HOW TO EVALUATE THE RELIABILITY OF THE REGIONAL INPUT-OUTPUT DATA? A CASE FOR CHINA

HAOYANG ZHAO [a], JIAN XU [a] and JING HE [b]

[a] *School of Economics and Management, University of Chinese Academy of Sciences, Beijing, People's Republic of China;* [b] *School of Management, Beijing University of Chinese Medicine, Beijing, People's Republic of China*

Abstract

Accurate statistic data are essential to a credible and cogent empirical analysis. For now, however, there is no mature and specialized methodology on how to evaluate accuracy of any input-output (IO) data. This research constructs a comprehensive yet relatively concise framework for evaluating regional IO data accuracy by including several indicators measuring all three quadrants. The framework examines regional IO data from following perspectives: time consistency and variation, coefficient correlation and whether it matches with national level data. A score indicating the overall accuracy as well as detailed information presenting concrete shortcomings of regional IO data could be offered after evaluating by this framework. As an example, the province-level IO data from 90 provinces IO tables in three consecutive session (2002, 2007, 2012) are being analyzed under the above framework. The main contribution and innovation of the research is building the applicable and exhaustive quality evaluation framework for regional IO data. This framework has meanings for both scholars and governments. It enables researchers realize potential flaws in IO data before utilizing it and government statistic agency improve data qualities by avoiding issues emerged in previous data quality evaluation.

Keywords: data quality; quality evaluation; China; regional IO data;

1. INTRODUCTION

As an increasing number of economic researches concern about structural issues, value-added in multi-sector international trade, for example, Input-Output (IO) data have become a prevalent data source and been used more frequently in empirical analyses than ever. Regarding this trend, an important issue naturally emerges, that whether the IO data is so accurate and reliable enough that researchers are able to utilize it to present credible empirical results. In fact, there has been relatively large quantity of researches dedicating to the quality of government statistics and it turns out that no socioeconomic statistical data is completely precise due to the following reasons: statistical regime defects (Xu, 1994; Holz, 2013), investigation and aggregation errors (Agafitei et al, 2015), and lack of independency in statistics agencies (Outrata, 2015). It is reasonable to presume that IO statistics, as a type of government statistical data, also suffer from some quality issues. Thus, the core problem is how to assess the current IO data and find out where the quality issues reside.

Before implementing evaluations, standards respecting statistics quality need to be specified. It is obvious that international organizations have major concerns for a long period over what standards statistical data should follow. The first organization that paid attention to the quality of statistical data was UN, which dated back to 1980. Different from other organizations, UN (2003) primarily focus on optimizing the structure of statistical agencies, arguing these agencies are supposed to obtain independency, relevance, credibility and respondent policy as their foundation. IMF (2013) require all the subscribers of SDDS, abbreviation for Special Data Dissemination Standard, to follow four statistics criteria, which are (1) data coverage, periodicity, and timeliness, (2) access by the public, (3) Integrity, and (4) Quality, which basically means the methodology and data are reasonable, like passing cross-check. Released also by IMF (2013), the GDDS, which is designed for relatively less developed government statistical systems, shares generally the same requirements. Compared with the above, OECD and Eurostat provide more detailed standards. OECD (2011) measures statistical data from eight dimensions including accuracy, coherence, timeliness and accessibility, etc. Eurostat (2011) presents a code constituted by 15 principles covering the institutional environment, the statistical production processes and the output of statistics, in which accuracy, coherence and comparability are included as well. Except for organizations, individual researchers also set several data quality standards including accuracy, timeliness and availability (Brackstone, 1999).

In terms of methods in data quality evaluations, all methods bifurcate into two branches, the data-driven one and the theory-driven one. The data-driven branch, consisting methods that purely are based on data itself and use statistics only, mainly focus on finding outliers in a group of data points and determine the quality of data with the number of outliers. For instance, Zhang (2003) introduces a statistical test on data to find outliers by assuming the distribution of data is exponential. Another example comes from machine learning, which offers various algorithms can be used to separate outliers from the remaining points, like support vector machine (James et al, 2014).

The methods in theory-driven branch, however, design indicators based on economic theories. These methods again fall in to two categories depending on how many indicator(s) they use. Some methods only use limited one or two statistics that closely link with data, which needs to be tested, as quality indicator(s) (Huenemann, 2001), while other methods construct a multi-indicator system (Klein & Özmucur, 2011; Ye, 2011) or use multi-variable regression method (Liu & Huang, 2009). Besides, variations and trends are common indicators when analyzing consistency of data, which are measured using time series analysis or simple comparison (Sinton, 2001). Differed from all these methods based on real calculation, Wang & Jin (2010) created a questionnaire including questions measuring respondents' subjective impressions of quality of statistics.

Several conclusions could be drawn from researches discussed above. An apparent issue is that none of these research is IO specified. The majority concentrate on GDP, and the remaining study transportation, energy and other particular areas except Input-Output data. A derivative problem is that although standards or principles remain the same, these methods are only compatible with single statistical data reflecting economic scale. IO data, on the other hand, consists hundreds of interrelated statistics which demonstrate economic scale and structure simultaneously. The delicate correlations between data indicate that a systematic method, or framework, need to be invented to evaluate the quality of IO data.

Therefore, the main contribution of this paper is to construct a plausible framework to evaluate regional IO data. The reason that assess regional instead of national data is that a benchmark is necessary during evaluation and usually national data possess better quality, that is, more appropriately serve as the benchmark.

To be precise, not all the principle of statistical data mentioned above will be implemented in the following IO data analysis. Since the objective of this paper is to evaluate the quality of data, standards like availability which mainly measure the quality of services of statistics agencies instead of data are omitted. Also, standards that are not applicable in IO data, like coverage, are also obliterated. Standards mainly measured in this paper are accuracy, coherence between regional and national data, and consistency in time series.

The remaining part of this paper is arranged as follows. In section 2, a framework evaluating regional IO data as a whole and individually is constructed. Section 3 is the empirical analysis using the framework constructed in section 2, taking China for example. Section 4 is the conclusion.

2. CONSTRUCTING THE EVALUATION FRAMEWORK

An IO table consists of hundreds even thousands of interrelated numbers. While structural information conveyed through interrelationships is highly favored by scholars and policy makers, the quantity of numbers and correlations between each other can be serious troubles compared with single-number data, like GDP, when evaluating data quality since it is simply not possible to find reference indicators outside the IO table for every single number. Therefore, when constructing the evaluation framework, two premises have been set as follows.

(1) the source of data used in an evaluation is the regional IO data itself;

(2) only limited but representative data will be evolved in analysis of IO data quality. Specifically, relatively important (large enough) direct input coefficients, or key coefficients (KC) will be representative numbers.

Then, it is reasonable to construct an indicator using the ratio of the number of aberrant KC(s) to the number of all KCs. The higher the ratio is, the worse the quality of an IO table is. Besides, it can also be used to examine the IO data quality in certain sector, or even data of all regions as a whole (national IO data system).

Next, a crucial question is how to define aberrant KC(s). From IO theory, direct input coefficients, as symbols of production technology, remain stable in a not too long term. Hence, once a mutation occurs, that is, a zero KC turns into a significant non-zero one in next year, or vice versa, this KC is categorized as aberrant. The reason of using the name "mutation" is that these kinds of changes usually indicates technology revolution in production, sectors that never existed before emerge or sectors died. Any of these changes can be regard as so tremendous that highly unlikely to happen in a short period. Nevertheless, the stableness KCs keep is not absolute, minor changes are inevitable between two accounting years. Those changes, however, are neither random nor without constraints. An assumption is that these changes are supposed to follow similar features, or trend for KCs within a sector, since the same national macroeconomic and industry policies, and similar market and technology conditions are shared by all KCs in a certain sector no matter what region the data represent. Accordingly, if some change(s) of KC(s) become outliers of all changes, these KC(s) are also considered as aberrant.

So, what standards are satisfied if this aberrant KC indicator is low enough? The apparent one is time-consistency, as the change is so minor that data in one year do not contradict one in another. Another standard is accuracy, though not directly. Imagine an alternative scenario, that quite a few numbers are erratic. In this case, data in at least one year are not accurate because data in one year challenge another. Conclusively, an ideal indicator of aberrant KC ratio does not necessarily mean accurate data, whereas a poor indicator certainly shows flaws in data. In other words, this indicator is the prerequisite, but not sufficient condition of accuracy.

A simplified example is shown below to illustrate aberrant KC indicator more clearly. Suppose a nation with 2 regions, 3 sectors. By some method, 8 direct input coefficients have been confirmed as KCs including all direct input coefficient except a_{12}. The locations of aberrant KCs are given in table 1 and the number of aberrant KCs have been displayed in table 2.

TABLE 1. An example of locations of aberrant KCs in an imaginary nation
(Aberrant KCs locate in shadow cells)

	Region 1			Region 2		
	Sector 1	Sector 2	Sector 3	Sector 1	Sector 2	Sector 3
Sector 1						
Sector 2						
Sector 3						

Be noted, when and only when counting the number regarding sectors, if a_{mn} is a KC, it is treated as KC both in sector m and sector n. And if $m = n$, a_{mn} is treated as 2 KCs in sector m (n).

TABLE 2. An example of the number of aberrant KCs in an imaginary nation

	Region 1	Region 2
The number of aberrant KCs	3	4
in Sector 1	2	2
in Sector 2	2	3
in Sector 3	2	3

Therefore, the KC indicators are calculated as follows:

$$\text{ratio of KC (Region 1)} = 3/8 = 0.375$$
$$\text{ratio of KC (Region 2)} = 4/8 = 0.500$$
$$\text{ratio of KC (Sector 1)} = (2+2)/(5+5) = 0.400$$
$$\text{ratio of KC (Sector 2)} = (2+3)/(5+5) = 0.500$$
$$\text{ratio of KC (Sector 3)} = (2+3)/(6+6) = 0.417$$
$$\text{ratio of KC (Nation)} = (3+4)/(8+8) = 0.438$$

Coherence in each sector and all regions together also need to be taken into consideration. From analyses above, evaluations for individual regions have to compare data with data from the same region in a different year to draw conclusions. Therefore, evaluations have no choice but to deal with a data package including two-year datasets and display results of two year together inseparably. However, as for each sector and all regions together, national data serves as a benchmark and results of single year is available.

To utilize national level statistics as benchmarks, a new indicator is introduced. A simple character of a coherent data system is the aggregation of regional level data approximately equals to the national one. Accordingly, a ratio of aggregation to national data is a reasonable measurement, regarding total output, consumption, capital formation, labor compensation, etc. If ratio of a sector is deviated from 1 significantly, the data of the sector is incoherent, as data contradicts each other even inside the system. Thus, the number of ratios of data of incoherent sectors (IS) to the numbers of all data is proper as another indicator which mainly aimed at data quality of each sector and all regions together. Similarly, the aggregation of total inward flows and outward flows (goods and

services imported from / exported to other region but not from / to foreign countries) should approximately equal. If the ratio of aggregation of inward flows to one of outward flows is larger or smaller than 1 significantly in a certain sector, this sector needs to be treated as IS as well.
Here is an example illustrating this indicator. Suppose again a nation with three sectors and data of ISs have been identified and noted in table 3.

TABLE 3. An example of IS indicator in an imaginary nation
(ISs locate in shadow cells)

	Sector 1	Sector 2	Sector 3
Total output			
Consumption			
Capital formation			
Labor compensation			
Inward / outward flow			

Therefore, the IS indicators are calculated as follows:

$$\text{ratio of IS (Sector 1)} = 2/5 = 0.400$$
$$\text{ratio of IS (Sector 2)} = 3/5 = 0.600$$
$$\text{ratio of IS (Sector 3)} = 1/5 = 0.200$$
$$\text{ratio of IS (Nation)} = (2+3+1)/15 = 0.400$$

To sum up, table 4 demonstrates the quality evaluation framework for regional IO data established above.

TABLE 4. Quality evaluation framework for regional IO data

Aberrant KC indicators	Mutation Trend outlier
IS indicators	Total output Consumption Capital formation Labor compensation Inward / outward flow

The data quality of each region is given by aberrant KC indicators, while quality of each sector as well as the whole nation is the average of individual aberrant KC indicator and IS indicator.[1]
For example, the data quality of the imaginary nation in examples above is

$$(0.438 + 0.400)/2 = 0.419$$

3. EVALUATION OF PROVINCE-LEVEL IO DATA IN CHINA

In this section, in order to put the framework established in the last section into practice, China's province-level IO data are analyzed as examples.

---

[1] Readers may notice that aberrant KC indicators are calculated for a two-year dataset package, while IS indicators are calculated in every single accounting year, which brings trouble for calculating the average. One possible solution is given in the next section.

3.1 Data

The IO accounting years in China ends with 2 or 7 (based on real IO survey) and 0 or 5 (update using general national accounting data). IO data in recent three consecutive accounting years based on real survey, 2002, 2007 and 2012, constitute the data source. The evaluation includes all provinces in China except Tibet and Taiwan. In short, 30 province-level table in 3 years, that is, 90 tables are included. Besides, the national IO table in these years are used for benchmarks. All data are available from National Bureau of Statistics.

All tables used here contain 42 sectors. However, there are some minor changes in sector classifications between any two years. Therefore, sector adjustment has been implemented to keep the sector classification consistent in time series. Therefore, all tables are modified into a 39-sector version. Specific changes on sectors and sector classification after adjustments are listed in appendix. Without further notification, calculations in this section are processed by R and Excel. R codes are available in supplemental information at the end of the paper.

3.2 Aberrant KC indicators

Choose Key Coefficients (KCs). It is true that there are several methods to choose KCs, but to simplify the calculation, a single rule has been adopted for choosing KCs that if a_{mn} is larger than 0.05 in two years of three with respect to national IO tables, a_{mn} is identified as a KC. After calculation, 87 coefficients (5.72% of all coefficients) satisfy the rule. The sum of these coefficients account for 47.39% (2002), 52.48% (2007) and 53.15% (2012) of sums of total direct input coefficients in three years. Figure 1 shows the sector distributions of KCs.

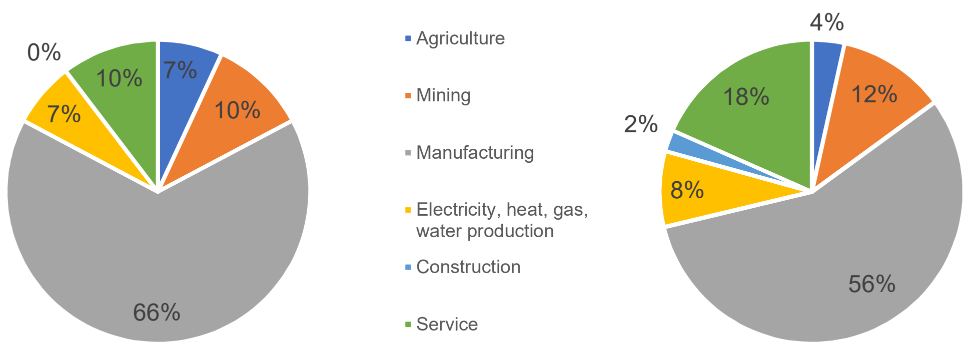

FIGURE 1. Sector distributions of key coefficients, viewed by row (left) and column (right)

Mutation. As mentioned above, mutation means a sudden change in number from zero to non-zero or vice versa. After examining 2,610 key coefficients, there are 66 (2002-2007) and 52 (2007-2012) mutations arise. Viewed by province, Qinghai possess the most mutations (13) in 2002 to 2007, and no other province own a mutation number over ten, no matter what year. Most mutations happened in provinces in middle and western China. As for sectors, most mutations happens in sector coal mining products, and other manufacturing, both 20 mutations in 2002-2007, and sector other manufacturing also have 20 mutations in 2007-2012, which is the most of the period, followed by 15 mutations in sector coal mining products and 14 mutations in gas production and supply. Detailed results of mutation, along with results of following evaluations can all be found in appendix.

Trend outlier. Normally the changes between two consecutive years shares some similarities, or trend mentioned in section 2. So, trend-breaker(s) are signs of flaws in data. However, how to identify these breakers, or outliers? Imagine a scatter plot which shows coefficients of a KC in all regions. Two axis represents values of coefficients in different years. The existence of a certain trend means normal data points should tend to concentrate to some extent. Outliers, however, do not concentrate with normal data points (hence the name). Figure 2 shows the general idea of this scatter plot. From this plot, it is transparent that point A and B are trend outliers.

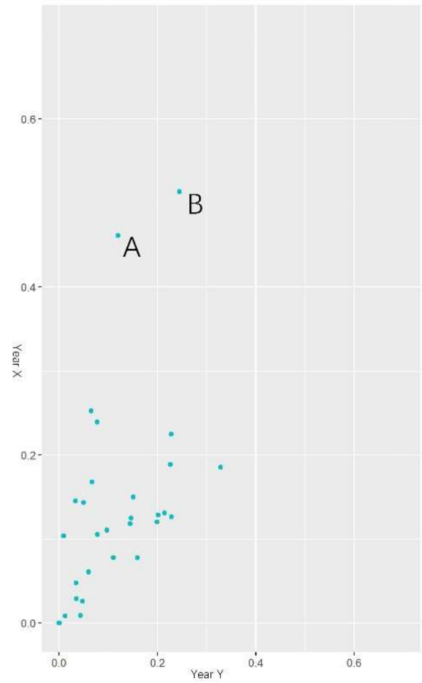

FIGURE 2. A general idea of scatter plot and trend outliers

Yet not all outliers could be identified so inarguably. Based on this scatter plot, an algorithm is introduced to help find outliers.

(1) Calculate the center of points in the plot using leave-one-out method, the coordinate of the center is given by arithmetic average of the coordinate of each point except the leaved-out one;

(2) Calculate the Euclidean distance between the center and each point except the leaved-out one, sum up all the distances;

(3) Repeat step (1) – (2) while change the leaved-out one. Stop repeating until all points are leaved out for one time;

(4) List all sums and use 2 times standard deviation rule to identify outliers in all sums;

(5) If a sum is identified as outlier, the corresponding point that was leaved out is the trend outlier.

Compared with mutations, trend outliers are relatively more in quantity. Individually, 209 and 194 trend outliers are identified in each period. In 2002-2007, Hainan had the most trend outliers (12), followed closely by Qinghai (11) and surprisingly, Beijing (11). Cases are similar in 2007-2012, although Qinghai (17) ranked first this time followed by Hainan (14) and Beijing (11). In terms of sectors, the two sectors own the most trend outliers are Chemical products (27, 27) and Metal smelting and rolling processing (18, 20) in both periods.

Summary. Figure 3 shows the aggregation of mutations and trend outliers, that is, all aberrant KCs in each province while figure 4 shows in each sector.

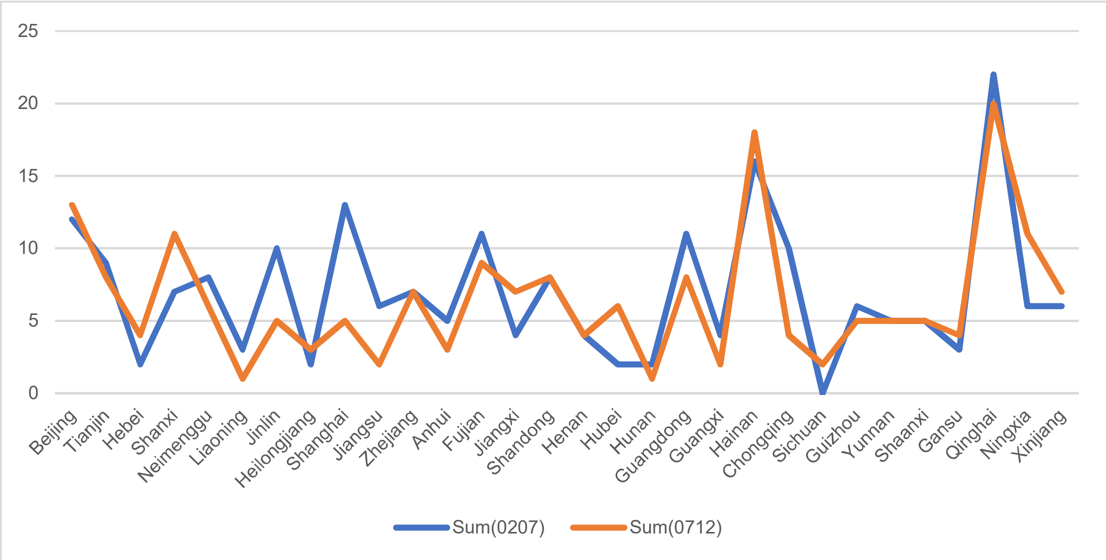

FIGURE 3. Aberrant key coefficients in each province

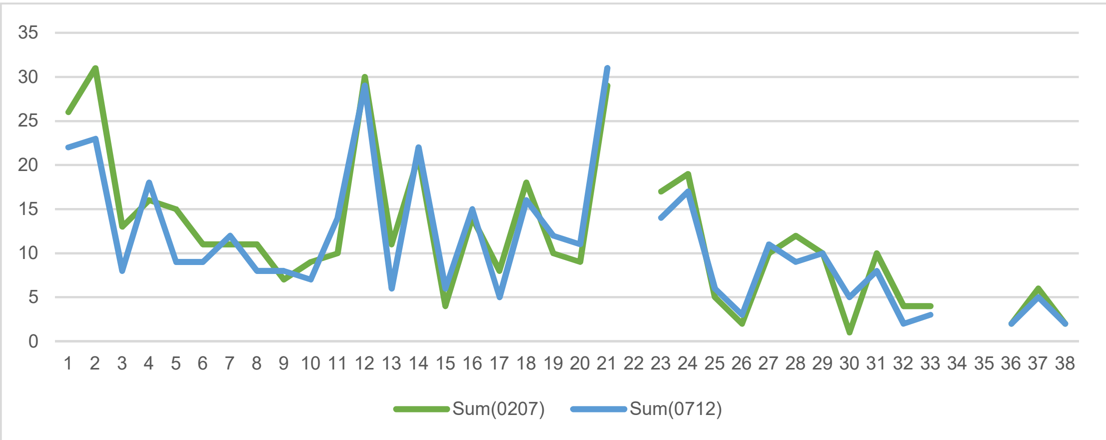

FIGURE 4. Aberrant key coefficients in each sector

From figures above, it is apparent that there is only a slight improvement in data quality in 2007-2012 compared with 2002-2007 when measured with aberrant KCs, and this conclusion does not hold true in all sectors and regions. Another transparent conclusion is the correlation between two periods. To be precise, the Pearson correlation coefficients and significance tests are calculated and showed in table 5. It turns out all KC indicators are positively correlated when significance level α is 0.05. In fact, expect for mutations, correlations of all indicators are statistically significant even when α equals to 0.01.

TABLE 5. The correlation between two periods regarding aberrant key coefficients

	Mutation	Trend Outlier	Sum
By province*	0.437 (0.016)	0.787 (0.000)	0.798 (0.000)

By sector	0.933 (0.000)	0.915 (0.000)	0.938 (0.000)

*Note: numbers in parenthesis are significance level.

As no IS indicator (discussed in the next subsection) is designed for regional evaluation, data quality in each province is given by aberrant KC indicators. Table 6 shows the first 5 best and worst quality provinces.

TABLE 6. Best and worst IO data quality regarding province in each period

Best data quality, 1-5			Worst data quality, 1-5		
Rank	Province	aberrant KCs*	Rank	Province	aberrant KCs
Period 2002-2007					
1	Sichuan	0.000	1	Qinghai	0.253
2	Hebei	0.023	2	Hainan	0.184
2	Heilongjiang	0.023	3	Shanghai	0.149
2	Hubei	0.023	4	Beijing	0.138
2	Hunan	0.023	5	Guangdong	0.126
Period 2007-2012					
1	Hunan	0.011	1	Qinghai	0.230
1	Liaoning	0.011	2	Hainan	0.207
2	Sichuan	0.023	3	Beijing	0.149
2	Guangxi	0.023	4	Shanxi	0.126
2	Jiangsu	0.023	4	Ningxia	0.126

*Note: aberrant KCs means the ratio of aberrant KCs to all KCs, lower is better.

3.3 IS indicators

The data quality in each sector has been evaluated and presented above. However, IS indicators still need to be calculated to assess the data quality of sectors and the whole nation's IO data system.

Ratios. To identify incoherent sectors (ISs), the ratio of aggregation of all province-level data to real national data have to be calculated first. Take total output for example.

$$\text{ratio of sector A} = \frac{\sum \text{total outputs of sector A in all province}}{\text{national total output in sector A}}$$

This formula is compatible with all IS indicators in table 4 except for inward / outward flow, which should apply the formula as follows,

$$\text{ratio of sector A} = \frac{\sum \text{total inward flow of sector A in all province}}{\sum \text{total outward flow of sector A in all province}}$$

In this analysis, data of inward / outward flows are only available in 2012. Figure 5 shows ratios of total output calculated using the formula above.

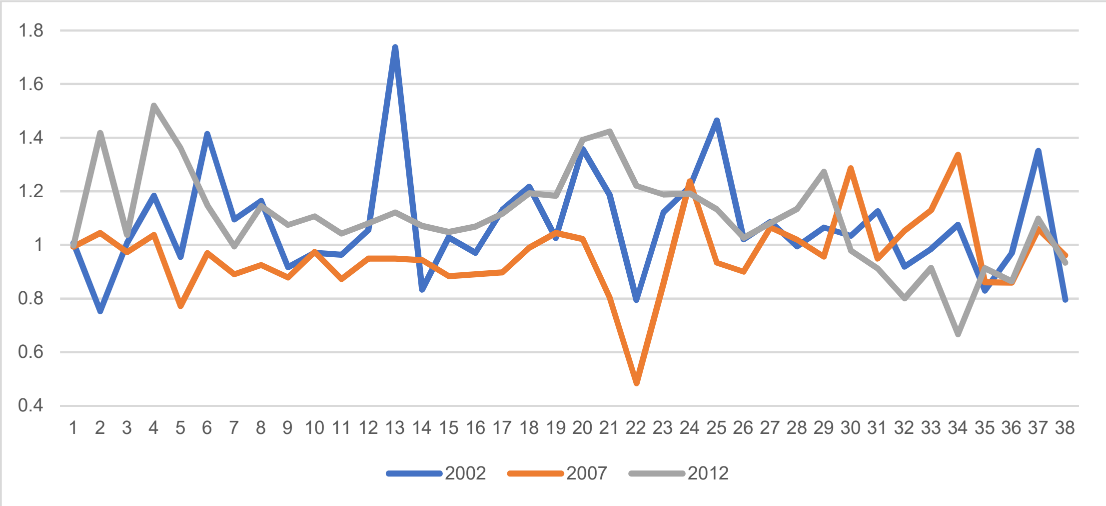

FIGURE 5. Ratios of the aggregation of total outputs in all regions to the real national one

Theoretically, all ratios are supposed to equals, or at least approximately equals to 1. But the norm is that these ratios may be a little bit larger or lesser than 1 for the following reasons:

(1) Price standard. Local producer prices are used in regional table instead of national one used in national table;

(2) Lack of data. In 2002 and 2007, Tibet did not conduct IO investigations and thus have no IO table.

(3) Statistical errors.

Despite all these reasons, differences between aggregated data and real national data still should be slight. For one thing, the economic scale of Tibet is pretty small, even compared with other middle and western China provinces, which themselves are less developed compared with eastern provinces. For another, issues of price levels and errors are usually minor. The price levels within a country should converge according to free market theory, and a national price level could be considered as an average. As for errors, a large statistical error itself is a sign of low data quality.

IS indicators. In following analysis, sectors with ratios larger than 1.2 or lesser than 0.8 are considered as incoherent sectors (ISs). From figure 5, the majority of total output ratios locate in this range, while there indeed are a few ratios too large or small. But in term of ratios of final demand (consumption, capital formation) and labor compensation, two features need to be stressed: (1) More peculiar ratios emerge. For instance, only 19.3% ratios of capital formation lie in the range 0.8 to 1.2; and (2) More extreme ratios emerge. Still, in ratios of capital formation, some ratios are negative and some ratios are larger than 30. This highly extreme ratios indicate that the national IO data system could not maintain coherence within it, and the reliability of data should be questioned.

3.4 Data quality of sectors and national IO system

With all indicators calculated, data quality of sectors and the whole system can be evaluated. Firstly, the quality of each sector is the arithmetic average of aberrant KC indicators and IS indicators.[2] The problem is that aberrant KC indicators are calculated in a two-year package. To solve this, aberrant

[2] Some sectors do not have key coefficients. In that case, the IS indicator singly determines the data quality of the sector.

KC indicators for 2002-2007 are treated as indicators for 2002, indicators originally for 2007-2012 are treated as indicators for 2012, while indicators for 2007 is the average of those two. Figure 6 shows the quality of individual sectors in three years.

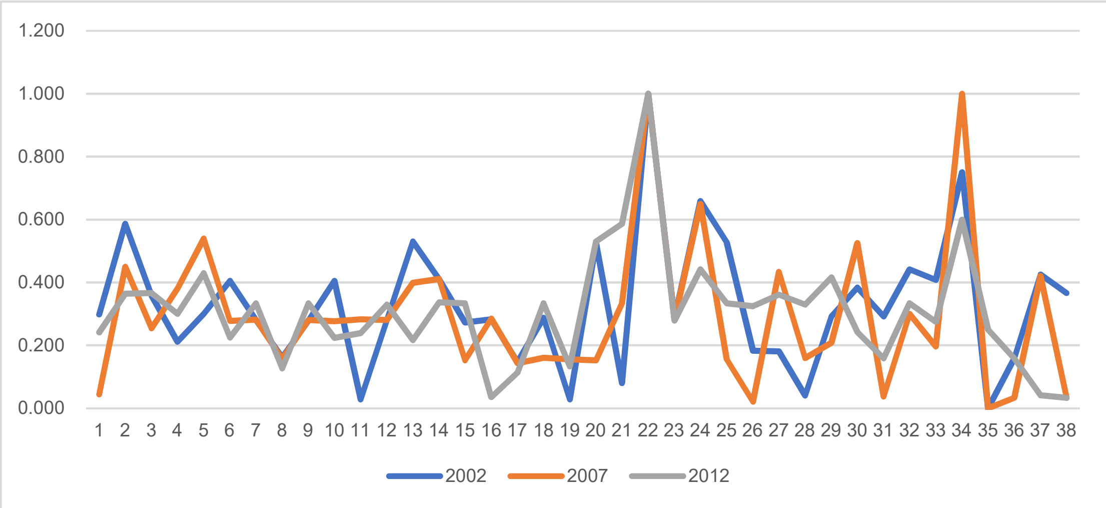

FIGURE 6. IO Data quality in each sector
(lower is better)

From figure 6, data qualities of scarp processing sector, Gas production and supply, and R&D and technical service are worst among all sectors in all three years, whereas sectors like agriculture and agricultural services, communication, computer and other electronic equipment, and education possess a relatively good data quality. The correlations of data quality of sectors in different years are also calculated. The results (showed in table 7) indicates that the data qualities in different years are strongly positively correlated.

TABLE 7. Correlations between data quality of individual sectors in different years

	2002 & 2007	2002 & 2012	2007 & 2012
correlation coefficients	0.695	0.544	0.637
test p-value	(0.000)	(0.000)	(0.000)

At the end of this section is the calculation of IO data quality of national IO data system, or the country as a whole. Method is the same with calculations of sectors and table 8 shows the result. For overall data quality, data quality in 2007 is only slightly better than other two years. While the least aberrant KC indicators occur in 2012, IS indicators in 2007 is better in a relatively large extent. Data quality in 2002 is the worst with respect to aberrant KC indicators and IS indicators.

TABLE 8. IO Data quality of national IO data system
(lower is better)

	2002	2007	2012
aberrant KC indicators	0.080	0.077	0.074
IS indicators	0.537	0.469	0.503
Data quality	0.309	0.273	0.289

4. Conclusion

In this paper, a framework for evaluating regional IO data has been constructed for the first time. It contains two types of indicators, aberrant key coefficients, which intend to measure accuracy and consistency in time series, and incoherent sectors, which are designed for examining coherence. Therefore, the framework takes the most important standards of data quality into account.

The framework possesses several features. Its structure is relatively simple compared with data quality measuring systems now existed. However, it covers various issues, including data quality of regions, sectors and the whole IO data system in a nation and utilize information from all three quadrants. Besides, evaluations under this framework do not need additional data from other sources, which, along with its simple structure, makes the framework easy to apply.

As a trail example, China's regional IO data is evaluated under the framework. Viewed by province, not only less developed provinces, like Qinghai, have quality issues of IO data, but relatively highly developed areas, like Beijing, also suffer from low IO data quality. As for sectors, extreme and unstable IS indicators expose that coherency problem cannot be ignored. In terms of the national IO data system, the overall data quality is not ideal enough, and there is no significant data quality improvement within 2002 to 2012.

Granted, the framework presented in this paper is far from perfect. A better, more accurate evaluation that reveals more information could be achieved by replenishing more data from additional data source or exploiting more interrelationships between numbers in IO table. It is true that for government statistic agencies, amelioration of this type in order to build a more comprehensive evaluation system is necessary for continuous improvements of data quality. However, for researchers and policy makers who would like to know about quality of the data they would like to use, this simply-constructed framework certainly provides an efficient and user-friendly instrument.

SUPPLEMENTAL INFORMATION

The R code for analyzing key coefficients in Section 3 can be found on GitHub:
https://github.com/zhaohaoyangruc/IO-data-quality

Reference

Brackstone, G. (1999). Managing data quality in a statistical agency. Survey Methodology, 25, 139--149.

Eurostat (2011). European Statistics Code of Practice. http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-32-11-955.

Holz, C. A. (2013). Chinese statistics: classification systems and data sources. Eurasian Geography and Economics, 54:5-6, 532-571

Holz, C. A. (2013). The quality of china's gdp statistics. A/ Working Papers, 30(35), 309-338.

Huenemann, R. W. (2001). Are china's recent transport statistics plausible?. China Economic Review, 12(4), 368-372.

International Monetary Fund (2013). The General Data Dissemination System Guide for Participants and Users.

http://www.imf.org/external/pubs/ft/gdds/guide/2013/gddsguide13.pdf.

International Monetary Fund (2013). The Special Data Dissemination Standard 2013 Guide for Subscribers and Users. http://dsbb.imf.org/Pages/SDDS/WhatsNew.aspx.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014). An Introduction to Statistical Learning: with Applications in R. Springer Publishing Company, Incorporated, 337-372.

Jin, Y. & Tao R. (2010). The Theoretical Research and Practice of Statistical Data Quality in China. Statistical Research, 27(1), 62-67. (in Chinese)

Klein, L. R., & Özmucur, S. (2011). The estimation of china's economic growth rate. Journal of Economic & Social Measurement, 28(4), 277-285.

Liu, H. & Huang, Y.(2009). An Evaluation Method of Statistical Data Quality Based on the Classical Econometric Model. Statistical Research, 26(3), 91-96. (in Chinese)

Mehrotra, A., & Pääkkönen, J. (2011). Comparing china's gdp statistics with coincident indicators. Journal of Comparative Economics, 39(3), 406-411.

OECD (2011). Quality Framework for OECD Statistical Activities. http://www.oecd.org/std/qualityframeworkforoecdstatisticalactivities.htm

Outrata, E. (2015). Influence of governance issues on the quality of official statistics. Statistical Journal of the Iaos, 31(4), 523-527.

Park, A., & Wang, S. (2001). China's poverty statistics. China Economic Review, 12(4), 384-398.

Rawski, T. G. (2001). What is happening to china's gdp statistics?. China Economic Review, 12(4), 347-354.

Reis, F., Gras, F., Kloek, W., Vâju, S., & Agafitei, M. (2015). Measuring output quality for multisource statistics in official statistics: some directions. Statistical Journal of the Iaos, 31(2), 203-211.

Sinton, J. E. (2001). Accuracy and reliability of china's energy statistics. China Economic Review, 12(4), 373-383.

United Nations Department of Economic and Social Affairs Statistics Division (2003). Handbook of Statistical Organization (Third Edition). https://unstats.un.org/unsd/dnss/hb/default.aspx.

Wang, H. & Jin, Y.(2010). Statistical Data Quality and Users' Satisfaction: Designing Evaluation Scale and Empirical Study. Statistical Research, 27(7), 9-19. (in Chinese)

Xu,X.(1994). The differences between the new Chinese national account system and SNA 1993 (Part II). Statistical Research, 11(6), 24-24. (in Chinese)

Ye, S. (2011). Study on the Evaluation Methods of Official Statistical Data Qualityand Its Application. (Doctoral dissertation, Hunan University). (in Chinese)

Zhao, X.,Wang, X. & Liu, K. (2011). Research on Quality of the China's Statistics: A Literature Review. Economic Review(1), 145-154. (in Chinese)

Appendix

TABLE A.1. Sector adjustment for province-level IO table in China

Merge into one sector	Move into sector "Other*"
General equipment , special equipment (2012)	Metal products, machinery and equipment repair services (2012)
Scientific research, Integrated technical services (2002)	Tourism (2002) Other social services (2002)
R&D, Integrated technical services (2007)	Water conservancy, environment and public facilities management (2007, 2012)

	Services to households and other services (2007) Services to households, repair and other services (2012)

*Note: All the sector not consistent in three years are moved into sector named other, which does not participate in evaluation.

TABLE A.2. Sectors after adjustment

No.	Sector	No.	Sector
1	Agriculture and agricultural services	21	Other manufacturing
2	Coal mining products	22	Scarp processing
3	Oil and gas products	23	Production and supply of electricity and heat
4	Metal mining	24	Gas production and supply
5	Non-metallic mining	25	Water production and supply
6	Food and tobacco	26	Construction
7	Textile	27	Wholesale and retail
8	Textile, leather and feather products	28	Transportation, storage and postal service
9	Wood products and furniture	29	Accommodation and catering
10	Paper, printing, stationery and sporting goods	30	Information transmitting, software and IT service
11	Petroleum, coking and nuclear fuel processing	31	Finance
12	Chemical products	32	Real estate
13	Non-metallic mineral products	33	Rental and business service
14	Metal smelting and rolling processing	34	R&D and technical service
15	Metal Products	35	Education
16	General and special equipment	36	Health and social work
17	Transportation equipment	37	Sports and entertainment
18	Electrical machinery and equipment	38	Public management, social security and social organizations
19	Communication, computer and other electronic equipment	39	Others
20	Instrumentation		

TABLE A.3. Aberrant KC indicators summed by province

	2002 to 2007			2007 to 2012		
	Mutation	Trend	Sum*	Mutation	Trend	Sum
Beijing	1	11	12	2	11	13
Tianjin	1	8	9	4	4	8
Hebei	0	2	2	1	3	4
Shanxi	0	7	7	2	9	11
Neimenggu	3	5	8	2	4	6
Liaoning	1	2	3	0	1	1
Jinlin	3	8	10	2	3	5
Heilongjiang	1	2	2	1	2	3

Shanghai	7	7	13	1	4	5
Jiangsu	0	6	6	0	2	2
Zhejiang	2	5	7	1	6	7
Anhui	0	5	5	0	3	3
Fujian	5	6	11	3	6	9
Jiangxi	1	3	4	1	6	7
Shandong	1	7	8	0	8	8
Henan	0	4	4	0	4	4
Hubei	0	2	2	1	5	6
Hunan	0	2	2	0	1	1
Guangdong	3	9	11	0	8	8
Guangxi	1	3	4	2	0	2
Hainan	5	12	16	5	14	18
Chongqing	4	6	10	0	4	4
Sichuan	0	0	0	0	2	2
Guizhou	3	3	6	2	3	5
Yunnan	1	4	5	1	4	5
Shaanxi	4	1	5	4	1	5
Gansu	1	2	3	1	4	4
Qinghai	13	11	22	4	17	20
Ningxia	3	3	6	7	4	11
Xinjiang	2	4	6	5	2	7
Nation Total	**66**	**150**	**209**	**52**	**145**	**194**

*Note: There are 7 (2002-2007) and 3 (2007-2012) overlapping aberrant KCs between mutations and trend outliers, which is the reason why sums are not always equal to the real sums of mutations and trend outliers.

TABLE A.4. Aberrant KC indicators summed by sector

	2002-2007			2007-2012		
Sector	Mutation	Trend	Sum*	Mutation	Trend	Sum
1	12	16	26	7	15	22
2	20	13	31	15	9	23
3	11	2	13	6	2	8
4	4	12	16	8	10	18
5	5	11	15	0	9	9
6	2	9	11	3	6	9
7	4	7	11	3	9	12
8	4	7	11	2	6	8
9	3	5	7	1	7	8
10	0	9	9	0	7	7
11	3	7	10	4	10	14
12	3	27	30	2	27	29
13	1	10	11	0	6	6
14	4	18	21	4	20	22
15	0	4	4	0	6	6

16	1	14	14	2	13	15
17	0	8	8	0	5	5
18	1	18	18	1	16	16
19	3	7	10	0	12	12
20	1	8	9	0	11	11
21	20	10	29	20	12	31
22**						
23	5	14	17	3	11	14
24	14	5	19	14	4	17
25	0	5	5	0	6	6
26	0	2	2	0	3	3
27	5	6	10	4	7	11
28	1	11	12	0	9	9
29	3	8	10	2	8	10
30	0	1	1	0	5	5
31	0	10	10	0	8	8
32	0	4	4	0	2	2
33	0	4	4	0	3	3
34						
35						
36	0	2	2	0	2	2
37	2	4	6	3	2	5
38	0	2	2	0	2	2
Total	**132**	**300**	**418**	**104**	**290**	**388**

*Note 1: Same with table A.3, overlapping between mutations and trend outliers has been eliminated.

**Note 2: The shadow cells indicate that there are no key coefficients in that sector.

TABLE A.5. IO data quality in each region

Province	2002-2007*	2007-2012	Province	2002-2007	2007-2012
Beijing	0.138	0.149	Henan	0.046	0.046
Tianjin	0.103	0.092	Hubei	0.023	0.069
Hebei	0.023	0.046	Hunan	0.023	0.011
Shanxi	0.080	0.126	Guangdong	0.126	0.092
Neimenggu	0.092	0.069	Guangxi	0.046	0.023
Liaoning	0.034	0.011	Hainan	0.184	0.207
Jinlin	0.115	0.057	Chongqing	0.115	0.046
Heilongjiang	0.023	0.034	Sichuan	0.000	0.023
Shanghai	0.149	0.057	Guizhou	0.069	0.057
Jiangsu	0.069	0.023	Yunnan	0.057	0.057
Zhejiang	0.080	0.080	Shaanxi	0.057	0.057
Anhui	0.057	0.034	Gansu	0.034	0.046
Fujian	0.126	0.103	Qinghai	0.253	0.230
Jiangxi	0.046	0.080	Ningxia	0.069	0.126
Shandong	0.092	0.092	Xinjiang	0.069	0.080

*Note: the meaning of numbers is the same with table 6, lower is better.

TABLE A.6. IS indicators summed by sectors

Sector	IS (2002)	IS (2007)	IS (2012)
1	0.500	0.000	0.400
2	1.000	0.750	0.600
3	0.500	0.333	0.600
4	0.333	0.667	0.500
5	0.500	1.000	0.800
6	0.750	0.500	0.400
7	0.500	0.500	0.600
8	0.250	0.250	0.200
9	0.500	0.500	0.600
10	0.750	0.500	0.400
11	0.000	0.500	0.400
12	0.500	0.500	0.600
13	1.000	0.750	0.400
14	0.750	0.750	0.600
15	0.500	0.250	0.600
16	0.500	0.500	0.000
17	0.250	0.250	0.200
18	0.500	0.250	0.600
19	0.000	0.250	0.200
20	1.000	0.250	1.000
21	0.000	0.500	1.000
22	1.000	1.000	1.000
23	0.500	0.500	0.500
24	1.000	1.000	0.600
25	1.000	0.250	0.600
26	0.333	0.000	0.600
27	0.250	0.750	0.600
28	0.000	0.250	0.600
29	0.500	0.333	0.750
30	0.750	1.000	0.400
31	0.500	0.000	0.250
32	0.750	0.500	0.600
33	0.750	0.333	0.500
34	0.750	1.000	0.600
35	0.000	0.000	0.250
36	0.250	0.000	0.250
37	0.750	0.750	0.000
38	0.667	0.000	0.000
National	**0.537**	**0.469**	**0.503**