

Research Report

“Fast, Clear and Accurate”: How Reliable Are Chinese Output and Economic Growth Statistics?

Carsten A. Holz

Abstract China’s statistics are widely viewed as unreliable, with data falsification in order to meet economic growth targets increasingly the norm. This report examines some of the most recent criticism of statistics on China’s industrial value-added and Gross Domestic Product, and shows this criticism to be unfounded as it is based on misunderstandings about the meaning and coverage of particular data. A lack of evidence on data falsification does not mean that China’s statistical system is necessarily honest in its statistical reporting, but recent developments in China’s statistical system further suggest that data falsification at the higher levels of the statistical bureaucracy is unlikely. Nevertheless, even if data are not being purposefully falsified by the National Bureau of Statistics, the margin of error in much of the published data is likely to be sufficiently large to allow the statistical authority a choice of final value from a relatively wide range of equally correct values.

Early evaluations of Chinese statistics were largely positive. Thus Li Chow-Ming in 1962 wrote cautiously about improvements following the statistical débâcle during the Great Leap Forward. Dwight Perkins in 1966 concluded that falsification of disaggregated data is highly improbable; in the case of aggregated data, falsification might remain unnoticed in the short run, but not in the long run, and in the end it may not be in the interest of the leadership. Thomas Rawski in 1976 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.” Gregory Chow in 1986 judged that “by and large Chinese statistics officials are honest.”¹

Yet in 1989 the National Bureau of Statistics (NBS) found it necessary to conduct a large-scale investigation of statistical work across the country, the first of four such investigations so far.² By the late 1990s Western researchers began to pay renewed attention to the quality and reliability of Chinese statistics. Some examples are the articles by

1. Li Choh-Ming, *The Statistical System of Communist China* (Berkeley: University of California Press, 1962); Dwight H. Perkins, *Market Control and Planning in Communist China* (Cambridge, MA: Harvard University Press, 1966), pp. 216f; Thomas G. Rawski, “On the reliability of Chinese economic data: discussion,” *Journal of Development Studies*, Vol. 12, No. 4 (1976), pp. 438–441 (quote from p. 440); Gregory C. Chow, “Chinese statistics,” *The American Statistician*, Vol. 40, No. 3 (August 1986), pp. 191–96 (quote from p. 193).

2. In 1998 the *Zhongguo tongjiju* changed its English name from State Statistical Bureau to National Bureau of Statistics (NBS). This report uses the current name NBS throughout.

Thomas Rawski and Robert Mead in 1998 on employment statistics, Cai Yongshun in 2000 on rural statistics, and Carsten Holz and Yi-min Lin in 2001 on industrial statistics.³ China’s prime minister Zhu Rongji publicly demanded “fast, clear, and accurate” (*kuai, jing, zhun*) statistics.⁴

Output data have come under particular scrutiny. Frequent falsification of such data in the countryside now appears beyond doubt. Cai Yongshun documents extensive upward falsification of rural output statistics by local leaders concerned about their promotion;⁵ Meng Lian and Wang Xiaolu provide similar evidence.⁶ Numerous articles in *Zhongguo tongji* (*China Statistics*), a magazine published by the NBS, give examples of data falsification by local leaders and reporting units. Following the 1995 industrial census, the NBS revised nation-wide output data of collective-owned and individual-owned enterprises for the years 1991–94 retrospectively downward by up to one-quarter.

Most recently the integrity of the NBS itself has been questioned. Since the early 1990s it may have come under increasing pressure to falsify aggregate nation-wide output data, such as industrial value-added and Gross Domestic Product (GDP), in order to meet the leadership’s continuously ambitious economic growth targets. The literature on Chinese statistics uses two procedures to show the falsification of output data. First, it compares output growth to the growth in other variables which one would expect to be strongly correlated with output. For example, Meng Lian and Wang Xiaolu contrast the growth rate of industrial value-added with the growth rate of industrial energy use, the average growth rate in the production quantity of 168 major products, and the growth rate of freight.⁷ They find that in particular since 1992, industrial value-added grew much more rapidly than any of these three control variables, and conclude that the growth rate of industrial value-added, and by extension GDP growth (industrial value-added accounts for almost

3. Thomas G. Rawski and Robert W. Mead, “On the trail of China’s phantom farmers,” *World Development*, Vol. 26, No. 5 (May 1998), pp. 767–781; Cai Yongshun, “Between state and peasant: local cadres and statistical reporting in rural China,” *The China Quarterly*, No. 163 (Sept. 2000), pp. 783–805; Carsten Holz and Yi-min Lin, “Pitfalls of China’s industrial statistics: inconsistencies and specification problems,” *The China Review*, Vol. 1, No. 1 (Fall 2001), pp. 29–71; Carsten Holz and Yi-min Lin, “The 1997–1998 break in industrial statistics: facts and appraisal,” *China Economic Review*, Vol. 12, No. 4 (2001), pp. 303–316.

4. NBS Division of Industry and Transportation, *Xinbian gongye tongji gongzuo zhinan* (*New Guide to Industrial Statistics*) (Beijing: Zhongguo tongji chubanshe, 1999), p. 3.

5. Zhao Peng, “Nongcun jingji tongji bu shi de yuanyin ji duice” (“Reasons why rural statistics are false and counter-policies”), *Nongcun hezuo jingji jingying guanli* (*Rural Co-operative Economy Management*), No. 6 (1998), pp. 42f, lists as reasons for data falsification in the countryside that leaders by nature want to only hear good news, that they are evaluated (promoted and remunerated) according to the economic performance of the locality, and that once data in one year have been falsified, going back to accurate data in the following years is almost impossible. On the incentive structures of local cadres also see Maria Edin, “Why do Chinese local cadres promote growth? Institutional incentives and constraints of local cadres,” *Forum for Development Studies*, Vol. 9, No. 1 (1998), pp. 97–127.

6. Meng Lian and Wang Xiaolu, “Dui Zhongguo jingji zengzhang tongji shuju kexindu de guji” (“An evaluation of the reliability of China’s statistics on economic growth”), *Jingji yanjiu* (*Economic Research*), No. 10 (October 2000), pp. 3–13.

7. *Ibid.*, and Wang Xiaolu and Meng Lian, “A reevaluation of China’s economic growth,” *China Economic Review*, Vol. 12, No. 4 (2001), pp. 338–346.

half of GDP), has been drastically exaggerated in recent years.⁸ Similarly, Gerard Adams and Chen Yimin contrast China's elasticity of energy consumption with respect to GDP (or industrial value-added) with that of other countries, and also come to the conclusion that China's GDP growth from the late 1970s to the mid-1990s has been exaggerated.⁹

Secondly, the quality of aggregate nation-wide GDP data can be checked by contrasting the results of three different approaches to derive aggregate nation-wide GDP: as the sum of value-added created across all sectors (production approach); as the sum of income across all sectors; and as the sum of expenditures (consumption, investment, government expenditures and net exports). In the derivation of the official aggregate nation-wide GDP, the NBS relies on the production approach for most economic sectors and on the income approach for some residual sectors. One alternative method, chosen by Thomas Rawski, is to piece together aggregate nation-wide GDP by relying solely on the income approach.¹⁰ Rawski concludes that real GDP growth in 1998 was not the official 7.8 per cent, but 5.7 per cent or even less. He believes that the "wind of falsification and embellishment" as noted in Chinese-language articles has been blowing particularly strongly since 1998, and he hints at intentional falsification of China's economic growth data (which would have to be done or at least condoned by the NBS). The doubts about China's most recent growth estimates are repeated in popular news sources such as *The Economist* (7 March 2001) and the *Asian Wall Street Journal* (22 November 2001), which suggest that the year 2000 GDP growth rate was overstated by one or two percentage points.

This report does not question the evidence on data falsification in the countryside; such evidence appears overwhelming. What it does is

8. Meng Lian and Wang Xiaolu, "An evaluation," find that prior to 1992 industrial value-added grew roughly in line with the other three indicators, but since 1992 much faster. They then estimate a production function for industry in which they measure total factor productivity growth separately for the years 1953 to 1997, 1978 to 1997 and 1992 to 1997. They conclude that the much higher total factor productivity growth since 1992, compared to the previous two periods, is a sign of exaggerated output data. Yet the link between total factor productivity and falsification of output data is far from logically compelling. Li Xiao-Ming, "The great leap forward, economic reforms, and the unit root hypothesis: testing for breaking trend functions in China's GDP data," *Journal of Comparative Economics*, Vol. 28, No. 4 (Dec. 2000), pp. 814–827, using rather sophisticated time series analysis methods to look for breaks in the development of GDP and its three sectoral components in the years 1953 to 1998, finds a 1960 and 1982 break for GDP, and a 1960 and 1991 break for industrial value-added. Meng and Wang's hypothesized 1992 break thus does not hold for GDP in total. Li's 1991 break for industrial value-added comes close to Meng and Wang's 1992 break, but Li argues that it is due to Deng Xiaoping's exhortation in 1992 for the country to quicken the reform pace and to invest and grow.

9. Gerard F. Adams and Chen Yimin, "Skepticism about Chinese GDP growth – the Chinese GDP elasticity of energy consumption," *Journal of Economic and Social Measurement*, Vol. 22, No. 4 (1996), pp. 231–240. They rule out the possibility that sectors with low energy intensity grew faster than those with high energy intensity, or that the efficiency of energy use has improved significantly over time. Thomas G. Rawski, "What's happening to China's GDP statistics," *China Economic Review*, Vol. 12, No. 4 (December 2001), pp. 347–354, also questions China's GDP data based on the severe discrepancy between GDP growth and the growth in energy use.

10. Thomas G. Rawski, "China by the numbers: how reform affected China's economic statistics," *China Perspectives*, No. 33 (January/February 2001), pp. 25–34.

question the evidence on the falsification of aggregate nation-wide output and economic growth data. The following two sections show that the meaning and coverage of the variables that are being used to double-check on the growth of industrial value-added are regularly misunderstood. Rawski's calculation of GDP via the income approach is only a partial calculation due to a lack of data; the resulting growth figure is little different from the official data once the later NBS revisions to the official data have been considered. What is left is problems in understanding the meaning and coverage of various Chinese statistics, and possibly technical difficulties for the NBS in collecting accurate statistics, but no evidence of data falsification at the national level.

But a lack of evidence on data falsification does not prove that the NBS does not report falsified aggregate nation-wide data. While one compelling piece of evidence of purposeful data falsification would be sufficient to establish wrongdoing by the NBS, a positive proof of honest data reporting is impossible. This report in the fourth section therefore proceeds, as a second step, with a plausibility argument; it argues why the NBS is unlikely to falsify data, and suggests that it has been making sincere efforts to improve data quality over time.

Questioning the Evidence on Falsification of Industrial Value-added

The following three sub-sections investigate the claim that the discrepancy in the growth rate of industrial value-added and the growth rates of three control variables (industrial energy use, production quantity of major products and freight) suggests falsification of industrial value-added data and, by extension, of GDP. The main counter-argument across all three cases is that the meaning and coverage of the control variables is misunderstood in the literature; in other words, control variables do not refer to the same set of production activities as the output data. Each sub-section begins with the available definitions of the control variable. (Some of the published definitions of the same variable contradict each other.) This is followed by a range of numerical double-checks on the probable coverage of the control variable, and then a brief summary conclusion as to why the particular control variable used in the literature as a double-check on industrial value-added is inappropriate.

One recurring theme is that the published data on the control variables may cover no more than a subset of all industrial enterprises. According to a lengthy manual on compiling industrial statistics, published by the industrial and transportation division of the NBS in 1999 and addressed to statistical personnel in charge of industry statistics, the NBS currently compiles five comprehensive annual reports. These cover aggregate output value and sales revenue (including product quantities), enterprise finances (including balance sheet and profit and loss account information), industrial energy use, technology levels, and product quantities of 235 products including production capacity. The first three reports cover only the directly reporting industrial enterprises; the last two are compiled by central line ministries or other relevant State Council

departments to cover their *xitong* – the government administration in charge of enterprises in a specific sector – only.¹¹ The periodic reporting system that forms the backbone of the published statistics thus does not cover industrial energy use and product quantities of *all* industrial enterprises. Data on those enterprises that do not directly report or belong to a central *xitong* thus are either not available to the NBS or obtained through surveys or “guesstimates.”¹²

The term “directly reporting industrial enterprises” prior to 1998 refers to all industrial enterprises with an independent accounting system at township level and above. Since 1998, it refers to all industrial state-owned enterprises (SOEs) with an independent accounting system plus all industrial non-SOEs with an independent accounting system and with annual sales revenue in excess of 5 million *yuan*. The pre-1998 administrative criterion for inclusion has thus been replaced by a size criterion. Due to the 1998 statistical break, 1998 and 1999 data on the directly reporting industrial enterprises are not comparable to pre-1998 data.¹³

Industrial energy use. The *Statistical Yearbook 2000* in a note to the overall nation-wide energy balance states that industrial energy consumption (or “use”) includes data on village-run industry.¹⁴ This statement could imply that energy consumption by at least private and individual-owned industrial enterprises is ignored. These units, however, increased their share in Gross Output Value of all Industry (GOVI) from 5.39 per cent in 1990 to 12.86 per cent in 1995 and 18.18 per cent in 1999 (industrial value-added data on these units are not available).¹⁵ By 1999, therefore, energy consumption of units responsible for almost one-fifth of industrial production may not have been accounted for in the official energy statistics.

On the other hand, the NBS in a statistical compendium on freight and energy use for the years 1949 to 1999 published identical energy data and claimed that the data on industrial energy consumption cover two groups

11. NBS Division of Industry and Transportation, *Xinbian gongye tongji*, pp. 4f. Industrial enterprises in a particular *xitong* are likely to all be state-owned enterprises, except in the second-light industry *xitong* (in as far as it still exists). State-owned enterprises are a sub-category of the directly reporting industrial enterprises.

12. For those industrial enterprises that are not “directly reporting industrial enterprises” and for all individual-owned enterprises (*getihu*), which by definition in the official statistics are not regarded as “enterprises,” the county-level statistical departments compile an annual report on the number of such enterprises, their gross output value, tax payments, equity (capital) and the year-end number of labourers. If gross output value data are not available, business income can be substituted; in general, estimations based on a “representative” sample are permissible (*ibid.* p. 38).

13. The statistical coverage of *all* industrial enterprises also changed in 1998, but the difference is likely to be exceedingly small. Prior to 1998, all industry comprised all industrial units, regardless of type of accounting system and subordination. Since 1998, all industry no longer comprises the industrial units with dependent accounting system which are part of a non-industrial unit. For more details see Holz and Lin, “The 1997–1998 break in industrial statistics.”

14. NBS, *Zhongguo tongji nianjian 2000* (*China Statistical Yearbook 2000*, in the following abbreviated as *Statistical Yearbook 2000* and similarly for other years) (Beijing: Zhongguo tongji chubanshe, 2001), p. 240.

15. *Statistical Yearbook 2000*, p. 409.

of enterprises.¹⁶ The first group consists of the directly reporting industrial enterprises, for which data are directly obtained from their statistical reports; the second group comprises all other industrial enterprises, with data obtained from the Township and Village Enterprise Bureau of the Agricultural Ministry combined with calculations based on their “development speed” (*fazhan sudu*).¹⁷ According to this definition, the energy data published in the *Statistical Yearbook* cover all industry.

A third, different, definition was offered by the chief statistician of the NBS, Long Hua, on 27 July 1994.¹⁸ At a conference of statistical department staff on the 1995 statistical work he explained that in the future data on energy consumption are to cover only industrial, construction and transportation enterprises that are located at county level or above.

The 1995 industrial census allows a first clarification as to the enterprise coverage of industrial energy consumption data. Such data in this census are explicitly limited to the directly reporting industrial enterprises, at that time the industrial enterprises with an independent accounting system at township level and above (that is, including one level below the county tier, but not the village level.)¹⁹ For coal consumption in 1995, the figure in the industrial census is equivalent to 99.56 per cent of the amount given in the *Statistical Yearbook* (which could be for all industrial enterprises, or all industrial enterprises except private and individual-owned units), for coke consumption the share is 88.53 per cent and for electricity consumption 169.52 per cent; if the industrial census data are viewed as more reliable, then these data question the consistency of the data on energy consumption reported in the *Statistical Yearbook*.²⁰ Furthermore, if even the industrial census did not collect energy consumption data for village-level, private and individual-owned industrial enterprises, then, in the absence of energy use surveys, any data on the energy use of

16. NBS Division of Industry and Transportation, *Zhongguo gongye jiaotong nengyuan 50 nian tongji ziliao huibian 1949–1999* (China Industry, Freight and Energy 50 Year Statistical Compilation 1949–1999) (Beijing: Zhongguo tongji chubanshe, 2000), p. 254.

17. Data on energy production are collected in the annual industrial product quantity report (*ibid.* p. 254). But this implies that since these product quantity reports cover only a subset of all production units, economy-wide energy data are likely to be underestimated. A minor issue is that total energy consumption (or production) excludes low energy combustibles (*direzhi ranliao*), presumably such items as wood, and also bioenergy and solar energy (*ibid.* p. 299).

18. Long Hua, “Guanyu 1994–1995 nian tongji zhidu fangfa gaige he shuju chuli fang’an de shuoming” (“Some explanations on the 1994–1995 reform of the statistical system and the handling of data”), 27 July 1994, pp. 1024–39 in *Tongji gongzuo zhongyao wenjian xuanbian 1991–1995* (Selected Major Documents On Statistical Work 1991–1995) (Beijing: Zhongguo tongji chubanshe, 1996).

19. NBS, *Zhonghua renmin gongheguo 1995 nian di san ci quanguo gongye pucha ziliao huibian* (Material on the PRC’s Third Industrial Census in 1995, in the following abbreviated as *Industrial Census 1995*) (Beijing: Zhongguo tongji chubanshe, 1997, three volumes), Vol. 1, p. 286.

20. Data are also available for a number of smaller energy items, but no “total” is available; the percentages are 100.46% for crude oil, 79.38% for gasoline, 72.25% for kerosene, 88.15% for diesel oil, 96.76% for fuel oil, and 110.66% for natural gas. Coal consumption (*meitan xiaofeiliang* in the *Statistical Yearbook*) in the industrial census is calculated as the sum of raw coal (*yuanmei*) and two types of washed coal (*xijingmei* and *xizhongmei*) consumption.

these enterprises would be pure guesswork. If the industrial energy consumption reported in the *Statistical Yearbook* were indeed to cover the whole economy, it would thus be a rather inaccurate figure.

For more recent years, provincial-level data allow a second check on the coverage of industrial energy consumption. Going through the 31 provincial statistical yearbooks published in the year 2000 with 1999 data, ten provinces report total industrial energy consumption (TIEC) in the energy section of the yearbook together with either a ratio that links energy use to output (the ratio of industrial energy use to industrial gross output value), or a direct output value.²¹ The explicit or implicit industrial output value given in the energy section as matching provincial TIEC can be checked against the industrial output value of different sets of industrial enterprises in the industry section of the yearbook. (See Table 1; output values from the industry section of the provincial statistical yearbooks that match the implicit or explicit ones in the energy section are in italics.)

In Hubei, Shaanxi and Qinghai, the reported TIEC in fact covers only directly reporting industrial enterprises, and in Xinjiang only industrial enterprises on township level and above. For these four provinces the match between the output value implied by the industrial energy use data and the independently reported output data of a particular sub-group of all industrial enterprises is perfect (in the case of Qinghai very close to perfect). In Yunnan, the reported TIEC is very probably also limited to directly reporting industrial enterprises. (The gross output value of these enterprises in 1990 prices, needed for a conclusive comparison, is not available. Some provincial statistical yearbooks note in their energy section that the corresponding gross output value is in 1990 prices; none explicitly matches energy consumption with gross output value using current prices.) In Jilin and Guangdong, reported TIEC appears to cover only a sub-category of all directly reporting industrial enterprises. Partial data on Jiangxi (see note to Table 1) suggest that TIEC in Jiangxi covers only the directly reporting industrial enterprises. In Shanxi, TIEC comes with an output value that exceeds that of directly reporting industrial enterprises, but falls significantly short of covering all industrial enter-

21. Sichuan data available in the year 2000 provincial statistical yearbook cover the year 1998, and therefore are not included in the 1999 analysis. It is never explicitly stated whether the ratio of industrial energy use to industrial gross output value is based on total industrial energy consumption, or on final industrial energy consumption. When available, total industrial energy consumption was used together with the ratio to derive the implicit industrial gross output value (this is the case for nine out of the ten provinces in 1999); otherwise final industrial energy consumption was used. For provinces for which data on both TIEC and final industrial energy consumption are available, the difference between the two data points is often (but not always) minor (minor with respect to the differences in reported output of various enterprise categories to which the output data derived from the energy data are compared). The provincial statistical yearbooks of Inner Mongolia and Chongqing provide detailed energy consumption by industrial sector with the corresponding implicit or explicit output value only for "key" enterprises (in Inner Mongolia defined as those consuming more than 10,000 tons of SCE per year). For Inner Mongolia no data on TIEC are available, while in Chongqing total industry consumes twice as much energy as the key industrial enterprises.

prises. In Guizhou and Ningxia, reported TIEC could be covering all industrial enterprises (figures on GOVI in 1990 prices are not available).²²

Neither the *Statistical Yearbook* nor the 1949–99 compendium on freight and energy use provide provincial data on industrial energy use, and not all provincial yearbooks contain an energy section. A direct comparison of nation-wide TIEC with the sum of the provincial TIEC is thus not possible. A round-about calculation is the following All industrial enterprises in the eight provinces (including Jiangxi) in which TIEC is available and covers only the directly reporting industrial enterprises, or some other subset of all industrial enterprises, accounted for 24.80 per cent of nation-wide industrial value-added in 1999.²³ Adjusting the TIEC of the eight provinces (of 253.681 m tons of SCE) upwards corresponding to their share in nation-wide industrial value-added yields nation-wide TIEC of 1,022.907 m tons of SCE ($253.681 / 0.2480$) in 1999. This is an approximate measure of the TIEC of the directly reporting industrial enterprises (or some other subset of all industrial enterprises) nation-wide.

Yet the *Statistical Yearbook* reports a nation-wide industrial TIEC value of only 907.97 m tons of SCE for *all* industrial enterprises. With the directly reporting industrial enterprises in 1999 accounting for only 61.66 per cent of industrial value-added, the economy-wide value of TIEC should be much larger, around 1,658.95 m tons of SCE ($1,022.907 / 0.6166$).²⁴ This is 82.71 per cent higher than the official nation-wide TIEC value reported in the *Statistical Yearbook*. The same pattern holds in 1997 and 1998, with the summary data presented in Table 2.²⁵

22. In Shanxi, Guangdong, Shaanxi, Qinghai and Xinjiang, data on industrial energy consumption provided in the provincial energy balance, or in a table that appears to be an abbreviated provincial energy balance table, fully or very closely match data provided for aggregate industry in a separate energy table which gives the link to output data and a breakdown of industrial energy consumption by industrial sector; in other words, for these provinces the value of TIEC reported in Table 1 is definitely claimed to cover all industry in this province. For Jilin, Yunnan and Ningxia no provincial energy balance is available, and for Hubei TIEC is only provided in the provincial energy balance. With Guizhou being one of the provinces in which TIEC matches all industry, the fact that the TIEC included in the overall energy balance is much smaller than the industrial energy use reported in the industry table of the energy section with the link to output – and TIEC very probably is the value that forms the basis for the nation-wide TIEC reported in the *Statistical Yearbook* – only strengthens the argument made in the text below.

23. *Statistical Yearbook 2000*, p. 61.

24. On the share of the directly reporting industrial enterprises in industrial value added see *Statistical Yearbook 2000*, pp. 53, 414.

25. The same pattern also holds if the industrial energy use data of the provinces in which TIEC covers only the directly reporting industrial enterprises are adjusted upward to obtain the estimate of directly reporting industrial enterprises' nation-wide TIEC using a different variable for the adjustment, namely the provincial value-added of the *directly reporting industrial enterprises* rather than the industrial value-added of the whole province. Thus, the directly reporting industrial enterprises in the eight provinces, including Jiangxi, in which TIEC is available and covers only the directly reporting industrial enterprises, or some other subset of all industrial enterprises, account for 27.314% of nation-wide value-added of directly reporting industrial enterprises; adjusting the TIEC of the eight provinces of 253.681 m tons of SCE upward corresponding to their share in nation-wide industrial value-added yields nation-wide TIEC of 928.755 m tons of SCE in 1999. The match with official TIEC in this case is even better (Table 2).

Table 1: Industrial Enterprise Coverage of Industrial Energy Consumption in 1999 Across Provincial Statistical Yearbooks

	Shanxi ^e	Jilin	Hubei	Guangdong	Guizhou ^f	Yunnan	Shaanxi	Qinghai ^h	Ningxia	Xinjiang
Energy section: industry										
Final ind. energy consumption ^a (10,000t SCE)	4,035.42	2,382.94	3,908.99	5,364.97	1,417.98	—	—	668.17	—	—
Total ind. energy consumption ^a (10,000t SCE)	4,958.07	2,588.49	4,489.03	5,670.76	2,415.72	2,388.30	1,898.29	—	1,086.71	1,854.69
Ratio of ind. energy consumption to GOV ^a	3.73	2.34	1.79 ^d	1.23	3.75	3.2	2.3	6.94	4.78	4.26 ^d
Implied GOV ^b (100 m yuan)	1,329.24	1,106.19	2,507.84	4,610.37	644.19	746.34	891.22	96.28 ^d	227.34	435.37
GOV reported in energy section	1,106.71	—	—	—	—	—	889.68 ^d	—	—	—
Industry section (100 m yuan)										
Directly reporting industrial enterprises ^c	A ^e	B	C	D	D	F	D	C	C	C
GOV (current prices)	1,096.83	1,366.92	2,831.70	10,538.17	551.93	1,002.88	1,042.58	160.77	197.66	631.84
GOV (1990 prices)	857.14	1,212.97	2,508.03	10,376.36	467.30	Match?	889.68	100.52	139.18	370.94
Value-added (current prices)	400.65	412.00	945.54	2,787.75	—	505.46	347.85	58.32	—	256.58
All industry	—	—	—	15,303.33	741.89	1,561.08	1,446.18	207.97	252.39	802.09
GOVI (current prices)	1,934.37	—	—	13,944.36	Match?	1,018.93	—	134.08	Match?	—
Value-added	650.62	552.34	1,692.46	3,705.88	284.16	680.01	487.26	69.96	80.70	318.00
Ind. enterprises on township level and above	—	—	—	—	—	—	—	—	—	—
GOV (1990 prices)	—	—	—	—	—	—	—	—	—	435.52

Notes:

SCE: Standard Coal Equivalent; GOV: Gross output value of industry (in total).

^a The Chinese terms are the following: final [industrial energy] consumption (*zhongdian xiaofei*), total [industrial energy] consumption (*nengyuan xiaofei zongliang*), ratio of energy use to GOV (*chanzhi nenghao*, or *mei wan yuan gongye zongchanzhi nengyuan xiaofei liang*). The unit of the ratio is tons of standard coal equivalent units per 10,000 yuan of GOV. This ratio together with total industrial energy consumption is used in the next line in the table to calculate the corresponding GOV.

b The industrial gross output value implicit in the energy tables is derived by dividing total industrial energy consumption by the ratio of energy use to gross output value; when total industrial energy consumption data are not available, final industrial energy consumption is used.

c Directly reporting industrial enterprise labels in specific provincial statistical yearbook (presumably all covering the same category of enterprises):

A: SOEs, large and medium-sized enterprises, and non-SOEs with annual sales revenue in excess of 5 m *yuan*.

B: Industrial enterprises above 5 m *yuan*.

C: All industrial SOEs and industrial non-SOEs with annual sales revenue in excess of 5 m *yuan*.

D: Above-norm (*guimo yishang*) industrial enterprises.

E: All SOEs and above-norm industrial enterprises.

F: All SOEs and all industrial non-SOEs with independent accounting system and sales revenue in excess of 5 m *yuan*.

d The source specifies that GOV is in 1990 prices.

e Shanxi's directly reporting industrial enterprises do not include industrial enterprises on village level and below.

f For comparison, gross output value of industrial SOEs, state-controlled shareholding companies, and collective-owned enterprises in Guangdong was 4,299.43*10,000 *yuan* in current prices, and 4,112.93*10,000 *yuan* in 1990 prices.

g Guizhou's total industrial energy consumption in the table listing energy use of each industrial sector and linking it to output is more than 70% above industrial energy consumption in the provincial energy balance table, and approximately equal to total provincial energy consumption.

h Qinghai's calculated GOV (from the energy consumption data) is based on final consumption. If total consumption data were available, GOV would be slightly higher.

Additional province Jiangxi: Industry consumption of coal (*meitan*) in 1999 was 22,2495 m tons as listed in the total coal consumption table (which carries no note on enterprise coverage). At the same time, according to a separate table with a note that it covers all industrial "enterprises with independent accounting system on township level and above plus all village and below-village level [industrial] enterprises with an annual sales revenue in excess of 5 m *yuan*," industrial consumption of raw coal (*yuanmei*) was 17,8113 m tons, of washed coal (*xijing mei*) 2,8082 m tons, of "other" washed coal (*qita xi mei*) 0.1752 m tons, and of coke 1,9103 m tons, which adds up to a total of 22,7050 tons of coal (pp. 124, 132f). In other words, total industrial coal consumption data appear limited in their coverage to directly reporting industrial enterprises, which may then also hold for total energy consumption.

Sources:

(Year 2000 particular provincial statistical yearbook): Shanxi: pp. 17, 142, 156, 158, 235, 245; Jilin: pp. 39, 103-5, 205, 217; Hubei: pp. 20, 106, 117, 196f; Guangdong: pp. 80, 196f, 200, 313, 315, 339; Guizhou: pp. 100, 111, 216, 218 (and *Statistical Yearbook 2000*, p. 61); Yunnan: pp. 237f, 241, 303; Shaanxi: pp. 52, 162, 167, 304f, 310; Qinghai: pp. 16, 122, 130, 207-9; Ningxia: pp. 37, 90f, 155, 158; Xinjiang: pp. 35, 361, 363, 368, 556, 569.

Table 2: Energy Consumption of Directly Reporting Industrial Enterprises Across Provinces versus Nation-wide Industrial Energy Consumption

	1997	1998	1999
Number of provinces for which matching energy consumption and industrial gross output value data are available	12	11	10
Number of provinces in which the energy section appears to cover only the directly reporting industrial enterprises, or some other sub-category of all industry	8	8	9
Of these, total industrial energy consumption (TIEC) data are available for the following number of provinces	7	6	8
Their combined TIEC is the following number of million tons of SCE (a)	165.915	189.824	253.681
The industrial value-added of these provinces is equivalent to the following percentage of the nation-wide total (b)	16.232	12.811	24.800
If all other provinces had the same ratio of industrial value-added to TIEC of the directly reporting industrial enterprises (or some other, similarly limited subset of all industry), all provinces together would have a combined TIEC of the following number of million tons of SCE; this figure approximately reflects nation-wide TIEC of the directly reporting industrial enterprises (a/b)	1,022.140	1,481.675	1,022.907
Actual nation-wide TIEC as reported in the <i>Statistical Yearbook</i> and defined in a note with the data as covering all industry (except, perhaps, the private and individual-owned economy), in million tons of SCE	1,000.80	944.09	907.97

A similar calculation can be made based on the industrial value-added of the directly reporting industrial enterprises:

The industrial value-added of the directly reporting industrial enterprises in these provinces is equivalent to the following percentage of the corresponding nation-wide total (c)	16.765	13.894	27.314
If all other provinces had the same ratio of industrial value-added of directly reporting industrial enterprises to TIEC of the directly reporting industrial enterprises (or some other, similarly limited subset of all industry), all provinces together would have a combined TIEC of the following number of million tons of SCE; this figure approximately reflects nation-wide TIEC of the directly reporting industrial enterprises (a/c)	989.633	1,366.259	928.755

Notes:

In 1997, the energy tables of three out of the twelve provinces for which the energy section provides matching output data cover only the directly reporting industrial enterprises (in two provinces this conclusion is based on final rather than total industrial energy consumption), and in five provinces only the industrial enterprises on township level and above; in four provinces the industrial output data in the energy section match output of all industry. In 1998, in seven out of the eleven provinces the energy tables cover only the directly reporting industrial enterprises (in one province this conclusion is based on electricity consumption), in one province only the industrial enterprises on township level and above (where this conclusion is based on final rather than total industrial energy consumption); in three provinces the industrial output data in the energy section match output of all industry. For 1999, see the text.

The seven provinces in 1997 for which TIEC (rather than only final industrial energy consumption) and matching output data are available in the energy section are Jilin, Fujian, Hubei, Guangxi, Yunnan, Shaanxi and Xinjiang. The six provinces in 1998 are Shanxi, Jilin, Hubei, Yunnan, Shaanxi and Xinjiang. For 1999 see the text.

Sources:

Statistical Yearbook 2000, p. 240; 2001, p. 230; Table 1, and similar tables for the years 1997 and 1998, which are available upon request.

The claim of the NBS in the 1949–99 compendium on freight and energy use that the TIEC data published – which are identical to those in the *Statistical Yearbook* – cover all industry thus can be refuted with near certainty. If the statement in the *Statistical Yearbook* that the coverage extends to village-level units implies that it does *not* extend to private and individual-owned units, then the comparison of TIEC with industrial output is immediately invalid. However, private and individual-owned units accounted for only 18.18 per cent of GOVI in 1999. The TIEC reported in the *Statistical Yearbook*, given its much larger shortfall of the estimated true aggregate nation-wide TIEC, is thus unlikely even to cover the set of all industrial enterprises at village-level and above.²⁶ The most likely coverage of the official TIEC (used by the critics) is the one suggested by the provincial-level data and by professional statistical practice: the directly reporting industrial enterprises, that is, those enterprises from which the statistical departments collect periodic reports and thus have available somewhat reliable data.²⁷

Product quantities. An argument similar to that for energy holds for the “output of major industrial products” reported in the *Statistical Yearbook*. The NBS in the introduction to the industry section of the *Statistical Yearbook* claims that it takes the product quantity data it receives from the directly reporting industrial enterprises and makes upward adjustments in order to account for all other enterprises.²⁸ But the NBS in a 1997 circular addressed to the provincial statistical bureaus, in which it explained the 1998 changes in the industrial statistical reporting system, states that the product quantity statistics “in principle” follow the new classification of the directly reporting industrial enterprises. Each province can go beyond this scope of enterprises for those products that are very important locally. Whatever choice is made, it should be maintained throughout 1998, and the quantities reported up to the NBS need not be adjusted.²⁹ This implies that the product quantities reported in the *Statistical Yearbook* could for one product cover total production across all enterprises in all provinces, or only production in directly reporting industrial enterprises across all provinces,

26. The private and individual-owned industrial enterprises furthermore account for a rapidly growing share of industrial output, and any comparison of aggregate industrial output with industrial energy use data that does not cover these units becomes progressively inappropriate over time.

27. Jonathan E. Sinton, “Accuracy and reliability of China’s energy statistics,” *China Economic Review*, Vol. 12, No. 4 (2001), pp. 373–383, severely questions the quality of energy data on a number of technical grounds, including a highly problematic compilation of coal statistics. The reliability of energy data appears to be particularly questionable in the most recent years. Jeffrey Logan, “Diverging energy and economic growth in China: where has all the coal gone?” *Pacific and Asian Journal of Energy*, Vol. 11, No. 1 (June 2001), pp. 1–13, raises similar questions about the quality of the reported energy data.

28. *Statistical Yearbook 2000*, p. 405.

29. NBS, “Guanyu tiaozheng gongye tongji baobiao zhidu bufen neirong de tongzhi” (“Circular on adjusting some of the content of the industrial statistical reporting system”), 6 October 1997, in NBS, *Zhongguo tongji gongzuo nianjian 1998* (*China Statistical Work Yearbook 1998*, in the following abbreviated as *Statistical Work Yearbook 1998* and similarly for other years) (Beijing: Zhongguo tongji chubanshe, 1998), pp. 123f.

or different sets of enterprises in different provinces; the enterprise coverage of different products need not be the same even within one province.

An employee of the Jiangxi statistical bureau confirms that only the directly reporting industrial enterprises (in 1998) regularly report product quantities. Thus they report the product quantities of 400 products in their monthly reports (up from 100 previously, presumably until 1997) and of 1,700 products in their annual reports. Yet the statistical bureau then is supposed somehow to compile the total product quantities of a group of 40 products across all industrial enterprises, a task that the employee views as close to impossible.³⁰ This implies that if any enterprise other than the directly reporting industrial enterprises were included in the product quantity statistics, their product quantities are at best “guesstimates.”

The *Industrial Yearbook* finally contains product quantity data which are identical to those reported in the *Statistical Yearbook*. The product quantity table in the *Industrial Yearbook* carries no note as to the coverage. But the section in which the data appear otherwise covers only the directly reporting industrial enterprises.³¹ This suggests that the product quantity data published in the *Statistical Yearbook* are across all years limited to the directly reporting industrial enterprises.

A first numerical double-check on the coverage of the product quantity statistics is to add up the quantities of particular products across provinces based on the data provided in the provincial statistical yearbooks, and then compare the sum across provinces with the reported nation-wide total in the *Statistical Yearbook*. Table 3 shows cement and steel production both as the sum of provincial data from the 31 provincial statistical yearbooks and as the nation-wide total reported in the *Statistical Yearbook*.

The sum across provinces and the nation-wide data as reported in Table 3 match almost perfectly. Yet the 1999 tables on product quantities in eleven provincial statistical yearbooks carry a note to the effect that these products are produced only by the directly reporting industrial enterprises. Only in two of the remaining 20 provincial statistical yearbooks, the ones for Shanxi and for Jiangxi, is a distinction made between total industrial production and production by the sub-category of directly reporting industrial enterprises; in Shanxi both cement and steel data cover all industry, while in Jiangxi cement data cover total industry, and

30. See the magazine *Zhongguo tongji* (*China Statistics*), No. 6 (1998), p. 24. The difficulties in compiling product quantities across all enterprises is widely echoed across articles in *Zhongguo tongji* (for example, in *Zhongguo tongji*, No. 2 (2000), p. 28, or No. 6 (2000), p. 35).

31. See NBS Division of Industry and Transportation, *Zhongguo gongye jingji tongji nianjian 1998* (*China Industrial Economy Statistical Yearbook 1998*, in the following abbreviated as *Industrial Yearbook 1998* and similarly for other years) (Beijing: Zhongguo tongji chubanshe, 1998), pp. 253–285, for 1997 data (with time series data on pp. 25–49), and *Industrial Yearbook 2001*, pp. 296–309, for 2000 data (with time series data on pp. 25–46). (No *Industrial Yearbook 1999* or *2000* has been published.)

Table 3: Production of Cement and Steel Across Provinces

	1997		1998		1999	
	Cement	Steel	Cement	Steel	Cement	Steel
Sum of product quantities across provinces, m tons	512.33	107.83	516.81	114.65	549.24	124.27
Nation-wide product quantity, m tons	511.74	108.94	536.00	115.59	573.00	124.26
Percentage shortfall in provincial sum from nation-wide total	-0.12	1.02	3.58	0.82	4.15	-0.01
Number of provinces for which a note to the product quantity table specifies that the table covers						
only the directly reporting industrial enterprises and, furthermore, that prior to 1998 it covered				10		11
(i) total production of all society				3		1
(ii) only the directly reporting industrial enterprises, too						1

total production of all society for some products output of all society, while for others only the directly reporting industrial enterprises	1	2	1
Number of provinces which provide data on both, all society, and the industrial enterprises on township level and above	5	1	1

Notes:

The provinces for which a note to the product quantity table specifies that only the directly reporting industrial enterprises are covered in 1998 were Heilongjiang, Zhejiang, Anhui, Fujian, Jiangxi, Guangdong, Hainan, Chongqing, Guizhou and Shaanxi; in 1999 they were Heilongjiang, Jiangsu, Zhejiang, Anhui, Fujian, Hubei, Hunan, Guangdong, Hainan, Chongqing and Guizhou. In 1997, the provinces which provided data on both, all society and the industrial enterprises on township level and above, were Beijing, Fujian, Hainan, Chongqing and Guizhou. (In calculating the sum across all provinces, whenever a provincial statistical yearbook provided data on all society and on the industrial enterprises on township level and above, the data on all society were used.)

Sources:

31 provincial statistical yearbooks of the year 1998 (with 1997 data), 1999, and 2000; *Statistical Yearbook 2000*, pp. 456f. (Qinghai's provincial statistical yearbook for 1998 steel output gives the same figure as for the previous product in the table, a figure which is several thousand per cent off the previous and next year's steel output value; therefore the mean of the almost equal 1997 and 1999 steel output values was used.)

steel production only covers the directly reporting industrial enterprises. In the other 18 provincial statistical yearbooks no explanation is offered.

If the 1999 data for these 18 provinces were to cover all enterprises, and if the nation-wide total were also to cover all industrial enterprises, as the *Statistical Yearbook* claims, then the provincial sum in the case of cement and steel should not be as close to the nation-wide total as it is (Table 3). The sum of provincial cement production in 1999 falls short of the nation-wide total by 4.15 per cent, while the sum of provincial steel production exceeds the nation-wide total by 0.01 per cent. The value-added of the directly reporting industrial enterprises in 1999 accounted for only 61.66 per cent of total industrial value-added. A mix of 12 (or 11, in the case of cement) provinces which report only approximately 61.66 per cent of their product quantities and 19 (20) provinces which could be reporting up to 100 per cent of their product quantities should yield an overall provincial total equal to at most 85 per cent of the nation-wide value, not a percentage which is 95.85 per cent or higher.

It is possible that steel is only produced in large enterprises with an independent accounting system located on the higher administrative levels, that is, in directly reporting industrial enterprises. That the sum of provincial steel output by directly reporting industrial enterprises equals the nation-wide total for industry then is not surprising. But much of cement production should occur in small-scale local enterprises not included in the directly reporting industrial enterprises. The nation-wide cement output reported in the *Statistical Yearbook*, if it covered all industry, thus should be significantly larger than the sum across provinces, but it is not.

Almost exactly the same provincial versus nation-wide pattern holds in 1998. But 1997 may be different. One of the 1998 provincial statistical yearbooks (with 1997 data) carries a note with the product quantity table that total product quantities of "all society" are reported, while five additional provinces report two figures, total production of all society, and product quantities of industrial enterprises on township level and above (regardless of accounting system).³² All other 1998 provincial statistical yearbooks do not clarify the coverage. This suggests that the 1997 coverage of the provincial data may extend to all industry. But a note to the product table in the year 2000 Hubei provincial statistical yearbook states that pre-1998 data cover only enterprises with an independent accounting system at township level and above; no such note was included with the identical data up to 1997 published in the 1998 Hubei

32. Across the five provinces which report both, product quantities of all industrial enterprises as well as of the industrial enterprises on township level and above, in the case of cement the output of the industrial enterprises on township level and above in 1997 accounted for 90.11% of the total output of all industrial enterprises, while in the case of steel the share was 99.99%. Much of cement production may well be concentrated in rather small industrial enterprises on township or county level; these enterprises were included in the group of directly reporting industrial enterprises prior to 1998 but not since 1998. The difference between the cement output of the directly reporting industrial enterprises and that of all industry then should be larger since 1998.

provincial statistical yearbook. This implies that at least before 1998 coverage was not uniform across all provinces. To judge from the sum of the provincial 1997 data (Table 3), the nation-wide figure may simply be the sum of whatever is reported by the provinces, in strident disregard of the variation in coverage between provinces.

A second double-check on the coverage of the product quantity statistics is to examine the changes in aggregate nation-wide product quantities over time. In 1997, output fell in 35 per cent of the product categories on which the *Statistical Yearbook* reports product quantities, in 1998 in 63 per cent and in 1999 in 19 per cent. The 1998 change in product quantities across all categories appears to differ from other years, potentially indicating a statistical break.³³

Table 4 reports on those product categories in which output in 1998 fell most; a cut-off point of a reduction by more than 10 per cent compared to the previous year is used. One would not expect salt production to fall by 27 per cent in one year, canned food production by 39 per cent or the output of vacuum cleaners by 35 per cent, especially not when output in most other years was stable. The simplest explanation is that the enterprise coverage changed in 1998 from a rough coverage of all industrial enterprises to coverage of only the directly reporting industrial enterprises. But the table also suggests that coverage could be changing frequently, or that in some years “guesstimates” on total output of society are made, while not in others. The fluctuations in the production of edible vegetable oil over time, for example, are highly unusual for a staple good, to the degree of not being credible. Another widespread change in coverage may even have occurred in 1995, when output in several categories where one would not expect drastic changes over time in fact experienced rapid growth.

Overall, the provincial-level data suggest that prior to 1998 the product quantities reported in the *Statistical Yearbook* could potentially be based on provincial-level data that more often than not (but not always) covered all industrial enterprises, while since 1998 the coverage is largely limited to the directly reporting industrial enterprises. Thus prior to 1998 the rule may have been to “guesstimate” total product quantities, while since 1998 the rule is to report only the output of directly reporting industrial enterprises. The nation-wide data confirm a 1998 pattern across product quantity categories that lacks credibility unless a statistical break has occurred. They also raise questions about the meaning of relative product quantity changes in other years. The available definitions and regulations as well as statements in Chinese journals confirm that aggregate product

33. *Statistical Yearbook 1998*, pp. 473–75; 1999, p. 445; 2000, pp. 459f. The 1999 and 2000 yearbook report data on 73 product categories for altogether the years 1997 through 1999. The 1998 yearbook reports data on 113 product categories for the years 1996 and 1997. All yearbooks report data on sub-categories for some products; four sub-categories in the earlier categorization turn into main categories in the more recent classification. The 1997 versus 1996 comparison is based on all 113 (main) product categories; if only the 73 (main) categories are considered on which data are available for 1998 and 1999 also, then the proportion of categories in which output fell in 1997 is only 23%, rather than 35%; this only strengthens the argument that the year 1998 is different.

Table 4: Relative Changes in Product Quantities (Products Whose Growth Rate in 1998 Was Below Negative 10%)

	1993/92	1994/93	1995/94	1996/95	1997/96	1998/97	1999/98	Remarks regarding 1998 growth rate
Cotton knitwear	0.86	0.89	1.14	0.84	1.06	0.64	1.10	New coverage?
Knitting wool	0.98	1.28	1.17	0.94	1.01	0.63	1.26	New coverage?
Silk	1.27	1.13	1.07	0.84	0.87	0.82	1.04	Long-term decline?
Woollen piece goods (wool fabric)	1.05	1.18	1.58	0.69	0.84	0.69	1.02	Long-term decline?
Machine-made paper and paperboard	1.11	1.12	1.32	0.94	1.04	0.78	1.02	New coverage?
Vacuum cleaners	1.08	1.17	2.50	1.09	1.11	0.53	1.11	New coverage?
Electric fans	1.20	1.11	1.51	0.79	0.79	0.82	0.92	Long-term decline? (Air conditioning grows rapidly)
Household ceramics	1.20	1.11	1.14	1.05	1.16	0.65	0.95	New coverage?
Salt	1.04	1.02	0.99	0.98	1.06	0.73	1.25	New coverage? Typo?
Canned food	1.03	1.07	1.26	0.91	0.90	0.61	1.08	New coverage?
Liquor	1.09	1.10	1.22	1.01	0.98	0.73	0.88	New coverage?
Edible vegetable oil	1.46	0.75	1.58	0.83	0.94	0.67	1.22	New coverage?
Traditional Chinese medicine	1.12	1.05	1.62	0.66	1.17	0.71	1.00	Extreme growth in 1995. New coverage?
Hi-fi stereo component players	1.37	0.94	1.80	0.83	1.02	0.82	1.28	Extreme growth in 1995. New coverage?
Mining equipment	0.99	1.24	1.42	0.57	1.07	0.67	0.85	Cyclical behaviour?
Equipment for chemical industry	1.14	1.14	1.65	0.85	1.02	0.67	0.85	New coverage?
Power generating equipment	1.14	1.14	1.65	0.85	1.02	0.67	0.85	New coverage?

Metal-cutting machine tools	1.15	0.79	1.91	0.45	1.05	0.64	1.19	New coverage? Extreme growth in 1995.
Motorcycles	1.73	1.51	1.54	1.11	1.13	0.80	1.18	New coverage?
Bicycles	1.02	1.05	1.02	0.75	0.89	0.77	1.04	New coverage?
Large and medium tractors	0.66	1.24	1.36	1.32	0.98	0.82	0.96	New coverage?
Internal combustion engines	1.65	1.08	0.96	1.40	0.93	0.78	1.11	New coverage? Extreme growth in 1995.
Railway passenger coaches	1.12	0.99	1.30	1.09	0.97	0.62	1.13	New coverage?
Railway freight wagons	1.34	1.30	0.99	0.88	0.95	0.75	0.79	Long-term decline?
Carrier communication equipment				0.38	0.56	0.86	1.07	Long-term decline?
Telephone sets	1.34	2.15	1.74	0.80	1.09	0.75	1.10	New coverage? (Mobile commun. equipm. grows fast)
Fax machines	-	-	-	1.01	1.18	0.79	1.24	New coverage?
Magnetic tape, converted into 6.3 mm	-	-	-	0.94	1.15	0.60	0.95	New coverage?
Floppy disks	-	-	-	0.61	0.66	0.75	1.14	New coverage?
Large semiconductor integr. circuit	-	-	-	0.82	1.61	0.54	2.52	New coverage? Typo (fall in 1998; this jump in 1999)?

Sources: *Statistical Yearbook 1995*, pp. 416-8, 1996, pp. 441-3 (revised product list), 1998, pp. 473-5, 1999, pp. 445f (revised product list), 2000, pp. 459f.

1991-99	
GDP	5
Industrial value-added	5
GOVI	1
GOVI minus GOV of private ind.	1
GOV of directly reporting ind. ent.	1

Notes:
 GDP: Gross Domestic Product; GOVI: Gross Output Value of Industry; GOV: Gross Output Value. Private industry in this table comprises private and individual-owned units in industry.
 GOV of directly reporting industrial enterprises is only available since 1979, the correlation coefficient for the period “1978-99” thus only covers 1980-99.
 “-” denotes insignificance (using 10% as the cut-off significance level).

Sources:
 Freight: *Statistical Yearbook 1992*, pp. 512f; 2000, p. 516. GDP and industrial value-added: *Statistical Yearbook 2000*, p. 55. GOVI (with separate data on private industry): *Statistical Yearbook 1992*, p. 407 (p. 406 for nominal data); 2000, p. 409 (for real and nominal data). GOV of the directly reporting industrial enterprises (available only since 1979): NBS, *Xin Zhongguo wushi nian tongji ziliao huibian (Statistical Materials on the 50 Years of New China)* (Beijing: Zhongguo tongji chubanshe, 1999), p. 146; *Statistical Yearbook 1992*, p. 414; 1993, p. 417; 1994, p. 378; 1995, p. 385; 1996, p. 411; 1997, p. 421; 1998, p. 441; 1999, p. 432; 2000, p. 414; GOVI in nominal and real terms is used to derive a deflator that is then applied to the nominal GOV of the directly reporting industrial enterprises, for which no real data are available.

quantity data may be of poor quality throughout; different provinces may be applying different coverage rules, and within one province coverage may differ from product to product. It appears that product quantity data are never suited as a double-check on industrial value-added.³⁴

Freight. The third assessment of the growth rate of industrial value-added is based on freight. Following the *Statistical Yearbook*, data on freight are compiled by the Transportation Ministry in co-operation with the NBS, and are based on comprehensive reporting by some transportation units and otherwise sample surveys. The private economy, including the self-employed (individual-owned enterprises), is explicitly covered.³⁵ Freight should thus correspond with the output of goods that are being transported. This is obviously not limited to industrial goods, and freight thus is not necessarily related to industrial output. If freight were nevertheless compared to industrial output (as done in the literature), the gross output value of industrial goods should be the relevant value for comparison, rather than value-added.³⁶

A first comparison of output growth to the growth in freight is possible using time series data. Table 5 reports the significance level of the correlation coefficient of five measures of output growth each with two measures of growth in freight (which in turn are also broken down into five exhaustive sub-categories). The two measures of freight are “freight traffic” (*huowu yunshu liang*), measuring the weight of freight, and “freight turnover” (*huowu zhouzhuan liang*), measuring weight times distance. The five output measures are GDP, industrial value-added, GOVI (of all industry), gross output value of all industry except the private sector, and gross output value of the directly reporting industrial enterprises.³⁷ (All comparisons are based on growth rates of freight and real growth rates of output.)

Except for the case of GDP, none of the output measures is significantly correlated with any of the two measures of freight over the

34. A *technical* reason why product quantity tables are not well suited as a control on output data is that bureaucratic inertia naturally leads to an emphasis on products which tend to be in the mature if not terminal phase of their life cycle, and on staples. The average growth in reported product quantities, even when there are no statistical breaks due to changes in enterprise coverage, thus will tend to be below that of (real) industrial value-added.

35. *Statistical Yearbook 2000*, p. 507.

36. Gross output value is more relevant than value-added because materials and intermediate goods also have to be transported. If an enterprise that previously produced a final product completely by itself closes down the production of some intermediate goods and begins to buy these intermediate goods from other enterprises, this adds to freight; yet aggregate value-added is unchanged (gross output value increases). Meng Lian and Wang Xiaolu, “An evaluation,” ignore these details and compare freight directly to industrial value-added to derive their conclusion that industrial value-added is likely to have been falsified since 1992.

37. Time series data on the value-added of directly reporting industrial enterprises are not available. What is available are net material product data from 1982 through 1992, and value-added data since 1992; the values of value-added in 1992 through 1994 appear highly unreliable (these values were originally reported in the *Statistical Yearbook* covering the particular year, and never retrospectively revised). A combined net material product/value-added growth rate time series yields results virtually identical to those of the gross output value of the directly reporting industrial enterprises.

period 1978 to 1999; GDP is significantly (and positively) correlated with total freight turnover only, and only at the 10 per cent significance level. Of the sub-categories, only waterway freight traffic and (civil) aviation are somewhat correlated with output measures, but then with output measures for the whole economy as well as for industry and a sub-category of industry, the directly reporting industrial enterprises. (Aviation and petroleum/gas pipelines account for only a very minor share of total freight, as shown in Table 6, and will therefore not be further discussed below.)

Due to a severe statistical break in 1979 and 1984 in highway freight data, which then also affects the total, Table 5 also reports the correlation coefficients for the period since 1985. It distinguishes between 1985 to 1990, and then 1991 to 1999, since Meng Lian and Wang Xiaolu conclude from the freight data on problems in industrial value-added statistics beginning in 1991. (The patterns of correlation coefficients are virtually the same if 1991/92 is used as breaking point instead of 1990/91.) In the period 1985 to 1990, none of the two total freight measures, nor the sub-categories railway, highway or waterway freight is significantly correlated with any of the output measures. In the period 1991 to 1999, on the other hand, freight is significantly correlated with output measures for the whole economy, for all of industry and for the directly reporting industrial enterprises. For the late 1980s these findings suggest that if one assumes freight data to be accurate, then output data may have been highly inaccurate; the good fit in the 1990s implies that falsification of nation-wide output data, if it ever were an issue, occurred in the late 1980s rather than in the 1990s. Alternatively, freight data could be inaccurate, or the logical relationship between freight and output weak. For the 1990s, the official output growth rates appear perfectly justified given the official freight growth data.

Figure 1 shows that the growth trends are not diverging systematically. GDP growth in the 1990s is (consistently) slightly above the growth in freight. The pattern of the 1990s furthermore replicates that of the early 1980s when industry was still very much under plan control and the data are thus perhaps quite reliable.

Of the different modes of freight, the most important ones are railway, highway and waterway freight (Table 6). Comprehensive railway data should be readily available to the NBS given the centralized nature of the railway system, and railway freight, especially railway freight turnover, then should be highly correlated with output, but it is not, neither in the earlier nor in the later period. Similarly, data on waterway freight could be relatively easy to obtain as much of waterway freight is concentrated in a few coastal harbours and thus could occur in a small number of enterprises (or administrative units) on which statistical reporting is complete, but there is virtually no correlation.³⁸ Part of the problem could

38. In 1999, Tianjin, Shanghai, Shandong, and Guangdong accounted for 64.02% of total waterway freight turnover, and Tianjin, Shanghai, Jiangsu, Zhejiang, and Guangdong for 67.42% of total waterway freight traffic.

Table 6: Relative Shares of Types of Freight (in %)

	Freight traffic				Freight turnover			
	1978	1985	1995	1999	1978	1985	1995	1999
Railway	44.23	17.53	13.43	12.93	54.38	44.25	36.02	31.70
Highway	34.22	72.15	76.16	76.62	2.79	10.36	13.14	14.13
Waterway	17.39	8.49	9.17	8.87	38.45	42.09	49.12	52.51
Civil aviation	0.00	0.00	0.01	0.01	0.01	0.02	0.06	0.10
Petroleum + gas pipelines	4.16	1.83	1.24	1.57	4.37	3.28	1.65	1.55

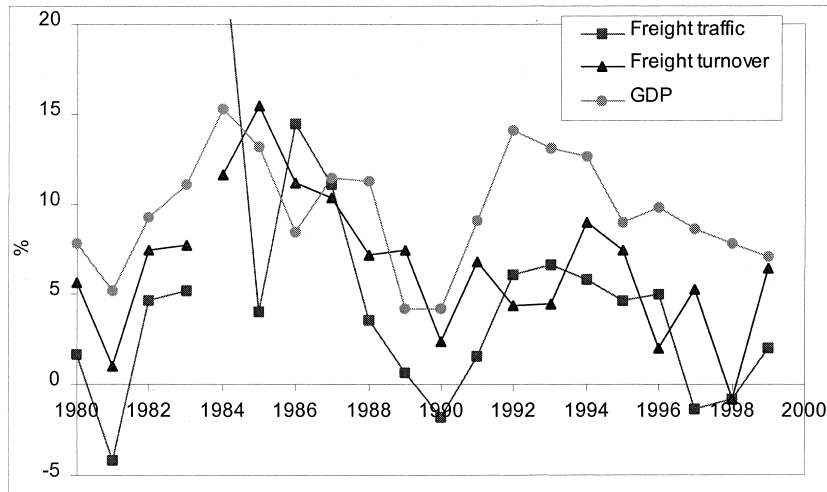
Notes:

The coverage of highway freight in 1979 was extended to all society (presumably it was limited to transportation by government transportation departments in previous years), with a 1979 growth rate in freight traffic of 335.58% (and of 171.76% in freight turnover), and since 1984 includes the private sector, with a 1984 growth rate in freight traffic of 32.88% (and of 41.70% in freight turnover).

Source:

Statistical Yearbook 2000, p. 516.

Figure 1: Growth of Freight and Real Output



Notes: Freight traffic and turnover experience a statistical break in 1984 (see note to Table 6). The 1984 growth rate of freight traffic is 24.27%. Sources: Freight: *Statistical Yearbook 1992*, p. 512, 2000, p. 516. GDP: *Statistical Yearbook 2000*, p. 55.

be that much of waterway freight appears to refer to overseas shipments, and exports and imports are not necessarily a constant fraction of output. Data on highway freight finally should be difficult to obtain because of the rapid growth of freight services provided by private sector units.³⁹ Furthermore, freight services provided by industrial enterprises themselves – presumably mainly in the form of highway freight – are by definition excluded from the statistics on freight which only cover registered transportation enterprises.⁴⁰ Nevertheless, the correlation between output and highway freight in the 1990s is rather strong.

Ralph Huenemann, examining the data on freight and passenger turnover, argues that road freight is likely to have become vastly underestimated during the 1990s; according to his calculations, petroleum consumption in the transportation sector between 1990 and 1995 grew by 11.2 per cent per year, compared to a growth rate for freight turnover

39. Thus the number of private trucks between 1985 (the first year for which the data are available) and 1998 (the last year for which comparable data below are available) increased from 0.2648 m to 19.203 m; the capacity of these private trucks between 1987 (the first year for which comparable data below are available) and 1998 increased from 15.372 m tons to 72.412 m tons. In the same periods, the number of trucks owned by the highway transportation department fell from 0.1480 m trucks to 0.1248 m trucks, and their capacity from 12.273 m tons to 9.002 m tons. I.e., by 1998 the number of private trucks and their capacity had widely surpassed that of the highway transportation department. In terms of all “civil” motor vehicles, the number of private trucks and their capacity had reached about one-third of the respective totals by 1998 and 1999, up from slightly more than one tenth in the mid-1980s (*Statistical Yearbook 2000*, pp. 530–32).

40. *Statistical Yearbook 2000*, p. 507.

of 6.4 per cent and for passenger turnover (passenger-kilometres) of 9.8 per cent.⁴¹ In the period 1995 to 1998 the growth rates were 14.0, 2.1 and 5.7 per cent. If the data on petroleum consumption in the transportation sector were correct, and if they were truly to cover all transportation services, including those provided by the private (and individual-owned) transportation sector and by industrial units themselves, then this implies that the official freight data are consistent underestimates of actual freight services provided.⁴²

A second check on the correlation between output growth and the growth of freight transportation is possible based on provincial data. Calculating the correlation coefficients of the growth rates of the five output measures with the growth rates of the two measures of freight (as well as of each of the three sub-categories for which provincial data are available, namely railway, highway and waterway freight) across the 31 provinces in one year at a time (1996, 1997, 1998 and 1999), none of the output growth measures is significantly correlated with freight growth ever, except occasionally GDP.⁴³

Provincial-level data appear highly problematic. The nation-wide total differs significantly from the sum across provinces in part because many transportation services cannot be clearly allocated to one particular province; thus a large volume of waterway freight and some highway freight are not classified by region.⁴⁴ Every year one or two provinces exhibit highly dubious (positive and negative) growth rates that can go up to several hundred per cent; the usual culprit is waterway freight, which may depend on harbour construction, but Tibet's highway freight data,

41. Ralph W. Huenemann, "Are China's recent transport statistics plausible?" *China Economic Review*, Vol. 12, No. 4 (2001), pp. 368–372.

42. One further indicator suggesting underestimation of freight is the difference in the development of the capacity of trucks in operation (measured in weight) and freight traffic (also measured in weight); between 1985 (the first year for which truck capacity data are available) and 1999, truck capacity increased by 293.83%, while highway freight traffic in the same period increased by only 184.08% (*Statistical Yearbook 2000*, pp. 516, 530).

43. For freight traffic, GDP is significantly correlated with total freight traffic in 1996 (10% significance level) and in 1999 (5%), with railway freight traffic in 1997 (10%), and with waterway freight traffic in 1998 (10%) and in 1999 (5%). For freight turnover, GDP is significantly correlated with total freight turnover in 1997 (1%) and in 1999 (5%), with railway freight turnover in 1997 (5%) and in 1999 (10%), and with waterway freight turnover in 1999 (10%). GDP is not significantly correlated with highway freight ever. Constructing real growth rates of the other four output measures in the case of one, at most two years involved some creative manipulations. (The details are too intricate to explain in brief.) In calculating the correlation coefficients all freight growth rates exceeding 50% in absolute terms were removed, since such growth rates are not credible and point to changes in definition or typos; this means that one to two outliers among the 31 provinces were removed every year, except for 1999 when the waterway freight data (and thus also total freight) were presumably re-defined and thus more provinces turned into outliers, 1999 data on waterway freight exploded by several hundred per cent across a dozen provinces, while unclassified waterway freight (not allocated to individual provinces) in 1999 dropped correspondingly. All provincial freight data are taken from the *Statistical Yearbook*.

44. In 1998, for example, freight turnover across all provinces following the provincial statistical yearbooks was 3138.99 b ton-kilometres, up 7.78% over the previous year. The *Statistical Yearbook*, however, gives a much larger annual value of 3784.07 b ton-kilometres, down 0.97% from the previous year; of this total, 1687.19 b ton-kilometres were not attributed to any province (*Statistical Yearbook 1998*, p. 543, 1999, p. 511).

for example, follow a incredible roller-coaster pattern. Finally, some provincial statistical yearbooks, such as that of Shanghai in 1997 and that of Heilongjiang in 1998, also include a note that some of the data cover only a sub-set of all freight enterprises, or that surveys were used in a particular year (but not in other years?).

Overall, the argument for a consistent relationship between freight and output measures is logically flawed. Too many other factors, such as the share of industry in GDP, the geographical concentration of economic growth, and the domestic and international trade patterns, matter. Freight data themselves appear of dubious quality, especially in waterway and highway freight. Freight data cover only transportation enterprises (while any transportation organized by a production unit itself is excluded), and it is unlikely that the statistical departments have available over time data of consistent coverage and quality on the private and individual-owned transportation sector.⁴⁵ Despite these complications, output and freight growth rates in the 1990s match rather well; this implies that freight data can provide no evidence to the effect that output data, either for the whole economy or for industry alone, are being falsified in the 1990s.

Questioning the Evidence on Falsification of GDP

The three approaches to the calculation of GDP yield three independent estimates of GDP. However, in China two of the three approaches are not used independently. According to a 1997 NBS compendium on how to compile GDP statistics, the production approach is dominant in some sectors of the economy (such as agriculture and industry) and the income approach in others (such as services). In those sectors where the production approach prevails, income GDP is simply set equal to production GDP; the reverse holds for the sectors in which the income approach is used.⁴⁶ In the aggregate, production GDP perfectly equals income GDP. The *Statistical Yearbook* presents nation-wide GDP, presumably production/income approach GDP, and, separately, nation-wide expenditure approach GDP. The two figures tend to differ slightly. Real growth rates are published only for production/income approach GDP.

Provincial-level data should allow a first double-check on nation-wide GDP. The *Statistical Yearbook* gives the presumably production/income approach GDP of each province with a break-down according to sector, separate provincial income-approach GDP and provincial expenditure approach GDP. Since 1995 the production/income approach provincial GDP figure equals the income approach provincial GDP figure across all provinces, and expenditure GDP matches these values in three-quarters of all provinces. This suggests that production and income approach

45. Private and individual-owned units are likely to underreport their services for income tax reasons. The official data will also not include smuggling activities, which are likely to change in volume over time as tariffs and local protectionist measures change.

46. NBS Division of National Income Accounting, *Zhongguo niandu guonei shengchan zongzhi jisuan fangfa* (*Calculation of China's Annual Gross Domestic Product*) (Beijing: Zhongguo tongji chubanshe, 1997), p. 12f.

GDP are indeed not derived independently, and that expenditure approach aggregates are often set equal to the production/income approach data, with some sub-category of the expenditure approach GDP then derived as residual.⁴⁷

The sum of production/income approach GDP across all provinces in most years exceeds the nation-wide total; the same is true for expenditure approach GDP. Accepted wisdom has it that provinces over-report output in order to meet provincial growth targets; the NBS consequently adjusts provincial data when deriving the nation-wide total.⁴⁸ The basis for these adjustments is not known; Liu Hong, then head of the NBS, in February 2000 claimed that the adjustments are made based on sample surveys and logical relationships between variables, without elaborating further.⁴⁹ Provincial data thus allow no double-check on nation-wide GDP data.

Thomas Rawski established an independent nation-wide income approach calculation based on the individual income components for 1997 and 1998.⁵⁰ He arrives at a 1998 real GDP growth rate of 5.7 per cent in contrast to the official 7.8 per cent. Yet these calculations are not unproblematic. Total GDP thus derived for 1998 equals only 83 per cent of official GDP, implying a fair number of omissions, which matter if the omitted variables grow at a different rate from the included variables. In addition, the calculations necessarily involve various assumptions, which on average may not have a neutral effect on the resulting growth figure. There are good reasons why Rawski's calculations may underestimate actual income approach GDP growth.⁵¹ Given these complications, the

47. For further details see Carsten A. Holz, "Institutional constraints on the quality of statistics in China," *China Information*, Vol. 16, No. 1 (2002), pp. 25–67.

48. According to *Zhongguo tongji*, No. 7 (1999), p. 12, GDP statistics are now adjusted downward by statistical departments on each tier (i.e., the county, municipal, and provincial tier) as they are reported up to the NBS. The provincial data themselves thus are already the result of revisions the foundation of which is not publicized.

49. See a 29 February 2000 news item titled "Zhongguo tongjiju juzhang 'bamai' tongji shuzi" ("NBS head 'gets a grip' on statistics") in China Infobank, an internet database (at <http://www.chinainfobank.com>).

50. Rawski, "China by the numbers."

51. For example, the profit of industrial township enterprises appears double-counted in 1997 (one time as part of the profit of the directly reporting industrial enterprises, and the second time as part of the profit of township- and village-level collective-owned enterprises), while in 1998 the profit of all small (annual sales revenue below 5 m *yuan*) urban non-state enterprises is omitted and the profit of all large (annual sales revenue below 5 m *yuan*) township- and village-level collective-owned industrial enterprises is double-counted; the profit of all rural private enterprises was excluded in 1997, and the profit of all small rural private enterprises is excluded in 1998 (by definition, the profit of private enterprises is "enterprise profit" and thus cannot have been included in the household surveys as household income). The effect of the different double-counting in the two years is unclear. But the unambiguous profit omissions are likely to be larger in 1998 than in 1997. The relevant data points to check the effects of the omissions are simply not available. (Rawski, "China by the numbers," p. 30, acknowledges these problems in general.) Another example is the adoption of official urban versus rural population data which are then multiplied with various urban versus rural per capita income data. Rawski and Mead, "On the trail," have reported 100 m missing farmers, and it is unlikely that these missing farmers' income is fully accounted for in rural household income survey data. If the migrant labour force grows relative to the total population, the nation-wide per capita household income is increasingly underestimated. A third example is the imputation of rent; Rawski's calculations cannot take into account

fact that Rawski’s 5.7 per cent income approach real GDP growth rate comes within 2.1 per cent of the official rate suggests that the official rate may not be very much off the mark, if at all.

Once the later revisions to the first-published official GDP data are considered, the discrepancy virtually vanishes.⁵² Thus, according to the *Statistical Yearbook 1999*, in 1998 nominal GDP grew by 6.6 per cent and real GDP by 7.8 per cent. But according to the *Statistical Yearbook 2000*, revised nominal GDP in 1998 grew by only 5.2 per cent, with the real growth rate still at 7.8 per cent.⁵³ The revised nominal GDP growth figure raises three possible conclusions. First, the implicit deflator of 0.989 in the *Statistical Yearbook 1999* is the true deflator (no GDP deflator is published separately); then the 1998 real growth rate needs to be revised downward to 6.4 per cent, rather closer to Rawski’s estimate of 5.7 per cent; this is the most likely scenario. Secondly, the published real growth rate of 7.8 per cent is the true value; then the new nominal GDP value for 1998 implies the need to revise the implicit GDP deflator downward to 0.976. Applying this deflator to Rawski’s partial income approach data (as he does using the earlier implicit deflator) yields a real growth rate using Rawski’s reconstructed income approach of 7.2 per cent, rather closer to the official real growth rate. A third possibility is that the GDP deflator needs only to be partly revised, leading to a combination of the above two cases. At any rate, following the later revisions to the official GDP data, the discrepancy between Rawski’s and the official real growth rate narrows to at most 0.7 percentage points; given Rawski’s necessarily incomplete and tentative calculations, this difference appears insignificant.⁵⁴

Yet another double-check on the official production/income approach GDP data is possible via the official expenditure data. Albert Keidel calculated the real GDP growth rate based on the expenditure approach and contrasted this growth rate with the official production/income approach real GDP growth rate.⁵⁵ The expenditure approach growth rate

footnote continued

improvements in housing quality over time, but these should be significant. (On the rental issue also see Albert Keidel, “Practical issues in using the Chinese statistical system,” *China Perspectives*, No. 33 (January/February 2001), pp. 35–41.)

52. First annual GDP data are published within about one month of the end of a year, and tend then to be repeated in the forthcoming *Statistical Yearbook*. The nation-wide nominal GDP data are typically revised downward (upward for 1999) in the next edition of the *Statistical Yearbook*.

53. *Statistical Yearbook 1999*, pp. 55, 57; *2000*, pp. 53, 55.

54. Assume that Rawski’s calculations were to yield the exact GDP figure for the production activities he covers (83% of all production activities). Using the first published official real GDP growth data, the sectors not covered by Rawski would have had to grow by 18.5% in 1998 for the official real GDP growth rate to be correct ($83\% * 1.0570 + 17\% * 1.1805 = 1.078$). Using the revised official nominal GDP growth data, the sectors omitted by Rawski only had to grow by 9.82% ($83\% * 1.0570 + 17\% * 1.0982 = 1.064$). A growth rate of 9.82% for the omitted sectors appears far from unrealistic. More recently, Rawski, “What’s happening,” guesses that China’s real GDP growth rate in 1998 was between -2.0% and $+2.0\%$, with similarly low rates in 1999 and slightly higher rates in 2000 and 2001. (Thomas Rawski himself qualified these figures as “guesses.”)

55. Albert Keidel, “China’s GDP expenditure accounts,” *China Economic Review*, Vol. 12, No. 4 (2001), pp. 355–367.

is 1.5 percentage points below the production/income approach 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 percentage points in 2000. Some discrepancies in the 1980s are larger (in both directions); if the expenditure approach GDP were the correct GDP value, these findings would contradict any argument to the effect that the NBS started falsifying data around 1992 (as Meng Lian and Wang Xiaolu suggest), or that falsification intensified in 1998 (as Thomas Rawski suggests).

But the expenditure approach GDP data are not without serious problems, and therefore are highly unlikely to constitute the “correct” GDP data. In the absence of real expenditure approach data, Albert Keidel applied various price indices to the different expenditure components, none of which is necessarily fully appropriate for the particular expenditure component. Furthermore, both the household survey data and the retail sales data which each could underlie the reported consumption expenditures (which in turn account for more than half of GDP) are highly problematic. Both probably represent an underestimate of aggregate consumption expenditures; households are likely to underreport income and expenditures (in order to avoid taxes), and retail sales statistics do not fully capture transactions outside the state and collective trading sector (because of either lack of data, underreporting for tax reasons or officially acknowledged omissions).⁵⁶ The NBS may have chosen at this stage of statistical development to focus on production/income approach GDP for good reason.

Neither the pure income approach nor the expenditure approach appear to yield a reliable alternative figure to the official real GDP growth rate and thus can conclusively show NBS falsification. Still, in case one were to trust these alternative results more than the official GDP growth rate, the differences between the official real GDP growth rate and the problematic alternatives appear minor if not insignificant.

Evidence Against Data Falsification by the NBS

Invalidating the key “evidence” on data falsification by the NBS does not amount to a proof that the NBS is honest in its statistical reporting. Such a positive proof is impossible to supply. However, it can be shown that the scope for NBS data falsification is limited, and that Chinese reports on data falsification fit a general pattern of bureaucratic and governmental conflict resolution that is more likely than not to imply a professional attitude by the NBS and lower-level tier statistical depart-

56. According to *Zhongguo tongji*, No. 9 (2001), p. 44, trade between farmers is not included in retail sales. Since 1993, retail sales to farmers which then use these goods as intermediate products are supposedly no longer included in the published retail sales data. (The article does not state how the volume of sales of in the context of agriculture intermediate goods is estimated; if provinces made this innovation at different points of time, as they so often do with statistical changes, retail sales time series data are not comparable over time.)

ments, rendering data falsification improbable. Finally, recent developments in the statistical system are likely to have reduced the opportunities for data falsification.⁵⁷

Scope for NBS data falsification. If the NBS were actually to have accurate data and then to publish falsified statistics, this would imply the existence of two sets of books, one with correct data and one with falsified data. Yet there is no evidence of this. Widespread data falsification would surely lead some employee of the statistical departments at some point to reveal the existence of a second set of data. Various internal publications that are occasionally available to researchers largely match the published data.⁵⁸

Composing a set of credible false data, furthermore, is likely to be exceedingly difficult. For example, detailed output data of the directly reporting industrial enterprises are available in most provincial statistical yearbooks, matching those published in the *Statistical Yearbook*. Provincial statistical yearbooks often provide a further breakdown according to municipalities and sectors. Aggregate provincial output of the directly reporting industrial enterprises equals the sum across sectors – approximately only, when data on the 40th sector, weapons and ammunition are not available – as well as the sum across municipalities. Data falsification by the NBS would require it to request provinces to change the data they report on some sectors or municipalities. Such falsification would furthermore have to be consistent over time in order not to arouse suspicion. While falsification by central order may be feasible on a limited scale, perhaps focusing on a few provinces, and sectors and municipalities within these provinces only, the larger the size of falsification and the wider the scope, the more likely that unexplainable inconsistencies emerge. The scope for falsification of local data by central order thus appears very small, especially for clearly defined variables with limited coverage, such as value-added of the directly reporting industrial enterprises.

The case of GDP is different. The NBS every year publishes an aggregate nation-wide GDP figure that is smaller (except in 1995) than the sum across provinces, claiming that the provincial data are inflated. Assuming that the provincial data are indeed inflated and that the NBS knows by how much, it can falsify data by not fully deflating, or by over-deflating the provincial data in deriving the aggregate nation-wide. However, the NBS may not know exactly by how much provincial data are inflated, and whether the haphazardly adjusted nation-wide value then are false or not cannot be determined. Furthermore, provincial output data

57. Data accuracy is an issue that is potentially separate from that of honest reporting. Even if data are reported without falsification, they may still be inaccurate for technical reasons. Holz, “Institutional constraints,” covers some technical issues.

58. See for example, the data presented in the internal monthly statistics *Shisan sheng shi guomin jingji zhuyao tongji zhibiao* (*Major Statistical Indicators on the Economy of 13 Provinces and Cities*). Rawski similarly sees no evidence for the existence of internal statistics that contradict the published data.

have probably been inflated all along – see the industrial census 1995 revisions to rural industrial output of the years 1991 to 1994 – when the NBS did not adjust the provincial GDP data in deriving the aggregate nation-wide as much as it has since 1997; this raises the possibility that today's aggregate nation-wide data are less false than earlier ones.

Another way for the NBS to “falsify” data is to phase in perfectly sensible adjustments to variable definitions and enterprise categorization as convenient. By changing the rules that cover the imputation of the rental value of self-owned housing in the countryside, for example, the NBS may easily change the value of GDP by several percentage points.⁵⁹ Xu Xianchun presents a number of items that, if implemented in the compilation of GDP, approximately cancel out overall, but the NBS could simply choose adjustments that all affect GDP in one particular direction.⁶⁰ (Xu Xianchun is affiliated with the national income accounting division of the NBS.) The exact extent of the difference in Chinese statistical practices and international statistical practices is unknown.⁶¹

Published later revisions to earlier data may incorporate some of the re-definitions, but deriving revised earlier data may not always be technically possible, even if the NBS desired to improve earlier data retrospectively.⁶² In 1994 the World Bank in its publications adjusted China's official per capita GDP upward by 34.3 per cent and continued to adjust official Chinese data in the following years. Xu Xianchun reports the World Bank's reasoning from an internal report, and also documents the changes in China's statistical system that have taken place between 1992 and 1998.⁶³

In 1999, the World Bank accepted the official Chinese GDP data for its own publications, presumably beginning with the 1998 data, thus effec-

59. Similarly, output data on those industrial enterprises that are not directly reporting to the statistical departments are frequently no more than “guesstimates.” Yet these enterprises account for approximately 40% of GOVI, and thus for approximately 20% of GDP. Changing the rules on how information on these enterprises is to be obtained (for example, multiply the output of the directly reporting enterprises in a locality by 1.8 rather than by 1.7, following the growth of private industry?) could easily raise an aggregate GDP growth rate of just below 8% to just above 8%. Such changes in data compilation may have good justifications.

60. Xu Xianchun, “Zhongguo guonei shengchan zongzhi hesuan zhong cunzai de ruogan wenti yanjiu” (“Some problems in the calculation of China's GDP”), *Jingji yanjiu*, No. 2 (February 2000), pp. 10–16. The transition in statistical practices is an ongoing process. Xu lists five areas in which further adjustments are needed: valuation of housing services, government subsidies, enterprise-internal social welfare activities, rural industrial statistics and agricultural statistics. While individual items are as large as 3.3% and negative 3.8% of GDP, the net effect of adjustments calculated by Xu for 1997 is an upward adjustment of GDP by only 0.2%. Xu suggests that further consideration on these issues is needed before the corrections are officially adopted.

61. Xu Xianchun, “Woguo GDP hesuan yu xianxing SNA de GDP hesuan zhijian de ruogan chayi” (“Some differences in China's GDP calculation and the current SNA GDP calculation”), *Jingji yanjiu*, No. 11 (November 2001), pp. 63–68, provides some details.

62. Also see Holz, “Institutional constraints,” and Keidel, “China's GDP expenditure accounts.”

63. Xu Xianchun, “Shijie yinhang gaogu Zhongguo GDP shuju” (“The World Bank estimates China's GDP to be higher”), *Zhongguo guoqing guoli (China's National Situation and Strength)*, No. 1 (January 1999), pp. 7–10; Xu Xianchun, “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*, No. 6 (June 1999), pp. 52–58.

tively certifying current official Chinese GDP statistics as sufficiently accurate by World Bank standards. Yet the reduction of the discrepancy of 34.3 per cent over presumably six years potentially implies that 5.04 percentage points of the annual growth in nominal GDP per capita as reported in official Chinese statistics is due solely to adjustments to China’s GDP calculation practices. Even if half of the adjustments the World Bank made to the 1992 data were to have been exaggerated – slightly above one-half of the GDP adjustments were due to the application of alternative prices – the remaining adjustments would still account for 2.67 percentage points of annual real growth. China’s official GDP growth rates thus need not always perfectly match actual growth in the underlying productive activities, without the reason necessarily being data falsification.

Re-interpretation of Chinese reports on data falsification. Reports on data falsification proliferated across the Chinese press in the late 1990s. Meng Lian, Wang Xiaolu, Rawski and Xiao Wei provide ample citations and references.⁶⁴ Reports tend to fall into two categories. One type presents concrete evidence of data falsification, typically regarding one local enterprise or government.⁶⁵ Such hard evidence is relatively scarce, and the degree of falsification in the particular instance often in the single-digit or lower double-digit percentage range. A second type of report remains very general, stating that data falsification takes place without offering proof, and often providing the standard reasons why it occurs (meeting growth or other targets, and achieving a higher evaluation of one’s job performance).⁶⁶

None of these articles offers an estimate of the share of enterprises in a particular locality (let alone nation-wide) that resort to data falsification, or on the relative number of local governments that pressure statistical departments into falsifying data. Similarly, no estimate of the degree of data falsification of a particular variable nation-wide is offered in these articles. The often highly sensational tone of Chinese reports further raises the suspicion that their motivation is not to reveal particular wrongdoing or to devise a useful measure of the degree of falsification.

Staff of local statistical departments in China have few means to induce enterprises into accurate reporting and are themselves exposed to latent if not explicit pressure from local governments. Higher-level statistical departments have no authority over local governments or enterprises. In such circumstances statistical staff concerned about the quality of Chinese statistics have little choice but to make use of a

64. Meng Lian and Wang Xiaolu, “An evaluation”; Rawski, “On the reliability”; Thomas G. Rawski and Xiao Wei, “China economic review symposium on Chinese economic statistics: introduction,” *China Economic Review*, Vol. 12, No. 4 (2001), pp. 298–302.

65. See, for example, *Zhongguo tongji*, No. 5 (1998), p. 21, or No. 7 (1998), p. 30.

66. See, for example, *Zhongguo tongji*, No. 11 (1998), pp. 21f, or No. 6 (1999), p. 22. Sloppy survey methods and poor quality of statistical personnel are also frequently cited. Yet these do not represent falsification as much as inaccurate methods of data collection.

powerful bureaucratic conflict resolution mechanism, and that is to attract the attention of the top Chinese leadership to their plight.

In 1995 when justifying revisions to the PRC Statistics Law, Zhang Sai, then head of the NBS, explained to the Standing Committee of the National People's Congress that the main purpose of the forthcoming revision to the PRC Statistics Law was to ensure the compilation of data of higher quality: "recently the phenomenon of false and deceptive reporting has spread in some localities and some units. The danger is large, the impact very negative."⁶⁷ Reports on data falsification then became standard fare in the NBS journal *Zhongguo tongji* in the late 1990s.⁶⁸

The top leadership was prodded into action. In 1997 the NBS cooperated with the Chinese Communist Party Central Committee (CCPCC) Disciplinary Commission and the CCPCC Organizational Department in drawing up a document attacking data falsification. The offices of the CCPCC and State Council in February 1998 formally issued the document. This was followed in 1999 by a NBS circular sharply criticizing data falsification in some localities.⁶⁹ Individual provinces followed suit, such as a circular by the government of Shaanxi province requesting an evaluation of the quality of GDP data, and assigning responsibility for the accuracy of GDP data to the heads of all statistical departments.⁷⁰ Local government leaders now have to weigh the benefits from data falsification against the costs of contravening an unambiguous central leadership order.

The course of events over the past few years fits well with an explana-

67. Zhang Sai, "Guanyu 'Zhonghua renmin gongheguo tongji fa xiuzheng'an (zao 'an)' de shuoming" ("Explanation of the (draft) scheduled revisions to the PRC Statistics Law"), 23 August 1995, in China Infobank.

68. The issue of data falsification has always been relevant. The increase in the frequency of such reports in the second half of the 1990s could reflect a new degree of openness, or a new degree of awareness of data complications (not necessarily falsification), or an increase in falsification. As an example of earlier reports on data falsification, the NBS Division of Industry and Transportation based on theoretical reckoning and survey findings in 1989 came to the conclusion that the "water content" in GOVI is about 1%. ("Gongye zengzhang sudu zhong de 'shuifen' you duoda?" ("How big is the 'water content' in industrial growth?"), *Jingji yanjiu cankao ziliao* (*Economic Reference Material*), No. 2042 (19 March 1989), pp. 2-5.)

69. The text of the two central regulations is not available. The titles, "Zhonggong zhongyang bangongting, guowuyuan bangongting guanyu jianjue fandui he zhizhi zai tongji shang nongxu zuojia de tongzhi" ("Circular issued by the offices of the CCPCC and the State Council on decisively opposing and curbing statistical data falsification") and "Guowuyuan bangongting zhuanfa guojia tongjiju guanyu yi xie difang zhongyao tongji shuju shishi wenti tongbao de tongzhi" ("Circular by the State Council passing on the NBS report on important statistical data inconsistencies in some localities") can be found in a China Infobank entry of 22 November 1999 titled "Heilongjiang sheng renmin zhengfu guanyu jiaqiang tongji gongzuo de jue ding" ("Decision by the Heilongjiang province government on strengthening the statistical work"). *Statistical Work Yearbook 1998*, p. 43, contains a short paragraph on the preparatory work. The CCPCC/State Council circular is mentioned in numerous sources, such as *Zhongguo tongji*, No. 6 (1998), p. 4.

70. Shaanxi Province Government, "Shaanxi sheng renmin zhengfu bangongting guanyu dui di shi guonei shengchan zongzhi shuju zhiliang shixing pinggu de tongzhi" ("Circular by the government office of Shaanxi province on evaluating the quality of municipal GDP data"), *Shaanxi zhengbao* (*Shaanxi Bulletin*), No. 21 (1999), pp. 27f.

tion based on established conflict resolution or policy-making practices within the Chinese political system.⁷¹ This explanation portrays the statistical bureaucracy as a largely professional organization concerned about the quality of the data it collects. If the NBS were intentionally to falsify data, the CCPCC and State Council decision in 1998 as well as the successive trickling down of this decision to local governments would make sense only as a propaganda measure meant to pacify a public sceptical about the quality of the official (falsified) statistics. While such reasoning cannot be invalidated, it appears rather far-fetched compared to the conflict resolution explanation.

In the 2001 issues of the *Zhongguo tongji* magazine, the wave of sensational reports on data falsification has largely given way to articles on technical and definitional problems of Chinese statistics. For example, an employee of the NBS Department for Statistical Development and Administration in an impressive review highlights some of the shortcomings (but also achievements) of Chinese statistical practice in recent years. The author at one point defends the aggregate growth data: “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 way to go.”⁷² In January 2002 the economist Fan Gang came to the defence of Chinese statistics, stating that he didn’t think Chinese economic data were exaggerated; on the contrary, there are good reasons why the official data may represent an underestimate of the actual state of the economy: both rich areas and poor areas prefer underestimates of their economic strength, the first in order to pay less taxes, the second in order to receive central government subsidies to poverty areas.⁷³

Reducing opportunities for data falsification. The 1995 industrial census appears a watershed in that it for the first time revealed widespread problems with the accuracy of rural industrial data. Following the industrial census of 1995, the NBS retrospectively revised data on the collective-owned and private economy for the years 1991 to 1994 down-

71. Also see, for example, Kenneth Lieberthal and Michael Oksenberg, *Policy Making in China: Leaders, Structures, and Processes* (Princeton: Princeton University Press, 1988), p. 30, on bureaucrats bringing issues to the attention of top leaders.

72. *Zhongguo tongji*, No. 11 (2001), pp. 8–10.

73. *Xinbao* (*Hong Kong Economic Journal*, a newspaper), 9 January 2001, p. 3. According to articles in *Zhongguo tongji*, many small, non-state enterprises underreport for tax reasons, but also to avoid becoming a member of the group of directly reporting enterprises which would lead to time-consuming reporting requirements. Statistical staff end up with a conflict of interest; on the one hand, they may wish to obtain accurate data, on the other hand, their job is not to police enterprises on behalf of the tax bureau. The default solution appears to be to accept whatever data the enterprise offers. The statistical departments also have problems establishing proper sampling frameworks for small enterprises, with as a result a tendency to underestimate these enterprises’ production activities. (*Zhongguo tongji*, No. 6 (1998), p. 17f; No. 6 (2000), p. 35; No. 11 (2000), pp. 16–18).

ward by up to 20 per cent.⁷⁴ The years since then have seen two developments. One is a gradual strengthening of statistical rules and regulations and the propagation of regular and more rigorous supervision over statistical reporting. The second is organizational changes in the collection and compilation of statistical data that are likely to reduce any scope for data falsification.

According to Zhang Sai, the 1983 PRC Statistics Law offered penalties for mis-reporting that were all too often too harsh actually to be implemented. The revised PRC Statistics Law of 1996 as well as the implementation instructions issued in 2000 consequently revised and expanded the range of penalties. The NBS in 1996 also clarified the procedures involved in handling violations of statistical regulations, and in 2001 issued a regulation on regularly examining the implementation of statistical rules and regulations across the statistics bureaucracy and the reporting units.⁷⁵

The NBS conducted nation-wide investigations in 1989, 1994, 1997 and 2001. An investigation typically lasts for three to five months, beginning with study sessions of recent statistical regulations, then self-investigation by all statistical departments and reporting units, followed by an investigation through central and local investigative teams of at least 10 per cent of all reporting units (and presumably all statistical departments). The 1997 investigation was organized by the NBS together with the Supervision Ministry and the State Council Law Office, while for the 2001 investigation the NBS was joined by the Supervision Ministry and the Justice Ministry.⁷⁶

The 1997 investigation revealed more than 60,000 violations of statistical regulations nation-wide, of which 70 per cent were self-reported. Of the 60,000 violations, 56.7 per cent concerned false reporting, 18.4 per cent refusals to report or late reporting, and 24.9 per cent other violations such as obstruction of the investigation or the implementation of unauthorized surveys. In the end, 15,000 violations were further pursued and led

74. Gross output value data of collective- and individual-owned enterprises for 1994 were retrospectively adjusted downward by 16% and 20%, with similar adjustments for the years 1991 to 1993 (see Holz and Lin, "Pitfalls of China's industrial statistics").

75. National People's Congress Standing Committee "Zhonghua renmin gongheguo tongji fa" ("PRC Statistics Law"), 15 May 1996, revision of original 8 December 1983 law, in China Infobank; NBS, "Zhonghua renmin gongheguo tongji fa shishi xize" ("Detailed implementation instructions for the PRC Statistics Law"), 2 June 2000, revision of original regulation of 19 January 1987, approved by the State Council, in China Infobank; NBS, "Guanyu renzhen gaohao tongji weifa anjian chachu gongzuo de tongzhi" ("Circular on conscientiously performing well the task of handling violations of statistical regulations"), 26 November 1996, in *Statistical Work Yearbook 1997*, pp. 139f; NBS, "Tongji zhifa jiancha guiding" ("Regulation on the examination of the enforcement of statistical rules and regulations"), 20 June 2001, PRC NBS Decree No. 6, *Zhongguo tongji*, No. 7 (2001), pp. 6–8. Since at least 1992 statistical departments on all tiers are also supposed formally to evaluate the quality of their data and to report their findings. These reports are not to be made public. (NBS, "Guanyu kaizhan zhuyao tongji zhibiao shuju zhiliang pinggu gongzuo de tongzhi" ("Circular on launching an appraisal of the quality of all important economic indicators"), 17 June 1992, in *Statistical Work Yearbook, 1993*, pp. 17–25.

76. The 2001 nation-wide inspection explicitly refers back to the 1998 CCPCC/State Council document on curbing statistical data falsification.

to disciplinary measures (in 0.4 per cent of these 15,000 cases), a formal warning (10.7 per cent), criticism published in a circular (26.3 per cent), an order to correct the data (7.5 per cent), fines (44.6 per cent) and other penalties (10.5 per cent). Individual cases of misconduct and the penalties were widely publicized. For example, exaggerating the county's GOVI by 61.57 per cent led to the dismissal of the county head.⁷⁷

Provincial-level inspections occur more frequently. Thus in Fujian province the provincial statistical office also conducted investigations in 1990, 1992 and 1995.⁷⁸ In many localities, inspections have become an ongoing process. Several provincial statistical bureaus have established permanent inspection offices (at the NBS this task falls to the policy and regulation division) and issued a corresponding regulatory framework. Almost all offices have by now issued their own statistical stipulations, incorporating the PRC Statistical Law but then expanding on penalties for data falsification. Henan's statistical bureau, for example, in 1996 in co-operation with the provincial supervision bureau issued a detailed catalogue of penalties depending on the degree of exaggeration. In 1993 Hunan's High People's Court established a liaison office for statistical issues, and similar offices were established by all municipalities and 95 per cent of all counties by 1998. A total of 19 cases were brought to court over the five-year period.⁷⁹ Central and provincial statistical bureaus have

77. On investigation issues see *Statistical Work Yearbook 1998*, pp. 25, 29, 38, 160, 171f, 330, the NBS website at [http://www.stats.gov.cn \(/xwkj/tjdt/200105230002.htm\)](http://www.stats.gov.cn (/xwkj/tjdt/200105230002.htm)), *Zhongguo tongji*, No. 3 (1998), pp. 15f, and the circular announcing the 1997 investigation (NBS, Supervision Ministry and State Council Law Office, “Guanyu kaizhan quanguo tongji zhifa da jiancha de tongzhi” (“Circular on beginning the nation-wide investigation into the implementation of statistical rules and regulations”), 4 May 1997, in *Statistical Work Yearbook 1998*, pp. 103–105). Data from individual provinces can be found in *Zhongguo tongji*, No. 1 (1998), p. 39; No. 5 (1998), p. 21; and No. 7 (1998), p. 30. In a municipality in Hunan province a central bank branch took the opportunity of the investigation of the statistical system to start an investigation into financial data, too (*Zhongguo tongji*, No. 7 (1998), p. 31). One example of data falsification with subsequent punishment in the 1994 investigation can be found in China Infobank under the title “Zhongguo dalu tongji fukua tanyuan” (“Exploration of statistical exaggeration in mainland China”), published on 30 June 1995. Both the scope of data falsification and penalties match those in the 1997 inspection. This article mentions that in some townships the township Party leader and the township head are promoted to the rank of deputy county-head if the gross output value of the township and village enterprises exceeds 100 m yuan. The preliminary results of the 2001 investigation are similar to those reported for 1997 in the text, with 62,000 violations, of which more than 15,000 as of December 2001 were being pursued further. (See article “Zhongguo guojia tongjiju juzhang Zhu Zhixin biaooshi tongji gongzuo yao jianchi shishi quishi yuanzi” (“Head of the NBS Zhu Zhixin says that statistical work has to adhere the principle of seeking truth from facts”) in China Infobank, 17 December 2001.)

78. *Fujian shengzhi tongjizhi (Fujian Provincial Almanacs Statistics Volume)* (Fuzhou: Fangzhi chubanshe, 2000), pp. 192–94.

79. See <http://www.stats.gov.cn/zwgk.htm> on the institutional structure of the NBS; Henan Statistical Bureau, “Weifan tongji falü fagui xingzheng chufen zanxing guiding” (“Temporary regulation on the administrative punishment for violations of the statistical laws and regulations”), 28 November 1996, issued together with the Henan supervision bureau, in *Statistical Work Yearbook 1998*, p. 153f, on Henan province; and *Zhongguo tongji*, No. 2 (1998), p. 38, on Hunan province. A county-level city in Hebei province has established a city-wide investigation network under the direction of the investigation division of the statistical department; its three staff together with eight hired helpers and nine staff seconded from the city government conduct two large-scale investigations every year (*Zhongguo tongji*, No. 5 (1998), p. 33). In Jilin province, the provincial People's Congress established inspection

since the early 1990s dispatched their own personnel to lower-level statistical offices (*tepaiyuan zhidu*) specifically to ensure that the higher-level bureau's regulations are implemented and to uncover violations of statistical regulations.⁸⁰

The second development in the late 1990s was a restructuring of the statistical reporting system. The innovations are most apparent in the industrial sector, but similar reforms apply at least to the construction sector and commerce. In industry, the NBS has moved to collect high-quality data on a small number of industrial enterprises that account for more than half of industrial output, while trying to obtain an accurate picture of all other industrial activities through sample surveys conducted by its own survey teams. Thus the 1998 change in the requirements for an industrial enterprise to be a "directly reporting industrial enterprise" with the switch from an administrative level as criterion for inclusion to a size criterion reduced by two-thirds the number of directly reporting industrial enterprises, while keeping aggregate output value of the enterprises covered approximately constant. The enterprises that dropped out are the smallest enterprises, those which are the least likely to have a reliable accounting system.⁸¹

A further streamlining of the industrial reporting system followed in 1999 and 2001.⁸² Beginning in 1999, the monthly statistical reports of the directly reporting industrial enterprises are sent to the local statistical office, verified by the provincial statistical bureau, and then reported to the NBS both individually and as provincial aggregates. Since April 2001 5,000 selected key industrial enterprises directly report their economic data to the NBS via the internet.⁸³ The tentatively chosen 5,000 enter-

footnote continued

units within the provincial, municipal and county statistical departments. These units are to check on the implementation of statistical laws and regulations, as well as on the accurate reporting of statistical reporting units. The inspectors, always in couples, have the right to enter the reporting unit and to demand presentation of the data sources behind the reported statistics. (Jilin People's Congress (Standing Committee), "Jilin sheng tongji guanli tiaoli" ("Jilin province stipulations on statistical administration"), 18 May 1997, art. 28–30, in China Infobank.)

80. At least in Fujian province the system of dispatching personnel to lower-level statistical offices extends all the way to the municipal tier. The NBS sent its first two staff to Fujian province in 1990, followed by another staff in 1993. The provincial statistical bureau in 1992 sent ten staff to seven municipalities, followed by another two staff to a further two municipalities in 1993. After 1993, the municipal statistical offices also began dispatching staff to the county tier (*Fujian shengzhi tongjizhi*, p. 191).

81. The statistical break also offered a chance to in one stroke eliminate all past exaggeration that may have had to be upheld to guarantee time consistency. The opportunity was apparently grasped by numerous localities, despite the occasional embarrassment (*Zhongguo tongji*, No. 9 (2001), pp. 15, 10). This and the next paragraph draw on *Zhongguo tongji*, No. 6 (1998) p. 23; No. 7 (1998), p. 29; No. 10 (1998), p. 20; No. 2 (2000), pp. 31f; No. 6 (2000), pp. 30, 35 and 38; No. 7 (2000), pp. 46f; and No. 2 (2001), p. 32.

82. NBS, "Guanyu zuohao tongji zhidu gaige de yijian" ("Suggestions on the reform of the statistical system"), 13 November 1998, issued together with the State Economic Commission and approved by the State Council General Office, in NBS, *New Guide to Industrial Statistics*, pp. 275–78.

83. For the starting date see an 11 April 2001 news item in China Infobank titled "Zhongguo wu qian jia gongye qiye lianwang zhibao xitong zhengshi kaitong" ("China's 5,000 industrial enterprise direct electronic reporting system officially began operation").

prises in 1999 accounted for 49 per cent of industrial assets, 43 per cent of industrial sales revenue and 83 per cent of industrial taxes and profit. The first reform reduces the potential for data manipulation by lower-level statistical departments, while the second reform eliminates all intermediaries. The NBS is thus adopting a core of important enterprises on which it can guarantee high quality data available quickly.

Hand in hand with the reduction in the number of enterprises on which the NBS collects detailed data goes a general switch to surveys on all activities that are not reliably covered through the traditional reporting system. In 1984 the NBS had already established urban and rural survey teams as administrative facilities directly under its control. Today it has rural survey teams in 857 of China's 2,109 counties, and urban survey teams in 226 of the 663 municipal or county-level cities. In 1994 the NBS began to add enterprise survey teams, which culminated in the establishment of 210 central city enterprise survey teams, linked by a computer network, in 1997 and 1998. The rural and urban survey teams conduct, among others, the annual urban and rural household surveys. The enterprise survey teams are in charge of the sample surveys of those enterprises that are not part of the group of directly reporting enterprises.⁸⁴

Dwight Perkins in 1966, also referring to previous work by Choh-Ming Li, concluded that “statistics are more accurate the greater the importance of having the data is to the regime, the fewer the number of units from which data have to be collected, the less backward these units are, and the greater the degree of control the state exercised over them. In addition, the political atmosphere, at least in 1958 and 1959, has had considerable influence on the willingness and ability of lower-level statistical workers to compile reliable data.”⁸⁵ From this perspective, the unambiguous signal by the CCPCC/State Council in 1998 to curb data falsification, the increasing codification of reporting processes and the expansion and strengthening of supervision, the focus of the NBS in industry on 5,000 core enterprises and approximately 160,000 directly reporting industrial

footnote continued

Similar reforms are taking place in other sectors of the economy, which have also moved to the new enterprise classification for directly reporting enterprises; in the real estate sector 3,000 enterprises now directly report to the NBS, and in wholesale and retail trade 1,200 enterprises (*Zhongguo tongji*, No. 12 (2000), p. 9).

84. On the issue of survey teams see NBS, “Guanyu jin yi bu gaohao qiye diaocha dui zuijian gongzuo de tongzhi” (“Circular on further implementing the enterprise survey team establishment”), 16 June 1997, in *Statistical Work Yearbook 1998*, p. 106; NBS, “Guanyu di yi pi zhongxin chengshi qiye diaocha dui jigou bianzhi fenpei fang’ an de tongzhi” (“Circular on the distribution of the first batch of central city enterprise survey teams”), 14 July 1997, in *Statistical Work Yearbook 1998*, p. 107; NBS, “Guanyu diaocha dui guanli tizhi gaige de shishi yijian” (“Suggestions on the administrative reform of the survey teams”), 8 December 1997, in *Statistical Work Yearbook 1998*, pp. 146–48; NBS, “Guanyu jin yi bu mingque quanguo qiye diaocha dui gongzuo renwu de tongzhi” (“Circular further clarifying the tasks of the nation-wide enterprise survey teams”), 11 December 1997, in *Statistical Work Yearbook 1998*, p. 141; NBS, “Guanyu di er pi zhongxin chengshi qiye diaocha dui jigou bianzhi fenpei fang’ an de tongzhi” (“Circular on the distribution of the second batch of central city enterprise survey teams”), 29 December 1997, in *Statistical Work Yearbook 1998*, p. 148; *Statistical Work Yearbook 1998*, pp. 46–49, 206–208; and <http://www.stats.gov.cn/zwgk.htm>. Total employment in the NBS survey teams runs to 80,000.

85. Perkins, *Market Control and Planning*, p. 222.

enterprises, all represent reforms that are likely to make the available data more reliable. The progress on the sample surveys appears slower, but unremitting.⁸⁶

Conclusions

How reliable then are today's output and economic growth statistics? They are both more and less reliable than the critics think. They are more reliable in that the NBS is unlikely to be purposefully falsifying data. As this report has argued, the evidence on falsification of aggregate nationwide output and economic growth data is invalid, or at least highly problematic and therefore unconvincing; the NBS as well as lower-level statistical departments furthermore appear to be making sincere efforts to prevent data falsification.

At the same time, today's output and economic growth statistics are also less reliable than one may think. The traditional statistical reporting system based on superordinate government bureaucracies was ill suited to capture the explosion of non-state units; the 1998 changes to the statistical reporting system address this problem, but it may take years before a fully functioning and reliable new economy-wide data collection system is in place. The adoption of international standards in measuring GDP is still ongoing. This implies that key aggregate nationwide data are not consistently measured over time, both in terms of enterprise coverage and in terms of variable meaning (definition).

These complications are not apparent in the published statistics. Much of the published data do not come with explanations as to which enterprises they cover or what the variables exactly mean. If explanations are available, they appear of dubious quality. The poor documentation could be a sign of severe negligence on the part of the statistical departments, in particular the NBS. But it could also be the case that not all 31 provinces uniformly apply the same set of statistical rules and regulations, perhaps due to local particularities or different time lags in adapting new rules issued by the NBS. The NBS itself consequently may not have available well-defined, time-consistent aggregate nation-wide values. Each aggregate nation-wide value should probably come with pages of explanations detailing individual provincial characteristics in data collection and compilation, characteristics the importance of which the provinces themselves may not even be aware of.

These are problems faced by every statistical system, but they appear more serious in China than in many other countries because of the

86. By mid-2001 an employee of the NBS claimed that steady progress had been made in the sample surveys of small industrial enterprises (those not reporting directly to the statistical departments), while sample surveys of the wholesale and retail trade as well as the catering business after four years of hard work are in place across all 31 provinces (*Zhongguo tongji*, No. 11 (2001), pp. 8–10). In the case of industry, the establishment of the appropriate sampling framework for the stratified sampling of approximately 8 m small industrial enterprises is a challenge in itself, in particular given the rapidly changing environment of economic transition with a fair amount of enterprise mergers, exits and new entries.

developing and transitional nature of China's economy. The recent reforms in the statistical system show that the Chinese leadership and the NBS are responding well to the challenge. First, recent reform measures are likely to have led to a strengthening of professional statistical practice across the bureaucracy; this may both curb the potential for data falsification and increase the spread of uniform statistical practices across the country. Secondly, the NBS itself now communicates directly with a small number of the largest enterprises in at least industry, construction and commerce. This guarantees a set of highly reliable statistics on, in terms of output, close to half of all production activities in these sectors. Data on the economy as a whole are likely to remain of much poorer nature until the survey system has stabilized and possibly even until economic transition has been completed. The unavoidably large margin of error in aggregate nation-wide data, such as industrial value-added or GDP, and the potential for perfectly justifiable adjustments to variable definitions and enterprise coverage imply that aggregate nation-wide data reported by the NBS should for the time being simply be regarded as one set of possible values out of a larger interval of equally correct data.