Institute for International Economic Policy Working Paper Series
Elliott School of International Affairs
The George Washington University

Characteristics and Implications of Chinese Macroeconomic Data Revisions

IIEP-WP-2012-09

Tara M. Sinclair
George Washington University

August 2012

Institute for International Economic Policy
1957 E St. NW, Suite 502
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: **iiep@gwu.edu**
Web: **www.gwu.edu/~iiep**

Characteristics and Implications of Chinese Macroeconomic Data Revisions

Tara M. Sinclair*
Department of Economics
and the Elliott School for International Affairs
The George Washington University
2115 G Street NW
Monroe Hall #340
Washington, DC 20052
tsinc@gwu.edu

Keywords: China, Real-Time Data, Data Revisions

JEL Codes: C82, E01, E30

August 21, 2012

ABSTRACT

Recent research examining U.S. macroeconomic data suggests that revisions may be much more important than traditionally assumed. This paper extends the analysis to Chinese data, where there has been substantial debate about data quality for some time. The key finding in this paper is that indeed the Chinese macroeconomic data revisions are not well-behaved, but that they are not much different from U.S. macroeconomic data revisions.

*The author thanks GW-CIBER and the Institute for International Economic Policy at the Elliott School of International Affairs for generous support of this project. Excellent research assistance was supplied by Caitlin Briglio, Shirley Hsieh, Xiaohan Ma, and Rajeev Patel. Dean Croushore, Yueqing Jia, Fred Joutz, Arun Malik, Herman Stekler, Simon van Norden, and participants at the 2012 G2 at GW conference provided helpful discussions and comments. All remaining errors are the responsibility of the author.

I. Introduction

Traditional macroeconomic analysis assumes data revisions are small and random and thus have no effect on economic modeling, policy, or forecasting (Croushore, 2011). Recent studies of U.S. macroeconomic data, however, suggest that revisions may be much more important than previously assumed (e.g. Aruoba, 2008; Croushore, 2008 and 2011; Dynan and Elmendorf, 2001; Kennedy, 1990; Sinclair and Stekler, 2012; Swanson and van Dijk, 2006). The recent global dependence on the macroeconomic performance of China,[1] as well as concerns about the quality of the data being released by Chinese statistical agencies,[2] suggest that a thorough study of the macroeconomic data revisions for China is an important endeavor.

Most macroeconomic data for any country are obtained from official government sources. Generally these data are considered trustworthy and valuable for firm, household, and policy decisions. However, the quality of data released by the National Bureau of Statistics of China (NBS), China's statistical authority, has regularly been questioned by both the media and researchers (e.g. *The Economist*, 2008; Holz, 2003 and 2006; Huang, 2011; Keidel, 2001 and 2007; Maddison, 1998 and 2006; Orlik, 2012; Rawski, 1976 and 2001; Rawski and Xiao, 2001; Ren, 2002; Wu, 2000; Xu, 2002; Young, 2003). For example, Heston (2001, pg. 3) claims that there are "winds of falsification" that surround the Chinese macroeconomic data, particularly in the late-1990s. Although there has been substantial research on the accuracy of Chinese macroeconomic data generally, there has been little analysis of the revisions of the Chinese data

[1] For example, see the New York Times article, "China, Driver of World Economy, May Be Slowing" (Bradsher, 2011) and the 2011 speech by Justin Lin, Chief Economist and Senior Vice President of the World Bank, where he claims that "growth in China has been a driving force for the recovery from the global crisis since 2009," and "in the decade beginning in 2000, China became the top contributor to the growth of global GDP."

[2] For a recent discussion of the literature on the quality of Chinese macro data, see Appendix 1-1 from Jia (2011) as well as the references discussed further below.

as released by the NBS. The small amount of evaluation of past macroeconomic data revisions for China has been used to argue that large data revisions suggest poor data quality for China.[3]

In contrast with Chinese data, even early releases of U.S. macroeconomic data are generally considered trustworthy and are carefully analyzed by the media as well as firms and policymakers the moment they are released. A growing body of research, however, has shown that revisions to U.S. macroeconomic data are substantial and "not well-behaved" (Aruoba, 2008). Aruoba shows that the early U.S. data show bias and have large and predictable revisions. In addition, it has been shown that for some U.S. macroeconomic data, large revisions are more likely in recessions than in expansions (Dynan and Elmendorf, 2001; Swanson and van Dijk, 2006). Furthermore, biases in data estimation have been identified for a number of U.S. macroeconomic series (Kennedy, 1990; Croushore, 2008; Sinclair and Stekler, 2012).[4] This research suggests that even the U.S. macroeconomic data experience substantial revision and bias, therefore the presence of revisions and biases in Chinese data would not be in itself evidence of lower data quality.

Households, firms, and policymakers rely on early macroeconomic data to make timely decisions. Knowing the properties of the historical revisions to the early data can help users of the data better interpret the early information they receive. Given the growing importance of the Chinese macroeconomy for global business, this paper fills a gap in the literature by providing a careful examination of China's macroeconomic data revisions. The methodology follows the analysis of revisions previously applied to U.S. data (see Croushore, 2011, for a recent survey).

[3] Two exceptions are Wu (2007) and Holz (2008) who both examine the impact of the first economic census, conducted in 2004, on the Chinese GDP data. Wu (2007) concludes that the 2006 benchmark revised GDP estimates based on the economic census are "not less questionable" than the data available before the census (pg. 2). Holz (2008) concludes that the benchmark revision "implies that Chinese statistics have to be taken with a *rock* of salt" (Holz's italics, pg. 163).

[4] A longstanding concern about data revisions is their impact on forecasts and forecast evaluation, for example see Stekler (1967), Joutz and Stekler (1998), Croushore and Stark (2002).

Furthermore, a comparison between revisions in China with those in the U.S. will give a relative measure of the performance of China's statistical authority as compared to what is typically considered the "gold standard" in data production. This will help develop a better understanding of the difference between typical challenges to data collection and those unique to China.

This paper proceeds as follows: first the data are introduced, including the new real-time macroeconomic dataset for China that is available as a companion to this paper.[5] Next in the results section the revisions to the Chinese macroeconomic data are evaluated and assessed relative to a comparable set of U.S. macroeconomic data revisions and according to Aruoba's (2008) requirements for well-behaved data revisions. Finally the conclusions suggest that although the Chinese macroeconomic data revisions are not well-behaved, they are not much different from U.S. macroeconomic data revisions.

II. Data

This paper focuses on the official annual macroeconomic data estimates as printed in the China Statistical Yearbook each year as part of the national accounts for China, published by the National Bureau of Statistics of China. These yearbooks have been released annually from 1981 through 2011 and cover the sample from 1978 to the year before the yearbook date.[6] The data consist of five series all in real terms and reported in annual growth rates: GDP, GNP, and the three main sectors (primary, secondary, and tertiary) underlying the production approach to measuring GDP. Primary is generally agriculture, secondary is the processing of primary products, and tertiary is broadly the service sector.[7] China's official approach is the production

[5] The dataset is available at http://home.gwu.edu/~tsinc/china_rtd.xlsx

[6] There are also quarterly data available for a portion of the sample, but the estimates of annual GDP printed in the official yearbooks are made separately from the quarterly estimates (Lequiller and Blades, 2007).

[7] Further detail on the NBS definitions of the industries is available on their website: http://www.stats.gov.cn/english/classificationsmethods/definitions/t20020419_402787584.htm

approach (Holz, 2008), unlike the U.S. that relies primarily on the expenditure approach to calculating GDP. The production approach is also considered to be the most reliable estimate for China (Lequiller and Blades, 2007). For this reason this analysis will focus on the three main sectors rather than the sub-components of the expenditure approach for evaluation.

a. Real-Time Dataset

As regularly noted in the China Statistical Yearbook (for example in 2006): "Data on gross domestic product and related indicators of the most recent year published in the yearbook are not final and are subject to changes when more information from financial data and administrative records are available." It was therefore necessary to construct a dataset that reflected the data that were available at different moments in time. The dataset for China was constructed in a similar manner to the "real-time dataset" for the U.S. introduced by Croushore and Stark (2001). Following the terminology of Croushore and Stark (2001), "vintages" were created which capture the data from each statistical yearbook. For example, the 2006 vintage is the data printed in the 2006 statistical yearbook. Data revisions result in different vintages having different values for the same category and date. Table 1 presents an example of the structure of the dataset.

For parts of the dataset, the secondary sector was further divided into industry and construction and the tertiary sector was further divided into transport and commerce. In this analysis, however, we will focus on the most frequently reported series to have the longest samples possible of real-time data.

b. Revisions

The revisions examined for each series in this paper are (1) the changes in the estimates from the first to the second time the data are reported in a statistical yearbook ("first revision") and (2) the changes from the first time the data are included in a statistical yearbook to the latest available data from July, 2012 ("final revision").[8] To construct these revisions we need to create three series: the "first release" which will be the main series we are analyzing, and two alternative series to compare with the first release: the "second release" and the "final" data. We create the "first release" series by collecting the last entry in each yearbook, i.e. the data for the year before the yearbook date. Table 2 presents an example of this series. Comparing this table to Table 1, it is clear why the first release data are sometimes called the "diagonals." The data reported in Table 2 appear on the bottom diagonal of Table 1. The "second release" is (the growth rate of) the data reported in the statistical yearbook two years after the year in question. The "final" or latest available data is simply the July, 2012, vintage of data (which should closely approximate the 2012 statistical yearbook which was not yet in print at the time of writing this article).

The sample for each series starts with the first available "real-time" release, which varies for each of the five series. For example, real GDP data were first reported starting in the 1994 statistical yearbook, therefore the first release sample for that variable starts in 1993. For all variables the first release sample ends in 2011. All first revisions series end in 2010 because we only have second releases through 2010. All final revisions series end in 2008 so that they incorporate the benchmark revisions from both of the national economic censuses conducted in

[8] There are releases between yearbook publications that are available on the NBS website, but the releases in the statistical yearbooks are the officially printed data each year and what will be focused on for this paper. Furthermore, there will most likely be further revisions after July of 2012, so the term "final" here only applies to the data available up to the writing of this paper. This terminology follows that of Aruoba (2008).

China, as discussed further below. Tables 3 through 6 provide descriptive statistics and sample coverage for the first release and latest available data as well as the revisions.

c. Economic Censuses and Benchmark Revisions

Most countries do regular benchmark revisions of their national income and product accounts incorporating definitional changes, infrequent census data, changes in international conventions, etc.[9] For China, there are two key economic censuses that were conducted during our sample that resulted in revisions beyond the regular annual revisions.[10] The First National Economic Census was conducted in 2004. In the 2006 statistical yearbook, we read "In 2005, the GDP figures for the year 2004 were recompiled in accordance with the *Programme of Compilation of GDP and National Accounts for the Year of Economic Census* and with the result from the first economic census. GDP data for earlier years were revised by trend deviation approach (trend of historical data were first calculated with data from the census, which was further revised by the ratio between the value of historical data and the trend of historical data, resulting in revised historical data series.). Data for 2005 are compiled on the basis of revised 2004 data." The 2006 statistical yearbook is the first one to include these revisions. The Second National Economic Census was conducted in 2009 and incorporated in the 2010 Yearbook data. Therefore, the latest available data include at least one benchmark revision relative to the first release data.

III. Results

Tables 3 and 4 present the descriptive statistics for the first release and the latest available vintage (as of July, 2012) of the real annual growth rates of Chinese GDP and GNP as well as the

[9] For example, see McCully and Payson's (2009) preview of the latest U.S. benchmark revisions released in July of 2009.

[10] There was also a first tertiary industry census that resulted in benchmark revisions in 1995, but these do not affect much of the sample investigated here.

value added for the primary, secondary, and tertiary sectors. Consistent with the assumptions of the OECD (Lequiller and Blades, 2007), the assumption in this paper is that the more heavily revised data are of higher quality than the earlier releases. It is the first release, however, that firms, households, and policymakers generally use in decision-making. Generally they cannot wait for more information to arrive. If anything, they are actually responding to earlier data and/or forecasts rather than waiting for the official yearbook to be printed more than halfway into the following year. As can be seen by comparing the means and medians from Tables 3 and 4, the average growth rate for each of the 5 series is less in the initial publication than in the latest available data. This is consistent with the overall view that in general the macroeconomic data for China have been revised upwards over time.

a. First Revisions

Table 5 presents the descriptive statistics of the first revisions, i.e. the difference between the real growth rates released in the statistical yearbook two years after and those from one year after the year of the economic activity they refer to. For both GDP and GNP the means are statistically significantly positive at the 10% level.[11] This suggests that the initial estimates are biased (Holden and Peel, 1990). Looking at the breakdown into the sectors it is clear where the bias is coming from – only the tertiary sector has a statistically significant mean, and it is again positive. In fact, the mean absolute revision for the tertiary sector is the same as the mean error. This indicates that revisions from the first release to the second release were never negative for the tertiary sector data. For the other two sectors the first releases appear to be unbiased estimates of the second releases, with larger absolute revisions for the secondary sector as compared to the primary sector. Real GDP has mean absolute revisions very similar to its mean

[11] Newey-West standard errors (Newey and West 1987) are reported throughout and used for all statistical tests due to the concern that the data and the revisions may be autocorrelated.

revisions whereas real GNP appears to have much larger mean absolute revisions. This difference between GNP and GDP is in large part due to the longer sample for GNP.

b. Final Revisions

Table 6 presents the descriptive statistics of the final revisions. It is important to recall that these revisions are measured from the first release to the latest available data, so that if the data are unchanged from the second release the revisions will be the same size as those used to construct Table 5. As expected, the mean absolute revision for each of the series is larger in Table 6 than in Table 5 since more information has been gathered from the first release to the latest available as compared to from the first release to the second release. Otherwise the results are generally similar to those reported in Table 5 with GDP, GNP, and the tertiary sector appearing to have a positive bias and primary and secondary appearing unbiased based on their estimated means.

It is interesting to note that the mean and mean absolute revision is the same when going from the first release of GDP to the latest available data, suggesting that there was never a negative revision from the first to latest available data over this sample for GDP. There were, however, some small negative revisions from the first to second release of GDP, but they were reversed in the most recent data. For GNP and for the primary and secondary sectors there are revisions in both directions for both the first revisions and final revisions. For the tertiary sector, however, there were never any negative revisions as reported in Tables 5 and 6. This may be in part due to the shorter sample for the tertiary sector as compared, in particular, to GNP, but it is also consistent with the discussion in Holz (2008) that the NBS had a sense that it was under reporting tertiary sector value added.

c. Comparison with U.S. Data

To give some context for interpreting the size of the Chinese revisions, Tables 7 and 8 provide descriptive statistics for what is generally considered to be high-quality data, i.e. U.S. real GDP.[12] Table 7 presents the descriptive statistics for U.S. real GDP growth data that would be comparable to the Chinese data. The first column is for the annual data released each July for the previous year. These data would be consistent with the Chinese data that are being called the first release in this paper. The second column of Table 7 is for the latest available data for the U.S., again from July of 2012. The sample is similar for the U.S. and for China since the U.S. began reporting real GDP data in real time in 1992. The U.S. has a much smaller real GDP growth rate on average, at about 2.4% compared to China's almost 10% growth rate. In fact, China's minimum growth rate for the first releases of its real GDP, at 7.10% is larger than the maximum growth rate for the comparable U.S. data. Similar to the Chinese data, the mean of the latest available data for the U.S. is higher than that of the first July release. Unlike China, however, the later U.S. data have a lower median with a lower minimum but a higher maximum as compared to the first release data. For China, fixing a comparable sample of 1993 – 2011 (not reported in the Tables), the mean, median, maximum, and minimum values are all larger in the latest available data as compared to first release data for real GDP growth.

Table 8 presents the descriptive statistics for the revisions to U.S. real GDP. Compared to the Chinese data, it is important to notice that the U.S. revisions appear to be unbiased, with means that are statistically insignificantly different from zero at the 10% level. Furthermore, the mean absolute final revision is smaller than for the Chinese data. For the first revision, however, the magnitudes are similar for the U.S. and China, with the U.S. actually having a slightly larger

[12] U.S. real GNP and the revisions are similar to U.S. real GDP and therefore not reported here. There are not comparable U.S. data for the sectors because U.S. data are not divided into sectors in the same way.

mean absolute revision. This is particularly surprising given the difference in the magnitudes of the growth rates. A more appropriate comparison would be to compare the size of the mean absolute revision relative to the mean growth rate. These are reported in Table 9 for both the U.S. and China. The striking thing to note is that the relative revisions for China are smaller than for the U.S. for both the first revision and the final revision.

d. Graphical Representation

An alternative way of viewing the impact of Chinese data revisions is to represent the data graphically. As can be seen from Figure 1, both the second release and latest available data are generally higher than the first release series, which suggests that most of the revisions to Chinese real GDP have been positive over time.[13] This finding is consistent with the positive and statistically significant means for both the first revisions and the final revisions of Chinese real GDP growth rates as reported in Tables 5 and 6. Similar results are seen in Figure 2 for GNP for the same sample where we have real-time GDP data available. In the early part of the GNP sample, however, there is more variation in the direction of the revision. For the primary sector data we can see in Figure 3 that there has been little revision in the later part of the sample, which is consistent with the discussion in Lequiller and Blades (2007). For the secondary sector data in Figure 4 we see that the latest available data is essentially the same as the second release data for 1994 through 2004. From 2004 through 2010 the latest available data for the secondary sector were generally revised upward from both the first and second releases. Perhaps the most dramatic figure, however, is Figure 5 which shows the data for the tertiary sector. We see a similar pattern here as what was presented in Figure 1 for real GDP, but the

[13] For comparison, Figure 6 shows the different release series for U.S. real GDP. In the later part of the sample, from approximately 2002, it is apparent that U.S. data have been primarily revised downward.

difference between the first release data, the second release data, and the latest available data is substantially larger for the tertiary sector as compared to real GDP.

e. Testing for Autocorrelation

Further exploration of the properties of the Chinese revisions is presented in Tables 10 and 11. If we treat the early estimates as similar to forecasts then we can evaluate their rationality. Following Berger and Krane (1985), we examine the revisions to see if the revisions are predictable from past revisions, i.e. the revisions are autocorrelated. The regression we estimate is:

$$revision_t = \alpha + \beta revision_{t-1} + e_t \quad (1)$$

If the coefficient on the lag of the revision, β, is statistically significant then the estimates would not be considered rational because the revisions to the estimates would be predictable. The inclusion of the constant, α, allows for the average revision to be nonzero. A statistically significant constant would suggest that the estimates are inefficient, but as discussed in Ashiya (2005), this might occur in forecasts if the economy is frequently hit with unpredictable shocks with a common sign. Similarly for the estimates, if the data only available with a delay have a faster (slower) growth rate than the immediately available data, then we may expect a positive (negative) constant.

The joint test for rationality is that $\alpha = 0$ and $\beta = 0$. The joint test rejects rationality for four out of the five first revisions series with the exception being the secondary sector. For GNP and Primary the coefficient on the lagged revision is statistically significant and negative. For GDP and the tertiary sector it appears that the failure of rationality comes from the positive bias of the early estimates. For GNP there also continues to be a positive bias as well as a significant coefficient on the lag, whereas for the primarily sector the failure of rationality appears to come

exclusively from the autocorrelation of the series. For final revisions, however, it appears that only three of the series reject rationality based on the joint test, and all of these failures appear to come from positive biases. Both the primary and secondary sector final revisions series fail to reject rationality. There is evidence of significant autocorrelation only for the secondary series at the 10% level. As discussed in Aruoba (2008), Regression (1) is one way of testing the "news hypothesis" for revisions. If revisions are errors from a rational estimate (i.e. they are "news"), then the revision should be uncorrelated with anything that was in the available information set at the time the estimate was made, including past revisions. Therefore for the series that had significant autocorrelation, i.e. first revisions of GNP and primary and final revisions (at the 10% level) of the secondary sector, there is already evidence to reject the news hypothesis for revisions. Autocorrelation, however, is not required to reject the news hypothesis. Anything that should have been available at the time the initial estimates were made should not be able to explain the revision. Further tests of this hypothesis will be presented below.

f. Testing for Incorporating Information About the State of the Economy

Next we investigate whether or not the early estimates incorporated information about the state of the economy. The state of the economy likely was not in the information set at the time of the initial estimates, but it has become common for researchers to consider if the revisions of macroeconomic data series are different in recessions as compared to expansions (see Croushore, 2011, and Sinclair and Stekler, 2012, for examples). Unlike the U.S., China does not have any negative growth rate observations in this sample that would match the typical description of a recession. China has, however, experienced "growth slowdowns." Following Jia (2011), we use the following dates for the slowdowns: 1980, 1984, 1989, 1996-1998 and 2008. We then apply a methodology similar to Sinclair et al (2010) and estimate the following regression:

$$revision_t = \alpha + \beta slowdown_t + e_t \quad (2)$$

where *slowdown* is an indicator variable that takes the value 1 in the slowdown years as listed above, and 0 otherwise. We will call this a modified Holden and Peel (1990) test because it allows for two different biases depending upon whether or not the economy is in a slowdown. The joint test is that $\alpha = 0$ and $\beta = 0$ so that there is no bias under the null. Tables 12 and 13 present the results for the first revisions and the final revisions, respectively. It should first be noted that the sample sizes are very small and not rejecting the null does not mean accepting the null. Therefore the focus will be on the cases where the null is rejected with the jury out on the other cases. . The results suggest that for first revisions both GDP and the tertiary sector exhibit bias, but none of the dummy coefficients are statistically significantly different from zero. For final revisions, however, both GDP and the tertiary sector show both evidence of bias and a statistically significantly negative coefficient on the slowdown dummy. Furthermore, for final revisions we also find evidence of bias for Chinese GNP. For these series in both the first revisions and the final revisions there is evidence that initial releases are too low in normal growth times and too high in slowdowns. This is similar to the finding for the U.S. for expansions versus recessions.

g. Testing for Predictability of Revision from Size of First Release

An alternative way to explore the role of both the state of the economy and possible predictability of the revisions is to use the value of the initial release as a possible explanatory variable for the revisions. Tables 14 and 15 present the results of regressing the revisions on the first release data. The question here is whether the size of the revision is related to the size of the first release. The regression is:

$$revision_t = \alpha + \beta FirstRelease_t + e_t \quad (3)$$

The joint test is that $\alpha = 0$ and $\beta = 0$. In this regression it appears to be the secondary sector where the first release most clearly plays a role in the size of the revision for both first and final revisions. There is some evidence for GDP first revision and the tertiary final revision also being affected by the size of the first release. For first revisions of GNP and tertiary and final revisions of GDP and GNP there is also the rejection of the joint restriction whereas neither the constant nor the coefficient on the first release is statistically significant. In all significant cases it appears to be a positive relationship between the size of the initial release and the size of the revision. This is consistent with the finding of positive revisions in stronger growth periods as reported in Tables 12 and 13.

h. Properties of "Well-Behaved" Revisions

Aruoba (2008) proposes three properties of "well-behaved" revisions: 1) the estimates are unbiased (i.e. the revisions have a zero mean), 2) the noise to signal ratio, measured as the ratio of the standard deviation of the revision over the standard deviation of the final value, should be "small", and 3) the forecast revisions are unpredictable. Tables 16 and 17 report the results of testing for these properties for the Chinese macroeconomic data as well as for U.S. real GDP. We can reject the null of no bias for both the first and final revisions of three of the five Chinese series: GDP, GNP, and the tertiary sector. Thus the first property is violated for the majority of the Chinese series, but not for U.S. real GDP.

For the second property, Aruoba (2008, pg. 325) argues that a "range from 0.05 to 0.94 with an average of 0.39" represents large numbers. Therefore, for the Chinese data, where we have a range of 0.13 to 0.38 with an average of 0.25, also appears large. Table 17 shows that the noise/signal ratio for the US final revisions is the same as what Aruoba (2008) reported in his Table 1 (page 325). The mean is statistically insignificant for the US real output in both cases,

but Aruoba finds substantial bias for other U.S. variables (as do others, for example, Sinclair and Stekler, 2012). Finally, for the third property, we find predictability of the revisions from either a lag of the revision or the value of the first release for either the first or final revision for all five Chinese series as well as for U.S. real GDP. According to these requirements, this analysis suggests that the first releases of the Chinese macroeconomic data in general fail on all three counts. As shown here and discussed further in Aruoba (2008), however, the U.S. also frequently fails these tests.

IV. Conclusions

Chinese macroeconomic data generally fail Aruoba's (2008) tests for well-behaved data revisions. The problems, however, are not at all unique for China, since even U.S. data fail most of Aruoba's tests. Therefore, as argued by Croushore (2011), it is not necessarily an indictment of the statistical authorities that the revisions are not well-behaved, but rather a call for further research to improve data gathering processes for all countries. It is also a reminder to firms and policymakers dependent on the early data to expect there to be substantial revisions in future releases.

Like the U.S., it appears the Chinese official data can serve as "a reliable guide" to the level and growth pattern of GDP (Lequiller and Blades, 2007) and should be the "first port of call" (Scheibe, 2003). Alternative data resources have not proved to be more precise (Holz 2006, Chow 2006). Furthermore, Klein and Ozmucur (2003) also found, based on principal components analysis, that the official real GDP estimates for China appear to be consistent with a set of indicators that captures the Chinese economy more broadly. Jia (2011) does not find evidence of any measurement errors or irregularity in the Chinese real GDP data. Curtis and Mark (2010) also find that the official Chinese macro data appear consistent with economic

theory. We agree with Curtis and Mark that "understanding the macroeconomics of China is too important to wait until the 'high-quality data' are available" (Curtis and Mark, 2010, pg. 4).

Table 1: Example of the Real Time Dataset Structure: Chinese Real GDP Growth Rates[*]

Yearbook Date / Activity Date	1994	…	2008	2009	2010	2011	2012[†]
1979	7.6	…	7.6	7.6	7.6	7.6	7.6
1980	7.8	…	7.8	7.8	7.8	7.8	7.8
…	…	…	…	…	…	…	…
2006	n/a	…	11.7	11.7	12.7	12.7	12.7
2007	n/a	…	11.9	13.0	14.2	14.2	14.2
2008	n/a	…	n/a	9.0	9.6	9.6	9.6
2009	n/a	…	n/a	n/a	9.1	9.2	9.2
2010	n/a	…	n/a	n/a	n/a	10.4	10.4
2011	n/a	…	n/a	n/a	n/a	n/a	9.2

Table 2: Example of the First, Second, and Final Releases: Chinese Real GDP Growth Rates

Activity Date	Yearbook Date	First Release	Second Release	Final Release
…	…	…	…	…
2007	2008	11.9	13.0	14.2
2008	2009	9.0	9.6	9.6
2009	2010	9.1	9.2	9.2
2010	2011	10.4	10.4	10.4

[*] The full dataset is available at http://home.gwu.edu/~tsinc/china_rtd.xlsx
[†] The 2012 vintage of data come from the NBS website as the yearbook has not yet been released.

Table 3: Descriptive Statistics of First Release Data (Growth Rates)

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Mean	9.60	9.66	4.15	11.32	8.71
Median	9.50	9.50	4.02	10.95	8.22
Maximum	13.41	15.11	7.49	17.38	12.55
Minimum	7.10	3.30	2.27	8.10	7.30
Std. Dev.	1.67	2.62	1.36	2.29	1.36
Skewness	0.49	-0.01	0.52	0.90	1.28
Kurtosis	2.70	3.40	2.70	3.73	4.31
Observations	19	30	27	18	19
Sample	1993-2011	1982-2011	1985-2011	1994-2011	1993-2011

Table 4: Descriptive Statistics of Latest Available Data (Growth Rates)

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Mean	9.93	9.92	4.61	11.50	10.92
Median	9.63	9.62	4.26	11.11	10.40
Maximum	15.18	15.25	12.95	21.14	19.29
Minimum	3.83	4.09	-1.41	1.87	2.32
Std. Dev.	2.75	2.75	2.69	4.44	3.42
Skewness	-0.17	-0.10	1.14	0.01	0.22
Kurtosis	2.92	2.85	5.52	3.11	3.84
Observations	33	33	33	33	33
Sample	1979-2011	1979-2011	1979-2011	1979-2011	1979-2011

Table 5: Descriptive Statistics of First Revisions (in percentage points)

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Mean (Newey-West SE)	0.27** (0.10)	0.24* (0.12)	-0.09 (0.11)	0.09 (0.12)	0.73*** (0.11)
Mean Absolute Revision	0.28	0.66	0.18	0.26	0.73
Median	0.16	0.19	0.00	0.00	0.60
Maximum	1.10	4.89	0.68	1.27	1.74
Minimum	-0.04	-1.83	-2.99	-0.30	0.09
Std. Dev.	0.34	1.14	0.61	0.45	0.49
Skewness	1.09	2.03	-4.20	1.62	0.45
Kurtosis	3.08	11.03	20.80	4.52	2.00
Observations	18	29	26	17	18
Sample	1993-2010	1982-2010	1985-2010	1994-2010	1993-2010

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 6: Descriptive Statistics of Final Revisions (in percentage points)

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Mean (Newey-West SE)	0.82*** (0.17)	0.65*** (0.15)	-0.18 (0.20)	0.19 (0.18)	2.25*** (0.19)
Mean Absolute Revision	0.82	0.89	0.27	0.37	2.25
Median	0.63	0.62	0.00	0.01	2.21
Maximum	2.23	1.99	0.68	1.67	3.43
Minimum	0.03	-1.16	-4.55	-0.33	0.76
Std. Dev.	0.55	0.87	0.95	0.56	0.79
Skewness	1.10	-0.34	-4.33	1.46	-0.29
Kurtosis	3.89	2.42	20.64	4.44	2.26
Observations	16	27	24	15	16
Sample	1993-2008	1982-2008	1985-2008	1994-2008	1993-2008

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 7: Descriptive Statistics of U.S. Real GDP Data (Growth Rates)

	First Release	Latest Available
Mean	2.40	2.44
Median	2.95	2.80
Maximum	4.30	4.80
Minimum	-2.60	-3.50
Std. Dev.	1.86	2.00
Skewness	-1.34	-1.39
Kurtosis	4.08	4.92
Observations	20	20
Sample	1991-2010	1991-2010

Table 8: Descriptive Statistics of US Real GDP Revisions (in percentage points)

	First Revision	Final Revision
Mean (Newey-West SE)	-0.11 (0.11)	0.10 (0.10)
Mean Absolute Revision	0.34	0.52
Median	-0.14	-0.01
Maximum	0.68	0.98
Minimum	-0.88	-0.85
Std. Dev.	0.42	0.61
Skewness	0.07	-0.03
Kurtosis	2.73	1.67
Observations	19	19
Sample	1991-2009	1991-2009

Table 9: Relative Comparison of China and U.S. Real GDP Mean Absolute Revisions

	U.S. First Revision	China First Revision	U.S. Final Revision	China Final Revision
Mean absolute Revision/Mean Initial Release	0.14	0.03	0.22	0.09

Table 10: Test of Rationality for First Revisions

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Constant (Newey-West SE)	0.19* (0.10)	0.34* (0.17)	0.01 (0.03)	0.01 (0.11)	0.83*** (0.25)
Coefficient on lag (Newey-West SE)	0.32 (0.23)	-0.47*** (0.14)	-0.12*** (0.01)	0.20 (0.17)	-0.10 (0.20)
p-value of F-test	0.01**	0.01**	>0.01***	0.52	>0.01***
Sample	1994-2010	1983-2010	1986-2010	1995-2010	1994-2010

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 11: Test of Rationality for Final Revisions

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Constant	0.54**	0.66***	0.01	0.08	2.03***
(Newey-West SE)	(0.20)	(0.12)	(0.03)	(0.14)	(0.35)
Coefficient on lag	0.35	-0.08	0.02	0.36*	0.08
(Newey-West SE)	(0.22)	(0.16)	(0.02)	(0.18)	(0.17)
p-value of F-test	>0.01***	>0.01***	0.50	0.14	>0.01***
Sample	1994-2008	1983-2008	1986-2008	1995-2008	1994-2008

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 12: Modified Holden Peel (1990) with Slowdown Dummy for First Revisions

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Constant	0.30***	0.30	-0.11	0.11	0.75***
(Newey-West SE)	(0.09)	(0.18)	(0.14)	(0.13)	(0.14)
Coefficient on dummy	-0.14	-0.30	0.08	-0.10	-0.10
(Newey-West SE)	(0.17)	(0.37)	(0.15)	(0.21)	(0.20)
p-value of F-test	0.02**	0.15	0.46	0.66	>0.01***
Sample	1993-2010	1982-2010	1985-2010	1994-2010	1993-2010

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 13: Modified Holden Peel (1990) with Slowdown Dummy for Final Revisions

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Constant	0.96***	0.68***	-0.22	0.26	2.52***
(Newey-West SE)	(0.18)	(0.20)	(0.25)	(0.21)	(0.18)
Coefficient on dummy	-0.56***	-0.15	0.17	-0.25	-1.06**
(Newey-West SE)	(0.16)	(0.22)	(0.26)	(0.23)	(0.38)
p-value of F-test	>0.01***	>0.01***	0.25	0.41	>0.01***
Sample	1993-2008	1982-2008	1985-2008	1994-2008	1993-2008

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 14: Regression of First Revisions on First Release

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Constant (Newey-West SE)	-0.53 (0.39)	0.23 (0.48)	0.55 (0.51)	-0.97*** (0.29)	0.14 (0.57)
Coefficient on first release (Newey-West SE)	0.08* (0.04)	0.00 (0.06)	-0.16 (0.15)	0.09*** (0.03)	0.07 (0.06)
p-value of F-test	0.01**	0.07*	0.55	0.01**	>0.01***
Sample	1993-2010	1982-2010	1985-2010	1994-2010	1993-2010

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 15: Regression of Final Revisions on First Release

	GDP	GNP	PRIMARY	SECONDARY	TERTIARY
Constant (Newey-West SE)	-0.42 (0.93)	0.94 (0.73)	0.81 (0.69)	-1.10*** (0.29)	0.22 (0.96)
Coefficient on first release (Newey-West SE)	0.13 (0.10)	-0.03 (0.07)	-0.24 (0.21)	0.11*** (0.03)	0.24** (0.10)
p-value of F-test	>0.01***	>0.01***	0.51	>0.01***	>0.01***
Sample	1993-2008	1982-2008	1985-2008	1994-2008	1993-2008

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 16: Summary of Well-Behaved Revisions Tests for First Revisions

	China					US
	GDP	GNP	PRIMARY	SECONDARY	TERTIARY	GDP
Mean (Newey-West SE)	0.27** (0.10)	0.24* (0.12)	-0.09 (0.11)	0.09 (0.12)	0.73*** (0.11)	-0.11 (0.11)
sd(revision) sd(second release)	0.17	0.38	0.26	0.14	0.32	0.21
Predictability of Revision?	First release*	Lag of revision***	Lag of revision***	First release***	n/a	n/a

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Table 17: Summary of Well-Behaved Revisions Test for Final Revisions

	China					US
	GDP	GNP	PRIMARY	SECONDARY	TERTIARY	GDP
Mean (Newey-West SE)	0.82*** (0.17)	0.65*** (0.15)	-0.18 (0.20)	0.19 (0.18)	2.25*** (0.19)	0.10 (0.10)
sd(revision) / sd(final data)	0.20	0.32	0.35	0.13	0.23	0.31
Predictability of Revision?	n/a	n/a	n/a	Lag of revision* First release***	First release**	Lag of revision**

*, **, *** Statistically significantly different from zero at the 10%, 5%, and 1% level respectively

Figure 1: Chinese Real GDP

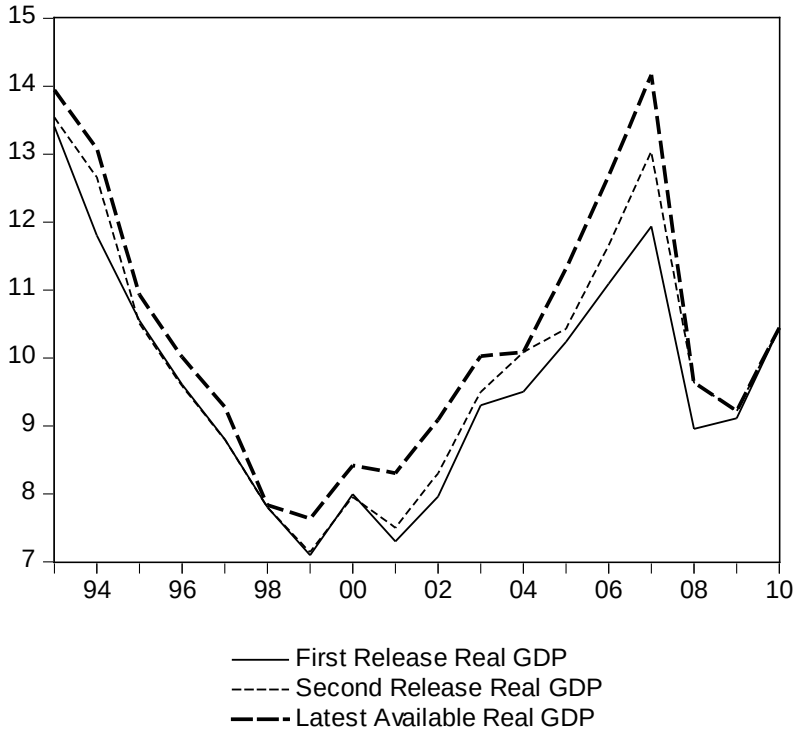


Figure 2: Chinese Real GNP

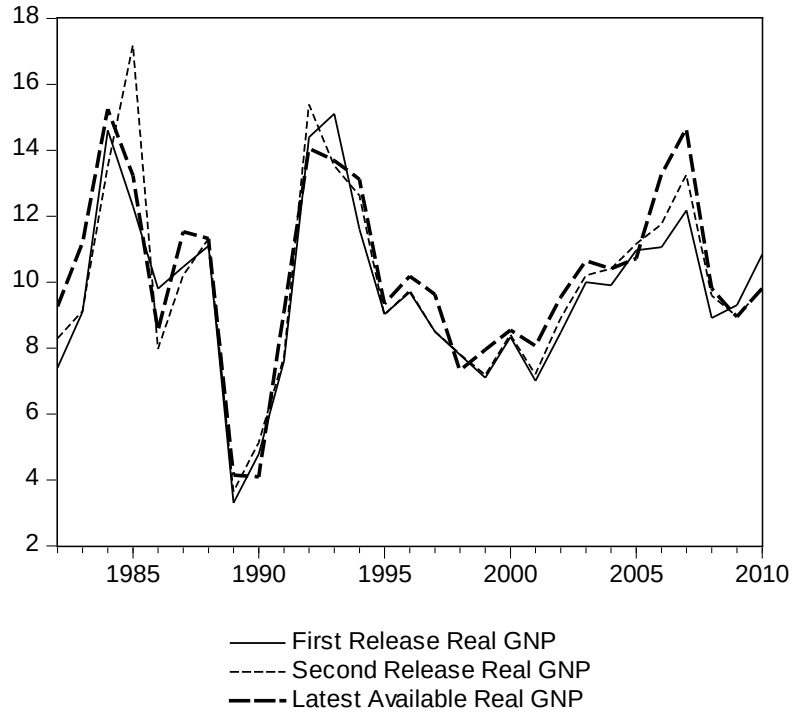


Figure 3: Chinese Primary Sector

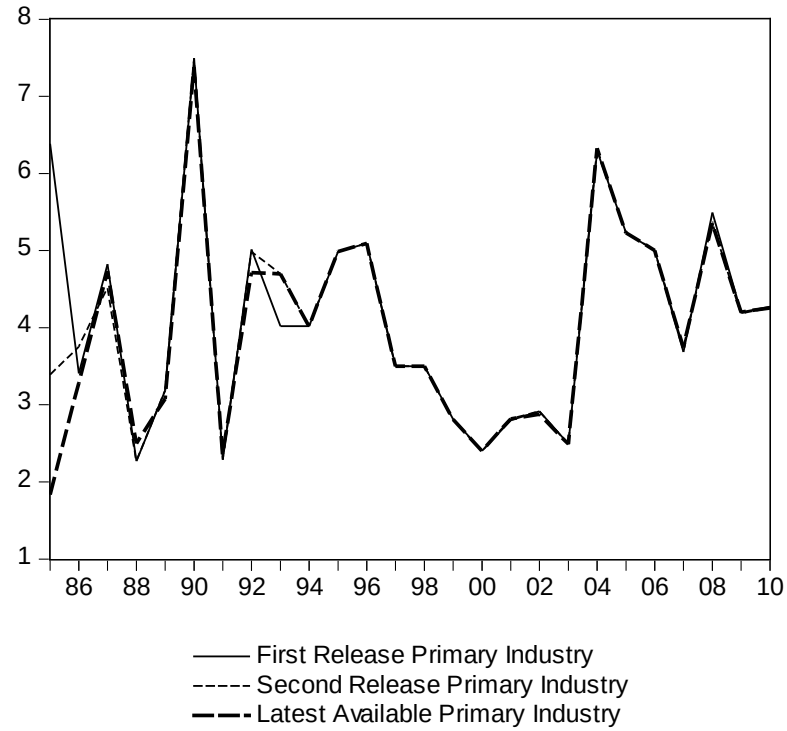


Figure 4: Chinese Secondary Sector

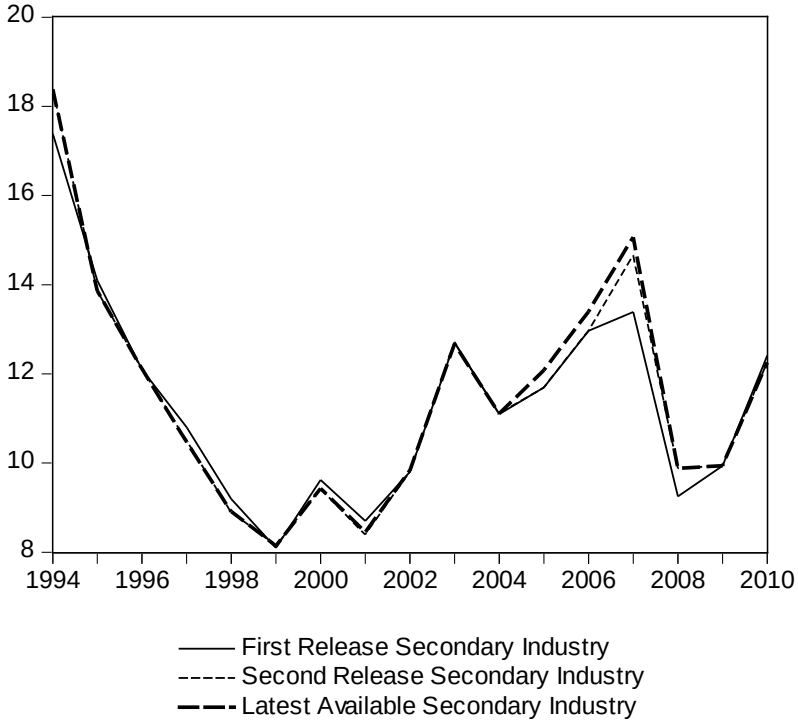


Figure 5: Chinese Tertiary Sector

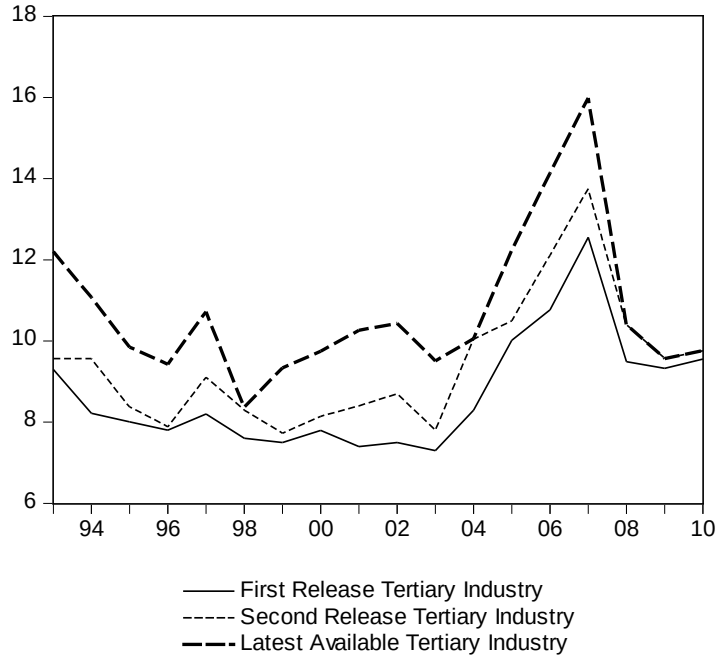
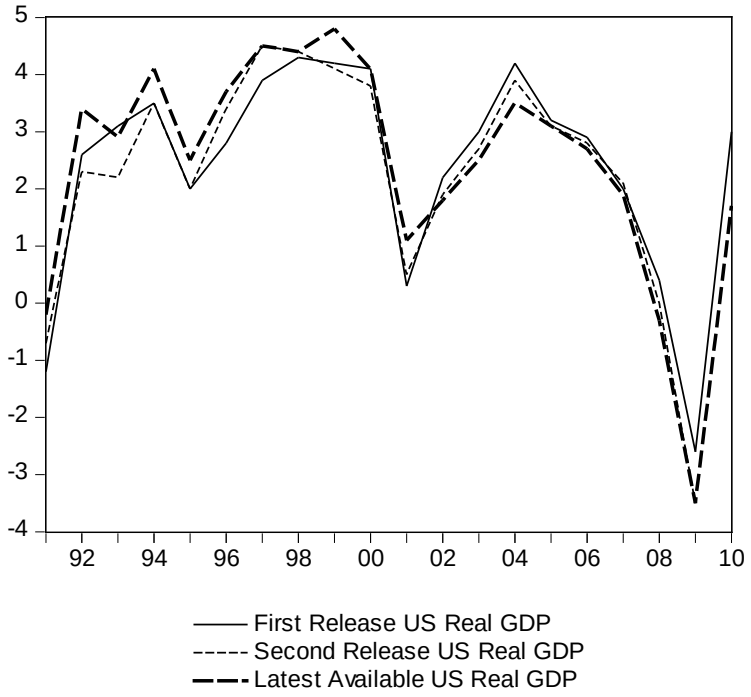


Figure 6: U.S. Real GDP



References

Ashiya, Masahiro. 2005. “Testing Rationality of Forecast Revision made by IMF and OECD.” Kobe University Discussion Paper No. 431 / 2005-02.

Berger, Allen N. and Spencer D. Krane. 1985. “The Informational Efficiency of Econometric Model Forecasts.” *The Review of Economics and Statistics*, Vol. 67, No. 1, pp. 128-134.

Bradsher, Keith. 2011. “China, Driver of World Economy, May be Slowing” *The New York Times*, September 23, 2011.

China Statistical Yearbook, various years, China Statistics Press. PDF files are available, with appropriate subscription, through http://chinadataonline.org/

Chow, G. 2006. “Are Chinese Official Statistics Reliable?” *CESifo Economic Studies*, Vol.52 2/2006, 396-414.

Croushore, D. 2011. "Frontiers of Real-Time Data Analysis," *Journal of Economic Literature*. 49(1): 72–100.

Croushore, D. 2008. "Revisions To PCE Inflation Measures: Implications for Monetary Policy" Philadelphia Fed working paper 08-8.

Croushore, D. and T. Stark. 2001. “A Real-Time Dataset for Macroeconomists.” *Journal of Econometrics* 105 (2001): 111-130.

Curtis, C. C. and N. Mark. 2010. “Business Cycles, Consumption and Risk-Sharing: How Different Is China?” NBER Working Paper No. 16154.

Dynan, K. E. and D. W. Elmendorf. 2001. "Do Provisional Estimates of Output Miss Economic Turning Points?" Federal Reserve Board Finance and Economics Discussion Series No. 2001-52.

The Economist. 2008. “An Aberrant Abacus: Coming to Terms With China’s Untrustworthy Economic Numbers” Finance and Economics. Economics Focus. May 1st 2008.

Heston, A. 2001. “Treatment of China in PWT 6” http://pwt.econ.upenn.edu/Documentation/China.pdf

Holden, K., & Peel, D. A. 1990. “On testing for unbiasedness and efficiency of forecasts.” *Manchester School*, 48, 120–127.

Holz, C. 2003. “’Fast, Clear and Accurate:’ How Reliable Are Chinese Output and Economic Growth Statistics,” *The China Quarterly*, 173: 122-163.

Holz, C. 2006. "China's Reform Period Economic Growth: How Reliable Are Angus Maddison's Estimates? Response to Angus Maddison's Reply." *Review of Income and Wealth,* 52(3): 471-5.

Holz, C. 2008. "China's 2004 Economic Census and 2006 Benchmark Revision of GDP Statistics: More Questions than Answers?" *The China Quarterly*, 193: 150-163.

Huang, Y., 2011. "Misinterpreting China's Economy" *Wall Street Journal*, August 25, 2011.

Jia, Y. 2011. "A New Look at China's Output Fluctuations: Quarterly GDP Estimation with an Unobserved Components Approach" George Washington University Research Program on Forecasting Working Paper No. 2011-006.

Joutz, Frederick & H. O. Stekler, 1998. "Data revisions and forecasting," *Applied Economics*, vol. 30(8), pages 1011-1016.

Keidel, Albert. 2001. "China's GDP Expenditure Accounts." *China Economic Review* 12 (4):355–67.

Keidel, Albert. 2007. "The Limits of a Smaller, Poorer China" *Financial Times*, November 14, 2007.

Kennedy, James E. 1990. "An Analysis of Revisions to the Industrial Production Index." Board of Governors of the Federal Reserve System, Economic Activity Section, Working Paper 109.

Klein, L. R. and S. Ozmucur. 2003. "The estimation of China's Economic growth rate" *Journal of Economic and Social Measurement*, 28(4): 187-202.

Lequiller, F. and D. Blades. 2007. "Chapter 13: China's National Accounts" *Understanding National Accounts***,** OECD, Pages: 365–382.

Lin, Justin Yifu. 2011. "China and the Global Economy" Luncheon Address at the 2011 Asia Economic Policy Conference at the Federal Reserve Bank of San Francisco. Transcript available at: http://www.frbsf.org/economics/conferences/aepc/2011/volume/Lin.pdf

Maddison, A. 1998. *Chinese Economic Performance in the Long Run*. Paris: Development Centre of the Organization for Economic Co-operation and Development.

Maddison, A. 2006. "Do Official Statistics Exaggerate China's GDP Growth? A Reply to Carsten Holz." *Review of Income and Wealth* 52 (1):121–26.

McCully, Clinton P. and Steven Payson. 2009. "Preview of the 2009 Comprehensive Revision of the NIPAs: Statistical Changes" *Survey of Current Business*, May, 2009, Volume 89, No. 5, Pages 6 – 16.

Newey, Whitney K. a dn Kenneth D. West. 1987. “A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica*, 55, 703-8.

Orlik, T. 2012. “Debating Lies, Damned Lies, and Chinese Economic Statistics.” *The Wall Street Journal* Thursday, July 26, 2012: C12.

Rawski T. G.1976. “On the reliability of Chinese economic data” *Journal of Development Studies*, 1743-9140, Volume 12, Issue 4, 1976, Pages 438 – 441.

Rawski, T. G. 2001. "What is happening to China's GDP statistics?" *China Economic Review*, 12(4): 347-354.

Rawski T. G and W. Xiao, 2001 “Round**Table** on Chinese Economic Statistics Introduction” *China Economic Review*, 12(2001) 298-302.

Ren, Rouen. 2002. “How Inflated Is China’s GDP Statistics?—Comment on Methodological Issues of Two Studies on China’s GDP Statistics.” *Economics Quarterly* 2 (1):37–52.

Scheibe, J. 2003. “Chinese Output Gap During the Reform Period 1978-2002” Department of Economics Discussion Paper No. 179, University of Oxford, December.

Sinclair, T. M. and H.O. Stekler, 2012. “Examining the quality of early GDP component estimates,” *International Journal of Forecasting,* forthcoming.

Sinclair, T. M., Joutz, F., & Stekler, H. O. 2010. “Can the Fed predict the state of the economy?” *Economics Letters*, 108(1), 28–32.

Stekler, H. O. 1967. “Data revisions and economic forecasting.” *Journal of the American Statistical Association*, 62, 470–483.

Swanson, N. R. and D. van Dijk. 2006. "Are Reporting Agencies Getting It Right? Data Rationality and Business Cycle Asymmetry. " *Journal of Business and Economic Statistics*, 24 (January 2006), pp. 24-42.

Wu, H. X. 2000. “China's GDP level and growth performance: alternative estimates and the implications.” *Review of Income and Wealth*, 46(4).

Wu, H. X. 2007. “The Chinese GDP Growth Rate Puzzle: How Fast Has the Chinese Economy Grown?” *Asian Economic Papers* 6(1): 1-23.

Xu, Xianchun. 2002.” Study on Some Problems in Estimating China’s Gross Domestic Product.” *Review of Income and Wealth* 48 (2):205–15.

Young, A. 2003. "Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period," *Journal of Political Economy*, 111(6): 1220-1261.