5)$$

In the equation above $p(x|y)$ is the conditional probability of x given y .

The conditional entropy of X given Y is given by

$$H(X | Y) = -E_Y[H(X | y)] = - \sum_{y \in Y} p(y) \sum_{x \in X} p(x | y) \log(p(x | y)) \quad (1.6)$$

A basic property of the conditional entropy is that $H(X|Y) = H(X, Y) - H(Y)$.

The *Kullback–Leibler divergence* (or relative entropy) is a way of comparing two distributions, a “true” probability distribution p , and an arbitrary probability distribution q . If we compress data in a manner that assumes q is the distribution underlying some data, when, in reality, p is the correct distribution, Kullback–Leibler divergence is the number of average additional bits per datum necessary for encoding the required information; mathematically, it holds

$$D_{KL}(p(X) \parallel q(X)) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}. \quad (1.7)$$

The Kullback–Leibler divergence is sometimes loosely called the “Kullback–Leibler distance” between q and p . However, it is not a true metric because it is not symmetric in q and p .⁷

Mutual information is a measure of how much information can be obtained about one random variable by observing another. The mutual information of X relative to Y , which represents conceptually the average amount of information about X that can be gained by observing Y , is

$$I(X; Y) = \sum_{y \in Y} p(y) \sum_{x \in X} p(x|y) \log \frac{p(x|y)}{p(x)} = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}. \quad (1.8)$$

A basic property of the mutual information is that $I(X; Y) = H(X) - H(X|Y)$. That is, knowing Y , we can save an average of $I(X; Y)$ bits in encoding X compared to not knowing Y . Mutual information is symmetric, that is $I(X; Y) = I(Y; X) = H(X) + H(Y) - H(X, Y)$.

Mutual information is closely related to the log-likelihood ratio test in the context of contingency tables and the multinomial distribution and to Pearson's χ^2 test: mutual information can be considered a statistic for assessing independence between a pair of variables, and has a well-specified asymptotic distribution.

III. THE DATA SET

The data are compiled from data ROSC mission reports.⁸ The Fund has fielded been 88 such mission so far. The DQAF indicators that are analyzed in this study are usually presented in the summary table of observance (Table 1) in each of the data ROSC mission reports. Data ROSC missions have visited some countries more than once and our data set includes all repeated data ROCS mission reports.

Data ROSC mission reports usually contain DQAF ratings for the following six data categories: national account statistics, producer price statistics, consumer price statistics, government finance statistics, external sector statistics, and monetary and financial statistics. In some cases, the data ROCS missions had more limited mandates and did not asses the entire spectrum of macroeconomic data. For the purposes of our analysis the existence of the

⁷ Formally, the Kullback-Leibler divergence is a pseudo-distance.

⁸ All data ROSC mission reports can be found on at <http://dsbb.imf.org/pages/dqrs/ROSCDataModule.aspx>.

mission with narrower mandates is irrelevant, since we pool the data from all data ROCS missions across data categories.

For the purpose of analysis, the data ROSC reports were pooled across all reports. This pooling results in approximately 88 observations per data set and per DQAF sub-dimension. For example, there are 88 observations that assess the observance of international methodologies for monetary and financial statistics.⁹

Table 1 presents the frequencies of different assessment for the six data sets. In total, the data set contains 9771 assessments. Government Finance Statistics and National Accounts Statistics are the two most frequently assessed data sets with 1818 and 1812 assessment respectively. Producer Price Statistics are least frequently assessed, with 1086 assessments.

Table 1: Frequencies of DQAF indicators by data sets

	NA	NO	LNO	LO	O	Assessed
National accounts statistics	124	5	215	694	898	1812
Monetary and financial statistics	206	6	75	608	1041	1730
External sector statistics	165	9	142	663	957	1771
Government finance statistics	118	20	229	728	841	1818
Producer price statistics	850	12	74	226	774	1086
Consumer price statistics	382	1	78	426	1049	1554
Total	1845	53	813	3345	5560	9771

IV. RESULTS

A. Computing the Sample Variance of Entropies

Using the Delta Method

Since the sample of the countries does not cover the whole population of the IMF members, we need to compute the sampling errors of the entropies and Kullback-Leibler divergences. We use the delta method and the properties of the multinomial distribution with four different outcomes to derive the approximate distribution of the sample variance of the entropies and the relative entropies in the sample of size n .

The delta method gives implies that the sample variance of entropy is given by the following formula

⁹ The DQAF codes after 2003 do not exactly match the original DQAF codes. While no substantive changes in the DQAF codes were made in the 2003 revision, the numbering of the codes changed. The matched codes are shown in Appendix II: Matching the Old and New DQAF Codes.

$$\text{Var}(H(\hat{p})) = \frac{1}{n} \sum_{i=1}^4 \sum_{j=1}^4 \left(\frac{\partial H(\hat{p})}{\partial p_i} \right) \text{Var}(\hat{p})_{ij} \left(\frac{\partial H(\hat{p})}{\partial p_j} \right). \quad (1.9)$$

A calculation which uses the properties of the multinomial distribution and the basic properties of sample averages and variances yields a simple and intuitive result—the sample variance of the entropy is equal to the variance of base two logarithm of the empirical probability distribution. Mathematically, it holds that

$$\text{Var}(H(\hat{p})) = \frac{1}{n} \left[\sum_{i=1}^4 \hat{p}_i (\log_2(\hat{p}_i))^2 - \sum_{i=1}^4 \hat{p}_i \log_2(\hat{p}_i) \right] = \frac{1}{n} \text{Var}_{\hat{p}}(\log_2(\hat{p})). \quad (1.10)$$

A straightforward extension of the above formula shows that the sample variance of the Kullback-Leibler divergence is

$$\text{Var}(D_{\text{KL}}(\hat{p}, \hat{q})) = \frac{1}{n} \left[\text{Var}_{\hat{p}}(\log_2(\hat{p})) + \text{Var}_{\hat{q}}(\log_2(\hat{q})) \right]. \quad (1.11)$$

Small Population Variance Adjustments

The data set contains data on approximately 88 countries out of 186 IMF members. The sample of such size (47 percent of all countries) is not small relative to the population. Because of the large relative size of the sample, we have to adjust downward the variance of the computed sample entropies.¹⁰ The correction can be derived by using the formulas for the correlations of draws without replacement from a population with finite size. The correction is

$$\text{Var}(X_{\text{finite}}) = \text{Var}(X_{\text{infinite}}) \left(1 - \frac{n}{N} \right). \quad (1.12)$$

This formula is intuitively appealing—if we sample the entire population, the variance of any random variable, defined on the population, must be equal to zero.¹¹

¹⁰ Intuitively, if we had data on all 186 IMF members, we would obtain a deterministic result with variance zero. The sample in our case is not as large as the population, but is still large enough that a finite population correction is needed.

¹¹ Care must be taken not to confuse the situation described in the text with the situation where we have a finite sample of countries, for example a subset of the population of the IMF members, and we estimate an econometric model on the sample. In that case, an application of the finite population would be erroneous, since in the population that is statistically relevant is not the population of all countries, but the population of all model residuals, which is infinite.

B. Entropy of Individual DQAF Indicators

The first set of results presents the estimates of entropy of individual DQAF indicators. The results are graphically presented in Figures 1 through 6. The length of each horizontal bar presents the average information content of the corresponding indicator, measured in bits of information. Since there are four possible outcomes for each DQAF indicator (O, LO, LNO, NO), it follows that the maximum information per indicator is two bits. Because the data set does not contain the entire population of the IMF member countries, the results are subject to statistical uncertainty. The statistical uncertainty is presented by the error bars, superimposed on each horizontal bar. Each error bar presents a 95 percent confidence interval. For example, in Figure 1, the indicator for the observance ethical standards contains about 0.4 bits of information and its 95 percent confidence interval is from 0.25 and 0.55. The results discussed above are also presented in the tables, which are listed in Section VII.

Figure 1 presents the information content of DQAF indicators for National Accounts Statistics (NAS). The most informative indicator is revision studies whose information content is approximately 1.65 bits. The least informative indicator is ethical standards with information content of only about 0.4 bits. Significant variation exists between the indicators. In addition, the information content of some indicators is rather low. This finding implies that the data ROSC questionnaires could be revised to increase the information obtained on data ROSC mission.¹²

Figure 2 presents the information contents for Monetary and Financial Statistics (MFS). The information content of the indicator of the observance of ethical standards is equal to zero—based on the information in our sample, nothing has been learned by asking the question about the observance of ethical standards in the case of Monetary and Financial Statistics. In addition, the information content of professionalism indicator is very low at around 1.18. This indicator is barely statistically significant at 95 percent level. Similarly, the information regarding the professionalism is barely statistically significant and its mean value is very small. The most informative indicator is the accessibility of metadata with information of around 1.6 bits, followed by revision policy and practice with information of around 1.4 bits.

Figure 3 presents the information content of the DQAF indicators for the External Sector Statistics (ESS).¹³ The results are broadly similar to the findings for National Accounts Statistics and Monetary and Financial Statistics. Metadata accessibility and revision studies are the most informative indicators. Their information content is approximately 1.7 bits, implying that they are approximately 85 percent informative relative to the maximum information content of an indicator with four possible outcomes. Two indicators in DQAF dimension one, ethical standards and professionalism, are almost entirely uninformative with the average information content of 0.10 and 0.15 respectively. In particular, the information

¹² The actual revisions of the DQAF should be subject to future research; the present study is not normative, but rather positive in nature.

¹³ Also abbreviated as BOP in the text below.

content of the ethical standards indicator is not significantly different from zero, which matches the result for the Monetary and Financial Statistics.

Figure 4 presents the information content of the DQAF indicators for the Government Finance Statistics (GFS). All DQAF indicators are statistically significant; however, their information contents vary broadly. The results are broadly similar to the findings described for the preceding three cases. Revision studies and metadata accessibility are the most informative indicators. Their information content is approximately 1.8 and 1.7 bits respectively. Two indicators in dimension one, ethical standards and professionalism are largely uninformative with the average information content of 0.15 and 0.6 respectively. In particular, the information content of the ethical standards indicator is barely significantly different from zero; its 95 percent confidence interval almost touches zero level.

Figure 5 presents the information content of the DQAF indicators for the Producer Price Index (PPI). Metadata accessibility and data accessibility are the most informative indicators. Their information content is approximately 1.5 bits, implying that they are approximately 85 percent informative relative to the maximum information content of an indicator with four possible outcomes. Two indicators in dimension one, ethical standards and professionalism, are uninformative with the average information content of 0.0 and 0.15 respectively. In addition, both indicators are not significantly different from zero.

Figure 6 presents the information content of the DQAF indicators for the Consumer Price Index (CPI). Resources and metadata accessibility are the most informative indicators. Their information content is approximately 1.5 bits, implying that they are approximately 85 percent informative relative to the maximum information content of an indicator with four possible outcomes. The least informative indicator is periodicity and timeliness, which is not statistically significant. It is likely that the insignificance of this indicator is a consequence of the common uniform practice of disseminating CPI data with monthly periodicity and timeliness of less than one month—the majority of countries observe these timeliness and periodicity requirements in the dissemination of CPI data.

Table 2: Average Information per Indicator in bits (average across data sets)

DQAF Dimension	Average Information per Indicator (in bits)	Rank
Prerequisites of quality	1.19	2
Assurances of integrity	0.60	6
Methodological soundness	1.13	4
Accuracy and reliability	1.19	2
Serviceability	1.08	5
Accessibility	1.38	1

Table 2 presents the averages of the information content of the six main DQAF dimensions across all countries in the sample and across all statistics. The most informative DQAF

dimension is accessibility with the average information content of 1.38 bits. The information content of prerequisites of quality and accuracy and reliability share the second place with 1.19 bits of information per indicator. The two least informative DQAF dimensions are serviceability and assurances of integrity with information content per indicator of 1.08 and 0.60 bits respectively.

C. Kullback-Leibler Divergences between Datasets

While the assessments of the informational contents of individual DQAF indicators the most important results of the present analysis, the informational “differences” between datasets are also of interest. For this reason, we calculate a table of average Kullback-Leibler divergences between different data sets. Table 3 presents the results.¹⁴ The left half of the table presents the divergences and the right half the standard errors of the divergences. The diagonal elements on the left are obviously zero. Certain common features emerge from the table. First, the divergences between the price statistics (i.e. CPI and PPI) are small, but statistically significant. The divergence between BPS and MFS is small and not statistically significant, presumably because in most countries the MFS and BPS are collected by the Central Banks. Similarly, the divergence between NAS and BPS is not statistically significant. The largest divergences exist between NAS and MFS, CPI and MFS and PPI and GFS; however, they are not statistically significant at the 5 percent level. Overall, the statistically significant divergences are those between CPI and BPS, CPI and PPI, MFS and CPI and PPI and CPI respectively. It must be stressed, that the interpretation of these results is not straightforward. The one-dimensional analysis of the information content of individual DQAF indicators is simpler to interpret and could potentially lead to more useful and more immediate practical applications.

Table 3: KL divergences between statistics, averages across DQAF indicators.

	Mean differences						Standard deviations					
	NAS	MFS	BPS	GFS	PPI	CPI	NAS	MFS	BPS	GFS	PPI	CPI
NAS	0.0000	0.3227	0.0984	0.1769	0.3038	0.1368	0.0000	0.4461	0.1024	0.2091	0.1933	0.1029
MFS	0.2591	0.0000	0.0876	0.2134	0.2621	0.1392	0.2301	0.0000	0.0512	0.1790	0.1031	0.0534
BPS	0.1061	0.0988	0.0000	0.1531	0.2711	0.1177	0.1044	0.0928	0.0000	0.1482	0.1325	0.0685
GFS	0.2670	0.2184	0.1540	0.0000	0.4214	0.2327	0.3103	0.2399	0.1722	0.0000	0.2817	0.1764
PPI	0.2003	0.2833	0.2070	0.4262	0.0000	0.0515	0.1222	0.1482	0.0946	0.3273	0.0000	0.0113
CPI	0.2054	0.3644	0.2194	0.3720	0.0909	0.0000	0.1277	0.1880	0.1074	0.2746	0.0160	0.0000

¹⁴ The table is not symmetric, since the KL divergence is not a proper distance, but rather a pseudo-distance.

V. THE DISCUSSION

The study presents an analysis of the information content of DQAF indicators. The results are informative with respect to the amount of information contained in the indicators. Significant differences exist in the quantity of information per DQAF indicator. In addition, there are significant differences in the quantity of information per DQAF dimension. The entropy analysis provides a ranking of the DQAF dimensions and sub-dimensions with respect to their information content. The most informative DQAF dimension is accessibility, followed by the prerequisites of quality and accuracy and reliability. The least informative DQAF dimensions are serviceability and assurances of integrity. Specifically, the ethics sub-dimension is very uninformative and often not statistically different from zero.

The implication of these findings is that the current DQAF indicators do not maximize the amount of information about data quality that could be obtained during data ROSC missions. An additional set of assessments that would refine and upgrade the existing DQAF indicators would be beneficial in maximizing the information gathered during data ROSC mission. For example, the assurances of integrity contain only 30 percent of the theoretical maximum of information. It is likely that additional sub-dimensions of this dimension could quickly and significantly increase the amount of information about data quality that can be obtained. While achieving the theoretical maximum is obviously impossible, substantial improvements in the information content of this DQAF dimension could be obtained.

In general, the information content, averaged across all DQAF dimensions and all data sets, is around 1.10 bits or approximately 55 percent of the theoretical maximum of two bits. There appears to be scope for improvement of the existing DQAF framework. However, the exact nature of improvements and/or additions to the framework is not immediately transparent and will require additional research.

The entropy of DQAF indicators could also be used in the construction of an index of data quality. It is intuitively appealing to assume that the weight of a component of an index of data quality should be proportional to its information content. Components with little or no information should weigh relatively less than more informative ones. It must be emphasized that the procedure to determine of the optimal weights and the content of individual components of a data quality index will require substantial additional investigations that are beyond the scope of this paper.

VI. REFERENCES

Hartwig, Jochen: *Trying to Assess the Quality of Macroeconomic Data – the Case of Swiss Labor Productivity Growth as an Example*, KOF Working Papers, No. 173 September 2007

Shannon, Claude E.: *A Mathematical Theory of Communication*, Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, 1948.

Morgenstern, O.: *On the Accuracy of Economic Observations*, Princeton University Press, 1950.

Bagus, P.: *The Problem of Accuracy of Economic Data*, Mises Daily, August 2006, (available online at <http://mises.org/story/2280>).

Morgenstern, O.: *On The Accuracy Of Economic Observations, In Activity Analysis Of Production And Allocation*, Proceedings Of A Conference edited by Tjalling C. Koopmans in cooperation With Armen Alchian, George B. Dantzig, Nicholas Georgescu-Roegen, Paul A. Samuelson, Albert W. Tucker, published by John Wiley & Sons, Inc., New York, Chapman & Hall, Limited, London, 1951 (available online at <http://cowles.econ.yale.edu/P/cm/m13/>).

VII. APPENDIX I: TABLES AND FIGURES

Figure 1: Entropy for DQAF Indicators—National Accounts Statistics

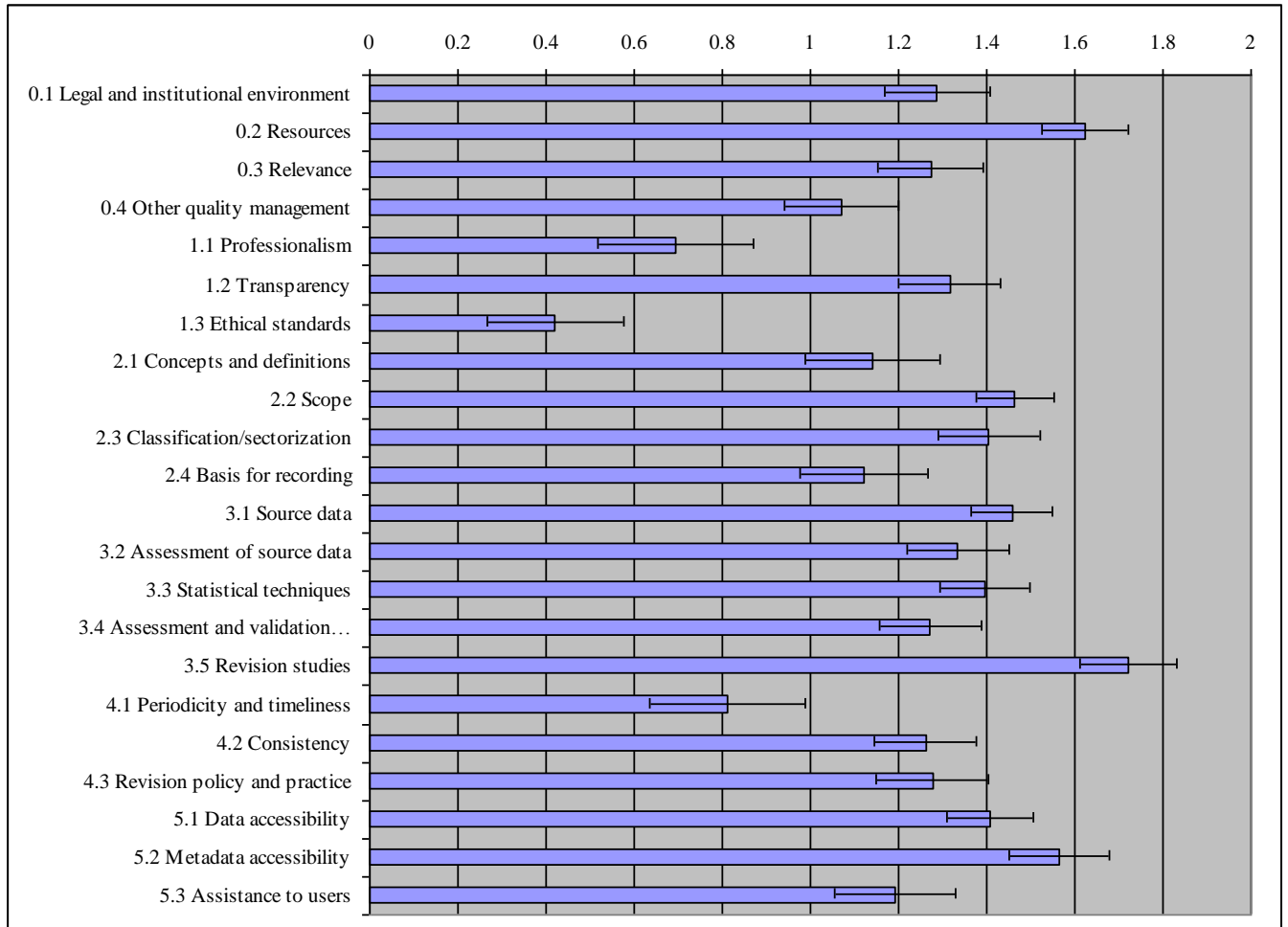


Figure 2: Entropy for DQAF Indicators—Monetary and Financial Statistics

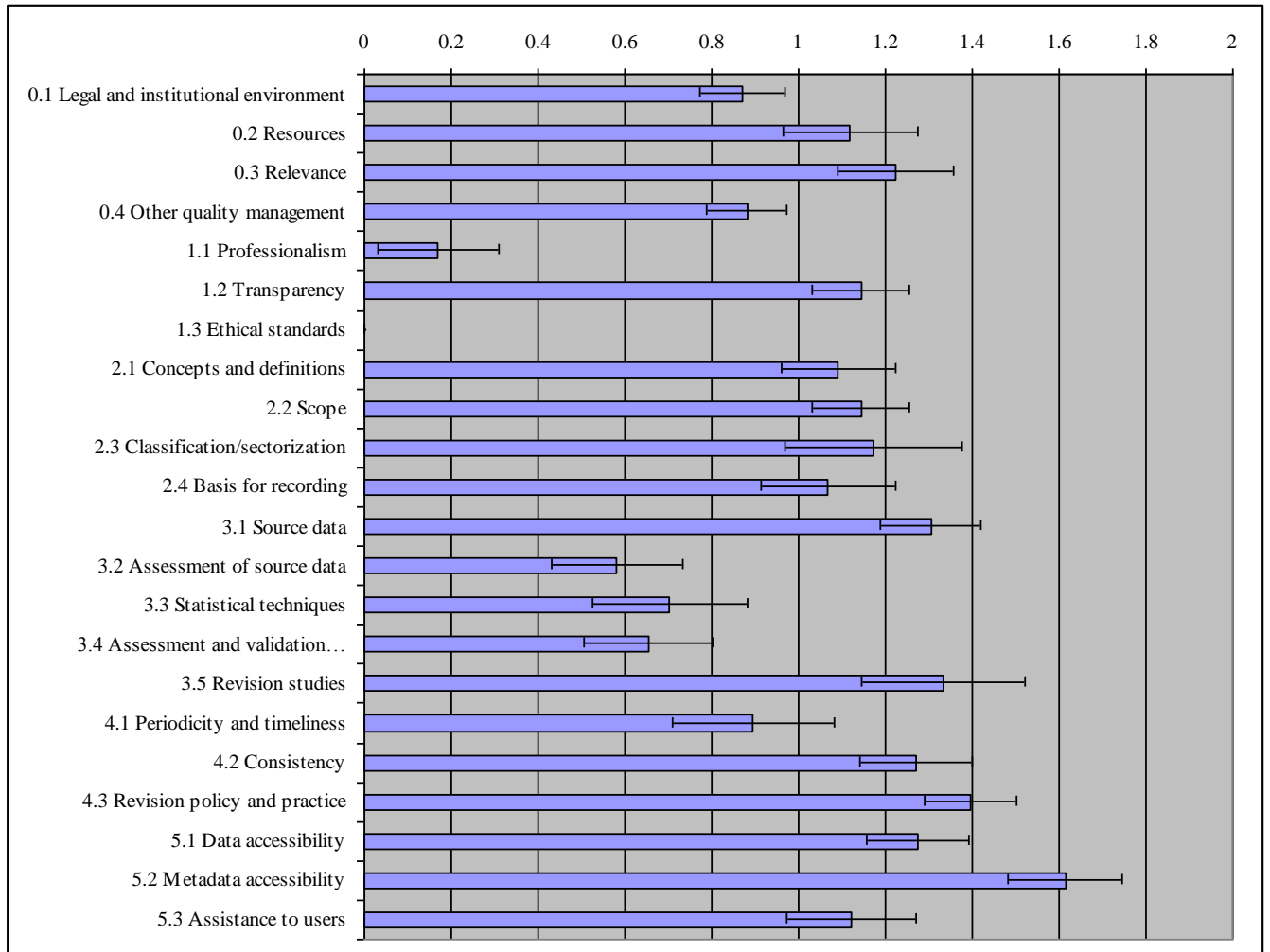


Figure 3: Entropy for DQAF Indicators—External Sector Statistics

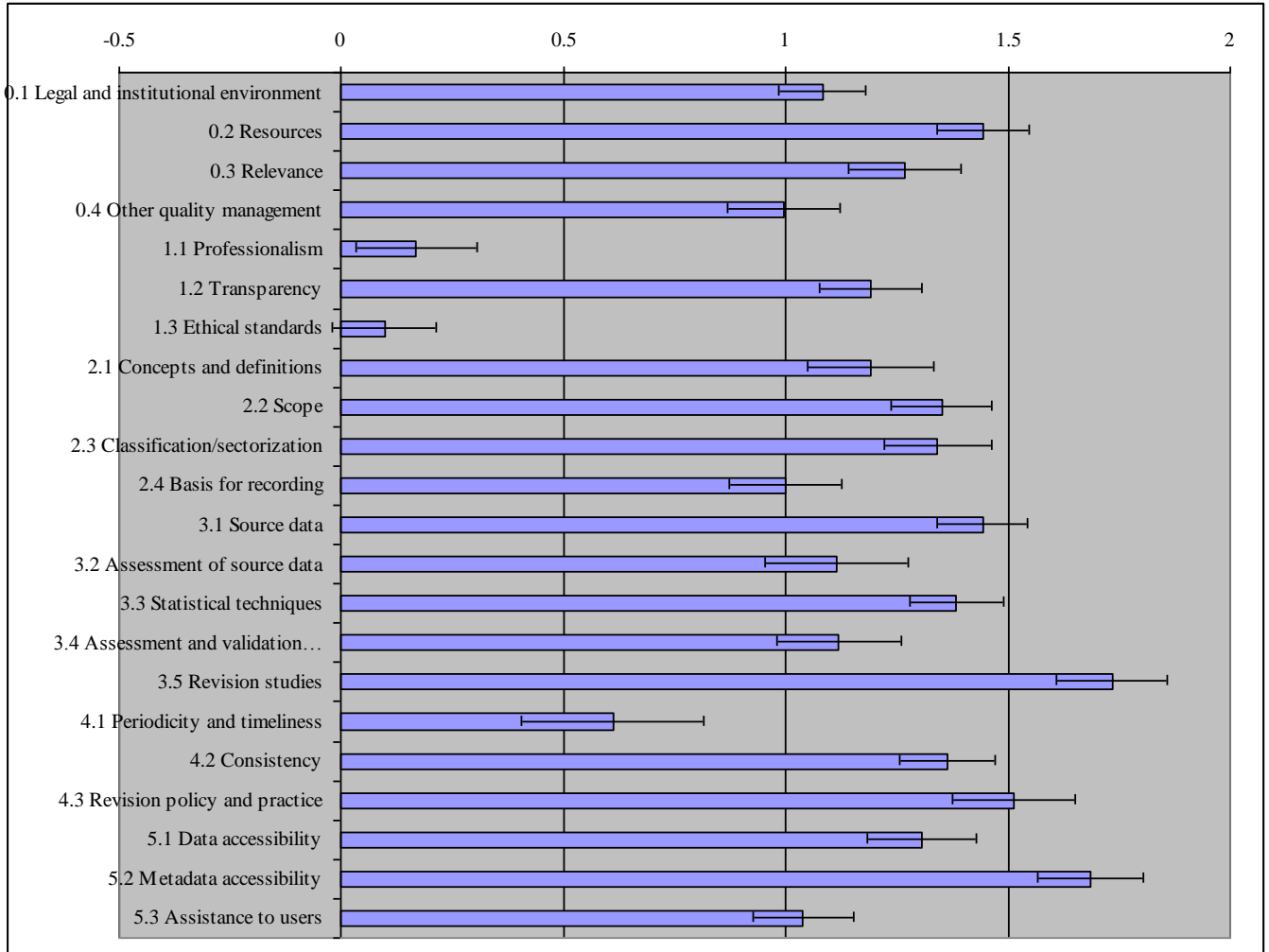


Figure 4: Entropy for DQAF Indicators—Government Finance Statistics

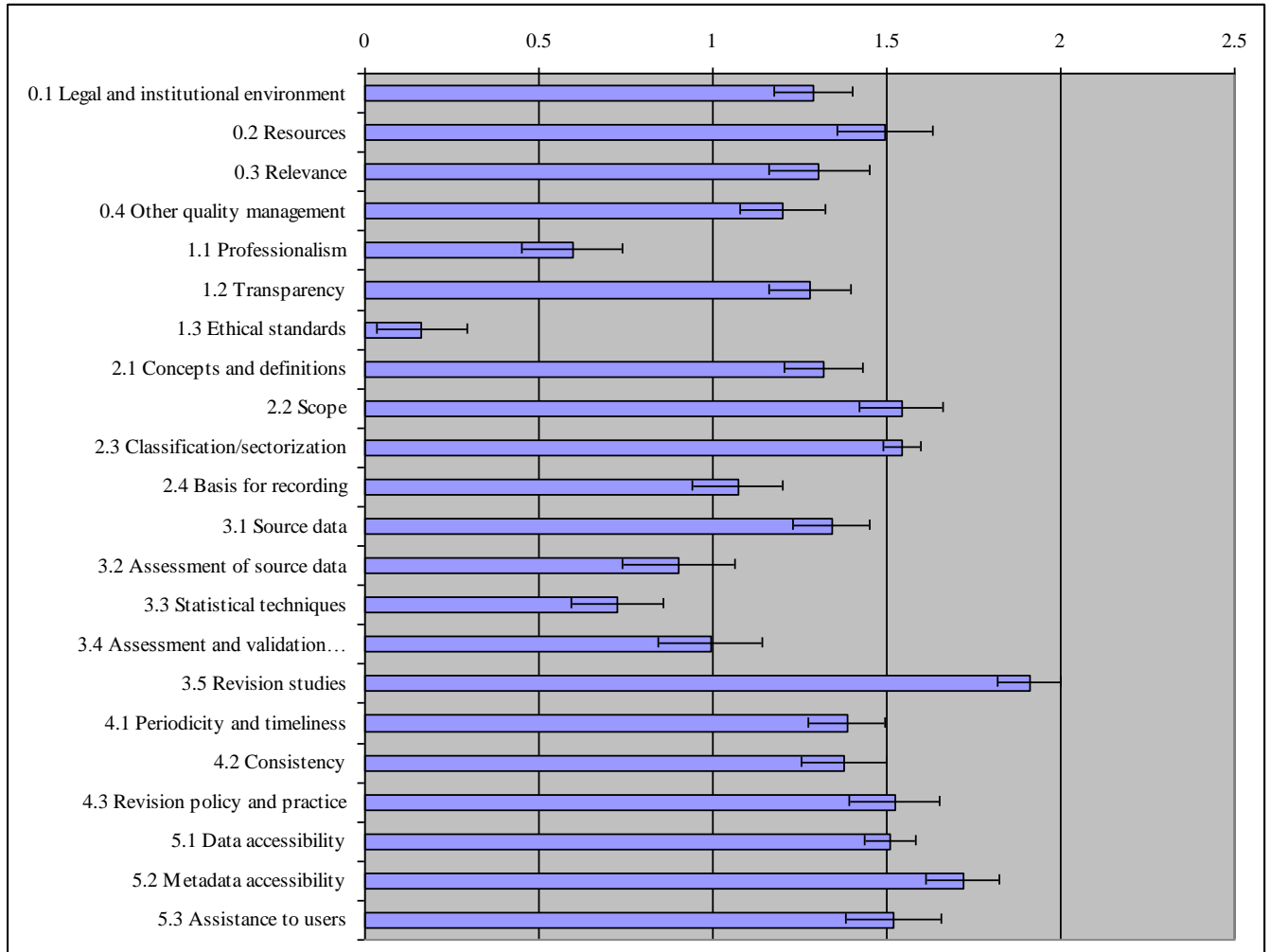


Figure 5: Entropy for DQAF Indicators—Producer Price Index

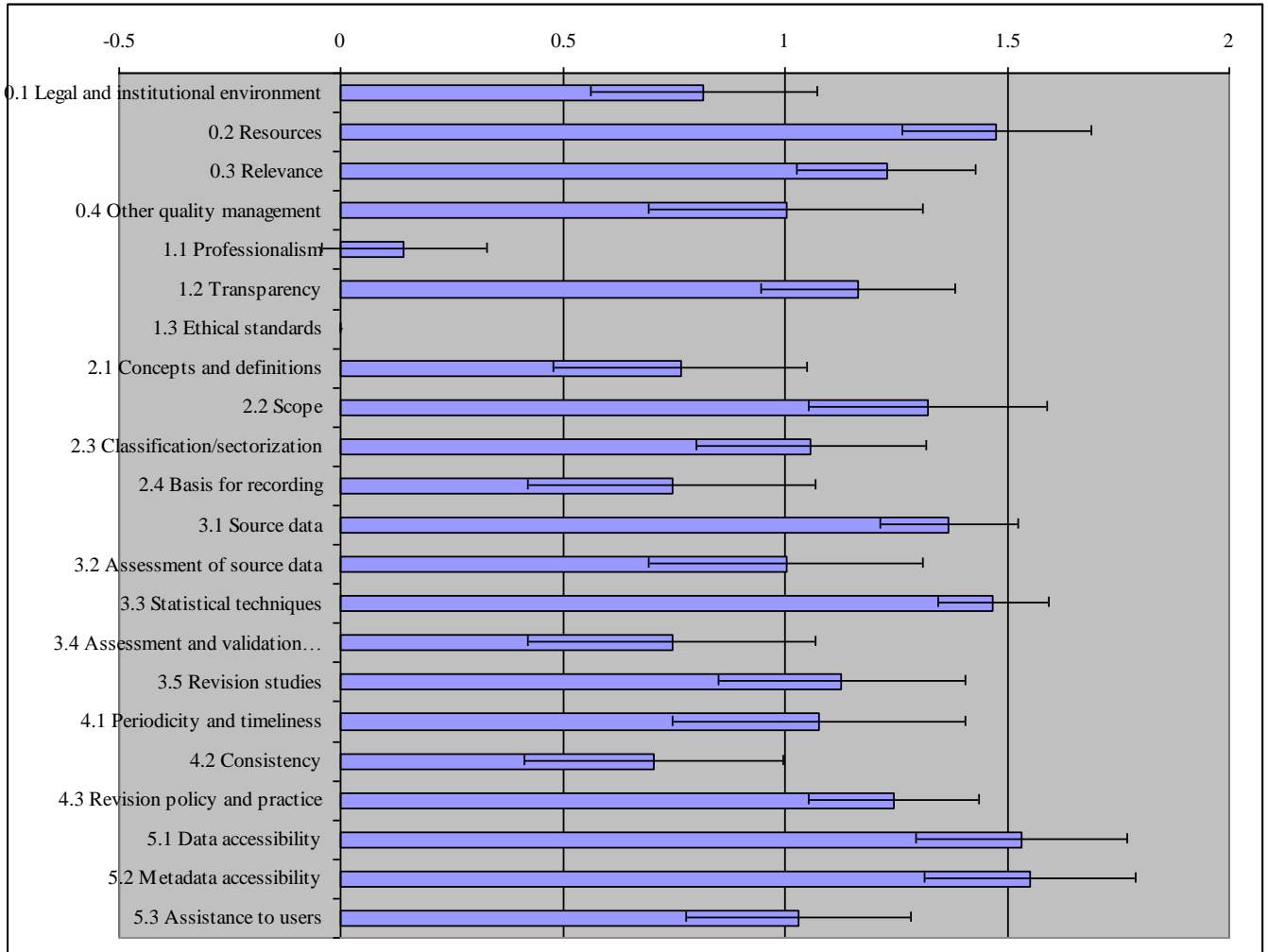


Figure 6: Entropy for DQAF Indicators—Consumer Price Index

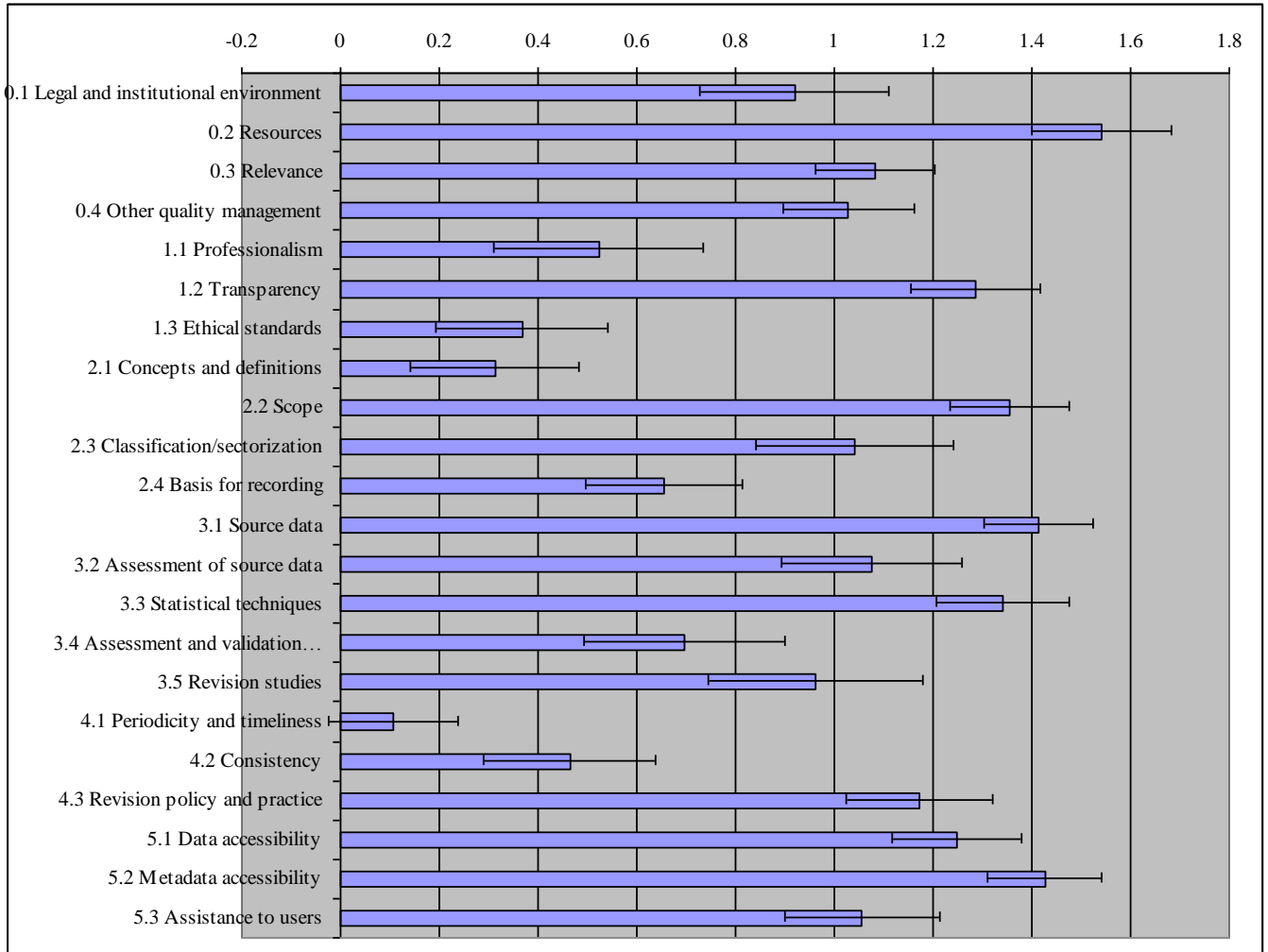


Table 4. DQAF Entropy Analysis for National Accounts Statistics

Indicator	Entropy	Stdev	p-val
0.1, Legal and institutional environment	1.2877	0.0604	0.0000
0.2, Resources	1.6249	0.0503	0.0000
0.3, Relevance	1.2726	0.0609	0.0000
0.4, Other quality management	1.0716	0.0656	0.0000
1.1, Professionalism	0.6928	0.0901	0.0000
1.2, Transparency	1.3164	0.0586	0.0000
1.3, Ethical standards	0.4208	0.0789	0.0000
2.1, Concepts and definitions	1.1418	0.0774	0.0000
2.2, Scope	1.4646	0.0453	0.0000
2.3, Classification/sectorization	1.4050	0.0587	0.0000
2.4, Basis for recording	1.1219	0.0741	0.0000
3.1, Source data	1.4574	0.0470	0.0000
3.2, Assessment of source data	1.3348	0.0588	0.0000
3.3, Statistical techniques	1.3958	0.0528	0.0000
3.4, Assessment and validation of intermediate data and statistical outputs	1.2721	0.0589	0.0000
3.5, Revision studies	1.7219	0.0564	0.0000
4.1, Periodicity and timeliness	0.8120	0.0904	0.0000
4.2, Consistency	1.2610	0.0589	0.0000
4.3, Revision policy and practice	1.2768	0.0653	0.0000
5.1, Data accessibility	1.4073	0.0503	0.0000
5.2, Metadata accessibility	1.5643	0.0585	0.0000
5.3, Assistance to users	1.1914	0.0700	0.0000

Table 5. DQAF Entropy Analysis for Monetary and Financial Statistics

Indicator	Entropy	Stdev	p-val
0.1, Legal and institutional environment	0.8702	0.0498	0.0000
0.2, Resources	1.1179	0.0790	0.0000
0.3, Relevance	1.2240	0.0676	0.0000
0.4, Other quality management	0.8813	0.0473	0.0000
1.1, Professionalism	0.1703	0.0706	0.0079
1.2, Transparency	1.1437	0.0565	0.0000
1.3, Ethical standards	0.0000	0.0000	NaN
2.1, Concepts and definitions	1.0921	0.0665	0.0000
2.2, Scope	1.1439	0.0574	0.0000
2.3, Classification/sectorization	1.1718	0.1037	0.0000
2.4, Basis for recording	1.0680	0.0791	0.0000
3.1, Source data	1.3049	0.0592	0.0000
3.2, Assessment of source data	0.5822	0.0776	0.0000
3.3, Statistical techniques	0.7038	0.0918	0.0000
3.4, Assessment and validation of intermediate data and statistical outputs	0.6550	0.0751	0.0000
3.5, Revision studies	1.3332	0.0953	0.0000
4.1, Periodicity and timeliness	0.8960	0.0951	0.0000
4.2, Consistency	1.2694	0.0657	0.0000
4.3, Revision policy and practice	1.3972	0.0542	0.0000
5.1, Data accessibility	1.2747	0.0594	0.0000
5.2, Metadata accessibility	1.6144	0.0675	0.0000
5.3, Assistance to users	1.1219	0.0766	0.0000

Table 6. DQAF Entropy Analysis for External Sector Statistics

Indicator	Entropy	Stdev	p-val
0.1, Legal and institutional environment	1.0818	0.0494	0.0000
0.2, Resources	1.4454	0.0526	0.0000
0.3, Relevance	1.2680	0.0651	0.0000
0.4, Other quality management	0.9945	0.0646	0.0000
1.1, Professionalism	0.1687	0.0696	0.0077
1.2, Transparency	1.1910	0.0584	0.0000
1.3, Ethical standards	0.0969	0.0591	0.0505
2.1, Concepts and definitions	1.1917	0.0720	0.0000
2.2, Scope	1.3515	0.0576	0.0000
2.3, Classification/sectorization	1.3420	0.0623	0.0000
2.4, Basis for recording	0.9995	0.0648	0.0000
3.1, Source data	1.4427	0.0516	0.0000
3.2, Assessment of source data	1.1144	0.0816	0.0000
3.3, Statistical techniques	1.3842	0.0542	0.0000
3.4, Assessment and validation of intermediate data and statistical outputs	1.1187	0.0713	0.0000
3.5, Revision studies	1.7337	0.0632	0.0000
4.1, Periodicity and timeliness	0.6107	0.1048	0.0000
4.2, Consistency	1.3652	0.0548	0.0000
4.3, Revision policy and practice	1.5120	0.0704	0.0000
5.1, Data accessibility	1.3069	0.0623	0.0000
5.2, Metadata accessibility	1.6863	0.0610	0.0000
5.3, Assistance to users	1.0386	0.0577	0.0000

Table 7. DQAF Entropy Analysis for Government Finance Statistics

Indicator	Entropy	Stdev	p-val
0.1, Legal and institutional environment	1.2888	0.0582	0.0000
0.2, Resources	1.4935	0.0704	0.0000
0.3, Relevance	1.3044	0.0737	0.0000
0.4, Other quality management	1.2009	0.0619	0.0000
1.1, Professionalism	0.5961	0.0737	0.0000
1.2, Transparency	1.2781	0.0601	0.0000
1.3, Ethical standards	0.1639	0.0669	0.0071
2.1, Concepts and definitions	1.3169	0.0575	0.0000
2.2, Scope	1.5424	0.0615	0.0000
2.3, Classification/sectorization	1.5439	0.0267	0.0000
2.4, Basis for recording	1.0716	0.0656	0.0000
3.1, Source data	1.3421	0.0565	0.0000
3.2, Assessment of source data	0.9035	0.0829	0.0000
3.3, Statistical techniques	0.7270	0.0678	0.0000
3.4, Assessment and validation of intermediate data and statistical outputs	0.9933	0.0764	0.0000
3.5, Revision studies	1.9110	0.0465	0.0000
4.1, Periodicity and timeliness	1.3861	0.0560	0.0000
4.2, Consistency	1.3774	0.0618	0.0000
4.3, Revision policy and practice	1.5229	0.0659	0.0000
5.1, Data accessibility	1.5113	0.0376	0.0000
5.2, Metadata accessibility	1.7182	0.0536	0.0000
5.3, Assistance to users	1.5186	0.0702	0.0000

Table 8. DQAF Entropy Analysis for Producer Price Statistics

Indicator	Entropy	Stdev	p-val
0.1, Legal and institutional environment	0.8157	0.1300	0.0000
0.2, Resources	1.4755	0.1084	0.0000
0.3, Relevance	1.2278	0.1020	0.0000
0.4, Other quality management	1.0012	0.1584	0.0000
1.1, Professionalism	0.1414	0.0951	0.0684
1.2, Transparency	1.1645	0.1109	0.0000
1.3, Ethical standards	0.0000	0.0000	NaN
2.1, Concepts and definitions	0.7641	0.1459	0.0000
2.2, Scope	1.3213	0.1369	0.0000
2.3, Classification/sectorization	1.0583	0.1316	0.0000
2.4, Basis for recording	0.7451	0.1651	0.0000
3.1, Source data	1.3687	0.0788	0.0000
3.2, Assessment of source data	1.0012	0.1584	0.0000
3.3, Statistical techniques	1.4688	0.0639	0.0000
3.4, Assessment and validation of intermediate data and statistical outputs	0.7451	0.1651	0.0000
3.5, Revision studies	1.1276	0.1419	0.0000
4.1, Periodicity and timeliness	1.0759	0.1674	0.0000
4.2, Consistency	0.7051	0.1486	0.0000
4.3, Revision policy and practice	1.2434	0.0981	0.0000
5.1, Data accessibility	1.5328	0.1218	0.0000
5.2, Metadata accessibility	1.5513	0.1206	0.0000
5.3, Assistance to users	1.0294	0.1295	0.0000

Table 9. DQAF Entropy Analysis for Consumer Price Statistics

Indicator	Entropy	Stdev	p-val
0.1, Legal and institutional environment	0.9192	0.0971	0.0000
0.2, Resources	1.5416	0.0719	0.0000
0.3, Relevance	1.0830	0.0619	0.0000
0.4, Other quality management	1.0285	0.0682	0.0000
1.1, Professionalism	0.5228	0.1086	0.0000
1.2, Transparency	1.2872	0.0671	0.0000
1.3, Ethical standards	0.3675	0.0889	0.0000
2.1, Concepts and definitions	0.3127	0.0875	0.0002
2.2, Scope	1.3553	0.0623	0.0000
2.3, Classification/sectorization	1.0409	0.1023	0.0000
2.4, Basis for recording	0.6554	0.0804	0.0000
3.1, Source data	1.4146	0.0562	0.0000
3.2, Assessment of source data	1.0758	0.0926	0.0000
3.3, Statistical techniques	1.3407	0.0691	0.0000
3.4, Assessment and validation of intermediate data and statistical outputs	0.6965	0.1046	0.0000
3.5, Revision studies	0.9620	0.1111	0.0000
4.1, Periodicity and timeliness	0.1068	0.0674	0.0565
4.2, Consistency	0.4645	0.0888	0.0000
4.3, Revision policy and practice	1.1736	0.0755	0.0000
5.1, Data accessibility	1.2487	0.0664	0.0000
5.2, Metadata accessibility	1.4266	0.0585	0.0000
5.3, Assistance to users	1.0562	0.0798	0.0000

VIII. APPENDIX II: MATCHING THE OLD AND NEW DQAF CODES

The DQAF codes after 2003 do not exactly match the original DQAF codes. While no substantive changes in the DQAF codes were made in the 2003 revision, the numbering of the codes changed. In order to use the entire sample of data ROSC missions, the DQAF codes need to be aligned. The alignments are presented in Table 10 and are color-coded—indicators of the old code and the matched new code are shown in the rightmost column.

Table 10. Correspondence of Old and New DQAF Codes

Current codes		Old codes		Matched new code
0.1	Legal and institutional environment	0.1	Legal and institutional environment	0.1
0.2	Resources	0.2	Resources	0.2
0.3	Relevance			4.1
0.4	Other quality management	0.3	Quality awareness	0.4
1.1	Professionalism	1.1	Professionalism	1.1
1.2	Transparency	1.2	Transparency	1.2
1.3	Ethical standards	1.3	Ethical standards	1.3
2.1	Concepts and definitions	2.1	Concepts and definitions	2.1
2.2	Scope	2.2	Scope	2.2
2.3	Classification/sectorization	2.3	Classification/sectorization	2.3
2.4	Basis for recording	2.4	Basis for recording	2.4
3.1	Source data	3.1	Source data	3.1
3.2	Assessment of source data	3.2	Statistical techniques	3.3
3.3	Statistical techniques	3.3	Assessment and validation of source data	3.2
3.4	Assessment and validation of intermediate data and statistical outputs	3.4	Assessment and validation of intermediate data	3.4
3.5	Revision studies	3.5	Revision studies	3.5
4.1	Periodicity and timeliness	4.1	Relevance	0.3
4.2	Consistency	4.2	Timeliness and periodicity	4.1
4.3	Revision policy and practice	4.3	Consistency	4.2
		4.4	Revision policy and practice	4.3
5.1	Data accessibility	5.1	Data accessibility	5.1
5.2	Metadata accessibility	5.2	Metadata accessibility	5.2
5.3	Assistance to users	5.3	Assistance to users	5.3

IX. APPENDIX III: DQAF DIMENSIONS

0. Prerequisites of Quality

- 0.1. Legal and institutional environment
- 0.2. Resources
- 0.3. Relevance
- 0.4. Other quality management

1. Assurance of Integrity

- 1.1 Professionalism
- 1.2. Transparency
- 1.3. Ethical standards

2. Methodological Soundness

Elements of methodological soundness include:

- 2.1. Concepts and definitions
- 2.2. Scope
- 2.3. Classification/ sectorization
- 2.4. Basis for recording

3. Accuracy and Reliability

- 3.1. Source data
- 3.2. Assessment of source data
- 3.3. Statistical techniques
- 3.4. Assessment and validation of intermediate data and statistical outputs
- 3.5. Revision studies

4. Serviceability

- 4.1. Periodicity and timeliness
- 4.2. Consistency
- 4.3. Revision policy and practice

5. Accessibility

- 5.1. Data accessibility
- 5.2. Metadata accessibility
- 5.3. Assistance to users