WEI CHEN
Chinese University of Hong Kong

CHANG-TAI HSIEH
University of Chicago

XILU CHEN
Chinese University of Hong Kong

ZHENG SONG
Chinese University of Hong Kong

A Forensic Examination of China's National Accounts

ABSTRACT China's national accounts are based on data collected by local governments. However, because local governments are rewarded for meeting growth and investment targets, they have an incentive to skew local statistics. China's National Bureau of Statistics (NBS) adjusts the data provided by local governments to calculate GDP at the national level. The adjustments made by the NBS have averaged about 5 percent of GDP since the mid-2000s. On the production side, the discrepancy between local and aggregate GDP is entirely driven by the gap between local and national estimates of industrial output. On the expenditure side, the gap is in investment. Local statistics increasingly misrepresent the true numbers after 2008, but there was no corresponding change in the adjustment made by the NBS. Using publicly available data, we provide revised estimates of local and national GDP by reestimating the output of firms in the industrial, construction, and wholesale and retail trade sectors, using data on value-added taxes. We also use several local economic indicators that are less likely to be manipulated by local governments to estimate local and aggregate GDP. These estimates also suggest that the adjustments by the NBS have been insufficient since 2008. Relative to the official numbers, we estimate that GDP growth from 2010 to 2016 was about 1.8 percentage

Conflict of Interest Disclosure: Chang-Tai Hsieh is the Phyllis and Irwin Winkelried Professor of Economics at the University of Chicago. Zheng (Michael) Song is a professor in the Department of Economics of the Chinese University of Hong Kong, where Wei Chen and Xilu Chen are also PhD students. Beyond these affiliations, Song received financial support from the General Research Council of Hong Kong. None of the authors received financial support from any firm or person for this paper or from any firm or person with a financial or political interest in this paper. The views expressed in this paper are those of the authors, and do not necessarily reflect the views of the University of Chicago or the Chinese University of Hong Kong.

points lower and that the investment and savings rate in 2016 was about 7 percentage points lower.

China's national accounts are primarily based on data collected by local officials. However, as documented by Wei Xiong (2018), local officials are rewarded for meeting growth and investment targets. Therefore, it is not surprising that local governments have an incentive to skew the statistics on local growth and investment. The statistical agency of the Chinese government, the National Bureau of Statistics, attempts to correct this bias using administrative data and other sources of data that it gathers directly. The accuracy of the final numbers of aggregate GDP and its components depends on the extent of misreporting by local officials, the data that NBS has at its disposal to correct the misreporting, and the effort it undertakes to do so.

Local GDP is measured via the production approach from three major surveys—of large industrial sector firms, large service sector firms, and "qualified" construction firms. These data are supplemented with surveys of smaller industrial firms and administrative data from other government departments to obtain a number for local GDP on the production side. On the expenditure side, local officials provide estimates of local consumption, investment, government spending, and net exports (vis-à-vis other localities in China and other countries). The two main sources are surveys of household income and expenditures (similar to the U.S. Consumer Expenditure Survey), from which they estimate local consumption, and survey data on investment projects, from which they estimate local investment. Because the sum of local consumption and investment typically exceeds local GDP measured on the production side, the remainder is attributed to local net exports.

The NBS does not simply add up the statistics reported by local governments to arrive at the national aggregates. The NBS also has access to the micro data of the surveys used by local governments, and it supplements these data with economic censuses and administrative data on such categories as land sales, vehicle registration, financial transactions, and foreign trade. Based on these data, the NBS produces its own numbers for national GDP and its components on the production and expenditure sides. The adjustment made by the NBS to the local statistics can be seen by the discrepancies between local GDP and national GDP.

In this paper, we check which of the numbers provided by local governments differ from their national counterparts and, hence, are likely to be

inaccurate. First, we show that the sum of local GDP frequently exceeds national GDP. Second, we compare the sum of the local consumption, investment, and net exports with national consumption, investment, and net exports reported by the NBS. We find little discrepancy between local and national consumption but large discrepancies between local and national statistics on investment and net exports. Third, we compare the sum of value-added taxes of sectors as reported at the local level with the same sectors at the national level. We find large discrepancies for the industrial sector and smaller gaps for the nonindustrial sectors.

We then use two approaches to determine the accuracy of adjustments to the local numbers made by the NBS. First, we adjust national GDP by the difference between value-added tax growth reported by the NBS and value-added tax revenue growth reported by the State Administration of Taxation in the sectors where the value-added tax is a major type of taxation. Our estimate suggests that the adjustments made by the NBS were roughly accurate until 2007–8 but that the adjustments made after this date no longer appear to be accurate. Our baseline estimate of GDP growth from 2010 to 2016 is about 1.8 percentage points lower than the official growth rate. Furthermore, our estimate of the aggregate investment and savings rate in 2016 is about 7 percentage points lower than the official numbers.

We use this same approach to adjust local production and expenditure GDP for each Chinese province. There is a positive relation between our adjustments to local GDP and investment across provinces. This evidence suggests that local governments inflate local GDP by overestimating both local production and local investment.

A second approach is to estimate a statistical model where we estimate the relationship between a set of economic indicators (which are less likely to be manipulated) and local GDP before 2008. We then use parameters of the estimated model, along with the same set of the indicators after 2008, to predict local GDP after 2008. The indicators include satellite night lights, national tax revenue, electricity consumption, railway cargo flow, exports, and imports. We use the method developed by Liangjun Su, Zhentao Shi, and Peter Phillips (2016) to control for hidden economic structural heterogeneities across regions. Using this method, we also find that the corrections made to national GDP no longer appear appropriate after 2008. Encouragingly, the adjustments to local GDP made using the two approaches are highly correlated. This provides additional support for our adjustments.

Our revised numbers for the Chinese national accounts thus indicate that the slowdown in Chinese growth since 2008 has been more severe than

suggested by the official statistics. At the same time, the true savings rate probably declined by about 11 percentage points from 2009 to 2016, with more than 80 percent of the savings decline showing up in the investment rate and the remainder in the external surplus. In this sense, our revised numbers for China's national accounts also indicate that Chinese growth is associated with consumption growth rather than investment and external surpluses.

I. China's GDP Accounting System

The Chinese national and local GDP statistics are compiled separately. Local statistical bureaus provide estimates of local GDP and its components on the production and expenditure sides. The NBS uses the same data collected and used by local governments, along with the data it collects independently, to arrive at a number for national GDP. The number provided by the NBS is the "official" number for Chinese GDP.

Although the local statistical bureaus are de jure branches of the NBS and are supposed to follow the statistical procedures set by the NBS, de facto they are branches of local governments. The budgets of the local statistical bureaus come from local governments, and officials of local statistical bureaus are evaluated and promoted by local governments. Because of this structure, local statistical bureaus are susceptible to pressure from local officials, who may have an incentive to report inaccurate statistics. The NBS is aware of this bias, and it therefore adjusts the numbers of local GDP provided by the local statistical bureaus.

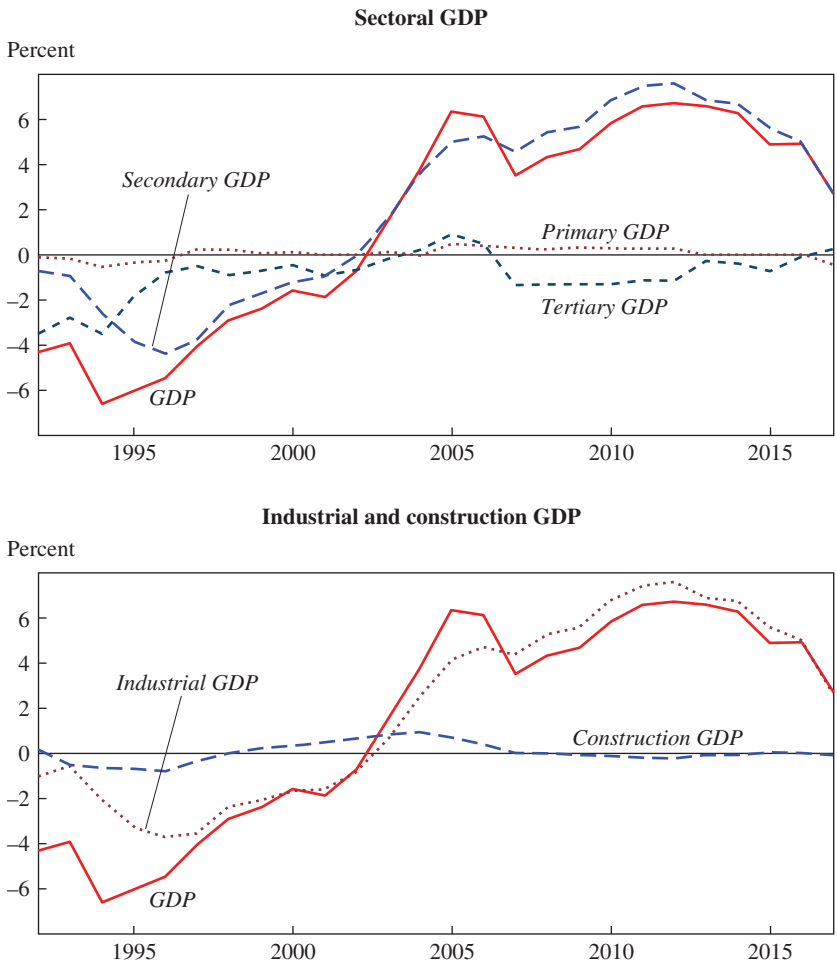
To assess the quality of the official numbers for local and national GDP, we proceed in three steps. First, we compare the sum of local GDP with aggregate GDP provided by the NBS (hereafter, we use the term "aggregate GDP" to refer to the number provided by the NBS). Second, we assess the data used to estimate GDP on the production side. And third, we assess the data used to construct GDP on the expenditure side.

I.A. Comparing Local GDP with Aggregate GDP

The solid line in the top panel of figure 1 shows the magnitude of the adjustment made by the NBS to the local statistics.[1] The figure shows

1. All the national accounts data between 1993 and 2017 were extracted from the NBS website on December 10, 2018. Some numbers—GDP in the primary and tertiary sectors in 2007–16—were updated in February 2019. The changes are very small, and our results are essentially unchanged with the updated numbers. The 1992 provincial data are from Hsueh and Li (1999).

Figure 1. The Gap between Local and Aggregate GDP, by Sector, 1992–2017[a]



Source: National Bureau of Statistics of China.

a. This figure plots the difference between the sum of local sectoral GDP and aggregate sectoral GDP as a percentage of the aggregate GDP for each sector.

the gap between the sum of GDP of each province and aggregate GDP provided by the NBS as a percentage of aggregate GDP. Local governments *understated* GDP relative to the NBS in the 1990s. The sum of local GDP was about 4 to 6 percentage points *lower* than aggregate GDP in the mid-1990s. This pattern changed after 2003. After this date, the sum of local GDP surpassed aggregate GDP and the gap was about

6 percentage points higher than aggregate GDP in 2006. The gap between these two numbers for China's GDP stabilized at about 5 percent of GDP after 2006.

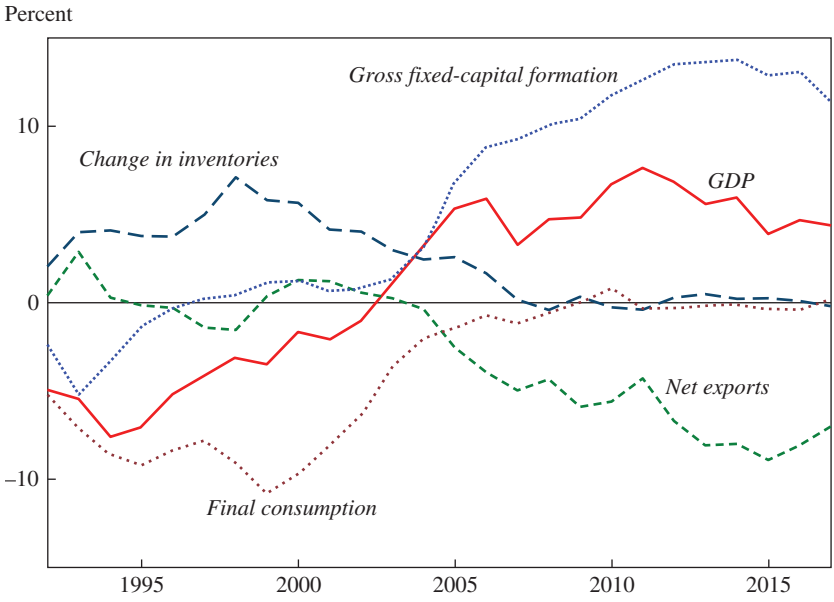
Local statistical authorities and the NBS also provide estimates of local and aggregate GDP by broad sectors. Figure 1 shows the gap between the sum of local GDP and aggregate GDP for each sector as a share of aggregate GDP (for all sectors). The top panel of figure 1 shows the ratio of the sum of local GDP to aggregate GDP for agriculture ("primary"), industry and construction ("secondary"), and services ("tertiary"). Before 2003, the sum of secondary GDP at the local level was lower than aggregate secondary GDP. Furthermore, from about 1997 to 2003 almost all the gap between local and aggregate GDP came from the gap in the industrial sector. Before 1997, some of the gap was due to the discrepancy between local and aggregate statistics for the service sector.

After 2003, all the discrepancy between local and aggregate GDP comes from the industrial sector. The bottom panel of figure 1 shows the comparison of industrial (mining, manufacturing, and public utilities) and construction GDP reported by local governments with that provided by the NBS. As can be seen, the gap between local and national statistics after 2003 is entirely in industry. This finding echoes the research of Carsten Holz (2014) and Ben Ma and others (2014), who also find that the inconsistency between provincial and national GDP mainly came from the industrial sector.

Figure 2 compares GDP expenditures provided by local governments and the NBS. On the expenditure side, there are substantial differences after 2003 in investment ("gross fixed-capital formation") and net exports reported by the two sources. The sum of local investment was close to the national level until 2002. After that date, the sum of local investment exceeded aggregate investment. In 2016, the gap in the two measures of investment reached about 13 percent of GDP. The mirror image is the growing discrepancy between the sum of local net exports and aggregate net exports This gap reached about –8 percent of GDP in 2016. In contrast, the national and local differences in final consumption and changes in inventory were essentially zero after the mid-2000s.[2]

2. Final consumption includes urban and rural household consumption and government consumption. The sum of each of the local consumption components is very close to its national counterpart.

Figure 2. The Gap between Local and Aggregate GDP, by Type of Expenditure, 1992–2017[a]



Source: National Bureau of Statistics of China.

a. This figure plots the difference between the sum of provincial expenditures and aggregate expenditures as a percentage of aggregate expenditures.

To summarize the main findings: First, the sum of provincial GDP is about 5 percent higher than national GDP after the mid-2000s. Second, after 2003 the NBS adjusts industrial GDP and investment downward and adjusts net exports reported by local governments upward. Third, the NBS does not adjust local consumption—the sum of local consumption is roughly the same as the data on national consumption provided by the NBS.

I.B. Production GDP

We do not know whether the adjustments to local GDP by the NBS are appropriate. To answer this question, we need to delve into the details of the data used by the local statistical offices and the data sources behind the adjustments that are made.

INDUSTRIAL GDP Remember that the gap between the local and aggregate numbers on the production side is entirely driven by the industrial

sector. The backbone source for the industry data is the Annual Survey of Industrial Firms (ASIF). These data are from a census of state-owned firms and privately owned firms with sales above 5 million yuan (until 2011) or 20 million yuan (after 2011). The Chinese statistical system calls the firms covered by the ASIF "above-scale" firms. Local statistical bureaus then add to the data from the ASIF an estimate of value added by industrial firms with sales below 5 million yuan (20 million yuan after 2011), which are referred to as "below-scale" firms in the Chinese statistical system, and businesses of self-employed individuals.[3]

We first investigate the data on value added in the ASIF. The micro data from this survey before 2007 have been widely used by researchers. After this date, however, the NBS clamped down on access to the micro data. There are good reasons to believe that the accuracy of this survey has declined over time. First, we can compare the sum of value added in the ASIF with aggregate industrial GDP reported by the NBS. This is shown in the solid line labeled "raw data" in figure 3. Aggregate value added in the ASIF should be lower than aggregate industrial GDP because the latter also includes output by small firms ("below-scale" firms) and the self-employed. However, the sum of value added in the survey exceeds the aggregate industrial GDP reported by the NBS in 2007. So the NBS must have adjusted value added in the ASIF downward.

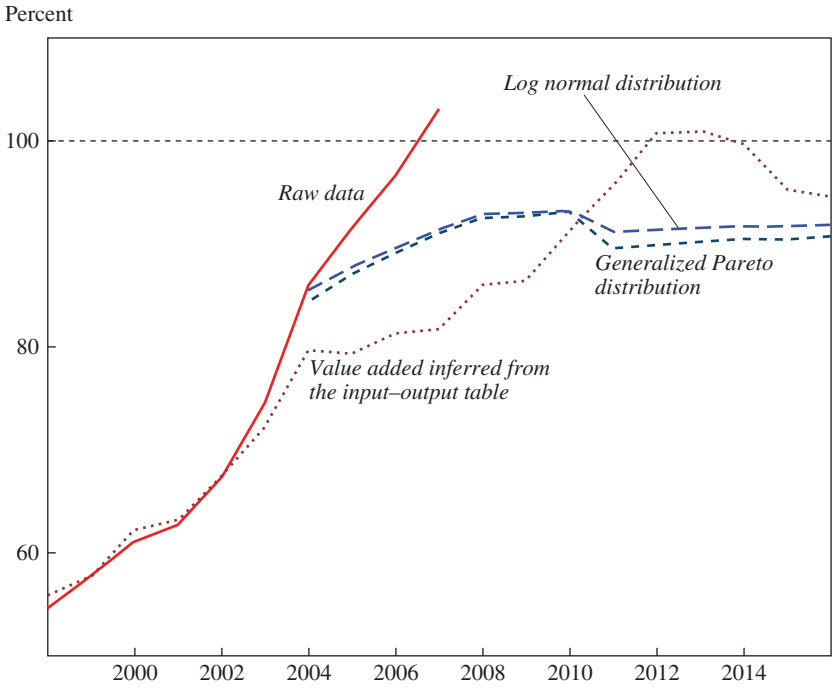
The ASIF does not report firm value added after 2008, so after this date local statistical bureaus used data on gross output in the survey to impute value added.[4] We do the same, using the ratio of gross output to value added given in the input–output (IO) tables.[5] Figure 3 presents aggregate value added imputed in this way from micro data on firm

3. The ASIF was conducted by local statistical bureaus until 2012. Orlik (2014) documents that a more centralized system was implemented nationwide in 2012, whereby firms would enter the statistics directly into an online database controlled by the NBS. Although the goal of this direct reporting system was to prevent local statistical officers from manipulating the data, local governments can still find ways to skew the data. See the case of data manipulation reported by Gao (2016) that is well known by the NBS.

4. For above-scale industrial firms and wholesale and retail firms below, we use their total sales revenue from the *China Statistical Yearbooks* to proxy total gross output.

5. The IO tables for 2002, 2007, and 2012 are from NBS (2006, 2009, 2015). The data for 2005, 2010, and 2015 are from the NBS website. For the years without data, the ratio is calculated by linear interpolation. In 2007, for example, the value-added share in industrial gross output is 0.23 and 0.29 in the IO table and ASIF, respectively. See figure A1 in the online appendix for the value-added shares between 2002 and 2015. The online appendixes for this and all other papers in this volume may be found at the *Brookings Papers* web page, www.brookings.edu/bpea, under "Past BPEA Editions."

Figure 3. Industrial GDP: The Aggregate of Micro Data versus the National Accounts, 1998–2016[a]



Sources: Authors' calculations, based on micro data from China's Industrial Survey, micro data from China's economic censuses, published tabulations from the *Input-Output Tables of China,* and data provided by the National Bureau of Statistics of China.

a. "Raw data" are the ratio of total value added by above-scale industrial firms in the Annual Survey of Industrial Firms (ASIF) micro data to industrial GDP. "Value added inferred from the input–output table" is the total value added inferred from total sales of above-scale industrial firms in the ASIF micro data adjusted by the ratio of value added to gross output in the input–output table. "Log-normal distribution" and "generalized Pareto distribution" are the total value added of above-scale firms inferred from fitting log-normal and Pareto distributions to the micro data (as a share of industrial GDP in the national accounts).

sales in the ASIF as a share of industrial GDP reported in the national accounts. The share exceeded 100 percent in 2012 and 2013. Again, the only explanation for this is that the NBS adjusted firm sales in the ASIF downward.

Remember that the ASIF only provides information for above-scale firms. For below-scale firms and the self-employed, the local statistical bureaus and the NBS rely on a survey of these two types of establishments (Xu 2004). However, the micro data from this survey are not publicly

Table 1. Aggregates in Census Micro Data versus National Accounts (percent)[a]

Year	*Above-scale firms in the census/ national accounts*	*Below-scale firms in the census/ national accounts*
Industrial firms		
2004	79.9	7.5
2008	88.1	5.6
Wholesale and retail trade firms		
2008	63.2	12.8

Sources: Authors' calculations, based on micro data from China's economic censuses and data provided by the National Bureau of Statistics of China.

a. This table reports total value added of above-scale (sales above 5 million yuan) and below-scale (sales below 5 million yuan) firms in the industrial sector (upper panel) and wholesale and retail sectors (bottom panel) as a percentage of the corresponding sectoral GDP.

available; nor is there information about the sampling and how aggregates are constructed from the survey.

We therefore take two approaches to measure the aggregate value added of small industrial firms and the self-employed. First, we use the micro data from the 2004 and 2008 economic censuses. These two censuses are a complete enumeration of all Chinese firms (including small ones), with the exception of the self-employed. The left column of table 1 shows that total value added of above- and below-scale firms in the micro data from the 2004 Economic Census is about 80 percent and 7 percent, respectively, of aggregate industrial GDP reported in the national accounts.[6] So if the 2004 national accounts are accurate, about 13 percent of industrial GDP in the national accounts is not in the census and should be attributed to the self-employed. The equivalent numbers for the 2008 Economic Census are about 88 percent and 6 percent of industrial GDP in the 2008 national accounts. The sharp increase in the output share of above-scale firms between 2004 and 2008 is consistent with the fast-growing economy, where a larger share of firms exceeds the threshold of 5 million yuan sales over time.

However, what is remarkable is that the increase in the share of above-scale firms between the 2004 and 2008 censuses reverses after 2008. The line labeled "Value added inferred from the input–output table" in figure 3 shows that the output share of above-scale firms fell by about

6. Instead of using self-reported firm value added (because of the fear that value added is inflated in the censuses, as it is in ASIF), we convert firm sales into value added by the ratio of value added to gross output in the IO tables.

6 percentage points from 2013 to 2016. If the national accounts data are accurate, the share of below-scale and the self-employed implied by the NBS's number for industrial GDP must have increased in recent years. These are precisely the firms for which the micro data are not available to the public, and this is also at odds with the trend whereby the share of above-scale firms rises over time with a growing industrial sector.

We can also estimate the importance of below-scale firms by making distributional assumptions. Specifically, we assume that firm sales follow either a log-normal or a Pareto distribution and estimate the parameters of the two distributions from the micro data from the economic census.[7] Because the economic census does not cover the self-employed, we assume that the value-added share of the self-employed in aggregate industrial GDP is about 13 percent in 2004 and about 6 percent in 2008 (see table 1). The share is linearly interpolated between 2004 and 2008 and is set to 6 percent for the post-2008 period.

Figure 3 shows the share of above-scale firms based on these two distributional assumptions. There are two main differences between the official and estimated output shares of above-scale industrial firms. First, the adjustment of the sales threshold in 2011 should generate a drop in the output share of above-scale firms in our estimates. However, there is no such drop in the official numbers.[8] Second, our estimates suggest a modest increase in the value-added share of above-scale firms since the sales threshold adjustment.[9] In contrast, the share declined after 2013 in the official numbers, which does not seem plausible.

Another way to gauge the accuracy of the NBS's estimate of industrial GDP growth is to use information on the growth of revenue from

7. We fit the two distributions by choosing parameters to fit the mean of log sales in each size percentile of industrial firms in the 2008 economic census. Specifically, we estimate the mean and standard deviation of the log-normal distribution and the mean and shape parameter of the generalized Pareto distribution to match the firm size distribution. We then assume the distribution in other years has the same standard deviation (for the normal distribution) or shape parameter (for the Pareto distribution) but a different mean parameter. We calibrate the mean parameters in the other years by targeting the average sales of above-scale industrial firms in each year. Applying the threshold of 5 million and 20 million yuan for sales before and after 2011, respectively, we can infer the output shares for above- and below-scale industrial firms.

8. In the online appendix, we present evidence that the 2010 ASIF covers fewer above-scale firms than it should. In other words, firm sales data are likely to be manipulated, disguising the otherwise discontinuous sales proportions of above-scale firms.

9. If we assume that the value-added share of individual businesses fell after 2008, as it did in the 2004–8 period, the increase in the estimated value-added share of above-scale firms would be more pronounced.

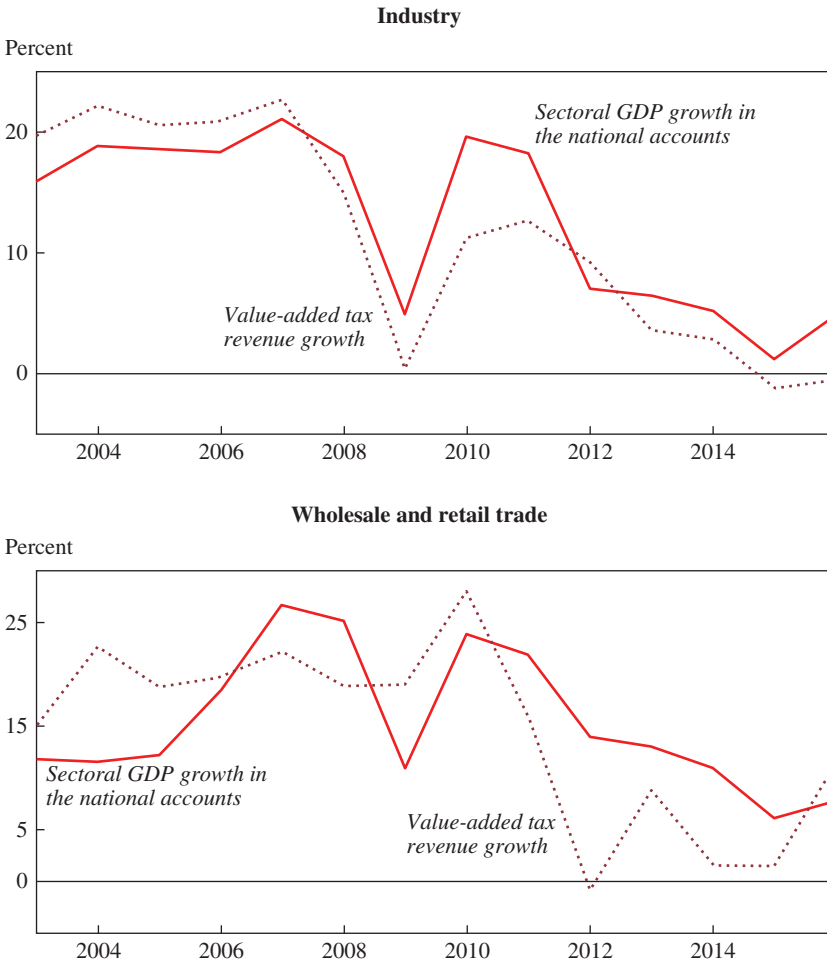
value-added taxes on industrial firms. China imposed a 17 percent value-added tax on essentially all industrial firms until 2018. There were three main exceptions. First, the value-added tax imposed on small firms (with annual sales below 0.5 million yuan) was 3 percent of their sales. Second, the value-added tax rate was 13 percent for a selected set of industrial goods. The online appendix shows that these different tax rates had negligible effects on value-added tax revenue growth. Third, a significant proportion of the domestic value-added tax (41 percent in 2015) was refundable through export tax rebates, which varied considerably across goods and over time.[10] To ensure that our estimates are not affected by tax rebates, we use data on revenues from value-added taxes *gross* of rebates for exports.

Furthermore, after 1994, there was little fraud vis-à-vis and evasion of the value-added tax. The State Administration of Taxation implemented the so-called Golden Taxation Project in 1994. A computerized taxation data network has also been in full operation since 2005, which allows the tax authorities to cross-check the input and output value-added tax at each stage of the production and distribution of goods and services. The effective value-added tax rate, defined as the ratio of the industrial value-added tax to industrial GDP net of the value-added tax, increased from 10.4 percent in 2001 to 12.9 percent in 2007, reflecting the improved tax enforcement in the period. There are several possible reasons why the effective tax rate was below the main statutory tax rate of 17 percent. The different tax rates mentioned above matter, but they cannot be quantitatively important. Another possibility is inflated industrial output. To account for this difference of nearly 4 percentage points, industrial GDP would need to be overestimated by nearly a quarter in 2007. Although tax evasion is hard to do for transactions within the industrial sector, it may be easier for industrial output sold to the sectors to which value-added tax does not apply.[11] But even if there is some tax fraud and evasion, as long as their degree does not increase, revenues from the value-added tax on industrial firms should be proportional to industrial GDP.

10. The domestic value-added tax revenue and the tax rebate for exports can be found in the *China Taxation Yearbook*.

11. The 2007 IO tables suggest that the share of industrial output to the construction and service sectors (excluding wholesale and retail) is 15.9 percent. If industrial firms hide all their output sold to the construction and service firms that do not have incentives to ask for value-added tax invoice, the effective value-added tax rate would be lowered by 2.7 percentage points.

Figure 4. The Growth in Value-Added Tax Revenues and GDP, 2003–16



Sources: National Bureau of Statistics of China; *China Taxation Yearbook.*

The top panel of figure 4 compares the growth rate of revenues from domestic value-added taxes with the growth rate of industrial GDP. The growth rate of revenues from value-added taxes exceeds that of industrial GDP before the mid-2000s, consistent with the improved enforcement of value-added taxes. However, after 2007, the growth rate of tax revenues is lower than the growth rate of industrial GDP. Furthermore, the gap has been widening over time. In 2010 to 2012, for instance, value-added tax

revenue growth is about two-thirds that of industrial GDP. The growth in tax revenues dropped to about half the growth rate of industrial GDP growth in 2013 and 2014, and even became negative in 2015 and 2016. Consequently, the effective value-added tax rate fell from about 12.9 percent in 2007 to about 9.3 percent in 2016.

A few tax policies introduced after 2007 may lower the effective value-added tax rate. The most relevant change is the value-added tax deduction on fixed-asset investment for domestic firms.[12] This policy was first introduced in three provinces of Northeast China in 2004, and was later extended to six provinces (Cai and Harrison 2018; Zhang, Chen, and He 2018). The central government unexpectedly increased the coverage to all provinces at the end of 2008, as part of the stimulus package in response to the global financial crisis. The nationwide policy became effective on January 1, 2009. Although the policy obviously reduced industrial value-added tax revenue in the transition period between 2004 and 2009, it is hard to estimate the extent to which value-added tax revenue growth was affected. The main obstacle is that we do not know how much of fixed-asset investment is deductible from the value-added tax.[13] A simple fix is to look at the value-added tax revenue growth after 2009.[14] The average industrial value-added tax revenue growth was about 5.3 percent between 2009 and 2016, which is about 3.4 percentage points lower than the average industrial GDP growth in the same period. The gap is similar to that of about 3.5 percentage points in the period 2007–16.

Another important policy change is the reform of replacing business tax with value added tax initiated in 2012 and completed in 2016. Because the purchase of service goods became deductible from value-added tax,

12. Foreign firms have always been eligible for the tax deduction.

13. The value-added tax deduction only applies to purchase of machinery, mechanical apparatus, means of transportation and other equipment, tools and fixtures related to production, and business operations (Ministry of Finance and State Administration of Taxation 2008a). According to the compositions in China's fixed-asset investment survey, which severely overestimated the level of fixed-asset investment, as is shown below, purchase of equipment and instruments accounts for about 40 percent of fixed-asset investment in the industrial sector. But the purchase of equipment and instruments includes items such as those for the newly built production department that are not eligible for value-added tax deduction.

14. The remaining concern is that the effect of the policy might be persistent by increasing industrial firms' investment rate in the subsequent periods. Using firm survey data from China's State Administration of Taxation, Chen and others (2019) found that the average firm investment rate increased by 2.6 percentage points in 2009 but then decreased by 0.7 percentage point in both 2010 and 2011. The diminishing effect on investment rate suggests that the policy should not lower value-added tax revenue growth in 2010 and onward.

the reform may also lower value-added tax revenue growth in the industrial sector. The pilot started in the transportation industry (excluding railway) and several "modern" service sectors in Shanghai from January 2012. The policy was extended to Beijing and seven other provinces and cities from August 2012 and then to all provinces from August 2013.[15] Railway transportation, the postal service, and telecommunications were added to the list of "modern" service sectors in 2014. In the online appendix, we identify 10 industries in the IO tables according to the description of "modern" service sectors in the documents. The transportation industry and all the modern service sectors accounted for 2.4 percent and 2.3 percent of industrial input, respectively, in 2012 (see the online appendix). For two main reasons, we assume that this reform did not affect industrial value-added tax revenue growth between 2010 and 2016. First, the purchase of transportation services was deductible from the value-added tax even before the reform started. Second, the input share of the modern service sectors was not big enough to generate a significant effect on industrial value-added tax revenue growth in the reform period.

Finally, the value-added tax policy for small taxpayers (defined as those with annual sales below a half million yuan) was adjusted twice after 2007. From August 2013 onward, taxpayers with annual sales below 240,000 yuan were exempted from the value-added tax. The cutoff was increased to 360,000 yuan in October 2014.[16] The effectiveness of the tax reform can be seen from the share of the value-added tax paid by small taxpayers, which is publicly available in the *China Taxation Yearbook*. In fact, the share was quite stable, at about 4 percent, between 2010 and 2016, and even increased from 3.7 percent in 2012 to 4.4 percent in 2013. One explanation is that firms and individual businesses with annual sales below the cutoff contribute little to total value-added tax revenue.

To summarize the main findings about the reliability of the NBS's estimate of industrial GDP: First, the micro data from the ASIF have overstated aggregate output at least since 2007. Second, the aggregate industrial GDP provided by the NBS implies an increasing share of below-scale firms and the self-employed in the industrial sector after 2012. Third, the growth rate of the aggregate industrial GDP has exceeded the growth rate of revenues from value-added taxes on industrial firms since 2008. Based on these three pieces of evidence, we conclude that despite the

15. See Ministry of Finance and State Administration of Taxation (2011).

16. See Ministry of Finance and State Administration of Taxation (2013, 2014). See also Lardy (2014) for the evolution of policies toward the private sector.

adjustments made by the NBS to local industrial GDP, the official numbers for the aggregate industrial GDP—and by extension, the aggregate GDP for all sectors—are likely to overstate the truth after 2007–8.

NONINDUSTRIAL GDP Turning to the nonindustrial sector, the NBS conducts surveys for all "qualified" construction firms, above-scale wholesale and retail firms, above-scale hotel and catering firms, and all real estate developers and operators.[17] We first look into the wholesale and retail trade sectors, which accounted for about 10 percent of the aggregate GDP in 2016. Though the published tabulations of the surveys provide total sales of above-scale wholesale and retail firms, value added is not reported. We thus convert total sales to value added, following the procedure used by Bai and others (2019).

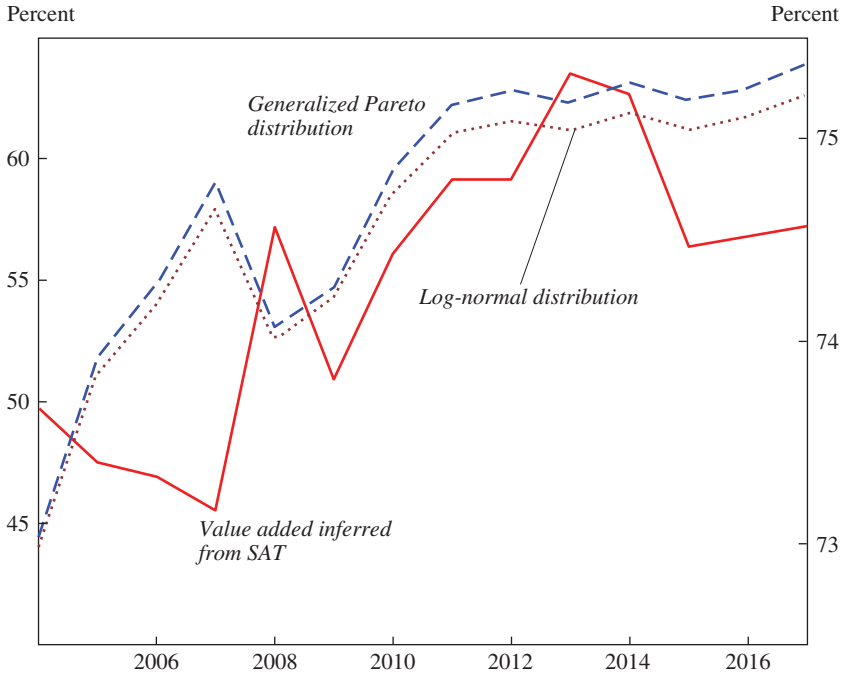
Based on this imputation, the solid line in figure 5 plots the value added of above-scale wholesale and retail firms in the published surveys as a share of official aggregate GDP in the wholesale and retail trade sectors. Using the same procedure described in the previous section, we estimate the parameters of the distribution of firm size in the wholesale and retail sectors in the 2008 Economic Census. We then calibrate the mean parameters of the log-normal and Pareto distributions to match average firm sales in each year. We further assume the value-added share of the self-employed in wholesale and retail GDP is fixed at 24 percent (the number suggested by the 2008 Economic Census; see table 1). The estimated models suggest that the share of above-scale wholesale and retail firms in aggregate GDP in these sectors has increased slightly in recent years. Like what we see for the industrial sector, this is also at odds with the dramatic drop in 2014 and 2015 in the official data.

The bottom panel of figure 4 compares domestic value-added tax revenue growth from the wholesale and retail sector with GDP growth in these sectors as provided by the national accounts.[18] Like what happened in the industrial sector, tax revenue outgrew sectoral GDP before the mid-2000s, but the pattern was reversed after 2010, except for 2016. The average difference from 2010 to 2016 between tax revenue and GDP growth is about 6 percentage points, suggesting that true wholesale and retail GDP is also likely to be overstated in the national accounts.

17. The sales threshold for wholesale and retail firms is 20 and 5 million yuan, respectively. The sales threshold for hotel and catering firms is 2 million yuan.

18. Value-added tax revenue accounts for about 40 percent of total tax revenue in the industrial and the wholesale and retail trade sectors. See figure A4 in the online appendix.

Figure 5. Wholesale and Retail GDP: The Aggregate of Micro Data versus the National Accounts, 2004–17[a]



Sources: Authors' calculations, based on micro data from China's economic censuses, published tabulations from the *Input-Output Tables of China*, and data provided by the National Bureau of Statistics of China.

a. This figure presents the ratio of the value added of above-scale retail and wholesale firms to wholesale and retail GDP in the national accounts. "Value added inferred from SAT" is the value added imputed from the total sales of above-scale wholesale and retail firms in the firm survey adjusted by the ratio of value added to sales in the data from the State Administration of Taxation (SAT). "Log-normal distribution" and "generalized Pareto distribution" are the total value added of above-scale firms inferred from fitting log-normal and Pareto distributions to the micro data.

The construction sector accounts for about 7 percent of GDP after 2010 (see the bottom panel of figure 1). Surprisingly, the output share of "qualified" construction firms has fallen in recent years (see figure A2 in the online appendix). Although no sales threshold applies to construction firms, larger construction firms are more likely to be qualified. For the same reason discussed above, pure economic forces are hard to reconcile with the observed output share change.

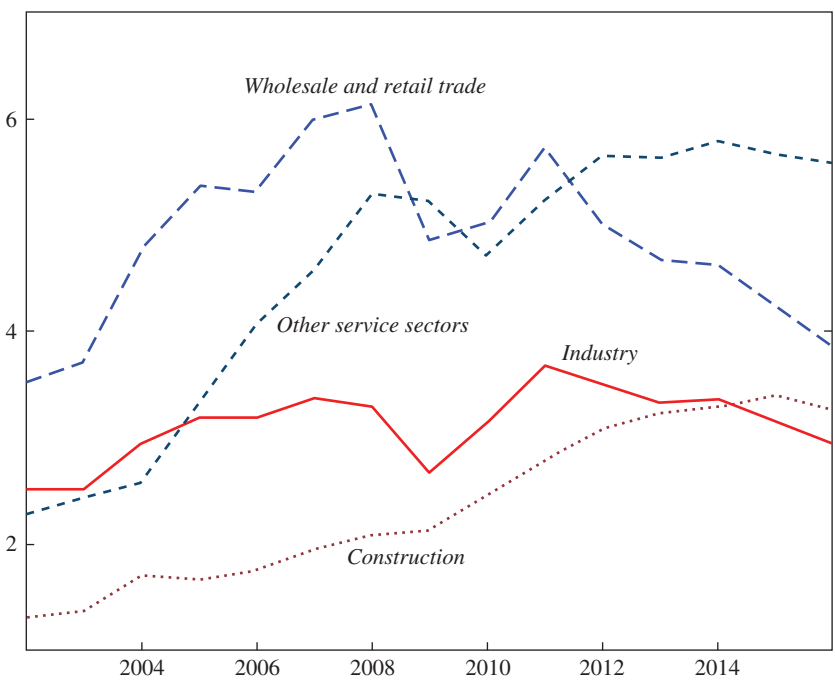
It is more difficult to examine the reliability of nonindustrial GDP because we do not have access to firm-level data other than the 2008

Economic Census. Furthermore, the value-added tax only applied to the industrial and wholesale and retail sectors before 2017. A close substitute is the corporate income tax. Like the value-added tax, a major proportion of corporate income tax revenue (60 percent) is paid to the central government.[19] Unlike the highly rigid value-added tax rate, there are many exemptions and special rates for corporate taxes. For example, there are special corporate income tax rates for labor-intensive and high-technology firms. The enforcement of corporate income taxes is also weaker than that of value-added taxes. Figure 6 plots sectoral corporate income tax revenue as a percentage of sectoral GDP for these four sectors: industry, construction, wholesale and retail trade, and services excluding wholesale and retail trade. The ratio of corporate income tax revenue to GDP increased in all the sectors before 2007. This is likely to be driven by both growing firm profitability and enhanced tax enforcement in the period. The ratio of corporate income tax revenue to GDP decreased dramatically in industry and in wholesale and retail trade after 2011, consistent with the growth slowdown in the two sectors. The ratio was fairly stable in the service sector excluding wholesale and retail, by about 2 or 3 percentage points higher than the ratio in industry and wholesale and retail trade in recent years. Construction is the only sector where the ratio of corporate income tax revenue to GDP kept increasing until 2015.

That said, corporate income tax revenue is still informative. Figure 7 compares corporate income tax revenue growth with GDP growth in the four sectors. For industry and wholesale and retail trade (respectively, the figure's top left and bottom left panels), the results are similar to those in figure 4: The sectoral GDP growth is above tax revenue growth in recent years. For construction (the top right panel), corporate income tax revenue growth is above GDP growth in most years. Given the fact that the ratio of corporate income tax revenue to GDP was very low in the construction sector in earlier years (figure 6), the strong corporate income tax revenue growth might be a consequence of much improved tax enforcement in that sector. Most interestingly, the bottom right panel of figure 7 shows that GDP growth seems in line with corporate income tax revenue growth in

19. The corporate income tax revenue paid to the central/local government is from the *China Taxation Yearbook*. A total of 75 percent and 50 percent of value-added tax revenue was paid to the central government before and after 2016. The sharing mechanism prevents local governments from inflating corporate income tax revenue, which would otherwise incur direct losses to local fiscal budgets. There is evidence that local governments manipulate business tax revenue, which applied to most service sectors and went entirely to local fiscal budgets (Lei 2017). The business tax was replaced by the value-added tax in 2017.

Figure 6. Corporate Income Tax Revenue as a Percentage Share of GDP, 2002–16[a]



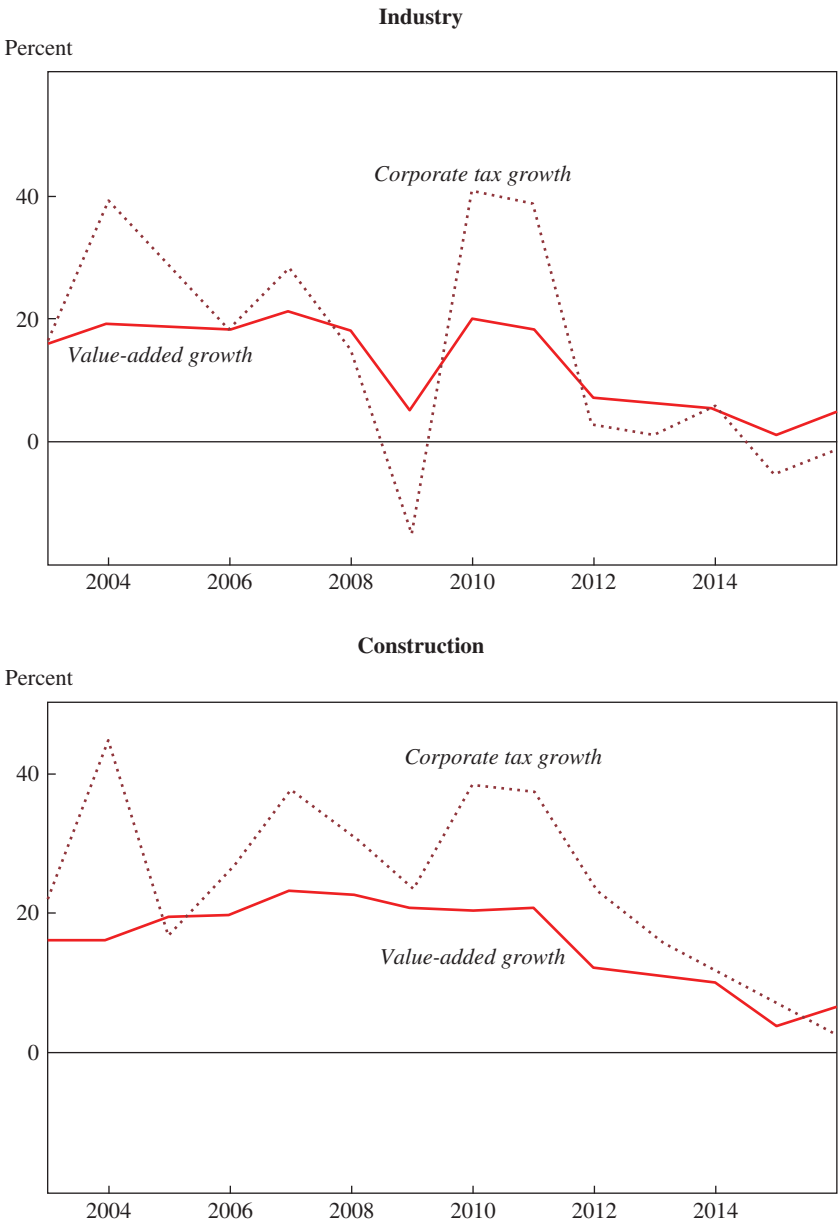
Sources: National Bureau of Statistics of China; *China Taxation Yearbook.*

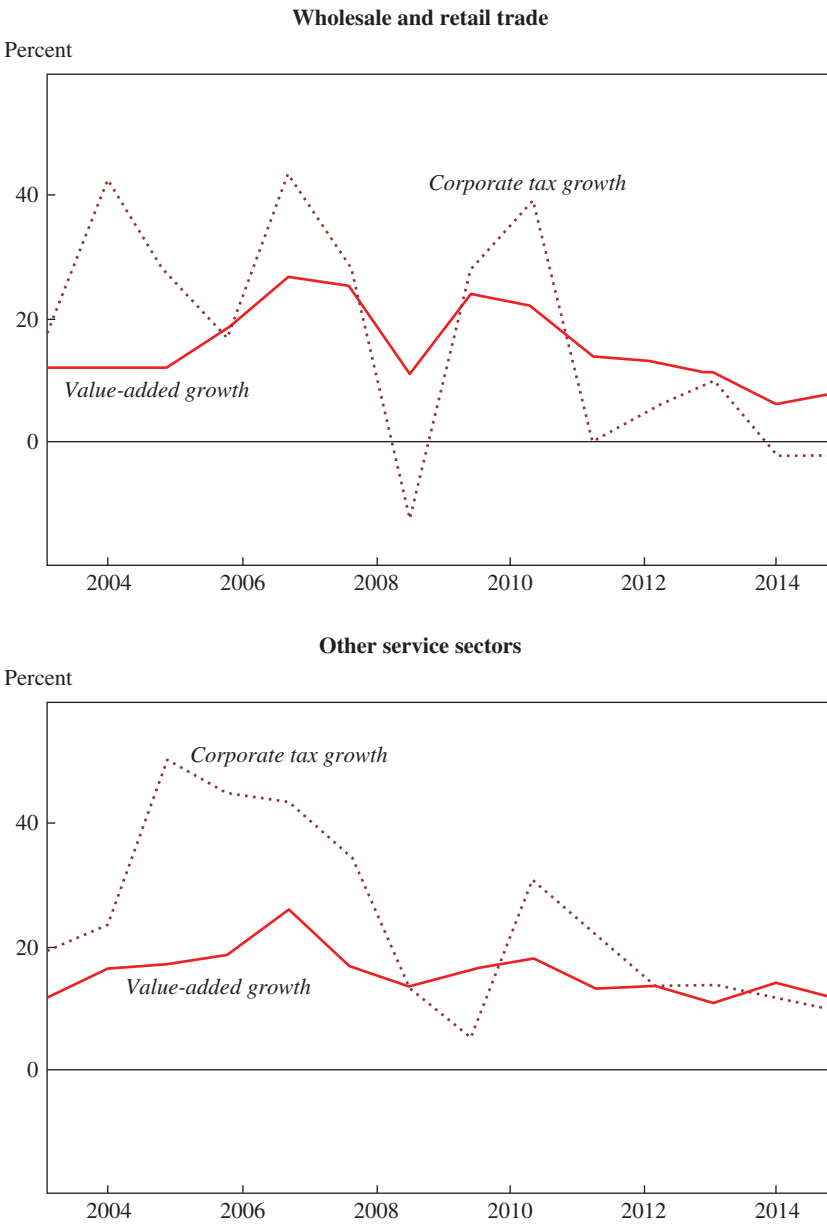
a. This figure plots the ratio of corporate income tax revenue to GDP in the industrial sector, construction, wholesale and retail trade, and other services excluding wholesale and retail.

the service sector excluding wholesale and retail trade. Because the ratio of corporate income tax revenue to GDP did not change much, we view the figure's bottom right panel as evidence that official estimates of GDP growth in the service sector (excluding wholesale and retail trade) in the national accounts are reliable.

In sum, while the growth in the wholesale and retail trade sectors is likely to be overstated in the official statistics, there is no evidence that the official statistics for the other service sectors are inaccurate. However, the effect of inaccuracies in the wholesale and retail trade sectors is important, given that these are two large sectors. Note also that figure 2 shows no gap between local and aggregate statistics for the service sector after 2003. Figure 2 simply tells us exactly where the NBS has adjusted the local numbers, not whether the adjustment or the absence of an adjustment is appropriate.

Figure 7. Sectoral GDP and Corporate Income Tax Revenue Growth, 2003–16





Sources: National Bureau of Statistics of China; *China Taxation Yearbook.*

I.C. Expenditure GDP

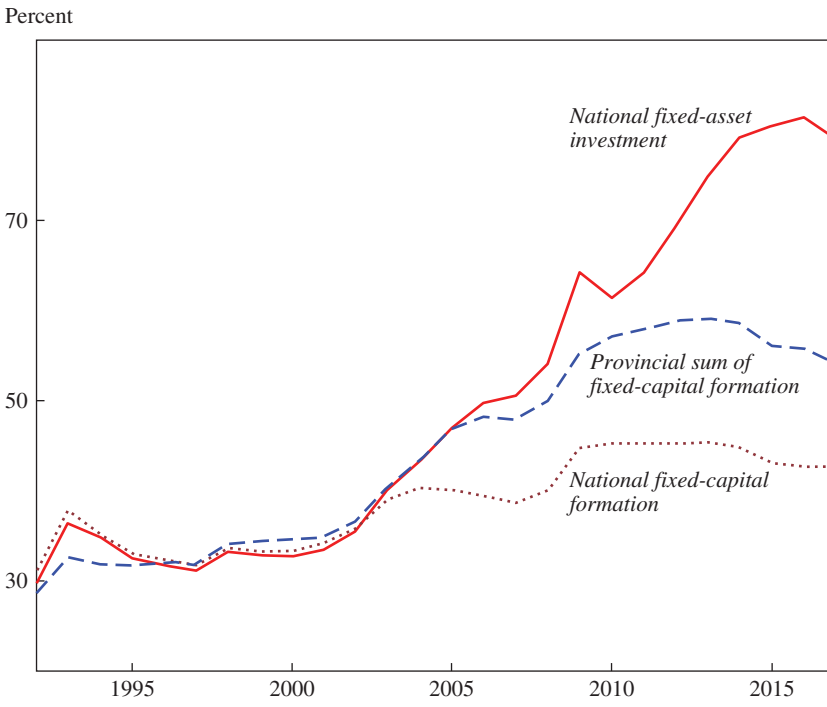
We now examine the underlying data used to construct GDP expenditures. As discussed above, government expenditures reported by local governments are consistent with those reported by the NBS. Furthermore, this information is based on administrative and verifiable data on public expenditures, so it is likely to be reliable. We therefore focus on household consumption, investment, and net exports.

The urban and rural household surveys are the backbone of aggregate household consumption. From these two surveys, the local statistical bureaus and the NBS directly take aggregates of household spending on food, clothing, household facilities, education, culture and recreation services, and miscellaneous goods and services. The other components of household consumption also use the household surveys but are adjusted for (1) accounting discrepancies (that is, medical expenditures paid by the government are not in the two household surveys but are included in final consumption); (2) biases in the surveys (that is, high-income households are underrepresented in the two household surveys). Xianchun Xu (2014) describes in detail how the NBS arrives at consumption aggregates by adjusting the data from the two household surveys. These adjustments are based on administrative data from the relevant government departments. For instance, the NBS uses social security income and expenditure data to adjust medical expenditures. Another example is to use the production, sales, and import data on automobiles from the Association of the Automobile Industry and the Department of Public Security (which registers all new automobiles) to adjust consumption data for transportation and communication. This helps correct the bias caused by underrepresented high-income households, which are more likely to purchase automobiles.

Investment spending is officially called fixed-capital formation (FCF) in the Chinese national accounts. These data are primarily based on reports of fixed-asset investment (FAI) by local governments. FAI measures gross investment spending, given that it includes expenditures on land purchases and used capital. Therefore, local statistical authorities use a survey of land purchases and used capital to subtract these two items from FAI in order to estimate net investment spending—that is, FCF.

However, there is abundant evidence that local data on gross investment have become more unreliable. In contrast with ASIF, which is based on a firm's financial statement, local administrative data on gross investment are based on reports of investment projects by local governments. There is no audit of these data; nor are there any consequences for misreporting

Figure 8. Fixed-Asset Investment versus Fixed-Capital Formation, 1992–2017



Source: National Bureau of Statistics of China.

this information. In addition to the incentives for local officials to misreport this number, tax considerations may also lead to the inflation of FAI.[20] In 2014, Xu Xianchun, a vice director of the NBS at the time, publicly stated that FAI is inflated by local statistical offices (Xu 2014). According to him, "Some regions tend to set unrealistic investment program targets, which are in turn assigned to lower-level governments as performance indicators."

Figure 8 shows that the gap between FAI and national FCF has increased since the early 2000s.[21] In 2015, the gap between aggregate FAI and net investment provided by the NBS (that is, FCF) reached about 38 percent

20. For instance, the Ministry of Finance and the State Administration of Taxation (2008b) introduced a policy that allows taxpayers to deduct fixed-asset investments from the value-added tax.

21. National and provincial FAI data are from the *China Statistical Yearbook*. The national FAI is identical to the sum of provincial FAIs.

of official GDP. In theory, the gap between the two measures of investment should only reflect land purchases and spending on used capital. The purchase of land and used capital does account for most of the difference in the early 2000s, but these two items are much too small to account for the gap in recent years (see figure A3 in the online appendix).[22] The enormous gap between FAI and FCF suggests that investment spending is overstated by local statistical offices and that the NBS has made large adjustments to these data to arrive at a number for aggregate net investment.

It is also evident that even local statistical bureaus adjust FAI downward when estimating local net investment. Figure 8 shows that the sum of provincial FAI exceeds provincial FCF by about 24 percent of GDP in 2015. This difference is, once again, too big to be reconciled by accounting discrepancies like purchase of land and used capital. But the extent to which local statistical bureaus adjust the data on FAI is obviously less than the adjustment by the NBS. The sum of FCF at the provincial level exceeds aggregate FCF by about 14 percent of GDP in 2015.

Note that the adjustment made by the NBS to investment spending provided by the local statistical bureaus is larger than the adjustment made to local estimates of industrial GDP. Because local GDP on the production side has to be equal to GDP on the expenditure side, local statistical bureaus use local net exports as the residual to balance production and expenditure GDP. This can be seen in figure 2, where the growing discrepancy between net exports and local net outflows is the mirror image of the gap between national and local FCF in figure 8.

The NBS completely disregards local estimates of net exports. Instead, it calculates aggregate net exports from data on net exports of goods in the customs data. For this reason, aggregate net exports in the national accounts are very close to net exports reported in the customs data. In contrast, local estimates of net exports are not based on any data and are simply a residual used to equalize local production and expenditure GDP.

To summarize the main findings: Local statistical bureaus inflate investment and, to a smaller extent, inflate output in the industrial and wholesale and retail trade sectors. Because investment data are easier to manipulate

22. Holz (2013, 2018) also documents the growing discrepancies between provincial and national investment and the widening gap between FAI and FCF. Liu, Zhang, and Zhu (2016) also show that the gap between FAI and FCF cannot be explained by land sales and purchases of used assets and buildings. Data for land sales and purchases of used assets and buildings are from the *China Land and Resources Statistical Yearbook* and the *Statistical Yearbook of the Chinese Investment in Fixed Assets*, respectively.

(the amount of investment is project-specific, and is disconnected from investing firms' financial statements), the misstatement of investment spending is more severe than the bias in GDP. The gap between the two is "reconciled" by the large net inflows of goods and services reported by local governments. In contrast, consumption data based on household surveys are more reliable.

II. Revised Estimates of GDP Growth

Given the findings presented above, this is the obvious question: What are the "true" estimates of China's GDP growth? Here, we make two efforts to come up with a number. First, we use alternative data from tax records to generate alternative measures of GDP on the production side. We then use these measures to reestimate aggregate investment as well as local GDP. Second, we take a data-fitting approach and use external data that are not likely to be manipulated by local governments to estimate GDP.

II.A. Adjusting National Accounts with Tax Data

Our first approach to estimate "true" GDP is built on three assumptions. First, we assume that industrial output reported by local statistical officers has not been reliable since the late 2000s. Second, we assume that nonindustrial output reported by local statistical officers is reliable. And third, we assume that industrial value-added tax revenue is proportional to true industrial value added.

The validity of the first assumption comes from the facts in the previous section. In particular, industry is the only major sector for which the NBS significantly adjusts locally reported output data. The second assumption is partly based on the evidence that corporate income tax revenue grew in tandem with value-added taxes in the service sector, and is partly made for practical reasons, given that we do not have reliable data to indicate true output in most nonindustrial sectors.[23] We relax the second assumption below. The third assumption is the strongest one, because it hinges on two institutional features discussed in the previous section. First, China has developed a sophisticated value-added taxation system to minimize tax fraud and evasion. Second, local governments do not

23. See also Bai and others (2019) for more evidence on the reliability of service data in the national accounts.

have incentives to overstate value-added tax revenue, because otherwise they would incur direct local fiscal losses. We have also discussed several tax policy changes that are likely to impair the proportionality between industrial GDP and value-added tax revenue. The effect of replacing the business tax is quantitatively small before 2016, and so is the effect of value-added tax exemption on small firms and individual businesses. Our adjustment begins with 2010, after the end of the value-added tax deduction on fixed-asset investment.

In the simplest case, our adjusted GDP assumes this equation:

$$\text{Adjusted GDP}_t = \text{Official GDP}_t - \Delta\,\text{Industrial GDP}_t \tag{1}$$

where $\Delta X_t \equiv \text{Official } X_t - \text{Adjusted } X_t$ denotes the adjustment in variable X and

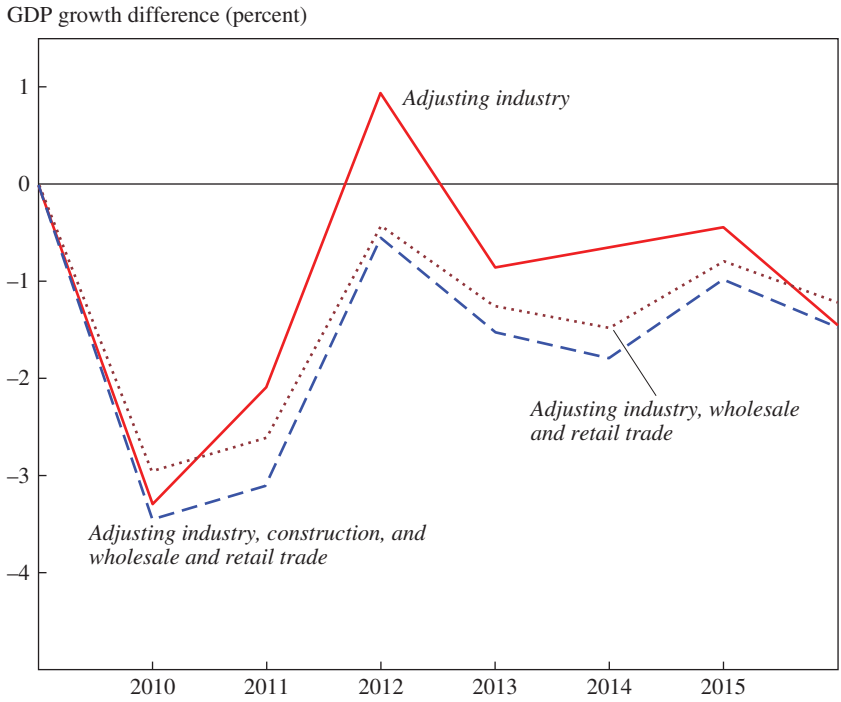
$$\text{Adjusted Industrial GDP}_t = \text{Adjusted Industrial GDP}_{t-1} \cdot \text{Industrial VA Tax Revenue Growth}_t.$$

The dotted line in figure 9 plots the difference between our adjusted GDP growth and the official nominal GDP growth (the solid line). The adjusted growth rate is always below the official growth rate, except in 2012. Figure 4 above shows that industrial value-added tax revenue growth is about 3.4 percentage points lower than official industrial GDP growth after 2009. The industrial sector accounts for roughly one-third of China's GDP. Therefore, correcting the overreporting of industrial output lowers GDP growth from 2010 to 2016 by about 1.1 percentage points; see the second column of table 2.

We relax the third assumption by also adjusting value added in wholesale and retail output growth. Because wholesale and retail value-added tax revenue growth is also below its GDP growth in the national accounts (figure 4), adjusting output in both the industrial and wholesale and retail sectors would further cut nominal GDP growth in recent years (the dashed line in figure 9). After we also adjust the growth rate of the wholesale and retail sectors, our estimate of the growth rate of nominal GDP from 2010 to 2016 is about 1.5 percentage points lower than the official rate (the third column of table 2).

We next look into expenditure-side GDP accounting. Based on the discussions in the previous section, we assume that the official statistics

Figure 9. The Gap between Adjusted and Official Nominal GDP Growth, 2009–16[a]



Sources: Authors' calculations, based on data provided by the National Bureau of Statistics of China and published tabulations from the *China Taxation Yearbook*.

a. This figure plots the difference between adjusted nominal GDP and official nominal GDP growth rates. The solid line adjusts industrial GDP growth only. The dotted line adjusts both industry and wholesale and retail trade GDP growth. The dashed line adjusts industry, construction, and wholesale and retail trade GDP.

on aggregate consumption and net exports are accurate. FCF is then obtained by

$$\text{Adjusted FCF} = \text{Adjusted GDP} - \text{Final Consumption} - \text{Inventory Change} - \text{Net Exports}.$$

However, this adjustment is incomplete. Most of the output of the construction sector is classified as investment on the expenditure side of GDP. Although the NBS does not adjust local estimates of the construction sector's output (the bottom panel of figure 1), our estimated investment

Table 2. The Adjusted Growth Rate of Nominal GDP[a]

		Adjustment by value-added tax			*Adjustment by statistical model on industrial GDP*		
Year	*Official data*	*Adjusting industrial GDP*	*Adjusting industrial and W&R GDP*	*Adjusting industrial, construction, and W&R GDP*	*All variables*	*Without light*	*Without light and national tax*
2009	9.25	9.25	9.25	9.25	7.09	7.97	5.26
2010	18.32	15.02	15.36	14.86	20.27	19.13	19.78
2011	18.47	16.37	15.86	15.36	19.42	18.17	17.76
2012	10.44	11.38	10.00	9.88	9.98	9.95	9.33
2013	10.16	9.29	8.89	8.63	10.12	9.51	10.08
2014	8.19	7.54	6.70	6.39		7.56	7.04
2015	7.00	6.55	6.20	6.02		5.98	4.83
2016	7.91	6.46	6.69	6.43			7.33
2017	10.93						10.19
Average, 2010–16	11.50	10.37	9.96	9.65			10.88

Sources: Authors' calculations, based on data provided by the National Bureau of Statistics of China and published tabulations from the *China Taxation Yearbook*.
a. W&R = wholesale and retail sector.

spending from the equation above suggests that the construction sector's output is also overstated. We therefore adjust the construction sector's GDP using this formula:

$$\Delta \text{ Construction GDP}_t = \frac{\gamma_t}{\theta^I_{con,t}} \cdot \Delta \text{ FCF}_t, \tag{2}$$

where $\theta^I_{con,t}$ denotes FCF per unit of construction GDP and γ_t is the proportion of construction FCF in total FCF.[24] The adjustment in the construction GDP leads to further adjustment in the aggregate GDP and, hence, another round of adjustment in the FCF and construction GDP. The full adjustment that balances the aggregate GDP, construction GDP, and FCF is given by:

$$\text{Adjusted GDP}_t = \text{Official GDP}_t + \frac{1}{1 - \gamma_t/\theta^I_{con,t}} \Delta \text{ Industrial GDP}_t. \tag{3}$$

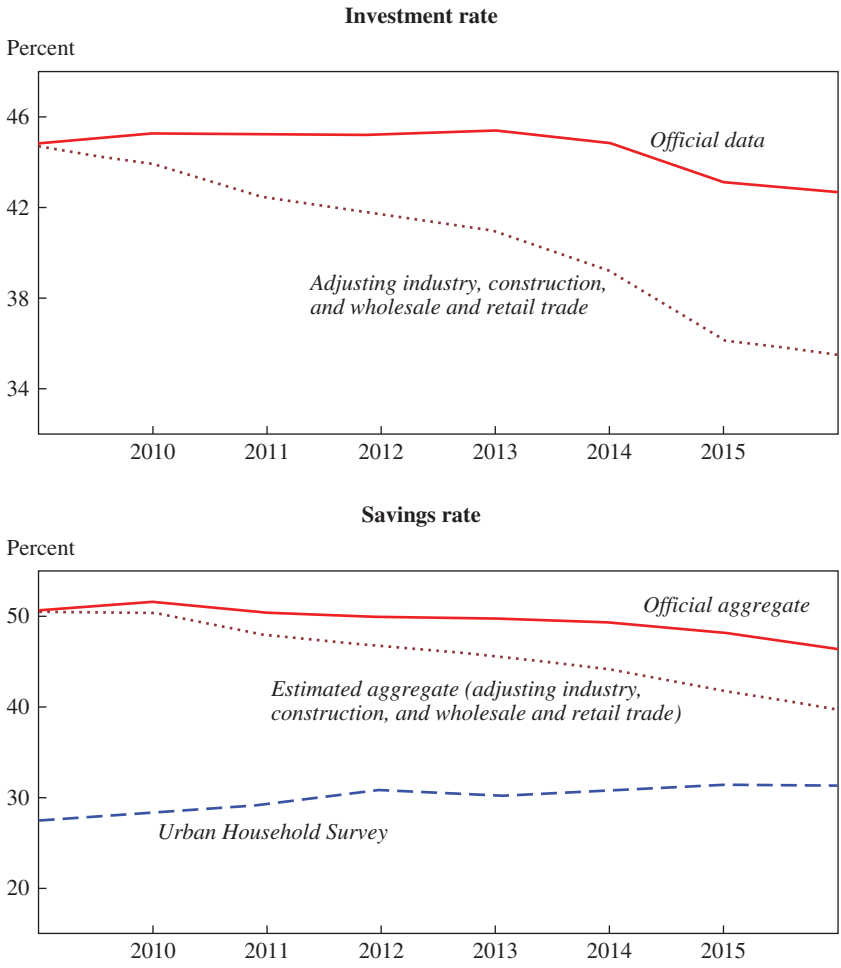
Compared with equation 1, the GDP adjustment in equation 3 is amplified by adjusting construction output. When we also adjust wholesale and retail GDP, Δ Industrial GDP_t in equation 3 should be replaced by Δ Industrial $\text{GDP}_t + \Delta$ WR GDP_t, where WR GDP denotes wholesale and retail GDP.

The results are shown in the top panel of figure 10. As can be seen, our estimate of the investment rate is significantly lower than the official numbers. In 2016, we estimate that the investment rate is about 35.6 percent of GDP—the official number is about 7 percentage points higher. Looking at the change since 2010, our estimate is that the investment rate fell from about 43.9 percent in 2010 to about 35.6 percent in 2016. The official statistic is that the investment rate decreased from about 45.2 percent to about 42.7 percent between these two years. Figure A6 in the online appendix plots the implied construction GDP growth.

The fourth column of table 2 reports the growth rate of nominal GDP after all three adjustments (industrial, wholesale and retail trade, and construction output). With all three adjustments, nominal GDP growth since 2013 has been about half the official growth rate of nominal GDP. During the 2010–16 period, our estimate of GDP growth is about 1.8 percentage points lower than the official growth rate.

24. Note that we do not need to adjust industrial output in a similar fashion. This is because industrial output can be exported, while construction output is for domestic use.

Figure 10. Official and Adjusted Investment Rates, 2009–16[a]



Sources: Authors' calculations; National Bureau of Statistics of China.

a. The solid line in the top panel plots the ratio of official fixed-capital formation (FCF) to official GDP. The dotted line plots the ratio of *adjusted* FCF and *adjusted* GDP (with industry, construction, and wholesale and retail trade GDP adjusted). For comparability with the adjusted FCF afterward, in 2009, the adjusted FCF is calculated as production-side GDP minus other expenditure components. The solid line in the bottom panel is the aggregate savings rate in the national accounts. The dotted line plots the adjusted aggregate savings rate. The dashed line is the average savings rate in the Urban Household Survey.

The bottom panel of figure 10 shows our estimate of the savings rate. Our estimate is that the savings rate fell significantly between 2010 and 2016, from about 50.4 percent to about 39.7 percent of GDP. The official numbers show a much smaller decrease, from about 51.5 percent to about 46.4 percent. Figure 10 also shows that our revised estimate of the savings rate is closer to the savings rate computed from the micro data from the Urban Household Survey. The smaller difference implies a more reasonable saving rate in the nonhousehold sector. Household income accounts for 62 percent of GDP in 2016.[25] To reconcile the official aggregate saving rate of about 46 percent and the household saving rate of about 31 percent, we would need a saving rate of about 70 percent in the corporate and government sectors. If, instead, the aggregate saving rate follows our estimate, the corporate and government sectors would have a saving rate of 54 percent, which is more reasonable than what is implied by the official aggregate saving rate.[26]

II.B. Adjusting Local GDP

A similar procedure can be applied to correct provincial GDP. The published data on revenues from value-added taxes do not break down revenues by province industries. However, value-added tax revenues from industry and wholesale and retail trade account for more than 90 percent of total value-added tax revenues before 2015 (see figure A5 in the online appendix). We use provincial value-added tax revenue growth to proxy the value-added tax revenue growth of industry and wholesale and retail trade in the provinces. The same benchmark adjustment for national GDP can then be used for provincial GDP.[27]

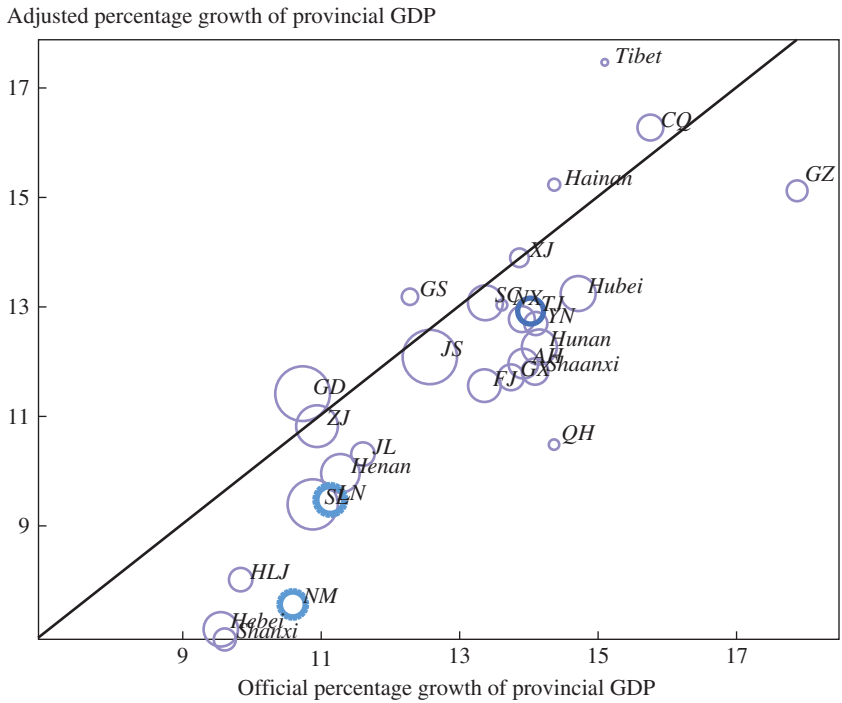
Figure 11 shows a scatterplot of our adjusted growth rate of provincial GDP against the official growth rate of provincial GDP. The majority of the provinces lie below the 45-degree line, indicating that the official growth rate of most provinces exceeds our adjusted estimates. The average

25. The data on household income and GDP are from the Flow of Funds Account in the *China Statistical Yearbook*.

26. The household saving rate may be underestimated in the surveys. Using the household saving rate of 36 percent in the 2016 Flow of Funds Accounts, the implied saving rate in the corporate and government sectors would be 62 percent and 46 percent by the official and our adjusted aggregate saving rate, respectively.

27. We drop Shanghai and Beijing for two reasons. First, these two provinces replaced the business tax with the value-added tax in 2012 and 2013, respectively. The reform had a significant effect on value-added tax revenue in both provincial-level cities, where the service sector is substantially larger than the industrial sector. Second, it is widely acknowledged that the two cities are among the regions with the most reliable GDP data.

Figure 11. Official and Adjusted Provincial GDP Growth, 2009–15[a]



Sources: Authors' calculations, based on data provided by the National Bureau of Statistics of China and published tabulations from the *China Taxation Yearbook.*

a. This is a scatterplot of the *adjusted* (*y* axis) versus official (*x* axis) provincial annualized GDP growth rate between 2009 and 2015. The three highlighted provinces are Liaoning (LN), Inner Mongolia (NM), and Tianjin (TJ), which were involved in GDP scandals. The solid line is a 45-degree line.

difference is about 1.2 percentage points. Guangdong (GD) and Zhejiang (ZJ) , however, are located on the 45-degree line. Among the provinces that are far below the 45-degree line are Liaoning (LN) and Inner Mongolia (NM). Local leaders in the two provinces were recently arrested in corruption crackdowns, and one of the official accusations was that these leaders had overstated local GDP. In addition, after the corruption crackdown, the local statistical bureaus in LN and NM issued new revised estimates of local GDP in 2016 and 2017, respectively.[28] The new numbers are

28. Tianjin (TJ) also acknowledged that the Binhai district overstated its GDP. But the Tianjin municipal government claimed that the district-level GDP overstatement did not affect Tianjin's GDP.

about 22 percent and about 11 percent lower than the official numbers in the previous year.[29] In comparison, our estimates show that the unadjusted official GDP in LN and NM was overstated by about 9 percent and about 15 percent in 2015, respectively. Furthermore, the official adjustment on industrial GDP accounts for about 70 percent of its adjustment on GDP in LN. In the case of NM, the local statistical bureau revised its estimate of total value added of above-scale industrial firms in 2016 downward by about 290 billion yuan, which accounts for the entire downward revision in GDP of NM that year.

Adjusting local FCF is more difficult. Unlike net exports at the national level, which are underpinned by customs data, provincial net outflows of goods and services are not based on any data. Therefore, the adjustment for national FCF cannot be applied to provincial FCF. We can, however, use this equation to indicate provincial FCF:

$$\text{FCF}_{jt} = \sum_i \theta^I_{ijt} Y_{ijt}, \tag{4}$$

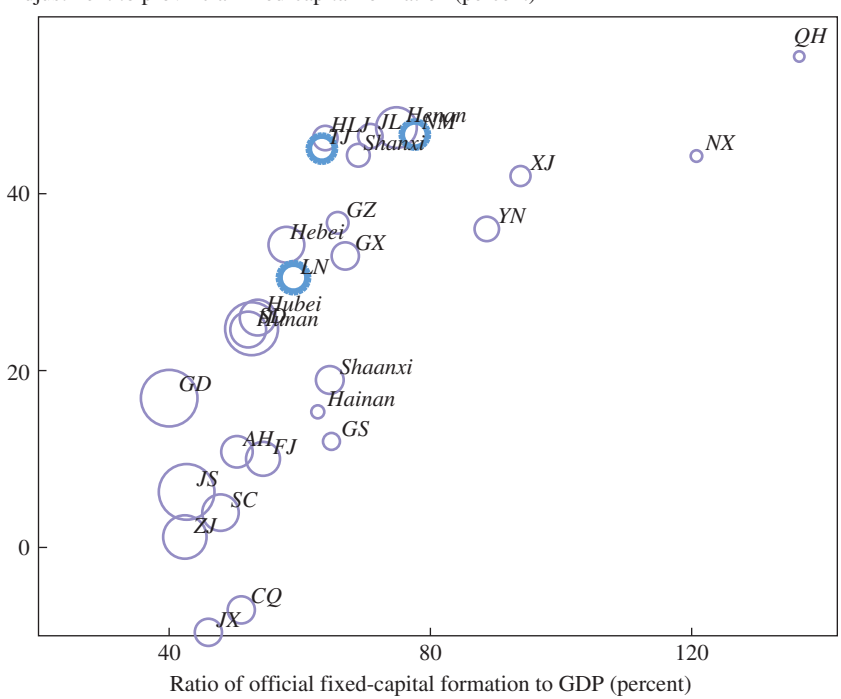
where j is the province index, and θ^I_{ijt} denotes the proportion of Y_{ijt}, sector i's value added in province j, that is converted to fixed capital in province j. Instead of using regional IO tables for θ^I_{ijt}, we assume $\theta^I_{ijt} = \theta^I_{it}$ and rely on the numbers in the national IO table. We then plot adjustment to provincial FCF against the official FCF GDP ratio in 2015 given in figure 12. We find that most provinces overreport FCF, and the extent of overreporting is increasing in the official investment rate. The overreporting of FCF is most severe in provinces such as Qinghai and Henan. The FCF GDP ratio was overstated by more than 50 percentage points in Qinghai. All three provinces discussed above, where local officials "confessed" to manipulating local statistics, are also associated with a severe overstatement of FCF. Their official FCF is about 30 to 50 percent higher than our estimates in 2015.[30]

Figure 13 shows a positive correlation between the extent of overreporting in provincial GDP and that overreporting in provincial FCF (the correlation is about 0.54). Although our estimated provincial GDP

29. We calculated these two numbers from the published tabulations in the *China Statistical Yearbook*.

30. We use the 2014 FCF data for LN because FCF in LN declined by about 30 percent in 2015. Without a big adjustment in GDP, LN's net exports jumped from –104 billion yuan in 2014 to 304 billion yuan in 2015. In other words, before its GDP adjustment in 2016, LN had scaled back its investment in 2015.

Figure 12. Adjustment to Provincial Fixed-Capital Formation in 2015[a]



Sources: Authors' calculations, based on data provided by the National Bureau of Statistics of China and published tabulations from the *Input-Output Tables of China* and the *China Taxation Yearbook.*

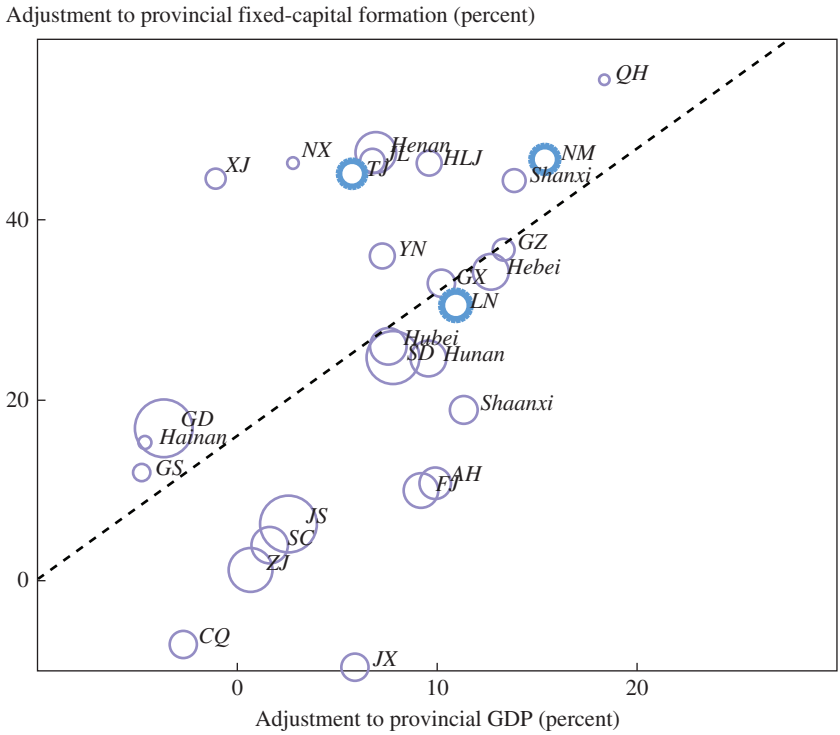
a. This is a scatterplot (FCF = fixed-capital formation) of (official provincial FCF – adjusted provincial FCF) / official provincial FCF (*y* axis) versus official provincial FCF / GDP (*x* axis).

and FCF are correlated by construction, there is no reason that the adjustments to provincial GDP and FCF should be correlated. If measurement errors in provincial GDP and FCF are large and independent, adjustments to the two variables would be uncorrelated. Figure 13 thus provides evidence that local governments overstate both GDP and FCF simultaneously.

II.C. Adjusting National Accounts with Statistical Models

A second approach is to explore the statistical relationship between GDP and a set of economic indicators outside China's national accounts. We first estimate a model using the provincial-level data before 2008, and we then use the estimated model and the indicators to predict provincial and national GDP after 2008. The success of the statistical approach

Figure 13. Adjustment to Provincial GDP and Fixed-Capital Formation in 2015[a]



Sources: Authors' calculations, based on data provided by the National Bureau of Statistics of China and published tabulations from the *Input-Output Tables of China* and the *China Taxation Yearbook*.

a. This is a scatterplot (FCF = fixed-capital formation) of (official provincial FCF – adjusted provincial FCF) / official provincial FCF (*y* axis) versus (official provincial GDP – adjusted provincial GDP) / official provincial GDP (*x* axis). The dashed line is a linear regression line.

depends on three conditions. First, the indicators are informative about local economies and are unlikely to be manipulated. Second, local GDP growth data before 2008 are more reliable than afterward. And third, the statistical model is flexible enough to capture the rich heterogeneity across Chinese provinces. We discuss these three conditions in order.

Our indicators include satellite night lights, national tax revenue, exports and imports, electricity consumption, railway cargo volume, and new bank loans.[31] National tax revenue is collected by local governments but is

31. Using bank loans (not new bank loans) delivers similar results.

directly paid to the central government. Cheating on national tax revenue would incur fiscal losses and, hence, is unlikely to happen. Exports and imports are from customs data, which are hard to manipulate due to the symmetry of the customs data from China's trading partners. Electricity consumption, railway cargo volume, and new bank loans are from the so-called Keqiang Index, which Li Keqiang, China's current premier, used to monitor local economic performance when he was the Communist Party secretary of Liaoning Province.

We understand that overreporting of local GDP started in the late 1990s. So local GDP growth data before 2008 cannot be entirely reliable. Yet we also understand that GDP overreporting has become more severe since 2008. What we identify from the next exercise is the difference in the degree of GDP overstatement between the period before 2008 and the post-2008 period. Consequently, when we rely on local GDP growth data before 2008, which is per se likely to be overstated, to estimate the subsequent growth, our adjustment needs to be a lower bound. The true GDP growth might be even lower than our estimates for the post-2008 period.[32]

In terms of the statistical model, we use the method developed by Su, Shi, and Phillips (2016) to control for hidden economic structural heterogeneities across regions. Consider this linear model:

$$y_{i,t} = \beta_i' X_{i,t} + v_i + \varepsilon_{i,t},$$

where y_{it} is log GDP of province i at year t, $X_{i,t}$ is a $p \times 1$ vector of logarithm of the indicators, β_i is a $p \times 1$ coefficient vector, v_i captures provincial fixed effects and $\varepsilon_{i,t}$ is the independent and identically distributed error term with mean zero. In the special case where $\beta_i = \beta$ the model reduces to the standard fixed effects regression. The more general model can capture heterogeneous economic structures across regions. Intuitively, β for the regions where the local economy relies heavily on resources might be very different from the others. Specifically, we assume β_i to be group-specific—that is, $\beta_i = \beta_k$ for all i in group k, where $i \in \{1, 2, \ldots, N\}$, $k \in \{1, 2, \ldots, K\}$, and $K \leq N$. Instead of grouping provinces by geographical or economic

32. Our approach differs from that of Fernald, Hsu, and Spiegel (2015) and of Clark, Pinkovskiy, and Sala-i-Martin (2017), who use data on exports to China from its trading partners and night lights as independent measures of China's economic activities. We instead train our statistical model by provincial industrial GDP data prior to 2008, when the overstatement of industrial GDP was much less evident compared with the post-2008 period. Also see Hu and Yao (2019), who use night-time lights data to estimate GDP in a number of countries.

characteristics, we implement the classifier least absolute shrinkage and selection operator (C-Lasso) method used by Su, Shi, and Phillips (2016). The method provides statistical inference for membership identification, which is totally data driven. We do not have to rely on prior knowledge about the number of groups or the number of provinces within each group. With the groups identified from C-Lasso, we can use the fixed-effects model to estimate the group-specific coefficients.

It is worth mentioning the rapid expansion of China's service sector. According to the national accounts data, service accounted for about 43 percent of GDP in 2007, and the share increased to about 52 percent in 2017. This is important because some of our indicators, such as electricity consumption and railway cargo volume, might be more relevant for industrial production than for service production. If $y_{i,t}$ includes service output, the ongoing structural transformation would imply time-varying β_i and, hence, invalidate our model. To address the concern, we use provincial industrial GDP as $y_{i,t}$ in the benchmark and then use provincial GDP as a robustness check. There are two reasons why we prefer provincial industrial GDP. First, the stationarity of β_i is more defensible for industrial GDP alone. Second, we have shown the evidence that GDP overstatement is larger in the industrial sector.

Our sample consists of annual observations from 30 Chinese provinces (excluding Tibet) between 2000 and 2017. GDP, electricity consumption, exports and imports, railway cargo volume, and new bank loans are all from the NBS;[33] national tax revenue is from the *China Taxation Yearbook*; and we use the Defense Meteorological Satellite Program–Operational Linescan System night-time lights data from the U.S. National Oceanic and Atmospheric Administration.[34] The time series are shorter for some variables. Satellite night-lights data end in 2013. National tax revenue data end in 2015 because the reform to replace the business tax with the value-added tax made national tax revenue not comparable before and after 2016.

Two remarks are in order. First, night-light data, electricity consumption, and railway freight are all in real terms. As a robustness check, we use

33. All the indicators were downloaded from the NBS website. Exports and imports (by place of destination or origin in China) were priced in dollars and were converted into yuan by annual averages of the exchange rate. A new bank loan is the annual difference of the outstanding bank loan in December.

34. The night-light data are not comparable before and after 2010 due to the satellite change. We use the average of the light growth in 2009 and 2011 to proxy the 2010 light growth for out-of-sample predictions.

GDP deflators to convert GDP, national tax revenue, exports and imports, and bank loans into real terms in the regressions (see also Clark, Pinkovskiy, and Sala-i-Martin 2017).[35] The estimated GDP will be converted back into nominal terms. The online technical appendix reports the results with price adjustments. The differences are small. Second, we can use more data in the earlier period to estimate the model, with the caveat that the estimated model might be less applicable to recent years due to structural changes. In the online appendix, we estimate the model with data from between 1995 (the year after implementation of the tax-sharing reform) and 2007. The main results are very similar.

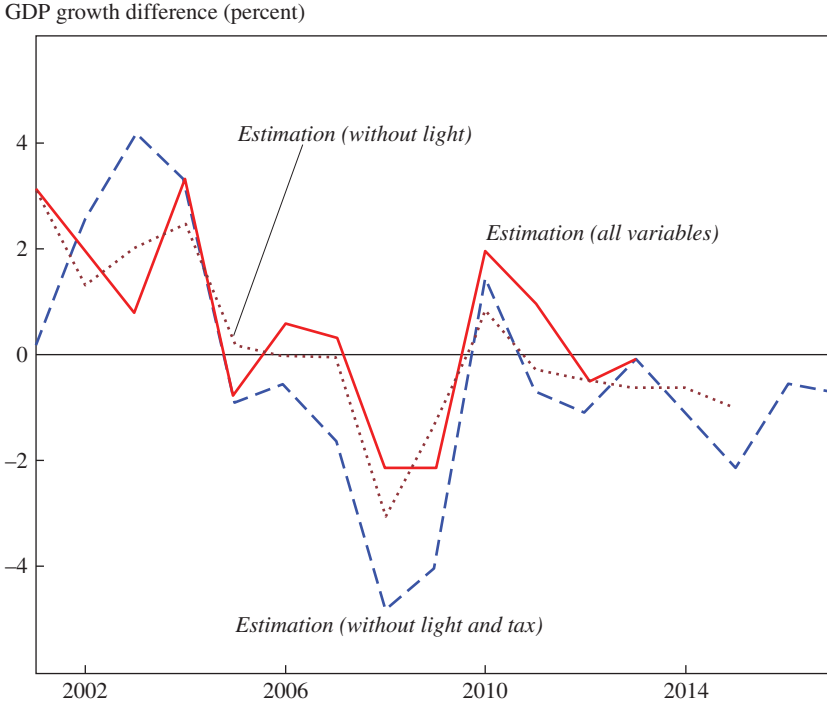
We first apply C-Lasso to the 2000–2007 data for model selection. K-fold cross validation, EBIC (Extended Bayesian Information Criterion), and data-driven penalty with heteroscedasticity (Belloni, Chernozhukov, and Hansen 2014; Belloni and others 2012, 2016) suggest keeping all the indicators except for new bank loans. Besides the statistical evidence, there is also an economic reason for us to drop bank loans. The "fiscal stimulus" launched by the Chinese government in late 2008 relaxed the borrowing constraint on local governments and led to a debt explosion afterward (Bai, Hsieh, and Song 2016). Much of the funds raised by local government financing vehicles are believed to finance infrastructure investment rather than production. This implies a structural change in the way that new bank loans contribute to GDP.

Our estimation is done in three steps. First, using the sample from before 2008, we run the C-Lasso estimation to classify provinces into different groups. Second, we estimate group-specific coefficients with post-Lasso ordinary least squares regressions. Finally, the estimated $\hat{\beta}_k$ and the same set of indictors are used to estimate provincial secondary industry value added throughout the whole sample period. Assuming that provincial agriculture, construction, and service GDP are reliable, we can estimate provincial GDP, which will be added together to obtain the aggregate GDP. Note that the estimated industrial value added after 2008 is out-of-sample prediction, while the estimation before 2008 is in-sample prediction.

When we use provincial industrial GDP, the C-Lasso procedure does not find statistical evidence for grouping, suggesting that the relationship between industrial GDP and these indicators is similar across provinces. As we show below, the result would be different if we replace provincial industrial GDP with provincial GDP. Because the satellite night-light

35. GDP deflators are inferred from the official real and nominal GDP growth.

Figure 14. The Gap between Estimated and Official Nominal GDP Growth, 2001–17[a]



Source: Authors' calculations.

a. This figure is based on the statistical model for provincial industrial GDP. It plots the difference between adjusted and official nominal GDP growth rates. "Estimation (all variables)" uses all variables. "Estimation (without light)" uses all variables except for night lights. "Estimation (without light and tax)" uses all variables except for night lights and national tax revenues. Adjusted GDP growth rates are in-sample estimates before 2009 and out-of-sample predictions after 2009.

data are not available after 2013, it can only be used for the out-of-sample prediction between 2008 and 2013. We rerun the C-Lasso and post-Lasso ordinary least squares regressions without night lights. The estimated model can make out-of-sample predictions for the post-2013 period.[36]

The out-of-sample predictions are shown in figure 14 and table 2.[37] Although the in-sample predictions are close to the official numbers, in

36. The tables with the regression coefficients are in the online appendix.

37. We aggregate provincial GDP growth by our estimated provincial GDP, which is based on the estimated provincial GDP growth and uses 2009 official provincial GDP as the benchmark.

recent years the out-of-sample predictions have been more volatile and lower than the official numbers. The estimated GDP growth is about 0.5 to 2.2 percentage points lower than the official GDP growth during 2014 and 2016 (see the dotted-and-dashed line in figure 14 and see the seventh and eighth columns of table 2).

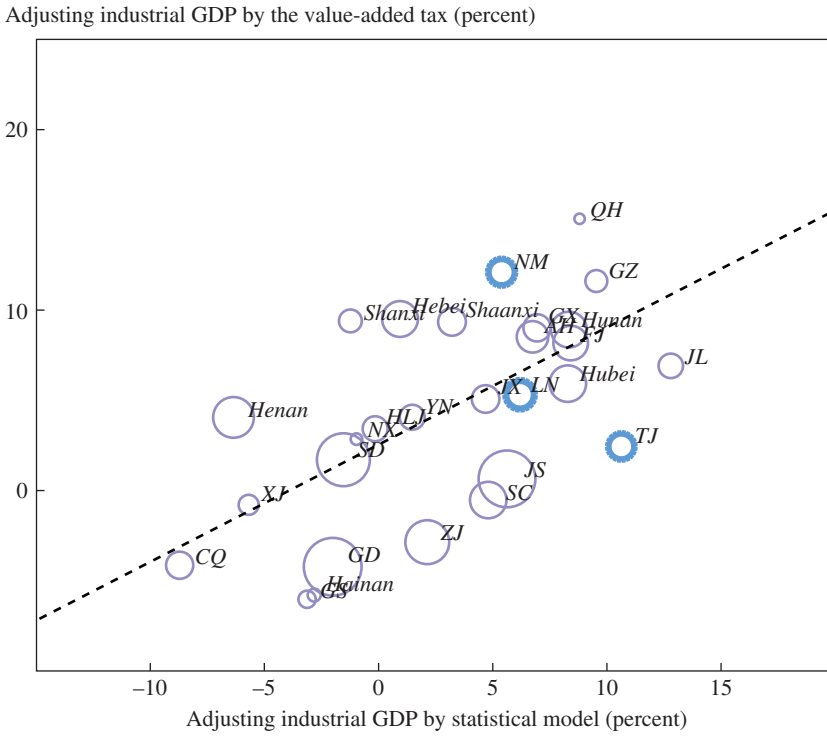
We note that although our two approaches are fundamentally different, they yield similar results in terms of the magnitude of overstatement of GDP. Table 2 shows that nominal GDP growth was overstated after 2010 and more so after 2013, and the magnitude of the overstatement after 2013 was about 1 to 2 percentage points.

One may wonder to what extent the tax revenue data used by the two approaches can explain their similar results on the recent overreporting of GDP. The first thing to notice is that tax revenue data are very different in the two approaches. National tax data include many taxes other than the value-added tax (for example, all consumption taxes and part of corporation income taxes), and only a fraction of value-added tax belongs to national taxes. Figure 15 plots the extent of GDP overstatement across provinces estimated by the first approach and the second approach with national tax revenue. Because the second approach only adjusts industrial GDP, we use the first approach that adjusts industrial GDP only to make the two approaches more comparable. The correlation is about 0.64. In other words, the different methods using different data sources deliver positively correlated estimates of provincial GDP overstatement.

We also run the regressions without national tax revenue. An advantage of dropping national tax revenue is to extend the estimation to the years after the completion of the reform to replace the business tax. The results are shown in figure 14 and in the last column of table 2. The overstatement of GDP growth after 2013 appears to be a robust finding, though its magnitude does depend on estimation method and variable selection.

We next replace provincial industrial GDP with provincial GDP for a robustness check. We drop both railway cargo volume and new bank loans, as suggested by Lasso. Given the huge disparity in GDP composition across provinces, not surprisingly, C-Lasso identifies two groups, with 16 provinces in group 1 and 14 provinces in group 2. See online appendix II for the detailed grouping results. Interestingly, Beijing, Shanghai, and Hainan, the three provinces with the highest service GDP share, are all in group 1. The fixed-effects regression results for each group are reported in the online appendix. Coefficients are indeed quite different across groups. We then run C-Lasso without night-light data, which also identifies two groups, with 11 and 19 provinces in groups 1 and 2.

Figure 15. Adjustment to Provincial GDP in 2015[a]



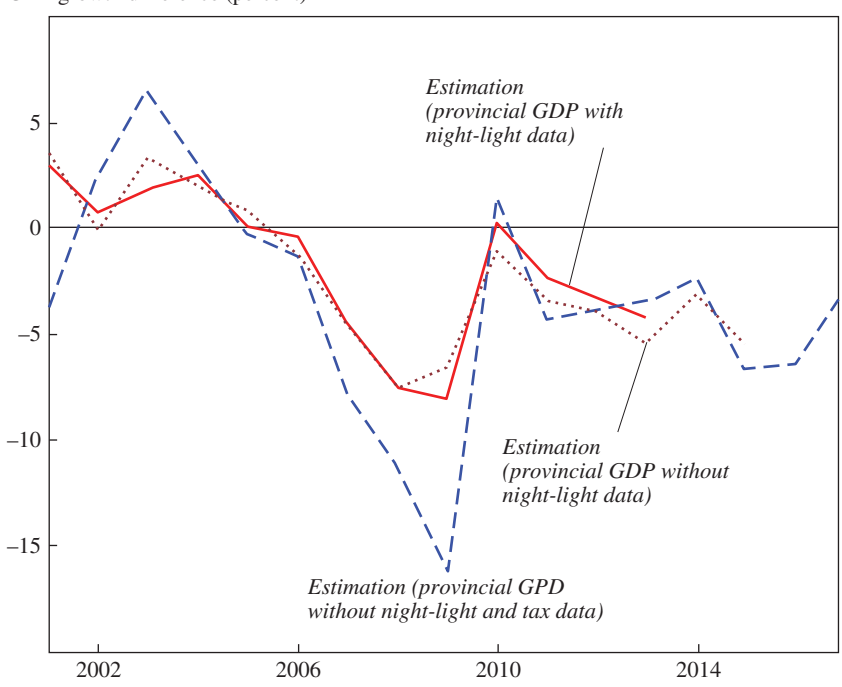
Source: Authors' calculations.

a. This figure is a scatterplot of (official provincial industrial GDP – estimated provincial industrial GDP by value-added tax) / official provincial GDP (*y* axis) versus (official provincial industrial GDP – estimated provincial industrial GDP by the statistical model with national tax revenue) / official provincial GDP (*x* axis). The dashed line is a linear regression line.

Online appendix II shows that 10 out of 11 provinces in group 1 are in group 1 identified by C-Lasso with night-light data. Again, Beijing, Shanghai, and Hainan are all in group 1.

Figure 16 compares the GDP growth rates from the official data, our estimates using provincial GDP with night-light data, provincial GDP without night-light data, and provincial GDP without night-light or tax data. Estimating provincial GDP directly implies a much bigger GDP overstatement. The difference between official GDP growth and our estimate was more than 5 percentage points in 2015. As discussed above, the caveat is the misspecification of the model that fails to capture how the rise of the service sector affects GDP growth.

Figure 16. Gap between Estimated and Official Nominal GDP Growth, 2001–17[a]



Source: Authors' calculations.

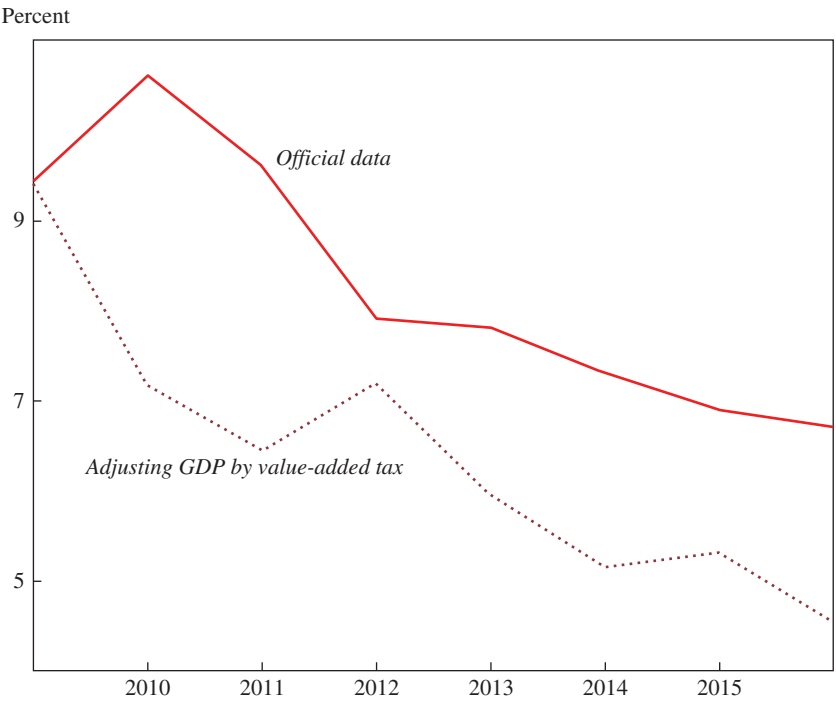
a. This figure is based on the statistical model for provincial GDP. See the note for figure 14. The only difference is that the estimation here is based on the statistical model for provincial GDP instead of industrial GDP in the province, as in figure 14.

III. Implications of Our Revised Estimates of China's National Accounts

To summarize the three main implications of our results: First, nominal GDP growth after 2010 and particularly after 2013 is lower than suggested by the official statistics. Second, the savings rate declined by about 11 percentage points between 2010 and 2016. The official statistics suggest the savings rate only declined by about 5 percentage points between these two years. Third, our statistics suggest that the investment rate fell about 8 percent of GDP between 2010 and 2016. Official statistics suggest that the investment rate fell 3 percent over this period.

We note that we do not have independent information on GDP deflators, so our statement is only about nominal GDP growth. The literature has

Figure 17. Real GDP Growth, 2009–16[a]



Sources: Authors' calculations; National Bureau of Statistics of China.

a. The adjusted real GDP growth is the difference between the adjusted nominal GDP growth rates (from the adjusted growth in the fourth column of table 2) and the growth rate of the official GDP deflator.

questioned the reliability of China's official price indexes, but we do not have independent information on the deflators.[38] Keeping in mind the caveats, we think it is useful to convert nominal output and input into real terms using the official GDP deflators and investment goods price index.

For real GDP growth, we calculate real GDP in the industrial, construction, and wholesale and retail trade sectors using our estimated nominal GDP (the first approach) and using the official GDP deflators for the three sectors. Adding adjusted real GDP in the three sectors to real GDP in the other sectors gives our adjusted real GDP, which is shown in figure 17. On average, the annual real GDP growth was overstated by about 2 percentage

38. See, for example, Brandt and Zhu (2010) and Nakamura, Steinsson, and Liu (2016).

points between 2010 and 2016. The official real GDP is about 13 percent above our estimate in 2016.

We now discuss the implications of our findings for capital returns, total factor productivity (TFP) growth, and the debt-to-GDP ratio. We begin with the return to capital. We use this equation to estimate returns to capital:

$$r(t) = i(t) - \hat{P}_Y(t) = \frac{\alpha(t)}{P_K(t)K(t)/P_Y(t)Y(t)} + \left[\hat{P}_K(t) - \hat{P}_Y(t)\right] - \delta(t),$$

where r denotes real returns to capital, i denotes nominal returns to capital, $\hat{P}_Y$ denotes the growth rate of output price, $\hat{P}_K$ denotes the growth rate of the capital goods price, α denotes the share of capital income in output, $P_K K/P_Y Y$ denotes the nominal capital-output ratio, and δ is the depreciation rate.

The results are plotted in figure 18.[39] The solid line uses the official data and replicates the earlier estimates made by Bai, Hsieh, and Yingyi Qian (2006) and the more recent ones by Bai and Qiong Zhang (2015). Recall that our adjustment of production GDP also lowers investment, which increases the ratio of output to capital. In either official or adjusted data, the dramatic decline in aggregate returns to capital in the post-2007 period turns out to be a robust phenomenon.

To estimate TFP, we assume this aggregate production function:

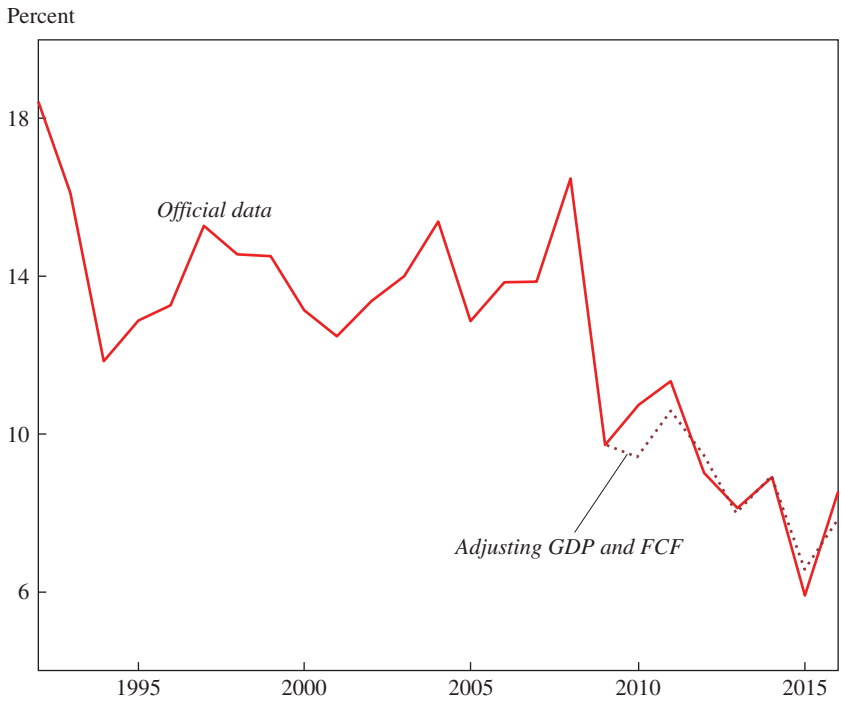
$$Y(t) = A(t)[K(t)]^{\alpha}[h(t)L(t)]^{1-\alpha},$$

where Y is real GDP, A is aggregate TFP, K is real capital, h is human capital per worker, and L is the number of workers.[40] The results are plotted in figure 19. The aggregate TFP growth rates given by our estimates appear to be more volatile than those given by official data. Yet it remains obvious that China's aggregate TFP growth slowed down substantially after 2007.

39. We discuss the details of the data used to estimate the return to capital in the online appendix.

40. We set $\alpha = 0.5$ (the results are similar if we use time-varying α calibrated in the online appendix for estimating returns to capital). We assume that $h = exp(s \cdot E)$, where E is the year of schooling and s is the return to schooling. The average year of schooling for workers in 2000 and 2005 is from the 2000 census and 2005 one-percent population survey data. We obtain the numbers between 2001 and 2004 by linear interpolation. For 2006 to 2016, we use the numbers from the labor force survey in the *China Population & Employment Statistics Yearbook*. For 1990 to 1999, we assume the annual growth of E to be its average growth from 2000 to 2005. We then use the 1 percent population survey data for 2005 to estimate returns to education by the Mincer earnings regression, which gives $s = 0.126$.

Figure 18. The Aggregate Return to Capital, 1994–2015[a]



Sources: Authors' calculations, based on the data provided by the National Bureau of Statistics of China and published tabulations from the *Statistical Yearbook of Chinese Investment in Fixed Assets, The Gross Domestic Product of China:* 1952–1995, and Hsueh and Li (1999).

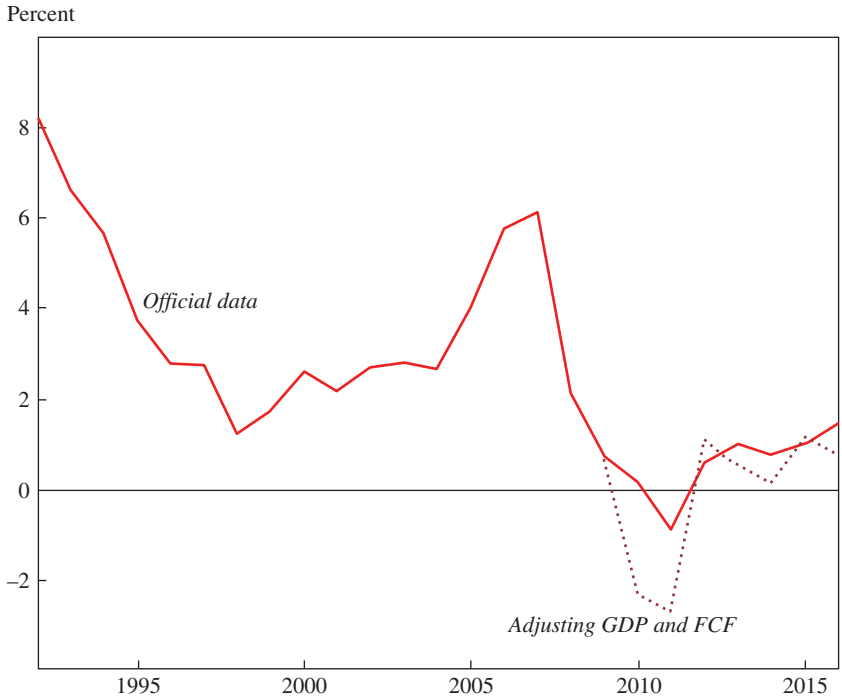
a. "Official data" indicate the return to capital computed from official GDP and fixed-capital formation (FCF). "Adjusting GDP and FCF" is the return to capital calculated from adjusted FCF and GDP (using the adjusted growth rate given in column 4 of table 2).

Finally, figure 20 shows the debt-to-GDP ratio with our revised estimate of nominal GDP. The estimation of debt follows that used by Song and Xiong (2018). The bottom line is that our revised numbers suggest that the debt-to-GDP ratio has increased by more than suggested by the official numbers. Our estimate of the debt-to-GDP ratio in 2016 is about 2.4—the official number is 2.1.

IV. Conclusion

A key institutional fact about China is that many administrative functions are controlled by powerful local governments. In another recent paper (Bai, Hsieh, and Song 2019), we argue that local governments have used this

Figure 19. Aggregate Growth of Total Factor Productivity, 1993–2015[a]



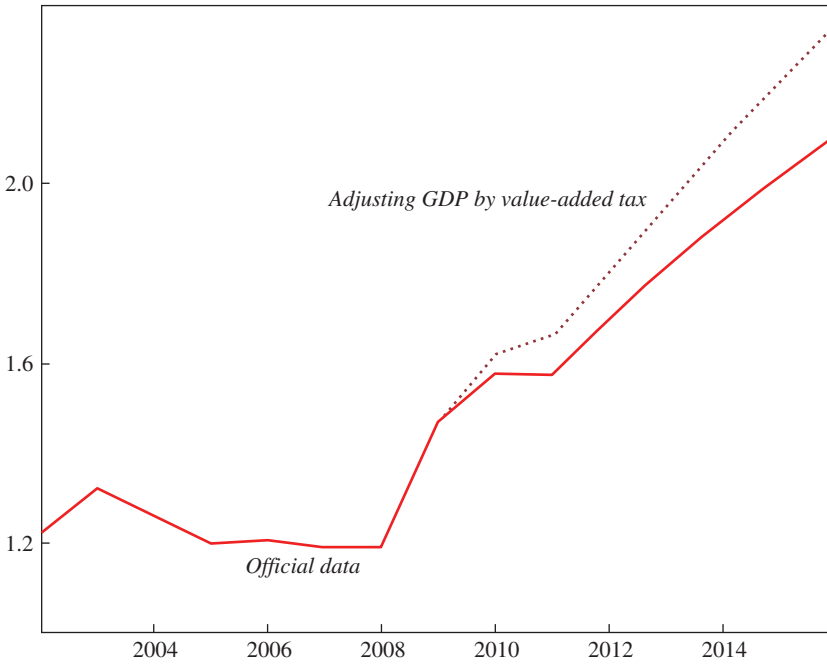
Sources: Authors' calculations, based on micro data from China's population censuses, data provided by the National Bureau of Statistics of China, and published tabulations from the *Statistical Yearbook of Chinese Investment in Fixed Assets* and the *China Population & Employment Statistics Yearbook.*

a. "Official data" indicate the return to capital computed from official GDP and fixed-capital formation (FCF). "Adjusting GDP and FCF" is the return to capital calculated from adjusted FCF and GDP (using the adjusted growth rate given in the fourth column of table 2).

power to support a large number of private businesses. The question in this current paper is what local governments choose to do with their power over local statistics.

We document that local governments have chosen to use this power to inflate local statistics on GDP, particularly by overstating industrial output and investment, particularly after the mid-2000s. As evidence, we show that the sum of local GDP has exceeded aggregate GDP since 2003. One possible explanation why they do this is the introduction of local economic performance in the evaluation of local officials by the Chinese Communist Party's Organization Department in the late 1990s. The official documentation of this policy change states that local officials will be evaluated based on "the speed, efficiency and potentials of economic

Figure 20. The Debt-to-GDP Ratio, 2002–16[a]



Sources: Song and Xiong (2018); National Bureau of Statistics of China.

a. The data on debt are from Song and Xiong (2018). "Official data" use official GDP in the denominator. "Adjusting GDP by value-added tax" uses adjusted GDP (using the adjusted growth given in the fourth column of table 2) in the denominator.

development, the growth of fiscal revenue, the improvement of people's living standards."[41] The revision intensified economic competition between local governments, and it seems likely that many local governments resorted to inflating local GDP numbers. Xiong (2018) provides a theoretical framework where competition between local governments results in the overstatement of both GDP and investment. And Changjiang Lyu and others (2018) present evidence that regional growth targets can be achieved by fabricating data.

The possibility that local governments misreport local GDP is well known, and the central government's National Bureau of Statistics adjusts the numbers reported by local governments. Before 2003, the NBS adjusted local GDP upward, but after 2003, it adjusted local GDP downward.

41. See Central Organization Department (1998).

However, our estimates suggest that the extent to which local governments exaggerate local GDP accelerated after 2008, but the magnitude of the adjustment by the NBS did not change in tandem. As a consequence, our best estimate is that the true growth rate of GDP is probably overstated by almost 2 percentage points from 2010 to 2016.

A final question is what tools and what incentives does the NBS have to report accurate statistics. We document that much of the underlying data behind the national accounts is out of the hands of the NBS. Furthermore, the question is what incentives does the NBS have to resist local officials who misreport data. Interestingly, although the NBS adjusts local statistics downward, it does not report the adjusted local statistics, perhaps out of a desire to not confront powerful local leaders. Given the NBS's weak position and the strong position of local leaders in the Chinese political system, it is not surprising that statistical data are potentially biased.

ACKNOWLEDGMENTS The authors thank David Dollar, Janice Eberly, Zhentao Shi, and Wei Xiong for helpful comments. Zheng Song acknowledges financial support from the Research Grant Council (Hong Kong) on "Re-Measuring China's Regional Investment," Project 14502718.

References

Bai, Chong-En, Chang-Tai Hsieh, and Yingyi Qian. 2006. "The Return to Capital in China." *Brookings Papers on Economic Activity*, Fall: 61–101.

Bai, Chong-En, Chang-Tai Hsieh, and Zheng Song. 2016. "The Long Shadow of China's Fiscal Expansion." *Brookings Papers on Economic Activity*, Fall: 129–81.

———. 2019. "Special Deals with Chinese Characteristics." *NBER Macroeconomics Annual*, forthcoming.

Bai, Chong-En, and Qiong Zhang. 2015. "The Return on Capital in China and Its Determinants." *China Economist* 10, no. 3: 20–37.

Bai, Chong-En, Xilu Chen, Zheng Song, and Xin Wang. 2019. "The Rise of China's Service Sector." Working paper.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Chris Hansen. 2012. "Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain." *Econometrica* 80, no. 6: 2369–2429.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *Review of Economic Studies* 81, no. 2: 608–50.

Belloni, Alexandre, Victor Chernozhukov, Christian Hansen, and Damian Kozbur. 2016. "Inference in High-Dimensional Panel Models with an Application to Gun Control." *Journal of Business & Economic Statistics* 34, no. 4: 590–605.

Brandt, Loren, and Xiaodong Zhu. 2010. "Accounting for China's Growth." Working Paper 4764. Bonn: Institute for the Study of Labor. http://ftp.iza.org/dp4764.pdf.

Cai, Jing, and Ann Harrison. 2018. "Industrial Policy in China: Some Unintended Consequences?" *Industrial and Labor Relations Review*, Forthcoming.

Central Organization Department. 1998. "Provisional Regulations on Evaluating Party and Government Cadres." Document 6. Chinese Communist Party, Beijing.

Chen, Zhao, Xian Jiang, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu. 2019. "Tax Policy and Lumpy Investment Behavior: Evidence from China VAT Reform." Working paper. https://economics.stanford.edu/sites/g/files/sbiybj9386/f/abstract_stanford.pdf.

Clark, Hunter, Maxim Pinkovskiy, and Xavier Sala-i-Martin. 2017. "China's GDP Growth May Be Understated." Working Paper 23323. Cambridge, Mass.: National Bureau of Economic Research. https://www.nber.org/papers/w23323.

Fernald, John G., Eric Hsu, and Mark M. Spiegel. 2015. "Is China Fudging Its Figures? Evidence from Trading Partner Data." Working Paper 2015-12. San Francisco: Federal Reserve Bank of San Francisco. https://www.frbsf.org/economic-research/files/wp2015-12.pdf.

Gao, Jie. 2016. "Bypass the Lying Mouths: How Does the CCP Tackle Information Distortion at Local Levels?" *China Quarterly* 228: 950–69.

Holz, Carsten A. 2013. "Chinese Statistics: Output Data." Working Paper 473. Stanford, Calif.: Stanford Center for International Development. https://globalpoverty.stanford.edu/sites/default/files/publications/473wp.pdf.

———. 2014. "The Quality of China's GDP Statistics." *China Economic Review* 30: 309–38.

———. 2018. "China's Investment Rate: Characteristics and Implications." Working paper. http://carstenholz.people.ust.hk/CarstenHolz-ChinaInvestmentRate-13May2018.pdf.

Hsueh, Tien-Tung, and Qiang Li. 1999. "China's National Income, 1952–1995." Boulder, Colo.: Westview Press.

Hu, Yinghao, and Jiaoxong Yao. 2019. "Illuminating Economic Growth." IMF Working Paper 19/77. Washington: International Monetary Fund. https://www.imf.org/en/Publications/WP/Issues/2019/04/09/Illuminating-Economic-Growth-46670.

Lardy, Nicholas R. 2014. *Markets over Mao: The Rise of Private Business in China*. Washington: Peterson Institute for International Economics.

Lei, Yu-Hsiang. 2017. "Quid Pro Quo? Government-Firm Relationships in China." Working paper. https://drive.google.com/file/d/181FwmoICvLQyWO5HOJhxu8_HcMLrLnHn/view.

Liu, Fang, Jun Zhang, and Tian Zhu. 2016. "How Much Can We Trust China's Investment Statistics?" *Journal of Chinese Economic and Business Studies* 14, no. 3: 215–28.

Lyu, Changjiang, Kemin Wang, Frank Zhang, and Xin Zhang. 2018. "GDP Management to Meet or Beat Growth Targets." *Journal of Accounting and Economics* 66, no. 1: 318–38.

Ma, Ben, Guojun Song, Lei Zhang, and David A. Sonnenfeld. 2014. "Explaining Sectoral Discrepancies between National and Provincial Statistics in China." *China Economic Review* 30: 353–69.

Ministry of Finance and the State Administration of Taxation. 2008a. "Detailed Rules for the Implementation of the Interim Regulations of the Value Added Tax." Document 50. Ministry of Finance and State Administration of Taxation, Beijing.

———. 2008b. "Notice of the Ministry of Finance and the State Administration of Taxation on Several Issues Concerning the National Implementation of Value Added Tax Reform." Document 170. Ministry of Finance and State Administration of Taxation, Beijing.

———. 2011. "Notice of the Ministry of Finance and the State Administration of Taxation on the Pilot Work of Levying the Value-Added Tax in Lieu of the Business Tax in the Transportation Industry and Some Modern Service Industries in Shanghai." Document 111. Ministry of Finance and State Administration of Taxation, Beijing.

———. 2013. "Notice of the Ministry of Finance and the State Administration of Taxation on Temporarily Exempting Certain Small- and Micro-Sized Enterprises

from the Value-Added Tax and the Business Tax." Document 52. Ministry of Finance and State Administration of Taxation, Beijing.

———. 2014. "Announcement of the State Administration of Taxation on Relevant Issues Concerning Exempting Small and Micro-Sized Enterprises from the Value-Added Tax and the Business Tax." Document 57. Ministry of Finance and State Administration of Taxation, Beijing.

Ministry of Natural Resources. Various years. *China Land and Resources Statistical Yearbook*. Beijing: Ministry of Natural Resources.

Nakamura, Emi, Jón Steinsson, and Miao Liu. 2016. "Are Chinese Growth and Inflation Too Smooth? Evidence from Engel Curves." *American Economic Journal: Macroeconomics* 8, no. 3: 113–44.

National Bureau of Statistics. 2006. *Input–Output Tables of China 2002*. Beijing: China Statistics Press.

———. 2009. *Input–Output Tables of China 2007*. Beijing: China Statistics Press.

———. 2015. *Input–Output Tables of China 2012*. Beijing: China Statistics Press.

———. Various years. *China Population & Employment Statistics Yearbook.* Beijing: China Statistics Press.

———. Various years. *Statistical Yearbook of Chinese Investments in Fixed Assets*. Beijing: China Statistics Press.

Orlik, Tom. 2014. "Reform at China's National Bureau of Statistics under Ma Jiantang, 2008–2013." *China Economic Review* 30: 304–8.

Song, Zheng, and Wei Xiong. 2018. "Risks in China's Financial System." *Annual Review of Financial Economics* 10: 261–86.

Su, Liangjun, Zhentao Shi, and Peter C. B. Phillips. 2016. "Identifying Latent Structures in Panel Data." *Econometrica* 84, no. 6: 2215–64.

Xiong, Wei. 2018. "The Mandarin Model of Growth." Working Paper 25296. Cambridge, Mass.: National Bureau of Economic Research. https://www.nber.org/papers/w25296.pdf.

Xu, Xianchun. 2004. "China's Gross Domestic Product Estimation." *China Economic Review* 15, no. 3: 302–22.

———. 2014. "Accurate Understanding of China's Income, Consumption and Investment." *Social Sciences in China* 35, no. 1: 21–43.

Zhang, Lei, Yuyu Chen, and Zongyan He. 2018. "The Effect of Investment Tax Incentives: Evidence from China's Value-Added Tax Reform." *International Tax and Public Finance* 25, no. 4: 913–45.

Comments and Discussion

COMMENT BY

DAVID DOLLAR This is an important paper. Every quarter, China's National Bureau of Statistics (NBS) reports data on GDP and its components, an event that is eagerly watched and that can move markets if the results are far from expectations. Buried toward the end of the Western press reports is usually a disclaimer. For example, the *New York Times* report on China's 2019 second-quarter growth stated that "economists widely doubt the veracity of the overall Chinese growth figure, which shows far more stability than comparable numbers from the United States and elsewhere" (Bradsher 2019). This paper aims to understand the main source of errors and to provide estimates of nominal GDP based on two different approaches—estimates that can be compared with official data.

This paper by Wei Chen and colleagues starts by examining how China's GDP is compiled. The NBS has branches all over the country, but these offices are integrated into local governments and have incentives to work with them. It has been known for some years that provincial GDPs, prepared by these offices, sum to significantly more than the national GDP reported by the central NBS. Clearly, the central NBS does not take the provincial estimates at face value and marks them down. This paper adds several pieces of useful information that help explain this process. First, overreporting of provincial GDP did not exist before about 2004; and second, it is almost wholly accounted for by overreporting of industry value added on the production side and by gross fixed-capital formation on the expenditure side. The authors tell a plausible story of local government officials since the mid-2000s being evaluated primarily based on implementation of big infrastructure projects and GDP growth, creating incentives for overreporting industrial value added.

The paper tries to come up with improved estimates based on two different methodologies. The first takes advantage of data on collection of the value-added tax (VAT). China has long had a VAT on goods, set at 17 percent. This is collected by finance bureaus at the local level, but they have no incentive to overreport it, given that most of it must be turned over to the central authorities. Starting in about 2008, the paper documents a tendency for VAT collected from industry to grow less rapidly than reported industrial value added. The authors use the VAT growth as a proxy for industrial value-added growth. This is the key input into their first alternative estimate. With this change, nominal GDP growth is reduced for the 2010–16 period from an official estimate of 11.50 percent a year to 10.37 percent a year (see Chen and colleagues' table 2, first and second columns). A similar adjustment to construction and wholesale and retail trade reduces estimated growth further, to 9.65 percent per year (the table's fourth column). Other services were taxed differently and in ways that changed over time, but the authors make a plausible case that estimates of service sector GDP were unbiased.

A potential problem with this first approach to adjusting GDP is that there were some changes in VAT rates over this period. The authors note that actual collection of VAT revenue increased from 10.4 percent of industrial value added in 2001 to 12.9 percent in 2007. The authors attribute this to improved tax collection. But it also likely reflects the fact that in practice, the tax rate was not always 17 percent for goods and that there were significant exemptions. In particular, China has long had VAT exemptions or rebates for exports, as is allowed under the World Trade Organization's rules. These rebates are not relevant because the authors use tax data gross of rebates. But in quite a few of China's export-processing zones, exports produced with imported components are exempted from the VAT. Shenzhen is a good example of a special zone that had this tax treatment (Ministry of Commerce 1986). Furthermore, China's policy toward the exemptions has changed over time. In the mid-2000s, as the overall trade surplus became very large and trade friction with the United States accelerated, China eliminated fully or partially the exemptions for many key exports. When the global financial crisis hit in 2008, the government reinstituted the VAT exemptions as a way of stimulating exports. The irregular application of the VAT was a constant complaint of the United States in economic dialogues with China at this time.[1] The issue

1. See, for example, USTR (2010, esp. 42–46).

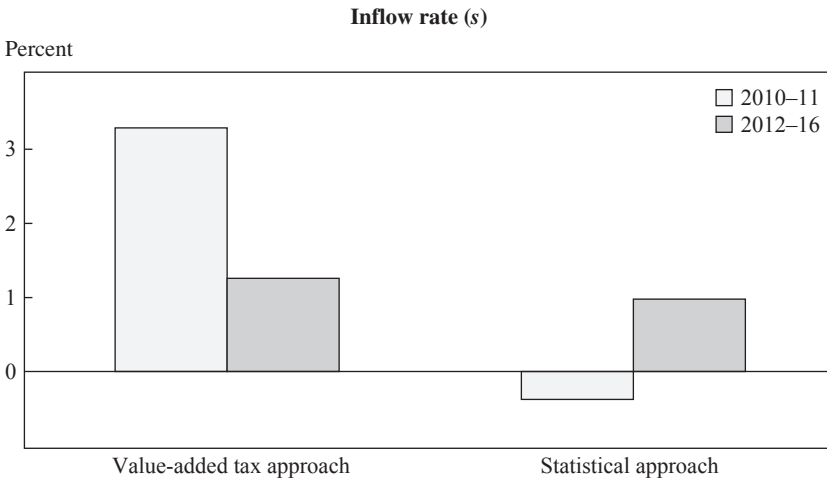
for the paper by Chen and colleagues is that there were changes in taxation policy and implementation right around the time of the global crisis.

Note in the top panel of Chen and colleagues' figure 4 that in the mid-2000s, VAT collected grew faster than industrial value added, but that starting with the global financial crisis, it grew more slowly. It is possible that some of the change resulted from changes in VAT rates for certain goods. In particular, as China started to recover from the global crisis, there were several years when industrial exports could increase without any increase in tax collection. In Chen and colleagues' figure 4, the largest discrepancy between the growth rate of VAT collection and of reported industrial value added is during 2009, 2010, and 2011, when China's stimulus was at its fullest.

A strength of the paper is that is has a second approach to estimating an alternative GDP series, which can then be compared with the results from the first approach. The second approach relies on regression analysis, specifically, a panel across provinces and over time. It assumes that industrial value-added data before 2008 were accurate, and it uses a set of explanatory variables whose data collection would plausibly not be subject to the same biases as the provincial reporting of industrial value added. These variables are national tax revenue collected from the province, satellite light data, electricity consumption, railway cargo, and trade. An equation is estimated for industrial value added with the data from 2000 to 2008. It is then used to predict provincial industrial value added until 2016. The aggregation of these provincial estimates is the estimate of national industrial value added. This estimate then replaces the official estimate of industrial value added to come up with a revised estimate of GDP. The light data are not available after 2013, and the tax data have the same potential problem already discussed. The authors include a reestimated GDP growth rate using this approach but dropping the light and tax data (Chen and colleagues' table 2, seventh column). For the period 2010–16, the reestimated growth rate averaged 10.88 percent a year, compared with the official figure of 11.50 percent a year and the reestimate from the first approach of 9.65 percent a year.

Thus, there is a pretty significant discrepancy between the results of the two approaches. The first approach suggests that GDP growth in China is overestimated by 1.8 percentage points a year; this is the number cited in the paper's abstract. The second approach finds a much smaller distortion, of 0.62 percentage point a year. It turns out that the difference in the estimates arises almost completely from the estimates for the crisis years, 2010 and 2011. For 2010–11, the VAT approach finds that official GDP

Figure 1. Overestimation of Official GDP Growth



Source: Author's calculations.

growth was overestimated by 3.3 percentage points a year; the statistical approach finds that official numbers *underestimated* actual growth by 0.4 percentage point a year (see my figure 1). Recall that these were years when the growth of VAT collection dropped sharply; but apparently the data from electricity consumption, rail cargo, and trade do not confirm such a sharp slowdown.

For the 2012–16 period, conversely, the two approaches provide very similar estimates. The VAT approach finds official GDP growth to be overestimated by 1.3 percentage point, compared with 1.0 according to the statistical approach. Hence, what is robust is that in recent years China's growth numbers have been overstated by about 1 percentage point.

A couple of empirical points from the paper change if the overestimation of GDP is taken from the statistical approach rather than the VAT tax approach. In particular, because the error in GDP on the expenditure side is in investment, the new estimates provide lower estimates of the investment rate and the savings rate than we see in official data. The paper reports that using the VAT tax-based estimates reduces the investment rate by 7 percentage points, to 35.5 percent of GDP in 2016. The savings rate is similarly reduced by about 7 percentage points of GDP, to 39.7 percent of GDP, because the balance of payments data are basically taken as accurate. If, alternatively, the revised estimates based on the statistical model

are used, the changes to the investment and savings rates would be significantly smaller. A back-of-the-envelope calculation suggests that, using the estimates from the statistical model, the revision to the investment rate is only 2.4 percentage points of GDP. Given the uncertainty about each set of revisions, a reasonable approach would be to average the estimates. This approach would reduce the investment rate in 2016 by 4.7 percentage points of GDP. In light of this large amount of uncertainty, we can only conclude that the investment and savings rates are likely overestimated in official data, but it is difficult to say precisely by how much. Both the investment and savings rates have been trending down in the official data, so these revised estimates suggest that there has been even more rebalancing in the Chinese economy than previously thought—rebalancing in the sense of a shift from reliance on investment to a greater reliance on consumption on the expenditure side.

Another strong feature of the paper is that it is able to look at the issue of GDP overestimation province by province. Chen and colleagues' figure 15 shows the cumulative adjustment to provincial industrial value added, by 2015, according to both methodologies. A couple of things jump out from this figure. First, the big coastal provinces that produce much of China's GDP appear to have pretty good data; that is, there is no large distortion according to either methodology. This is true for Guangdong, Zhejiang, Jiangsu, and Shandong. Second, overall, there is a positive correlation between the results from the two approaches: the provinces found to have exaggerated GDP according to the first approach also have exaggerated GDP according to the second approach. Third, much of the overall distortion depends on 10 provinces in the upper-right quadrant of the figure, ones whose GDP is overestimated by at least 5 percent in 2015 according to both approaches. This group includes the rust belt of the Northeast (Heilongjiang, Jilin, and Liaoning) and the interior provinces (Qinghai, Inner Mongolia, Guizhou, Hunan, Hubei, Guangxi, and Anhui). A number of these provinces have had corruption scandals that have involved, among other things, falsified economic data. The point here is that if China wants to clean up this problem and improve its data, it knows where to start. The fact that the coastal provinces have generally reliable GDP data shows that China is capable of producing decent statistics.

In summary, this paper by Chen and colleagues is a useful examination of the likely errors in China's GDP statistics and GDP's major components on the production and expenditure sides. It makes a plausible case that local officials in some cases are able to get national acceptance of overestimation of provincial industrial value added and investment. A strength

of the paper is that is uses two different approaches to reestimate industrial value added, and hence GDP. The two approaches yield significantly different estimates of the distortion in the GDP data, and the presentation emphasizes the larger distortion, when in fact the other approach is probably more defensible. What is robust in the two approaches is that the growth rate of nominal GDP has been overestimated by about 1 percentage point a year since 2010, that the savings and investment rates have come down more than is shown in official data, and that the errors are particularly large in 10 provinces of the Northeast and the interior.

REFERENCES FOR THE DOLLAR COMMENT

Bradsher, Keith. 2019, "China's Economic Growth Hits 27-Year Low as Trade War Stings." *New York Times*, July 14. https://www.nytimes.com/2019/07/14/business/china-economy-growth-gdp-trade-war.html.

Ministry of Commerce. 1986. "Supplementary Regulations for Tax Reduction and Exemption for Enterprises in Shenzhen Special Economic Zone, May 8." Ministry of Commerce, People's Republic of China, Beijing.

USTR (Office of U.S. Trade Representative). 2010. *Report to Congress on China's WTO Compliance.* Washington: U.S. Government Publishing Office.

COMMENT BY

WEI XIONG This paper by Wei Chen and his colleagues provides a fascinating study that systematically "corrects" China's GDP statistics. My discussion focuses on three aspects related to China's questionable GDP statistics: (1) the economic mechanism that leads to systematic misreporting; (2) the approach taken by Chen and colleagues to correct the misreporting; and (3) the potential consequences of such misreporting.

THE ECONOMIC MECHANISM OF OVERREPORTING Chen and colleagues provide compelling evidence of substantial inconsistencies between China's national and provincial GDP, mostly in the reported industrial output. Such inconsistencies reveal systematic overreporting in provincial statistics as well as the National Bureau of Statistics' (NBS) effort to correct the overreporting in national statistics. What causes the overreporting in provincial statistics? This overreporting problem is deeply rooted in the career incentives of provincial officials because (1) provincial bureaus of statistics report provincial economic statistics under strong influence from the provincial governments, and (2) some provincial statistics, such as provincial GDP and industrial output, are important measures in the central government's performance evaluations of provincial officials.

Even though the Chinese government has long abandoned central planning, it continues to play a central role in an increasingly market-driven economy. China has a complex government system, in which the central government works with regional governments at several levels: province, city, county, and township. Regional governments play key roles in China's economic development. Regional governments carry out over 70 percent of fiscal spending and are responsible for developing economic institutions and infrastructure at the regional level—for example, opening new markets and constructing roads, highways, and airports. Despite their autonomy in economic and fiscal issues, regional government leaders are appointed by the central government rather than being elected by the local electorate. As a key mechanism to incentivize regional leaders, the central government has established a tournament among officials across regions at the same level that uses economic performance to determine their career advancement. This system has greatly stimulated China's economic growth by giving local officials both fiscal budgets and career incentives to develop local economies. However, such powerful incentives may also lead to short-termist behaviors of local officials, such as overreporting local economic statistics, especially those that are most relevant for the performance evaluation.

My 2018 paper systematically examines such short-termist behaviors by developing a "Mandarin model" to account for the agency problems between China's central and local governments in affecting the Chinese economy. This model builds on the growth model of Robert Barro (1990) with a number of regions. In each region, the representative firm has a Cobb–Douglas production function with three factors: labor, capital, and local infrastructure. By creating more infrastructure in the region, the local government can boost the productivity of the local firm. Infrastructure investment represents the key channel for the local government to directly stimulate the local economy. The aforementioned tournament helps to mitigate the local government's tendency to underinvest in infrastructure relative to the social optimum. As more investment in infrastructure improves regional output, the tournament generates an implicit incentive for each region's governor to invest in infrastructure through the "signal-jamming mechanism" coined by Bengt Holström (1982), given that the central government is unable to fully determine whether regional output is due to the governor's ability or infrastructure investment. The powerful career incentives may also lead local governments to overreport regional output. This mechanism is similar in spirit to overreporting of earnings by executives of publicly listed firms (Stein 1989).

Specifically, the Mandarin model outlines a trade-off between the provincial government's GDP overreporting and the tax transfer to the central government. For the sake of argument, suppose that a provincial government collects tax revenue at a tax rate of τ of the province's GDP, Y_t, in year t, and needs to transfer a fraction of the tax revenue to the central government at a rate τ_c of the reported GDP Y'_t, with the transfer rate τ_c being lower than the gross rate τ. Then, the residual tax revenue for the provincial government is $T_t = \tau Y_t - \tau_c Y'_t$. It is clear that overreporting ($Y'_t > Y_t$) reduces the local fiscal budget, which in turn disciplines the overreporting. This trade-off leads provincial leaders with greater career incentives to overreport more. In addition, the Mandarin model implies that career incentives would also lead provincial leaders with strong career incentives to aggressively use leverage to boost their fiscal budgets at the expense of future debt burdens. Xiong (2018) provides a scatter plot of the ratio of provincial GDP overreporting to GDP, which is based on the estimation by Chen and colleagues and on the local debt-to-GDP ratio in 2015. The plot shows that these two types of short-termist behaviors are correlated across provinces. This curious association likely reflects the same mechanism driving these short-termist behaviors, as shown by the Mandarin model.

The dynamics of provincial GDP reporting are more nuanced than simply always overreporting. Commentators sometimes argue that fast-growing provinces like Guangdong have sufficient margins to meet their growth targets and may choose to underreport, rather than overreport, their GDP, because underreporting may help to reduce tax transfers to the central government and keep a strategic buffer for the future.

The inconsistencies between national and provincial GDP also reflect the fact that the central government is fully aware of the incentives of provincial governments to overreport their economic statistics and has made an effort to correct the overreporting in the national statistics. Interestingly, the NBS does not provide any breakdown on its assessments of overreporting by individual provinces, possibly because it does not want to publicly embarrass the provincial leaders, some of whom are already members of the Politburo and some of whom will eventually become national leaders. Furthermore, the correction by the NBS is constrained by its own data limitations, along with its own incentives. The NBS may not have a bias toward overreporting national statistics, yet it may be reluctant to report statistics that fall substantially short of the economic targets preannounced by the central government. In this sense, the NBS has also another set of incentives to manage the national statistics.

ESTIMATING OVERREPORTING Chen and colleagues use value-added tax revenue as the basis for correcting the overreporting in the GDP statistics. The premise of this approach is that local governments would not overreport value-added tax revenue because doing so would lead to greater shares of the actual tax revenues being transferred to the central government. This estimation approach is appealing and powerful and leads to a set of interesting and relevant corrections. Yet it is useful to note that although it is costly to overreport tax revenue, it may nevertheless happen in practice. In recent years, several regions—such as Liaoning, Tianjin, and Inner Mongolia—have officially acknowledged the previous overreporting of their GDP and tax revenues. These confessions typically happened after the previous regional leaders lost their prominence due to corruption allegations, and they helped the current leaders to obtain bigger fiscal subsidies or transfers from the central government. The confessed overreporting of tax revenue was substantial, even though it might have also been exaggerated due to the current leaders' incentives to plead for central government transfers.

CONSEQUENCES OF OVERREPORTING Economic statistics are an important source of information that helps policymakers, firms, and individuals adjust their policy, production, and investment decisions in response to time-varying economic conditions. It is difficult to directly estimate the potential distortions induced by the misreporting of economic statistics. China's Great Famine in 1959–61 was an extreme example of the deadly consequences of overreporting regional agricultural output. Xin Meng, Nancy Qian, and Pierre Yared (2015) provide forceful evidence that the severity of the famine was driven by overprocurements of grain during this period. Ziying Fan, Xiong, and Li-An Zhou (2016) point out a surprising observation: that the central government provided no famine relief during the first two years of the famine. To the contrary, China had exported a large quantity of grain to other countries, including sending food aid to African countries, in 1959 and 1960, after the famine had already spread across all of China. It is important to note that this devastating famine was also accompanied by widespread overreporting of grain output by regional leaders, due to their desire to support Mao Zedong's plan to quickly increase the country's agriculture production in support of the Great Leap Forward. Such overreporting might have led to excessive procurements of grain, which left insufficient grain to support local civilians. More profoundly, the overreporting might have also encouraged the central government's radical policies of moving more workers from agricultural production to industrial production just when grain output was rapidly falling.

In the modern era, there are many other sources of information for the central government and the general public to gauge the economic conditions in China. Thus, we do not expect this kind of fatal information breakdown to occur again. Nevertheless, the overreporting of economic statistics documented by this paper deserves close scrutiny.

REFERENCES FOR THE XIONG COMMENT

Barro, Robert J. 1990. "Government Spending in a Simple Model of Endogenous Growth." *Journal of Political Economy* 98, no. 5: S103–S125.

Fan, Ziying, Wei Xiong, and Li-An Zhou. 2016. "Information Distortion in Hierarchical Organizations: A Study of China's Great Famine." Working paper. https://wxiong.mycpanel.princeton.edu/papers/Famine.pdf.

Holmström, Bengt. 1982. "Managerial Incentive Problems: A Dynamic Perspective." In *Essays in Economics and Management in Honor of Lars Wahlbeck*. Helsinki: Swedish School of Economics.

Meng, Xin, Nancy Qian, and Pierre Yared. 2015. "The Institutional Causes of China's Great Famine, 1959–1961." *Review of Economic Studies* 82, no. 4: 1568–1611.

Stein, Jeremy C. 1989. "Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior." *Quarterly Journal of Economics* 104, no. 4: 655–69.

Xiong, Wei 2018. "The Mandarin Model of Growth." Working paper, Princeton University. http://wxiong.mycpanel.princeton.edu/papers/Mandarin.pdf.

GENERAL DISCUSSION Eswar Prasad observed that Chinese nominal GDP growth fluctuates while real GDP growth is smooth, which suggests both a problem with the deflators and any real GDP numbers derived using these deflators. Prasad noted a large change in the composition of Chinese GDP, of which the service sector accounts for more than 50 percent. This service sector growth has been largely in areas where the value-added tax (VAT) is much harder to collect. These two observations lead Prasad to believe that there might be an undercounting of GDP. He explained that a few years ago, the sampling frame for the service sector was rebased, and a few extra percentage points of GDP were found. He praised the authors for their analysis of the industrial and investment sectors, and noted that the Chinese household surveys give a reasonable sense of what is happening with household savings. In addition, he observed that there is already a reasonable measure of corporate and government savings, suggesting that it must be possible to reconcile these numbers with those generated by the authors to get the current account balance right. Finally,

he mentioned his time working on China, where he became aware that even those working for the People's Bank of China were skeptical of the official GDP numbers. He recalled a former senior official who at one time became very concerned about the numbers of the Chinese National Bureau of Statistics (NBS) and came up with his own measure. He presented these new numbers—which differed significantly from the official numbers—to the State Council. Two weeks later, this same official received a letter from the NBS informing him that it was illegal for any other institution besides the NBS to collect data on GDP.

Jason Furman remarked that there have been two hypotheses about Chinese data. The first is the smoothing hypothesis, whereby China reports its GDP sooner than anyone else and does not revise its numbers. The second hypothesis is that China is in fact overstating its own growth. Furman wondered what forensics Chen and others had devised about the first hypothesis, because it could be confused with the second. Indeed, an unrealized permanent negative shock to the growth rate would result in the smoothing of GDP. Finally, Furman asked whether Chen and others had looked at the deflators issue, given that he has heard it plays into both hypotheses.

Tom Orlik described a time in 1998 when Premier Zhu Rongji noticed some falsification and embellishments in the Chinese statistical system. Orlik understood that after 1998, the Central Government and the NBS became more powerful relative to the local statistical bureaus, and he asked the authors whether his understanding was correct. If so, he wondered whether something had changed in the most recent period that had impeded or prevented the NBS from cross-checking the local area statistical agencies' numbers.

Orlik observed that in the last decade, China's credit growth has been rapid and that most of this growth has gone toward paying for investment spending. At face value, Orlik noted that this contradicts Chen and others' thesis that investment is being overreported in the national accounts. Orlik noticed that this contradiction is lessened by David Dollar's observation that credit growth is not going toward anything and that little investment is happening. However, Orlik questioned whether this contradiction is completely diminished. Finally, Orlik pointed out compositional issues with the VAT data and wondered whether it serves as a good proxy for growth.

Elaine Buckberg commented that the government officials in China and the provinces are powerfully responding to incentives for promotion to higher office. However, Chinese officials are also incentivized to attract transfers to the provinces to make sure they remain in office. Buckberg

asked at what point the incentive to retain their position offsets the incentive to overstate GDP. Similarly, she agreed with Prasad's observation about the rapidly growing, largely private, service sector; and she pointed out that service sector underreporting will increasingly offset government officials' incentive to overstate GDP.

John Haltiwanger wondered whether it would be possible to examine the external validity of the paper, noting that in the United States, there are times when the survey data and administrative data do not match, particularly during periods of large cyclical variation. As such, Hatiwanger asked whether Chen and his colleagues would be able to replicate their analysis using another country. Haltiwanger then suggested replicating how the U.S. Bureau of Economic Analysis measures investment, instead of using survey data. Specifically, the bureau calculates investment flows by taking the nominal production of the capital goods industry, subtracting exports, and adding imports.

Stanley Fischer observed the small difference between the official numbers and what Chen and his colleagues find. Fischer commented that the difference seemed especially small, considering existing research that finds an underreporting of GDP of about 15 percent due to black market activity in Western economies.[1]

Jonathan Wright reflected on Yingyao Hu and Jiaxiong Yao's paper, which uses satellite measurements of electricity as a true output measure.[2] Hu and Yao also compared their generated numbers with those of the NBS, and observed a misreporting number roughly double that of Chen and colleagues. Wright observed that Hu and Yao do not use any Chinese data to generate this number.

Richard Cooper mentioned that the prime minister of China usually preannounces GDP growth to the National People's Congress in a speech. Cooper contemplated the pressure the NBS must feel to line up its numbers with this official government narrative. For example, Cooper described that two years ago, the prime minister announced that targeted GDP growth would be between 6.5 and 7 percent. Later, the head of the NBS clarified that its official 6.7 percent number rounded up to 7 percent. On a separate note, Cooper wondered about the validity of the official Chinese deflator numbers and advised the authors to address this issue in their paper.

1. Paulina Restrepo-Echavarria, "Macroeconomic Volatility: The Role of the Informal Economy," *European Economic Review* 70 (2014): 454–69, https://ideas.repec.org/a/eee/eecrev/v70y2014icp454-469.html.

2. Yingyao Hu and Jiaxion Yao, "Illuminating Economic Growth," working paper, 2019, http://www.econ2.jhu.edu/people/hu/paper_HUandYAO.pdf.

Jay Shambaugh described a time in about 2009 when the U.S. government was especially concerned that China was increasing its VAT rebates as a way to subsidize both exports and exporting industries. He wondered whether Chen and colleagues could look to see how much these VAT rebates on exports were increased, as an adjustment to their own adjustment.

Chang-Tai Hsieh clarified that their VAT number is the gross number before rebates. He noted that they chose to look at the industrial sector, in addition to wholesale and retail trade, because these are the two sectors to which the VAT tax applies. He noted that they decided to focus on these two industries because they were reasonably confident that there was relatively little evasion on the tax side and that any tax exemptions took the form of rebates on the export side. Although he wished they could do the same thing for the other service sectors, he noted that the VAT does not apply to these other sectors and that other data—such as corporate income tax data—were of poor quality and had more exemptions. As a result, he and his coauthors decided to use the official number for the other service sectors. He thanked Haltiwanger for his comment on a different way to measure investment data and noted that this method was implemented in their paper.

Hsieh explained that they used the official deflators and hoped that someone else would investigate whether the official deflator numbers were right. In response to whether the data on credit growth are consistent with their revised numbers for investment, he observed that an increase in credit does not necessarily translate into a one-to-one increase in investment, due to the substantial amount of substitution in sources of financing. In terms of the smoothing of GDP, he explained that nominal GDP becomes smoother after 2007. He noted that they did not feel confident that this smoothness really told them anything, due to the NBS's adjustments.

In response to comments made on whether the adjustments are big or small, Hsieh noted that it is hard to dispute that Chinese economic growth has not been high. He clarified that their paper only claims that GDP growth might be about 1 or 2 percent lower per year. Indeed, finding anything more—for example, that real GDP growth has been zero—would not be consistent with the obvious: that the Chinese economy has been growing. Hsieh clarified that he does not believe that the Chinese statistics are that different from the true number. Moreover, he explained that this discrepancy in the official numbers is similar to those in other countries. He referred to a paper he wrote on the Singaporean National Accounts, where

he also found inconsistencies in the official data.[3] However, he noted that, although the difference found from the official number is small, he still believes it to be valuable information.

Finally, Hsieh observed the incredible amount of power held by local Communist Party secretaries. He noted that none of the central government ministers—such as the minister of finance—are members of the Politburo of the Communist Party. This implies that the party secretaries of Shanghai, Beijing, and Guangzhou, for example, are significantly more powerful than any finance minister or central bank governor. Of course, this power structure also means that the head of the NBS is vastly outranked by these local Communist Party leaders. He concluded that though this structure has contributed to the success of the Chinese economy, it also explains discrepancies found in the data and elsewhere.

3. Chang-Tai Hsieh, "What Explains the Industrial Revolution in East Asia? Evidence From the Factor Markets," *American Economic Review* 92, no. 3 (2002): 502–26.