The quality of China's GDP statistics ☆

Carsten A. HOLZ *

Stanford Center for International Development, Stanford University, 366 Galvez Street, Stanford, CA 94305-6015, United States

A R T I C L E   I N F O

A B S T R A C T

Since the 1998 "wind of falsification and embellishment," Chinese official statistics on gross domestic product (GDP) have repeatedly come under scrutiny. This paper evaluates the quality of China's GDP statistics in four stages. First, it reviews past and ongoing suspicions of the quality of GDP data and examines the evidence. Second, it documents the institutional framework for data compilation and concludes on the implications for data quality. Third, it asks how the Chinese National Bureau of Statistics could possibly go about credibly falsifying GDP data without being found out. Fourth, it examines if the first- and second-digit distributions of official GDP data conform to established data regularities (Benford's Law). The findings are that the supposed evidence for GDP data falsification is not compelling, that the National Bureau of Statistics has much institutional scope for falsifying GDP data, and that certain manipulations of nominal and real data would be virtually undetectable. Official GDP data, however, exhibit few statistical anomalies (conform to Benford's Law) and the National Bureau of Statistics thus either makes no significant use of its scope to falsify data, or is aware of statistical data regularities when it falsifies data.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Chinese official statistics have a bad name. It doesn't seem to matter which Chinese statistics one looks at. There is the discrepancy between the official Beijing air quality measurements and the air quality measurements conducted by the U.S. Embassy in Beijing atop its embassy building: one day in October 2011, the Chinese official reading was "slightly polluted" while the U.S. reading was "crazy bad" (later revised to "beyond index").[1] China's premier Li Kejiang in 2007 called China's gross domestic product (GDP) figures "'man-made' and therefore unreliable," with all data other than electricity consumption, volume of rail cargo, and amount of loans disbursed being "for reference only."[2] Energy data came under suspicion in 2012 with a report that "local and provincial government officials have forced [power] plant managers not to report to Beijing the full extent of the slowdown."[3]

[1] See *Los Angeles Times*, 29 October 2011. The U.S. Embassy's readings are fed into twitter on an hourly basis (justified by diplomatic rules which require that information regarding health and security risks are being made available). The Chinese Foreign Ministry complained in July 2009 that "because the U.S. data conflicted with China's, they were causing 'confusion' and undesirable 'social consequences.'".
[2] *New York Times* (22 June 2012) and the *Wall Street Journal* (6 December 2010). The source is an American diplomatic cable released by WikiLeaks.
[3] *New York Times*, 22 June 2012.

What often goes unnoticed are the details: the U.S. Embassy tracks particles of smaller size than Beijing's official measurements do. The U.S. readings thus do not prove the Chinese data wrong. (The U.S. data might, however, be more meaningful for health matters than the Chinese data.) Premier Li Kejiang's statement on the reliability of Chinese statistics goes back to his time as Party Secretary of Liaoning province, and his evaluation is explicitly focused on Liaoning data. There is abundant evidence of data falsification at the *local* level, typically attributed to cadres' evaluation being linked to local economic performance.[4] Falsification of energy data in 2012 is questioned by a Western expert on Chinese energy data as hardly feasible at the five biggest electricity generation companies that together produce half of China's electricity, and while potentially possible at smaller producers such data would eventually have to be reconciled with all fuel, electricity, and financial accounts.

If anything, Li Kejiang's statement reveals the difficulty that China's National Bureau of Statistics (NBS) faces in publishing meaningful data. GDP data are by definition 'man-made' — the very fact that Li Kejiang criticizes. Value-added (VA) isn't apples growing on trees, to be counted and added up. For example, across the tertiary sector, the NBS in compiling national GDP has to rely on income approach data, and for many of these data, the NBS has to rely on sample surveys and imputations. Li Kejiang's preference for energy, rail cargo, and loan data presents a fallback to the planned economy era with output measured in physical units (such as tons and cubic feet) rather than in value terms, and with a focus on "material production."

The suspicions of Chinese data are a relatively recent phenomenon, starting in the late 1990s. To be sure, there were the data excesses of the "Great Leap Forward,"[5] and there have always been suspicions of some local "water content" and potentially (small) adjustments to national data,[6] but up through the 1980s researchers tended to issue a clean bill of health for Chinese statistics. Thus, in 1962, Li wrote cautiously about improvements following the statistical debacle during the "Great Leap Forward." Perkins concluded in 1966 that falsification of disaggregated data was highly improbable; in the case of aggregated data, falsification might have remained unnoticed in the short run, but not in the long run, and such falsification in the end would not have been in the interest of the Chinese leadership. In 1976, Rawski argued that "most foreign specialists now agree that statistical information published in Chinese sources provides a generally accurate and reliable foundation on which to base further investigations" (p. 440). In 1986, Chow wrote that "by and large Chinese statistics officials are honest" (p. 193).

This paper evaluates the quality of China's GDP statistics in four stages. First, it reviews past and ongoing suspicions of the quality of Chinese GDP data and examines the evidence. Second, it documents the institutional framework of data compilation in China and concludes on the implications for data quality. Third, it asks how the NBS could possibly go about credibly falsifying China's GDP data without running the danger of being detected, and evaluates the practical feasibility and likelihood of such data falsification. Fourth, it examines if the first- and second-digit distributions of official Chinese data conform to established data regularities (Benford's Law) or if they exhibit statistical anomalies.

2. Questions about Chinese data quality, and technical limitations

The problems in compiling accurate data in transition economies have been widely documented for Eastern Europe and Russia, but much less so for China.[7] Since the beginning of the economic reforms in 1978, economic transition, economic growth, and structural change have created severe challenges for the compilation of accurate statistics in China. These challenges include:

- Rapid growth in production units outside the traditional reporting system.
- Adoption of novel statistical concepts and variables with a transition from physical measures to value-based national income accounting that covers the complete economy.
- Increasing data falsification at local levels (supposedly because cadres' evaluation became linked to economic performance).
- Potentially declining interest of reporting units and relevant government departments to report to the statistics *xitong* (bureaucracy).[8]

While these challenges provide no rationale for data falsification at the national level, they make it potentially difficult for the NBS to compile accurate data. By the second half of the 1990s, researchers began to point out data inconsistencies.

---

[4] On local data falsification see Zhao (1998), Edin (1998), Cai (2000), or a NBS report of 14 June 2013 posted at http://www.stats.gov.cn/bgt/, accessed 14 August 2013. (I am grateful to Daniel Rosen for bringing the NBS report to my attention.).

[5] For a summary see, for example, Orlik, 2012, pp. 3f. For detailed accounts, see Yang (2012). (I am grateful to Andrew Walder for pointing out this latter source.)

[6] In 1989, the NBS Industry and Transportation Division asked how big the "water content" (*shuifen*), i.e., the degree of exaggeration, in industrial growth statistics was. The answer was "a little bit" (*yidian*), and the explanation for the exaggerated output statistics was predominantly a technical one focusing on compilation methods. Wang and Fewsmith (1995, p. 59) document falsification of the national economic growth rate in 1975 by the State Planning Commission, to which the NBS was directly subordinate at the time: two Politburo members on the evening of December 31, 1975, called the responsible officials of the State Planning Commission and demanded that the reported negative growth rate be immediately changed to a positive growth rate. The State Planning Commission then issued a positive growth rate of 0.056% and the *Renmin ribao* newspaper on the next day reported that the economy had grown in the previous year. (How anyone can know the growth rate of the current year's economy on the last day of the current year, down to the third decimal of a percentage point, is a mystery.) GDP, 1952–2004 (2007) (p. 5) reports a real GDP growth rate for 1975 of 8.7% (and a real GDP growth rate of negative 1.6% for 1976, which makes one wonder if Wang Lixin and Joseph Fewsmith mistook the year); a caveat: at the time, in 1975 (and 1976), the official economic growth figure was not measured using GDP.

[7] For the case of Eastern Europe and Russia, see, for example, Bartholdy (1997), Koen (1996), or Powers (1992). Xu (2000a, 2002) explains some of the continuing imperfections in China's GDP calculations.

[8] In the market (-oriented) economy, paying personnel to report to the statistics *xitong* is not in an enterprise's interest. Similarly, central government departments on which the NBS relies to collect statistical data in their *xitong* have little interest in collecting data on behalf of the NBS, paid for out of their own budgets.

2.1. The 1998 wind of falsification and embellishment

On 23 August 1995, Zhang Sai, then head of the NBS, in justifying revisions to the People's Republic of China Statistics Law explained that "recently, the phenomenon of false and deceptive reporting has spread in some localities and some units. The danger is large, the impact very negative." Reports on data falsification then became standard fare across the Chinese press and in the monthly NBS journal *Zhongguo tongji* (China Statistics). The Chinese slogan of *jiabao fukuafeng*, "wind of falsification and embellishment," soon made the rounds.[9]

The reports on data falsification came in two types. One type presented concrete evidence of data falsification, usually regarding one local enterprise or government. The figures were typically in the single-digit or lower double-digit percentage range, with no aggregate data for any geographic entity.[10] A second type of report remained very general, stating that data falsification takes place, but offering only the standard reasons for why it occurs (meeting growth or other targets, or achieving a higher evaluation of one's job performance).[11]

In 1998, the issue of data falsification came to a head, with first the Central Committee Discipline Commission and Organization Department and then the State Council throwing their weight behind NBS attempts to curb local data falsification. A campaign unfolded: circulars and reports were issued and investigative teams dispatched to provinces and central departments. At no point did the campaign target the NBS — the NBS spearheaded the campaign. By 2001, the wave of sensational reports on data falsification in *Zhongguo tongji* ceased.[12]

Researchers picked up on the 1998 wind of data falsification and embellishment.[13] Rawski (2001a) pieced together a national income measure and found real income growth in 1998 to be 5.7%, in contrast to the official real GDP growth rate of 7.8%. Official nominal GDP was later revised, but not the real growth rate, implying a larger implicit deflator. Rawski had applied the earlier (official) implicit deflator to the nominal income data. Applying the revised implicit deflator to Rawski's nominal income data yields a revised Rawskian real income growth rate of 7.2% compared to the official real GDP growth rate of 7.8%, a difference that seems well within the range of imperfections inherent in China's GDP calculations and in a by necessity incomplete and tentative income calculation.[14]

Researchers have also double-checked the growth rates of real GDP and industrial VA against the growth rates of energy use, physical output quantities of industrial products, and freight transportation.[15] In the late 1990s, the alternative variables grew at rates significantly below those of GDP or industrial VA. But data on these alternative variables are problematic. Chinese data on energy use and on the output quantities of individual industrial products cover only a subset of the economy, typically the directly reporting enterprises. Comparisons between 1997 and 1998 are especially dangerous because the subset of directly reporting enterprises was redefined in 1998. In the case of industry, the 1998 output value for the redefined group of directly reporting industrial enterprises (DRIEs) was lower than that of the (differently defined) group of DRIEs in 1997, and thus it is not astonishing if energy consumption and output of individual products—with coverage de facto limited to the DRIEs—fell. The use of freight data as a double-check for industrial VA (or even GDP) fares no better. According to official data, nationwide freight transportation stagnated at a time when petroleum consumption in the transportation sector was rising at double-digit growth rates, a reflection of the fact that the NBS has available data only on the decreasing share of freight transportation that occurs under the administration of the relevant government departments.[16]

In the end, neither a double-check via the income calculation of GDP nor a double-check via alternative variables yields compelling evidence of data falsification. Other double-checks confirm rather than question the official GDP real growth rates of the late 1990s, as Lardy (2002/2003) argues for the case of imports, tax revenues, and credit.[17]

---

[9] Rawski popularized the term "wind of falsification and embellishment" in the Western literature on Chinese statistics. For specific Chinese sources on data falsification and for their interpretation, see, for example, Rawski and Xiao (2001), and Rawski (2001a,b).

[10] See, for example, *Zhongguo Tongji*, no. 5, May 1998, p. 21, or no. 7, July 1998, p. 30. A more recent report (NBS, 14 June 2013. uanyu zhongshan shi henglan zhen zai tongji shang nongxuzuojia ji zhenggai qingkuang de tongbao [Report on the fraudulent statistics in Henglan Township, Zhongshan Municipality, and their rectification], http://www.stats.gov.cn/bgt/, accessed 14 August 2013) again covers very narrow ground: the four-fold exaggeration of GOV refers to 71 enterprises (out of 249) in the *township* (i.e., at very low administrative level) Henglan in Guangdong province.

[11] See, for example, *Zhongguo Tongji*, no. 11, November 1998, pp. 21f, or no. 6, June 1999, p. 22. Sloppy survey methods and poor quality of statistics personnel were also frequently cited; these do not constitute falsification as much as unreliable methods of data collection.

[12] The focus shifted to technical and definitional problems of Chinese statistics, where shortcomings are openly acknowledged. For example, an employee of the NBS Department for Statistical Development and Administration in November 2001 provided a comprehensive review of the shortcomings (but also achievements) of Chinese statistical practices in recent years and concluded that "China's economic growth rate has been questioned by some international organizations and individuals. I basically feel that our growth rate objectively reflects China's economic development, but when it comes to international standards and common methods, we still have quite some ways to go." (*Zhongguo Tongji*, no. 11, November 2001, pp. 8–10.) For details on the anti-falsification campaign see Holz (2003).

[13] See, for example, Adams and Chen (1996), Meng and Wang (2000), Wang and Meng (2001), Rawski (2001a,b), Rawski and Xiao (2001), *The Economist* (7 March 2001), the *Asian Wall Street Journal* (22 Nov. 2001), and an examination of the evidence in Holz (2003).

[14] Wu (2007) pointed to the 1998 growth rate again as a particular case. He investigated the 2006 benchmark revision to GDP data undertaken by the NBS following the 2004 economic census and noted that the NBS appears to have intentionally given the 1998 real GDP growth rate special treatment in the revisions by *not* adjusting it upward in line with the adjustments to the growth rates of other years.

[15] See, for example, Adams and Chen (1996) or Meng and Wang (2000).

[16] The comparison of freight transportation and petroleum consumption in the transportation sector is by Huenemann (2001). Energy use, output quantities, or freight transportation are not even variables to which the NBS accords particular importance among the thousands of variables on which it collects data. In Western economies, statistics on output quantities and freight transportation are often not available at all, while energy use may come with a range of caveats, including limitation to a subset of the economy.

[17] Movshuk (2002), however, questions Lardy's findings. He argues that part of the increase in (reported) imports is due to a clamp-down on smuggling starting in 1998, and part of the increase in government revenues is due to an anti-tax evasion campaign, also starting in 1998.

2.2. Censuses and benchmark revisions

The NBS regularly conducts censuses that may then trigger revisions to economic data. Since the beginning of the 1990s—the compilation of national income and product accounts data starts in the early 1990s—the following output-related censuses have been undertaken[18]

- 1993 tertiary sector census, followed by a benchmark revision of tertiary sector nominal VA and real growth rates of the years 1978–1993 (1993: nominal tertiary sector VA + 32.04%, nominal GDP + 9.99%).
- 1995 industrial census, followed by a benchmark revision of industrial gross output value (GOV) for 1991–1994 (downward by 6, 7, 9, and 10%), but no changes to industrial VA and GDP.[19]
- 1996 agricultural census, with no benchmark revision of earlier years' GDP data.
- 2004 economic census, followed by a 2006 benchmark revision of 1993–2004 nominal VA (GDP 2004: +16.8%) and real growth rates (and revisions to *1978–1992* tertiary sector values).
- 2008 economic census, followed by a 2010 benchmark revision of 2005–2008 nominal VA (GDP 2008: +4.4%) and real growth rates.

The benchmark revisions following the 1995 industrial census and the 2004 economic census are potentially problematic. The downward revisions to industrial GOV of 1991–1994 following the 1995 industrial census should have triggered corresponding revisions to VA. The fact that they didn't, implies that intermediate inputs were revised by exactly the same absolute amount, which is not plausible. But it is not clear if this is an issue of intentional falsification of industrial VA (and GDP), or perhaps an issue of problematic GOV data for non-state owned enterprises—where all the revisions happened—with the NBS possibly never having calculated their VA as the difference of GOV and intermediate inputs.

In the 2006 benchmark revision following the 2004 economic, the nominal VA of all sectors in 1993–2004 and the real growth rates of tertiary sector VA were revised, but not the real growth rates of the primary and secondary sectors. First, the retention of the original primary and secondary sector real growth rates is not plausible. Real growth rates of the secondary sector would remain constant if the 1993 nominal values had been revised by the same percentage as the 2004 values, but that is not the case. Industry and construction VA of 1993 changed by only +0.3 and −0.8%, but VA of 2004 changed by +3.8 and −9.2% (with a continuous trend in the years in between). The implication of not changing real growth rates is that the NBS raised the implicit deflator for industry and lowered that for construction. But the 2004 economic census collected no price data, and the NBS offered no explanations of why and how it revised the sectoral deflators.

Second, the NBS's retention of the unchanged pre-economic census *secondary sector* real growth rates is not plausible. The secondary sector real growth rate is a weighted average of the real growth rates of industry and construction. Retaining the pre-economic census secondary sector real growth rates implies that the NBS did not change the weights of industry and construction in the calculation of secondary sector real growth rates. This is despite the *increase* in nominal VA of industry and the *decrease* in nominal VA of construction.[20]

The upshot is that a number of alternative average annual real GDP growth rates can be calculated for the period 1992 through 2004:

- Based on first published real growth rates[21]: 9.2%.
- Based on most updated pre-economic census real growth rates (from the *Statistical Yearbook 2005*): 9.4%.
- Based on post-economic census real growth rates: 9.9%.
- Applying the *first published* implicit sectoral deflators of the primary, secondary and tertiary sectors to the post-economic census *revised* nominal sectoral values, and aggregating into real GDP growth using the Törnqvist index: 10.9%.

The last value is 1.7 percentage points higher than the annual average of the first published real growth rates, and 1.0 percentage points higher than the annual average of the revised (most updated) official real growth rate. In other words, post-economic census 2004, the official real GDP growth rate for the period 1992 through 2004 was revised upward by 0.7 percentage points per year, but it could equally well have been revised upward by more than twice as much.[22]

---

[18] For details on the various censuses and their impact on output measures see Holz (2013b). In 2007, a minor benchmark revision of 1978–2004 primary sector VA (upward) and tertiary sector VA (downward by the identical absolute amount) occurred as part of the adoption of a new sectoral classification system (Holz, 2013a).

[19] For details, see Holz and Lin (2001). GOV data of collective-owned enterprises, individual-owned and private enterprises, and enterprises in other (non-state) ownership for 1994 were retrospectively adjusted downward by 19%, 25%, and 16%, with similar or lesser revisions for the years 1991 through 1993. That data on state-owned enterprises (SOEs) were not retrospectively revised downward may be due to the fact that *all* SOEs were regularly reporting to the statistics *xitong*, whereas not all enterprises in non-state ownership did.

[20] These changes were sufficiently large to at least in some years change the real growth rate of the secondary sector, calculated with one decimal point and using a Törnqvist index or previous-year weights. The NBS can justify not revising secondary sector real growth rates by using a Laspeyres volume index with decade (since 2000: five-year) base years to calculate real growth rates, but the outcome is not economically plausible.

[21] The "first published" real growth rate of each year is the one published in the *Statistical Yearbook* of the following year (the first *Statistical Yearbook* issue to cover the year under consideration). Similarly, the first published implicit deflator is obtained from the two most recent annual nominal values and the most recent real growth rate published in the *Statistical Yearbook*.

[22] The much smaller 2010 benchmark revision of 2005–2008 data incorporated some changes to real growth rates but again included (now minor) changes to the implicit deflators. For the period 1992–2011, the average annual real GDP growth rate based on annual real growth rates as first published is 9.6%, based on annual real growth rates as most recently published 10.3%, and based on most recent nominal data and first published implicit deflators 11.1%.

2.3. Annual revisions

Every year, the NBS in the *Statistical Yearbook* provides a set of revised nominal GDP and sectoral VA for the last previously published annual data (and similarly for expenditures).[23] The NBS does not report deflators; an *implicit* deflator can be calculated by dividing the growth rate of the nominal values (1.X) by the real growth rate (1.Y). Real growth rates are typically not revised. Yet, in as far as the first published (implicit) deflators constitute the final estimates of the true deflators (details further below), they are applicable not only to the first published GDP data but also to the later revised nominal GDP data.

The first column of Table 1 reports the *annual* revisions to nominal VA, ignoring the 2006 and 2010 benchmark revisions of several earlier years' values. Apart from the tertiary sector census year 1993 and the economic census years 2004 and 2008 (when the annual revisions coincided with the benchmark revisions), the annual retrospective revisions to GDP in most years tend to be minor. The extremes are an upward correction in 1994 of 3.9% and a downward correction in 1998 of 1.3%.

The first column of real GDP growth rates in Table 1, for the years 1987–2011, reports real GDP growth rates as first published. The second column reports real GDP growth rates as most recently published prior to the 2006 benchmark revision, in the *Statistical Yearbook 2005*; these values incorporate the benchmark revision following the 1993 tertiary sector census. Throughout the early 1990s, first and later published real GDP growth rates differ by up to approximately one percentage point. Between 1995 and 2003, real growth rates were rarely revised. The 2006 and 2010 benchmark revisions (reflected in the *Statistical Yearbook* issue of *2012*, next column in Table 1) bring further growth rate revisions (for 1993–2004 and 2005–2008), never more than one percentage point per year, except in 2006 and 2007 with 1.1 and 2.3 percentage points.

The second-to-last column of Table 1 reports real GDP growth rates obtained by combining the first published implicit deflators with the most recently published nominal GDP values (incorporating both the 2006 and 2010 benchmark revisions), and the last column, proceeding similarly and with almost identical results, reports the Törnqvist index based on sectoral real growth rates. The annual revisions to nominal data consistently yield increases in real growth rates if one retains the first published implicit deflator (except for 1996).

The resulting (derived) real growth rates for 1992–1994 are substantially higher than the official first published or revised ones, with a maximum difference in 1992 of 4.4%. In the subsequent years, the difference can reach up to close to two percentage points, except in 2001 and 2007, when the derived real growth rates are 3.2 and 4.9 percentage points higher than the first published one. Real growth in 2007 is (an almost incredible) 16.8% if the first published implicit deflator is used, compared to a first published real growth rate of 11.9% that was later raised to 14.2%.[24] The NBS's handling of the implicit deflator clearly matters and is taken up in more detail below, in Section 4.2.[25]

2.4. Three approaches to calculating GDP

When the NBS announces official GDP values, it reports the production approach data (VA), supplemented by income data. The focus on the production approach—rather than aggregate expenditures, as in the U.S.—can be traced back to the socialist economy's Material Production System that shaped China's pre-reform statistical system. By definition, production approach GDP equals aggregate expenditures equals national income. How close the values come under real world circumstances can be an indicator of the quality of this set of statistics.

At the national level, China's nominal GDP can be up to five percent lower and up to three percent higher than nominal aggregate expenditures. The ratio of GDP to aggregate expenditures does not follow any particular pattern, except that there seems to be a tendency for GDP to over the years be marginally lower than aggregate expenditures (Fig. 1, using two different sets of data, one consisting of first published data, the second of data published in the *Statistical Yearbook 2012* and incorporating all benchmark revisions for all years).

Complete income data are only collected at the provincial level. Checking the years 1993 (the first year for which the *Statistical Yearbook* reported provincial expenditures and income), 1995, 2000, 2003, 2007, and 2010, gross regional product (VA) equals income across all provinces; these two approaches appear to not be conducted independently. Aggregate expenditure values differ (slightly) in about one-third of all provinces.

2.5. Summed provincial vs. national values

Theoretically, (national) GDP equals summed provincial VA (gross regional product). But starting 1997, coinciding with the wave of reports on local data falsification, summed provincial VA routinely exceeds GDP (Table 2). By 2004, summed provincial VA was 19.3% larger than (pre-economic census) GDP, with the largest discrepancy in the tertiary sector. Upward revisions to GDP and

---

[23] For example, the *Statistical Yearbook 2012*, published in September 2012, provides (final) revised nominal GDP data for 2010 together with preliminary GDP data for the latest year, 2011.

[24] The growth rate of (published) nominal GDP is 22.9%, and the CPI is 4.8% (*Statistical Yearbook 2012*, pp. 44, 311). The official real GDP growth rate of 14.2% comprises the following sectoral real growth rates (in percent): agriculture 3.7, industry 14.9, construction 16.2, and services 16.0, which in turn comprises transport, storage and post 11.8, wholesale and retail trade 20.2, hotels and catering services 9.6, financial intermediation 27.6, real estate 24.4, and "others" 11.3 (Statistical Yearbook 2012, pp. 47, 52).

[25] Sinclair (2012) examines the annual and the benchmark revisions of GDP data by the NBS and finds that these revisions generally fail tests for "well-behaved" data revisions — but so do revisions of U.S. data, which are generally considered the "gold standard." Sinclair concludes that the Chinese official data "can serve as a 'reliable guide' to the level and growth pattern of GDP."

Table 1
Annual revisions to national value-added, real GDP growth rates.

	Nominal: annual revised value/first published value						Real GDP growth rates (in %)				
							Official values			Constructed	Törnqvist index
							As first published	Statistical Yearbook			
	Total (GDP)	Primary sector	Secondary sector	# Ind.	# Constr.	Tertiary sector		2005	2012		
1987							10.4/11.1	11.6	11.6		10.9
1988							10.4/11.2	11.3	11.3		11.7
1989							3.5/4.3	4.1	4.1		3.3
1990							5.0/3.9	3.8	3.8		3.9
1991							7.8/8.0	9.2	9.2		12.4
1992							13.2	14.2	14.2	17.6	17.8
1993	1.104	1.035	1.011	1.000	1.085	1.335	13.4	13.5	14.0	15.5	15.8
1994	1.039	1.002	1.052	1.055	1.039	1.043	11.8	12.6	13.1	17.0	17.0
1995	1.004	1.000	1.013	1.015	1.000	0.992	10.5	10.5	10.9	11.6	11.4
1996	0.990	0.997	1.000	1.000	1.000	0.968	9.6	9.6	10.0	9.4	9.1
1997	0.996	1.017	1.012	1.021	0.959	0.958	8.8	8.8	9.3	9.6	9.1
1998	0.987	0.997	0.998	0.999	0.994	0.964	7.8	7.8	7.8	8.1	7.6
1999	1.002	1.001	1.004	1.003	1.005	1.000	7.1	7.1	7.6	8.9	8.6
2000	1.001	1.029	0.988	0.987	0.995	1.007	8.0	8.0	8.4	9.6	9.4
2001	1.014	1.055	0.993	0.995	0.987	1.028	7.3	7.5	8.3	10.5	10.6
2002	1.004	1.000	0.990	0.988	1.000	1.027	8.0	8.3	9.1	10.1	10.2
2003	1.001	0.990	1.000	1.000	1.000	1.008	9.3	9.5	10.0	10.7	10.9
2004	1.168	1.009	1.021	1.038	0.908	1.487	9.5	9.5	10.1	10.5	11.0
2005	1.004	1.000	1.004	1.004	1.000	1.006	10.2		11.3	11.3	11.4
2006	1.005	0.972	1.000	1.000	1.000	1.021	11.1		12.7	13.3	13.4
2007	1.031	1.019	1.028	1.030	1.018	1.038	11.9		14.2	16.8	16.9
2008	1.044	0.991	1.019	1.009	1.098	1.090	9.0		9.6	10.2	10.2
2009	1.001	1.000	1.000	1.000	1.000	1.003	9.1		9.2	9.2	9.2
2010	1.001	1.000	0.999	0.999	0.998	1.003	10.4		10.4	10.5	10.4
2011							9.3		9.3	9.3	9.3

All ratios of annual revisions through 2003 are obtained as national value published in the *Statistical Yearbook 2005* divided by the national value as first published in the then current year *Statistical Yearbook*. After 2003, all data are the second-most recent year's value in a *Statistical Yearbook* issue divided by the most recent value in the previous issue.

Year 1993 GDP was retrospectively revised upward largely because of a 33.45% upward revision to tertiary sector VA following the 1993 tertiary sector census. Tertiary sector VA was first revised in the *Statistical Yearbook 1995*, and then a second time (which is unusual), by a minor amount, in the *Statistical Yearbook 1996*. Similarly, 2004 and 2008 are economic census years.

As first published: individual issues of the *Statistical Yearbook* series. The starting year is 1987 because the *Statistical Yearbook 1988* (with most recent data for 1987) is the first issue with GDP data. For 1987–1991, GDP data are not available: the first figure represents the weighted average of the real growth rates of the three main economic sectors (using previous-year nominal values as weights), the second figure is the one published retrospectively in the *Statistical Yearbook 1993*, pp. 31f.

Constructed series: first published implicit deflator, *Statistical Yearbook 2012* nominal values. To compile the Törnqvist index, the nominal growth rates of the three main economic sectors are divided by the corresponding first published implicit deflators; the resulting sectoral real growth rates are then aggregated into real GDP growth rates using a Törnqvist index with the nominal shares as weights.

national sectoral VA following the 2004 economic census (GDP + 16.8%) and the 2008 economic census (GDP + 4.4%) lowered the discrepancy somewhat, but only temporarily.

A number of rationales have been advanced to explain the discrepancy. In 2004, the then NBS commissioner, Li Deshui, argued that provinces use 1990 base year prices when calculating industrial real growth, while the NBS makes adjustments to this procedure based on a price index (and starting in 2004 the NBS, for agriculture and industry, relies only on a price index); provinces double-count cross-provincial economic activities; provinces still use (presumably questionable) report forms for industrial enterprises with annual sales revenue below 5 million yuan (non-DRIEs); provinces use the opportunity of the as yet incomplete measurement of tertiary sector activities to adjust tertiary sector output upward such that the sectoral data add up to a desired aggregate output value; and provinces have incentives to exaggerate output (due to growth targets, comparisons of different localities by their output growth rates, and the use of statistics to measure local cadres' "achievements").[26]

To the extent that the NBS uses provincial data in deriving *national* GDP, how does the NBS adjust the provincial data? In February 2000, Liu Hong, the then NBS commissioner, offered two comments. The NBS contrasts provincial GDP data with key economic data obtained through sample surveys in each province. The NBS also has available data on variables related to GDP and assumes that the

---

[26] Some Chinese authors attribute the discrepancy to data falsification at lower tiers and to problems in calculating provincial-level GDP (let alone municipal- and county-level GDP). In the production approach (supplemented by income data), alleged local exaggeration in the tertiary sector is attributed to a lack of local data, with nationwide data presumably collected through centrally organized, nationwide sample surveys. See, for example, Pan and An (2003).
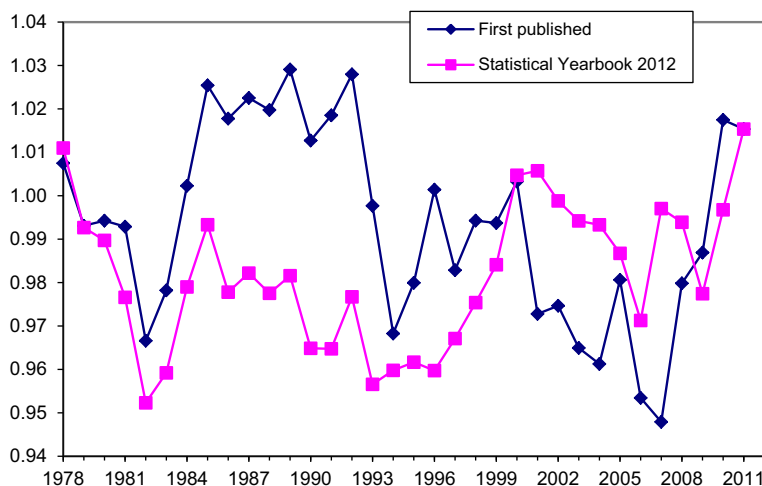
Fig. 1. Nominal GDP divided by aggregate expenditures.
Sources: first published values from *Statistical Yearbook 1995* (pp. 32, 36) for all years through 1994, and then each subsequent year's *Statistical Yearbook*. (Aggregate expenditure data became available in the *Statistical Yearbook 1995* only, with data from 1978 onward. This source does not report VA values of 1979 and 1981–83, but the *Statistical Yearbook 1998*, p. 55, does, with identical production approach values for 1978, 1980, and 1984 as in the *Statistical Yearbook 1995*).

Table 2
Summed provincial value-added divided by national GDP.

	Total (GDP)	Primary sector	Secondary sector	# Industry	# Construction	Tertiary sector
1991	0.969	1.002	0.982	0.977	1.024	0.918
1992	0.983	0.998	0.983	0.977	1.025	0.972
1993[a]	1.091	1.028	1.004	1.001	1.019	1.307
1994	1.013	0.982	1.013	1.024	0.946	1.019
1995	0.989	0.996	0.957	n.a.	n.a.	1.034
1996	1.000	1.004	0.950	0.960	0.886	1.077
1997	1.029	1.047	0.981	0.994	0.898	1.093
1998	1.043	1.018	0.997	0.998	0.994	1.123
1999	1.070	1.011	1.008	1.003	1.043	1.194
2000	1.087	1.044	1.007	0.998	1.065	1.232
2001	1.113	1.064	1.016	1.005	1.087	1.283
2002	1.126	1.006	1.038	1.023	1.139	1.316
2003	1.156	1.005	1.084	1.069	1.178	1.336
2004	1.193	1.006	1.139	1.128	1.213	1.370
2005	1.080	0.997	1.113	1.111	1.135	1.067
2006	1.096	1.000	1.121	1.123	1.103	1.093
2007	1.105	1.017	1.140	1.149	1.070	1.087
2008	1.088	0.993	1.142	1.154	1.048	1.051
2009	1.073	1.000	1.141	1.165	1.000	1.017
2010	1.089	1.000	1.173	1.202	1.001	1.019
2011	1.103	0.999	1.197	1.230	1.000	1.025

Data are first published values, in the *Statistical Yearbook* of the year following the year covered by the data, with the provincial data until (including) 1994 first published in the *Statistical Yearbook* of two years later. (1995 provincial values are taken from GDP, 1952–96, 1998.)

  [a] The 1993 provincial values are the first published values, available in the *Statistical Yearbook 1995*, incorporating the findings of the tertiary sector census of 1993. The 1993 national value is the first published value, available in the *Statistical Yearbook 1994*, *not* incorporating the findings of the tertiary sector census of 1993. (2004 and 2008 values are pre-economic census values, for provincial and national data.)

values of these variables cannot grow at a speed that is much different from that of GDP.[27] But the NBS is increasingly unlikely to make much use of provincial data and instead relies on the economic censuses, annual data from directly reporting units, and sample surveys.

2.6. Alternative data series

In 1994, the World Bank in its publications adjusted China's official 1992 per capita GDP upward by 34.3% and continued to adjust official Chinese data until in 1999 it accepted the official Chinese GDP data (presumably of 1998). If the World Bank's calculations

---

  [27] *China Infobank*, 29 February 2000. Li Deshui's 2004 reasoning on the differences suggests that, in addition, the NBS also considers a price index and survey data on small industrial enterprises, and controls against double-counting.

yielded a correct GDP value for 1992, then the disappearance of the discrepancy of 34.3% over six years implies that 5.04 percentage points of annual growth in China's nominal GDP as reported in official Chinese statistics is due solely to adjustments to China's GDP calculation practices. With half of the adjustments that the World Bank made to the 1992 GDP data not relevant for *real* GDP growth—18.3 of the 34.3 percentage point adjustment to 1992 GDP was due to the application of alternative prices—the adjustments account for just above two percentage points of annual real GDP growth. Xu (1999a,b), currently deputy-commissioner of the NBS (and head of the national income accounts division of the NBS), argues that not all corrections undertaken by the World Bank were justified.[28]

Another attempt to adjust Chinese GDP statistics was made by Angus Maddison in a 1998 OECD study in which he adjusted official Chinese GDP values of 1952–1995. His resulting average annual real GDP growth rate of 7.5% for the period 1978–1995 is 2.4 percentage points lower than the official value of 9.9% (Maddison, 1998, Table C.7, p. 160). The difference between Angus Maddison's real GDP growth rate and the official value is mainly due to two adjustments: the assumption of zero labor productivity growth in "non-productive services" (he then measures real VA growth of these services by employment growth), and the substitution of a measure of growth in product quantities for industrial real VA growth (using data from an early version of Wu, 2002). These two adjustments, on a variety of in the literature carefully documented grounds, are not tenable. The assumption of zero labor productivity growth in the service sector, for example, not only fails in appropriate cross-country comparisons but will also immediately strike as implausible anyone who has exchanged foreign currency into Renminbi in a Chinese bank in the early 1980s and then again in the mid-1990s. The substitution of the growth rate in industrial product quantities for real VA growth likewise fails in cross-country comparisons (Fig. 2) and is equally implausible given its assumption of no quality change over time (realistic for TV sets and cars?), no introduction of new, possibly high-growth products (DVD players, video cameras, computers, or cell phones), and the fact that product quantity data in China in that period cover a decreasing subset of the economy.[29]

2.7. Value-added of a subset of industry exceeds value-added of all of industry

Data on industrial VA are available for the DRIEs (in the industry statistics) and for all of industry (in the sectoral breakdown of GDP in the national income and product accounts). The share of DRIEs in the VA of all of industry fell from above 95% in 1979 to 75% in 1992 (with a problematic 1993 value) and to a low of 61% in 1997 (Fig. 3). After a statistical break in the definition of the DRIEs in 1997–1998, the DRIE share rose from 58% in 1998 to 87% in 2004, 99.7% in 2006, and 109.2% in 2007, after which publication of DRIE VA data ceased. The 2007 share, in excess of 100%, is logically not possible: a subset of industry cannot produce more VA than all of industry (unless other subsets produce negative VA, which is not plausible). In contrast, the employment data are plausible (Fig. 4). The share of DRIEs in industry employment fell from 0.67 in 1980 to 0.44 in 1998 and a low of 0.40 in 2001, before rising again, to 0.51 in 2007 and 0.57 in 2010. (The 2011 value is 0.54; the definition of the DRIEs changed in 2011.)

At least in 2007, either DRIE VA or VA of all of industry (and thus the GDP data)—or both—are problematic. With industry VA comprising DRIEs and non-DRIEs and little information available on the non-DRIEs, the focus in the following is on the DRIEs. DRIEs send a variety of monthly and annual report forms to the statistics *xitong* (NBS, 2003, pp. 215–224). These include data on VA and VA real growth (monthly), GOV (monthly, annual), GOV in constant prices (annual), and sales revenue (monthly, annual).

Data on DRIE VA could be problematic for three reasons: first, it is a national income accounting rather than an enterprise accounting concept and enterprises do not maintain an account labeled "value-added" (but can calculate VA as the difference of GOV and intermediate inputs, where each of these two items involves manipulations of several accounts).[30] Second, enterprises under pressure from local governments may misreport VA. Third, enterprises may choose to misreport data in order to protect their commercial interests. The NBS is in gross breach of Article 25 of the 2009 Statistics Law (NPC, 2009) which strictly prohibits the release of data that allows the identification of individual reporting units — yet complete datasets on the DRIEs for many years are now publicly available, with the datasets identifying firms by name and address and providing data on output, balance sheet, and profit and loss account items.[31]

The NBS can calculate VA from GOV (or, almost identically, sales revenue) in two ways.[32] First, it can obtain VA as the difference of GOV and intermediate inputs. However, the NBS is unlikely to have data on intermediate inputs. None of the dozen

---

[28] Xu reports the World Bank's original reasoning, questions it, and documents the changes in China's statistical system that took place between 1992 and 1998.

[29] For details, see Carsten Holz (2006a). Angus Maddison (2006) in a response disagrees with the criticism, and Carsten Holz (2006b) then disagrees with Angus Maddison's response. This argument also has implications for the Penn World Table data on China because in the Penn World Tables the official Chinese data are adjusted taking into account Angus Maddison's adjustments.

[30] Industrial enterprise accounting regulations do not include an account "value-added." Relevant for payment of the value-added tax, they include the accounts "taxes to be paid" (*ying jiao shuijin*) and "taxes to be subtracted" (*dai kou shuijin*, taxes paid on inputs that are not yet subtracted from taxes charged on outputs) in the balance sheet, and "commodity sales taxes and surcharges" (*chanpin xiaoshou shuijin ji fujia*) in the profit and loss account. (See Finance Ministry (General Office), 1999, pp. 418 ff.)

[31] Article 25 specifies that "no unit or individual may provide to the public, disclose, or use for purposes other than statistics, material obtained in statistical data collection from which one can identify or infer the identity of an individual reporting unit." Not only are complete annual datasets of the DRIEs for many years in the hands of numerous researchers, but also are available for purchase at www.taobao.com. While the public availability of confidential data collected by the NBS does not prove that it is the NBS which violated the Statistics Law, it is at least very likely: the formal reporting channels for enterprise data imply that only the NBS is in a position to create this dataset.

[32] Sales revenue data may be more reliable as this is an accounting concept and enterprises would seem to have no incentives to misreport sales revenue. Sales revenue plus net additions to inventories (another accounting concept) approximately equals GOV, with, in practice, a difference between sales revenue and GOV on the order of one percent. In detail, GOV equals sales revenue, plus self-produced self-utilized products and equipment, plus already completed but not yet delivered products, plus changes to the inventories of semi-finished products. (*Economics Dictionary: Statistics*, 1996, p. 168).
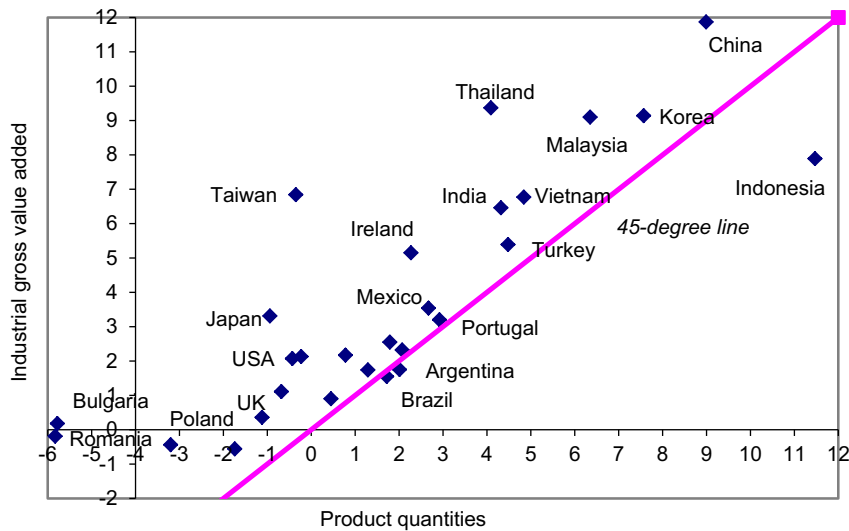
Fig. 2. Average annual growth rates of product quantities vs. real value-added. The period covered is 1978–1997. The countries constitute a selection of highly developed countries and of developing countries in Latin America, Europe, and Asia. All transition countries for which product quantity data are available in the two years (1978 and 1995) are included. The country selection was made prior to any calculations. The period through 1997 is chosen because Wu (2002), in an early version of which Maddison's work is based, conducts his analysis through 1997. The product quantity data are from the United Nations Industrial Commodity Production Statistics Database (http://unstats.un.org/unsd/snaama, accessed on 10 October 2004, with industrial VA of Taiwan in constant prices from http://www.stat.gov.tw/bs4/nis/enisd. htm, accessed on 9 February 2004).) For each country, all products on which data are available are used. The number of products varies from 17 in the case of Taiwan to 129 in the case of China and 380 in the case of Japan.
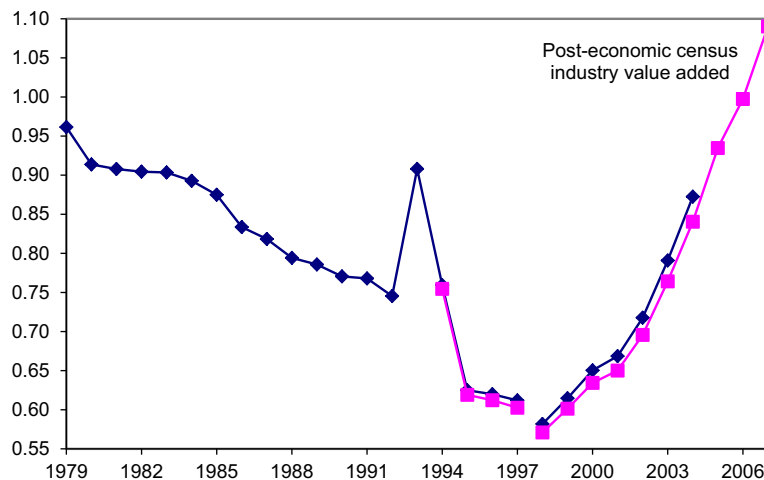

Fig. 3. DRIE share in value-added of industry. The chart uses two sets of data for industry VA in the national income and product accounts: the first set was published just before the results of the 2004 economic census were incorporated into the national income and product accounts, the second set is based on the (since 1993) following the 2004 economic census retrospectively (upward) revised industry VA. Coverage of the DRIEs changed in 1998. Data on VA of the DRIEs for the years prior to 1992 are not available. For 1980 and 1982–1991, the ratio of industry VA to industry net material product is applied to the net material product values of the DRIEs. For the years 1979 and 1981, when DRIE net material product values are not available, DRIE VA is derived based on the ratio of DRIE GOV to constructed DRIE VA in 1980 and 1982; for 1979 and 1981, the ratios of 1980 and 1982 are linearly interpolated and then multiplied with the 1979 and 1981 DRIE GOV values. (A 1978 DRIE GOV value is not available.) Linear interpolation is justified by the trend in the ratio during the period 1980 and 1982 through the 1990s. VA of the DRIEs is not revised following the 2004 economic census (the *Statistical Yearbook 2008*, p. 492, reports values for 1998–2004 that are unchanged from the previously published ones).
Sources: Industry VA: pre-economic census: *Statistical Yearbook 2005*, p. 51; post-economic census: *Statistical Yearbook 2008*, p. 37. Industry net material product (1978–1992): *Statistical Yearbook 1993*, p. 33. DRIE VA: *Statistical Yearbook 1993*, p. 417, *1994*, p. 378, *1995*, p. 388, *1996*, p. 414, *1997*, p. 424, *1998*, p. 444, *1999*, p. 432, *2000*, p. 414, *2001*, p. 410, *2002*, p. 432, *2003*, p. 468, *2005*, p. 488, *2007*, p. 508, *2008*, p. 492. DRIE net material product (1980, 1982–1992): *Statistical Yearbook 1984*, p. 216, *1986*, p. 278, *1987*, p. 263, *1988*, p. 320, *1989*, p. 292, *1990*, p. 419, *1991*, p. 399, *1992*, p. 411, *1993*, p. 417, *Industrial Yearbook 1986*, p. 21; DRIE GOV (1980, 1982–1992): *Seventeen Years*, p. 146.
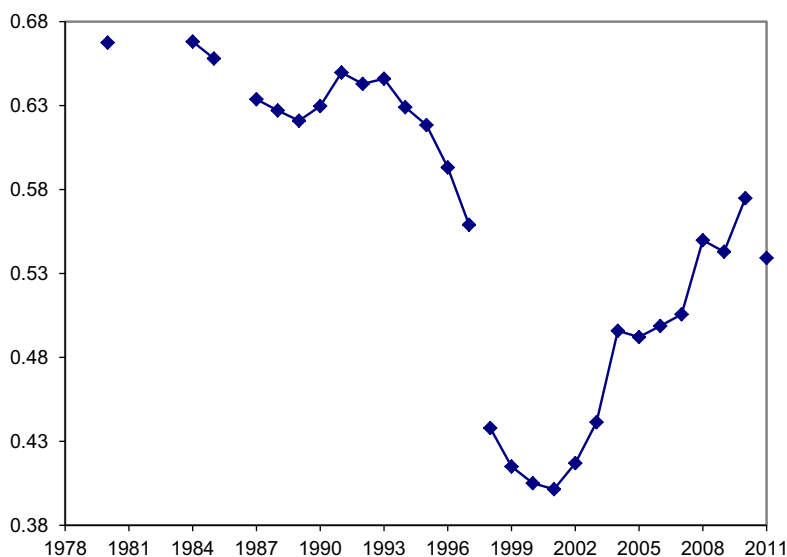
Fig. 4. DRIE share in industry employment. All values are mid-year values. Industry employment: secondary sector employment times the share of industry in secondary sector employment. Secondary sector employment values of the years prior to 1990 are adjusted (see Holz, 2013c). Values for the share of industry in secondary sector employment are available from the population censuses (1982, 1990, 2000, and 2010) and population surveys (1987, 1995, 2005). Share values of other years are interpolated; identical annual growth rates are assumed for all years between two adjacent census/survey dates. The census and survey values for the years through 1990 are of 1 July, and since then of 1 November of that year. The census value of 1 November 1995 was adjusted to turn it into a 1 July 1995 value. To obtain annual values for the years since 1996, the annual growth rate from 1 November in the first year of the interval to 1 November in the last year of the interval was applied to each year's 1 July value (in this interval). Employment values for the DRIEs are "average annual total employment" in 1994–2011, with the values for 1994–1997 obtained as DRIE VA divided by (published) DRIE labor productivity (which yields an identical value to the published average annual total employment value for 1998, the first year for which this particular employment series is published). For the years prior to 1994, DRIE employment values are "average annual staff and workers," in the one overlapping year 1993 identical up to a difference of 0.0006% to the "average annual total employment." Coverage of the DRIEs changed significantly in 1998 and in 2011.
Sources: Secondary sector employment: *Statistical Yearbook 2012*, p 128. Shares of industry in secondary sector employment: *Population Census* and *Population Survey* volumes of the corresponding years. DRIE employment: 1998–2011: *Statistical Yearbook 2012*, p. 510; 1994–1998: *Statistical Yearbook 1995*, pp. 385, 401, *1996*, pp. 411, 427, *1997*, pp. 414, 437, *1998*, 444, 459, *1999*, 432, 437 (labor productivity and VA); 1978–1994 (as available): *Industrial Statistical Yearbook 1993*, p. 90, *1994*, p. 81, *1995*, p. 79.
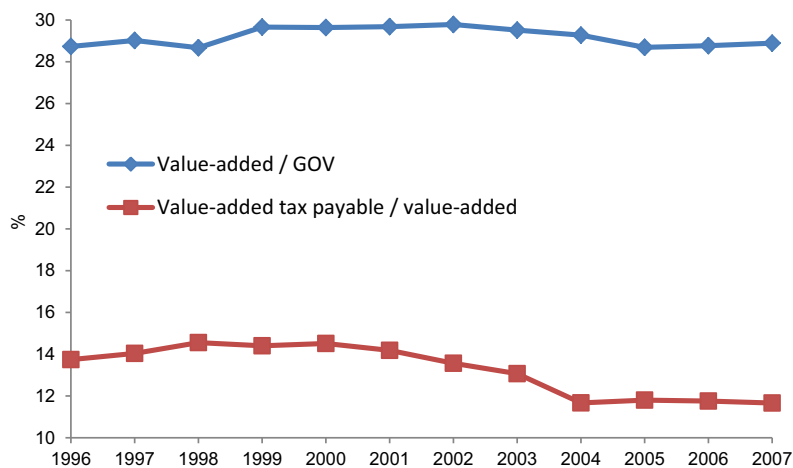

Fig. 5. DRIE value-added, value-added tax payable, GOV. Data on VAT payable start in 1996. Data on VA end in 2007.
Source: Each year's *Industrial Yearbook*, supplemented by the industry section in the *Statistical Yearbook* when needed.

report forms includes such an item.[33] Second, the NBS can multiply GOV by a fixed ratio to obtain VA — which it is known to do for the monthly DRIE data.[34] The NBS has held the ratio of VA to GOV constant over time at around 29% (Fig. 5). But China's economy is likely to be increasingly specializing, which would imply a fall in the ratio of VA to GOV, and thus a lower value of DRIE VA.

---

[33] Not even the 2004 economic census collected data on intermediate inputs (or VA) (NBS, 2007, pp. 113–5).
[34] Orlik (2012, p. 25) writes that monthly VA is obtained by multiplying GOV by the ratio of VA to GOV of the previous year.

In the input-output table of 1997, the ratio for the aggregate of all of industry is 30%, in 2007 it is 23% (*Input–output Tables 1997, 2007*).[35]

Another double-check is possible using data on the value-added tax (VAT) payable, available starting 1996. The VAT rate is 17% (with exemptions).[36] In 1998, VAT payable was equal to 14.6% of VA (the highest actual rate observed), before it fell to 11.7% by 2007 (Fig. 5).[37] If one accepts the absolute data on VAT payable as correct and assumes that its ratio to VA were *constant* at the (highest, 1998) 14.6% level through 2007 and adjusts 2007 DRIE VA correspondingly, the 2007 DRIE share in VA of all industry falls by about twenty percentage points ((11.7-14.6)/14.6) from its 109.2% level to approximately 90%.

A third piece of evidence comes from the 2008 economic census. Industry VA in the national income and product accounts covers all industrial enterprises (DRIEs and non-DRIEs) plus all industrial non-enterprises. In 2008, all industrial enterprises produced GOV of 53.4 trillion yuan (VA data are not available); DRIEs accounted for 50.7 trillion (94.9%). Data are also available for a *subset* of all non-enterprises, the licensed sole proprietorships, with the numbers of such proprietorships and of their employees. If employees in licensed sole proprietorships produced as much GOV per employee as the non-DRIEs did (which is about one-seventh the value of the DRIEs), then GOV of all industrial enterprises and all licensed sole proprietorships amounted to 54.6 trillion yuan, and the share of DRIEs in that new total would be 92.9%. A comparison of *employment* figures reveals that the aggregate of all industrial enterprises plus licensed sole proprietorships falls short of total industry employment by approximately one-sixth.[38] I.e., the DRIE share in the true total industrial GOV is lower than 92.9%.[39] The upshot of all these considerations is that the official DRIE VA figure of 2007 is probably exaggerated, and that the share of DRIEs in industry VA is likely to be around or below 90%.

2.8. Evaluation of official GDP compilation methods

NBS explanations of 1997 and 2002 on how VA is calculated, sector by sector, allow a subjective estimate of the potential margin of error in GDP data (Table 3). The first columns of the table show the extent to which the production and income approach are combined to derive official GDP. Only in agriculture is the production approach the sole approach used. In industry and construction the production approach is applied to one group of enterprises, namely the DRIEs, while the income approach, in part in combination with GOV calculations, is used for those enterprises not reporting directly to the statistics *xitong*. In the tertiary sector, the income approach dominates throughout. Across almost all sectors (or subsectors), ratios of VA to GOV or other standardizing variables are obtained from units that directly report to the statistics *xitong* (or other government departments), or from surveys, the tertiary sector census, or the input–output table, and then applied to the standardizing variable (typically GOV) obtained or estimated for those units that do not report directly.

Table 3 includes subjective judgments of the quality of VA data compiled in each sector (subsector) based on VA data for 2000, the year for which the official explanations on data compilation are likely to provide the closest match. The directly reported data are judged as being of highest quality. (The previous section suggests that VA of the DRIEs *by 2007* has become problematic.) Data obtained through various approximations or via unreliable institutions are categorized either as "somewhat" reliable or unreliable. Of total GDP in China in 2000, only 45.03% is likely to be highly reliable, while a second part of 11.07% is somewhat reliable, and the remainder, namely 43.90%, is of poor quality.[40]

If unreliable data were up to one third too high or too low, this implies an approximately 15% margin of error in the GDP data. As long as no clear bias is involved, errors could cancel out. In 2000, Xu (2000a) suggested that the agricultural and industrial portion are likely to be systematically upward biased, while the real estate portion is likely to be systematically and gravely downward biased (due to imperfect imputations of the service value of owner-occupied housing). The benchmark revision following the 2004 economic census confirmed the downward bias in the earlier published real estate VA (but virtually no bias in agricultural and industrial VA). Even if there are clear measurement biases, as long as these biases are consistent over time, growth rates could still be quite accurate, with a margin of error of at most around 1 percentage point (the difference between using a 15% inflated vs. non-inflated GDP value in both years multiplied by an annual growth rate of about 8%).

---

[35] Economy-wide, the ratio of VA to GOV fell from 41.5% in 1990 to 36% in 2000 and 32.2% in 2010 (*Zhongguo tongji*, no. 8, August 2013, p. 8).

[36] Small enterprises are subject to different tax arrangements, which should not affect the bulk of the (by definition relatively large) DRIEs.

[37] VA includes VAT. Thus, a ratio of VAT to VA of 0.146 (the 1998 value) implies a VAT rate (VAT relative to VA net of taxes) of 17.1%, obtained as 0.146/(1–0.146).

[38] In 2008 (following the economic census), employment in the DRIEs was 88.3763 million, in all industrial enterprises 120.0731 million, in the licensed sole proprietorships 14.0269 million, and in all industry (based on *Statistical Yearbook* secondary sector employment figures, adjusted to cover only industry) 160.7491 million. All employment values are mid-year values.

[39] A fourth piece of evidence is that NBS commissioner Ma Jiantang estimates that the 700,000 "three-above" enterprises account for "90% or more" of industrial GOV, VA, and sales revenue. See http://www.stats.gov.cn/tjdt/gjtjjdt/t20120217_402785065.htm, accessed 11 November 2013, with a middle element of /tjdt./ on 27 November 2013. The "three-above" enterprises denotes above-norm enterprises in industry (i.e., the DRIEs), in trade, and in construction, plus, in addition, for more correctly a "three-above plus one" set of enterprises, all real estate development enterprises.

[40] The subjective evaluation can be updated to other years. If DRIE data have become less reliable over the years, then the share of poor quality data may well have increased since 2000.

Table 3

Reliability of GDP data compilation methods (2000).

	Two groups?[a] ("Direct" vs. "indirect")	Approach to calculating value-added — Production approach	Income approach	Need data from "Direct" for "indirect" group?	Data sources other than direct reporting (where mentioned)	Share of this sector in GDP, year 2000, in %	Share in % of sector (sub-sector) value-added that is[d] — Highly reliable	Somewhat reliable	Unreliable
Farming, forestry, husbandry, fisheries		Yes				16.35			100.00
Industry	Yes	For direct	For indirect[c]		Survey	43.66	66.49		33.51
Construction	Yes	(Yes[b])	(Yes[b])	Yes		6.58	56.74	43.26	
Tertiary sector						33.41			
Transport (and storage), post and telecommunications	Yes	Theoretically[b]	For direct; indirect[c]	Both approaches		(6.05)	60.00		40.00
Wholesale/retail trade, catering	Yes	(Yes)	For direct; indirect[c]	Production approach	Survey, tertiary census, I/O table	(8.18)	48.62		51.38
Banking and insurance						(5.83)	80.00	20.00	
Banking	Yes	(Yes)	Yes[c]	Both appr.					
Insurance		(Yes)	Yes						
Real estate		(Yes[b])	Yes		Survey	(1.86)			100.00
Government, Party, social organizations			Yes		I/O table	(2.62)		100.00	
All other services			Yes		Tertiary census, I/O table	(8.87)		50.00	50.00
Services for farming, etc.						[0.26]			
Geological prospecting, water cons.						[0.37]			
Social services						[3.63]			
Health care, sports, social welfare						[0.92]			
Educ., culture, arts, radio, film, TV						[2.67]			
Scientific research, polytechnic serv.						[0.70]			
Others						[0.31]			
Total: share in GDP, 2000, in %							45.03	11.07	43.90

A "Yes" *in parenthesis* means that this approach is not the main approach. The 2000 (pre-economic census) GDP value is 8944.22 billion yuan. *Statistical Yearbook 2002*, p. 51.

Sources: Xu (2000b), NBS (1997). *Statistical Yearbook 2001*, pp. 410, 465f., 551, 559; *2002*, pp. 51, 55.

  [a] Are the production units in this sector (or subsector) split into two groups, i.e., units reporting directly to the statistics *xitong* or some other government department vs. units not reporting directly? The directly reporting units carry labels such as "above-norm enterprises" or "designated units," with the definitions being sector-specific.

  [b] The source describes how the production approach is used in this sector (subsector), but describes only the calculation of GOV. It then mentions that VA is obtained through the income approach, and GOV minus income approach VA yields the value of intermediate inputs. In the case of construction, VA of those units not reporting directly to the statistics *xitong* are based on GOV and investment statistics combined with VA ratios obtained from the group of directly reporting units (where VA was obtained using the income approach).

  [c] In the calculation of the VA of the "indirect" group in industry, income approach VA is obtained through sampling, with the ratio of (income approach) VA to GOV in the sample then applied to the GOV "estimated" for the whole group of enterprises not reporting directly to the statistics *xitong* or government departments. In the case of transport, post and telecommunications, for some units within the "indirect" group the income approach is supported with ratios obtained from the "direct" group, while for other units, GOV is obtained and then multiplied by the ratio of VA to GOV in the "direct" group. In wholesale/retail trade and catering, a ratio of VA to GOV is obtained from the last tertiary sector census or the latest input–output table and then applied to the GOV of the "indirect" group. For approximately 80% of the banking sector (measured in terms of loans), data on income approach VA are available; ratios obtained from this group are then applied to total loans.

  [d] The categorization of data as highly reliable, somewhat reliable, or unreliable is subjective. In industry and construction, the split into two groups is based on the share of sectoral VA produced in directly reporting (above-norm) units. No data on transportation, post and telecommunications occurring outside the *xitong* are available (apart from aggregate sectoral VA); in waterway and highway transportation this share may be high. In wholesale/retail trade and catering, the split into two groups was based on the share of sales units "above designated size" (directly reporting units). In banking and insurance, the split into two groups was based on the share in total loans of those banks on which the NBS has balance sheet and profit and loss account data available (to derive income approach VA).

2.9. Summary

While the 1998 campaign against the "wind of falsification and embellishment" triggered a long-lasting suspicion of the quality of Chinese official statistics, attempts to show (national-level) data falsification in the late 1990s never succeeded. Annual revisions as well as benchmark revisions following censuses appear plausible, with few exceptions: (i) the NBS seems reluctant to retrospectively revise real growth rates and rather accepts—typically upward—revisions to the implicit deflator, and (ii) the 2007 VA real growth rate of 14.2 to 16.8% appears implausibly high. Different approaches to calculating GDP yield no striking discrepancies, while the discrepancies between summed provincial VA and national GDP are well known and have no further effect on the national data. The World Bank in 1999 abandoned its alternative data series of six years, and Angus Maddison's alternative series are based on indefensible assumptions. By the second half of the 2000s, DRIE VA data became openly untenable, suggesting great difficulties in collecting accurate data from basic statistical units. A subjective evaluation of the NBS's data compilation methods suggests that nominal GDP may be quite inaccurate, but as long as it is consistently so over time will have little impact on real growth rates. Overall, there is no evidence of blatant data falsification, neither in form of a significant one-time event nor in form of a systematic bias.[41]

3. Institutional scope for data falsification

Official statistics in China are compiled by the NBS and by other institutions with approval of the NBS. The Statistics Law of 1996 and its revision of 2009 assign responsibility for organizing, directing and coordinating statistics work throughout the country to the NBS (NPC, 1996, Art. 4; and NPC (National People's Congress), 2009, Art. 27). The NBS collects data through direct reporting, surveys, and censuses. It also receives data from approximately one hundred other institutions, including other government departments such as the People's Bank of China with financial sector data, the Finance Ministry with fiscal sector data, and the Customs General Administration with foreign trade data.

The NBS is a bureau directly under the administrative leadership of the State Council, which also appoints major personnel and provides funding. The NBS has very little authority over provincial statistics bureaus (business leadership) or over the statistics divisions of other central government departments (business guidance).[42] The NBS has direct control only over its survey teams.

When the NBS relies on other government departments for the collection of some statistics, the choice, quality and coverage of these statistics are dictated by the relevant government departments' own data needs and their departmental reach. Some of these data are limited in scope. For example, in the transport and communications sector, GOV data are collected by the Railway Ministry (Bureau), the Communications (Transportation) Ministry (Bureau), the Aviation Bureau, the Post and Telecommunications Ministry (Bureau), the Township Enterprise Bureau of the Agriculture Ministry, and the Industry and Transport Department within the Finance Ministry — covering enterprises and other units under their jurisdiction. For the compilation of the sectoral VA, the NBS has to guesstimate GOV of road and water transportation by transport and communications enterprises/units that are not operating under these departments.

The low rank of the NBS, half a rank below that of a ministry, may prove a hindrance to the collection of accurate data. For example, the NBS appears to make little use of tax bureau data. The tax bureau belongs to a different *xitong*. Data from the tax bureau could be particularly helpful in the case of small enterprises on which the NBS has little reliable data — should the tax data themselves be accurate.[43]

The NBS commissioner Ma Jiantang is simultaneously the Party secretary of the NBS. The first deputy commissioner of the NBS is also the (one) deputy-secretary of the Party. One of the five regular Party cell members heads the Party Disciplinary Commission within the NBS, while three further members serve as three of the five deputy-commissioners.[44] I.e., there is a near-perfect congruence between the Party leadership and the bureaucratic leadership within the NBS, with only one Party cell member not simultaneously a deputy-commissioner, and only one deputy-commissioner not simultaneously a Party cell member. The commissioner is typically a political appointee, i.e., someone not trained for a career in statistics work.[45]

A notion of formal independence of the NBS can be found neither in regulations nor in the statements of individuals. The NBS is a typical government department under direct control of government and Party. It primarily serves government and Party leaders. A NBS "work regulation" of 16 November 1995 explicitly states that the NBS is to implement "important decisions and instructions of

---

[41] An online appendix (Appendix B, at http://ihome.ust.hk/~socholz/QualityGDP.html) examines the coverage of military value-added, which could be underestimated. As long as any mis-estimation of military value-added occurs consistently over time, it should have virtually no impact on real growth rates of GDP.

[42] Provincial statistics personnel are appointed by the provincial Party Committee and government, and statistics personnel in other central government departments are appointed by those departments. Central government budgetary funds are intended to finance the operations of the NBS as well as to support statistical work on lower tiers. In principle, local statistics departments receive central funding to cover work-related costs (*shiye bianzhi de jingfei*), while local funding covers basic administrative costs (*xingzheng jingfei*) — the NBS thus may have some leverage over local data compilation practices through its control over part of the funding.

[43] NBS access to sales tax data could provide a convenient double-check on the retail sales data collected from the commercial system. The quality of China's retail sales data, which, for example, figure prominently in the calculation of consumption in the expenditure approach to GDP, is questionable (Holz, 2004). As of the 2000s, the NBS appears to receive data on tax payments by individual-owned units from the State Tax Administration.

[44] See http://www.stats.gov.cn/jgzn/2006_lingdao.htm, accessed 28 March 2013.

[45] The current (2013) commissioner of the NBS, Ma Jiantang, in this position since September 2008, previously held Party and government positions in Qinghai Province. Immediately prior to assuming the position of NBS commissioner, he was the deputy Party secretary and deputy governor of Qinghai Province. He earlier worked in the State Council Development Research Office and in the State Economic and Trade Commission.

the Chinese Communist Party Central Committee and the State Council." Zhang (2001), NBS commissioner from 1984 through 1997, clarified the tasks and the responsibilities of the statistics *xitong*: "the government statistics organization primarily serves the needs of macroeconomic decision-making of Party and government leaders at each administrative level, and is responsible to the Party and government leaders at each administrative level" (p. 319). The NBS commissioner apparently sees no need to include the public in the list of who the NBS serves, and responsibility is to leaders rather than to institutions such as the parliament of China.

The Statistics Law lists as the "fundamental task of statistics work" to conduct statistical examination of the implementation of the national economic and social development plan, to analyze the statistics, to provide statistics material and statistics advice and suggestions, and to supervise through the use of statistics (NPC, 1996, Art. 2, and similarly for the revised Statistics Law in NPC, 2009, Art. 2). I.e., statistics work serves primarily a policy and supervision purpose. Consequently, the National Development and Reform Commission has access to NBS data before publication of the data.[46] At the local level, key data compiled by local statistics departments, such as GDP data, need approval by a local government leader before they can be reported up to the next higher-level statistics department.[47]

The Statistics Law warns against data falsification, but omits any reference to central leaders: "statistics personnel must seek truth from the facts, strictly abide by professional standards [*daode*], and have the necessary professional knowledge that qualifies them to do statistics work," and "the leaders of localities, government departments, or other units may not order or suggest to statistics departments and statistics personnel to tamper with statistics material or to compile false data" (NPC, 1996, Art. 24 and 7, and in slightly different form in NPC, 2009, Art. 6 and 29). If a local government leader were to, in violation of the Statistics Law, "request" higher economic growth rates, the Statistics Law of 1996 requires the local statistics department to—quite unrealistically—refuse to cooperate with its immediate superior (the local government leader), and there the matter ends (NPC, 1996, Art. 7). The Statistics Law of 2009 (NPC, 2009) does not contain such a requirement but has a much expanded sixth section on legal duties that specifies disciplinary procedures and pecuniary penalties when the law is violated. The appointing institution or the supervisory institution is to take disciplinary action "in accordance with the law" (Art 37). How information on prohibited behavior is to reach the appointing institution or the supervisory institution, or what is to happen if the appointing institution or the supervisory institution is the originator of the violation of the Statistics Law, is not specified.

The NBS, in cooperation with the Ministry of Supervision and the Bureau of Legislative Affairs of the State Council, conducted nationwide inspections on the implementation of statistics regulations in 1987, 1989, 1994, 1997, and 2001 (and not since).[48] In the 2001 inspection, misreporting by enterprises accounted for almost 60% of the 60,000 violations of the Statistics Law that were revealed, with other violations consisting of refusal by enterprises to report data or late reporting, but not of misbehavior by statistics (or other government) officials.[49] No information about violations within the statistics *xitong* is available, even though statistics work of government departments was also investigated.

Three conclusions emerge. First, the regulatory framework to ensure accurate reporting by statistics units to the statistics *xitong* and in the final instance to the NBS appears extraordinarily weak, as does enforcement of this regulatory framework. Second, much of data compilation is outside the control of the NBS. The NBS has no authority over data collection by other central government departments and by provincial statistics bureaus. Even in the case of censuses, while the NBS may design the censuses, implementation requires involvement by a slate of other government departments. The only data over which the NBS has a considerable amount of control are its own survey data. Third, the NBS has no formal degree of independence of the central Party/state apparatus, and serves individuals in leadership positions rather than institutions.

The implications are two-fold. First, with no reliable, comprehensive framework for the collection of annual data in place, the NBS in deriving annual GDP data will have to improvise. One option could be to base its annual GDP data on economic census values extrapolated to non-economic census years by relying on the annual change in the values of various reported data, and on surveys. Perhaps understandably, the NBS does not publish sufficient data to allow a researcher to retrace its GDP calculations.

Second, the final, official GDP values may be rather haphazard values. Not only are there large technical limitations to data quality, but the NBS also has no institutional independence. How the final GDP values are obtained is shrouded in secrecy. In all likelihood, the choices that lead to the final official GDP values are known only to a very small number of people in the NBS. Quite possibly, these choices are made by only a handful of people, such as the Party cell or the near-identical leadership group of NBS commissioner and deputy-commissioners, in an environment of implicit or even explicit expectations raised by top leaders.

---

[46] See the International Monetary Fund's General Data Dissemination Standard (GDDS) metadata homepages (at http://dsbb.imf.org/Applications/web/gdds/gddscountrycategorydiapreport/?strcode=CHN&strcat=175, accessed 16 April 2004). Access to the data produced and disseminated by the NBS before release is provided to senior officials of the State Development and Planning Commission, the predecessor to the National Development and Reform Commission, as well as to the (by now defunct and partly integrated into the State Development and Reform Commission) State Economic and Trade Commission.

[47] Pan and An (2003, p. 8.)

[48] Nowadays, the NBS provides a telephone number and an e-mail address on its homepage (http://www.stats.gov.cn) to which violations of statistics laws and regulations can be reported.

[49] For information on inspections, see the China webpages in the International Monetary Fund's GDDS (http://dsbb.imf.org/Applications/web/gdds/gddshome/) under the heading "Table C. Data Integrity and Access by the Public," accessed 16 April 2004. As of April 2008, this information is no longer available at the GDDS website. The NBS website carries no up-to-date report on inspections (though a number of local reports). For details on the 2001 inspection, see also http://www.stats.gov.cn/xwkj/tjdt/20010523002.htm, accessed 8 Feb. 2002 and as of 2013 no longer available.

4. How to credibly falsify China's official GDP statistics

If the NBS wanted to publish falsified GDP statistics—perhaps under Party pressure to produce sufficiently high, or at times of overheating sufficiently low figures—what kind of falsification could possibly remain undetected?

There is no evidence of two sets of data, an internal set of correct data and a published one of falsified data. Various internal publications that are occasionally available to researchers largely match the published data.[50] If the NBS does not maintain two sets of data, falsification of GDP data has to rely on manipulation of those data that are not publicly available, or on not specifying (problematic) choices made when aggregating data.

4.1. Scope for falsification of nominal GDP

Data falsification is easily accomplished when a published datapoint is based on published data *plus* unpublished data. For example, industrial VA (and thereby nominal GDP) is based on published data such as DRIE output *plus* output of non-DRIEs, with no data published on the non-DRIEs. The same pattern applies to construction and other sectors. It is the unpublished data on the units that do not report directly to the statistics *xitong* that can be manipulated. To abstract from sample survey data to the complete set of all units that do not report directly to the statistics *xitong* requires assumptions about the total number of such units, the representativeness of the survey, and the appropriate method to translate whatever data can be collected, such as sales revenue data, into VA data. The NBS can freely vary its (unpublished) assumptions, and a wide range of values may be equally valid. Since 2008, even the nominal VA of some of the directly reporting units—those in industry—are subject to (unpublished) manipulations, with the NBS since 2008 limiting its publication of industrial VA to an aggregate figure for all of industry.

Double-checks on the size of that part of the economy on which the NBS does not publish individual data appear impossible. Indicators such as energy, product quantities, or freight transportation do not allow a double-check on total industrial output because the coverage of these indicators is limited to a subset of the economy (such as the DRIEs). The economic censuses and population censuses offer a potential double-check via employment (as explored above in Section 2.7), but the employment data are hard to interpret: for example, the 2000 population census counts as employed anyone who has worked for one hour or more in the previous week — much of employment especially in the units on which no output data are available could be part-time. Furthermore, such employees may have available very little physical capital and labor productivity thus may be far lower than in the DRIEs, making informed output guesstimates difficult if not impossible.

The NBS's unpublished data that find their way into GDP are subject not only to NBS manipulations and NBS technical limitations, but also to attempts by economic units to escape recording. The larger issue at hand is the "shadow economy:" the shadow economy comprises illegal economic activities, unreported economic activities, and other economic activities not captured in official statistics (typically because these activities are too small-scale and too dispersed).[51] The size of China's shadow economy matters for two reasons. First, to the extent that the shadow economy is not included in GDP, official GDP underestimates actual GDP. I.e., the current year's GDP value may be inaccurate. Second, the larger the shadow economy, the greater the potential for changes to the official GDP value due simply to a changing degree to which the NBS succeeds in capturing all economic activities. This also affects the accuracy of historical data trends.

Across OECD countries, official GDP is estimated to underreport actual GDP by 8 to 30%, with a low of 8–10% for countries such as the U.S. and Japan, and a high of 24–30% for countries such as Greece, Italy and Spain (Schneider & Enste, 2000). According to a "high-level official" in the NBS in 2003, the NBS may miss economic activities equivalent to approximately 10% of GDP.[52] The deputy editor of the Hongqi Publishing House, Huang Weiting, in 2003 estimated the missing economic activities to be equivalent to approximately 15–20%, but "in recent years this may already have come down a bit, to approximately 15%." Various studies find that the peak of the underground economy occurred around 1990, with the VA of the underground economy equivalent to 20–25% of official GDP, then dropping to single-digit percentage levels by the late 1990s.[53] These studies require a number of strong assumptions and thus at best provide a rough indication of the size of the shadow economy and of its development over time.[54]

What if the benchmark revisions simply reflected newly captured economic activities (previously part of the shadow economy) thanks to improved NBS data compilation methods? It is officially acknowledged that the 2006 benchmark revision following the

---

[50] Data presented in the internal monthly publication *Shisan sheng shi guomin jingji zhuyao tongji zhibiao* (Major statistical indicators on the economy of thirteen provinces and cities) is consistent with the published data. Rawski (1976, 2001) similarly sees no evidence for the existence of internal statistics that contradict the published data. Sub-sets of the published data add up to the total, across sectors and/or provinces (and sectors and/or municipalities within provinces), except in clearly identifiable instances with plausible rationales — summed provincial gross regional product not being equal to GDP, or DRIE VA exceeding industry VA in 2007.

[51] For this definition of the "shadow economy" see Xia and Wan (1993), who claim that all three elements are relevant for China. The NBS does not consider illegal activities in its calculation of GDP, but attempts to capture all other economic activities as much as possible (Xu, 2001, p. 64). In a survey of the literature on the shadow economy, Schneider and Enste (2000) offer a "commonly used working definition" of the shadow economy as "all economic activities that contribute to the officially calculated (or observed) gross national product but are currently unregistered" (p. 78).

[52] In 2003, the NBS conducted a (unpublished) study to estimate the extent of the underground economy.

[53] See, for example, China Infobank (7 June 2003), Zhang, Ding, and Dan (2004), and Guo and Yi (2006). One study, starting with the year 1994, arrives at an underground economy equivalent to a more extreme 36% of GDP in 1994, followed by a gradual decline to eight percent in 2003 (Liu, Zhang, & Liao, 2007).

[54] For example, they require the assumption of no shadow economy in 1978, and often the assumption that the amount of currency in circulation can serve as indicator of the size of the shadow economy.

2004 economic census indeed extended the coverage of GDP to include (i) economic activities previously ignored, such as those occurring in subordinate units of an enterprise and outside the main business of the enterprise, and (ii) economic activities captured through statistical compilations outside the economic census (and previously not included in GDP), such as homeowners renting out housing, home teaching, or childcare services.[55] One may wonder to what extent these innovations should equally affect all earlier years, rather than the NBS fully accounting for these changes only in 2004, and reducing their impact to zero in 1992 (or 1978, in case of the tertiary sector).

The size of benchmark revisions may give an indication of the potential degree of historical mis-estimation of GDP:

- Following the 1993 tertiary sector census, tertiary sector VA of 1993 was revised upward by 32.04% (an increase in 1993 GDP by 9.99%). The values of earlier years starting 1978 were also revised, but tertiary sector VA of 1978 increased by only 4.37% (and thereby 1978 GDP by 1.00%).
- Following the 2004 economic census, tertiary sector VA of 2004 was revised upward by 48.71% and GDP by 16.81%. The values of earlier years starting from 1993 were also revised, but tertiary sector VA of 1993 increased by only 5.91% (and GDP by 2.02%).
- Following the 2008 economic census, tertiary sector VA of 2008 was revised upward by 9.01% and GDP by 4.45%, with lesser revisions for the years 2005–2008; 2004 values were not revised.[56]

What if these newly found economic activities had existed in equal proportion to official tertiary sector VA and GDP in earlier years, but were captured by the NBS only in the censuses of 1993, 2004, and 2008?[57] In this case, the 1993 tertiary sector census should have led to a 32.04% upward revision in tertiary sector VA of all earlier years, and the 2004 and the 2008 economic censuses should have led to 48.71 and 9.01% upward revisions in the tertiary sector VA of all earlier years. The upshot is that 1978 tertiary sector VA would then have had to be revised upward by 114.03% (1.3204 times 1.4871 times 1.0901) rather than the NBS's 4.37% upward revision. Overall, 1978 GDP should have been revised upward by 34.13% (1.0999 times 1.1681 times 1.0445) rather than the NBS's 1.00% upward revision. Perhaps revisions to nominal values on such a scale can constitute an upper bound for the potential size of plausible revisions of historical values (under the assumption that current-year nominal values are correct.) Such a large upper bound gives the NBS great freedom in choosing historical GDP values.

Separately, the definition of GDP itself as excluding (unpaid) home production may imply that relatively low GDP values in the early years are technically correct. In the socialist economy, the scale of home production—from cleaning to childcare, household repairs and maintenance—may have been large, and would by definition not be counted in GDP. As these services became increasingly outsourced in the reform period, they were counted in GDP. Thus, while the change in GDP over time may appear implausibly fast as it would seem to exaggerate the growth in productive activities, in fact it may correctly reflect the growth in those productive activities which are to be counted.[58]

4.2. Scope for falsification of real GDP growth rates

Movements in China's real GDP are not significantly different from movements in related data. Thus, Klein and Ozmucur (2002–2003) in principal component analysis find that movements in their principal components are consistent with China's real GDP, both using annual data for fifteen real measures for 1980–2000, as well as monthly data for twenty real measures for February 1992 through April 2002. Mehrotra and Pääkkönen (2011) use factor analysis to summarize information from a variety of macroeconomic indicators (using quarterly data from 1997 to 2009)—including energy production and the output of industrial products—and find that the estimated factors match the GDP dynamics well. Fernald, Malkin, and Spiegel (2013) in principal component analysis find that official real GDP (quarterly data 2000 through third quarter 2009 with an out-of-sample period from the 4th quarter of 2009 through 2012) is systematically related to alternative indicators of Chinese economic activity.[59] While these findings suggest that *movements* in real GDP are accurate, they cannot show that the *level* of China's real GDP (or the level of China's real GDP growth rate) is accurate.

---

[55] See Xu (2006, pp. 17f). Xu (2009, p. 457) provides a 6-item breakdown of how the production coverage of GDP was expanded in the year of the economic census to newly include: (i) some enterprise service activities (that are not covered in annual statistics), (ii) service activities of administrative and public institutional units (that are not covered in the annual statistics), (iii) production activities of those household unincorporated enterprises that did not register with the authorities, (iv) establishments whose activities differ from the principal activities of the enterprises to which they belong, (v) housing rent services from households, and (vi) domestic and personal services produced by employing paid domestic staff that had not previously been estimated. The 2006 benchmark revision following the 2004 economic census further incorporated three innovations in the calculation of sectoral VA, implemented to comply with international practices: (i) interest on household savings deposits was previously counted as financial sector VA and is now attributed to the individual sectors that produced the particular VA, (ii) depreciation on residential housing was previously based on construction costs, but is now based on current market values, and (iii) the economic census collected data on income from computer software sales, which allows the NBS to include approximate expenditures on computer software in gross fixed capital formation.

[56] For the data, see *Statistical Yearbook 1994*, p. 32; *1995*, p. 32; *2005*, p. 51; *2006*, p. 57; *2009*, p. 37; *2010*, p. 38. Tertiary sector VA and GDP of the years 1978 through 1992 also underwent (minor) revisions.

[57] The Material Production System in use through the early 1990s focused on material production: agriculture, industry, construction, transport and communication, and wholesale and retail trade, thus ignoring a large part of the tertiary sector. It may be difficult to retrospectively create complete tertiary sector data.

[58] Such a transition could be reflected in a rising labor force participation rate. However, there has been no significant, long-run increase in the labor force participation rate in the reform period. Focusing on the years with population censuses, the labor force participation rate, defined as the employed divided by the population in the age group 15–64 years, increased from 0.8486 in 1982 to 0.8558 in 1990 before declining to 0.8118 in 2000 and 0.7622 in 2010 — some of the latter decline could be due to an increase in the duration of compulsory education and higher education, and home production in the pre-reform and early reform period may have been predominantly provided during free time (while holding a full-time job) or by retirees (those older than 64 years). *Statistical Yearbook 2007*, pp. 108, 130, and *2012*, pp. 105, 125, with pre-1990 employment values adjusted to account for the 1990 statistical break in the official employment data (Holz, 2013c).

[59] One set of alternative indicators includes electricity production, rail cargo shipments, and loan disbursements; another set includes an index of consumer sentiment, construction of new floor space, an index of raw materials usage, air passenger volume, and the nominal value of new residential real estate construction; a third set includes Chinese imports and exports reported by its trading partners.

The NBS's handling of the benchmark revisions of nominal values has potential implications for real growth rates. The NBS chose to have it both ways: only minor revisions to nominal VA of the earlier years, and largely unchanged real growth rates, thereby letting the implicit deflator bear the brunt of the revisions to nominal data. But revisions to implicit deflators are not plausible. Price indices are final in the year in which they are published (neither annual nor monthly price indices have ever been revised), as are deflators derived from the data on the directly reporting enterprises. Revisions to nominal data are largely, if not exclusively, based on revised data from those statistical units that do not report directly to the NBS: these data almost certainly comprise only nominal values (no constant-price output values), in many instances not even of direct output measures but of related measures such as sales revenue. If the first published implicit deflators are correct, this has two possible implications. First, if the official real GDP growth rates are correct, the benchmark revisions of nominal data should have come with larger increases of the values of earlier years (previous section). Second, if the benchmark revisions of nominal data are correct, official real GDP growth rates are systematically underestimated.

Taking nominal values as given, the key issue for real growth rates is the quality of the deflators. Klein and Ozmucur (2002–2003) argue that China's real GDP growth could well be significantly underestimated given that (unobserved or disregarded) quality changes lead to an overestimate of the deflator. In the U.S., the Boskin Commission concluded that the 1995–96 CPI was overestimated by 1.1 percentage points, and the subsequent changes to the consumer price index (CPI) compilation practices of the Bureau of Labor Statistics led to substantial increases in U.S. real GDP growth rates. If quality corrections were necessary in China, the effect on Chinese real growth rates would likely be larger than in the U.S.

On the other hand, some authors argue that China's deflators and price indices are underestimated. Thus Movshuk (2002), using proxies of sectoral price indices for 1991–1999, argues that the official implicit deflators are underestimates of the true price development and China's real GDP growth is therefore exaggerated by approximately two percentage points per year.[60] Young (2000) finds that replacing implicit sectoral deflators by proxies of sectoral price indices reduces real GDP growth of 1978–1998 by 1.7 percentage points (2.5 percentage points if agriculture is excluded), with all of the downward adjustment in the growth rate being brought about by the deflator substitution *after 1986*. However, an early double-check of real GDP growth rates against real growth in aggregate expenditures (requiring a number of assumptions on deflators in deriving real expenditures) yielded no systematic differences (Keidel, 2001).[61]

In the following, the nominal values are taken as given and an alternative set of deflators is constructed from available price indices. In the production approach to the calculation of GDP, these are sectoral price indices. In the expenditure approach, these are price indices specific to individual expenditure categories. The intention in reconstructing the official (revised implicit) deflators is to follow as closely as possible NBS explanations on how the NBS proceeds in obtaining deflators—explanations which are entirely plausible—while also exploring alternative deflators. A background paper (Appendix A) provides details on how the NBS obtains sectoral and expenditure deflators, and on how to use the available data on price indices to independently construct sectoral and expenditure deflators.[62] The following two subsections summarize the process of establishing an alternative set of deflators and report the key findings.

4.2.1. Sectoral real VA (and GDP)

NBS explanations are available on how the NBS obtains sectoral VA real growth rates or sectoral constant-price VA. The procedure changes over time. The case of agriculture and industry are explained in some detail here; all sectors are covered in Appendix A.

In agriculture, the NBS obtains real VA as the difference between real GOV and the real value of intermediate inputs (with VAT added separately in those years when not included in GOV). Real GOV until approximately 2004 was obtained by first multiplying individual output quantities by constant prices and then applying the resulting aggregate constant-price real growth index to the base year value of agricultural GOV; since then, real GOV is obtained by dividing the nominal values by a price index. The real value of intermediate inputs is obtained by applying 14 (since approximately 2004, 12) price indices to the current-price values of 14 (12) different categories of intermediate inputs.[63]

For industry, prior to approximately 2004, the NBS obtained real growth rates by applying a deflator to current-price industrial VA. Current-price VA of all industry is the sum of two separate datapoints. In the case of the DRIEs, VA is supposedly obtained as the difference between GOV and the value of intermediate inputs, plus, since 1995 (when GOV no longer included

---

[60] Movshuk (2002) presents two pieces of evidence: (i) the implicit GDP deflator for industry is approximately two percentage points below the survey-based producer price index; (ii) applying various price indices to the primary sector, industry, construction, three tertiary sector subsectors, and a tertiary sector residual in order to obtain alternative sectoral real growth rates suggests that real GDP growth is exaggerated by approximately two percentage points per year. Movshuk argues that a main mechanism for the underestimation of deflators is the base year pricing method. Enterprises may report identical nominal and real output values when no official base year price for a particular product exists, or if the use of base year prices appears too cumbersome and enterprise practices are not subject to stringent controls. Thus, Movshuk finds that in the period 1991–1999 the implicit deflator for the output of industrial SOEs is significantly higher than that of the *less regulated* industrial collective-owned enterprises (mean 5.5% vs. 1.7%, with a mean implicit GDP deflator for all industry of 5.1%). The fact that the output share of industrial collective-owned enterprises has fallen drastically over time should alleviate underestimation of the deflator in more recent years; furthermore, in the early 2000s the NBS adopted the use of producer price indices to deflate industrial VA.

[61] Keidel's aggregate expenditure real growth rate is 1.5 percentage points below the official real GDP growth rate in 1997, 0.6 percentage points lower in 1998, and 3.1 percentage points lower in 1999; on the other hand, it is higher by 2.0 percentage points in 1994, by 0.9 percentage points in 1995, by 0.2 percentage points in 1996, and by 0.8% age points in 2000.

[62] Appendix A is available at http://ihome.ust.hk/~socholz/QualityGDP.html.

[63] See, for example, NBS (1997, pp. 21–30), or Xu (2000b, pp. 75f). For more recent years, see NBS (2007).

VAT), VAT applicable to the products produced.[64] In the case of all other industrial enterprises, a sample survey collects data on GOV and on the income components of VA; the ratio of VA to GOV in this sample is then applied to the GOV of all non-DRIEs.[65]

The deflator for industrial VA is a GOV deflator with adjustments, obtained in a two-step procedure. First, enterprises—and presumably these are the DRIEs only—price their output at (base year) product-specific constant prices provided by the NBS (or a department designated by the NBS) and revised approximately every decade.[66] Aggregating across products yields a constant-price GOV time series and thereby a constant-price (real) growth index. Since some (base year) constant prices may differ from the base year *market* prices, the NBS applies this constant-price growth index to the base year's current price GOV to obtain a real GOV time series in (market) base year prices. Contrasting this time series to the series of current-price GOV of these enterprises yields the deflator. For the purpose of deflating VA, adjustments are made to this deflator depending on the development of the raw materials price index. This adjusted deflator is applied to the VA of *all* industry in order to obtain the real growth rate of industrial VA. Since approximately 2004, the NBS calculates real VA of each industrial sector by dividing nominal VA by the corresponding component (sub-index) of the ex-factory price index of industrial goods, and then aggregates across industrial sectors to obtain industry VA.[67]

In construction, real VA is obtained using a price index. In the tertiary sector, most subsectors use price indices, but real VA of some activities in 'transport and communication' is obtained using constant prices, while for other activities, base year VA is multiplied by an output quantity growth factor.[68]

Fig. 6 through Fig. 8 show a selection of alternative deflators—justified by NBS explanations on how the NBS obtains sectoral VA real growth rates—for the primary sector, industry, and the tertiary sector. The figure for construction shows only small differences between alternative series and is omitted. (Whenever data are aggregated, a Törnqvist index is used.) In comparison to the official implicit deflator (calculated from the nominal and real growth data published in the *Statistical Yearbook 2012*, or using the first published implicit deflators), alternative price series in the primary sector can diverge significantly in individual years, whether that is the rural retail price index (with GOV deflated by price indices specific to each subsector, and one price index for all agricultural intermediate inputs) or the agricultural procurement price index. In industry, the ex-factory price index is significantly higher in 1985, 1989, 1993 and 1994 (and thereby driving Movshuk's and Young's findings).[69] In the tertiary sector, using approximated price indices for six tertiary sector subsectors (with no significantly different results if a more detailed breakdown available for some years is used), this alternative tertiary sector deflator tends to be slightly higher than the official implicit deflator.

---

[64] As mentioned in Section 2.7, the NBS report form system in effect in 2003 does not include intermediate inputs. the NBS, thus, in contrast to its own explanations as to how it obtains VA of the DRIEs, may have had to apply some ratio to GOV all along in order to obtain DRIE VA.

[65] See Xu (2000b, pp. 23–26). The applicable VAT is derived from the value of VAT actually paid, taking into consideration timing (time of sales when the VAT accrues vs. time of production) and corrections for tax reimbursements (for example, in the case of exports) and VAT paid on bought-in products. Small enterprises may simply apply a flat 6% rate to sales revenue (which not necessarily equals VA) to obtain the value of VAT. The GOV of some small enterprises may be approximated with business revenue. (For the complications of the statistical break in the GOV series in 1995, see Holz & Lin, 2001) NBS (1997, pp. 32–7) provides similar explanations: the DRIEs at that point in time were defined differently and this source provides further details for the case of the non-DRIEs. The *GDP Manual*, 2001, p. 21, states that the ratio of VA to GOV which is to be applied to the non-DRIEs' GOV is provided through sample surveys implemented by the (NBS's) enterprise survey teams. For the years since 2004, NBS (2007, pp. 16f), explaining economic census procedures, proceeds along similar lines as Xu (2000b) and NBS (1997); the economic census provides census data for 2004 so that no mentioning of surveys is made.

[66] In the early 1990s, the constant-price list applied to all industrial enterprises at the township level and above. It explicitly did not apply to village and below village-level enterprises or to the urban or rural self-employed (NBS, 30 Sept. 1990). The constant-price list consists of two parts. One is a national list that covers approximately 1500 industrial products in 45 categories. The national list reprinted in NBS (1995, pp. 1055–1132) does not contain actual prices. It names, for each category, a government department or company that (presumably) determines the price. The second part consists of a supplemental list at the provincial level; it may not substitute for or replace items included in the national list, and it is to be filed with the NBS. (The literature has mentioned the existence of comprehensive price manuals in use by statistical departments if not firms, apparently, however, without such a price manual ever having found its way into a library accessible to Western researchers.) A number of special cases is discussed in the regulation (NBS, 30 Sept. 1990). For example, if an enterprise's product differs in specification to one included in the list, then the constant price of that enterprise's product is obtained by the local statistics department based on the product's 1990 relative price compared to the item included in the list. Newly introduced products are to be reported up to the relevant government department, which is to determine the constant price; with approval of the local statistics department, a temporary procedure is to use as the constant price the ex-factory price when the product is first produced (which would appear the basis for any claims that a deflator is underestimated).

[67] See NBS (2007, p. 21). One scenario that appears plausible, given the explanations on data compilation as well as the data themselves, is that the NBS used the GOV deflator of the enterprises at the township level and above through sometime around 1997 (when this series ends), although the stated adjustment to the GOV deflator based on the raw materials price index appears to not have happened given that such an adjustment would in many years have moved the GOV deflator further away from the official implicit deflator of industrial VA. Sometime after 1997, the NBS may have switched to the ex-factory price index of industrial products. The *GDP Manual* of 2001, an internal document on the compilation of GDP data, states that the GOV deflator, which is applied to current-price VA in order to obtain real growth rates of VA, is in fact the ex-factory price index compiled by the urban survey team (p. 67). This contrasts with the constant price method outlined above. Similarly, a statement by former NBS commissioner Li Deshui regarding the reasons for the discrepancies between summed provincial and national GDP suggests that the price index method (without specifying the use of the ex-factory price index) became the standard method at the national level in 2004. The *Statistical Yearbook* no longer reports GOV in base year prices for the DRIEs (or any other enterprise group) starting with the year 2004; all output values in the industry section of the *Statistical Yearbook* now are nominal values. The *Statistical Yearbook 2005*, exceptionally, on the first page of the industry section, reports real VA growth for the DRIEs and various subgroups of the DRIEs for 2004.

[68] For details, see, for example, Xu (2000b, pp. 77–80), or the *GDP Manual*, 2001, pp. 67–74. For data since 2004, see NBS (2007).

[69] Not astonishingly, Angus Maddison (1998) in his desire to achieve lower real growth rates for industry, without presenting evidence, states that "there are two official price indices which provide a more realistic measure of the pace of inflation" than official implicit deflators; these are the producer price index for industrial products (in official terms, the "ex-factory price index of industrial products"), and the retail price of industrial products in rural areas (p. 140).
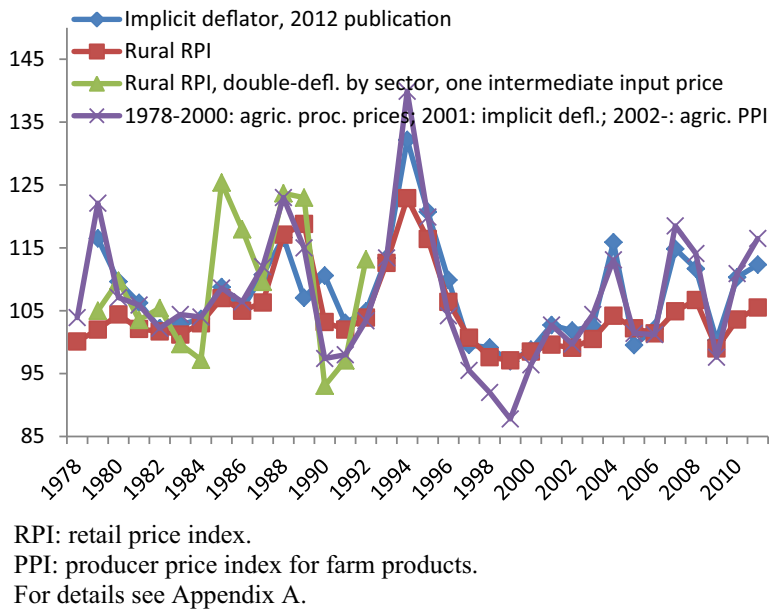
RPI: retail price index.
PPI: producer price index for farm products.
For details see Appendix A.

Fig. 6. Primary sector deflators. RPI: retail price index. PPI: producer price index for farm products. For details see Appendix A.
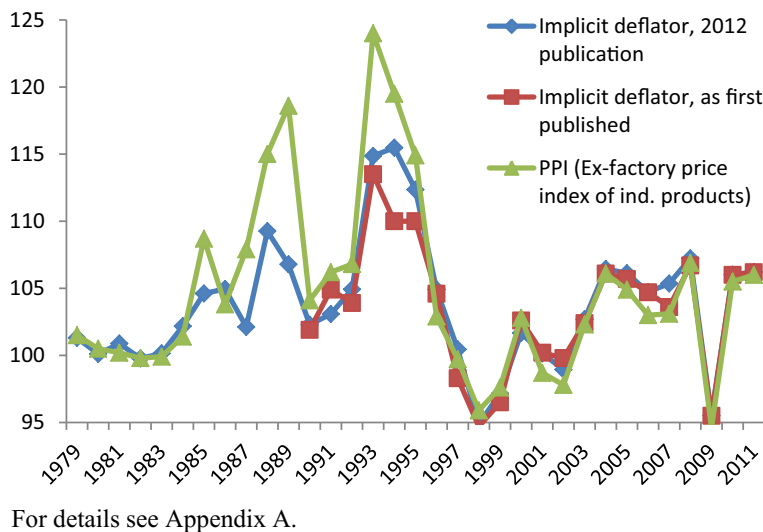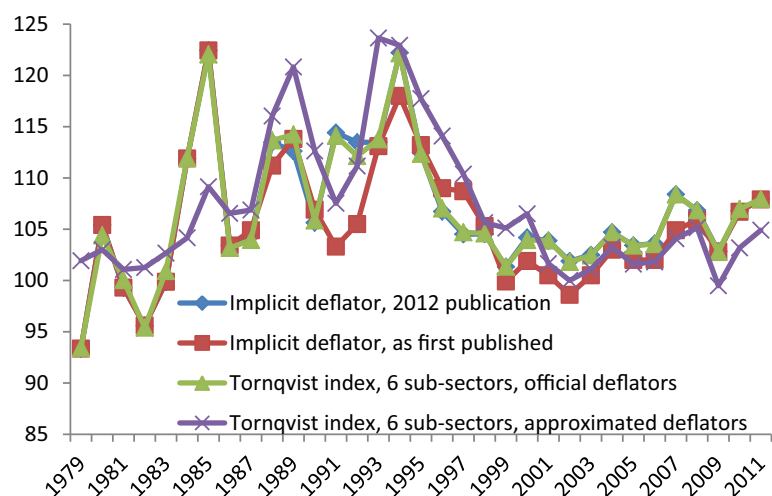

For details see Appendix A.

Fig. 7. Industry deflators. For details see Appendix A.

Table 4, using the most recently published nominal data, reports average annual real GDP growth rates for the period 1978–2011 in four deflator scenarios:

- With the official implicit GDP deflator, the average annual real growth rate is 9.8%.
- With the *first published* official implicit deflator, it is 10.5%.
- If each sector's nominal VA is deflated using one of the lowest (plausible) alternative price indices, it is 11.0%.
- If each sector's nominal VA is deflated using one of the highest (plausible) alternative price indices, it is 9.1%.

In other words, while the official (most recently published) implicit average annual real growth rate is 9.8%, it could reasonably be anywhere between 9.1% and 11.0%.

Fig. 9 charts the time series of the deflators for these four scenarios. What is striking in the low-growth (high-deflator) scenario is that the alternative deflator significantly exceeds the official implicit deflator only in certain years, 1988–1989 and 1993–1994, the two high-inflation periods in the reform period. Except for 1989, these are also years of high real growth, and the NBS thus should not have had incentives to under-estimate inflation when calculating real GDP growth. In contrast to the low-growth scenario, the high-growth (low-deflator) scenario exhibits no singular striking discrepancies to the official implicit deflators.

For details see Appendix A.

Fig. 8. Tertiary sector deflators. For details see Appendix A.

Table 4
Average annual real GDP growth rates.

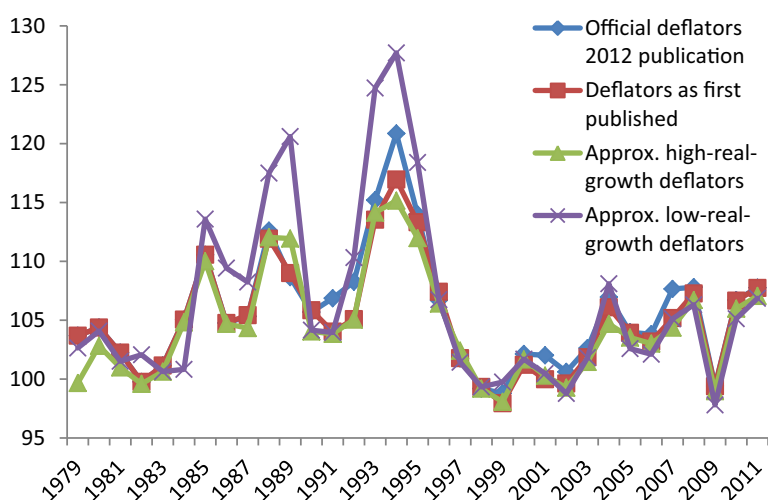|	Implicit deflator:		Reasonable deflator scenario for	
	Statistical Yearbook 2012 "official"	As first published (1987/90-, earlier years: official)	High real growth	Low real growth
Primary sector			Rural RPI	Rural RPI, double-deflated by sector, one intermediate input price
Industry			As first published	PPI
Construction			Investment in fixed assets price index: construction and installation sub-index (−1990: official)	Industrial products purchasing price index: building materials sub-index (−1988: official)
Tertiary sector			As first published	Törnqvist index of approx. deflators for exhaustive six tertiary sector sub-sectors
Average annual real GDP growth rate 1978–2011, in %	9.8	10.5	11.0	9.1

4.2.2. Aggregate real expenditures

In the expenditure approach, the NBS reports the real growth rates of final consumption with its different components as well as of gross capital formation with its two components through 2004. It does not report real growth rates of net exports or of aggregate expenditures.[70]

As in the case of GDP, details of how the NBS derives implicit deflators are relegated to Appendix A. In brief, household consumption is deflated by product (service) category using sub-indices of the CPI, except that the investment in fixed assets price index is applied to the imputed value of owner-occupied housing services, and for financial services enters a weighted deflator together with the CPI.[71] Government consumption in form of fixed asset depreciation is deflated with the investment in fixed assets price index, and all other government consumption with the CPI. Gross fixed capital formation is deflated with the investment in fixed assets price index, while inventory investment is deflated by sector using a variety of methods. Exports and imports of goods and services are deflated with price indices that incorporate information from the Customs Office, other countries, and the service component of the CPI; these price indices are not published, and neither are real growth rates or real value series (which would allow the calculation of implicit deflators).

---

[70] Real growth rates of inventory investment (published) and of net exports (not published) are problematic in that when one year's values are close to zero (absolute-value), growth rates of this and the following year can be extremely large. (The NBS does not publish real growth rates of inventory investment in 1962 and 1963; in 1962, inventory investment was near-zero.) The volume series (not published) are still meaningful.

[71] Prior to 2004, household consumption was deflated using the corresponding components of the CPI with a number of exceptions, of which the major ones are: rural self-produced self-consumed production was deflated using the corresponding agricultural GOV deflator, owner-occupied housing using the investment in fixed assets price index, financial services using the investment in fixed assets price index and the retail price index, and urban consumption of agricultural sideline products using the agricultural products market transaction price index (*nongmao shichang nongchanpin chengjiao jiage zhishu*). See NBS (1997, pp. 157–61).

For details see Appendix A.

Fig. 9. Alternative GDP deflators. For details see Appendix A.

To construct alternative real growth rates by expenditure category, and then to calculate an alternative real growth rate series for aggregate real expenditures, a number of alternative price indices are considered (and detailed in Appendix A). In the case of exports and imports of goods and services, the available nominal net export values are split into exports and imports using the Balance of Payments data available since 1982. Price indices for exports (and imports) of goods are then constructed using the export and import breakdown by sub-category available in the trade statistics, combined with corresponding price subindices of such price indices as the ex-factory price index and the purchasing price index for industrial producers. Exported services are deflated using the Chinese service price index, and imported services using the U.S. services subindex in the U.S. urban consumer CPI.

Table 5 reports the results for three scenarios (and in comparison to the official GDP real growth rate):

- a semi-official series using the available official real growth rates of consumption and gross capital formation available through 2004 combined with the here calculated real values for net exports;
- an alternative series "Constructed A" using derived deflators throughout, except in the case of gross fixed capital formation where the official implicit deflators are used for the years through 1989, followed by the investment in fixed assets price index available starting 1990;
- an alternative series "Constructed B" using derived deflators throughout, with the ex-factory price index of industrial products used to deflate gross fixed capital formation.

The starting year 1983 is due to the ability to reliably break down net exports into exports vs. imports starting in 1982 only, and the choice of 2004 as one possible end year is determined by the availability of official real growth rate data for consumption and gross capital formation.

Through 2004, the constructed series have a lower average annual real growth rate than the semi-official series or the official GDP real growth rate series, with a difference of no more than 0.5 percentage points. In a comparison of the alternative series and the official GDP real growth rate series over the complete reform period, the alternative series yield marginally higher results (up to 0.2 percentage points higher). Throughout, there is no difference if one calculates annual real growth rates of aggregate expenditures by deflating each year's nominal value to the previous year's prices in order to obtain the annual real growth rate, or if one uses a Törnqvist index of the annual real growth rates of the different expenditure components.

Table 5
Average annual real growth rates of aggregate expenditures.

	Semi-official		Constructed A		Constructed B		*Reference*:
		as Törnqvist index		as Törnqvist index		as Törnqvist index	GDP
1983–2004	9.9	9.9	9.4	9.4	9.5	9.5	10.0
1979–2004	9.6	9.6	9.2	9.2	9.3	9.3	9.6
1979–2011			10.0	10.0	10.1	10.1	9.9
1983–2011			10.3	10.3	10.3	10.3	10.2

For definitions of the three (six) series see the text.
The data for the reference GDP series are the official real growth rates reported in the *Statistical Yearbook 2012*.

5. Benford's Law

With unpublished data on economic activities in small units, potentially varying coverage in face of a moving boundary between the recorded economy and the shadow economy, and a choice of plausible deflators and price indices, the NBS has much scope in choosing how to determine the final nominal and real growth data. But does the NBS go beyond choosing calculation methods to choosing final figures? If the NBS falsifies the final data to fit its purposes, it may yet violate certain statistical data regularities.[72]

5.1. Introduction to Benford's Law

Benford's Law, also named "first digit law," states that the frequency distribution of first digits in many—but not all—real world data is not distributed uniformly, but according to the widths of gridlines on a logarithmic scale: the first digit "1" occurs with frequency 30.1%, the first digit "2" with frequency 17.6%, etc., with continuously declining frequencies all the way to the first digit "9" (4.6%).[73]

To illustrate Benford's Law, assume that a quantity doubles every year. Starting with the value 100, for one whole year the leading digit will be 1, until at the end of the year the value 200 is reached. The leading digit will then be 2 for approximately seven months, and 3 for the remaining approximately five months of the second year. In the third year, the leading digits will be 4, 5, 6, and 7, with each digit being the leading digit for an ever shorter period of time. In sum, the digit 1 occurs for a longer period (more frequently) than the digit 2, which in turn occurs more frequently than the digit 3, etc.

Smith (2013) explains that by focusing on first digits one is in fact *constructing* the Benford distribution: the focus on first digits implies that each number under examination is being multiplied or divided by ten until it is between 1 and 9.9. But multiplying or dividing by ten means changing the numbers in a pattern based on powers of ten — in other words, taking the anti-logarithm (of log to the base ten). Not astonishingly, the processed data look logarithmic.[74]

Mathematically, Benford's Law is stated as the probability of a particular digit value d occurring:

$$p(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}\left(1 + \frac{1}{d}\right). \tag{1}$$

The equation also applies to a particular *string* n of digits. For example, if d is the string 251, the probability that a number starts with this string is $\log_{10}(1 + 1/251) = 0.001727$. This can be extended to the probability that a particular digit occurs at a given position within a number. For example, the probability that the digit 3 occurs as second digit is

$$\log_{10}\left(1 + \frac{1}{13}\right) + \log_{10}\left(1 + \frac{1}{23}\right) + \dots + \log_{10}\left(1 + \frac{1}{93}\right) = 0.10433. \tag{2}$$

The distribution of the i_{th} digit, as i increases, approaches a uniform distribution with 10% for each of the ten possible numbers 0, … 9.

The Benford distribution has two properties: it is base invariant and scale invariant. Base invariance means that Benford's Law can also be expressed in bases other than the decimal base, as long as the base is larger or equal to 2. Scale invariance means that Benford's Law holds independent of the scale of the variable (for example, GDP expressed in the home currency or in a foreign currency). The Benford distribution is the only distribution with these properties (Hill, 1995a,b, 1998; Pinkham, 1961).[75]

To use Benford's Law to check for data falsification requires the assumption that the true data follow the Benford distribution. In the early literature, the focus was on datasets that may comply with Benford's Law, while more recently the focus has shifted to *distributions* that qualify. Two families of processes are known for which the data converge to the logarithmic Benford distribution as limiting distribution (Michalski & Stoltz, 2013): (i) random samples of random distributions (Hill, 1995c), and (ii) processes with exponential growth.[76] The observed data—either taking the complete sample of available observations, or a random sample thereof—need to cover several orders of magnitude in order to approach the Benford distribution. Nye and Moul (2007) suggest that the data should span at least two orders of magnitude, while Fewster (2009) reports that with four orders of magnitude and a sample size of 100 it is "difficult for a distribution *not* to look reasonably Benford" (p. 31).

---

[72] I am grateful to Rodney Chun for bringing Benford's Law to my attention, probably as early as the mid-2000s.

[73] For an easily accessible overview of Benford's Law see Browne (2013), Weisstein (2013), and the entry in Wikipedia on Benford's Law (accessed 15 January 2013). The regularity now called Benford's law (following Benford, 1938), was earlier noted by Newcomb (1881).

[74] Raimi (1976) provides a selection of early reasonings as to why data may follow the Benford distribution.

[75] Consequently, if a set of numbers does not follow the Benford distribution, multiplication by some constant may move the set further away or closer to the Benford frequencies, but no constant exists that turns a non-Benford distribution into a Benford distribution (Raimi, 1969).

[76] This also implies that the data are not subject to artificial truncation, psychological barriers (for example, prices concentrated at specific values such as $1.99), or structural shifts.

A number of tests are available to test for conformity of observed data to the Benford distribution.[77] The most widely used test is the Chi-squared test:

$$D^2 = n \sum_{d=1}^{9} \frac{(h_d - p_d)^2}{p_d} \tag{3}$$

where n denotes the number of available observations, h is the observed empirical frequencies of the digits d element $\{1, ..., 9\}$—in the case of second digits d element $\{0, ..., 9\}$—and p is the theoretical frequencies (such as in Eq. (1) for first digits). As n goes toward infinity, the statistic D^2 converges in distribution to a Chi-square distribution with 8 (9) degrees of freedom for first (second) digits. The Chi-squared test has low statistical power in small samples (i.e., tends to fail to correctly reject the null hypothesis of no difference between the observed and theoretical frequencies). Power improves with sample size to the point that it may become too "easy" to reject the null hypothesis. Michalski and Stoltz (2013) suggest a minimum sample size of approximately 100.[78]

Since the Chi-squared test is sensitive to quadratic deviations, an alternative is to test for a difference in the means of the observed data and the Benford distribution using the standard normal statistic with $T_{\overline{d}_1} = \sqrt{n} \frac{\overline{d}_1 - 3.940}{\sqrt{6.057}}$ for the first digit and $T_{\overline{d}_2} = \sqrt{n} \frac{\overline{d}_2 - 4.687}{\sqrt{8.254}}$ for the second digit, where the mean of the first digit of the Benford distribution is 3.940 (with variance 6.057) and of the second digit 4.687 (with variance 8.254), and the mean of the observed first digits is obtained as $\overline{d}_1 = \sum_{d_1=1}^{9} (d_1 + 0.5) h_{d_1}$ and for the second digits as $\overline{d}_2 = \sum_{d_2=0}^{9} (d_2 + 0.5) h_{d_2}$. The standard normal statistic can also be used to test for the compliance of *individual digits* with Benford's Law, with a test statistic of $T_d = \sqrt{n_d} \frac{h_d - p_d}{\sqrt{p_d (1 - p_d)}}$.

Three further tests are the Kuiper test (a modified Kolmogorov–Smirnov test), the m (max) statistic and the d (distance) statistic. With H_d and P_d as the *cumulative* frequencies of the first digit d in the observed data and in the Benford distribution, the Kuiper statistic is $T_K = (D_n^+ + D_n^-) \left[\sqrt{n} + 0.155 + \frac{0.24}{\sqrt{n}} \right]$, where $D_n^+ = \sup[H_d - P_d]$ and $D_n^- = \sup[P_d - H_d]$. Leemis et al. (2000) proposed the m statistic $m = \max_{d=1, ..., 9} |h_d - p_d|$, and Cho and Gaines (2007) the d statistic $d = \sqrt{\sum_{d=1}^{9} (h_d - p_d)^2}$. Morrow (2010) provides Benford-specific test values at the 10, 5, and 1% significance level for the Kuiper test as well as for modified m and d tests, with the modification consisting of multiplication by $\sqrt{n}$.[79] Morrow notes that the small-n performance of these tests is fairly good in that the simulated test statistics are very close to the asymptotic values, especially for $n \geq 80$.

Smith (2013) proposes a "Ones Scaling Test." If one multiplies each number in the set by a constant, the distribution of first digits will remain unchanged as a consequence of the scale invariance of the Benford distribution (and *only* of the Benford distribution). He explores 696 multiplications by 1.01 (in the final instance this implies multiplication by 1000) and plots the fraction of numbers with digit one. If Benford's Law is followed, the plot shows the same frequency (for any particular first digit) in all 697 instances.

Benford's Law has been applied in a wide variety of circumstances. Of most interest in the context here are the applications to macroeconomic data series.[80] For example, Rauch et al. (2011) examine annual national and financial accounts data (156 variables) for each EU country from 1999 to 2009 and rank countries by the adherence of their macroeconomic data to Benford's Law (Greece's data show the greatest deviation). Nye and Moul (2007) find that the World Bank International GDP data and the purchasing power parity corrected Penn World Table data for OECD countries conform well to Benford's Law, but some subsets of the data, especially GDP data from the developing world, do not, suggesting deliberate manipulation of the underlying series. Michalski and Stoltz (2013) examine the quarterly balance of payments data of 103 countries in the period 1989 to 2007 to determine under which conditions (such as having a fixed exchange rate regime) countries' data tend to violate Benford's Law. Gonzalez-Garcia and Pastor (2009) examine five annual time series in each of the areas banking, balance of payments, national accounts, and government finance statistics for the period 1945–2007 for 80 countries (nested in six country groups) and find that most macroeconomic series for most country groups conform with Benford's Law. Rejections of Benford's Law, however, are not correlated with the corresponding data quality ratings compiled by experts based on the IMF's Data Quality Assessment Framework (thus questioning the accuracy either of the

---

[77] The following description of tests draws, among others, on Stephens (1970) for the correction factor in the case of the Kuiper test, Leemis, Schmeiser, and Evans (2000), Cho and Gaines (2007), Nye and Moul (2007), Günnel and Tödter (2009), Morrow (2010), Rauch, Göttsche, and Brähler (2011), and Michalski and Stoltz (2013).

[78] The distribution of D^2 is close enough to the Chi-square distribution when $n \geq 30$ and $n h_i \geq 5$ for all $i \in \{1, ... 9\}$; the latter condition of $n \geq 5 / h_9$ yields the 100-criterion (with 9 being the least frequent first digit).

[79] At the 10, 5, and 1% significance level, the three values for the Kuiper test are 1.191, 1.321, and 1.579; for the modified m test 0.851, 0.967, and 1.212; and for the modified d test 1.212, 1.330, and 1.569.

[80] Benford's Law has also been used to detect or examine unusual patterns in firm reported earnings (Thomas, 1989), tax fraud (Nigrini, 1996) and accounting fraud (Durtschi, Hillison, & Pacini, 2004; Nigrini, 2012), psychological product pricing (el Sehity, Hoelzl, & Kirchler, 2005), price manipulations in eBay auctions (Giles, 2007), in-kind campaign donations (Cho & Gaines, 2007), fraud in scientific data (Diekmann, 2007), the quality of survey data (Judge & Schechter, 2009), and election fraud (Deckert, Myagkov, & Ordeshook, 2011). Varian (1972) proposed to use it as a criterion to evaluate a model's performance: if the input data obey Benford's Law but the output data do not, that would raise questions about the model. He examines the input and output (forecast) data of a forecasting model and finds both sets of data to be in accordance with Benford's Law.

expert opinions or of the Benford test). The authors also find that rejection of Benford's Law may result from structural shifts in data series.

Two authors show the importance of extending the analysis to the second digit. Diekmann (2007) in examining the first and second digits reported in published statistical estimates of research papers and in a similar set of faked data finds that both the first and second digits in published statistical estimates approximately match the Benford distribution, while only the first digits of faked statistical estimates do. In the economic realm, Günnel and Tödter (2009) for forecasts of GDP growth and CPI inflation in Germany find that the second digits of forecasts exhibit an excess of zeros and fives.

A key issue is how to interpret violations of Benford's Law in the observed data. This could imply that the underlying data distribution violates Benford's Law, that the data do not cover sufficient orders of magnitude, or that the data have been falsified.[81] Rather than conclude on outright falsification, some of the literature proceeds to compare or rank different sets of data (such as different countries' data on one and the same variable or set of variables) as to their conformity to Benford's distribution. I.e., conformity to Benford's Law is benchmarked against other comparable datasets; a lower degree of conformance is then interpreted as potentially lower data quality.[82]

5.2. Benford's Law and Chinese national income and product accounts data

To examine Chinese national income and product accounts data, four sets of nominal data for the years 1978–2011 are selected:

- (production approach) GDP (with data as first published, and as most recently published);
- sectoral VA for the four exhaustive sectors agriculture, industry, construction, and services (with data as first published, and as most recently published);
- aggregate expenditures; and
- five exhaustive components of aggregate expenditures.

U.S. data on aggregate expenditures in their totality, as well as with a breakdown into 8 and 14 components, are considered for comparison. (The U.S. does not publish production approach GDP data.) The number of observations for GDP or aggregate expenditures is small, with only 34 observations, and covering only two orders of magnitude. In the case of sectoral VA or the components of aggregate expenditures, the number of observations is above 100, typically covering three orders of magnitude.

The key question is why these data should conform to Benford's Law?[83] For nominal GDP, the case rests on exponential growth: for exponential growth processes, data converge to the logarithmic Benford distribution as limiting distribution. Nominal GDP in China grew at an average 15.89% per year between 1978 and 2011. As long as annual deviations from this average annual growth rate are random (the standard deviation of annual growth rates is 7.07 percentage points), one would expect Chinese nominal GDP data to conform to Benford's Law.[84]

The case for real growth rates and deflators is harder to make. Real growth rates need no further consideration: if nominal GDP as well as deflators follow Benford's Law, then so do real growth rates. The crux then is the deflators. One approach is to interpret price indices as exponential growth series. The GDP deflator in China grew at an average annual 5.51% per year between 1979 and 2011 (with a standard deviation of 4.82%).[85] However, unlike in the case of nominal GDP—where factors such as population growth, capital accumulation, human capital growth, and technical progress can be viewed as driving nominal GDP growth, and a case can be made that this is an ongoing and relatively stable process—growth in a deflator likely depends much more on singular decisions—such as monetary policy decisions—that can be adjusted at short notice, and the argument for stable exponential growth therefore is much weaker. Another approach is to focus on second digits of annual deflators (not an index series) and argue that while the first digit may not be random—monetary policy may, for example, target low inflation rates—the second digit could be anything. The Chinese central bank has not issued specific numeric inflation targets, and thus while there may be an intention to achieve a first digit in the lower single-digit range, the second digit is not subject to any monetary policy objective and is likely random. This is the approach taken here: in the case of deflators (and, by implication, real growth rates), second digits of non-fraudulent data are expected to follow the Benford distribution.

Table 6 reports the results of the test statistics introduced above in terms of the significance levels (10, 5, and 1%) of any significant deviations of the actual data from the Benford distribution.[86] As the top half of the table shows, the distributions of Chinese and U.S. nominal data do not differ significantly from the Benford distribution at the 10% significance level (except for two isolated instances), neither for first digits nor for second digits. Fig. 10 provides the corresponding charts for a selection of nominal data.

---

[81] Conversely, Benford's Law cannot detect if all data have been falsified equally by the same percentage. Benford's Law will also not detect the very occasional instance of data falsification, and the size (magnitude) of any such falsification has no impact on the likelihood of it being discovered.

[82] The larger the sample of observed data given a fixed *share* of falsified data, the more likely that statistical tests indicate data falsification; an issue for comparability across datasets therefore is to compare datasets of equal size.

[83] Across the literature, nominal GDP, real GDP growth rates, and the CPI have been assumed to follow Benford's Law. (See, for example, Nye & Moul, 2007, or Günnel & Tödter, 2009.).

[84] For agriculture, industry, construction, and services, the average annual growth rates of nominal value-added and the standard deviations are: agriculture 12.32% (standard deviation 8.22 percentage points), industry 15.53 (8.67), construction 17.93 (12.26), and services 17.99 (8.74) (*Statistical Yearbook 2012*, p. 44).

[85] By sector, the average annual deflator values (and standard deviations) are 7.14 (7.44), 3.72 (4.69), 6.67 (6.69), and 6.84 (6.37) (*Statistical Yearbook 2012*, p. 48).

[86] The Ones Scaling Test results in only a graphical representation and these results are not reported. For each set of data, the data were multiplied ninety-nine times by 1.05, and the (first- and second-digit) frequencies of the digits 1, 2, 3, and 4 were then plotted for the 100 results (original plus ninety-nine additional results). The results never yield straight frequency lines but undulating shapes. A likely culprit of the undulating shape is the relatively small order of magnitudes in the data.

Table 6
Significance levels of deviations of actual data from the Benford distribution.

Nominal data	n	D^2	$T_{\bar{d}}$	T_d	T_K	m*	d*	D^2	$T_{\bar{d}}$	T_d	T_K	m*	d*
		First digit						Second digit					
China													
GDP as first published	34												
GDP (*Statistical Yearbook 2012*)	34												
Sectoral VA (4 sectors) as first published	136											5	
Sectoral VA (4 sectors) (*Statistical Yearbook 2012*)	136	10											
Aggregate expenditures (*Statistical Yearbook 2012*)	34												
Aggregate expenditures: 5 components (*Statistical Yearbook 2012*)	170												
U.S.													
Aggregate expenditures	34												
Aggregate expenditures: 8 components	272												
Aggregate expenditures: 14 components	476												

Deflators and real growth rates: second digit	n	D^2	$T_{\bar{d}}$	T_d	T_K	m*	d*	D^2	$T_{\bar{d}}$	T_d	T_K	m*	d*
		Deflators						Real growth rates					
China (implicit deflators, published real growth rates)													
GDP as first published	25							1	1	a	1	1	1
GDP (*Statistical Yearbook 2012*)	33			10				5	1	b	5	10	10
GDP constructed high-growth scenario	33									10	10		
GDP constructed low-growth scenario	33									10			
Sectoral VA (4 sectors) as first published	94							1	10		5		5
Sectoral VA (4 sectors) (*Statistical Yearbook 2012*)	132							10					5
Sectoral VA (4 sectors) constructed high-growth scenario	132	1		c	1	1	1						
Sectoral VA (4 sectors) constructed low-growth scenario	132	5			5		10		10				
U.S.													
Aggregate expenditures	33							10					10
Aggregate expenditures: 7 components	231							5	5			5	5
Aggregate expenditures: 13 components	429							5	5	c	5	1	5

No entry implies no significance at the 10% level.

For both China and the U.S., nominal values cover the years 1978–2011, while deflators and real growth rates cover the years 1979–2011. Exceptionally, Chinese first published implicit deflators and real growth rates in agriculture and services cover the years 1987–2011, and in industry and construction 1990–2011; those of GDP cover the years 1987–2011 with real growth rates of 1987–1991 constructed from sectoral real growth rates (primary-agriculture, secondary, and tertiary-service sector) using previous-year nominal weights.

n: Number of observations.

Test statistics: D^2: Chi-squared statistic. $T_{\bar{d}}$: Difference in means standard normal statistic (two-sided test). T_d: Difference in individual digits (specific digit in actual vs. Benford distribution) standard normal statistic (two-sided test). T_K: Kuiper statistic. m*: Modified Leemis et al. (2000) m-statistic (m-statistic times $\sqrt{n}$). d*: Modified Cho and Gaines (2007) d-statistic (d-statistic times $\sqrt{n}$).

China GDP as first published: GDP through 1991, by necessity, obtained as summed sectoral VA; values through 1987 from *Statistical Yearbook 1988* (the first source of such data).

China sectoral VA (4 sectors): GDP breakdown into the four exhaustive sectors agriculture, industry, construction, and services.

China aggregate expenditures: 5 exhaustive components: household consumption, government consumption, gross fixed capital formation, inventory change (always positive and larger than ten units (100 million yuan)), and *absolute values* of net exports *times 100*.

U.S. aggregate expenditures: 8 (7) exhaustive components: personal consumption of goods, personal consumption of services, fixed gross private domestic investment, *absolute value* of change in private inventories (in gross private domestic investment) *times 100* [no deflator and real growth rate values available], exports, imports, federal government consumption and gross investment, state and local government consumption and gross investment.

U.S. aggregate expenditures: 14 (13) exhaustive components: in personal consumption: durable goods, non-durable goods, services; in gross private domestic investment: non-residential structures, non-residential equipment and software, residential, *absolute value* of change in private inventories *times 100* [no deflator and real growth rate values available]; in exports: goods, services; in imports: goods, services; in government consumption expenditures and gross investment: federal national defense, federal non-defense, state and local.

GDP constructed high-/low-growth scenario refers to the deflators and real growth rates constructed for these scenarios in the previous section.

Note a (b,c): second digit is 0 at 1% (5%, 10%) significance level. (In all cases, the digit 0 has more observations than the Benford distribution would suggest.)

The bottom half of Table 6 reports the findings for the case of deflators and real growth rates for 1979–2011.[87] First published real growth rates do not start until 1987 or 1990, depending on sector (using the real growth rate of a particular year as first published, and not using retrospectively constructed series for the early reform years). No official real growth data have been published for aggregate expenditures. Newly added are data for the high- and low-growth deflator scenarios presented in the previous section. For the official data, deflators are (by necessity) implicit deflators, derived by dividing nominal growth rates by real growth rates. For the high- and low-growth scenarios, deflators are the ones constructed in the previous section, and real growth rates follow by deflating the official

---

[87] For simplicity of automated calculation, in order to avoid having to manually identify the first digit of negative numbers and decimals, Chinese deflator and real growth rate statistics are based on the *absolute values* of the original percentages, times 1000. U.S. values were multiplied by 10,000. In the Chinese low-growth deflator scenario for 4-sector VA, one value was multiplied by 1,000,000.
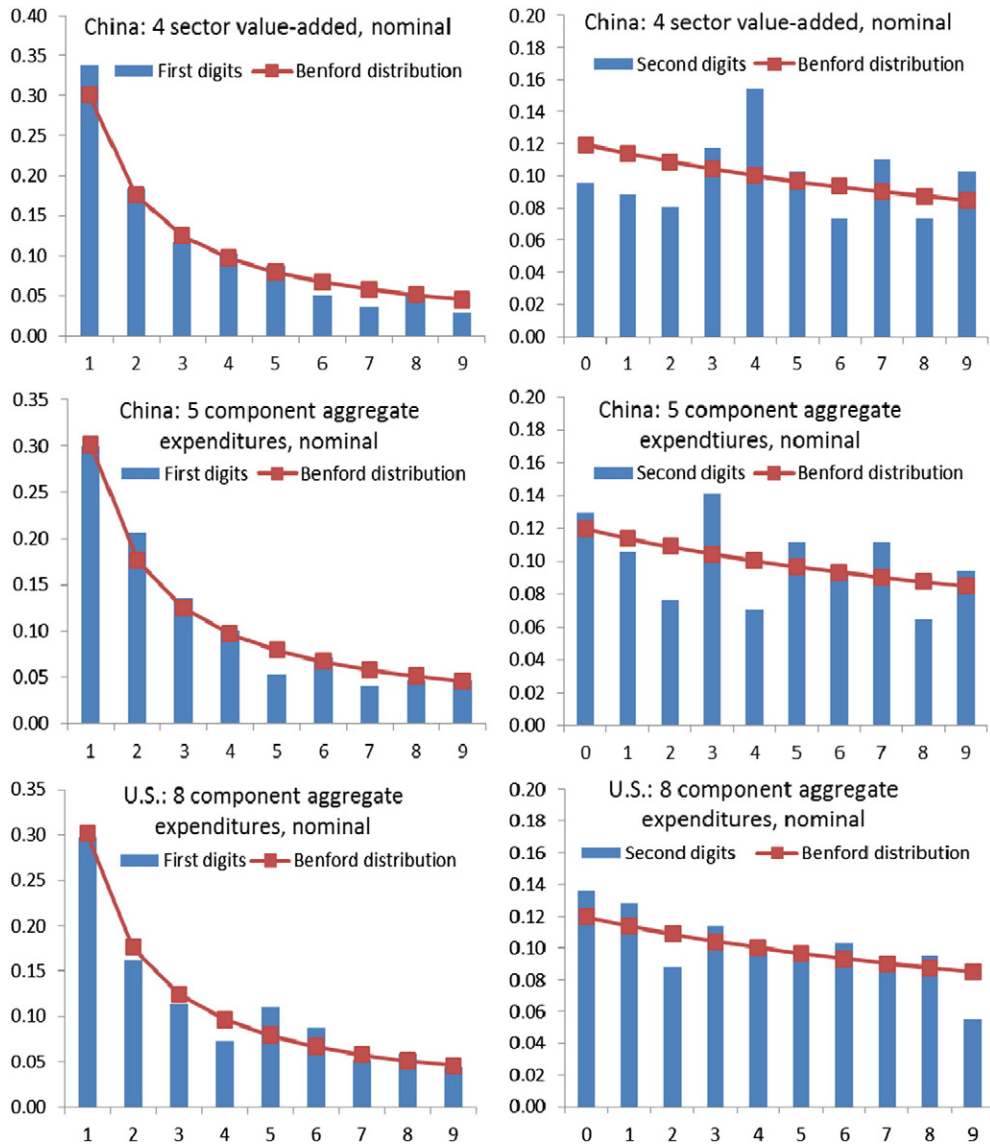
Fig. 10. Nominal data and Benford's Law.

nominal data. The U.S. separately publishes price and real growth data, with no such data available for the expenditure component 'changes in inventories.'

The test statistics suggest that the *sectoral* high- and low-growth deflators constructed in the previous section do not follow the Benford distribution (though the economy-wide values do). Chinese official deflators (in the aggregate and by sector) follow the Benford distribution, as do U.S. deflators.

It is for the real growth rates that test statistics consistently raise the alarm. Both, official Chinese real growth rates and U.S. real growth rates tend to *not* follow the Benford distribution. For China, the sectoral VA data fare better than the overall GDP data; for the U.S., aggregate expenditures in their totality fare better than the data on expenditure components.

Fig. 11 presents a selection of the data (keeping in mind that the number of observations in the case of GDP is small and the statistical tests therefore accept large deviations). The figures show that for both the U.S. and China the problem with the real growth rates is that the second digit is zero with too high a frequency. The fact that there are too many zeros suggests rounding by the statistical authority. This could be intentional, accurate rounding for the sake of not pretending a higher degree of accuracy than the data can provide (with a penchant for rounding to the zero), or intentional rounding for the sake of falsifying data. One could also argue that deflators may not follow Benford's Law, and the fact that there are too many zeros then cannot be interpreted as rounding or data falsification. In this case, however, it is striking that both China and the U.S. exhibit the same second-digit pattern. The quality of Chinese deflators, thus, appears vouched for by U.S. data.
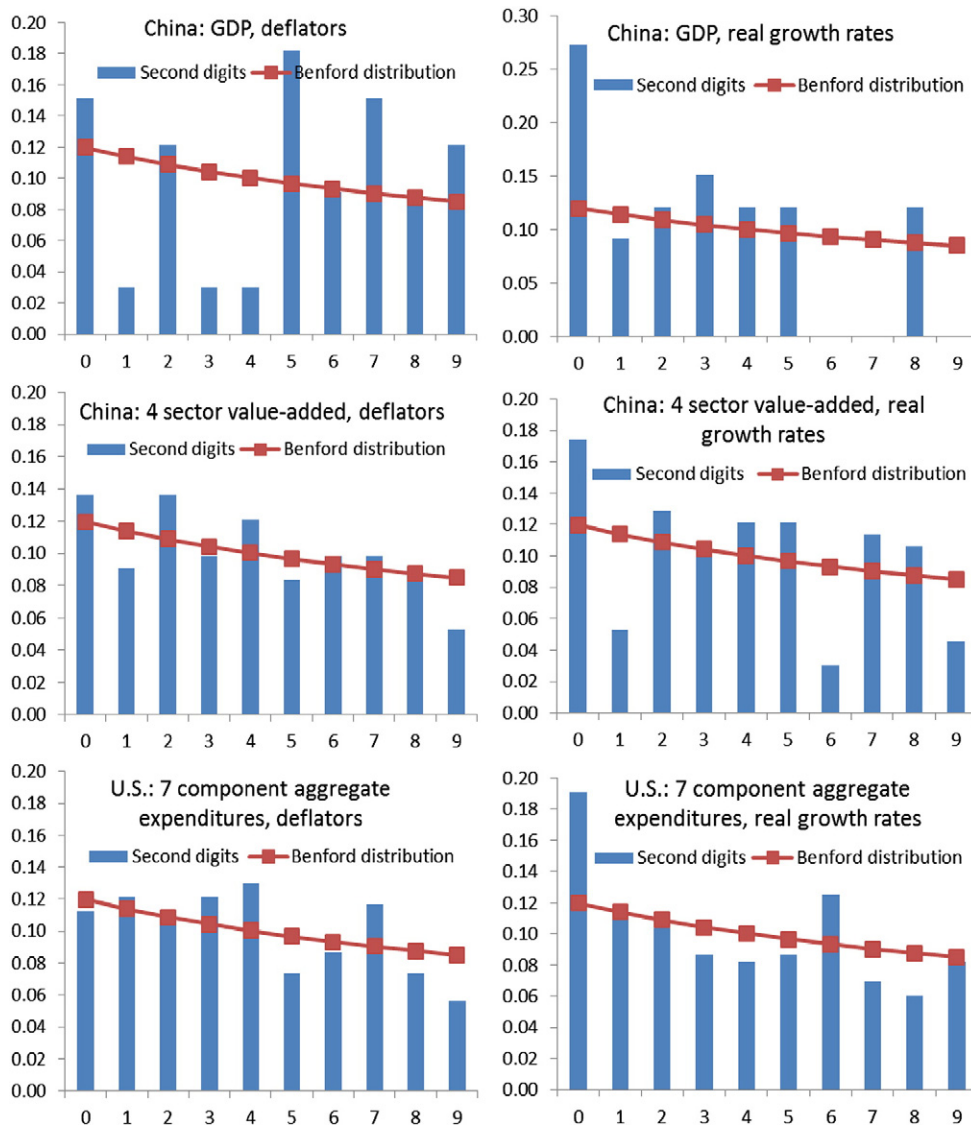
Fig. 11. Deflators, real growth rates and Benford's Law.

6. Conclusions

Reviewing past and ongoing suspicions of the quality of Chinese GDP data, one is hard pressed to find evidence of data falsification. One may wonder if the 2006 benchmark revision following the 2004 economic census should not have led to *higher* real growth rates. One cannot but question the quality of the 2007 DRIE VA data (and then wonder how the NBS obtains data on industrial VA). And one may ponder if the 2007 real GDP growth rate is economically feasible. But there is no clear evidence of data falsification.

The institutional arrangements of data compilation in China raise major questions. The NBS has no degree of formal independence and serves the Party and the government. How the NBS obtains the final nominal GDP value and real GDP growth rate is anybody's guess. The limitations of publicly available data do not allow one to double-check the accuracy of the official (final) data.

The NBS has scope for intentional manipulations in order to achieve a politically desirable GDP value or real GDP growth rate. The NBS can falsify nominal data, with implications for real growth rates, by falsifying data on that part of the economy on which it does not regularly publish data, such as small economic units that are covered through surveys only. It can also "falsify" nominal data by expanding its data compilation method to capture an increasing share of the shadow economy. (To its credit, the NBS appears to have implemented such changes to its data compilation method in benchmark revisions that came with retrospective revisions to earlier data.) The NBS can falsify real growth rates through its choice of deflation method, as well as through its choice of retrospective revisions to nominal values and implicit deflators in benchmark revisions.

In the case of real growth rates, neither a double-check on the deflators of VA nor a double-check on the deflators of expenditure components yields significantly different aggregate (GDP) real growth rates. Plausible alternative deflators may yield up to one

percentage point higher as well as lower average annual real growth rates for the reform period. That does not appear to be a clear bias, nor an exorbitant margin of error.

The real growth rates are as good as the deflators and the aggregation methods are. The deflators are as good as the constant price method is or the individual price indices are; the individual price indices, in turn, depend on a vast array of individual price data. All deflation/aggregation methods come with inherent deficiencies, potentially giving too much weight to high-growth sectors due to high initial prices (leading to high GDP growth rates), or neglecting quality change and thus causing deflators to be too high and growth rates to be too low. On balance, there is no clear overall effect. If the underlying individual price data, however, were incorrect, all bets on real growth rates are off. Chinese consumers will readily attest that China's official CPI underestimates true consumer price inflation—but so will European and American consumers for their respective CPIs.

A test of Chinese statistics against Benford's Law does not find egregious data falsification. The only finding is that the second digit of real growth rates is 'zero' with too high a frequency — as is the case for the U.S. That does not rule out the very occasional data manipulation (which Benford's Law cannot detect).

In the end, one is left with an issue of trust: does one trust that the NBS in the many data manipulations that are needed to calculate GDP has no other purpose than to obtain an accurate estimate of GDP? Does the NBS simply work with large margins of error and an atrocious lack of transparency? Or is there a — Benford's Law suggests at most *occasional* — political imperative that guides NBS manipulations of the underlying data? NBS explanations of how it obtains nominal VA and real growth rates are not encouraging as they appear at times inaccurate or contradictory, and the necessary data for double-checks data are not being made available.

Things may not be much better in other countries. In the U.S., the Boskin Commission overnight changed real growth rates on grounds that may be perfectly plausible but only serve to illustrate the wide range of plausible inflation values, and thereby real growth values.[88] In an economy with rapidly changing product characteristics and product variety, such as China, measuring inflation (and thereby real growth) accurately is potentially a challenging task.

Political interference likewise is no anomaly: "President Lyndon Johnson would review the GNP reports before their release. If he did not like the report, he would keep sending the GNP estimates back to the Commerce Department until they got the numbers 'correct.'" Under the first Bush administration, a senior member of the Executive Branch approached a large computer company with the request to in its report to the BEA exaggerate computer sales: "thanks to the heavy leverage of computer deflation, reported GDP growth enjoyed an artificial spike."[89] In contrast to the U.S., data falsification in China—if it occurs at the national level at all—may yet tread on lighter ground.

In the final instance, when political regimes—in the East and in the West—derive their legitimacy from economic growth and when real GDP growth is difficult to measure accurately, reliable GDP figures may be an illusory commodity. When, furthermore, GDP by definition is an incomplete measure of productive activities and a flawed measure of well-being, perhaps it is time to de-emphasize the importance of GDP and to reflect on the ideological framework in which we operate.

References for China literature

Adams, F. G., & Chen, Y. (1996). Skepticism about Chinese GDP growth — The Chinese GDP elasticity of energy consumption. *Journal of Economic and Social Measurement, 22*(4), 231–240.

Bartholdy, K. (1997). Old and new problems in the estimation of national accounts in transition economies. *Economics of Transition, 5*(1), 131–146.

Bureau of Economic Analysis (2013). 2013 comprehensive revision of the national income and product accounts. (Available at). http://www.bea.gov/national/an1.htm#2013comprehensive ((especially first item, "Briefing"), accessed 20 September 2013).

Cai, Y. (2000). Between state and peasant: Local cadres and statistical reporting in rural China. *The China Quarterly, 136*, 783–805.

China Infobank (2003, 7 June). *Yuce zhongguo 'bu ke cesuan de jingji' zhan GDP de bizhong yue wei 10% [The 'economic activities that cannot be estimated' are estimated to be equivalent to 10% of GDP] (at http://www.chinainfobank.com).*

Chow, G. C. (1986). Chinese statistics. *The American Statistician, 40*(3), 191–196.

Economics Dictionary: Statistics (1996). *Jingji da cijian tongjijuan [Big economics dictionary, statistics volume].* Shanghai: Shanghai cishu chubanshe.

Edin, M. (1998). Why do chinese local cadres promote growth? Institutional incentives and constraints of local cadres. *Forum for Development Studies, 1*, 97–127.

Fernald, J., Malkin, I., & Spiegel, M. (2013). On the reliability of Chinese output figures. *FRBSF Economic Letter, 2013–08.*

Finance Ministry (General Office) (1999). *Xianxing caiwu kuaiji zhidu quanshu [Complete current financial accounting system].* Two volumes. Beijing: Zhongguo caizheng jingji chubanshe.

GDP 1952–96 (1998). *Zhongguo guonei shengchan zongzhi hesuan lishi ziliao (zhaiyao) [Historical data on China's gross domestic product 1952–1995 (Summary)].* Beijing: Zhongguo tongji chubanshe.

GDP 1952–2004 (2007). *Zhongguo guonei shengchan zongzhi hesuan lishi ziliao 1952–2004 [Historical data on China's gross domestic product 1952–2004].* Beijing: Zhongguo tongji chubanshe.

GDP Manual (2001). *Zhongguo guonei shengchan zhongzhi hesuan shouce [Manual for GPD calculation in China].* Compiled by the NBS National Income Accounts Division. Beijing: Zhongguo tongjiju guomin jingji hesuansi.

Guo, T., & Yi, X. (2006). Zhongguo dixia jingji guimo jiben cesuan he shizheng fenxi [Basic estimation of the size of China's underground economy and a positive analysis]. *Gongye jishu jingji, 25*(11), 129–132.

Holz, C. A. (2003). 'Fast, clear and accurate:' How reliable are Chinese output and economic growth statistics? *The China Quarterly, 173*, 122–163.

Holz, C. A. (2004). Deconstructing China's GDP statistics. *China Economic Review, 15*(2), 164–202.

Holz, C. A. (2006a). China's reform period economic growth: How reliable are Angus Maddison's estimates? *Review of Income and Wealth, 52*(1), 85–119.

Holz, C. A. (2006b). China's reform period economic growth: How reliable are Angus Maddison's estimates? Response to Angus Maddison's reply. *Review of Income and Wealth, 52*(3), 471–475.

Holz, C. A. (2013a). Chinese statistics: Classification systems and data sources. *Stanford Center for International Development Working Paper No. 471.* Stanford University.

Holz, C. A. (2013b). Chinese statistics: Output data. *Stanford Center for International Development Working Paper No. 473.* Stanford University.

---

[88] As to real growth rates, the Bureau of Economic Analysis in 2013 adopted new practices such as capitalizing R&D expenditures (business R&D was previously categorized as intermediate expenditures), which led to a 0.1 percentage point upward revision to average annual real growth rates of 1929–2012 (0.2 percentage points for 2007–2012).

[89] On deflator and GDP falsification in the U.S., see http://www.shadowstats.com/article/gdp-benchmark-revisions, accessed 24 August 2013.

Holz, C. A. (2013c). Measuring productivity growth in China, 1952–2011. Mimeo, Stanford Center for International Development.
Holz, C. A., & Lin, Y. (2001). Pitfalls of China's industrial statistics: Inconsistencies and specification problems. *The China Review, 1*(1), 29–71.
Huenemann, R. W. (2001). Are China's recent transport statistics plausible? *China Economic Review, 12*(4), 368–372.
Input–output Tables 1997 (1999). *Zhongguo touru chanchubiao 1997 [Input–output tables of China 1997].* Beijing: Zhongguo tongji chubanshe.
Input–output Tables 2007 (2009). *Zhongguo touru chanchubiao 2007 [Input–output tables of China 2007].* Beijing: Zhongguo tongji chubanshe.
Keidel, A. (2001). China's GDP expenditure accounts. *China Economic Review, 12*(4), 355–367.
Klein, L., & Ozmucur, S. (2002–2003). The estimation of China's economic growth rate. *Journal of Economic and Social Measurement, 28*(4), 187–202.
Koen, V. (1996). Russian macroeconomic data: Existence, access, interpretation. *Communist Economies & Economic Transformation, 8*(3), 321–333.
Lardy, N. (2002/2003). Evaluating economic indicators in post-WTO China. *Issues and Studies, 38*(4)/39(1), 249–268.
Li, C. (1962). *The statistical system of communist China.* Berkeley: University of California Press.
Li, D. (2004). Guanyu GDP de ji dian sikao (Some considerations on GDP). *Jingji yanjiu, 4,* 26–28.
Liu, H., Zhang, W., & Liao, F. (2007). Jiyu xianjin bilu moxing gaijin de wo guo dixia jingji guimo cesuan [Estimate of the size of China's underground economy based on an improved cash ratio model]. *Wuhan ligong daxue xuebao, 29*(1), 138–140.
Maddison, A. (1998). *Chinese economic performance in the long run.* Paris: Development Centre of the Organisation for Economic Co-operation and Development.
Maddison, A. (2006). Do official statistics exaggerate China's GDP growth? A reply to Carsten Holz. *Review of Income and Wealth, 52*(1), 121–126.
Mehrotra, A., & Pääkkönen, J. (2011). Comparing China's GDP statistics with coincident indicators. *Journal of Comparative Economics, 39,* 406–411.
Meng, L., & Wang, X. (2000). Dui zhongguo jingji zengzhang tongji shuju kexindu de guji [An evaluation of the reliability of China's statistics on economic growth]. *Jingji yanjiu, 10,* 3–13.
Movshuk, O. (2002). The reliability of China's growth figures: A survey of recent statistical controversies. *The Journal of Econometric Study of Northeast Asia, 4*(1), 31–45.
NBS (National Bureau of Statistics) (1990, Sept. 30). 1990 nian gongye chanpin bubian jiage [1990 constant prices of industrial products]. *In NBS (1995)* (pp. 768–772).
NBS (National Bureau of Statistics) (1995). *Tongji zhidu fangfa wenjian xuanbian 1987–1993 [Selected documents on the methods of the statistical system 1987–1993].* Beijing: Guojia tongji tongji sheji guanlisi (Design and Administration Department of the NBS) (no specific publisher, 1995).
NBS (National Bureau of Statistics) (1997). *Zhongguo niandu guonei shengchan zongzhi jisuan fangfa [Calculation method of China's annual gross domestic product].* NBS National Income Accounts Division. Beijing: Zhongguo tongji chubanshe.
NBS (National Bureau of Statistics) (2003). *Guojia tongji diaocha zhidu 2003 [National statistical survey system 2003].* (No publishing house information or date; presumably an internal volume).
NBS (National Bureau of Statistics) (2007). *Zhongguo jingji pucha niandu guonei shengchan zongzhi hesuan fangfa [Calculation method of China's annual gross domestic product from China's economic census].* NBS National Income Accounts Division. Beijing: Zhongguo tongji chubanshe.
NBS Industry and Transport Division (1989). Gongye zengzhang sudu zhong de 'shuifen' you duoda? [How big is the 'water content' in industrial growth?] *Jingji yanjiu cankao ziliao*(42), 2–5.
NPC (National People's Congress) (1996, May). *Zhonghua renmin gongheguo tongji fa [People's Republic of China Statistics Law].* Revision of original 8 December 1983 law. China Infobank.
NPC (National People's Congress) (2009, June). Zhonghua renmin gongheguo tongji fa [People's Republic of China Statistics Law]. (revision of 15 May 1996 law. At). http://www.npc.gov.cn/npc/xinwen/lfgz/zxfl/2009-06/27/content_1508548.htm (accessed 11 September 2009).
Orlik, T. (2012). *Understanding China's economic indicators: Translating the data into investment opportunities.* Upper Saddle River, NJ: FT Press Science.
Pan, Z., & An, Y. (2003). Yi wan yi de chaju cong he er lai: dui guojiaji, shengji hesuan shuju chaju de sikao [Where is the one-trillion difference from? Some thoughts on the difference between national and provincial accounts data]. *Zhongguo tongji*(11), 8f.
Perkins, D. H. (1966). *Market control and planning in communist China.* Appendix A: Reliability of Chinese statistics. Cambridge, Massachusetts: Harvard University Press, 215–225.
Powers, S. (1992). Statistical needs in Eastern Europe. *Monthly Labor Review, 115*(3), 18–28.
Population Census 2010 (2012). *Zhongguo 2010 nian renkou pucha ziliao [Materials on China's 2010 population census].* Three volumes, compiled by the population census office under the State Council and the NBS Department of Population and Employment Statistics. Beijing: Zhongguo tongji chubanshe.
Rawski, T. G. (1976). On the reliability of Chinese economic data: Discussion. *Journal of Development Studies, 12*(4), 438–441.
Rawski, T. G. (2001a). China by the numbers: How reform has affected China's economic statistics. *China Perspectives, 33,* 25–34.
Rawski, T. G. (2001b). What is happening to China's GDP statistics? *China Economic Review, 12*(4), 347–354.
Rawski, T. G., & Xiao, W. (2001). China economic review symposium on Chinese economic statistics: Introduction. *China Economic Review, 12*(4), 298–302.
Schneider, F., & Enste, D. H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature, 38*(1), 77–114.
Sinclair, T. M. 2012. Characteristics and implications of Chinese macroeconomic data revisions. Mimeo, Department of Economics and Elliott School for International Affairs, The George Washington University.
Wang, L., & Fewsmith, J. (1995). Bulwark of the planned economy: The structure and role of the state planning commission. Chapter 5. In C. L. Hamrin, & S. Zhao (Eds.), *Decision-making in Deng's China: Perspectives from insiders* (pp. 51–64). Armonk, New York: M.E. Sharpe.
Wang, X., & Meng, L. (2001). A reevaluation of China's economic growth. *China Economic Review, 12*(4), 338–346.
Wu, H. X. (2002). How fast has Chinese industry grown?—Measuring the real output of Chinese industry, 1949–97. *Review of Income and Wealth, 48*(2), 179–204.
Wu, H. X. (2007). The Chinese GDP growth rate puzzle: How fast has the Chinese economy grown? *Asian Economic Papers, 6*(1), 1–23.
Xia, X., & Wan, A. (1993). *Zhongguo dixia jingji wenti yanjiu [Study of China's underground economy].* Zhengzhou: Henan renmin chubanshe.
Xu, X. (1999a). Shijie yinhang gaogu zhongguo GDP shuju [The World Bank overestimates China's GDP]. *Zhongguo guoqing guoli, 1,* 7–10.
Xu, X. (1999b). Shijie yinhang dui zhongguo guanfang GDP shuju de tiaozheng he chongxin renke [The official Chinese GDP figures as adjusted and approved by the World Bank]. *Jingji yanjiu, 6,* 52–58.
Xu, X. (2000a). Zhongguo guonei shengchan zongzhi hesuan zhong cunzai de ruogan wenti yanjiu [Some problems in the calculation of China's GDP]. *Jingji yanjiu*(2), 10–16.
Xu, X. (2000b). *Zhongguo guonei shengchan zongzhi hesuan [Calculation of China's GDP].* Beijing: Beijing daxue chubanshe.
Xu, X. (2001). Woguo GDP hesuan yu xianxing SNA de GDP hesuan zhijian de ruogan chayi [Some discrepancies in China's GDP compilation in comparison to the current SNA GDP compilation methods]. *Jingji yanjiu, 11,* 63–68.
Xu, X. (2002). Study on some problems in estimating China's gross domestic product. *Review of Income and Wealth, 48*(2), 205–215.
Xu, X. (2006). Guanyu jingji pucha niandu GDP hesuan de bianhua [Changes in the calculation of annual GDP in the economic census]. *Jingji yanjiu, 2,* 16–20.
Xu, X. (2009). The establishment, reform, and development of China's system of national accounts. *Review of Income and Wealth, 55*(1), 442–465.
Yang, J. (2012). *Tombstone: The great Chinese famine, 1958–1962.* New York: Farrar, Straus and Giroux.
Young, A. (2000). Gold into base metals: Productivity growth in the People's Republic of China during the Reform Period. *NBER Working Paper Series, 7856.*
Zhang, S. (1995, 23 August). *Guanyu 'Zhonghua renmin gongheguo tongjifa xiuzheng'an (zao'an)' de shuoming [Explanation of the (draft) scheduled revisions to the People's Republic of China Statistics Law].* China Infobank.
Zhang, S. (2001). *Zhongguo tongji gaige fazhan zhanlue de lilun yu shijian [Theory and practice of China's statistical system development strategy].* Beijing: Zhongguo tongji chubanshe.
Zhang, Z., Ding, H., & Dan, F. (2004). Zhongguo dixia jingji de guimo guce [An estimate of the size of China's underground economy]. *Lilun xintan*(12), 22f.
Zhao, P. (1998). Nongcun jingji tongji bu shi de yuanyin ji duice [Reasons why rural statistics are false and counter-policies]. *Nongcun Hezuo Jingji Jingying Guanli*(6), 42f.

References for literature on Benford's Law

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society, 78*(4), 551–572.
Browne, M. W. (2013). Following Benford's Law, or looking out for no. 1. (At). http://www.rexswain.com/benford.html (last accessed 15 January 2013).

Cho, W. K. T., & Gaines, B. J. (2007). Breaking the (Benford) Law. *The American Statistician*, *61*(3), 218–223.

Deckert, J., Myagkov, M., & Ordeshook, P. C. (2011). Benford's Law and the detection of election fraud. *Political Analysis*, *19*(3), 235–268.

Diekmann, A. (2007). Not the first digit! Using Benford's Law to detect fraudulent scientific data. *Journal of Applied Statistics*, *34*(3), 321–329.

Durtschi, C., Hillison, W., & Pacini, C. (2004). The effective use of Benford's Law to assist in detecting fraud in accounting data. *Journal of Forensic Accounting*, *V*(1), 17–34.

el Sehity, T., Hoelzl, E., & Kirchler, E. (2005). Price developments after a nominal shock: Benford's Law and psychological pricing after the Euro introduction. *International Journal of Research in Marketing*, *22*(4), 471–480.

Fewster, R. M. (2009). A simple explanation of Benford's Law. *The American Statistician*, *63*(1), 26–32.

Giles, D. E. (2007). Benford's Law and naturally occurring prices in certain eBay auctions. *Applied Economics Letters*, *14*(3), 157–161.

Gonzalez-Garcia, J., & Pastor, G. (2009). Benford's law and macroeconomic data quality. *IMF Working Paper WP/09/10*.

Günnel, S., & Tödter, K. (2009). Does Benford's Law hold in economic research and forecasting? *Empirica*, *36*(3), 273–292.

Hill, T. P. (1995a). Base-invariance implies Benford's law. *Proceedings of the American Mathematical Society*, *123*(3), 887–895.

Hill, T. P. (1995b). The significant-digit phenomenon. *The American Mathematical Monthly*, *102*(4), 322–327.

Hill, T. P. (1995c). A statistical derivation of the significant-digit law. *Statistical Science*, *10*(4), 354–363.

Hill, T. P. (1998). The first digit phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. *American Scientist*, *86*(4), 358–363.

Judge, G., & Schechter, L. (2009). Detecting problems in survey data using Benford's Law. *The Journal of Human Resources*, *44*(1), 1–24.

Leemis, L. M., Schmeiser, B. W., & Evans, D. L. (2000). Survival distributions satisfying Benford's Law. *The American Statistician*, *54*(4), 236–241.

Michalski, T., & Stoltz, G. (2013). Do countries falsify economic data strategically? Some evidence that they might. *Review of Economics and Statistics*, *95*(2), 591–616.

Morrow, J. (2010). *Benford's Law, families of distributions and a test basis.* Mimeo.

Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, *4*(1/4), 39–40.

Nigrini, M. J. (1996). A taxpayer compliance application of Benford's Law. *Journal of the American Taxation Association*, *18*(1), 72–91.

Nigrini, M. J. (2012). *Benford's Law: Applications for forensic accounting, auditing, and fraud detection.* Hoboken, N.J.: John Wiley & Sons.

Nye, J., & Moul, C. (2007). The political economy of numbers: On the application of Benford's Law to international macroeconomic statistics. *The B.E. Journal of Macroeconomics*, *7*(1,17) (14 pp.).

Pinkham, R. S. (1961). On the distribution of first significant digits. *The Annals of Mathematical Statistics*, *32*(4), 1223–1230.

Raimi, R. A. (1969). On the distribution of first significant figures. *The American Mathematical Monthly*, *76*(4), 342–348.

Raimi, R. A. (1976). The first digit problem. *The American Mathematical Monthly*, *83*(7), 521–538.

Rauch, B., Göttsche, M., & Brähler, G. (2011). Fact and fiction in EU-governmental economic data. *German Economic Review*, *12*(3), 243–255.

Smith, S. W. (2013). The scientist and engineer's guide to digital signal processing. (Available at). http://www.dspguide.com/ch34.htm.

Stephens, M. A. (1970). Use of the Kolmogorov–Smirnov, Cramér–Von Mises and related statistics without extensive tables. *Journal of the Royal Statistical Society*, *32*(Series B, Methodological, 1), 115–122.

Thomas, J. K. (1989). Unusual patterns in reported earnings. *The Accounting Review*, *64*(4), 773–787.

Varian, H. R. (1972). Benford's Law (in section Letters to the Editor). *The American Statistician*, *26*(3), 65f.

Weisstein, E. W. (2013). Benford's Law. (At). http://mathworld.wolfram.com/BenfordsLaw.html (last accessed 14 January 2013).