China Economic Review 30 (2014) 353–369

Contents lists available at ScienceDirect

China Economic Review

Explaining sectoral discrepancies between national and provincial statistics in China

Ben MA [a,*], Guojun SONG [a], Lei ZHANG [a], David A. SONNENFELD [b]

[a] School of Environment and Natural Resources, Renmin University of China, Beijing 100872, China
[b] Department of Environmental Studies, State University of New York College of Environmental Science and Forestry, Syracuse, NY 13210, USA

ARTICLE INFO

Article history:
Received 24 March 2014
Received in revised form 14 July 2014
Accepted 14 July 2014
Available online 22 July 2014

JEL classification:
C43
C82
O43
O53
Q43
R58

Keywords:
Index decomposition
Institutional arrangements
Data discrepancy
Gross domestic product (GDP)
Energy consumption

ABSTRACT

This paper examines sectoral contributions to discrepancies between China's national aggregate statistical values and the sum of provincial figures. In institutional terms, the paper then explores the sources of principally sectoral discrepancies. We find that the industrial sector has been the major contributor to discrepancies in both gross domestic product (GDP) and total energy consumption in recent years. Technical aspects such as statistical coverage, data collection method, and double-counting cannot explain the discrepancy. For the industrial sector, limited data accessibility undermines external checks and balances from the general public. As the primary bodies in collecting industrial data, the Provincial Bureaus of Statistics (PBSs) are not subject to effective internal checks and balances from other governmental divisions. To out-compete counterparts and get promoted, provincial leaders have explicit incentives to overstate provincial GDP, with industrial added value being the first statistic to be affected. This dynamic further extends to industrial energy consumption, which is over-reported as well.

1. Introduction

While there is little doubt about China's fast economic growth over the last three decades, many domestic analysts and international observers are skeptical of the quality of China's official economic statistics. The doubts and criticisms concerning the quality of China's officially released economic data became evident in 2010, when China's gross domestic product (GDP) ranked second and its total energy consumption and fuel-induced carbon emission ranked first in the world (British Petroleum, 2011; IEA, 2012). Regarding GDP and total energy consumption, there have been long-standing inconsistencies between the nationwide aggregate statistical data and the sum of the corresponding provincial values. As shown in Fig. 1, the sum of the provincial GDP exceeded the national figure by 10.27% in 2011, that is, 4.86 trillion Chinese yuan (CNY), roughly three times the total GDP of Beijing or two thirds of that of South Korea. Meanwhile, the provincial energy consumption was in excess of the national figure by 743 million tons

* Corresponding author at: Office Room 302, School of Environment and Natural Resources, Renmin University of China, Haidian District, Beijing 100872, China. Tel.: +86 134 8889 3576.
E-mail addresses: mbruc@163.com, maben@ruc.edu.cn (B. Ma), songguojun@ruc.edu.cn (G. Song), lei.leizhang@gmail.com (L. Zhang), dsonn@esf.edu (D.A. Sonnenfeld).

http://dx.doi.org/10.1016/j.chieco.2014.07.004
1043-951X/

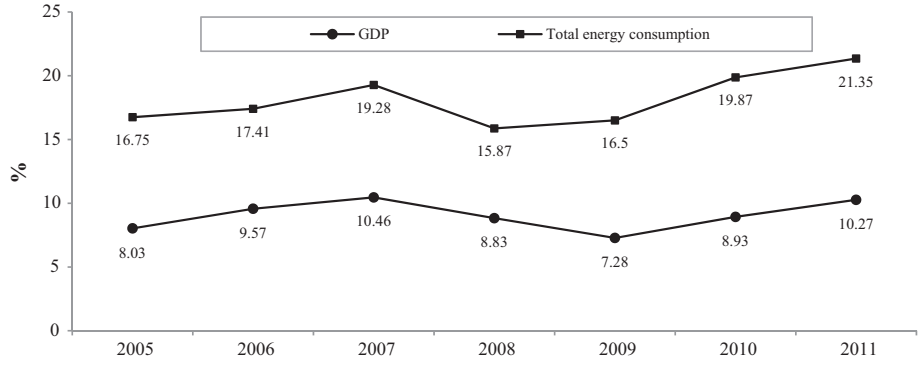

Fig. 1. The overall discrepancies for GDP and total energy consumption in China. The discrepancy is defined as the percentage that the sum of provincial figure exceeds the corresponding national aggregate value (see Section 2.1 for details). The "first confirmed" values are used.
Data sources: China Statistical Yearbook and China Energy Statistical Yearbook 2006 to 2012.

standard coal equivalence in 2011, which is twice that for Shandong Province, the largest energy consumer among the Chinese provinces. This data inconsistency has become a well-known issue after continuous coverage by the mass media,[1] undermining the public's basic trust in China's statistical analyses (Pan & An, 2003). Furthermore, the performance evaluation of provincial leaders by the central government in terms of energy and carbon intensity relies primarily on statistical data at the provincial level. Therefore, there is a need to carefully examine the discrepancy and its underlying causes to address this issue and facilitate the genuine achievement of national energy-conservation and carbon-reduction targets, reliable public decision-making and the improvement of the credibility of official Chinese statistics.

To date, the existing literature discussing the issues involving China's economic data mainly address GDP data quality (Holz, 2004a; Koch-Weser, 2013; Mehrotra & Pääkkönen, 2011; Rawski, 2001; Wang & Meng, 2001), with only a few studies primarily focusing on the data quality of energy consumption (Sinton, 2001; Song & Ma, 2013). To explore the data quality, the benchmark revisions following China's economic census are generally used based on the logic that census data are of higher quality than China's annual values obtained by routine statistics due to the former's comprehensive statistical coverage (Holz, 2008; Wang & Chandler, 2011). The value-added discrepancy in China is typically used as evidence by some analysts to question the added-value data quality (e.g., Holz, 2004b; Koch-Weser, 2013). Although the overall and, in some cases, the sectoral original discrepancies in the value-added data have been revealed, the overall discrepancy has not been decomposed into sectoral contribution shares or combined with the energy consumption data to conduct mutual cross-checks. Few studies report the discrepancy in energy consumption (Wang & Chandler, 2011; Zhang, 2010), let alone its decomposition by specific energy source and subsector. As a result, the quantitative contribution rates of each sector to the overall discrepancies in GDP and energy consumption remain unknown. With China's on-going statistical reform and its restructured economy and energy mix, the contribution percentages by sector will probably change steadily over time. Using the latest data, tracing the sectoral contributions to the overall discrepancy will be essential in further exploring the underlying causes.

Generally, the direct cause of the discrepancies between China and its provinces lies in the separate calculation regimes; that is, the statistical bureau at each level independently calculates its corresponding GDP and energy consumption based on its own data sources (NBS, 2013a). The separated calculation regime has been regarded as a welcome change, as is China's decentralization reform (Holz, 2005). However, the separated calculation approach, in other words, the systematic errors, cannot account for the long-standing marked discrepancies alone,[2] which has even increased in recent years, as shown in Fig. 1. Seeking to identify the fundamental causes, some experts, such as Pan and An (2003), Holz (2005) and Koch-Weser (2013), realized that institutional arrangements could play an important role in explaining the discrepancy. However, these studies took the institutional perspective more broadly, without specifying incentives or constraints with regard to the sectoral discrepancies of added values and energy consumption, and thus failed to reveal the underlying causes of the data inconsistencies.

To identify the fundamental sources, this paper first constructed index decomposition models for GDP and total energy consumption to divide these overall discrepancies by sector, leaving no residuals. Then, it investigated several technical explanations, including statistical coverage, statistical methods and double-counting. From an institutional perspective, this study investigated incentives leading to falsely reported economic data, especially at the local (or provincial) level. Furthermore, the effectiveness of checks and balances was carefully examined by sector to reveal the institutional causes of sectoral data discrepancies comprehensively. The findings show that the industrial sector accounts for most of the overall value-added discrepancies. The discrepancy in total energy use mainly comes from raw coal consumption, within which the industrial sector also contributes the most. The results of the two cases, GDP and energy use, are fully compatible and support one another. Through careful examinations, however, it is determined that technical

[1] For instance, immediately after the release of GDP for the first three quarters of 2013, the Beijing News (*Xin Jing Bao*), a mainstream mass media, reported the GDP discrepancy between the country and its provinces, which had been widely forwarded by other mass media on the Internet.

[2] If systematic errors serve as the main causes, both downward- and upward-biased local figures would be expected. The biases of aggregate national data would be largely balanced; thus, the data discrepancy would not exhibit a statistically significant trend. However, the current data discrepancies in China in recent years are much larger than the acceptable upper bound caused by systematic errors, that is, 5% (Pan & An, 2003).