Can We Trust The Numbers?

By Carsten A. Holz

For years analysts have questioned the accuracy of China's economic data, and especially its GDP figures. A careful look at the evidence shows that Chinese statistics get the big picture right. The key problem is not the quality of the data, but the government's lack of transparency.

The world has a love-hate relationship with China's data on its gross domestic product. While markets closely follow every release of the GDP figures, it is also commonly believed that the data cannot be trusted. China's own premier Li Keqiang was quoted as saying that the GDP figures are "man-made and therefore unreliable," and that he instead watched indicators such as electricity consumption, rail cargo volumes and bank loans. And every few years a popular argument emerges that China's GDP data are uniquely distorted and untrustworthy.

Here I make the case that while not every single data series is of the highest quality, nonetheless the official data get the picture of China's economy broadly right. This finding is based on an exhaustive review of the arguments for data falsification, none of which stands up to close scrutiny. And it validates an insight from nearly half a century ago by the eminent Harvard development economist Dwight Perkins, who persuasively argued that falsified data could not persist for long, since it was not in the interests of the Chinese leadership. It is hard to believe that China could have prospered so obviously for so long if policy makers were relying on

Carsten A. Holz is professor of economics in the Social Science Division at Hong Kong University of Science and Technology. Home page: http/ihome.ust.hk/~socholz.

systematically faulty data. As for Premier Li, his statement—made in 2007 when he was party secretary of Liaoning province—reveals more about his own ignorance than about Chinese data. GDP data are man-made as well as unreliable all around the world, and his focus on the physical measures of the planned economy betrays a limited vision.

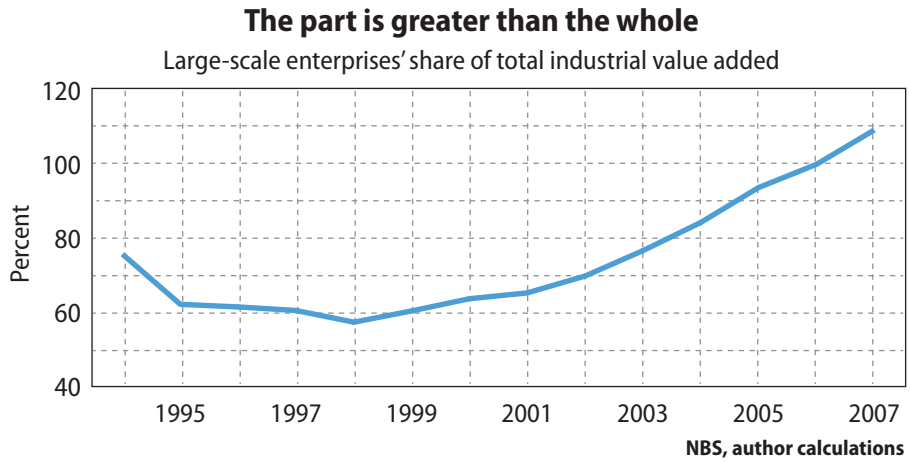
A parade of skeptics

Skepticism about Chinese statistics has a long history. In 1994, the World Bank adjusted its estimate of China's 1992 per-capita GDP up by 34%, based in part on purchasing-power parity estimates of non-tradable goods and services. It continued to adjust Chinese economic data until 1999, when it accepted the official numbers. If one assumes that the World Bank's 1992 number and Beijing's 1998 number were both correct, this would imply that average annual real GDP growth per capita in 1992-98 was two percentage points lower than the Chinese government reported. But Xu Xianchun, now deputy commissioner of the National Bureau of Statistics (NBS) and the head of its national accounts department, argued that some of the World Bank adjustments were not justified, and that other factors worked in the opposite direction.

In a 1998 study for the OECD, economic historian Angus Maddison estimated that Chinese real GDP growth in 1978-95 averaged 7.5% a year, well below the official reading of 9.9%. Maddison arrived at his figures by assuming zero labor productivity growth in "non-productive services," and by rejecting Beijing's figures on value-added growth in industry in favor of growth in product quantities. These adjustments are hard to justify. Anyone who used banking services in China in 1978 and in 1995 will readily attest that labor productivity growth in banking was far above zero. And reliance on product quantities ignores both improvement in product quality (again evident to anyone who bought household appliances in China in the 1970s and 90s) and the introduction of new products.

Anyone who used banking services in China in 1978 and in 1995 will readily attest that labor productivity growth in banking was far above zero

But the meme of China's statistical unreliability really caught on after 1998, thanks to a series of articles in Chinese government statistical publications denouncing a "wind of falsification and embellishment" by local authorities. This critique was picked up and elaborated by Thomas Rawski of the University of Pittsburgh, who pieced together a national income measure and came up with a growth rate for 1998 of 5.7%, well below the



official GDP growth rate of 7.8% (See Thomas G. Rawski, "The Credibility Gap: China's Recent GDP Statistics," *China Economic Quarterly*, March 2001). But Rawski relied on the official deflator of the National Bureau of Statistics (NBS) to arrive at his real income growth rate, and once the NBS revised that deflator a year later, Rawski's real income growth rate rose to 7.2%, much closer to the government's GDP estimate. Subsequent research (using imports, tax revenues, electricity production, and bank loans as proxies) essentially confirmed the official GDP growth rates of the late 1990s. It is also worth noting that the problems in the late-1990s data were first flagged not by skeptical foreign analysts, but by the government's own statisticians. This fact undermines the notion that Beijing's data-minders are engaged in an elaborate conspiracy to hoodwink the world about China's economic performance.

Better roughly right than precisely wrong

Even though none of the past critiques of China's GDP has stood up under sustained analysis, suspicion remains that China's economic numbers cannot be trusted. This suspicion is routinely validated by the facts that the sum of provincial economic output always exceeds national GDP, and most or all provinces tend to report economic growth rates above the national figure. There is no doubt that China's economic data are problematic, and not always internally consistent over time. But the important question is whether on the whole they are roughly right or precisely wrong. The evidence supports a "roughly right" verdict for national-level data.

The discrepancy between provincial and national data owes at least in part to the incentives for local cadres to exaggerate local output growth in order to boost their promotion prospects. NBS officials are aware of this

tendency and therefore largely sidestep the local data. At a more technical level, big firms that operate nationally may have their output claimed by multiple provinces, and these double-counts are netted out when national figures are compiled.

But even the national data have real problems. A good example lies in the measurement of industrial value-added. NBS splits industrial enterprises into two groups: large scale and small scale (规模以上/以下). The criterion is revenue from the principal business, which had to be at least Rmb5m for large-scale firms through 2010, rising to Rmb20m in 2011. By 2007, large-scale enterprises produced more value added than did the industrial sector as a whole. At first glance, this implies either that small-scale industrial enterprises somehow produced negative output, or that official figures understated total industrial value-added and hence GDP. Neither explanation is correct.

The likely truth is more technical and boring: NBS just never updated the formulas it used to translate enterprise gross output value (approximately equal to sales revenue) into value-added plus intermediate inputs. (One can show this using the national input-output table and value-added tax data.) NBS's response to this problem was uninspiring: it stopped publishing value-added data for the large-scale enterprises, and continued publishing only the industry-wide value-added as part of GDP.

There is another reason to be wary of industrial data. Article 25 of the 2009 Statistics Law strictly prohibits the release of data that allows the identification of individual reporting units. Yet the data of all individual large-scale industrial enterprises, covering several years, can be bought on Chinese websites, with the datasets identifying firms by name and address together with their output, balance sheet, and profit and loss account data. As an enterprise manager, would you report wholly truthful data to a statistical authority that you know operates in gross breach of the Statistics Law and spills your internal information to the competitors and the public? The bottom line is that industrial value-added, which accounts for half of GDP, is likely to have become a rather rough approximation of reality.

Another issue, not unique to China, is the "shadow economy": economic activities not recorded in official statistics, whether because they are illegal, or simply beyond the ability of statisticians to capture. Across OECD countries, official GDP is estimated to underreport actual GDP by 8-30% percent, with a low of 8-10% for countries such as the US and Japan, and a high of 24-30% for countries such as Greece, Italy and Spain. Estimates for China range from 10-25%, and many analysts believe the size of the shadow economy may have peaked around 1990. It may be that part of apparent "economic growth" in the last two decades was simply the

statistical authorities doing a progressively better job of capturing a wider range of economic activity.

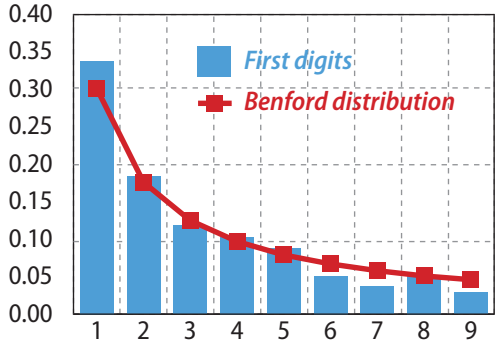
To catch a cheater: Benford's Law

The above observations suggest that China's GDP data could be somewhat unreliable either for technical reasons (eg the shadow economy problem, or the construction of industrial value-added estimates) or because of deliberate falsification by NBS. To assess the likelihood of falsification, we can apply Benford's Law. The Law states that the first digit of numbers under a wide variety of circumstances (including exponential growth series, such as China's nominal GDP) are distributed in a very consistent way: 1 is the first digit 30% of the time, 2 is the first digit 18%, and so on down to 9 which is the first digit less than 5% of the time. (The law similarly applies to the distribution for second digits, useful in analyzing series like growth rates.) If the distribution of initial digits in a data series differs sharply from this pattern, then it is likely that the data have been tampered with. For instance, Greece's data routinely and significantly conform less to Benford's Law than the data of other European countries, and this supports the widespread view that Greece's data are not to be trusted.

I subjected China's nominal GDP numbers to a first-digit Benford test and its deflators and real growth rates to a second-digit test. I then performed the same tests on US data for comparison. The results are clear. For nominal GDP, neither Chinese nor US nominal data differ significantly from the Benford distribution. Thus, even though the NBS compiles nominal GDP data in utter secrecy, the quality of these data is no worse than that of the comparable US data.

China's GDP data is as reliable…

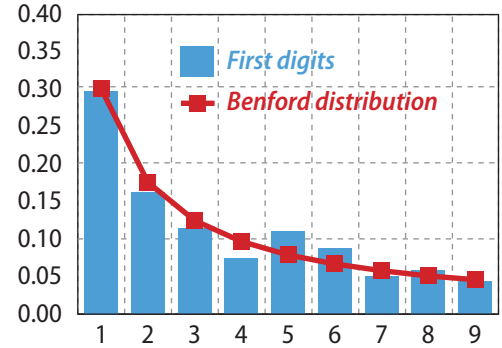
Distribution of first digits in nominal value-added figures for four major sectors*, China (1978-2011)



*Agriculture, industry, construction, services
Author calculations

…as America's

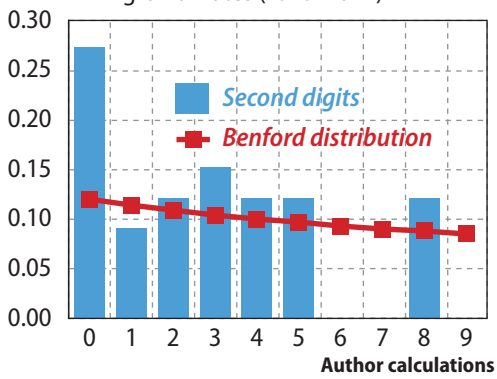
Distribution of first digits in nominal figures for the eight major expenditure components of GDP, US (1978-2011)



Author calculations

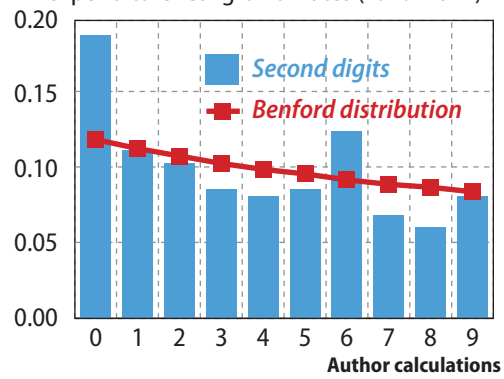
China's growth rates have problems…

Second-digit distribution of China real GDP growth rates (1979-2011)



…just like America's

Second-digit distribution of US aggregate expenditure real growth rates (1979-2011)



We can thus feel comfortable at least that China's overall nominal GDP figures are not manipulated. The second-digit test on the deflators for GDP and individual sectors produces a similar result. China's deflators follow the Benford distribution with no statistically significant deviation, as do US prices. For real growth rates, the statistical significance tests suggest that *both* the Chinese and US statistical authorities engage in some rounding, because both countries report growth rates in which the second digit is zero much more frequently than Benford's Law predicts. But this appears to be the only serious problem.

So Chinese data definitely have problems, but they are not systematically falsified. We have no grounds for thinking that the nominal GDP data are substantially more inaccurate than in any other large economy. The question then becomes how big a margin of error should we apply to the GDP growth rate, which is the number most people pay attention to? Here the issue is the accuracy of the deflators or price indices that NBS uses to convert nominal GDP figures into real growth rates. The Benford analysis suggests the deflators are not systematically manipulated, but there is little doubt that NBS occasionally uses deflators as a "fudge factor" (see box, "Grossly distorted deflators?"). And much external criticism of China's GDP data boils down to claims about problems in price data.

My approach here is to create a range of estimates for average real GDP growth rates in 1978-2011, using different deflators. I came up with four such estimates:

1. The latest available official data published in the China Statistical Yearbook.

Grossly distorted deflators?

I argue that China does not systematically falsify its economic data, but this does not mean one can believe every number that comes out of Beijing. A classic problem in Chinese statistics, intensely annoying to detail-oriented analysts but only marginally relevant for practical purposes, is the use of the implicit GDP deflator as a fudge-factor. (The GDP deflator is the ratio of nominal to real GDP growth, and in principle is the weighted average of all price changes in an economy—in other words, an economy-wide inflation rate.)

NBS publishes nominal values and real growth rates both for GDP as a whole and for the constituent individual economic sectors—agriculture, industry, construction and services. When it revises the nominal values and does not proportionally change the real growth rates, the result is a change in the implicit deflator. Occasionally these changes are suspiciously large, and not justified by any actual new price data.

A good example came in 2008. NBS originally reported a nominal GDP value of Rmb 30.1 trn and a real growth rate of 9.0%. Using the previous year's nominal GDP data, one could calculate an implicit GDP deflator of 7.2%. A year later, nominal GDP for 2008 was revised up to Rmb 31.4 trn, and the real growth rate was revised up to 9.6%. These values yielded an implicit GDP deflator of 7.8%, far higher than the previous figure and without any backing in individual price series (for instance, consumer price index values remained unchanged).

What happened is that NBS collected new nominal data for 2008 in its economic census that year. The new 2008 nominal GDP value implied a 10.2% growth rate over 2007 using the old implicit deflator, rather than the earlier reported 9.0%. For whatever reason, NBS judged that this real growth rate was too high. So it split the difference: it revised the real growth rate up by just 0.6 percentage points, and—utterly implausibly—permitted the implicit deflator to rise by an equal amount to make things work out mathematically.

The year 2008 is special, because an economic census provided a lot of new nominal data. In most years the fudge-factor is much smaller: for instance, revisions to 2010 GDP resulted in a 0.1 percentage point change to the implicit deflator. These sorts of fudges are far from statistical best practice, but they also do not materially change the view of China's long-run economic performance.

2. Using latest available nominal GDP data but applying the original implicit GDP deflators.
3. Using the latest available value-added data for each sector, deflated using one of the *lowest* plausible alternative price indices.

4. Using the latest available value-added data for each sector, deflated using one of the *highest* plausible alternative prices indices.

The results of this exercise are presented in the chart "How fast does China grow?" The highest estimate of average annual real GDP growth over the past quarter-century is 11.0% and the lowest is 9.1%. In other words, the current official average growth figure (9.8%) is subject to a margin of error of about one percentage point on either side. That is not a very large margin.

Summing it all up, evidence that China systematically distorts its GDP data is weak, and for practical purposes Chinese GDP data is as useful as that from other major economies. Most of the problems in Chinese economic data result from easily identifiable technical problems, which are a headache in any economy, and particularly in a developing economy like China undergoing rapid structural change. It is reasonable to apply a margin of error of one percentage point to Chinese GDP growth rates, but larger adjustments are not warranted.

In the end the issue is one of trust and transparency. The evidence suggests that China's data is broadly reliable, and the Benford test indicates that if there is any political interference in the published GDP numbers, it is only so occasional that statistical significance tests cannot detect it. But unlike statistical agencies in many other countries, NBS operates behind a shroud of secrecy and is formally accountable only to a small set of even more secretive leaders, not to the broader public. So long as these conditions persist, so will doubts about the quality of Chinese statistics.

How fast does China grow?

Estimates of China's GDP growth using alternative price series (1978-2011)

	Sectoral deflators				Average real GDP growth (1978-2011)
Scenario	**Primary**	**Industry**	**Construction**	**Tertiary**	
Official 1	Implicit deflators, using latest data revisions				9.8%
Official 2	Implicit deflators, using original unrevised data for each year				10.5%
High-growth	Rural retail price index	As first published	Construction/ installation sub-index of FAI price	As first published	11.0%
Low-growth	Rural retail price index, double-deflated by intermediate price index	Producer price index	Building materials sub-index of industrial PPI	Index of approximated deflators for six major tertiary sub-sectors	9.1%

Author calculations from official data