Review of Income and Wealth
Series 50, Number 3, September 2004

CHINA'S STATISTICAL SYSTEM IN TRANSITION: CHALLENGES, DATA PROBLEMS, AND INSTITUTIONAL INNOVATIONS

BY CARSTEN A. HOLZ*

Hong Kong University of Science and Technology

While China's official statistics are often regarded as of questionable quality, critics are rarely aware of just how difficult it is to compile accurate statistics in a developing and transition economy. This paper traces the challenges economic reforms pose 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 challenges and data problems are being addressed through institutional innovations in data compilation. Analysis of China's data compilation methods allows broad judgments on data quality. Special attention is given to GDP data as the aggregate measure of productive activities in China.

1. 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 of how China's economy operates, develops, or undergoes structural change is regularly 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. For example, the fixed asset values of industrial enterprises—needed in productivity analysis—suffer from the inclusion of non-productive fixed assets and the lack of adjustment for inflation (Chen Kuan *et al*., 1988).[1] A decade later, researchers routinely explore the meaning of

*Correspondence to: Carsten A. Holz, Social Science Division, Hong Kong University of Science and Technology, Hong Kong.

[1]Some of the later literature on adjustments to fixed asset values includes Gary Jefferson (1992), Wing-Thye Woo *et al*. (1994), and Gary Jefferson *et al*. (2000).

specific official Chinese data. The large and growing body of literature covers every area of Chinese statistics from agricultural labor force data to alternative real growth rates in industry.[2]

While much of the literature on Chinese statistics focuses on the meaning of specific official data, a recent strand of literature concentrates on the falsification of official statistics. Cai Yongshun (2000) reports on data falsification in the countryside. Thomas Rawski (2001a, 2001b, 2002b) and others question the reliability of GDP statistics in the most recent years, especially since 1998.[3] Albert Keidel (2001) documents the non-systematic differences between the official real GDP growth rates based on the production approach and real GDP growth rate calculations based on the expenditure approach.

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, but not for China.[4]

This paper moves beyond the exploration of individual areas of statistics and of data falsification issues toward a broader understanding of Chinese statistics: the challenges to data compilation posed by economic reforms in China since 1978, the resulting data problems, and the response of China's National Bureau of Statistics (NBS) to these challenges and problems.[5] Analysis of the institutional reforms initiated by the NBS, in turn, allows broad conclusions on data quality.

The following section explores the challenges posed by economic reforms to official data compilation in China. Economic reforms have 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 estab-

[2]Thomas Rawski and Robert Mead (1998) note that the agricultural labor force data appear to have been derived as a residual by subtracting employment in all other sectors from the total work force, resulting in a likely over-reporting of farm employment by 100 million workers. Thomas Rawski (2002a), Dorothy Solinger (2001, 2002), and Liu Ta and Chan Kam Wing (2001) investigate the coverage of employment, unemployment and migration data. Thomas Scharping (2001) explains the meaning and limitations of Chinese official population statistics, Albert Park and Wang Sangui (2001) of poverty statistics, Ralph Huenemann (2001) of transport statistics, Jeffrey Logan (2001), Jonathan Sinton (2001), and Jonathan Sinton and David Fridley (2000, 2002) of energy statistics, and Martin Ravallion and Chen Shaohua (1999) as well as Chris Bramall (2001) of household income survey data. Carsten Holz and Yi-min Lin (2001a, 2001b) examine the problems of industrial statistics. Vaclav Smil (1999) documents the under-estimation of China's agricultural land and James Nickum (1995) takes the measure of irrigation. Harry Wu (1993) and Angus Maddison (1998) provide alternative long-run GDP estimates, Harry Wu (2000) an overview over different GDP estimates and their implications, and Harry Wu (2002) alternative real growth rates for industry. Thomas P. Lyons (2003) in a historical study shows how to interpret China's trade statistics of 1859–1948.

[3]Other critics include Gerard Adams and Chen Yimin (1996), and Meng Lian and Wang Xiaolu (2000). Carsten Holz (2003a), on the other hand, questions the evidence that purports to show data falsification in the nationwide aggregate official statistics.

[4]For the case of Eastern Europe and Russia, see, for example, Kasper Bartholdy (1997), Vincent Koen (1996), or Susan Powers (1992). Xu Xianchun (2000a, 2002) explains some of the imperfections in China's GDP calculations *today*. Carsten Holz (2002) elaborates on the institutional limitations to the compilation of GDP in China.

[5]Until 1997, the National Bureau of Statistics (*guojia tongji ju*) used the English name "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.

lishes and illustrates a typology of these data problems in official Chinese statistics today. We have no means to directly evaluate the severity of these problems. However, some of the procedures employed by the NBS in its data compilation process are public knowledge. Changes in these procedures, especially if they address the challenges and data problems listed in the second and third section, as they indeed do, can then be evaluated as to their likely impact on data quality (fourth section). The fifth section briefly reviews and evaluates some of the arguments why official Gross Domestic Product (GDP) data are of poor quality and evaluates the quality of aggregate GDP based on the data compilation methods examined in the fourth section. The final section concludes.

2. Economic Reforms and the Implications for Statistical Work

The compilation of statistics in China occurs under the aegis of the NBS at the central level and the corresponding statistical departments at provincial, municipal (prefectural), and county level. While the institutional arrangements have changed little since the beginning of economic reforms in 1978, rapid economic growth and structural change have led to the emergence of severe challenges to the compilation of statistics in China.

2.1. *Rapid Growth in Productive Units Outside the Traditional 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 sector or task, organized from central to county government). They simultaneously reported the statistics 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 1 million villages with 191 million 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," also called "individual-owned enterprises" (*getihu*, employing less than eight persons), and then, in 1993, in the form of "private enterprises" (*siying qiye*, employing more than seven persons), led to a 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 12 million 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 percent of which were state-owned and 75.98 percent collective-owned. The total number of industrial enterprises rose to 437,200 in 1984 thanks to an increase in the number of collective-owned enterprises. The inclusion of village-level collective-owned industry and individual-owned industry in the statistics in 1985 then led to a twelve-fold rise in the total number of industrial enterprises to 5,185,300 in 1985. Individual-owned industrial enterprises accounted for just 1.85 percent of gross output value of industry in 1985, but for 64.56 percent of the total number of industrial 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 percent in 1999, the last year for which these data are available.

Only a fraction of these small enterprises report directly and regularly to the statistical authority; those that do, report detailed balance sheet and profit and loss account as well as output and employment data. 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 fell from presumably close to 100 percent in the early 1980s to only 5.65 percent in 1986. This much lower share remained stable until a statistical break in the category of directly reporting industrial enterprises in 1998, when it fell further to 2.07 percent. The directly reporting industrial enterprises also account for a decreasing share of industrial gross output value, falling from above 90 percent in 1979 to less than 60 percent in 1999.[6]

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 missing out on a growing share of the economy, the NBS began to make efforts to include the missing enterprise categories but 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 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.

2.2. *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

[6]For the data reported in the above two paragraphs see numerous issues of the *Statistical Yearbook*.

apparatus which needed data on physical inputs and outputs, machinery and technology levels. In a more market-oriented economy, macroeconomic policy 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 aggregate production approach GDP statistics (in some sectors, until today, supplemented by income approach data). In 1989 it for the first time experimented with expenditure approach GDP statistics, and in 1992 it began to calculate GDP independently, rather than to 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 approach data for earlier years were published retrospectively to cover first 1978–84, and then 1952–77; expenditure approach GDP data are available for the years since 1978.[7]

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 approach calculated as the sum of value-added across the economy (and supplemented by income approach data). The concept of value-added (*zengjiazhi*) replaced the concept of net material product (*jingchanzhi*).[8]

While the switch from the MPS to the SNA in the case of the material production sectors is primarily a task of asking production units to manipulate accounting items in new ways, some tertiary sector activities, on the other hand, were newly captured by the GDP statistics. The difficulties in newly compiling tertiary sector value-added turned out to be formidable. The tertiary sector census of 1992–93 led to major revisions of tertiary sector value-added in the years 1978–93,

[7]See NBS National Income Accounts Division (1997b, 1998) for the official data, Xu Xianchun and Tian Xiaoqing (1997) for explanations on the derivation of the retrospectively derived official data for 1952–1977, Cai Zhizhou (1994) or Xu Xianchun (2000) for further explanations on the adoption of the SNA, and Hsueh Tien-tung and Li Qiang (1999) for alternative data and explanations on their derivation.

[8]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, 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.

with a maximum impact on GDP (upward revision) of 9.99 percent in 1993.[9] Going back to earlier years, a gradual decline in the percentage corrections to almost zero in 1978 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.

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 often 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). The 2003 tertiary sector census (due to SARS postponed until 2004 and incorporated into a newly designed one-time "economic census") is expected to yield yet another major (upward) revision of tertiary sector GDP data.

While tertiary sector statistics may be particularly problematic, problems with the switch to the SNA exist in all areas of Chinese statistics. Xu Xianchun (2000a, 2002), currently head of the NBS National Income Accounts Division, 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 lifestock products. 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.[10]

2.3. *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 to a market-oriented economy. Yet they also have implications for statistical work.

With the introduction of the new accounting system 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 led to the redefinition of a number of variables; one example is sales revenue, which since 1994 is reported net of the newly comprehensive value-added tax. Balance sheet and profit and loss account variables are defined by the Finance Ministry, with utter disregard for time consistency in statistical data. Statistical authorities in other countries may face similar problems, but

[9]For the tertiary sector sub-sector of wholesale and retail trade and catering, these upward revisions were as high as 88.71 percent in 1992, with this sub-sector in 1993 accounting for 9.00 percentage points of the tertiary sector's 31.79 percent share in GDP. (The other two sub-sectors are transport, post and telecommunications with 5.97 percentage points and "other sectors" with 16.83 percentage points.) For the revised data see *Statistical Yearbook 1995*, p. 32; *2000*, p. 51. For the original data see *Statistical Yearbook 1993*, pp. 31f., *1994*, p. 32.

[10]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 percent in order to newly account for the shadow economy (*The Economist*, May 1, 1997).

changes in accounting and tax systems are likely to be 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, although 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 year 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.

2.4. *Data Falsification at Lower-level Tiers*

Over-reporting of output data has been an issue for many years. 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 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 in part rely on estimates. For example, in the case of industrial non-state enterprises with annual sales revenue below 5 million RMB and of all self-employed in industry, the statistical departments attempt to collect data on the number of such enterprises (units), their gross output value (or, lacking that, sales revenue), tax remittances, paid-in capital, and the year-end number of laborers. The data collection method comprises simple report cards, "processing of related data" (*genju xiangguan ziliao jiagong zhengli*), and sampling. The quality of these data are likely to be low, often no more than guesstimates, and easily subject to manipulation.

Data falsification appears to mostly be 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.[11] The provincial data themselves thus are frequently already the result of revisions, the foundations of which are not publicized.

[11]See the NBS publication *Zhongguo tongji* No. 7/1999, p. 12. The statistical bureau's section in the *Shaanxi Yearbook 2001* (p. 310) announced that the provincial statistical bureau did not have to correct the 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). For numerous instances of data falsification by lowest-level governments in the countryside, see Cai Yongshun (2000).

Nevertheless, at the national level, the sum of production approach GDP across all provinces in most years still exceeds the nationwide total; the same is true for expenditure approach GDP. The NBS, thus, adjusts provincial data when deriving the nationwide total. Table 1 shows the extent of adjustments in production 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 2001. In the production 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 percent of industrial value-added. In the expenditure approach, the NBS systematically revises provincial household consumption upward and gross capital formation, particularly inventories, downward.[12]

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 which is much different from that of GDP.[13]

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 25 percent.[14]

In 1997 the issue of data falsification 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, February 16, 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. The NBS immediately followed up with it own circular (NBS, March 19, 1998) in which the CCPCC/State Council document was billed the third big step in the fight

[12]The large downward adjustments to provincial production 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.

[13]See *China Infobank*, February 29, 2000. 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.

[14]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 percent, 25 percent, and 16 percent, with similar or lesser revisions for the years 1991–93 (see Carsten Holz and Yi-min Lin, 2001a). 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.

TABLE 1

SUM OF PROVINCIAL VALUE-ADDED DIVIDED BY NATIONWIDE VALUE-ADDED

Production approach

	Total (GDP)	Primary Sector	Secondary Sector	# Industry	# Construction	Tertiary Sector
1993	1.091	1.028	1.004	1.001	1.019	1.307
1994	1.008	0.982	1.013	1.024	0.946	1.019
1995	0.989	0.996	0.957	n.a.	n.a.	1.034
1996	1.000	1.004	0.950	0.960	0.886	1.077
1997	1.029	1.047	0.981	0.994	0.898	1.093
1998	1.043	1.018	0.997	0.998	0.994	1.123
1999	1.070	1.011	1.008	1.003	1.043	1.194
2000	1.087	1.045	1.007	0.998	1.065	1.232
2001	1.113	1.064	1.016	1.005	1.087	1.283

Expenditure approach

	Total (GDP)	Final Consumption	# Household Consumption	## Rural Households	## Urban Households	# Government	Gross Capital Formation	# Fixed Capital Formation	# Change in Inventories	Net Exports
1993	0.990	0.911	0.924	0.978	0.870	0.868	1.038	0.916	1.803	–0.665
1994	0.979	0.909	0.920	0.989	0.856	0.867	1.084	0.911	2.998	0.811
1995	0.979	0.892	0.887	0.933	0.842	0.909	1.106	0.964	2.063	1.086
1996	0.999	0.924	n.a.	n.a.	n.a.	n.a.	1.121	n.a.	n.a.	0.841
1997	1.009	0.922	0.894	0.924	0.864	1.041	1.160	0.995	2.640	0.849
1998	1.035	0.942	0.912	0.955	0.873	1.059	1.210	1.034	3.450	0.709
1999	1.061	0.942	0.902	0.950	0.861	1.095	1.245	1.061	6.564	1.196
2000	1.088	0.960	0.910	0.948	0.879	1.143	1.278	1.067	–17.365	1.471
2001	1.080	0.985	0.912	n.a.	n.a.	1.240	1.211	1.056	10.020	1.417

Notes: # 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).

Source: *Statistical Yearbook 1994*, pp. 32, 35; *1995*, pp. 32–4, 36–40; *1996*, pp. 42–44, 46–50; *1997*, pp. 42, 44, 46–50; *1998*, pp. 55, 62f., 67–71; *1999*, pp. 55, 63, 67–71; *2000*, pp. 53, 60f., 65–9; *2001*, pp. 49, 57, 60–65; *2002*, pp. 51, 59, 63–7; NBS National Income Accounts Division (1998), p. 95 (for year 1995 production approach provincial-level value-added data), p. 106 (for year 1996 expenditure approach provincial-level value-added data).

against data falsification since a 1962 CCPCC/State Council circular on strengthening statistical work and the passing of the 1996 Statistics Law. The NBS together with other government departments also dispatched 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, October 25, 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 spate 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, giving way to articles on technical and definitional problems of Chinese statistics.

2.5. *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.[15] A distinction between the public and non-public economy constitutes 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."[16]

The category collective-owned enterprises continues to include all stock cooperatives independent of their actual ownership. The category "state-owned enterprises" covers:

- 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;

[15]For example, the 1993 Company Law brought about the establishment of limited liability companies and stock companies. 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.

[16]See NBS Statistical Design and Administration Division (1999), p. 9, quoting Jiang Zemin's sentence "Public ownership . . . ," or *Jinrong shibao* [Financial Times] Sept 22, 1997 for the full speech including the complete passage quoted here.

- 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.[17]

A further, new SOE category of "state-owned and state-controlled enterprises" comprises all of the above sub-categories except the last. It is clearly larger than the pre-1998 SOE category on which publication of data ceased in 1998. The share of the new SOE category in the value-added of the directly reporting industrial enterprises consequently increased from 46.35 percent in 1997 to 57.03 percent in 1998, although some of this increase could be due to a simultaneous redefinition of the group of directly reporting industrial enterprises.[18] While the changes in the 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 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.

2.6. *Shifting Interests of Reporting Units and Changes in Local Government Administration*

In the socialist economy, filing regular reports with the statistical authority was 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 are not made at all.

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 want to 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.

[17]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).

[18]For the data see *Statistical Yearbook 1998*, p. 444; *1999*, p. 432.

3. Typology of Data Problems

The challenges arising from economic reforms, i.e. the rapid growth in productive 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. The resulting data complications can be summarized as follows:

(1) Definitions of variables and coverage of productive 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 as variables are abandoned and new ones adopted
  (1.5) meaning of variable improves over time
(2) Errors
  (2.1) random measurement errors
  (2.2) systematic measurement errors
  (2.3) sampling errors
  (2.4) shortcomings in data manipulation
(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

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.[19]

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 go utterly unnoticed. In statistical compendia presenting (limited)

[19]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 percent between 1997 and 1998. See *Statistical Yearbook 1998*, p. 432, and *1999*, p. 422.

time series data on the directly reporting industrial enterprises, for example, *Seventeen Years of Reform* or the *Industrial Yearbook*, statistical breaks are not noted.

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.[20] 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. 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.

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 themselves are becoming more meaningful over time as the NBS gradually adopts further elements of the SNA (Xu Xianchun, 2001). The 1993 revision of the accounting system introduced balance sheet terminology as used in market economies.

Some variables are more prone to errors, a second type of data problems, than others. Statistical staff may make every effort to accurately calculate nationwide agricultural output, but inevitably cannot obtain the "true" value. For other variables, such as the number of enterprises, adding up the number of enterprises registered at the Industrial and Commercial Administration yields a 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 systematically 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 percent.

Most productive activities on which the NBS reports data cannot be compiled through report forms collected from *all* productive units. The NBS increasingly resorts to sample surveys. 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 a much less professional manner.

Some of the NBS data manipulations appear no longer appropriate. One example is the use of official base-year prices of individual products over a period of ten years to obtain a measure of (real) gross output at base-year prices, with

[20]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.

the ratio of nominal to real gross output value then used as a deflator to deflate value-added in order to obtain real value-added growth. In China's current market-oriented economy with more rapid product innovation and incomplete coverage of productive units, the resulting deflator is likely to be of poor quality.[21] 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 5 million 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 taxes. On the other hand, output data from the countryside are likely to be exaggerated.

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.[22]

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 *Statistical Yearbook* at times report the original data, and at other times the corrected data, without making this explicit.[23]

Since the mid-1990s the NBS has stepped up its efforts to explain what particular variables mean and which productive units are covered. For example, the NBS National Income Accounts Division in 1997 issued a series of six books on the implementation of the SNA in China,[24] 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.

But 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

[21]The problems with this deflator range from substitution bias to the reporting of output value at current instead of constant prices, especially for new products for which no base-year prices exist—apart from the fact that the gross output value deflator is not necessarily an appropriate deflator for value-added. On the other hand, rapid quality improvements are likely to go unnoticed. Lawrence Klein and Suleyman Ozmucur (2003) suggest that annual real GDP growth in China may have to be adjusted upwards by more than half a percentage point due to quality improvements.

[22]See, for example, Carsten Holz (2003a) for the aggregation of provincial output of individual industrial products.

[23]For details, see Carsten Holz and Yi-min Lin (2001a, p. 49 and note 51).

[24]NBS National Income Accounts Division (1997a), included in the references, is one of the six books.

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 reconstructs these variables. When the surveys are not random sample surveys, as few are, errors are likely to be on the same scale as when data are collected by the NBS. 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, who may publish the enterprise-specific data, and lie to 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.[25] (Two internal consistency checks and their limitations are briefly discussed in Section 5.) The NBS and provincial statistical bureaus themselves conduct regular, clearly prescribed double-checks for major economic variables. An NBS regulation of January 22, 1999, for example, lists in great detail the double-checks that are currently applied.[26]

A third, previously unexplored, indirect approach 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 on the general quality of Chinese data. This is the subject of the following two sections.

4. Changes in Data Collection Methods

The revised Statistics Law of 1996 (NPC, May 15, 1996) introduced major institutional innovations in that it drastically reduced the role of the traditional reporting system while stressing the importance of censuses and sample surveys: "Statistical investigation should collect and compile statistical material through regular censuses as the basis [*jichu*], routine sample surveys as mainstay [*zhuti*], and unavoidable [*biyao de*] statistical reporting, key [*zhongdian*] investigations, and comprehensive analysis as supplement [*buchong*]" (Art. 10).[27]

4.1. *Regular Reporting System*

The regular reporting system is most developed in the industrial sector. The traditional direct reporting system in industry covered all industrial enterprises

[25]The evidence on data falsification is discussed and critically reviewed by, for example, Carsten Holz (2003a), and in part by Nicholas Lardy (2003).

[26]The NBS on June 17, 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.

[27]No such statement was included in the original 1983 PRC Statistics Law (NPC, December 8, 1983).

with independent accounting system at township level and above.[28] 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 5 million RMB.[29] 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 which 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, November 13, 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.[30] In addition, 5000 selected key industrial enterprises since April 2001 directly report their economic data to the NBS via the internet.[31] The in 1999 tentatively chosen 5000 enterprises accounted for 49 percent of industrial assets, 43 percent of industrial sales revenue, and 83 percent of industrial taxes and profit. The first reform reduces the potential for data manipulation at 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.

[28]The administrative criterion (of being located or registered at township level and above) implies that all SOEs were included.

[29]The self-employed (individual-owned enterprises) are excluded from the directly reporting enterprises throughout since they are not officially regarded as enterprises. Similar to the case of industry, 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 *xitong*.

[30]As an anonymous referee has pointed out, the same "super aggregation" occurs at least in commerce and catering for wholesale enterprises with annual sales revenue in excess of 20 million RMB, retail enterprises with annual sales revenue in excess of 5 million RMB, and catering enterprises with annual sales revenue in excess of 2 million RMB.

[31]For the starting date see an April 11, 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. A similar process is in effect for 3000 key real estate enterprises and 1000 key wholesale and retail enterprises.

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, while sample surveys capture the 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 a small group of industrial enterprises which have accurate data and account for approximately 60 percent of industrial output, plus sample surveys to cover all other industrial enterprises.

4.2. *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:[32]

- 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 percent retrospective downward corrections to the gross output value of non-state industrial enterprises of the years 1991 through 1994, on the other hand, rang warning bells about output exaggeration. 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, exits, and new entries; a proper sampling frame is a pre-condition for meaningful survey results.

4.3. *Sample Surveys*

The recent years have seen a broad 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,

[32]NPC, May 15, 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 December 31, 1998 established a census center in Beijing to be in charge of improving census design and implementation and to link censuses to ongoing sample survey work (*Zhongguo tongji* no. 2/1999, p. 43). A one-time "economic census" is to be conducted in 2004; it incorporates the 2003 tertiary sector census which was postponed due to SARS.

in 1997 and 1998. The rural and urban survey teams conduct, among others, the annual urban and rural household surveys; the urban survey team also collects price data. The enterprise survey teams are in charge of sample surveys of non-state industrial enterprises with annual sales revenue below 5 million RMB and the self-employed, 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.[33]

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 percent 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's 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 use of sample surveys of small industrial enterprises (those not reporting directly to the statistical departments) for the compilation of GDP statistics, 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.[34] These are recent additions to the already established surveys of, among others, agricultural production, household income and expenditures, urban prices, and population size.

The development of a system of sample surveys is still incomplete. The organizational arrangement with three different types of central survey teams appears wasteful and there are signs that they may be merged in the future into a single survey team. Central survey teams are often complemented by local survey teams (set up by the local statistical authority), apparently without much coordination. The coming years will see further developments in the field of surveys, especially of tertiary sector surveys.

4.4. *Evaluation*

The revisions to the data compilation system address three of the six challenges identified in the second section. The switch to sample surveys provides

[33]On the issue of survey teams see various NBS regulations in the *Statistical Work Yearbook 1993* and *1998*. *Fujian shengzhi tongjizhi* (2000), pp. 15–18, has details on the survey teams in Fujian Province until 1995.

[34]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 December 27, 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 April 25, 2001. Progress reports on sample surveys also appear in almost every issue of *Zhongguo tongji*. At the local level, the switch to sample surveys of industrial enterprises is more gradual. For example, Xi'an Municipality in Shaanxi Province in 2002 switched to sample surveys of industrial enterprises with independent accounting system and with annual sales revenue below 5 million RMB, but the other municipalities in Shaanxi are unlikely to already have made the switch (*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 5 million RMB) and the self-employed that was to be conducted in Jiangsu Province in 2001.

higher-quality data on the productive units that do not report directly to the statistical authority. Sample surveys also obliterate the effects of data falsification at the local level. Finally, they lighten the reporting burdens of small productive units and bypass local government departments which may have had few incentives to collect accurate data.

As regards the three other challenges, the process of adopting novel statistical concepts and variables, the fourth challenge, is close to completion. Redefinition of economic variables by other government departments is an issue the NBS cannot resolve on its own, but it may have become less severe as the accounting and tax system have stabilized. Political considerations remain a relevant challenge to professional data collection.

In terms of specific categories of data problems, the procedural innovations directly address the size of errors, data falsification and aggregation problems. Definitional problems have to some extent been addressed, such as through the re-definition of state-ownership in the industrial sector (in the process of streamlining the reporting system), but the lack of *public* definitions and explanations remains severe. Outright mistakes in data presentation are not addressed by the institutional innovations discussed here.

In as far as the institutional innovations have properly addressed the key challenges to the compilation of reliable data and the key data problems resulting from these challenges, the margin of error in the official data has been reduced. How big the remaining margin of error is depends on the individual variable; the following section provides an estimate of the margin of error of GDP data. One challenge that the NBS cannot address, but which potentially exerts a significant influence on data quality, is political considerations; one positive sign is that the recent change in political leadership has led to a continuation if not intensification of the focus on professionalism. Sloppiness in the presentation of data to the public and the dearth of public documentation are issues related to political considerations: the NBS as a government department in a socialist state serves primarily the Party and the government, not the public.

5. Issues in GDP Calculation

The value of China's GDP is not a perfectly reliable number. How reliable it is can be evaluated by examining the procedures used to collect GDP data and by noting the innovations to these data collection procedures. Numerous internal control procedures are already in place. Publicized revisions to official GDP data include "benchmark" revisions, such as the one following the tertiary sector census of 1993 which lead to major retrospective changes in GDP,[35] and a regular annual process of revisions.

The NBS in the *Statistical Yearbook* series usually offers one revision of annual GDP data. Thus, each edition of the *Statistical Yearbook* contains GDP data for the most recent year (already reflecting adjustments to the provincial-level

[35]The revisions to industrial gross output value following the industrial census of 1995, in contrast, did not lead to a revision of GDP (which it should have).

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 *Statistical Yearbook*; these earlier revised data are reprinted in the current *Statistical Yearbook*.[36] The first block of data in Table 2 reports the ratio of a particular year's production approach GDP as given in the *Statistical Yearbook 2002* (most up-to-date data) divided by the production approach GDP as first reported in the corresponding year's *Statistical Yearbook*; the second block covers expenditure approach GDP.[37] Annual retrospective revisions to GDP were all minor, with the largest one a downward correction by 1.32 percent of 1998 expenditure approach GDP. The varying but often large revisions to inventory investment (in the expenditure approach) suggest considerable problems in obtaining accurate 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 a few percentage points higher or lower growth. Thus, Thomas Rawski (2001a) conducts an internal consistency check by using 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 percent real GDP growth rate for 1998, in contrast to the official 7.8 percent. 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 percent. (Real growth rates are usually not revised.) This is not much different from 5.7 percent. If one assumes the revised nominal GDP data imply the need to solely revise the implicit GDP deflator (since the published real growth rate was not retrospectively revised), then Thomas Rawski's 5.7 percent real growth rate needs to be revised to 7.2 percent, as his 5.7 percent real growth rate was based on the implicit original deflator when it should have used the implicit later revised deflator. A revised income approach figure of 7.2 percent in Thomas Rawski's calculation is not much different from the official real GDP growth rate of 7.8 percent.[38]

[36]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 *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' *Statistical Yearbook* (OECD, *National Accounts for China*, 11).

[37]Year 1993 production approach GDP was retrospectively revised upward by 10 percent, as already noted above, largely because of a 33.45 percent upward revision to tertiary sector value-added following the 1993 tertiary sector census. Tertiary sector value-added was first revised in the *Statistical Yearbook 1995*, and then a second time (which is unusual), by a very minor amount, in the *Statistical Yearbook 1996*. (The value in the latter has been reprinted in later editions.)

[38]The deflator underlying real GDP most likely is based on only a subset of the whole economy, largely those productive 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 productive units, such as the banking system, insurance businesses, the railways, and civil aviation; these productive 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 2

REVISIONS TO ANNUAL NATIONWIDE VALUE-ADDED

Production approach

	Total (GDP)	Primary Sector	Secondary Sector	# Industry	# Construction	Tertiary Sector
1993	1.104	1.035	1.011	1.000	1.085	1.335
1994	1.039	1.002	1.052	1.055	1.039	1.043
1995	1.004	1.000	1.013	1.015	1.000	0.992
1996	0.990	0.997	1.000	1.000	1.000	0.968
1997	0.996	1.017	1.012	1.021	0.959	0.958
1998	0.987	0.997	0.998	0.999	0.994	0.964
1999	1.002	1.001	1.004	1.003	1.005	1.000
2000	1.000	1.029	0.988	0.987	0.995	1.006

Expenditure approach

	Total (GDP)	Final Consumption	# Household Consumption	## Rural Households	## Urban Households	# Government	Gross Capital Formation	# Fixed Capital Formation	# Change in Inventories	Net Exports
1994	1.005	1.005	1.005	1.028	0.984	1.005	1.017	0.969	1.546	0.720
1995	0.984	0.962	0.968	0.970	0.966	0.939	1.012	0.988	1.179	1.105
1996	0.998	0.996	0.987	1.007	0.966	1.035	1.000	1.000	1.000	1.000
1997	0.985	0.973	0.965	0.972	0.958	1.009	0.996	0.979	1.153	1.041
1998	0.989	1.000	1.000	1.000	1.000	1.000	0.972	0.981	0.865	1.000
1999	1.003	1.001	0.998	1.003	0.995	1.010	1.007	1.000	1.201	1.000
2000	1.003	1.000	1.000	1.000	1.000	1.000	1.008	1.000	0.336	1.000

Notes: 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*. For further notes and data sources see previous table.

A second internal consistency check was introduced by Albert Keidel (2001) who contrasts production approach real GDP growth with expenditure approach real GDP growth. The expenditure approach throughout all years yields slightly different nominal data; the NBS does not publish real growth data. Albert Keidel (2001) calculated the expenditure approach real GDP growth rate 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 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 real GDP growth rate data are not systematically biased but come with a certain amount of error.[39]

An alternative way to evaluate the quality of GDP data which avoids internal consistency checks altogether 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, i.e. by focusing on the data collection procedures employed by the NBS. Table 3 shows the extent to which the production and income approach are combined in deriving production approach GDP; income approach data presumably come with a relatively large margin of error. 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 3 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 are judged as being of highest quality, while those data not directly reported but obtained through various approximations (including the income approach) or through unreliable institutions are categorized either as

[39]For the data see various issues of the *Statistical Yearbook* and Albert Keidel (2001). 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 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).

"somewhat" reliable or unreliable. Of total GDP in China in the year 2000, only a component equal to 45.03 percent is likely to be highly reliable, while a second part of 11.07 percent is somewhat reliable, and the remainder, namely 43.90 percent, is of poor quality.

At the level of purely subjective evaluation, data in this latter category of unreliable data could well be one third too high or too low, which implies an approximately 15 percent margin of error in GDP. As long as no clear bias is involved, these errors could cancel out. In fact, as Xu Xianchun (2000) 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 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, with a margin of error of at most around 1 percentage point (the difference between using a 15 percent inflated vs. non-inflated GDP value in both years times the published growth rate of about 8 percent). For the accuracy of growth rates what matters more are redefinitions of GDP, such as new rules on imputations, which appear to be implemented gradually but are not publicly documented.

6. Conclusions

Economic transition in China had a severe impact on the compilation of statistical data. Traditional data collection methods were unable to capture the rapid growth in productive units outside the direct reporting system. New statistical concepts and variables had to be adopted while some variables were redefined by other government departments in the course of economic transition. Incentives for data falsification, political prerogatives, and shifting interests of reporting units did nothing to facilitate the compilation of accurate statistics. Consequently, Chinese official data suffer from a large number of complications ranging from statistical breaks to various errors.

The NBS responded to these challenge through significant changes in its data compilation methods. The revised Statistics Law of 1996 dramatically reduced the role of the traditional reporting system in favor of censuses ("as the basis") and sample surveys ("as mainstay"). 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 development of a reliable survey system remains the last, still incomplete institutional innovation in China's statistical system.

The recent and ongoing innovations in data compilation methods suggest that the margin of error inherent in the nationwide aggregate data, short of outright, centralized data falsification, is shrinking. In the case of GDP, if all erroneous data were to diverge from the correct value with the same sign, the margin of error could still be as large as 15 percent of GDP (somewhat unlikely). But if the bias were consistent over time, GDP growth rates are rather reliable with a margin of error of at most around 1 percentage point.

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

TABLE 3

RELIABILITY OF GDP DATA

	Two Groups?[a] ("Direct" vs. "Indirect")	Approach to Calculating Value-added (VA)		Need Data from "Direct" for "Indirect" Group?	Data Sources Other than Direct Reporting (where Mentioned)	Share of this Sector in GDP, Year 2000, in %	Share in % of Sector (Sub-sector) Data that Are[d]		
		Production Approach	Income Approach				Highly Reliable	Somewhat Reliable	Unreliable
Farming, forestry, animal husb., fishery		Yes				16.35			100.00
Industry	Yes	For direct	For indirect[c]		Survey	43.66	66.49		33.51
Construction	Yes	(Yes[b])	(Yes[b])	Yes		6.58	56.74	43.26	
Tertiary sector						33.41			
Transport (and storage), post and telecommunications	Yes	Theoretically[b]	For direct; indirect[c]	Both approaches		(6.05)	60.00		40.00
Wholesale/ retail trade, catering	Yes	(Yes)	For direct; indirect[c]	Production approach	Survey, tertiary census, I/O table	(8.18)	48.62		51.38
Banking and insurance						(5.83)	80.00	20.00	
Banking	Yes	(Yes)	Yes[c]	Both appr.					
Insurance		(Yes)	Yes						
Real estate		(Yes[b])	Yes		Survey I/O table	(1.86)			100.00
Government, Party, social organizations			Yes			(2.62)		100.00	
All other services			Yes		Tertiary census, I/O table	(8.87)		50.00	50.00
Services for farming						[0.26]			
Geol. prospecting, water conservancy						[0.37]			
Social services						[3.63]			

Health care, sports, social welfare	[0.92]			
Educ., culture, arts, radio, film, TV	[2.67]			
Scientific research, polytechnic serv.	[0.70]			
Others	[0.31]			
Total: share in GDP, year 2000, in %		45.03	11.07	43.90

Notes: A "Yes" in parenthesis means that this approach is not the main approach.

[a]Are the productive 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 is 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 sector census or the latest input-output table and then applied to the gross output value of the "indirect" group. For approximately 80 percent 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).

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

be astonishing to see another round of retrospective revisions to GDP data in about 2004 or 2005, when the results of the 2004 one-time economic census (which incorporates the tertiary sector census originally scheduled for 2003) are in and the survey system has stabilized. This would also be a good opportunity to introduce numerous improvements, from technically superior imputations to a switch to a new set of deflators in GDP calculation. At the same time, detailed definitions of all variables on which the NBS reports data and documentation of the statistical breaks in these variables would go a long way towards improving our understanding of Chinese data.

References

Adams, F. Gerard and Chen Yimin, "Skepticism about Chinese GDP Growth—The Chinese GDP Elasticity of Energy Consumption," *Journal of Economic and Social Measurement*, 22(4), 231–40, 1996.

Bartholdy, Kasper, "Old and New Problems in the Estimation of National Accounts in Transition Economies," *Economics of Transition*, 5(1), 131–46, May 1997.

Bramall, Chris, "The Quality of China's Household Income Surveys," *The China Quarterly*, 167, 689–705, September 2001.

Cai Yongshun, "Between State and Peasant: Local Cadres and Statistical Reporting in Rural China," *The China Quarterly*, 163, 783–805, September 2000.

Cai Zhizhou, "Cong guonei shengchan zongzhi zhanghu dao guomin jingji hesuan" (From GDP to national income accounts), *Jingji yanjiu* [Economic Research], 2, 71–7, February 1994.

CCPCC/SC (Chinese Communist Party Central Committee/State Council), February 16, 1998, "Zhonggong zhongyang bangongting, guowuyuan bangongting guanyu jianjue fandui he zhizhi zai tongji shang nongxu zuojia de tongzhi" (CCPCC/State Council offices' circular on resolutely opposing and curbing the falsification of statistics), in *Zhongguo gongchandang dangnei fagui xuanbian* (Selection of Chinese Communist Party rules and regulations), Falu chubanshe, Beijing, pp. 408–10, 2001.

Chen Kuan, Gary H. Jefferson, Thomas G. Rawski, Wang Hongchang, and Zheng Yuxin, "New Estimates of Fixed Investment and Capital Stock for Chinese State Industry," *The China Quarterly*, 114, 243–66, June 1988.

China Infobank (Internet database on, among others, news items, and laws and regulations: http://www.chinainfobank.com).

———, February 29, 2000, "Zhongguo tongjiju zhang 'bamai' tongji shuzi" (The head of China's statistical office 'gets a grip' on the statistics).

———, December 27, 2000, "2001 nian zhongguo tongji zhidu fangfa gaige kewang shixian xin tupo" (Reform of China's statistical methods can attain a new break-through in 2001).

———, April 11, 2001, "Zhongguo wu qian jia gongye qiye lianwang zhibao xitong zhengshi kaitong" (China's 5000 industrial enterprises' direct electronic reporting system officially began operation).

———, April 25, 2001, "Jiangsu nantong shi tongji ju dadan tupo gesi tongji" (Nantong municipality in Jiangsu Province has achieved a bold breakthrough in the statistics on private and individual-owned units).

Chow, Gregory C, "Chinese Statistics," *The American Statistician*, 40(3), 191–6, 1986.

Fujian shengzhi tongjizhi (Fujian Provincial Almanacs, Statistics Volume), Fangzhi chubanshe, Fuzhou, 2000.

Holz, Carsten A., "Institutional Constraints on the Quality of Statistics in a Developing and Transitional Economy: The Case of China," *China Information*, 16(1), 25–67, 2002.

———, "'Fast, Clear and Accurate:' How Reliable Are Chinese Output and Economic Growth Statistics?," *The China Quarterly*, 173, 122–63, March 2003(a).

———, *China's Industrial State-owned Enterprises between Profitability and Bankruptcy*, World Scientific Publishing Company, 2003(b).

Holz, Carsten A. and Yi-min Lin, "Pitfalls of China's Industrial Statistics: Inconsistencies and Specification Problems," *The China Review*, 1(1), 29–71, Fall 2001(a).

———, "The 1997–1998 Break in Industrial Statistics: Facts and Appraisal," *China Economic Review*, 12(4), 303–16, 2001(b).

Hsueh Tien-tung and Li Qiang, *China's National Income, 1952–1995*, Westview Press, Boulder, CO, 1999.

Huenemann, Ralph W., "Are China's Recent Transport Statistics Plausible?," *China Economic Review*, 12(4), 368–72, 2001.
Jefferson, Gary H., "Growth and Productivity Change in Chinese Industry: Problems of Measurement," in M. Dutta (ed.), *Research in Asian Economic Studies: A Research Annual. Asian Economic Regimes: An Adaptive Innovation Paradigm*, Volume 4 (Part B), JAI Press Inc., Greenwich, CT, 427–42, 1992.
Jefferson, Gary H., Thomas G. Rawski, Wang Li, and Zheng Yuxin, "Ownership, Productivity Change, and Financial Performance in Chinese Industry," *Journal of Comparative Economics*, 28(4), 786–813, December 2000.
Keidel, Albert, "China's GDP Expenditure Accounts," *China Economic Review*, 12(4), 355–67, 2001.
Klein, Lawrence and Suleyman Ozmucur, "The Estimation of China's Economic Growth Rate," Manuscript, March 2003.
Koen, Vincent, "Russian Macroeconomic Data: Existence, Access, Interpretation," *Communist Economies & Economic Transformation*, 8(3), 321–33, September 1996.
Lardy, Nicholas, "Evaluating Economic Indicators in Post-WTO China," *Issues and Studies*, 38/39(4/38–1/39), 249–68, December 2002/March 2003.
Liu Ta and Chan Kam Wing, "National Statistics on Internal Migration in China," *China Information*, 15(2), 75–113, 2001.
Logan, Jeffrey, "Diverging Energy and Economic Growth in China: Where Has All the Coal Gone," *Pacific and Asian Journal of Energy*, 11(1), 1–13, June 2001.
Lyons, Thomas P., *China Maritime Customs and China's Trade Statistics 1859–1948*, Willow Creek Press, Trumansburg, NY, 2003.
Maddison, Angus, *Chinese Economic Performance in the Long Run*, Development Centre of the Organisation for Economic Co-operation and Development, Paris, 1998.
Meng Lian and Wang Xiaolu, "Dui zhongguo jingji zengzhang tongji shuju kexindu de guji" (An evaluation of the reliability of China's statistics on economic growth), *Jingji yanjiu*, 10, 3–13, October 2000.
NBS (National Bureau of Statistics), June 17, 1992, "Guanyu kaizhan zhuyao tongji zhibiao shuju zhiliang pinggu gongzuo the tongzhi" (Circular on launching an appraisal of the quality of all important economic indicators), in *Statistical Work Yearbook 1993*, pp. IX, 17–25.
———, March 19, 1998, "Guojia tongji ju guanyu renzhen guanche luoshi 'Zhonggong zhongyang bangongting, guowuyuan bangongting guanyu jianjue fandui he zhizhi zai tongjishang nongxu zuojia de tongzhi' de tongzhi" (On conscientiouisly implementing the CCPCC/State Council offices' circular on resolutely opposing and curbing the falsification of statistics), in NBS (2001), pp. 425–8.
———, November 13, 1998, "Guanyu zuohao tongji zhidu gaige de yijian" (Suggestions on the reform of the statistical system), issued jointly with the State Economic and Trade Commission and approved by the State Council General Office, in NBS Industry and Transport Division (1999), pp. 275–8.
———, January 22, 1999, "Guojia tongjiju he shengji tongjiju dui zhuyao tongji zhibiao shuju zhiliang pinggu de shishi banfa" (Implementing instructions for the NBS and provincial statistical bureaus to evaluate the quality of main statistical indicators), in NBS (2001), pp. 560–86.
———, October 25, 1999, "Guojia tongjiju guanyu hensha fukuafeng quebao tongji shuju zhiliang de tongzhi" (On stopping the wind of exaggeration dead in its track and guaranteeing the quality of statistical data), in NBS (2001), pp. 704–6.
———, 2001, *Tongji gongzuo zhongyao wenjian xuanbian* (Selection of important documents on statistical work), Zhongguo tongji chubanshe, Beijing.
NBS (National Bureau of Statistics) National Income Accounts Division, *Zhongguo niandu guonei shengchan zongzhi jisuan fangfa* (Calculation method of China's annual gross domestic product), Zhongguo tongji chubanshe, Beijing, 1997(a).
———, *Zhongguo guonei shengchan zongzhi hesuan lishi ziliao 1952–1995* (Historical data on China's gross domestic product 1952–1995), Dongbei caijing daxue chubanshe, Dalian, 1997(b).
———, *Zhongguo guonei shengchan zongzhi hesuan lishi ziliao (zhaiyao) 1952–1996* (Abstract of historical data on China's gross domestic product 1952–1996), Zhongguo tongji chubanshe, Beijing, 1998.
NBS (National Bureau of Statistics) Industry and Transport Division, "Gongye zengzhang sudu zhong de 'shuifen' you duoda?" (How big is the 'water content' in industrial growth?), *Jingji yanjiu cankao ziliao* [Economic Research Reference Material], 42, 2–5, March 19, 1989.
———, *Xinbian gongye tongji gongzuo zhinan* (New guide to industrial statistics), Zhongguo tongji chubanshe, Beijing, 1999.

NBS (National Bureau of Statistics) Statistical Design and Administration Division, *Guanyu tongji shang ruhe fanying suoyouzhi jiegou wenti de yanjiu* (How to reflect the issue of ownership structure in the statistics), Zhongguo tongji chubanshe, Beijing, 1999.
Nickum, James E., "Dam Lies and Other Statistics: Taking the Measure of Irrigation in China, 1931–91," East-West Center Occasional Papers, 1995.
NPC (National People's Congress and its Standing Committee).
———, December 8, 1983, "Zhonghua renmin gongheguo tongji fa" (PRC Statistics Law), in *China Infobank*.
———, May 15, 1996, "Zhonghua renmin gongheguo tongji fa" (PRC Statistics Law), revision of original 8 December 1983 law, in *China Infobank*.
OECD (Organization of Economic Co-operation and Development), *National Accounts for China: Sources and Methods*, OECD, Paris, 2000.
Park, Albert and Wang Sangui, "China's Poverty Statistics," *China Economic Review*, 12(4), 384–98, 2001.
Perkins, Dwight H., *Market Control and Planning in Communist China*, Harvard University Press, Cambridge, MA, 1966.
Powers, Susan, "Statistical Needs in Eastern Europe," *Monthly Labor Review*, 115(3), 18–28, March 1992.
Ravallion, Martin and Chen Shaohua, "When Economic Reform Is Faster than Statistical Reform: Measuring and Explaining Income Inequality in Rural China," *Oxford Bulletin of Economics and Statistics*, 61(1), 33–56, February 1999.
Rawski, Thomas G., "On the Reliability of Chinese Economic Data: Discussion," *Journal of Development Studies*, 12(4), 438–41, 1976.
———, "China by the Numbers: How Reform Has Affected China's Economic Statistics," *China Perspectives*, 33, 25–34, January/February 2001(a).
———, "What Is Happening to China's GDP Statistics?," *China Economic Review*, 12(4), 347–54, 2001(b).
———, "Where Will the Jobs Come From?," *China Economic Quarterly*, 6(4), 39–43, 2002(a).
———, "Measuring China's Recent GDP Growth: Where Do We Stand," Mimeo, August 29, 2002(b), at http://www.pitt.edu/~tgrawski/papers2002/.
Rawski, Thomas G. and Robert W. Mead, "On the Trail of China's Phantom Farmers," *World Development*, 26(5), 767–81, 1998.
Scharping, Thomas, "Hide-and-Seek: China's Elusive Population Data," *China Economic Review*, 12(4), 323–32, 2001.
Seventeen Years of Reform, *Gaige kaifang shiqi nian de zhongguo diqu jingji* (China's Regional Economy in Seventeen Years of Reform and Opening), Zhongguo tongji chubanshe, Beijing, 1996.
Shaanxi Yearbook, *Shaanxi nianjian* (Shaanxi Yearbook), Shaanxi nianjianshe, Xi'an, various issues.
Shaanxi Province Government, "Shanxi sheng renmin zhengfu bangongting guanyu dui di shi guonei shengchan zongzhi shuju zhiliang shixing pinggu de tongzhi" (Circular by the government office of Shaanxi Province on evaluating the quality of municipal GDP data), *Shaanxi zhengbao* [Shaanxi Bulletin], no. 21, 27f., 1999.
Sinton, Jonathan E., "Accuracy and Reliability of China's Energy Statistics," *China Economic Review*, 12(4), 373–83, 2001.
Sinton, Jonathan E. and David G. Fridley, "What Goes Up: Recent Trends in China's Energy Consumption," *Energy Policy*, 28(10), 671–87, 2000.
———, "A Guide to China's Energy Statistics," *The Journal of Energy Literature*, 8(1), 22–35, 2002.
Smil, Vaclav, "China's Agricultural Land," *The China Quarterly*, 158, 414–29, June 1999.
Solinger, Dorothy J., "Why We Cannot Count the 'Unemployed'," *The China Quarterly*, 167, 671–88, September 2001.
———, "The View from Wuhan: China's Uncountable Unemployed," *China Economic Quarterly*, 6(4), 34–9, 2002.
Statistical Work Yearbook, *Zhongguo tongji gongzuo nianjian* (China Statistical Work Yearbook), Zhongguo tongji chubanshe, Beijing, various years.
Statistical Yearbook, *Zhongguo tongji nianjian* (China Statistical Yearbook), Zhongguo tongji chubanshe, Beijing, various issues.
Township Enterprise Yearbook, *Zhongguo xiangzhen qiye nianjian* (China Township Enterprise Yearbook), Nongye chubanshe, Beijing, various years.
Wiemer, Calla and Tian Xiuhua, "The Measurement of Small-scale Industry for China's GDP Accounts," *China Economic Review*, 12(4), 317–22, 2001.
Woo Wing-Thye, Hai Wen, Jin Yibiao, and Fan Gang, "How Successful Has Chinese Enterprise Reform Been? Pitfalls in Opposite Biases and Focus," *Journal of Comparative Economics*, 18(3), 410–37, 1994.

Wu, Harry X., "The Real Chinese Gross Domestic Product (GDP) for the Pre-Reform Period 1952–77," *Review of Income and Wealth*, 39(1), 63–87, March 1993.
———, "China's GDP Level and Growth Performance: Alternative Estimates and the Implications," *Review of Income and Wealth*, 46(4), 475–99, December 2000.
———, "How Fast Has Chinese Industry Grown? Measuring the Real Output of Chinese Industry, 1949–97," *Review of Income and Wealth*, 48(2), 179–204, June 2002.
Xi'an Yearbook, *Xi'an nianjian* (Xi'an Yearbook), Xi'an chubanshe, Xi'an, various issues.
Xu Xianchun, "Zhongguo guonei shengchan zongzhi hesuan zhong cunzai de ruogan wenti yanjiu" (Some problems in the calculation of China's GDP), *Jingji yanjiu*, 2, 10–16, February 2000(a).
———, *Zhongguo guonei shengchan zongzhi hesuan* (Calculation of China's GDP), Beijing daxue chubanshe, Beijing, 2000(b).
———, "Wo guo GDP hesuan yu xianxing SNA de GDP hesuan zhijian de ruogan chayi" (Some discrepancies between the calculation of GDP and the calculation of GDP in the SNA), *Jingji yanjiu*, 11, 63–8, November 2001.
———, "Study on Some Problems in Estimating China's Gross Domestic Product," *Review of Income and Wealth*, 48(2), 205–15, June 2002.
Xu Xianchun and Tian Xiaoqing, "Zhongguo guonei shengchan zongzhi lishi shuju (1952–1977 nian) de ziliao shouji he cesuan fangfa" (Compilation and estimation of data used in establishing China's GDP (1952–1977)), *Jingji yanjiu*, 9, 65–71, September 1997.
Yue Ximing and Zhang Shuguang, "Woguo fuwuye zengjiazhi de hesuan wenti" (Problems in the calculation of the value-added of services in China), *Jingji yanjiu*, 12, 51–9, 2002.
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, 42f., 1998.
Zhongguo tongji (China statistics), monthly journal published by the NBS.