Research on China's economy makes frequent use of official Chinese statistics. Yet the data used are often poorly understood, or of poor quality to begin with. The overarching theme of this paper is to better understand the quality of official Chinese data. It traces the challenges economic reforms posed for the development of China’s statistical system, establishes a typology of the resulting data problems in official Chinese statistics today, and examines how these data problems are being addressed. While numerous examples are provided along the way to illustrate the challenges to China’s statistical system arising from economic reforms and the current data problems, special attention is given to GDP data as the aggregate measure of production activities in China.
Introduction

Research on China’s economy makes frequent use of official Chinese statistics, both in quantitative analysis as well as in predominantly qualitative arguments. Our understanding of China’s economy, including how China’s economy operates, develops, or undergoes structural change is frequently based on the analysis of numerical data. Our evaluation of economic transition in China or of the urgency of various economic problems ranging from state-owned enterprise reform to rural development all hinge on numerical data. Yet the data are often poorly understood, or of poor quality to begin with.

Early evaluations of the quality of Chinese statistics were largely positive. Thus Li Chow-Ming in 1962 wrote cautiously about improvements following the statistical debacle 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 such falsification in the end may not be in the interest of the Chinese 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” (p. 440). Gregory Chow in 1986 judged that “by and large Chinese statistics officials are honest” (p. 193).

Yet, by the late 1980s, Western researchers in their economic analyses increasingly noted the limitations of Chinese statistics. One example are the fixed asset values of industrial enterprises—needed in productivity analysis—which suffer from the inclusion of non-productive fixed assets and the lack of adjustment for inflation (Chen Kuan et al., 1988). Another example are the inadequate and inaccurate statistics on private rural enterprises (Ole Odgaard, 1990/91).


---

1 Some of the further literature on adjustments to fixed asset values includes Gary Jefferson (1992), Wing-Thye Woo et al. (1993), and Gary Jefferson et al. (2000).

2 In a much earlier article, Kenneth Walker (1982) tried to make sense of Chinese grain consumption.

The compilation of meaningful statistics is not facilitated by the ongoing process of economic transition and development. The problems in compiling accurate data in a transition economy have been widely noted for Eastern Europe and Russia by, for example, Kasper Bartholdy (1994, 1997), Vincent Koen (1996), or Susan Powers (1992). Carsten Holz (2002), for the case of China, elaborates on the institutional constraints on data quality in a transition as well as rapidly developing economy. Xu Xianchun (2000a, 2002) explains some of the imperfections in China’s GDP calculations today.

The conclusion that emerges from this literature is that many Chinese data are problematic. In the absence of perfect information about machinations of potentially tens of thousands of accounting staff in enterprises and statistical staff in government departments, let alone systematic nationwide alternative (independently collected) data on the state of China’s economy, it is impossible to prove beyond doubt that the aggregate Chinese official data speak the “truth.” However, it is equally impossible to show official data falsification at the nationwide level through internal consistency checks as long as one does not fully understand the nature (meaning) and quality of the data one wishes to check on, or of the data used to double-check on the data whose accuracy is being questioned. The discussion on, for example, the “true” real growth rate of China thus, while interesting and fruitful in promoting an increased understanding of Chinese data, is unlikely to be resolved in the short run.

This paper moves beyond the exploration of individual areas of statistics and tries to bypass the question of the presence and extent of data falsification in the case of nationwide data. The overall purpose of this paper is to provide a general appraisal of the quality of official Chinese data and to increase awareness of the various factors that impact on data quality. The paper starts out by documenting just how daunting the challenges to official data compilation in China have become in the course of the economic reform period. Economic reforms affected the quality of statistical data through no less than six different channels. A consequence of the economic reform challenges is a range of data problems. The third section proceeds to establish a typology of these data problems in official Chinese statistics today, illustrating each type of problem with examples.

3 Until 1997, the National Bureau of Statistics (guojia tongji ju) in English called itself “State Statistical Bureau.” In the following the new English name is used throughout. The terms “statistical authority,” “statistical system,” or “statistical departments” are used interchangeably to denote the totality of all statistical offices in China ranging from the center (NBS) to the county-level statistical department and the township statistics official.
But China’s statistical system has not stood still. The fourth section examines how some of the data problems have already been addressed while others are scheduled to be addressed in the near future through further innovations to statistical data compilation. The frequently noted and critically commented upon Gross Domestic Product (GDP) data are the subject of the fifth section, which makes a broad attempt to evaluate the quality of these data. The final section concludes.

**Economic Reforms and The Implications for Statistical Work**

According to the PRC Statistics Law of 1983 and (the revised version of) 1996, the “basic task” (jiben renwu) of statistical work in China comprises the following standard tasks of a statistical authority: to undertake statistical investigation (tongji diaocha) and statistical analysis (tongji fenxi) of the national economic and social development, to provide statistical data (and, since 1996, statistical advice), and to supervise all statistical work (NPC 8 Dec. 1983 and 1 May 1996, Art. 2). The administration of statistical work follows the principle of unified leadership with responsibility allocated to all levels of government; the National Bureau of Statistics (NBS) as an institution of the State Council exercises unified leadership and coordinates statistical work nationwide, while all levels of government, government departments, enterprises and administrative facilities are to establish statistical organizations or to designate statistical personnel in accordance with the need to fulfill statistical tasks (Art. 4). The formal tasks and organizational arrangements have remained unchanged for the past two decades. Yet, with the introduction of market-oriented economic reforms in China beginning in 1978 and the subsequent rapid economic growth, the nature of compiling data on the national economic and social development has changed dramatically over time.

Economic reforms have posed a number of challenges for the statistical system, including the following: (i) rapid growth in new, often small and private enterprises outside the official statistical reporting system; (ii) the adoption of novel statistical concepts and variables in response to new government policy needs or following the adoption of international standards; (iii) changes in the meaning of variables on which the NBS collects data due to institutional innovations originating in other government departments (outside the NBS); (iv) data falsification from the bottom up in response to changing incentive structures for local reporting units and governments; (v) changes in the classification of statistical units in response to changing political considerations of the Chinese leadership; and (vi) the gradual undermining of some of the statistical reporting structures due to shifting incentives of reporting units and changes in local government administration.

**Rapid growth in statistical units outside the traditional reporting system**

The introduction of market-oriented economic reforms in the countryside in the late 1970s and then the industrial reforms in the mid-1980s posed a severe challenge to the traditional statistical reporting system. The pre-reform statistical reporting system was based on comprehensive periodic reporting that encompassed the communes in agriculture and the predominantly state- and collective-owned production units in the other four material production sectors. Government departments in charge of agriculture and enterprises reported the statistics on their production units up to the next higher-level government department within the same xitong (bureaucracy in charge of a particular area, organized from central to
county government), and simultaneously to the local statistical bureau, which then passed the data on to the next higher-level statistical bureau within the statistical xitong.

By the mid-1980s, reforms in the countryside had led to the abandonment of the commune system and thereby the replacement of the (in 1978) 53,000 communes as basic statistical reporting units by close to 1m villages if not 191m rural households in 1985 (Statistical Yearbook 1981, p. 131; 1986, p. 23). The gradual acceptance of private entrepreneurs first, in 1982, in the form of the “self-employed” (getihu, employing less than eight persons), in the following also called “individual-owned enterprises,” and then, in 1993, in the form of “private enterprises” (siying qiye, employing more than seven persons), led to the proliferation of individual-owned and private enterprises. Much of a ten-fold increase in the number of rural enterprises between 1978 and 1985 to approximately 12m is due to the rise in individual-owned enterprises (Township Enterprise Yearbook 1978-1987, p. 570). At the same time, the opening to the outside world with the establishment of four special economic zones in 1979 and 1980 and the opening of fourteen coastal cities to foreign investment in 1984 led to a rapid rise in the number of foreign-funded enterprises in China.

Consequently, the statistical reporting system struggled to measure the economic activities of several hundred million farmers and the many newly established enterprises. For example, in industry, in 1978 the NBS compiled basic data—including the number of enterprises, their gross output value, net material product (since 1993 value-added), and employment—on 348,400 enterprises, 24.02% of which were state-owned and 75.98% collective-owned (Table 1). The total number of industrial enterprises rose to 437,200 in 1984 thanks to an increase in the number of collective-owned enterprises. Data for the years prior to 1985 do not include village-level collective-owned industry and individual-owned industry.

The inclusion of village-level collective-owned industry and individual-owned industry in 1985 led to a more than ten-fold rise in the total number of industrial enterprises from the 437,200 in 1984 to 5,185,300 in 1985. Individual-owned industrial enterprises (on which separate data are available) accounted for just 1.85% of gross output value of industry in 1985, but for 64.56% of the total number of industrial enterprises. The retrospectively adjusted output data suggest that individual-owned enterprises were non-existent in 1978, and the case may not be much different for village-level collective-owned enterprises. The number of individual-owned industrial enterprises further doubled between 1985 and 1987, as did their (still small) share of gross output value of industry. Since then the number of individual-owned industrial enterprises has not changed dramatically, but their output share rose to 18.18% in 1999, the last year for which these data are available.

Table 1 about here

---

4 The 1982 constitution of the PRC referred to the “self-employed” as a “supplement” to the socialist public economy. The 1993 constitution granted private enterprises the same status. The 1999 constitution finally upgraded the status of both the self-employed and private enterprises to “important constituent elements of the socialist market economy.”

5 Village-level collective-owned enterprises in 1986, the first year for which separately listed data on these enterprises are available, accounted for 7.49% of gross output value of industry and 8.42% of the total number of industrial enterprises (Statistical Yearbook 1988, 324). Gross output value data are reported since value-added data are not available across all ownership forms.
Only a fraction of these small enterprises report regularly to the statistical authority; those that do report detailed balance sheet and profit and loss account, as well as output and employment data. The directly reporting enterprises in the industrial sector up through 1997 comprised all industrial enterprises with independent accounting system registered or organized at township level and above. Since 1998 the directly reporting industrial enterprises comprise all industrial state-owned enterprises (SOEs) plus all non-state industrial enterprises with independent accounting system and annual sales revenue in excess of 5m Renminbi (RMB). The self-employed (individual-owned enterprises) are excluded from the directly reporting enterprises throughout since they are not officially regarded as enterprises.

Table 1 also reports the number of directly reporting industrial enterprises since 1986, with no earlier data available; given the invariability of these numbers over time (from 1986 through 1997, before the definition of this category changed), the numbers for the early 1980s are likely to be quite similar. One may thus surmise that in the early 1980s almost all industrial enterprises in China were included in the group of directly reporting industrial enterprises. With the inclusion of village-level collective-owned industry and individual-owned industry in the statistics on all industrial enterprises in 1985, the share of the directly reporting industrial enterprises in all industrial enterprises in 1986 was only 5.65%. In other words, the share of industrial enterprises from which the statistical authority directly collects detailed data fell drastically in the mid-1980s. The much lower share remained stable until the statistical break in the category of directly reporting industrial enterprises in 1998, when it fell further to 2.07%. The directly reporting industrial enterprises also account for a decreasing share of industrial gross output value, falling from above 90% in 1979 to less than 60% in 1999 (Figure 1).

The fact that the NBS did not collect data on village-level collective-owned and individual-owned enterprises before 1985 reveals some of the difficulties the NBS encountered in adjusting its data compilation practices to the reform economy. When it became clear that the official data were beginning to miss out on a growing share of the economy, the NBS began to make efforts to include the missing enterprise categories, with retrospective adjustments to earlier gross output value data but not enterprise numbers. The new coverage since 1985 implies that by the mid-1980s the NBS had to increasingly rely on other sources than direct reports to compile aggregate nationwide data on industry. Furthermore, with many of the directly reporting industrial enterprises of small size and with poor or incomplete accounts, the aggregate quality of the reported data could only fall. In the early years, when the directly reporting enterprises still accounted for the very largest part of
industrial output in China, this may not have mattered much, but by the late 1980s, and then particularly since the early 1990s, collecting high-quality data on those many millions of industrial enterprises not reporting directly to the NBS gained increasing urgency.

Adoption of novel statistical concepts and variables

The process of transition also involves changes in the variables on which data are collected. Statistical compilation in China traditionally served a planning apparatus which needed data on physical inputs and outputs, machinery and technology levels. In a more market-oriented economy, macroeconomic policy, replacing much of central planning, requires data on such variables as GDP, economy-wide employment and unemployment, price indices, or on social security matters, which previously were handled by rural communes and individual enterprises rather than by the government. The government may also be interested in the capacity of the tertiary sector to absorb workers laid-off by industrial SOEs; in the Material Production System (MPS), which guided statistical work prior to 1993, some services were ignored altogether, while others were subsumed in the five material production sectors (agriculture, industry, construction, commerce, and communication).

In 1985, the NBS began work on compiling tertiary sector as well as overall production/income approach GDP statistics, in 1989 it for the first time experimented with expenditure approach GDP statistics, and in 1992 it began to calculate GDP independently, rather than obtain it through manipulations of national income accounts data under the MPS. In 1993, the NBS ceased to compile national income accounts data under the MPS and switched to the United Nation’s newly revised “1993” System of National Accounts (SNA). Production/income approach data for earlier years were published retrospectively to cover first 1978-1984, and then 1952-1977; expenditure approach GDP data are available for the years since 1978 only.

The SNA asks for the compilation of data on numerous new variables that neither statistical personnel nor reporting units were previously familiar with, as well as for the redefinition of some variables already in use in China prior to the switch to the SNA. The key example is GDP itself as a measure of national income, in the production/income approach calculated as the sum of value-added across the economy. The concept of value-added (zengjiazhi) replaced the concept of net material product (jingchanzhi). The sum of net material product from all five material production sectors (agriculture, industry, construction, commerce, communication) equals national income as defined in the MPS. The value-added of all productive activities across the economy, on the other hand, adds up to GDP.

Although in a material production sector, such as industry, both indicators, net material product and value-added, are derived in the production approach by deducting intermediate inputs from gross output value, what is included in intermediate inputs differs. For the case of industry,

---

7 The production and income approach calculations of GDP in China are not conducted independently. For those sectors in which the production approach is not feasible, income approach data are used. See more details below.

net material product = gross output value – intermediate inputs,
where intermediate inputs include depreciation but exclude service payments to non-
material production units (such as interest payments); and

value-added = gross output value – intermediate inputs,
where intermediate inputs exclude depreciation but include payments for services.

Industrial value-added, in contrast to industrial net material product, thus includes
depreciation but excludes all services.

While the switch from the MPS to the SNA in the case of the material production sectors
is primarily a task of getting production units to manipulate accounting items in new ways,
some tertiary sectors, on the other hand, were newly captured in the GDP statistics and the
hitherto unfamiliar income approach to the calculation of value-added had to be established
first.

The difficulties in compiling value-added may be most evident in the tertiary sector, on
which aggregate data were collected only beginning in the mid-1980s. The tertiary sector
census of 1992-93 led to major revisions of tertiary sector value-added for the years 1978-93.
Table 2 reports the retrospective percentage upward revisions to tertiary sector value-added,
to the value-added of the three major tertiary sector components, and to aggregate GDP. For
the tertiary sector sub-sector of wholesale and retail trade and catering, these upward
revisions were as high as 88.71% in 1992, with a maximum impact on GDP of 9.99% in
1993. Going back to earlier years, the gradual decline in the percentage corrections suggests
that while accurate revisions could be made for the most recent years, perhaps 1992 and
1993, the revisions to earlier years’ data may have been somewhat arbitrary, with an
assumption that only the most minor revisions needed to be made to 1978 data.

Table 2 about here

Yue Ximing and Zhang Shuguang (2002) list a number of factors which they think
contribute to an underestimation of tertiary sector value-added until today. These include
housing provided by work units at below-market rates, unreported income in the tertiary
sector (where value-added data are based on the income approach), and a lack of data on
small-scale services and on only recently introduced services (such as accounting or stock
broker services). According to an unofficial source, the World Bank expects the 2003 tertiary
sector census to yield yet another major (upward) revision of tertiary sector GDP data.

While tertiary sector statistics may be problematic until today, problems with the switch
to the SNA exist in all areas of Chinese statistics. China’s GDP still does not fully accord to
SNA stipulations (as probably no other country’s GDP calculations do). Xu Xianchun,
currently head of the NBS National Income Accounts Division, in a year 2000 article
describes five areas where China’s GDP statistics are still far from perfect, namely in the
imputation of rental values to housing, the treatment of fiscal subsidies, the underpricing of
welfare services provided by work units, the overstatement of rural industrial gross output
value, and the overestimation of livestock products (Xu Xianchun, 2000a, with an English
version published in 2002). In a 2001 article he further lists eleven differences between SNA
stipulations and Chinese practice, such as the inability of the NBS to impute the value of illegal activities.\(^9\)

**Redefinition of economic variables by other government departments**

The adoption of the SNA was accompanied by the introduction of a new accounting system in 1993 and a new tax system in 1994. Both represent progress towards institutions better suited for a market-oriented economy. Yet they also have implications for statistical work.

The introduction of the new accounting system marked the transition from “sources” and “uses” of funds to the type of balance sheets used in market economies (FM 30 Nov. 1992a/b). Many traditional concepts, such as fixed-quota working capital and above-quota working capital were simply abandoned, while variables such as assets, long-term investment, or intangible assets were newly introduced. The new tax system which took effect on 1 January 1994 led to the redefinition of a number of variables, including sales revenue, which since 1994 is reported net of the newly comprehensive value-added tax.

The definition of variables such as balance sheet and profit and loss account items are beyond the control of the NBS. The NBS simply collects data on these variables as defined by the Finance Ministry, with utter disregard for time consistency in statistical data. Statistical authorities in other countries are likely to face the same problem, but changes in accounting and tax systems are likely to be much less severe in developed countries than in a developing country undergoing rapid economic transition.

Since 1993/94 both the accounting and tax system have remained remarkably stable, yet some of the changes introduced in 1993/94 were not implemented immediately, such as the proper valuation of land in the balance sheet, but spread over several years. The 2000 revisions to the accounting system were minor compared to the 1993 revisions, and the tax system has experienced no further major changes after 1994, so that balance sheet and profit and loss account items are gradually stabilizing.

**Data falsification at lower-level tiers**

Over-reporting of output data has been an issue for many years.\(^{10}\) As early as 1989, the NBS Industry and Transportation Division asked how big the “water content” (shuifen) in industrial growth statistics is? The answer was “a little bit” (yidian), and the explanation for the exaggerated output statistics predominantly a technical one focusing on compilation methods. By the late 1990s, over-reporting of output data has become an issue of outright data falsification, well documented for lower-level tiers in the countryside. It is viewed as an immediate result of economic reforms that have led to local leaders being evaluated, at least

\(^9\) While many of the differences between SNA stipulations and Chinese practice appear rather minor, illegal activities, if included in GDP, might lead to a drastic increase; Italy in 1987 adjusted its GDP figures upwards by 18% in order to newly account for the shadow economy (The Economist, 1 May 1997).

\(^{10}\) Data falsification is an issue not only for output data. Enterprise balance sheet and profit and loss account items may be equally affected, but the NBS has no means to evaluate and adjust these data. On the over time varying incentives to misreport profit, see, for example, Carsten Holz (2003b, Chapter 3) for the case of industrial SOEs.
in part, according to their economic achievements. Thus Zhao Peng (1998) 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.

Output exaggeration is most prevalent in the countryside because the official in charge of statistics at the township level and the statistical department at the county tier often have to rely on estimates rather than comprehensive reporting. For example, in the case of industrial non-state enterprises with annual sales revenue below 5m RMB, and all self-employed in industry, the statistical departments attempt to collect data on the number of such enterprises (units), their gross output value, tax remittances, paid-in capital, and the year-end number of laborers. If data on gross output value are not available, sales revenue may be substituted. The data collection methods comprise simple report cards, “processing of related data” (genju xiangguan ziliao jiagong zhengli), and sampling. The quality of these data are likely to be very low, often no more than guesstimates, and easily subject to manipulation.

The evidence on data falsification suggests that it is mostly committed by non-statistical personnel (such as a reporting unit), or by a local leader overruling the local statistical bureau. County, municipal, and provincial statistical bureaus then have all been reported to “squeeze the water content” out of the data they receive from the immediate lower-level government’s statistical bureau. The provincial data themselves thus are frequently already the result of revisions, the foundations of which are not publicized.

Nevertheless, at the national level, the sum of production/income approach GDP across all provinces in most years still exceeds the nationwide total; the same is true for expenditure approach GDP. Provinces supposedly over-report output (or under-adjust exaggerated lower-level tier data) in order to meet provincial growth targets; the NBS consequently adjusts provincial data when deriving the nationwide total. Table 3 shows the extent of adjustments in production/income approach GDP and expenditure approach GDP.

In both approaches the extent of downward revisions to the sum of provincial GDP by the NBS has increased from 1997 through 2000. In the production/income approach, the NBS is revising downward provincial primary sector and in particular tertiary sector value-added; value-added in industry, on the other hand, appears to be reported by the provinces rather reliably, presumably thanks to the rigorous reporting system in place for the directly reporting industrial enterprises which account for approximately 60% of industrial value-added. In the expenditure approach, the NBS systematically revises provincial household consumption upward and gross capital formation, particularly inventories, downward. Some of the upward revisions to household consumption could be due to the NBS’ reliance on nationwide retail sales statistics to calculate part of household consumption, while provincial statistical bureaus may rely primarily on household income surveys which tend to be more problematic in terms of capturing total consumption. For details on production/income and expenditure calculations see Xu Xianchun (2000).

---

11 For the report form see NBS Industry and Transport Division (1999, p. 38).
12 For numerous instances of data falsification by lowest-level governments in the countryside see Cai Yongshun (2000). On the incentive structure of local cadres also see Maria Edin (1998).
13 See Zhongguo Tongji, No. 7/1999, p. 12. The statistical bureau’s section in the Shaanxi Yearbook 2001 (p. 310) proudly announced that the provincial statistical bureau did not have to correct he GDP data reported up for the first six months of 2000 in the case of half of the municipalities in Shaanxi. A year later the provincial GDP data were almost identical to the sum of the municipal data (Shaanxi Yearbook 2002, p. 337).
14 Nationwide pure income approach GDP data are not available; only provincial-level data are.
15 Some of the upward revisions to household consumption could be due to the NBS’ reliance on nationwide retail sales statistics to calculate part of household consumption, while provincial statistical bureaus may rely primarily on household income surveys which tend to be more problematic in terms of capturing total consumption. For details on production/income and expenditure calculations see Xu Xianchun (2000).
downward adjustments to provincial production/income approach GDP in 1993 are due to the fact that 1993 provincial-level data, published a year late, already incorporate the retrospective upward revisions to GDP following the tertiary sector census, while the nationwide data do not.

Table 3 about here

While the NBS offers no detailed explanation on how it adjusts provincial GDP data, Liu Hong, the then head of the NBS in February 2000 stated in just two sentences that the NBS adjusts provincial GDP data based on two sets of information. It contrasts provincial GDP data with key economic data obtained through sample surveys in each province. The NBS also has available data on variables related to GDP, and assumes that the values of these variables cannot grow at a speed that is much different from that of GDP.\(^{16}\) (China Infobank, 29 February 2000)

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 gross output value data on the collective-owned and private economy for the years 1991 through 1994 downward by up to 20%.\(^ {17}\) That data on SOEs were not retrospectively revised downward may be due to the fact that all SOEs are regularly reporting to the statistical authority, whereas not all enterprises in non-state ownership are. Data on those enterprises not subject to the regular reporting system appear to have until the late 1990s all too often simply been guesstimates, thereby opening the doors to data falsification.

Zhang Sai, then head of the NBS, on 23 August 1995 explained to the Standing Committee of the National People’s Congress that the main purpose of the upcoming 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.” Reports on data falsification then became standard fare in the NBS journal Zhongguo Tongji (China Statistics) in the late 1990s.

In 1997 the issue of data falsification finally came to a head. 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 (CCPCC/SC, 16 Feb. 1998), warning local officials not to, because of a discrepancy between reality and the targets they are responsible to achieve, falsify statistics or command the falsification of statistics; Party and administrative disciplinary measures are threatened. Party committee and government of each province and all central government departments are to conduct one-time investigations to uncover the problems in the statistical work of their province/department with reports due at the Party center and the State Council by the end of 1998.

\(^{16}\) Assuming that the provincial data are indeed inflated, and that the NBS knows by how much, the NBS at this occasion has the opportunity to falsify data by not fully deflating, or by over-deflating the provincial data in deriving the aggregate nationwide values.

\(^{17}\) Gross output value data of collective-owned enterprises, individual-owned and private enterprises, and enterprises in other (non-state) ownership for 1994 were retrospectively adjusted downward by 19%, 25%, and 16%, with similar or lesser revisions for the years 1991 through 1993. (See Carsten Holz and Yi-min Lin, 2001a.)
The NBS immediately followed up with its own circular (NBS, 19 March 1998) in which the CPCC/State Council document was billed the third big step in the fight against data falsification since a 1962 CCPCC/State Council circular on strengthening statistical work and the passing of the 1996 Statistics Law. It urges provincial statistical bureaus and State Council departments to concentrate their efforts on finding and publicizing a few big cases of data falsification; statistical work especially at the county and the township level is to be strengthened. A final report on local and departmental efforts was due at the NBS by end-October 1998. The NBS together with other government departments also announced the dispatch of investigative teams to the provinces and departments in the second half of 1998. The findings triggered an (unpublished) criticism by the NBS and the State Council of some localities in 1999, and the further dispatch of NBS investigative teams in October and November 1999 (NBS, 25 Oct. 1999).  

Individual provinces responded to the central circulars by issuing their own instructions to investigate and prevent the falsification of statistics, such as a circular by the government of Shaanxi Province (1999) requesting an evaluation of the quality of provincial GDP data, and assigning responsibility for the accuracy of GDP data to the heads of all statistical departments. Following the spade of regulations and presumably re-evaluations of the quality of their statistics by all lower-level tier governments and government departments, the issue of data falsification largely disappeared in the Chinese press. In the 2001 and the 2002 issues of Zhongguo Tongji the wave of sensational reports on data falsification has given way to articles on technical and definitional problems of Chinese statistics.  

Impact of political considerations on statistical work

In 1998, the NBS together with the Industrial and Commercial Administration restructured the registration-based enterprise classification system in order to better take into account recent developments, including the evolution of new organizational forms, the growth of foreign direct investment in China, the growth of township, private and individual-owned enterprises, and the complexity of joint operations and mergers between enterprises across regions and economic sectors. For example, the 1993 Company Law brought about the establishment of limited liability companies and stock companies; less formal stock cooperatives (employee shareholding companies) also began to take hold. In the published statistics, however, the NBS continued to adhere to an ownership-based classification system with some but not all of the formal companies in a separate category.

In 1998, the NBS accommodated the changes in the real economy by adopting a two-tier categorization system. A distinction between the public and non-public economy constitutes

---

18 The unpublished criticism is entitled “Guowuyuan bangongting zhuanying guojia tongjiju guanyu yi xie difang zhongyao tongji shuju shishi wenti tongbao the tongzhi” (Circular by the State Council passing on the NBS report on important statistical data inconsistencies in some localities), as reported in a China Infobank entry of 22 Nov. 1999. The Statistical Work Yearbook 1998, p. 43, contains a short paragraph on the preparatory work.

19 In January 2002 the economist Fan Gang came to the defense of Chinese statistics, stating that he did not think Chinese economic data were exaggerated; to 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. (Xinbao [Hong Kong Economic Journal, a daily newspaper], 9 Jan. 2001, p. 3).
the first level of classification, with a further ownership-based second-tier classification; clear rules were established on how to handle different organizational forms. The ownership classification system furthermore was augmented by a separately published new category of SOEs defined in a way that maximizes the size of the state-owned sector. This is explicitly in response to General Secretary Jiang Zemin’s declaration at the 15th Chinese Communist Party Congress in fall 1997 that “the meaning of public ownership needs to be fully recognized. Public ownership comprises not only state ownership and collective ownership, but also the state and collective share in mixed ownership forms.”

The category collective-owned enterprises continues to include all stock co-operatives independent of their actual ownership, and the category of SOEs now turned into “state-owned and state-controlled enterprises.” The following equation summarizes the coverage of this newly defined SOE category:

“State-owned and state-controlled enterprises” =  
pre-1998 definition SOEs, i.e.,  
- pure SOEs (established or organized in accordance with the 1988 SOE Law),  
- SOE-SOE joint operation enterprises, and  
- solely state-owned limited liability companies,  
+ all (other) shareholding companies (i.e., limited liability companies and stock companies) in which the state has a controlling share,  
+ all (other) enterprises in which the state has a stake (or companies in which the state has less than a controlling stake), with the economic data counted towards the SOE category in proportion to the state’s equity share, where the share of legal persons in paid-in capital is ignored for the purpose of the calculation.

The new SOE category (“state-owned and state-controlled enterprises”) thus is as large as possible without unduly distorting the meaning of the SOE ownership category. It is clearly larger than the pre-1998 SOE category (on which publication of data ceased in 1998). The share of the SOE category in the value-added of the directly reporting industrial enterprises consequently increased from 46.35% in 1997 to 57.03% in 1998, although some of this increase could be due to the simultaneous redefinition of the group of directly reporting industrial enterprises. While the changes in the registration-based enterprise classification system reflect the necessity to respond to new institutional developments in the economy—the introduction of the formal company system in 1993, and the evolution of the stock co-operatives through the privatization or semi-privatization of some state- or collective-owned

---


21 The term “(other)” means that solely state-owned limited liability companies are not double-counted in the group of pre-1998 definition SOEs and then as shareholding companies in which the state has a controlling share; the same holds for the SOE-SOE joint operation enterprises vis-a-vis the enterprises (non-companies) in which the state has a stake. For further elaboration on the changes in the registration-based classification and the content of the different ownership categories on which data are published, including the detailed definition of “state-controlled,” see Carsten Holz and Yi-min Lin (2001b).

22 The Industrial Yearbook 2001 and 2002 published very limited data on the “pure” SOEs, a sub-category of the pre-1998 definition SOEs, for the years since 1999.

23 See Statistical Yearbook 1998, p. 444; 1999, p. 432. The data on other ownership groups are not correspondingly reduced. State-controlled shareholding companies and other enterprises in which the state has a (minority) stake are still included in one of the various non-SOE categories. The (other) state-controlled shareholding companies and the (minority) state of the state in all other enterprises thus are double-counted since 1998. For details, see Carsten Holz and Yi-min Lin (2001b).
enterprises—the continued adherence to the simple ownership classification when reporting data on, for example, industrial units, reflects a political choice, with the redefinition of the SOE category being a drastic example of how political considerations shape the organization of statistics in China.

**Shifting interests of reporting units and changes in local government administration**

In the socialist economy, filing regular reports with the statistical authority may simply have been one of many bureaucratic tasks enterprises faced. In a market economy, however, paying enterprise personnel to report to the statistical authority is not in an enterprise’s immediate interest. The quality of such reports is likely to fall, or reports may not be made at all. The revised Statistics Law of 1996—but not the earlier one of 1983—explicitly recognized the limitations of imposing statistical reporting tasks on enterprises: “Regular reporting by all primary-level units must be strictly limited.” (Art. 10) All data that can be collected through other means, such as surveys, may not be collected through regular comprehensive reporting.

Similarly, as local fiscal budgets became ever tighter, local governments have been reported to save money on statistical personnel. Some townships, which according to the Statistics Law (of 1983 and 1996) are supposed to either employ a statistician or have some person fulfill the tasks of a statistician along with other tasks, no longer bother with statistical work. Central government departments that the NBS is relying on to collect statistical data in their xitong have little interest in doing statistical work for the NBS paid for out of their own, tight budgets. As the strict centralism of the socialist bureaucracy has weakened over time, so have the traditional channels of statistical data compilation.

The NBS has a monopoly on authorizing the compilation of statistics, which includes a monopoly on authorizing the release of statistics (enshrined in the revised Statistics Law of 1996, Art. 14). Competition in the compilation of accurate data thus is non-existent in China. Government employment furthermore still tends to be permanent, with dismissal for poor performance unlikely. Consequently, published data are often presented sloppily, with typos and errors common, and statistical breaks not noted or explained. It also means that the speed and degree of innovation in the statistical system tends to be low.

**Typology of data problems**

The challenges arising from economic reforms, i.e., the rapid growth in statistical units outside the traditional reporting system, the adoption of novel statistical concepts and variables, the redefinition of economic variables by other government departments, data falsification at lower-level tiers, political considerations, and shifting interests of reporting units and changes in local government administration all impact on the quality of official Chinese data. Some of the resulting data problems have already been mentioned above to illustrate the effect of particular reforms on data compilation. The following is summary categorization of the various data problems, followed by examples:

1. Definitions of variables and coverage of statistical units for a particular variable:
   1.1 lack of definition, or unexplained redefinition
   1.2 definition such that variable not suitable for use in economic research
(1.3) meaning of variable hollowed out over time due to changes in economy
(1.4) time inconsistency due to abandonment of old variables and adoption of new variables
(1.5) meaning improves over time
(2) Errors:
(2.1) random measurement errors
(2.2) systematic measurement errors
(2.3) sampling errors
(2.4) shortcomings in the manipulation of statistics
(3) Data falsification
(3.1) by reporting unit
(3.2) within data collection system
(3.2.1) by the statistical authority’s institutional hierarchy (xitong)
(3.2.2) by other government departments collecting data on behalf of the NBS
(4) Aggregation across provinces of provincial-level data that measure different things in different provinces
(5) Mistakes in data presentation

The lack of definitions is a perennial problem of Chinese statistics. The Statistical Yearbook, for example, has never offered a concise definition of the since 1998 newly introduced SOE category “state-owned and state-controlled enterprises.” Labels may even remain unchanged while the content changes and no definition is available ever. For example, the “per capita cash expenditure of rural households” in the NBS monthly publication China Monthly Economic Indicators of August 2001 was redefined without this being apparent.24

All too often, much needed explanations are simply lacking. The industrial employment data provided in the industry section of the Statistical Yearbook in recent years match neither total industry nor the directly reporting industrial enterprises, even though they are reported in the section that covers (only) these two groupings. Until 1997, furthermore, these data included all employees, independent of whether they were actually working or not. This became apparent in 1998 only, when an explanatory note in the Statistical Yearbook (Statistical Yearbook 1999, p. 140) stated that beginning in 1998 employment figures are limited to those employees who are actually working in their work unit (zai gang zhigong). Employment in industrial SOEs fell by 32.65% between 1997 and 1998.25

When definitions are provided, the NBS does not always make changes to variable definitions explicit, even in cases where some of the decision to redefine variables rests with the NBS. One example of when the redefinition of a variable was made explicit is the statistical break of industrial gross output value in 1995. The Statistical Yearbook 1997

24 According to the NBS publication China Monthly Economic Indicators 17, no. 8 (2001), p. 72, per capita cash expenditure of rural households fell from 2338 RMB in 1999 to 2140 RMB in 2000, without any indication of the cause of this decline. The identical two data points can be found in the respective Statistical Yearbook (Statistical Yearbook 2000, p. 330, and 2001, p. 322). But the more recent volume besides the 2000 value also reports a revised 1999 value of 1917 RMB. The tables in the Statistical Yearbook furthermore contain a breakdown of the cash expenditure into sub-categories. A comparison of the two tables in the two different yearbooks reveals that the variable per capita cash expenditure of rural households was redefined in 2000 to exclude the sub-category “expenditure for savings and credit.”

25 For the data see Statistical Yearbook 1998, p. 432, and 1999, p. 422. Part of the reduction can also be due to enterprise closures and partial lay-offs. A further complication is that the length of the workweek has changed over time. For example, it was reduced from 48 hours per week to 44 hours per week on 1 March 1994, and further to 40 hours on 25 March 1995 (Jefferson et al., 2000).
marked the 1995 and 1996 data as following “new stipulations” (the key change being the new exclusion of the value-added tax from the gross output value data).

In the case of the directly reporting industrial enterprises, on which the *Statistical Yearbook* reports data only once, in each issue for the most recent year, statistical breaks tend to go utterly unnoticed, such as the statistical break in sales revenue in 1994. In statistical compendia presenting (limited) time series data on the directly reporting industrial enterprises, for example, *Seventeen Years of Reform* or the *Industrial Yearbook*, statistical breaks are not noted. While the redefinition of industrial gross output value or sales revenue may in part reflect some choice of the NBS in response to changes imposed by other institutions (in this case, the Finance Ministry), in many instances the NBS itself does not even have the means to understand changes in variables determined by other institutions; the continuous revaluations of net fixed assets or redefinitions of profit are beyond documentation.

When the definition of variables, or of the coverage of variables, is entirely under NBS control and the NBS actually provides a definition, the explanations are sometimes wrong. Labor productivity across all industrial enterprises with independent accounting system at township level and above, as reported in the *Statistical Yearbook*, fell from 47,824 RMB of value-added per person in 1993 by 63% to 17,648 RMB in 1994. In both years labor productivity was defined as value-added per average annual employment, when the 1993 value was apparently still based on gross output value, as in the earlier years.\(^{26}\)

At other times, variables turn out to have little or no economic meaning. For example, “rural household savings” (*nonghu chuxu*) by unpublished definition denotes all household deposits at the Rural Credit Cooperatives, rather than the savings of rural households. This institution-based definition does not consider that the state commercial banks (in particular the Agricultural Bank of China) and other formal as well as informal financial institutions have branches (or are located) in the countryside; farmers furthermore may increasingly save at the branches of financial institutions in the cities.\(^{27}\) This is also a prime example of how the meaning of a variable has been hollowed out over time due to changes in the economy. In the pre-reform economy, with during some periods the People’s Bank of China and the Rural Credit Cooperatives as the only financial institutions in China and farmers rarely if ever traveling to cities, the Rural Credit Cooperatives were indeed the sole repository of rural household savings.

Adoption of the SNA is a prime cause of inconsistencies in various time series. For example, gross output value was redefined in 1995 to, among others, newly exclude the value-added tax. This creates a statistical break in the gross output value time series, but does not affect the reported value-added data to which tax data are now added separately. Of three further changes to the definition of gross output value in 1995, only one affects value-added, and the NBS deems it of minor importance.\(^{28}\)

\(^{26}\) See *Statistical Yearbook 1994*, pp. 397, 417; *1995*, pp. 401, 421. In the NBS publication *Seventeen Years of Reform* this misreporting is consistently repeated for each province (not necessarily in the same year), even though the note to each provincial table presenting data from 1978 through 1995 states that labor productivity data are based on value-added. For example, Shanghai experienced a more than 70% “drop” in labor productivity in 1992 (p. 422).

\(^{27}\) See, for example, *Financial Yearbook 2001*, pp. 376, 481. The NBS itself does not collect financial sector data but relies on the People’s Bank of China to fulfill this task.

\(^{28}\) See, also, NBS National Income Accounts Division, 1997a, p. 36, and for more details in the case of industry Carsten Holz and Yi-min Lin (2001a).
Various accounting time series are also not consistently defined across the reform period. Net fixed assets of industrial enterprises, for example, while not being adjusted annually for inflation, are periodically re-valued, such as in the early and mid-1990s when the State Council undertook a major endeavor to determine the accurate value of state assets across first all SOEs, and then enterprises in other ownership forms. The definition of profit also changes with every revision of the accounting system (such as in 1984/85, 1989, 1993, and 2000). The volume of profit is further affected by any change in depreciation or tax rates; depreciation rates were adjusted several times in the reform period.29

Over time, more or newly meaningful variables have been added, such as GDP, including tertiary sector value-added, while statistics are no longer compiled on other variables, such as net material product. GDP data are also becoming more meaningful over time as the NBS gradually adopts the last elements of the SNA (Xu Xianchun, 2001). Balance sheet data for the directly reporting industrial enterprises only became available in 1993, with the in previous years reported variables on sources and usage of funds not comparable to the later balance sheet items. As the accounting system was revised in 1993 and 2000 and regulations were tightened, balance sheet items have become more meaningful over time.

Measurement errors are unavoidable in data collection. Some variables are more prone to measurement errors than others. Thus, statistical staff may make every effort to accurately calculate nationwide agricultural output, but inevitably cannot obtain the “true” value due to the vastness of this task. For other variables, such as the number of enterprises, adding up the number of enterprises registered at the Industrial and Commercial Administration is likely to yield a highly reliable result, as enterprises (by definition with more than seven employees) are unlikely to be able to escape the registration requirement. In many instances, the measurement problem is simply one of random errors. Given the resources available to the statistical authority and given the types of variables to be measured, the official data come with a particular margin of error. When measurements are consistently skewed in one direction, the measurement problem is—in addition to the problem of random errors—one of bias. Thus the imputed rental value of rural housing in the GDP calculation appears to be consistently underestimated due to the assumption of a rather low depreciation rate of 2%.30

Most productive activities on which the NBS reports data cannot be compiled through report forms collected from all statistical units. The NBS increasingly resorts to sample surveys, such as for the small enterprises with annual sales revenue below 5m RMB and for much of the catering and retail sectors. The quality of the sampling procedure directly affects the quality of the data. While (limited) random sampling is being used in the rural and urban household surveys, other surveys are likely to be conducted in much less professional manner.

29 For details on the innovations introduced by the accounting and tax reforms in the industrial sector see Carsten Holz and Yi-min Lin (2001a). For details on the revaluation of net fixed assets in industry and on changes in the definition of profit, with the impact on actual profit in industry, see Carsten Holz (2003b, Chapter 3).
30 For example, the imputed expenditures on owner-occupied rural dwellings equal the value per single room at the end of the accounting year times the average number of rooms owned by the sampled rural household times the number of households in the rural population times 2% (depreciation rate); the value of the room is based on construction costs. The data are collected in the rural household survey. (OECD, 2000, p. 52f.) The depreciation rate for urban areas is 4%. Xu Xianchun (2000a) notes five consistent biases in GDP calculation, which, however, partially cancel out.
Some of the NBS data manipulations appear no longer appropriate. Thus, in the pre-reform period it may have been appropriate to use official base-year prices of individual products over a period of ten years to obtain a measure of output at base-year prices and to use this measure together with the nominal gross output value to obtain a measure of inflation, which, once applied to nominal value-added, yields real value-added growth. But in a market-oriented economy with more rapid product innovation and no complete coverage of all statistical units the resulting deflator is likely to be of poor quality. The problems with the current practice range from substitution bias to the reporting of output value at current in the place of at constant prices, especially for new products for which no base-year prices exist. On the other hand, rapid quality improvements in one and the same product are likely to go unnoticed. The NBS still relies on base-year prices of individual products to deflate agricultural and industrial output, which together account for more than half of GDP.

The third issue in the typology of data problems is data falsification. In the case of some variables, reporting units have incentives to systematically misreport data. For example, small enterprises are likely to underreport their output when their sales revenue approach 5m RMB annually in order to avoid having to fill in cumbersome monthly report forms for the statistical authority; profit is likely to be underreported when managers wish to avoid high corporate income tax payments. On the other hand, output data from the countryside are likely to be exaggerated.

While evidence of data falsification by lower-level tier statistical departments is scarce, the quality of data collected not by the statistical authority itself, but by various other government departments, is often questionable. Thus financial sector data are collected by the People’s Bank of China, and much of the data on the countryside by the Agricultural Ministry. These departments may have incentives to present their constituencies in the most favorable light, which may bias their data if not induce outright falsification within the government department or at the reporting or estimation level. Since the industrial census 1995, the output data on township and village enterprises reported by the Township and Village Enterprise Bureau of the Agricultural Ministry are notorious for their exaggeration.

Fourth, to the degree to which the NBS relies on provincial-level data to obtain nationwide aggregate data, further inconsistencies are introduced. Lower-level tier data may not only be falsified and of low measurement quality, but different provinces furthermore tend to differ in the speed at which they adopt new concepts (including new variable definitions), as well as in the extent of errors and data falsification; data thus may not even be consistent across provinces at any one point of time, with the nationwide aggregate figure representing a hodgepodge of different concepts of a variable in different provinces.

For example, in 1999 the sum of cement output reported in the 31 provincial statistical yearbooks fell short of the nationwide aggregate cement output reported in the Statistical Yearbook 2000 by 4.15%; the sum of provincial steel output matched the nationwide total. Yet the product tables reporting cement and steel output in 11 of the 31 provincial statistical yearbooks came with a note saying that the data covered only the directly reporting industrial enterprises, while in one province the data covered total output across all enterprises, and in one further province some of the product data covered only the directly reporting industrial enterprises while other product data covered total output across all enterprises; the remaining

Lawrence Klein and Suleyman Ozmucur (2003) suggest that annual real GDP growth in China may have to be adjusted more than half a percentage point upwards due to quality improvements.
18 provincial statistical yearbooks offered no explanations as to the coverage.\textsuperscript{32} The close match between the sum of provincial and nationwide data suggest that the NBS may just be taking the data it receives from the provinces and adding it up to obtain the nationwide total, disregarding any definitional differences between provinces of the items being added up.

The last type of data problems is mistakes in the presentation of the data. Very often this appears to simply be negligence. For example, following the industrial census of 1995, the gross output value of industry was retrospectively revised downward for the years 1991 through 1994; later volumes of the \textit{Statistical Yearbook} at times report the original data, and at other times the corrected data, without making this explicit.\textsuperscript{33}

Since the mid-1990s the NBS has stepped up its efforts to improve data quality. This includes an increasing openness about what particular variables mean and their coverage of statistical units. For example, the NBS National Income Accounts Division in 1997 issued a series of six books on the implementation of the SNA in China,\textsuperscript{34} the NBS Industry and Transport Division in 1999 a book on industrial statistics, the NBS Statistical Design and Administration Division also in 1999 a booklet on ownership definitions in the statistics, and Xu Xianchun in 2000 a book on the calculation of GDP in China. Double-checks on the explanations provided in these books using the actual data do not necessarily always confirm the correctness of the claims made in the books, but at least the NBS is making an effort to explain what it is doing. The NBS maintains its own website at http://www.stats.gov.cn, although a few clicks reveal the rather shallow depth; repeated visits show that non-numerical information are not being accumulated but turned over. The website also includes a list of publications that purport to explain Chinese official statistics and statistical work.

Even with carefully defined variables, it remains difficult to evaluate the effect of errors, data falsification, inconsistencies across provinces, and mistakes on the quality of particular data points, let alone to derive a summary statement on the quality of all Chinese statistics. We have no means to know the true state of China’s economy, nor do we have alternative nationwide (or provincial) data that could be contrasted with the official data.

One solution is to conduct one’s own surveys. Yet when these surveys collect data on standardized variables (such as value-added, fixed assets, or profit), the same definitional problems come in, unless the researcher completely reconstructs these variables and finds interviewees who are willing to spend the amount of time necessary to establish a consistent time series. When the surveys are not random sample surveys, as very few non-NBS surveys are, errors are likely to be as large as when data are collected by the NBS, or perhaps even larger. Data falsification by the statistical authority or government departments can in the case of private surveys be safely ruled out, but not data falsification by the surveyed unit. Why should an enterprise tell the truth to a researcher rather than the NBS?

A second solution is internal consistency checks of Chinese official statistics. Yet, internal consistency checks of Chinese official statistics conducted by researchers often fail because of a poor understanding of the meaning and coverage of the data that are being checked, or of the data that are used to double-check. The NBS itself, as well as provincial statistical bureaus, conduct regular, clearly prescribed double-checks for major economic

\textsuperscript{32} For details, see Carsten Holz (2003).
\textsuperscript{33} For details, see Carsten Holz and Yi-min Lin (2001a, p. 49 and note 51).
\textsuperscript{34} NBS (1997), included in the references here, is one of the six books.
variables. An NBS regulation of 22 Jan. 1999, for example, lists in great detail the double-checks that are currently applied.\textsuperscript{35}

A third solution is to try to get a feeling for the margin of error inherent in the data one wishes to use and to find out their limitations by trying to understand what the data exactly mean and how they are collected. The numerous references to the literature given in the introduction reflect the attempts by an increasing number of researchers to come to terms with different areas of official statistics.

While it seems close to impossible for outsiders to ascertain the quality of the data presented by the NBS, one indirect approach that is available is to examine the procedures the NBS has adopted in the compilation of its statistical data and to evaluate the implementation of these procedures in practice. By understanding the process of data compilation and its changes over time we may be able to note improvements and to reach judgments as to the likely changes in the margin of error over time.

**Changes in data collection methods**

The pre-reform statistical system of the PRC was almost solely based on regular reporting by statistical units through their superordinate departments to the statistical authority. The since the beginning of the economic reforms proliferation of new, often small enterprises means that an increasing share of the economy is not captured through the regular reporting system. The revised Statistics Law of 1996 (NPC, 15 May 1996) marks a watershed in data collection: “Statistical investigation should collect and compile statistical material through regular censuses as the basis [\textit{jichu}], routine sample surveys as mainstay [\textit{zhuti}], and unavoidable [\textit{biyao de}] statistical reporting, key [\textit{zhongdian}] investigations, and comprehensive analysis as supplement [\textit{buchong}]” (Art. 10). No such statement was included in the original 1983 PRC Statistics Law (NPC, 8 Dec. 1983).

This new approach drastically reduces the role of the traditional reporting system while stressing the importance of censuses and sample surveys. In as far as the traditional reporting system was unable to capture an increasing swathe of the economy, and may in some localities have become dominated by local government interests, moving away from the sole reliance on regular reporting by basic statistical units represents an improvement. In the following, the changing role of regular reporting forms, censuses, and then random sample surveys are examined with a view to evaluating data quality.

**Regular reporting system**

The regular reporting system is most developed in the industrial sector.\textsuperscript{36} The traditional reporting system in industry covered all industrial enterprises with independent accounting

\textsuperscript{35} The NBS on 17 June 1992 issued a first regulation on the type of double-checks to be conducted for a series of important variables. NBS Industry and Transport Division (1999), pp. 13f., for the case of industrial value-added lists a number of control variables ranging from investment in fixed assets to energy consumption and transportation volume.

\textsuperscript{36} Similarly, in construction ”grade four and above” enterprises report regularly and directly, in commerce and catering the units “above designated size,” and in transportation and communication those units which are part of a \textit{xitong}. 
system at township level and above; the administrative criterion (of being located or registered at township level and above) implies that all SOEs were included. It was only in 1998 that the NBS reclassified the group of directly reporting industrial enterprises as all industrial SOEs plus all non-SOEs with independent accounting system and annual sales revenue in excess of 5m RMB. The NBS thereby reduced by two thirds the number of reporting enterprises while keeping the aggregate output volume covered approximately constant. This change in coverage had four consequences, all of which constitute unambiguous improvements in data quality.

First, the enterprises that dropped out of the category of directly reporting industrial enterprises are the small enterprises, mostly township-run enterprises, which are noted for their poor data quality. Many of them may never have had a proper accounting system. Second, the statistical break offered a chance to in one stroke eliminate all past exaggeration that may have had to be upheld to guarantee time consistency. Third, the directly reporting industrial enterprises are now supposed to directly report to the statistical bureau, reducing the opportunities for data falsification; previously, in many localities the state-owned industrial enterprises reported to the statistical departments through their superordinate government departments or the economic and trade commission, and the collective-owned industrial enterprises through the second light industry department or the township enterprise department. Fourth, a by-product of the change in coverage is that data on small enterprises may now be more accurate than before. Enterprises prefer to not have to report to the statistical departments every month, as the directly reporting industrial enterprises do, and thus have an incentive not to exaggerate output if they come close to the size criterion.

In late 1998, the State Council approved a joint reform proposal of the NBS and the State Economic and Trade Commission to streamline the industrial reporting system (NBS, 13 Nov. 1998). 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. In addition, 5000 selected key industrial enterprises since April 2001 directly report their economic data to the NBS via the internet. The in 1999 tentatively chosen 5000 enterprises accounted for 49% of industrial assets, 43% of industrial sales revenue, and 83% of industrial taxes and profit. The first reform reduces the potential for data manipulation at

37 In 1997, the directly reporting industrial enterprises, following the old definition, comprised 468,506 enterprises; 319,438 of these were collective-owned, with presumably a majority of often relatively small township-run enterprises (Statistical Yearbook 1998, p. 444). In 1998, the directly reporting industrial enterprises, following the new definition, numbered 165,080; only 47,745 of these were collective-owned (Statistical Yearbook 1999, p. 432). In one county-level city in Jiangxi, the number of directly reporting enterprises fell from 505 to 67, including only 6 township-run enterprises. But not only did the enterprises with the probably poorest data drop out, the statistical departments now furthermore have the time to ensure that the township-run enterprises still included no longer simply report target output figures as results, and to check on the reports of other enterprises. (See Zhongguo tongji no. 2/2000, p. 28.) On the issues in this paragraph see, also, Zhongguo tongji no. 6/1998, p. 23; no. 2/2000, pp. 31f; no. 6/2000, pp. 30, 35 and 38; no. 7/2000, pp. 46f; no. 2/2001, p. 32.

38 Guangxi Province, for example, in its 1999 provincial statistical yearbook retrospectively (and without any explanation) revised downward its GDP figures for 1993-97 by approximately 10%. Changes in industrial value-added accounted for the bulk of the corrections.

39 On the enterprises that do not meet the size criterion, limited data are regularly collected by the township enterprise bureau, the industrial and commercial administration, and the tax office.

40 For the starting date see a 11 April 2001 news item in China Infobank. The NBS’ website (http://www.stats.gov.cn) contains a link for these 5000 industrial enterprises to enter their data.
lower-level tier statistical departments, while the second reform eliminates the possibility of data manipulation during the transfer of data from the enterprise to the NBS altogether.

The direct reporting system thus is far from being abandoned. The recent reclassification of the directly reporting industrial enterprises and the selection of 5000 industrial enterprises which submit their data directly to the NBS suggest that the NBS is limiting direct reporting to those instances where it expects direct reporting to yield accurate data, with some of them available in a very short time span. Direct reporting, where it still occurs, thus is becoming more meaningful. A second pillar, currently still under construction, is the sample surveys designed to capture the very large group of industrial enterprises not reporting regularly to the statistical authority. Industrial statistics thus moved from an all-comprehensive reporting system with guesstimates for an increasing share of industry to a two-class data compilation system with regular direct reporting for one group of industrial enterprises and sample surveys to cover all other industrial enterprises.

**Censuses**

The new “basis” of data collection since 1996 is censuses. China currently conducts five censuses, of which four every ten years, and one every five years:41

- Population census (in years with last digit 0)
- Tertiary sector census (in years with last digit 3)
- Industrial census (in years with last digit 5)
- Agricultural census (in years with last digit 7)
- Census of basic statistical units (in years with last digit 1 or 6)

In the reform period, censuses have come to serve a powerful role as double-check on the comprehensiveness and accuracy of China’s statistical work. The tertiary sector census of 1993 led to large upward revisions to tertiary sector value-added and thus GDP, including retrospective revisions to data for 1978 through 1993. The industrial census of 1995 with the ensuing up to 20% retrospective downward corrections to the gross output value of non-state industrial enterprises of the years 1991 through 1994 (an increase in total 1994 industrial gross output value of 10%), on the other hand, rang warning bells on output exaggeration; GDP statistics, however, were never retrospectively revised. This could be an omission, which would mean that GDP in 1994 was overestimated by approximately 5%. It could also imply that data on industrial value-added are not derived from gross output value, but are approximations based on multipliers applied to the accurate data available for SOEs (where the industrial census did not lead to data revisions) and perhaps for the other directly reporting industrial enterprises (on which data are published only once, for the current year, and the extent of revisions thus can not be ascertained).

With four of the censuses only conducted every ten years, and regular monthly or annual reporting covering an ever decreasing share of the economy only, sample surveys are the only means to bridge the gaps between censuses for that part of the economy that is not regularly reporting to the statistical authority.

---

41 NPC, 15 May 1996, Art. 10. Also see *Zhongguo tongji* no. 12/1999, pp. 27f., *Fujian shengzhi tongjizhi*, and several issues of the Statistical Work Yearbook. The NBS on 31 December 1998 established a census center in Beijing, in charge of improving census design and implementation, and to link censuses to ongoing sample survey work (*Zhongguo tongji* no. 2/1999, p. 43).
Sample surveys

The recent years have seen a general switch to surveys on all activities that are not reliably covered through the traditional reporting system. Already in 1984 the NBS established urban and rural survey teams as administrative facilities directly under its control. Today it has rural survey teams in 857 of China’s 2109 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. Survey teams number about one dozen staff each, which makes them equal in size to the county statistical department. The rural and urban survey teams conduct, among others, the annual urban and rural household surveys. The enterprise survey teams are in charge of sample surveys of enterprises on village-level and below, and check on the validity of important enterprise indicators in general. But they in 1997 also conducted surveys on enterprise reform, and on losses in large and medium-sized enterprises.42

These survey teams allow the NBS to by-pass the lower-level tier statistical departments either to check on the quality of data or to obtain selective data quickly. The head of the NBS in 1997 claimed that 60% of all statistical information going to the general offices of the Chinese Communist Party Central Committee and the State Council was compiled by survey teams. Presumably data collected by the survey teams play a major role in adjusting provincial GDP figures. The fact that nationwide GDP data differ significantly from the sum of provincial GDP data only since 1997 may reflect the NBS’ recently acquired ability to question provincial data (rather than the absence of provincial output exaggeration prior to 1997).

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 sector after four years of hard work are finally in place across all 31 provinces.43 He Keng and Zheng Jingping (2001, p. 58) report that in the case of industry the survey system “will” cover 300,000 units in four different categories: (i) non-state industrial enterprises with annual sales revenue below 5m RMB, located or registered at township level and above; (ii) non-state industrial enterprises with annual sales revenue below 5m RMB, located or registered at village level and below; (iii) the rural self-employed in industry; (iv) the urban self-employed in industry. Two surveys are to be conducted each year, one to obtain data for the months January through September, and one for annual data. Key variables are the number of laborers, gross output value, and sales revenue. In Shaanxi Province, Xi’an Municipality in 2002 switched to sample surveys of industrial enterprises with independent

43 See Zhongguo tongji no. 11/2001, pp. 8-10. On the awareness of the increasing importance of sample surveys, see, for example, numerous reports in China Infobank, such as on 27 December 2000 with excerpts of a speech by Zhu Zhixin, head of the NBS, at the national meeting of statistical department heads, or a report on 25 April 2001. Progress reports on sample surveys also appear in almost every issue of Zhongguo tongji.
accounting system and with annual sales revenue below 5m RMB, but the other municipalities in Shaanxi are unlikely to already have made the switch.\textsuperscript{44}

The development of a system of sample surveys is still very much ongoing. The organizational arrangement with three different types of central survey teams, each established over time as a particular need for accurate data became urgent, appears wasteful, and there are signs that they may be merged in the future into a single survey team. With provincial and even lower-level tier statistical departments also establishing their own survey teams, some counties host both a central and a local survey team, while others do not have a single survey team. Attempts have been made to avoid setting up both central and local survey teams in one and the same county, but the division of labor between central and local survey teams needs further clarification.

\section*{Issues in GDP Calculation}

The value of China’s GDP is far from being an exact number. “Benchmark” revisions such as the one following the tertiary sector census of 1993 lead to major retrospective changes in GDP. But even the regular annual process of revisions is able to yield far from negligible results.

The NBS in the \textit{Statistical Yearbook} series usually offers one revision of annual GDP data. Thus each edition of the \textit{Statistical Yearbook} contains GDP data on the most recent year (already reflecting adjustments to the provincial-level data) and revised GDP data for the second-most recent year. GDP data going back more than two years were already revised once in previous editions of the \textit{Statistical Yearbook}; these earlier revised data are reprinted in the current \textit{Statistical Yearbook}.\textsuperscript{45} The first block of data in Table 4 reports the ratio of a particular year’s production/income approach GDP as given in the \textit{Statistical Yearbook 2002} (revised data) divided by the production/income approach GDP as first reported in the corresponding year’s \textit{Statistical Yearbook}; the second block covers expenditure approach GDP. Year 1993 production/income approach GDP was retrospectively revised upward by 10%, as already noted above, largely because of a 33.45% upward revision to tertiary sector value-added following the 1993 tertiary sector census.\textsuperscript{46}

\begin{table}[h]
\centering
\caption{Table 4 about here}
\end{table}

Later revisions to GDP were all minor, with the largest one a downward correction by 1.32% of 1998 expenditure approach GDP. The varying but often large revisions to inventory investment (in the expenditure approach) suggest considerable problems in obtaining accurate

\textsuperscript{44} See \textit{Xi’an nianjian 2002}, p. 216. Calla Wiemer and Tian Xiuhua (2001) provide details on a pilot study of small industrial enterprises (enterprises with annual sales revenue below 5m RMB) and the self-employed that was to be conducted in Jiangsu Province in 2001.

\textsuperscript{45} The first GDP estimates of every year are published in “China’s Statistical Communiqué” in February of the following year, and then in more detail in “A Statistical Survey of China” published in May. “First confirmed” estimates are published in the \textit{Statistical Yearbook} in September. In the second year, “A Statistical Survey of China” carries the “second confirmed” (and last) estimates, which are then also published in the second year’s and the following years’ \textit{Statistical Yearbook}. (OECD, \textit{National Accounts for China}, 11)

\textsuperscript{46} Tertiary sector value-added was first revised in the \textit{Statistical Yearbook 1995}, and then a second time (which is unusual), by a very minor amount, in the \textit{Statistical Yearbook 1996}. (The value in the latter has been reprinted in later editions.)
data on inventory investment. But since inventory investment accounts for only a small share of GDP, the impact of these revisions on aggregate GDP is small.

Although the later revisions to GDP data tend to be minor, they still matter when the discussion about China’s real growth rate focuses on one percentage higher or lower growth. Thus Thomas Rawski (2001a) uses widely dispersed official data to construct an independent income approach estimate of GDP in China for the years 1997 and 1998 when he suspects the transition from somewhat reliable data to a “tornado of deception.” His income approach calculations yield a 5.7% real GDP growth rate for 1998, in contrast to the official 7.8%, which is based on the production approach for most sectors and the income approach in sectors such as catering. But if one assumes that the original (implicit) GDP deflator is correct, then the later revised nominal GDP value for 1998 implies a real growth rate of 6.4%. (Real growth rates are usually not revised.) This is not much different from 5.7%. If one assumes the revised nominal GDP data imply the need to solely revise the implicit GDP deflator (since the real growth rates published in the Statistical Yearbook are not revised when the nominal data are), then Thomas Rawski’s 5.7% real growth rate needs to be revised to 7.2%, since his 5.7% real growth rate was derived using the implicit original deflator when it should have used the implicit later revised deflator. A revised income approach figure of 7.2% in Thomas Rawski’s calculation is not much different from the official real GDP growth rate of 7.8%.

The deflator underlying real GDP most likely is based on only a subset of the whole economy, largely those statistical units on which relatively complete nominal and real data are available for the first published GDP data. Nominal GDP data are later revised once final data are available on a number of residual statistical units, such as the banking system, insurance businesses, the railways, and civil aviation; these statistical units are likely to only provide nominal data. In as far as the first published (implicit) deflator reflects the best possible estimate of the true deflator for the whole economy, it is applicable not only to the first published GDP data but also to the revised nominal GDP data. Therefore the official real GDP growth rate rather than the deflator is likely to require correction in response to the revised nominal GDP growth rate.

Table 5 presents the revised production/income approach real GDP growth rates for the years since 1990, assuming that the original GDP deflator is the correct one and real GDP growth rates are wrongly not adjusted following the revisions to nominal GDP. For the years 1990 through 1994 the correction method based on later nominal GDP data are inapplicable because in this one instance real growth rates (together with the nominal data) were revised retrospectively following the 1993 tertiary sector census. In the years since 1994, the revised (implicit) real growth rate was significantly lower than the published real growth rate (based on unrevised nominal GDP data) in two years, 1996 and 1998, in each year by slightly more than one percentage point, and otherwise very similar to the official real GDP growth rate.

Table 5 about here

In theory, the three approaches to calculating aggregate GDP should all yield identical results. But in Chinese practice, the production and income are not independent of each other. The expenditure approach throughout all years yields slightly different nominal data; real
growth data are not provided. Albert Keidel (2001b) calculated the real GDP growth rate based on the expenditure approach by making assumptions about the appropriate price deflators. The resulting expenditure approach real GDP growth rate is 1.5 percentage points below the official 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. These non-systematic differences suggest that, if the expenditure approach at the national level were truly independent, China’s GDP data are not systematically biased but come with a certain amount of error.

One possibility to get a better feeling for the margin of error is to break down GDP into its components and to judge the quality of each component based on the available information on how each component is constructed. Table 6 shows the extent to which the production and income approach are combined in deriving production/ income approach GDP. Only in agriculture is the production approach the sole approach used. In industry and construction the production approach is applied to one group of enterprises, namely the directly reporting industrial enterprises, while the income approach, in part in combination with gross output value calculations, is used for those enterprises not reporting directly to the statistical authority. In the tertiary sector the income approach dominates throughout. Across almost all sectors (sub-sectors), ratios of value-added to gross output value or other standardizing variables are obtained from units that directly report to the statistical authority or other government departments, or from surveys, the tertiary sector census, or the input-output table, and then applied to the standardizing variable obtained (estimated) for those units that do not report directly.

Table 6 also provides a subjective judgment on the quality of value-added data compiled in each sector (sub-sector) based on value-added data for the year 2000, the latest year for which detailed tertiary sector value-added data are available. The directly reported data is judged as being of highest quality, while those data not directly reported but obtained through various approximations or through unreliable institutions are categorized either as “somewhat” reliable or unreliable altogether. Of total GDP in China in the year 2000, only a component equal to 45.03% is likely to be highly reliable, while a second part of 11.07% is probably somewhat reliable, and a component almost as large as the highly reliable one, namely 43.90%, is of poor quality.

At the provincial level the expenditure approach is frequently not independent from the production/ income approach. This is presumably the case because provincial-level exports to and imports from other provinces as well as abroad are not easy to measure and may be derived as residual by subtracting all other expenditure approach categories from production/ income approach GDP. See, also, Carsten Holz (2002). At the national level, export and import data only refer to international trade and may be relatively easy to obtain. The inventory adjustment then could be calculated as a residual by subtracting all other expenditure approach data from production/ income approach GDP.

The expenditure approach GDP data are not without serious problems and therefore are highly unlikely to capture the “true” GDP values. In the absence of real expenditure approach data, Albert Keidel applied various price indices to the different expenditure components; none of these price indices is necessarily fully appropriate for the particular expenditure component. Furthermore, both the household survey data and the retail sales data which 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 (due to the lack of data, underreporting for tax reasons, and officially acknowledged omissions).

47 At the provincial level the expenditure approach is frequently not independent from the production/ income approach. This is presumably the case because provincial-level exports to and imports from other provinces as well as abroad are not easy to measure and may be derived as residual by subtracting all other expenditure approach categories from production/ income approach GDP. See, also, Carsten Holz (2002). At the national level, export and import data only refer to international trade and may be relatively easy to obtain. The inventory adjustment then could be calculated as a residual by subtracting all other expenditure approach data from production/ income approach GDP.

48 The expenditure approach GDP data are not without serious problems and therefore are highly unlikely to capture the “true” GDP values. In the absence of real expenditure approach data, Albert Keidel applied various price indices to the different expenditure components; none of these price indices is necessarily fully appropriate for the particular expenditure component. Furthermore, both the household survey data and the retail sales data which 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 (due to the lack of data, underreporting for tax reasons, and officially acknowledged omissions).
At the level of purely subjective evaluation, data in each of the sub-categories of this latter category of unreliable data could well be one third too high or too low, which implies a 15% margin of error in GDP. As long as no clear bias is involved, these errors could cancel out. In fact, as Xu Xianchun (2000a) suggests, the agricultural and industrial portion are likely to be systematically upward biased, while the real estate portion is highly likely to be systematically and gravely downward biased (due to highly imperfect imputations of the service value of owner-occupied housing). Even if there are clear measurement biases, as long as these biases are consistent over time, growth rates could still be quite accurate. For the accuracy of growth rates one big danger are redefinitions of GDP, such as new rules on imputations, which appear to be implemented gradually but not publicly documented.

Conclusions

Over the past 25 years, China’s statistical system, responding to challenges posed by reforms in the real economy, has evolved from a repository of statistical reports within the planning system into a modern statistical agency relying on three channels for data compilations as laid out in the revised Statistics Law of 1996, namely censuses, sample surveys, and limited direct reporting. While the issue of data quality has captured the attention of Western researchers since the late 1990s, official Chinese data, such as GDP, are likely to be of better quality today than ever before. Given the recent wave of research on Chinese data we are also increasingly better positioned to judge the quality of a wide variety of Chinese official statistics.

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.” (p. 222) 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 5000 core industrial enterprises and approximately 160,000 directly reporting industrial enterprises, all represent reforms that are likely to make the available data more reliable.

The development of a reliable survey system appears the last, still incomplete innovation in China’s statistical system. The regular reporting system in its current form was adopted in 1998 and is now fully functional. Most censuses have by now been conducted twice and are turning into routine statistical work. The census of basic statistical units is crucial for surveys in that every five years it establishes a basic sampling frame for surveys, much needed given the rapidly changing nature of China’s economy, with a fair number of enterprise mergers, exist, and new entries; a proper sampling frame is a pre-condition for meaningful survey results. As surveys become regular features of statistical data compilation in the coming years, the quality of Chinese data is likely to increase further. It would not be astonishing to see another round of retrospective revisions to GDP data in about 2005 or 2006, when the results of the 2003 tertiary sector census (and possibly of the 2005 industrial census) are in and the survey system has stabilized. The NBS could also use this opportunity to switch to improved imputation methods.
For the user of Chinese statistics, the best approach in evaluating the quality of Chinese data appears to be the application of a fair dose of common sense. Thus, data on the directly reporting industrial enterprises (in any sector of the economy) are better than data on all enterprises. Balance sheets must at least balance, and profit and loss account items need to add up to sales revenue. In terms of individual items, sales revenue of the directly reporting industrial enterprises are hard data straight out of the corresponding enterprise accounts (except for the statistical break in 1994). Circulation taxes similarly are hard data. Profit is a soft data point as accountants are given numerous choices to make in the process of its derivation. Gross output value is also somewhat soft in that it involves imputations for additions to inventories and self-produced investment goods. Value-added if calculated as gross output value minus the value of intermediate inputs is likely to be soft, but at the enterprise level it can also be calculated using the income approach, which then provides a double-check. Net fixed assets are relatively hard data in low-inflation periods in that the purchasing price is given and annual depreciation, once the depreciation period has been chosen, is fixed. (Special write-offs in the early 1990s led to time inconsistencies then.) Employment data are hard but frequently not usefully defined (not matching the enterprise coverage of other variables or covering also those laborers not on their post).

Data on large enterprises are likely to be more reliable than data on small enterprises which may lack a proper accounting system. Raw data are better than indicators that manipulate raw data; indicators are not always defined, and definitions that change not always properly documented. Data on state-owned units and urban collective-owned units are likely to be more accurate than data on foreign-funded units or rural collective-owned units; data on individual-owned and private units are probably the least accurate, as long as sample surveys are not firmly established. Finally, perhaps most important is some reflection as to what data the NBS can plausibly collect; apparently economy-wide data may in fact only cover certain reporting enterprises, with the enterprise coverage furthermore changing over time.

The quality of the data also differ depending on where they are reported, and by whom. At the peak of the quality pyramid are the various censuses when statistical staff made a concerted effort to capture all relevant activities across the country following central instructions on how to proceed. Data collected by the NBS itself and reported in the Statistical Yearbook comes next. If offers at least some rudimentary definitions and gives the main sources of the data in each section. Statistical breaks are frequently noted in the Statistical Yearbook of the year when the break occurs, but then no longer in the following years. The quality of data in peripheral NBS publications, such as Seventeen Years of Reform, deteriorates quickly. Data published by other government departments, such as the banking system or the Agricultural Ministry (for example, in the Township Enterprise Yearbook), all too often comes with no explanations and thus allows redefinitions and adjustments to go undetected.

The NBS clearly has much further scope for improvements. It has made great efforts in recent years to establish and to improve sample surveys, but this process is still incomplete. The NBS still does not provide definitions for all variables on which its reports data, let alone always correct definitions in the cases where it provides definitions. It all too often does not explain which enterprises/units are covered and how the data were collected. If the NBS remains reluctant to explain the meaning of its statistics or offers explanations that upon
double-checking turn out to be vacuous, it need not be astonished to find its credibility questioned domestically and abroad.
References


CCPCC/ SC (Chinese Communist Party Central Committee/ State Council).


China Infobank. (Internet database on, among others, news items, and laws and regulations.)

22 Nov. 1999. “Heilongjiang sheng renmin zhengfu guanyu jiaqiang tongji gongzuo de jueding” (Decision by the Heilongjiang province government on strengthening the statistical work).


11 April 2001. “Zhongguo wu qian jia gongye qiye lijing yu xitong zhengshi kaitong” (China’s 5000 industrial enterprises’ direct electronic reporting system officially began operation).

25 April 2001. “Jiangsu nantong shi tongji ju danan tuopo gei tongji” (Nantong municipality in Jiangsu Province has achieved a bold breakthrough in the statistics on private and individual-owned units).


_____. China’s Industrial State-owned Enterprises between Profitability and Bankruptcy. Singapore: World Scientific, forthcoming in 2003(b)

NBS (National Bureau of Statistics).
22 Jan. 1999. “Guojia tongji ju guanyu hensha fukuafeng quebao tongji shuju zhiliang de tongzhi” (Implementing instructions for the NBS and provincial statistical bureaus to evaluate the quality of main statistical indicators). In NBS (2001), pp. 560-86.


NPC (National People’s Congress and its Standing Committee).


“China by the Numbers: How Reform Has Affected China’s Economic Statistics.”
China Perspectives, no. 33 (Jan./Feb. 2001a): 25-34.


Seventeen Years of Reform. Gaige kaifang shiqi nian de zhongguo diqu jingji (China’s Regional Economy in Seventeen Years of Reform and Opening). Beijing: Zhongguo tongji chubanshe, 1996.

Shaanxi Yearbook. Shaanxi nianjian (Shaanxi Yearbook). Xi’an: Shaanxi nianjianshe, various issues.

Shaanxi Province Government. “Shanxi sheng renmin zhengfu bangongting guanyu di shi guonei shengchan zongzhi shuju zhiliang shixing pinggu de tongzhii” (Circular by the
government office of Shaanxi Province on evaluating the quality of municipal GDP data. 


Xi’an Yearbook. *Xi’an nianjian* (Xi’an Yearbook). Xi’an: Xi’an chubanshe, various issues.


_____., “Zhongguo guonei shengchan zongzhi hesuan zhong cunzai de ruogan wenti yanjiu” (Some problems in the calculation of China’s GDP). *Jingji yanjiu*, no. 2 (Feb. 2000a): 10-16.


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 Cooperative Economy Management], no. 6 (1998): 42f.

Zhongguo tongji (China statistics). Monthly journal published by the NBS.
Table 1. Numbers and Output Shares of Industrial Enterprises

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of enterprises (thousand)</th>
<th>All industrial enterprises</th>
<th>Share in GOVI (%)</th>
<th>Directly reporting industrial enterprises</th>
<th>Number</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>SOEs</td>
<td>COEs</td>
<td>IOEs</td>
<td>Others</td>
<td>SOEs</td>
</tr>
<tr>
<td>1978</td>
<td>348.4</td>
<td>83.7</td>
<td>264.7</td>
<td></td>
<td>77.63</td>
<td>22.37</td>
</tr>
<tr>
<td>1979</td>
<td>355.0</td>
<td>83.8</td>
<td>271.2</td>
<td></td>
<td>78.47</td>
<td>21.53</td>
</tr>
<tr>
<td>1980</td>
<td>377.3</td>
<td>83.4</td>
<td>293.5</td>
<td>0.4</td>
<td>75.97</td>
<td>23.54</td>
</tr>
<tr>
<td>1981</td>
<td>381.5</td>
<td>84.2</td>
<td>296.8</td>
<td>0.5</td>
<td>74.76</td>
<td>24.62</td>
</tr>
<tr>
<td>1982</td>
<td>388.6</td>
<td>86.0</td>
<td>301.9</td>
<td>0.7</td>
<td>74.44</td>
<td>24.82</td>
</tr>
<tr>
<td>1983</td>
<td>392.5</td>
<td>87.1</td>
<td>304.6</td>
<td>0.8</td>
<td>73.35</td>
<td>25.74</td>
</tr>
<tr>
<td>1984</td>
<td>437.2</td>
<td>84.1</td>
<td>352.1</td>
<td>1.0</td>
<td>69.09</td>
<td>29.71</td>
</tr>
<tr>
<td>1985</td>
<td>5185.3</td>
<td>93.7</td>
<td>1742.1</td>
<td>3347.8</td>
<td>1.7</td>
<td>64.86</td>
</tr>
<tr>
<td>1986</td>
<td>7474.1</td>
<td>97.6</td>
<td>1819.3</td>
<td>5553.3</td>
<td>3.9</td>
<td>62.27</td>
</tr>
<tr>
<td>1987</td>
<td>8105.6</td>
<td>99.1</td>
<td>1853.0</td>
<td>6148.1</td>
<td>5.4</td>
<td>59.73</td>
</tr>
<tr>
<td>1988</td>
<td>7980.7</td>
<td>102.3</td>
<td>1747.0</td>
<td>6124.2</td>
<td>7.2</td>
<td>56.80</td>
</tr>
<tr>
<td>1989</td>
<td>7957.8</td>
<td>104.4</td>
<td>1668.5</td>
<td>6176.0</td>
<td>8.8</td>
<td>56.06</td>
</tr>
<tr>
<td>1990</td>
<td>8079.6</td>
<td>104.7</td>
<td>1577.2</td>
<td>6386.7</td>
<td>10.8</td>
<td>54.60</td>
</tr>
<tr>
<td>1991</td>
<td>8079.6</td>
<td>104.7</td>
<td>1577.2</td>
<td>6386.7</td>
<td>10.8</td>
<td>56.17</td>
</tr>
<tr>
<td>1992</td>
<td>8612.1</td>
<td>103.3</td>
<td>1640.6</td>
<td>6854.0</td>
<td>14.2</td>
<td>51.52</td>
</tr>
<tr>
<td>1993</td>
<td>9911.6</td>
<td>104.7</td>
<td>1803.6</td>
<td>7971.2</td>
<td>32.1</td>
<td>46.95</td>
</tr>
<tr>
<td>1994</td>
<td>10017.1</td>
<td>102.2</td>
<td>1863.0</td>
<td>8007.4</td>
<td>44.5</td>
<td>37.34</td>
</tr>
<tr>
<td>1995</td>
<td>7341.5</td>
<td>118.0</td>
<td>1475.0</td>
<td>5688.2</td>
<td>60.3</td>
<td>32.61</td>
</tr>
<tr>
<td>1996</td>
<td>7986.5</td>
<td>113.8</td>
<td>1591.8</td>
<td>6210.7</td>
<td>70.2</td>
<td>28.48</td>
</tr>
<tr>
<td>1997</td>
<td>7922.9</td>
<td>98.6</td>
<td>1772.3</td>
<td>5974.7</td>
<td>77.3</td>
<td>25.52</td>
</tr>
<tr>
<td>1998</td>
<td>7974.6</td>
<td>57.3</td>
<td>1797.8</td>
<td>6033.8</td>
<td>85.7</td>
<td>21.57</td>
</tr>
<tr>
<td>1999</td>
<td>7929.9</td>
<td>52.1</td>
<td>1659.2</td>
<td>6126.8</td>
<td>91.8</td>
<td>20.31</td>
</tr>
</tbody>
</table>

SOEs: State-owned enterprises.
COEs: Collective-owned enterprises.
IOEs: Individual-owned enterprises (the self-employed).
GOVI: Gross output value of industry.

Data on the number of enterprises (but not GOVI) for the years prior to 1985 exclude village-level industry and individual-owned industry. Gross output value data are retrospectively revised data; data on individual-owned industry prior to 1985 thus are likely to be somewhat inaccurate.

The 1998 and 1999 data on state-owned enterprises were calculated as residual due to a statistical break in 1998. Since 1998 the data also exclude those industrial units with dependent accounting system belonging to a statistical unit outside the industrial sector. Industrial units with dependent accounting system belonging to an industrial enterprise with independent accounting system continue to be included as part of that latter enterprise. (For details on the 1998 statistical break see Carsten Holz and Yi-min Lin (2001a and 2001b).)

Gross output value of industry was redefined in 1996 to, among others, exclude the value-added tax; the redefinition applies across all ownership forms.

Data on all industrial enterprises for the years since 1999 are not available.


Directly reporting industrial enterprises: Industry, Communication, and Energy 1949-1999, pp. 74, 94. The data reported in this source are identical to those reported in the Statistical Yearbook series. For the years prior to 1986 the Statistical Yearbook series does not provide data, and those in Industry, Communication, and Energy 1949-1999 appear to cover a group of enterprises slightly different from the directly reporting industrial enterprises as defined prior to 1998.
In terms of value-added, with data available for the directly reporting industrial enterprises since 1993, the directly reporting industrial enterprises in 1993 through 1999 accounted for the following percentage shares of total industrial value-added: 90.80, 75.93, 62.49, 62.61, 61.20, 58.17, 61.46; the decline between 1993 and 1995 appears unrealistically steep, perhaps reflecting the gradual improvement in the estimation of value-added of those industrial enterprises that are not reporting directly to the statistical authority.


Figure 1. Share of Directly Reporting Industrial Enterprise in Gross Output Value of Industry, in %
Table 2. Upward Revisions in GDP Data Following Tertiary Sector Census, in %

<table>
<thead>
<tr>
<th>Year</th>
<th>GDP</th>
<th>Tertiary sector</th>
<th>Transport, post and telecommunications</th>
<th>Wholesale and retail trade, catering</th>
<th>Other services</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>1.00</td>
<td>4.37</td>
<td>0.00</td>
<td>0.00</td>
<td>9.32</td>
</tr>
<tr>
<td>1979</td>
<td>1.00</td>
<td>4.86</td>
<td>0.00</td>
<td>0.00</td>
<td>9.52</td>
</tr>
<tr>
<td>1980</td>
<td>1.07</td>
<td>5.20</td>
<td>0.00</td>
<td>0.00</td>
<td>9.56</td>
</tr>
<tr>
<td>1981</td>
<td>1.83</td>
<td>8.96</td>
<td>0.00</td>
<td>12.20</td>
<td>11.12</td>
</tr>
<tr>
<td>1982</td>
<td>2.17</td>
<td>10.83</td>
<td>0.00</td>
<td>24.44</td>
<td>11.44</td>
</tr>
<tr>
<td>1983</td>
<td>2.55</td>
<td>12.50</td>
<td>0.00</td>
<td>35.32</td>
<td>11.71</td>
</tr>
<tr>
<td>1984</td>
<td>3.50</td>
<td>15.90</td>
<td>0.00</td>
<td>44.65</td>
<td>12.63</td>
</tr>
<tr>
<td>1985</td>
<td>5.12</td>
<td>20.62</td>
<td>0.00</td>
<td>52.24</td>
<td>11.94</td>
</tr>
<tr>
<td>1986</td>
<td>5.31</td>
<td>21.17</td>
<td>0.00</td>
<td>58.10</td>
<td>12.36</td>
</tr>
<tr>
<td>1987</td>
<td>5.80</td>
<td>22.99</td>
<td>0.00</td>
<td>62.32</td>
<td>13.21</td>
</tr>
<tr>
<td>1988</td>
<td>6.07</td>
<td>23.36</td>
<td>0.00</td>
<td>65.10</td>
<td>10.72</td>
</tr>
<tr>
<td>1989</td>
<td>5.70</td>
<td>20.30</td>
<td>0.00</td>
<td>66.70</td>
<td>8.78</td>
</tr>
<tr>
<td>1990</td>
<td>4.80</td>
<td>17.17</td>
<td>2.68</td>
<td>67.56</td>
<td>8.49</td>
</tr>
<tr>
<td>1991</td>
<td>7.08</td>
<td>24.66</td>
<td>10.39</td>
<td>67.56</td>
<td>13.90</td>
</tr>
<tr>
<td>1992</td>
<td>9.33</td>
<td>33.11</td>
<td>9.50</td>
<td>88.71</td>
<td>21.68</td>
</tr>
<tr>
<td>1993</td>
<td>9.99</td>
<td>32.04</td>
<td>11.69</td>
<td>73.40</td>
<td>24.75</td>
</tr>
</tbody>
</table>

% share in revised 1993 GDP: 31.79, 5.97, 9.00, 16.83

Xu Xianchun (2000b, pp. 96f.) provides a similar, but less complete table. He labels the category “other services,” here obtained as a residual, “non-material services.”

A double-check for the primary and secondary sectors revealed that no revisions had taken place following the tertiary sector census, except a very small revision for 1993, the most recent year, which presumably reflects a standard revision in the second year after the first publication of the 1993 data.

Sources:

Pre-revision data: Statistical Yearbook 93, pp. 31f., and 1994, p. 32. (The latter edition is the main source. It is lacking data for 1979, 1981, and 1982, which are taken from the earlier source; the data for 1978, 1980, 1983, and 1984 were checked as to whether they are identical in both sources, and they are.)

Post-revision data: Statistical Yearbook 95, p. 32, and 2002, p. 51. (The earlier edition is the main source. It is lacking data for 1979 and 1981-83, which are taken from the later source; the data for 1978, 1980, and 1984 were checked as to whether they are identical in both sources, and they are.)
Table 3. Sum of Provincial Value-added Divided by Nationwide GDP

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Primary sector</th>
<th>Secondary sector</th>
<th># Industry</th>
<th># Construction</th>
<th>Tertiary sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>1.0907</td>
<td>1.0281</td>
<td>1.0036</td>
<td>1.0012</td>
<td>1.0194</td>
<td>1.3068</td>
</tr>
<tr>
<td>1994</td>
<td>1.0084</td>
<td>0.9822</td>
<td>1.0131</td>
<td>1.0237</td>
<td>0.9458</td>
<td>1.0187</td>
</tr>
<tr>
<td>1995</td>
<td>0.9891</td>
<td>0.9961</td>
<td>0.9572</td>
<td>n.a.</td>
<td>n.a.</td>
<td>1.0339</td>
</tr>
<tr>
<td>1996</td>
<td>0.9999</td>
<td>1.0400</td>
<td>0.9497</td>
<td>0.9596</td>
<td>0.8858</td>
<td>1.0771</td>
</tr>
<tr>
<td>1997</td>
<td>1.0292</td>
<td>1.0466</td>
<td>0.9808</td>
<td>0.9939</td>
<td>0.8982</td>
<td>1.0931</td>
</tr>
<tr>
<td>1998</td>
<td>1.0426</td>
<td>1.0184</td>
<td>0.9973</td>
<td>0.9979</td>
<td>0.9940</td>
<td>1.2331</td>
</tr>
<tr>
<td>1999</td>
<td>1.0703</td>
<td>1.0114</td>
<td>1.0084</td>
<td>1.0030</td>
<td>1.0432</td>
<td>1.1944</td>
</tr>
<tr>
<td>2000</td>
<td>1.0873</td>
<td>1.0445</td>
<td>1.0065</td>
<td>1.0648</td>
<td>1.0648</td>
<td>1.2315</td>
</tr>
<tr>
<td>2001</td>
<td>1.1129</td>
<td>1.0637</td>
<td>1.0160</td>
<td>1.0867</td>
<td>1.2826</td>
<td>1.2826</td>
</tr>
</tbody>
</table>

# denotes a sub-category, ## a sub-sub-category. The list of sub-categories is complete, as is the list of sub-sub-categories. 1993 is the first year in which GDP data calculated according to the expenditure approach became available.

Provincial-level expenditure data for 1993 and 1994 were published only in the Statistical Yearbook 1995 and 1996, i.e. one year late; this implies that provincial-level data could be revised data (while nationwide data are those as first published). In 1993, in Beijing, Fujian, Shandong and Hunan the sum of final consumption, gross capital formation and net exports was different from aggregate provincial expenditures by -18.07%, 1.91%, 1.58%, and -2.82%. In Shaanxi, the sum of the two components of gross capital formation exceeded provincial aggregate gross capital formation by 15.15%. For a few other provinces in 1993, various checks of the sum of sub-categories yielded differences below 1%. 1994 data show two very minor discrepancies, plus a 7.73% discrepancy between gross capital formation and the sum of its components in Guangxi. 1997 data have one minor discrepancy.

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Primary sector</th>
<th>Secondary sector</th>
<th># Industry</th>
<th># Construction</th>
<th>Tertiary sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>1.1037</td>
<td>1.0349</td>
<td>1.0113</td>
<td>1.0003</td>
<td>1.0854</td>
<td>1.3345</td>
</tr>
<tr>
<td>1994</td>
<td>1.0390</td>
<td>1.0200</td>
<td>1.0524</td>
<td>1.0545</td>
<td>1.0378</td>
<td>1.0434</td>
</tr>
<tr>
<td>1995</td>
<td>1.0037</td>
<td>1.0000</td>
<td>1.0129</td>
<td>1.0150</td>
<td>1.0000</td>
<td>0.9919</td>
</tr>
<tr>
<td>1996</td>
<td>0.9897</td>
<td>0.9971</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9683</td>
</tr>
<tr>
<td>1997</td>
<td>0.9959</td>
<td>1.0174</td>
<td>1.0123</td>
<td>1.0208</td>
<td>0.9587</td>
<td>0.9582</td>
</tr>
<tr>
<td>1998</td>
<td>0.9868</td>
<td>0.9968</td>
<td>0.9981</td>
<td>0.9987</td>
<td>0.9942</td>
<td>0.9643</td>
</tr>
<tr>
<td>1999</td>
<td>1.0019</td>
<td>1.0010</td>
<td>1.0035</td>
<td>1.0032</td>
<td>1.0051</td>
<td>1.0001</td>
</tr>
<tr>
<td>2000</td>
<td>1.0004</td>
<td>1.0293</td>
<td>0.9879</td>
<td>0.9868</td>
<td>0.9950</td>
<td>1.0059</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Final consumption</th>
<th># Household cons.</th>
<th>## Rural households</th>
<th>## Urban households</th>
<th># Government formation</th>
<th># Fixed capital formation</th>
<th># Change in inventories</th>
<th>Net exports</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>1.0045</td>
<td>1.0053</td>
<td>1.0284</td>
<td>0.9836</td>
<td>1.0052</td>
<td>1.0167</td>
<td>0.9694</td>
<td>1.5462</td>
<td>0.7197</td>
</tr>
<tr>
<td>1995</td>
<td>0.9842</td>
<td>0.9620</td>
<td>0.9695</td>
<td>0.9663</td>
<td>0.9393</td>
<td>1.0123</td>
<td>0.9877</td>
<td>1.1794</td>
<td>1.1045</td>
</tr>
<tr>
<td>1996</td>
<td>0.9976</td>
<td>0.9958</td>
<td>0.9866</td>
<td>1.0070</td>
<td>0.9663</td>
<td>1.0354</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>1997</td>
<td>0.9845</td>
<td>0.9734</td>
<td>0.9650</td>
<td>0.9719</td>
<td>0.9582</td>
<td>1.0086</td>
<td>0.9963</td>
<td>0.9788</td>
<td>1.1526</td>
</tr>
<tr>
<td>1998</td>
<td>0.9894</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9720</td>
<td>0.9805</td>
<td>0.8646</td>
</tr>
<tr>
<td>1999</td>
<td>1.0030</td>
<td>1.0008</td>
<td>0.9984</td>
<td>1.0030</td>
<td>0.9945</td>
<td>1.0099</td>
<td>1.0067</td>
<td>1.0000</td>
<td>1.2011</td>
</tr>
<tr>
<td>2000</td>
<td>1.0027</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0076</td>
<td>1.0000</td>
<td>0.3362</td>
</tr>
</tbody>
</table>

All ratios are obtained as annual nationwide value as published in the *Statistical Yearbook 2002* divided by the annual nationwide value as first published in the then current year *Statistical Yearbook*. # denotes a sub-category, ## a sub-sub-category. The list of sub-categories is complete, as is the list of sub-sub-categories. 1993 is the first year in which GDP data calculated according to the expenditure approach became available.

Provincial-level expenditure data for 1993 and 1994 were published only in the *Statistical Yearbook 1995* and 1996 (one year late); this implies that provincial-level data could be revised data (while nationwide data are those as first published). On some individual expenditure-approach data problems see previous table.

Sources: see previous table.
Table 5. Real GDP Growth, 1990-2001

<table>
<thead>
<tr>
<th>Year</th>
<th>Official real GDP growth rate</th>
<th>Albert Keidel’s real GDP growth rate (expenditure approach)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>as first published</td>
<td>as in <em>Statistical Yearbook 2002</em></td>
</tr>
<tr>
<td></td>
<td>(production/ income approach)</td>
<td>calculated based on revised nominal GDP growth rate</td>
</tr>
<tr>
<td></td>
<td>assuming original implicit</td>
<td>assuming original implicit GDP deflator was correct</td>
</tr>
<tr>
<td></td>
<td>GDP deflator was correct</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>3.9</td>
<td>6.8</td>
</tr>
<tr>
<td>1991</td>
<td>8.0</td>
<td>10.1</td>
</tr>
<tr>
<td>1992</td>
<td>13.2</td>
<td>10.9</td>
</tr>
<tr>
<td>1993</td>
<td>13.4</td>
<td>10.7</td>
</tr>
<tr>
<td>1994</td>
<td>11.8</td>
<td>14.6</td>
</tr>
<tr>
<td>1995</td>
<td>10.5</td>
<td>11.4</td>
</tr>
<tr>
<td>1996</td>
<td>9.6</td>
<td>9.8</td>
</tr>
<tr>
<td>1997</td>
<td>8.8</td>
<td>7.3</td>
</tr>
<tr>
<td>1998</td>
<td>7.8</td>
<td>7.2</td>
</tr>
<tr>
<td>1999</td>
<td>7.1</td>
<td>4.0</td>
</tr>
<tr>
<td>2000</td>
<td>8.0</td>
<td>8.8</td>
</tr>
<tr>
<td>2001</td>
<td>7.3</td>
<td>8.8</td>
</tr>
</tbody>
</table>

Official real GDP growth rates were first published in the *Statistical Yearbook 1993*, for the years up through 1992. Since then each issue of the *Statistical Yearbook* has reported real GDP growth rates for the years up through one year before the publication date (the date in the title).

In 1995 the nominal and real growth rates of the years through 1993 underwent major corrections primarily in response to the findings of the 1993 tertiary sector census. Since then only the 1994 real GDP growth rate was ever retrospectively revised, while nominal GDP (and thus, implicitly, the nominal GDP growth rate) continues to as before be revised once, in the year following its first publication.

### Table 6. Reliability of GDP Data

<table>
<thead>
<tr>
<th>Two groups?a</th>
<th>Approach to calculating value-added (VA)</th>
<th>Need data from “direct” for “indirect” group?</th>
<th>Data sources other than direct reporting (where mentioned)</th>
<th>Share of this sector in GDP, year 2000, in %</th>
<th>Share in % of sector (sub-sector) data that are unreliable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(“Direct” vs. “indirect”)</td>
<td>Two groups?</td>
<td>Production approach</td>
<td>Income approach</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farming, forestry, animal hush., fishery</td>
<td>Yes</td>
<td>For direct</td>
<td>For indirect</td>
<td>Survey</td>
<td>16.35</td>
</tr>
<tr>
<td>Industry</td>
<td>Yes</td>
<td>For direct</td>
<td>For indirect (Yesb)</td>
<td>Yes</td>
<td>43.66</td>
</tr>
<tr>
<td>Construction</td>
<td>Yes</td>
<td>Theoretically b</td>
<td>For direct; indirect</td>
<td>Both approaches</td>
<td>(6.05)</td>
</tr>
<tr>
<td>Tertiary sector</td>
<td>Yes</td>
<td>(Yes)</td>
<td>(Yes)</td>
<td>Survey, tertiary census, I/O table</td>
<td>(8.18)</td>
</tr>
<tr>
<td>Transport (and storage), post and telecommunications</td>
<td>Yes</td>
<td>Theoretically b</td>
<td>For direct; indirect</td>
<td>Production approach</td>
<td>(5.83)</td>
</tr>
<tr>
<td>Wholesale/ retail trade, catering</td>
<td>Yes</td>
<td>(Yes)</td>
<td>(Yes)</td>
<td>Both appr.</td>
<td>(1.86)</td>
</tr>
<tr>
<td>Banking and insurance</td>
<td>Yes</td>
<td>(Yes)</td>
<td>Yes</td>
<td>Both appr.</td>
<td>(2.62)</td>
</tr>
<tr>
<td>Banking</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Survey</td>
<td>(8.87)</td>
</tr>
<tr>
<td>Real estate</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>I/O table</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Government, Party, social organizations</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>I/O table</td>
<td>(0.37)</td>
</tr>
<tr>
<td>All other services</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Tertiary census, I/O table</td>
<td>(3.63)</td>
</tr>
<tr>
<td>Services for farming, etc.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>I/O table</td>
<td>(0.92)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.67)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.70)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.31)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A “Yes” in parenthesis means that this approach is not the main approach.

a. Are the statistical units in this sector (or sub-sector) split into two groups, i.e., units reporting directly to the statistical authority or some other government department vs. units not reporting directly? The directly reporting units carry labels such as “above-norm enterprises” or “designated units,” with the definitions being sector-specific.
b The source describes how the production approach is used in this sector (sub-sector), but describes only the calculation of gross output value. It then mentions that value-added is obtained through the income approach, and gross output value minus income approach value-added yields the value of intermediate inputs. In the case of construction, value-added of those units not reporting directly to the statistical authority are based on gross output value and investment statistics combined with value-added ratios obtained from the group of directly reporting units (where value-added was obtained using the income approach).

c In the calculation of the value-added of the “indirect” group in industry, income approach value-added are obtained through sampling, with the ratio of value-added to gross output value in the sample then applied to the gross output value “estimated” for the whole group of enterprises not reporting directly to the statistical authority or government departments. In the case of transport, post and telecommunications, for some units within the “indirect” group the income approach is supported with ratios obtained from the “direct” group, while for other units gross output values are obtained and then multiplied with the ratio of value-added to gross output value in the “direct” group. In wholesale/retail trade and catering, a ratio of value-added to gross output value is obtained from the last tertiary census or the latest input-output table and then applied to the gross output value of the “indirect” group. For approximately 80% of the banking sector (measured in terms of loans) direct income approach value-added data are available; ratios obtained from this group are then applied to total loans.

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

Sources: Xu Xianchun (2000), NBS (1997a). Statistical Yearbook 2001, pp. 410, 465f., 551, 559; 2002, pp. 51, 55. Year 2000 data are used because this is the most recent year for which details on tertiary sector value-added are available.