

## **Deconstructing China's GDP Statistics**

China's National Bureau of Statistics has on repeated occasions explained in great detail how its GDP statistics are derived from underlying data. Based on these explanations, this article reconstructs Chinese official household consumption, which accounts for half of GDP. The findings are condemning. Not only do the various official explanations offered between 1997 and 2001 differ from each other, but none allows the researcher to accurately reconstruct household consumption. The relationship between the GDP component household consumption and the underlying data, furthermore, varies from year to year, which suggests that time series comparisons of Chinese GDP may be invalid.

JEL classification:

C80 – Data Collection and Data Estimation Methodology; Computer Programs --- General

O53 - Asia including Middle East (within O5 - Economywide Country Studies)

P24 - National Income, Product, and Expenditure; Money; Inflation (within P2 – Socialist Systems and Transitional Economies)

P30 - Socialist Institutions and Their Transitions --- General

Keywords: Chinese statistics, national income accounting, System of National Accounts (SNA), household surveys, population statistics

Carsten A. HOLZ

E-mail: cholz@stanford.edu

(Visiting Scholar at the Asia/Pacific Research Center, Stanford University, 8/03–7/04)

*Mailing address:*

Carsten Holz, Asia/Pacific Research Center, Stanford University, 616 Serra Street, Encina Hall  
E301, Stanford, CA 94305-6055, USA  
Tel. (650) 724-9632

(Associate Professor, Social Science Division, Hong Kong University of Science & Technology,  
Clear Water Bay, Kowloon, Hong Kong)

28 October 2003

## **Deconstructing China's GDP Statistics**

Revised version, 28 October 2003

### **1. Introduction**

Chinese statistics have in recent years come under increasing criticism, with researchers questioning the meaning of individual statistics and raising the specter of widespread data falsification.<sup>1</sup> Three books published by China's National Bureau of Statistics (NBS) between 1997 and 2001 explain how China's GDP is calculated. This article takes the official instructions on how GDP is calculated at face value and tries to implement them in an attempt to reconstruct the household consumption component of GDP from the underlying data. Household consumption accounts for approximately one half of GDP throughout the reform period.<sup>2</sup>

The results are astounding. First, the different official explanations on how the various (household) consumption components are calculated are rarely identical; occasionally they are inappropriate on logical grounds. Second, implementing the (differing) official explanations on how to derive the various consumption components from the underlying data, the outcome of the calculations rarely matches the official consumption data. Third, time series comparisons of the consumption data with the underlying data reveal roller-coaster relationships that stretch credulity.

The implications are severe. If the underlying data were correct, the published GDP value of any particular year is not comparable to that of any other year as the compilation process—of at least consumption—has not been held constant, and changes in the compilation process do not lead to retrospective revisions of earlier published data. Thus, an official annual GDP growth rate of X% in a particular year does not imply that final demand (the sum of consumption, investment, and net exports) in that year is X% higher than in the previous year.

That the publicized procedures for the derivation of expenditure approach GDP in the case of household consumption do not allow the reconstruction of official data violates basic principles of official statistics as outlined in 1994 by the United Nations Statistical Commission: “To retain trust in official statistics, the statistical agencies need to decide according to strictly professional

considerations, ... The laws, regulations and measures under which the statistical systems operate are to be made public.”<sup>3</sup> The fact that the procedures published in an internal GDP manual are of no further help suggests that clear-cut procedures may not even exist within the NBS, or that procedures are being continuously modified.

The focus of this article is household consumption as the major component of expenditure approach GDP, rather than total GDP, for two reasons. First, space constraints do not allow the in-depth examination of more than this largest single component of expenditure approach GDP. Second, in the case of household consumption the underlying data are somewhat accessible, in contrast to other expenditure approach GDP components or production approach GDP components.

The following section introduces basic definitions and data sources. The third section contrasts the various components of household consumption as part of GDP with the reportedly underlying data. The fourth section points out some of the implications of the problems in reconstructing the official data, and the fifth section elaborates on an additional complication due to the population data used for aggregation. The last section concludes.

## 2. Data

GDP in the national income accounts (NIA) can be calculated as the sum of household and government consumption, investment, and net exports. Details on the derivation of the individual items are provided in (i) an NBS compendium on how to calculate GDP in China (NBS, 1997), (ii) a book by Xu Xianchun (2000), who is currently the head of the National Income Accounts Division of the NBS, on the calculation of GDP in China, and (iii) an internal year 2001 handbook of the NBS on the calculation of GDP in China, in the following labeled *GDP Manual*.<sup>4</sup> Of the three sources, the internal *GDP Manual* is the shortest and least comprehensive. Its content does not differ systematically from that of the public versions; however, the process of data compilation at times appears simplified in that certain items supposedly included in GDP according to the public sources are omitted.

NIA household consumption comes with a breakdown into the two groups rural and urban. Data for each group can be further broken down into four (three) exhaustive categories:

“own account” consumption (in the case of rural households only),  
commodity consumption,  
consumption of cultural and personal services, and  
housing (housing, water and electricity).

NIA household consumption is derived from rural and urban household surveys and retail sales data. The rural and urban household surveys each offer numerous categories. There is no direct correspondence between NIA household consumption categories and household survey categories, neither for the rural nor for the urban case.

NIA household consumption data, rural and urban, aggregate and per capita (needed for the comparison with the per capita household survey data), are provided in two sources. One source is the *Statistical Yearbook* beginning with the 1995 issue which published retrospective aggregate data for the years since 1978, plus the *Statistical Abstract* (a preliminary publication of previous-year data and revised selected earlier data, issued approximately five months before the comprehensive *Statistical Yearbook* of that year) to fill some missing values in the *Statistical Yearbook*. The *Statistical Yearbook* and the *Statistical Abstract* do not provide data on the individual four (three) components of NIA household consumption.<sup>5</sup>

The second source of NIA household consumption data is *GDP 1952-96*, a compendium published in 1998.<sup>6</sup> Its per capita (as well as aggregate) household consumption data are identical to those published in the *Statistical Yearbook* for the years 1978 through 1993 but slightly larger in the years 1994-96 (for all households, rural households, and urban households). There is no explanation of the difference in the years 1994-96. *GDP 1952-96* contains data on the four (three) components of NIA household consumption for the years 1978 through 1996.

Figure 1 and Figure 2 show the relative shares of the four (three) categories in total rural and urban NIA consumption from 1978 until 1996. In the case of both rural and urban household consumption, housing accounts for only a very small share of total consumption, at 5% or less throughout the reform period. Cultural and personal services accounted for less than 10%

throughout the 1980s, reaching 11.55% in the rural case by 1996 and 18.37% in the urban case. The largest component is commodity consumption. In the urban case, throughout the 1980s, almost 90% of total household consumption consisted of commodities, falling to 76.49% in 1996. In the rural case, commodity consumption ranged from a minimum of 45.75% of total household consumption in 1978 to a maximum of 67.29% in 1988, with a final value of 56.21% in 1996. In the rural case, own account consumption constituted the rest, with a maximum of 47.25% in 1978 and a minimum of about 25% in the early 1990s. Commodity consumption (plus own account consumption in the rural case) thus are the key factors driving NIA household consumption and ultimately expenditure approach GDP.

*Figure 1 and Figure 2 about here*

Household survey data come according to two different classifications. One distinct rural and one distinct urban classification system were in use in the 1980s (in the rural case until 1992, and in the urban case until 1991). A new, identical classification system for both rural and urban areas was introduced in the early 1990s (subject to minor differences in labeling), with limited retrospective data for the early reform years. A distinction is always made between monetary and non-monetary expenses, with, for rural areas, data published on total (monetary and non-monetary) expenses as well as, separately, on monetary expenses, while for urban areas data are published only on the monetary expenses.

Table 1 provides details on the different classifications (early vs. recent, rural vs. urban) with the (exhaustive) main categories within each classification on which data are regularly reported. Table 2 expands on the housing-related categories in the household surveys, with no data available for the various sub-categories. (“No data available,” here and below, means that the author is not aware of any publication containing these data.) These tables will be frequently referred to below.

*Table 1 and Table 2 about here*

Household survey data are obtained from three different sources. The key source is the livelihood section of the *Statistical Yearbook*. The most recent *Rural Survey Yearbook 2002*

reports data from 1978 through 2001 according to the new classification; it is the only source for retrospectively reorganized data (into the new classification) for all years since 1978, and is consulted for the years 1979, 1981-84, and 1986-89, for which the *Statistical Yearbook* does not provide data according to the new classification.<sup>7</sup> The *Urban Survey Yearbook* provides a more detailed breakdown with consumption sub-categories for urban household surveys, which is useful at one occasion below. Retail sales data are drawn from the *Statistical Yearbook*.

### **3. NIA consumption components and underlying household survey data**

Aggregate NIA household consumption comprises “own account” consumption (in the case of rural households only), commodity consumption, consumption of cultural and personal services, and housing (housing, water and electricity). Each of these NIA household consumption components is officially derived from, or with the help of, household survey data. Data on the individual NIA household consumption components have been published for the years 1952 through 1996. The focus here is on the reform period (the years since 1978).

#### *3.1. Rural own account consumption*

The rural household (living expenditure) survey consists of two surveys, one of total consumption, and one of consumption paid for using money. According to NBS (1997, p. 142), the difference between total and monetary consumption constitutes “self-provided consumption” (*zijixing xiaofei*). This difference is virtually zero in the four household survey categories medicine and health care; transport, post and telecommunications; culture, education and recreation; and other goods and services. (For the household survey categories see Table 1.) As regards the four remaining categories, where the difference between total and monetary consumption is significant, some sub-categories have to be subtracted out to obtain own account consumption. This leads to the following definition of NIA rural own account consumption by the NBS in terms of rural *non-monetary* household survey categories (NBS, 1997):

foods, less the sub-category food processing costs (*shipin jiagong fei*),

- + clothing, less the sub-category clothing processing and mending costs (*yizhuo jiagong xiulifei*)
- + fuels within the category residence, i.e., residence, less the sub-categories construction materials, rent, electricity, housing repair services, and other residence expenses,<sup>8</sup>
- + household equipment, less the sub-category household equipment processing and repair costs (*shebei yongpin jiagong xiulifei*).

In the first, second, and fourth groups of commodities, the items to be subtracted out are service components—the only service components in these categories—and they should be exceedingly small. The third group, fuels, in the recent household survey classification is a sub-category of residence, and data are not available for sub-categories. In the early household survey classification, fuels constitutes a separate category of its own.

In contrast to NBS (1997), Xu Xianchun (2000, p. 62) simply defines NIA own account consumption as comprising the difference between total and monetary expenditures for the four rural household survey categories food, clothing, residence, and household equipment. He does not subtract out sub-categories, nor reduces the category residence to only fuels. The *GDP Manual* (p. 78) defines own account consumption as total minus monetary consumption. The coverage is not limited to four categories, and the category residence is not reduced to fuels.<sup>9</sup>

Table 3 contrasts these definitions with the actual data. If the NBS (1997) definition were correct, the ratio of NIA rural own account consumption to the sum of food, clothing, equipment and an approximation of fuels (derived reliably as outlined in the notes to the table) should be close to unity. It could be slightly below unity since the data are not available to subtract out the service components from the categories food, clothing, and equipment. Using the data from the household survey with the recent classification, the ratio indeed was 0.98 in 1978, but then declined almost continuously to 0.83 in 1986 before rising again to 0.90 in 1989, falling drastically in 1990 to 0.74, and then jumping from 0.79 in 1994 to 1.00 in 1995.

*Table 3 about here*

Variations in the ratio of up to 25% (between 1994 and 1995) are implausible. The 1995 and 1996 values of unity suggest that the definition is correct in these years (except that services may in practice not be subtracted out, or are very small), but incorrect in all previous years. The wide

swings in the ratio prior to 1995 suggest that the definition has changed repeatedly. These swings are not without effect on GDP. For example, if the ratio in 1995 had been the same as in 1994, namely 0.79, nationwide expenditure approach GDP in 1995 would have been 1.49% lower.<sup>10</sup> In as far as a ratio below unity suggests that NIA rural own account consumption is underestimated, official 1994 GDP reflects a corresponding underestimate of the true GDP value and the official 1995 GDP growth rate an overestimation of the actual growth rate.

Table 3 also reports the ratios corresponding to Xu Xianchun's definition (according to which NIA rural own account consumption includes all of the category residence, not just fuels) and to the *GDP Manual*'s definition (which is based on all eight non-monetary rural household survey consumption categories). The results are virtually identical since the non-fuel part of the category residence as well as the other four categories are negligibly small. It is only when fuels are arbitrarily (and mistakenly) excluded, in violation of all three definitions, that the ratio moves closer to unity (but then exceeds unity in 1995 and 1996).

The early household survey distinguishes between commodity consumption and service consumption in the first instance, and then for commodity consumption between monetary and non-monetary commodity consumption (Table 1); i.e., non-monetary consumption by definition covers only goods.<sup>11</sup> The five (non-monetary) commodity consumption categories are foods, clothing, housing, fuels, and articles for use. Because NBS (1997) explicitly includes fuels but rules out other items that are comprised in housing in the recent classification (Table 2), NIA own account consumption should match the sum of the four non-housing categories in the early classification, namely the sum of foods, clothing, fuels, and articles for use. Table 3 shows this ratio to be almost identical over time to the one in the case of the recent classification, with the same wide swings. Trying similar variations as in the case of the recent classification, to match the other definitions, also yields no new results.

One additional complication is the fact that the rural household survey following the early classification offers two different sets of data points for the year 1990, one published in 1991, and the other published in all later years. In 1990, the prices applied to self-provided consumer

goods (but not services) were revised. To obtain the first set of 1990 values, survey teams applied state procurement prices to all types of self-provided goods for which state procurement existed, and (large-quantity) market prices to all types of self-provided goods for which there was no state procurement. The second set of 1990 values is based on *comprehensive contract prices* for all types of self-provided goods for which state procurement existed, with the comprehensive contract price being a combination of the quota procurement price and the above-quota procurement price, and (large-quantity) market prices for all types of self-provided goods for which there was no state procurement.

Presumably all pre-1990 survey values of self-provided goods were compiled using the same method as that used for the first published 1990 data. This implies that if rural household survey values prior to 1990 indeed constituted the basis for the calculation of NIA rural own account consumption, GDP in all years prior to 1990 is likely to have been underestimated in as far as imputed prices were below market prices. The application of new prices raises the value of self-provided goods in 1990 by 28.53%, almost completely due to revisions to the value of self-provided *foods*.

Between 1991 and 1997, the NBS compiled survey values of self-provided goods using the method applied in the case of the later published 1990 data.<sup>12</sup> The significantly lower ratio of NIA rural own account consumption to the relevant household survey categories in 1990 using the new survey data—but also in 1991 and 1992 (the last years for which data following the early classification are available)—compared to the earlier years suggests that NIA rural own account consumption data in these years could be based on old imputation prices, thereby underestimating rural consumption (and thus GDP) in the early 1990s. The recent classification also exhibits a statistical break in 1990, with identical ratios as in the case of the early classification for 1990 through 1992, at 0.74, 0.75, and 0.78, and then, further, similarly low values of 0.76 and 0.79 in 1993 and 1994; these ratios are down from 0.90 in 1989 (first data column in Table 3). The fact that the values went up to unity in 1995 suggests that the NIA

imputations may not have adapted the new imputation prices in 1990 but only much later. This would imply that NIA rural own account consumption was underestimated all through 1994.

The NIA own account consumption data, thus, appear riddled with inconsistencies: a questionable time pattern throughout the 1980s, very likely (severe) statistical breaks in 1990 and 1995, and an imperfect match of NIA and survey data in all years except 1995 and 1996. Two clearly identifiable effects are a tendency of official GDP to underestimate actual GDP throughout all years until 1994, and to overestimate GDP growth between 1994 and 1995.

### 3.2. *Commodity consumption*

According to NBS (1997, pp. 140f.) and Xu Xianchun (2000, pp. 61f.), rural and, separately, urban NIA (household) commodity consumption can be obtained in two ways. The first is to add up the various commodity categories in the rural (monetary) and urban household surveys, and to multiply the per capita values by the rural and urban population numbers to obtain aggregate rural and urban commodity consumption. The second is to set aggregate NIA commodity consumption equal to the economy-wide value of retail sales of consumer goods to residents; in a separate step, this value can be split into a rural and an urban component by using the ratio of aggregate rural to urban household survey values of commodities. Relying solely on household survey data (the first method) is in both sources reported to yield values which are “too small.” Therefore the NBS relies on the second method, setting NIA commodity consumption equal to the retail sales of commodities to residents (*jumin*) for consumption. These are obtained as:

- total retail sales of consumer goods (which include sales between residents)
- retail sales of consumer goods to social entities (*shehui jituan*)
- construction materials used by residents to construct or renovate houses.<sup>13</sup>

In order to derive the values of the rural and urban components, aggregate NIA commodity consumption is split according to the relative size of rural vs. urban commodity consumption in the household surveys. To derive the rural share, in a first step rural per capita household consumption of commodities is obtained from the rural *monetary* household survey by

subtracting the various services from the eight categories and then adding up the remaining commodity components in these categories. According to NBS (1997, pp. 140f.),<sup>14</sup> the relevant rural commodity consumption items in the rural household survey are

category foods: foods – food processing costs,  
category clothing: clothing – clothing processing and mending costs,  
category residence: construction materials used to repair housing + fuels,  
category equipment: equipment – equipment processing and repair costs – “others,”  
category medicine and health care: medicine and health care goods + “others,”  
category transport, post and telecommunications: transport equipment and  
telecommunications equipment,  
category culture, education and recreation: culture, education and recreation goods, and  
category “others:” (other) commodity expenses.

The relevant *urban* commodity consumption items according to NBS (1997, p. 141) and Xu Xianchun (2000, p. 63) are derived in similar fashion (with no further details provided).

The explanations in Xu Xianchun (2000, pp. 61-63) differ in two respects. In the case of residence, fuels are excluded rather than included.<sup>15</sup> In the case of household equipment, the sub-category “others” is not subtracted out. The *GDP Manual* (pp. 75-8, 81-84) contains the same list as NBS (1997) and also explicitly lists the individual relevant items for the *urban* case, where the exact classification within categories sometimes differs slightly from that in the rural household survey (the detailed urban list is omitted here).

One significant difference between NBS (1997) and the *GDP Manual* is that according to the latter, in the calculation of retail sales of commodities to residents for consumption, *only* retail sales of consumer goods to social entities are subtracted from total retail sales of consumer goods (not mentioned/ subtracted are construction materials used by residents to construct or renovate houses). This is furthermore done by reducing the amount of total retail sales of consumer goods by 15%, a value gained from past experience, with “appropriate adjustments according to the actual situation” in each year. NBS (1997, p. 141), in contrast, provides an estimate of 9-11% for the share of retail sales of consumer goods to social entities; the figure is presented as a descriptive estimate rather than as a ratio to be exactly applied in the derivation process.

The household survey data are per capita values. Aggregate rural and urban household survey consumption of goods are obtained by multiplying with the rural and urban population numbers. In a final step, the NBS applies the relative shares of aggregate rural and urban household survey consumption of goods to total retail sales of consumer goods to residents in order to split the latter into the rural and urban component of NIA commodity consumption.

Table 4 first contrasts aggregate NIA household commodity consumption with an approximation of retail sales of commodities to residents, and then NIA urban (rural) household commodity consumption with the corresponding household survey values (which were multiplied by the population data implicit in the official aggregate and per capita NIA consumption data in order to obtain comparable, aggregate values). Since no data are available on retail sales of construction materials used by residents to construct or renovate houses, retail sales of commodities to residents for consumption are calculated in two ways. One way is to only subtract the retail sales of consumer goods to social entities from total retail sales of consumer goods—as outlined in the *GDP Manual* but in contrast to NBS (1997) and Xu Xianchun (2000) who further subtract retail sales of construction materials used by residents to construct or renovate houses. Retail sales of consumer goods to social entities accounted for, on average, 10.66% of all retail sales of consumer goods in the years 1978 through 1993, with a standard deviation of 0.4470 percentage points. An alternative calculation is to simply adjust retail sales of consumer goods 15% downward every year to obtain an approximate value of retail sales of consumer goods to residents, as the *GDP Manual* suggests; this percentage being larger than the percentage of retail sales of consumer goods to social entities, it could also cover retail sales of construction materials used by residents to construct or renovate houses.

Economy-wide aggregate NIA household commodity consumption should be approximately identical to retail sales of consumer goods to residents (with the difference possibly accounted for by construction materials used to construct or renovate houses). The ratio of aggregate NIA commodity consumption to total retail sales of consumer goods to residents is indeed close to unity throughout all years. If the retail sales data are correct, NIA commodity consumption data

are accurate. Similarly, a comparison with 85% of all retail sales of consumer goods suggests that NIA consumption data are quite accurate, without any trend over time. A separate comparison for urban and rural areas yields no further insights.<sup>16</sup>

*Table 4 about here*

But while the comparison of NIA commodity consumption to retail sales is favorable, a comparison of the NIA urban and rural commodity consumption data directly to the underlying household survey data is not. Table 4 also reports the ratios of urban (rural) commodity consumption to the consumption of consumer goods in the urban (rural) household survey, based on both the recent and the early classification; various approximations needed to derive household survey commodity consumption data are explained in the notes to Table 4. NIA commodity consumption by far exceeds the underlying urban and rural household survey consumption data for goods.<sup>17</sup> Based on the recent classification, the ratio in urban areas ranges from 1.19 in 1985, the first year for which the data are available, to 1.31 in 1996, the last year for which the data are available; the variation over time appears minor. In the rural case, the ratio varies from 1.14 in 1983 to 1.40 in 1985, ending with 1.21 in 1996. The variation between individual years is not always credible. For example, the rise in the ratio from 1.21 in 1991 to 1.33 in 1992 is not credible. If the 1992 ratio had been the same as in 1991, nationwide expenditure approach GDP in 1992 would have been 1.62 % lower.<sup>18</sup>

The findings are similar when the household survey consumption data based on the early classification are used. One estimate of the household survey consumption value of consumer goods (Version I) due to a lack of more detailed data includes the category housing and construction materials in full (labeled “housing” in the rural case; see Table 2), rather than only construction materials used to *repair* houses. A second estimate excludes this category (Version II); this second estimate is likely to be more accurate since the value construction materials used to repair houses may well be negligibly small. The urban ratio (at between 1.08 and 1.25) is closer to unity than the rural ratio, and with relatively little variation over the years. The rural ratio can vary significantly from year to year, with, in the first version, a low of 1.01 and a high

of 1.23 (ignoring the extraordinarily high outlier value of 1.41 in 1978), and, in the second version, a low of 1.24 and a high of 1.53. Such large variations, in the second version even in the course of just two years, are not credible.

To explain the discrepancy between the retail sales and the household survey data, a direct component-by-component comparison is possible between retail sales of consumer goods to households *and* social entities vs. the household survey data on consumer goods consumption in the years 1981 through 1991. The statistics on retail sales of consumer goods come with a breakdown into product categories similar to those used in the urban and rural household surveys; common categories can be established by reducing the various classifications to four categories, namely foods, clothing, fuels, and all else. Figure 3 shows that the retail sales values of foods, clothing and fuels, which accounted for 55.03%, 16.45%, and 3.07% of total retail sales of consumer goods in 1991, exceed the corresponding values in the household surveys by approximately 30% to 110%. Some annual variations in all three time series are also not credible.

*Figure 3 about here*

The question that arises is which data are more reliable, retail sales data or household survey data? If retail sales data are accurate, then NIA commodity consumption is quite accurate throughout all years. If retail sales data are inaccurate and household survey data are accurate, then NIA commodity consumption is exaggerated at varying degrees over time (with perhaps an upward trend in the urban case). There is no escaping the implication that at least one of these two data series is inaccurate. It would seem that household survey data collected from tens of thousands of households with each household reporting data over a three-year time period should provide ample time consistency, though not necessarily complete coverage if households neglect to report some consumer items. If household survey data are time consistent, then NIA commodity consumption data, at least in the rural case, are in many years time inconsistent. That households routinely consume one-quarter to one-third more consumer goods, in terms of value,

than they report, especially in the rural case, is also not plausible (questioning the level of NIA commodity consumption).

On the other hand, official retail sales data could well reflect an *underestimate* of actual retail sales. First, retail sales are biased downward because they omit sales from farmers to farmers. Second, the 1993 tertiary sector census led to a retrospective upward revision of tertiary sector value-added for earlier years. In the sub-sector wholesale and retail trade plus catering, value-added for 1981, the earliest year for which these data were revised, was retrospectively revised upwards by 12.20%, and for 1993 by 73.40%. The published retail sales data, however, were not revised. If the bias in the value-added of this tertiary sector sub-category of wholesale and retail trade plus catering were representative of a bias in retail trade by itself, and if the bias in (retail) sales were the same as in retail sales value-added, then NIA commodity consumption should be larger than reported at least in the years 1981 through 1993.<sup>19</sup>

In general, retail sales data appear to be a combination of reliable and unreliable data. Thus, the retail sales data on consumer goods sold by state-owned units (for example, state-owned department stores) are likely to be highly reliable, while retail sales of consumer goods by individual-owned units, or by “other” units (which include, and may well largely consist of, retail sales by farmers to non-farmers), are likely to be much less reliable. The state-owned share of retail sales of consumer goods fell from 37.52% in 1993 to 20.65% in 1998 (the only period for which these data are available) while the shares of individual-owned units and other units rose from 28.45% and 15.98% to 37.05% and 25.16%. With less than 40% of retail sales by 1998 occurring in state-owned units, the margin of error in the aggregate retail sales of consumer goods data could be large, and even have become larger over time.<sup>20</sup> The margin of error in the retail sales data translates directly into a corresponding margin of error in the NIA household commodity consumption data.

### *3.3. Cultural and personal service consumption*

NIA cultural and personal services comprise the processing and repair services in the first seven of the eight household survey categories (all categories except “others”),<sup>21</sup> plus further, specific service items in the fifth through eighth household survey categories. (NBS, 1997, pp. 142f.; for the classification see Table 1.) The specific service items are:

- category medicine and health care: services;
- category transport, post and telecommunications: transport, post and telecommunications fees;
- category culture, education and recreation: in the rural case, services; in the urban case, the complete sub-category education (except textbooks and reference books, and “others”) and cultural and entertainment fees (within the culture and entertainment sub-category);<sup>22</sup>
- category “others:” in the rural case, services; in the urban case, the two items travel and services within the sub-category personal consumption, plus the sub-category other services.<sup>23</sup>

The definitions in the three sources are virtually identical with only minor discrepancies; these are more likely due to carelessness than to systematic differences.<sup>24</sup>

Lacking a complete set of detailed data following the recent classification, household survey service consumption is derived by taking that share of an individual household survey category that does not constitute goods, except that the category residence is omitted in total. In the early classification, a distinct category services is available. (For details see notes to Table 5.)

Table 5 reports the ratios of NIA cultural and personal service consumption to the underlying household survey data. In the urban case, following the recent classification, the NIA data exceed the household survey data in all years, with a drastic rise in the ratio between 1985 and 1992, before leveling off at approximately twice the value of the household survey data. The values following the early classification, where the service data in the rural and urban household surveys should be highly reliable since they are explicitly reported in the official statistics, are still higher. In the rural case, following the recent classification, the ratio is slightly above 0.5 in the years 1980 through 1986, then rises continuously to 1.31 in 1996. Following the early classification, the ratios are almost twice as high in the early 1980s, but then fall to reach the same level as that in the recent classification in 1990/1991.

*Table 5 about here*

While in the urban case the ratios across all years consistently exceed unity, in the rural case, thus, they only do so since the early 1990s. There is a clear upward trend in the urban ratios, and a clear upward trend following the recent classification in the rural case since the late 1980s. In other words, if the household survey data were correct, NIA urban cultural and personal service consumption reflects a large amount of overestimation, and increasingly so, while in the rural case underestimation may have turned into overestimation in the early 1990s.

The changes in the ratios between individual years are often far from credible. For example, the urban ratio following the recent classification fell from 2.37 in 1992 to 2.09 in 1993. If the 1993 ratio had been the same as that in 1992, 1993 GDP would have been 0.52% larger.<sup>25</sup> Or, following the early classification, had the 1990 urban ratio of 2.66 been the same as that in 1989 (2.27), 1990 GDP would have been 0.58% smaller.<sup>26</sup> If the long-term trends in the ratios were correct (i.e., household surveys increasingly underestimate services), time inconsistencies in the urban and rural NIA cultural and personal service consumption time series would add a half-percentage point range of flexibility to annual aggregate GDP growth.

### *3.4. Housing, water, and electricity consumption*

The NIA category “housing, water, and electricity consumption” (in the following abbreviated as “housing”) captures three underlying items, namely (i) rent, (ii) water, electricity, and gas charges, and (iii) the value of the services rendered by owner-occupied housing. The three compendia differ in their instructions on how the underlying data are obtained. Table 6 reviews the explanations offered in the three sources.

*Table 6 about here*

In the rural case, only Xu Xianchun includes fuels which leads to double-counting if, as reported, NIA commodity consumption is derived from retail sales and fuels are included in retail sales.<sup>27</sup> The value of the services rendered by owner-occupied housing is measured in the form of depreciation. According to the NBS and Xu Xianchun, depreciation is based on the value per room of housing (presumably estimated by local statistical bureau staff), while the *GDP*

*Manual* focuses on construction costs per square meter. The *GDP Manual* uses mid-year values for housing area and population, whereas the other two compendia use year-end values, and it allows adjustments to the depreciation rate of 2%, where the NBS does not specify the depreciation rate, and Xu Xianchun simply states 2%.

In the urban case, Xu Xianchun further includes repair services (which he, consequently, does not include in the previous section's calculation of the value of services). The *GDP Manual* lists different types of energy expenses to be included. The value of housing services are uniformly based on construction costs, across all three compendia, with the depreciation rate varying from “2-4%,” to “4%,” and “around 4%.”

The differences in treatment that are likely to matter most in terms of overall impact on the value of housing are the inclusion of fuels by Xu Xianchun, the differing calculation methods of housing services of owner-occupied housing in the rural case, and the absence of corrections to the latter value for housing that has been rented out. Sufficient data to check which of the three definitions is actually applied to derive official NIA household housing consumption are not available. The recent household survey classification offers no useful household survey category, only “residence,” which includes an unknown amount of construction materials and fuels (and no imputations for services of owner-occupied housing). In the early classification, in the urban case, the three categories rent, water and electricity, and gas, are of immediate relevance; no categories in the early rural classification are relevant.

In the following, NIA housing consumption is compared to the imputed value of housing services of *all* housing in China. It is assumed that the value of annual services derived from a particular house equals the value of rent this house will achieve if rented out; rent, thus, need not be added (since it is already captured by the imputations). Data on water, electricity and gas charges are only available for the urban case in the early classification; expressing these charges as a percentage of imputed housing values, an annual average percentage can be derived and applied in those years (and the rural case) where data on these charges are not available.

Table 7 reports (i) the urban and rural NIA housing consumption data, (ii) the corresponding imputed housing values, (iii) the available water, electricity, and gas charges, and (iv) the (urban and rural) ratios of NIA housing consumption to the sum of imputed housing values and water, electricity and gas charges. In the urban case, imputed housing values can be derived only based on construction costs, with a distinction as to whether the area completed in each year is valued at the construction costs of that year, or whether the housing stock in total is valued at current-year construction costs. Lacking construction costs for the years prior to 1982, the total 1982 housing area was valued at 1982 construction costs. (This means the value of housing in 1982 may be overestimated. See notes to the table for details on the derivation.) The separate household survey item rent, in the years in which it would have been available, is ignored since it has been replaced by the aggregate imputations.<sup>28</sup>

Comparing NIA housing consumption to the imputed time series, both imputed urban time series are significantly lower than the NIA data, and increasingly so. While in 1982 imputed values were equivalent to 71.12% of the NIA value (10.16/ 14.27), by 1996 that share had fallen to 14.36% (cumulative method of imputation) or 45.47% (current-year method). Including water, electricity and gas charges with the imputations, the ratio of NIA housing consumption to the sum of imputed housing values and the three charges was relatively constant throughout the 1980s but then began to rise drastically in the early 1990s (with water, electricity, and gas charges since 1992 estimated based on the average rate of the previous years; also see notes to the table). The ratio in 1996 reached 2.81 in the cumulative case, up from approximately unity in the 1980s, and 1.50 in the current-year case, up from approximately 0.7 to 0.8 in the 1980s. The ratios are consistently close to one half smaller if different assumptions—which are less likely to be correct—are made regarding the relevant measure of housing area (for details see notes to Table 7), but the time trend and variation remain.

*Table 7 about here*

In the rural case, the annual service (consumption) value of all rural housing can be calculated via the available data on (i) square meters of living space per capita and the value of

the average square meter of living space (in the household surveys), (ii) the value of the average room, the number of rooms per household, and the number of persons per household (also in the household surveys), (iii) square meters of living space per capita (in the household surveys) and annual construction costs of new housing (part of the construction statistics), using the cumulative method where additional housing each year is valued at that year's current construction costs, or (iv) by applying the current-year construction costs to the living area per capita independent of when the housing was constructed; the last two methods are similar to the urban case (for which no household survey data on housing are available). The second method yields identical results to the first method in the years since 1993, but has fewer data points and more questionable data for the earlier years (for details see notes to Table 7). Table 7 reports the results of the first, third, and fourth methods.

The value of imputed rural housing consumption based on household surveys (first method) tends to be higher than the value obtained using the cumulative construction cost method (except in the early years, perhaps due to the overestimated 1982 value in the construction cost method), but lower than the current-year construction cost method.<sup>29</sup> The two construction cost methods in the 1990s, thus, provide a lower and upper bound. The household survey method and the cumulative construction cost method yield values that are consistently smaller than NIA housing consumption; the current-year construction cost method yields imputed housing consumption values that exceed NIA housing consumption, even before considering the water, electricity and gas charges.

The ratio of NIA rural housing consumption to the sum of imputed rural housing services and the three (estimated) charges is much lower than in the urban case, mostly below unity, which suggests that rural housing consumption could well be underestimated throughout. When rural imputed housing is based on square meter values (survey values), the ratio is slightly above unity in 1982 and 1983 only, then falls as low as 0.74 in 1989, before gradually rising to 0.94 in 1996. When imputed housing is based on cumulative construction costs, the ratio fluctuates around 0.8 for many years before rising to 1.25 in 1996; with current-year construction costs, the ratio stays

below unity throughout, but also shows a rise in recent years. Annual changes are occasionally quite large, such as a rise from 0.76 in 1990 to 0.94 in 1991 in the case of cumulative construction costs, or from 0.84 in 1993 to 1.12 in 1995 (and then 1.25 in 1996); in the case of current-year construction costs, the ratio rises from 0.57 in 1985 to 0.73 in 1986, then falls back to 0.57 in 1987. Such drastic changes of more than 20% per year are not realistic.

The rapid rise in both the urban and rural ratios in the early 1990s could suggest that the NBS is consciously redefining GDP to cover more housing services, with, in the urban case, by 1996, NIA housing consumption well exceeding the underlying values. Both the urban and rural depreciation rates at approximately 4% and 2% appear by far too low and would suggest even further scope for increases in the ratio, especially in the rural case, where NIA housing consumption has by 1996 merely reached the level of the underlying data. Annual revisions to the ratio of NIA housing consumption to the underlying data would imply that GDP data are not comparable over the years.

If Xu Xianchun's definition of NIA housing consumption were correct, fuels would have to be included. Data specifically on fuels are available for the rural case in the early classification with, for the most recent year, 1992, a monetary value of 7.94 RMB (and a value of self-provided fuels of 20.95 RMB, which was presumably included in NIA own account consumption). Adding the monetary expenditures on fuels to the denominator would further lower the lower the ratios of NIA housing consumption to the sum of imputed housing, the three charges, and now fuels. But even lower ratios than the ones already obtained without fuels do not appear credible.

### *3.5. Additional NIA consumption items*

The three compendia list further items to be included in NIA consumption. While NBS (1997) does not distinguish explicitly between rural and urban areas, both Xu Xianchun (2000) and the *GDP Manual* do.

The first item is banking and insurance services consumed by households. NBS (1997, p. 144) and the *GDP Manual* (pp. 80f., 86f.) cover this item, Xu Xianchun (2000) does not. These services are calculated as a share of the gross output value of the banking sector plus a share of the gross output value of the insurance sector. The first share is determined as the sum of rural (or urban) household savings deposits interest receipts plus interest expenses (on loans obtained for “living purposes”), divided by the sum of banking sector interest receipts plus interest expenses; the second share is determined as the insurance premiums paid by rural (or urban) households relative to all insurance sector premium receipts. Ignoring insurance services, a rough estimate of the value of rural banking services in the year 2000 is equivalent to approximately 1% of NIA rural household consumption; the urban value is equivalent to a slightly smaller share of NIA urban household consumption.<sup>30</sup>

Three further items, listed by NBS (1997, pp. 144f.), implicitly limited to *urban* areas, and by Xu Xianchun (2000, p. 64) and in the *GDP Manual* (pp. 85f.) explicitly for urban areas, are in-kind income consumption, public health care consumption, and collective welfare consumption. In-kind income consumption refers to consumer goods received from the work unit, self-produced agricultural products or agricultural products received free of charge from farmers, consumer goods brought back from abroad, “etc.” Public health care consumption refers to health care costs born by state-owned, urban collective-owned, and “other types of” units for their employees, and the health care costs of retired personnel (presumably of these units). Collective welfare consumption refers to the costs of collective welfare equipment and subsidies plus recreation, sports and propaganda costs, also as borne by the same three types of units.<sup>31</sup>

All the additional items should indeed be included in NIA household consumption, but the components own account consumption (in the rural case only), commodity consumption, cultural and personal services, and housing, water and electricity consumption in both the rural and the urban case, as published in *GDP 1952-96*, already add up exactly to the total rural and urban NIA consumption data. It is also not possible that the data in the *Statistical Yearbook*, which does not offer a breakdown into the three (four) components, comprise the additional items, since

the aggregate values in the two sources are identical in all years up through 1993. In the years 1994 through 1996, the data—total, rural, and urban—published in *GDP 1952-96* even exceed the corresponding values in the *Statistical Yearbook*; this means that not only do the *Statistical Yearbook* data not incorporate the additional items, but part of the basic three (four) components reported in *GDP 1952-96* went missing in the *Statistical Yearbook* in these three years (in all issues up through 2002).<sup>32</sup> The earliest that the additional items could have been included is 1997. The fact that Xu Xianchun (2000) does not mention banking and insurance services suggests that at least this item may not have been included in more recent years.

Inclusion of the additional items in the original three (four) NIA household consumption components would contradict the definitions in all three compendia. The data confirm that this is not happening. For example, urban in-kind consumption of consumer goods could theoretically have been included in NIA commodity consumption. But the match between NIA commodity consumption and retail sales is so close that there appears no scope for NIA commodity consumption to include anything other than retail sales of consumer goods to residents. With retail sales of consumer goods to social entities on average accounting for 10.66% of total retail sales, and assuming these were all passed on to households as in-kind income consumption (which they presumably are), the ratios obtained earlier of NIA commodity consumption to *retail sales to residents* only, if the in-kind consumption were included in NIA commodity consumption, should be clearly above unity.<sup>33</sup>

If the additional items were to have been included in NIA household consumption in more recent years (after 1996), the NIA household consumption time series (and thereby the GDP time series) would experience a statistical break in the year in which they were included for the first time, with the growth rates in this first year being artificially inflated. If the additional items were introduced gradually over time, NIA household consumption and GDP data are not comparable over time.

#### **4. Implications of Aggregate Data**

One approximate, aggregate measure of the relationship between NIA household consumption and the underlying household survey data are simply the ratios of rural and urban NIA household consumption to the corresponding total rural and urban living expenditures in the household surveys. Figure 4 shows the time series. The denominator of this measure incorporates the complete category residence in the household surveys (rather than subtracting out construction materials for new construction or renovation) but lacks imputed housing consumption values; since the share of housing in NIA consumption is rather small (Figure 1 and Figure 2), its impact on the ratio should be negligible. The denominator also misses out on that share of commodity consumption which is part of retail sales but not captured in the household surveys. There is no reason why the impact of these simplifications on the ratios should not be constant over time.

*Figure 4 about here*

In the rural case, except in 1983 and 1990, the ratio is always at or above unity. The urban ratio is always far above unity, as well as above the rural ratio. The difference in the two aggregate ratios is driven by the fact that NIA rural own account consumption largely adopts the rural non-monetary household survey data, while NIA urban cultural and personal service consumption as well as NIA urban housing consumption by far exceed the corresponding household survey data (Table 3, Table 5, and Table 6).

Both series exhibit a distinct pattern over time. The rural ratio first declined from 1.19 in 1978 to 0.99 in 1983, then rose until 1986, fell back with some variation until 1990 (to 0.98), and rose monotonously to 1.24 in 2001, before falling back to 1.21 in 2002. This degree of variation over time is considerable. For example, if the year 2000 ratio had been the same as that in 1991, namely 1.00, rather than 1.22, then year 2000 GDP would have been 3.87% lower.<sup>34</sup> The urban ratio started out at 1.23 in 1981 and declined slightly to 1.15 in 1986, then rose steeply to first 1.43 in 1993, and, after two shallow dips, 1.48 in 2000, followed by a sharp drop to 1.35 in 2002. The maximum difference is between 1.15 (in 1986) and 1.48 (in 2000); if the year 2000 ratio had been the same as that in 1986, then year 2000 GDP would have been 5.92% lower.<sup>35</sup>

In the 1990s, the ratios are characterized by a clear upward trend. In the rural case, the ratio has since the early 1990s increased steadily (although only to recover its level of the beginning of the reform period); the urban ratio started a long-term gradual increase in 1987 that was broken only in 2001. The impact of changes in these aggregate rural and urban ratios on GDP are far from trivial. Over the long run, if the initial 1981 ratios had been applied to rural and urban household survey consumption (with 1981 being the first year for which both, rural and urban ratios can be calculated), aggregate consumption in 2000 would have been 15.93% smaller, and GDP 7.65%.<sup>36</sup> Assuming that the approximations inherent in using these ratios are constant in relative terms over time, such systematic increases over several years, especially during the 1990s when China gradually adopted the System of National Accounts and then improved its compilation practices, could indicate the conscious, gradual phasing in of new GDP components or imputation elements. Alternatively, the coverage of the household surveys could have deteriorated drastically throughout the 1990s, or NIA household consumption been inflated artificially.

Figure 5 and Figure 6 go one step further. Figure 5 contrasts the annual changes in the rural ratio (of NIA rural household consumption to rural household survey living expenditures) with the annual real growth of aggregate NIA rural consumption. Since no official real data for expenditure approach GDP and its components are available, the nominal data are deflated using the rural consumer price index (which is available only since 1986). Figure 6 does the same for the urban case, with the urban consumer price index available since 1979 (but household survey data only since 1981).

*Figure 5 and Figure 6 about here*

In the rural case (Figure 5), the two time series move very much in step. In other words, a large percentage *increase* in the ratio of NIA rural consumption to rural household living expenditures goes hand in hand with a high NIA rural consumption real growth rate. The correlation coefficient across the 17 years for which the data are available (1986-2002) of 0.6262 is significant at the 1% level. In the urban case (Figure 6), the match is almost equally close; the

correlation coefficient over the whole period of 22 years (1982-2002) of 0.4311 is significant at the 5% level. The urban correlation coefficient for the years since 1989, avoiding a large counter-movement in the mid-1980s, yields a correlation coefficient of 0.7777, significant at the 0.1% level. Since NIA rural (urban) consumption is derived from the rural (urban) household survey data, the rural (urban) household survey data are prior. In other words, changes in the ratio over time cause corresponding changes in the NIA consumption data. In as far as changes in the ratio reflect the gradual phasing in of new GDP components or imputation elements (or other manipulations), the real growth of NIA consumption and thus of expenditure approach GDP are not due to changes in the value of goods and services consumed but due to changes in accounting practices.

## **5. Problems with Underlying Population Data**

Aggregate NIA household consumption, total, rural, and urban, if divided by the corresponding NIA per capita consumption, yields the implicit population data underlying the NIA consumption statistics. Table 8 shows that the implicit population data from NIA consumption statistics published in the *Statistical Yearbook* (and *Statistical Abstract*) and in *GDP 1952-96* are identical in all years except for small rounding differences.

*Table 8 about here*

But these implicit population data are quite different from the official population data in the population section of the *Statistical Yearbook*. While the implied total population vs. the official total population differs less than 1% in all years (1978-2002), the NIA consumption statistics, compared to the official population statistics, over time increasingly overestimate the rural population and correspondingly underestimate the urban population. The overestimation of the rural population rises from less than 1% in 1978 and 1980 to 20.9% in 2002; the underestimation of the urban population rises from less than 4% in 1978 to 33.4% in 2002.

Aggregate NIA household consumption is obtained as the sum of rural and urban NIA household consumption. Each of these, in turn, is derived as commodity consumption (in the

aggregate based on the retail statistics) plus the other NIA household consumption components; aggregate values for the other NIA household consumption components are explicitly obtained by multiplying per capita household survey living expenditures (and imputed housing costs) by the rural and urban population numbers (*GDP Manual*, pp. 78-85). Except for the case of commodity consumption, the per capita rural and the per capita urban consumption data thus come first; the aggregate rural and aggregate urban NIA household consumption data second, and the nationwide aggregate NIA household consumption data third.<sup>37</sup>

The problem then is that the per capita rural and urban consumption data of the non-commodity component in the derivation of aggregate consumption values are multiplied by population numbers that differ from the official ones. The *GDP Manual* repeatedly specifies the source of the population data (for example, on p. 77) as the public security's "national county and city population statistical material" (*quanguo fen xian shi renkou tongji ziliao*); these population data could well be the traditional household registration data. The official population statistics, on the other hand, are based on the regular population censuses interspersed with sample data (*Statistical Yearbook 2002*, p. 112), where the urban-rural distinction follows complex definitions of relevant areas.<sup>38</sup>

The rural household survey data are explicitly limited to villages. The implicit rural population in the NIA statistics covers, in Chinese, *nongcun*, which would also suggest villages (matching the household survey data). The term "rural" in the *Statistical Yearbook* population section—in Chinese, *xiangcun*—implies "villages and rural townships." But the population section's population in villages *and* rural townships is *smaller* than the NIA section's implicit population in villages. This is logically impossible.<sup>39</sup>

If the official population data (in the population section of the *Statistical Yearbook*) were correct, then even these rural population data, although already smaller than those used in the NIA statistics, are still too large to use in aggregating NIA values since they cover both villages and rural townships, rather than only villages. Furthermore, the official data on the rural population (in villages and rural townships) are more likely to err on the higher than the lower

side as migrant labor to the cities is difficult to capture. Since urban per capita consumption is much larger than rural per capita consumption, the above issues imply that aggregate NIA household consumption is currently underestimated.

The NIA household consumption statistics make sense if the population that is registered in rural areas but living in urban areas has a consumption level equal to that of the population living in rural areas. On the other hand, if the former had a consumption level similar to the population registered at the public security organs as urban population, then the official population statistics would be the more appropriate population data by which to multiply rural and urban per capita consumption to obtain aggregate nationwide consumption.

If the official population data were applied to per capita rural and urban NIA consumption in 2002, aggregate nationwide consumption and thereby GDP would rise significantly. Aggregate rural consumption in 2002 would have been 1738.515b RMB (per capita rural consumption of 2222 RMB times official rural population of 782.41m) rather than 2102.69b RMB, while aggregate urban consumption would have been 4086.253b RMB (8138 RMB times 502.12m) rather than 2720.00b RMB. Rural and urban aggregate consumption together (i.e., nationwide aggregate consumption) would have been 5824.768b RMB rather than 4822.69b RMB, i.e., 20.78% larger than the NIA statistics report. Consequently, 2002 GDP would have been 9.41% larger.<sup>40</sup> The “correct” aggregate consumption data are probably larger than the official NIA consumption data, although not as much larger as 20.78% since the NIA *commodity* consumption data are obtained in the aggregate from the retail sales statistics, and since not all of the missing urban population in the NIA statistics may be consuming as much as the average urban population.

## **6. Conclusions**

Following up on official statements as to how NIA consumption is derived from the underlying data reveals a number of problems. First, official explanations differ in details across three different, independent official sources on how China compiles GDP data, all of which can

make a strong claim for credibility. Theoretically, all three sources could be correct, each at the time of its publication (1997, 2000, 2001); this would mean that the official annual GDP data are derived in ways that differ almost year-by-year, without this being made apparent by the NBS when publishing the official data. While none of the definitional differences are likely to have a major impact on GDP values or GDP growth rates, the fact that official sources contain a significant degree of variation in details suggests that either the NBS is unwilling to reveal its actual practices or incapable of accurately specifying the practices it follows. Reading one of these publications provides a false sense of reliability of the data compilation process in use.

Comparing NIA consumption data to the underlying household survey data reveals numerous problems, from the two slightly differing datasets for NIA consumption in the years 1994 through 1996 to statistical breaks in household survey data when calculating NIA consumption (such as the 1990 statistical break in rural household survey in-kind consumption), and changes in the practice of calculating particular components of NIA consumption (such as in the case of rural own account consumption in 1995 and 1996). In the year in which banking and insurance services, urban in-kind income consumption, urban public health care consumption, and urban collective welfare consumption were included in NIA household consumption (in any year after 1996, if at all), real growth in expenditure approach GDP was exaggerated.

Questions also arise from the built-in large scope for manipulating the derivation of NIA consumption data, such as when in the derivation of NIA commodity consumption the underlying retail sales data are reduced by 9-11%, or by 15% (depending on the source of the explanations), or when the depreciation rate applied to the value of urban housing is “4%, adjustable according to the actual situation.” Furthermore, the de facto rules for deriving NIA consumption from the underlying data may be changing frequently, as suggested by the fact that the ratio of the various NIA consumption components to the corresponding underlying household survey values fluctuates significantly between many years. In the aggregate, the year-on-year fluctuations in the 1990s occur around a long-term upward trend, which raises questions of its own, such as whether the relative growth of NIA consumption compared to the underlying data

is due to an increasingly better grasp of what should be included in NIA consumption, to decreasing coverage of the household survey data, or unwarranted manipulations. No matter what the causes of the adjustments are, *relative to* year 2000 NIA consumption (but less so relative to year 2001 NIA consumption), the early NIA consumption values are underestimated (or vice-versa).

The population data underlying the NIA nationwide consumption calculations indicate that even *current* GDP data could be underestimated; in these population data the share of the low-consumption rural population may well be overestimated, and the share of the high-consumption urban population correspondingly underestimated. If the separately published, official population data were correct, GDP in 2002 could have been several percentage points larger than officially reported. Furthermore, imputations such as for the services rendered by owner-occupied housing, especially in the countryside, at a 2% depreciation rate, appear consistently too low.<sup>41</sup>

The fact that changes in the ratio of rural (or urban) NIA components to the underlying data are highly correlated with the official real growth rate has significant implications. The year-on-year changes in the ratios of NIA rural or urban consumption (and its components) to the corresponding underlying data often appear arbitrary, if not due to outright mistakes (such as the possible reliance on old imputation prices in the derivation of rural own account consumption for some years after 1989). In as far as the changes in ratios translate into corresponding annual changes in the aggregate NIA rural (urban) household consumption, the real growth rates of rural (urban) NIA household consumption and GDP reflect unexplained annual decisions on how to calculate GDP rather than changes in the value of consumption in one year. Real growth in part then is an accounting artifact.

China's official GDP real growth rate is derived based on the production/ income approach. Documenting the highly unreliable nature of the household consumption component of expenditure approach GDP thus need not immediately question the quality of the official real growth rate figures. It could, for example, be the case that the NBS obtains reliable GDP data from the production/ income approach and then manipulates expenditure approach GDP until it

approximately matches production/ income approach GDP. However, if the data in the expenditure approach are of the poor quality that the investigations here suggest, how is the NBS able to obtain higher-quality data in the production/ income approach where it faces such difficult tasks as measuring value-added in the individual-owned catering industry? Production/ income approach data are unlikely to be of any better quality than expenditure approach data.<sup>42</sup>

The conclusion on the quality of NIA household consumption data as part of expenditure approach GDP is not all bleak. First, if retail sales data were accurate, the largest component of NIA household consumption is accurate to within lower single-digit percentage points. Second, with the ratio of NIA rural own account consumption to rural self-produced-self-consumed goods equal to unity in 1995 and 1996, at least this component of NIA household consumption corresponds fully to the underlying data (which, in turn, could be of rather poor quality since these are imputed data). Third, rough estimations of the NIA household consumption component ratios to the underlying data for the years after 1996, based on very limited provincial-level data, suggest that the annual variations have abated in the second half of the 1990s.<sup>43</sup> The signs, thus, are of stabilization in ratios and thereby NIA household consumption compilation methods, rather than of increasing divergence and instability.

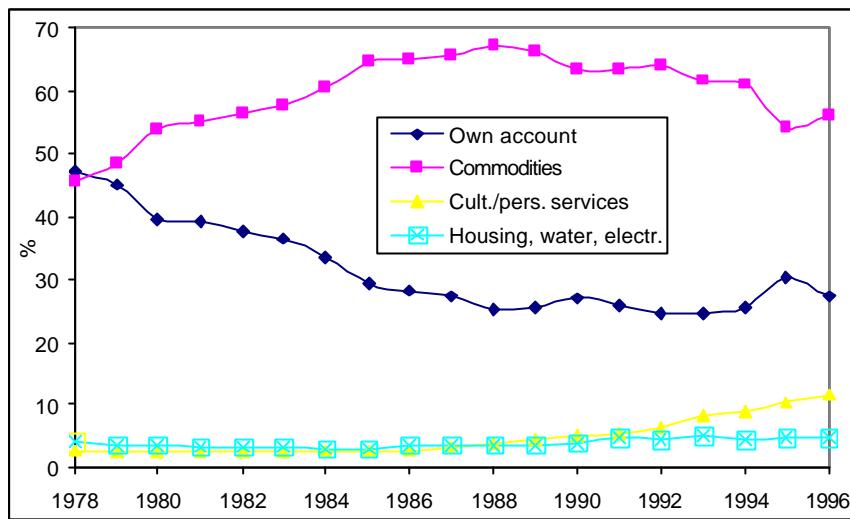
This paper could be extended to cover government consumption, investment, and net exports in an attempt to completely reconstruct expenditure approach GDP. An attempt could also be made to reconstruct production/ income approach GDP, although the data to do so for many sectors or sub-sectors are not available. A final issue is the transition to real values (real growth rates); China's GDP deflator is of questionable quality.

By focusing on just one item in expenditure approach GDP, accounting for half of GDP, this paper was able to conduct an as exhaustive and conclusive investigation of the quality of these data as possible given the explanations on how GDP data are derived and given the available data. The deficiencies unearthed—from a variety of competing explanations on how GDP data are derived to frequent statistical breaks in the derivation practice and wide margins for discretionary judgment—are likely to apply to all national income accounting in China.

## References

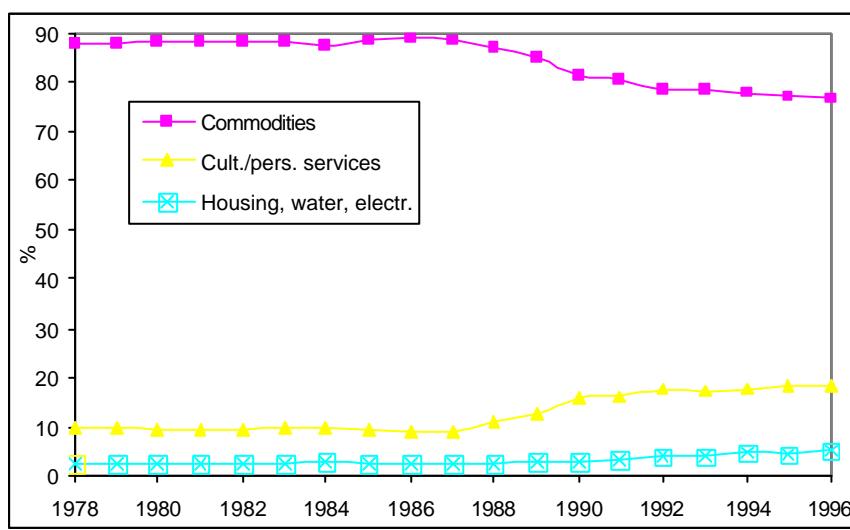
- Adams, F.G. & Chen Y., 1996, Skepticism about Chinese GDP Growth – the Chinese GDP Elasticity of Energy Consumption, *Journal of Economic and Social Measurement* 22, no. 4, 231-40.
- Cai Y., 2000, Between State and Peasant: Local Cadres and Statistical Reporting in Rural China, *The China Quarterly* 163, 783-805.
- Financial Yearbook*, various years, *Zhongguo jinrong nianjian* (Almanac of China's Finance and Banking) (Zhongguo jinrong chubanshe, Beijing).
- GDP 1952-95*, 1997, *Zhongguo guonei shengchan zhongzhi hesuan lishi ziliao 1952-1995* (Historical data on China's gross domestic product 1952-1995) (Dongbei Caijing daxue chubanshe, Dalian).
- GDP 1952-96*, 1998, *Zhongguo guonei shengchan zongzhi hesuan lishi ziliao, zhaiyao* (Historical data on China's gross domestic product, abstract) (Zhongguo tongji chubanshe, Beijing).
- GDP Manual*, 2001, *Zhongguo guonei shengchan zhongzhi hesuan shouce* (Manual for GPD calculation in China) (Compiled by the NBS National Income Accounts Division; Zhongguo tongji guomin jingji hesuansi, Beijing).
- Holt, T., 1998, The Fundamental Principles and the Impact of Using Statistics for Administrative Purposes, *Statistical Journal* 15, no. 3/4, 203-12.
- Holz, C.A., 2002, Institutional Constraints on the Quality of Statistics in a Developing and Transitional Economy: the Case of China, *China Information* 16, no 1, 25-67.
- \_\_\_\_\_, 2003, 'Fast, Clear and Accurate:' How Reliable Are Chinese Output and Economic Growth Statistics?, *The China Quarterly* 173, 122-63.
- Hsueh T.-T. & Li Q., 1999, *China's National Income, 1952-1995* (Westview Press, Boulder, CO).
- Keidel, A., 2001, China's GDP Expenditure Accounts, *China Economic Review* 12, no. 4, 355-67.
- Liu C., Liu K., & Jin Z., 2000, *Ruhe shiyong tongji nianjian* (How to Use the Statistical Yearbook) (Zhongguo tongji chubanshe, Beijing).
- Meng, L. & Wang X., 2000, Dui zhongguo jingji zengzhang tongji shuju kexindu de guji (An evaluation of the reliability of China's statistics on economic growth), *Jingji yanjiu* 10, 3-13.
- NBS (National Bureau of Statistics), 1997, *Zhongguo niandu guonei shengchan zongzhi jisuan fangfa* (Method of calculating China's annual GDP) (Issued by the NBS National Income Accounts Division, Zhongguo tongji chubanshe, Beijing).
- Rawski, T.G., 2001(a), China by the Numbers: How Reform Has Affected China's Economic Statistics, *China Perspectives* 33, 25-34.
- \_\_\_\_\_, 2001(b), What Is Happening to China's GDP Statistics?, *China Economic Review* 12, no. 4, 347-54.
- \_\_\_\_\_, 2002, Measuring China's Recent GDP Growth: Where Do We Stand, Mimeo, 29 August 2002, at <http://www.pitt.edu/~tgrawski/papers2002/>.

- Rural Statistical Yearbook*, various years, *Zhongguo nongcun tongji nianjian* (China Rural Statistical Yearbook) (Zhongguo tongji chubanshe, Beijing).
- Rural Survey Yearbook*, various years, *Zhongguo nongcun zhuhu diaocha nianjian* (China Rural Household Survey Yearbook) (Zhongguo tongji chubanshe, Beijing).
- Shaanxi Statistical Yearbook*, various years, *Shaanxi tongji nianjian* (Shaanxi Statistical Yearbook) (Zhongguo tongji chubanshe, Beijing).
- Statistics Manual*, 1990, *Zhongguo jingji shiyong daquan* (Practical Manual on Economic Statistics of China) (Zhongguo renmin daxue chubanshe, Beijing).
- Statistical Abstract 2003*, 2003, *Zhongguo tongji zhaiyao 2003* (China Statistical Abstract) (Zhongguo tongji chubanshe, Beijing).
- Statistical Yearbook*, various years, *Zhongguo tongji nianjian* (China Statistical Yearbook) (Zhongguo tongji chubanshe, Beijing).
- Urban Survey Yearbook*, various years; issues through 1994: *Zhongguo chengzhen jumin jiating shouzhi diaocha ziliao* (China Urban Resident Household Income and Expenditure Survey Material); issues since 1995 and until 1999: *Zhongguo wujia ji chengzhen jumin jiating shouzhi diaocha tongji nianjian* (China Material Price and Urban Resident Household Income and Expenditure Statistical Yearbook); issues since 2000: *Zhongguo jiage ji chengzhen jumin jiating shouzhi diaocha tongji nianjian* (China Price and Urban Resident Household Income and Expenditure Statistical Yearbook) (all: Zhongguo tongji chubanshe, Beijing).
- Xinjiang Statistical Yearbook*, various years, *Xinjiang tongji nianjian* (Xinjiang Statistical Yearbook) (Zhongguo tongji chubanshe, Beijing).
- Xu X., 2000, *Zhongguo guonei shengchan zongzhi hesuan* (Calculation of China's GDP) (Beijing daxue chubanshe, Beijing).



Source: GDP 1952-96, pp. 60, 62.

**Figure 1. Composition of Rural Consumption**



Source: GDP 1952-96, p. 62.

**Figure 2. Composition of Urban Consumption**

**Table 1. Survey Data Classification of Living Expenditures in Official Statistics**

Rural	Urban
<i>Recent classification</i>	
(a) Living expenditures paid for in money ( <i>huobixing xiaofei</i> ) <ul style="list-style-type: none"> <li>(i) foods (<i>shipin</i>)</li> <li>(ii) clothing (<i>yizhuo</i>)</li> <li>(iii) residence (<i>juzhu</i>)</li> <li>(iv) household equipment, i.e., household equipment, articles and services (<i>jiating shebei yongpin ji fuwu</i>)</li> <li>(v) medicine and health care (<i>yiliao baojian</i>),</li> <li>(vi) transport, post and telecommunication (<i>jiaotong tongxun</i>)</li> <li>(vii) culture, education and recreation (<i>wenjiao yule yongpin ji fuwu</i>)</li> <li>(viii) other goods and services (<i>qita shangpin jifuwu</i>)</li> </ul>	(i) foods ( <i>shipin</i> ) (ii) clothing ( <i>yizhuo</i> ) (iii) household equipment, i.e., household equipment, articles and services ( <i>jiating shebei yongpin ji fuwu</i> ) (iv) medicine and health care ( <i>yiliao baojian</i> ) (v) transport, post and telecommunication ( <i>jiaotong tongxun</i> ) (vi) recreation, education and cultural services ( <i>yule jiaoyu wenhua fuwu</i> ) (vii) residence ( <i>juzhu</i> ) (viii) other goods and services ( <i>qita shangpin jifuwu</i> ) On a separate survey of in-kind income (consumption), conducted at least in recent years, but possibly also in earlier years, no data are available; the categories in the in-kind consumption survey are the same as above.
(b) In-kind consumption ( <i>shiwuxing xiaofei</i> ) <ul style="list-style-type: none"> <li>(i) – (viii) identical to above in (a)</li> </ul>	
<i>Early classification</i>	
(a) Living expenditures for goods ( <i>shenghuo xiaofeipin zhichu</i> ) <ul style="list-style-type: none"> <li>(a.1) goods which are commodities, i.e., paid for in money (<i>shangpinxing xiaofei</i>)               <ul style="list-style-type: none"> <li>(i) foods (<i>shipin</i>)</li> <li>(ii) clothing (<i>yizhuo</i>)</li> <li>(iii) housing (<i>zhufang</i>)</li> <li>(iv) fuels (<i>ranliao</i>)</li> <li>(v) articles for use and others, i.e., largely consumer durables (<i>yongpinji qita</i>)</li> </ul> </li> <li>(a.2) own-account consumption, i.e., self-produced-self-consumed goods (<i>zijixing xiaofei</i>).               <ul style="list-style-type: none"> <li>(i) – (v) identical to above in (a.1)</li> </ul> </li> </ul>	(a) Living expenditures for goods ( <i>goumai shangpin zhichu</i> ) <ul style="list-style-type: none"> <li>(a.i) food (<i>shipin</i>)</li> <li>(a.ii) clothing (<i>yizhuo</i>)</li> <li>(a.iii) basic daily-use commodities (<i>riyongpin</i>)</li> <li>(a.iv) cultural and recreational goods (<i>wenhua yule yongpin</i>)</li> <li>(a.v) books, newspapers and magazines (<i>shubao zazhi</i>)</li> <li>(a.vi) medicine and medical goods (<i>yaoji yiliao yongpin</i>)</li> <li>(a.vii) housing and construction materials (<i>fangwu ji jianzhu cailiao</i>) [in <i>Statistics Manual</i> only identified as construction materials]</li> <li>(a.viii) fuels (<i>ranliao</i>)</li> <li>(a.ix) other goods (<i>qita shangpin</i>).</li> </ul>
(b) Living expenditures for non-goods, i.e., services ( <i>fei shangpin zhichu</i> ) <ul style="list-style-type: none"> <li>(b.1) cultural services (<i>wenhua fuwu</i>)</li> <li>(b.2) living services (<i>shenghuo fuwu</i>)</li> </ul>	(b) Living expenditures for non-goods, i.e., services ( <i>fei shangpin zhichu</i> ) <ul style="list-style-type: none"> <li>(b.i) rent (<i>fangzu</i>)</li> <li>(b.ii) water (<i>shuifei</i>)</li> <li>(b.iii) electricity (<i>dianfei</i>)</li> </ul>

- 
- (b.iv) gas (*meiqi*)
  - (b.v) city transportation (*shinei jiaotongfei*)
  - (b.vi) other transportation (*qita jiaotongfei*)
  - (b.vii) post and telecommunications (*youdianfei*)
  - (b.viii) medicine and health care (*yiliao baojianfei*)
  - (b.ix) education (*xuezafei*)
  - (b.x) child care (*baoyufei*)
  - (b.xi) culture and recreation (*wenyufei*)
  - (b.xii) repair services (*xiuli fuwufei*)
  - (b.xiii) other non-goods (*qitafei shangpin*)
- 

The urban household survey statistics as published in the *Statistical Yearbook* cover only monetary living expenditures (*shenghuo xiaofei zhichu*); Xu Xianchun (2000), p. 60, provides a separate urban model survey form labeled “in-kind income survey” (*shiwu shouru diaocha biao*), organized according to the same classification scheme as the monetary survey. The category of “recreation, education and culture” in the urban case, following the recent classification, appears to cover only services, while in the rural case it covers such *services and goods*; the difference is due to incorrect labeling only, as the urban service category has a subcategory of durable goods (for example, *Statistical Yearbook 2002*, p. 324).

In the early urban survey, for most years the categories (b.ii) and (b.iii) are combined, as are the categories (b.v) and (b.vi). With the detailed early classification in the urban household survey one would expect to be able to establish a correspondence to the values of the categories in the recent classification for the two years for which data on both classifications are available, 1985 and 1990, but this is not the case. Only the two categories food and clothing are identical in both classifications, in both years. Otherwise, only medicine and health care (in the recent classification) can be reconstituted as the sum of a.vi and b.viii, and recreation, education, and cultural services as the sum of a.iv, a.v, b.ix, and b.xi., and for both categories only in 1990.

The *Urban Survey Yearbook* series offers a more detailed breakdown than the *Statistical Yearbook* (used here due to its longer time series) with further sub-categories. Xu Xianchun (2000) has a yet more complete list of both rural and urban survey items (without data).

Sources: Actual data in *Statistical Yearbook*; Xu Xianchun (2000), pp. 58-60, provides three model survey forms (recent classification): rural monetary survey, urban (implicitly monetary) survey, and urban in-kind income survey. Liu Chengxiang, Liu Ke, and Jin Zhaofeng (2000), pp. 120-2, 133-6, provide definitions of the individual items in the recent classification; *Statistics Manual*, pp. 1245f., 1254f., provides a list of categories and sub-categories as well as detailed definitions for the earlier classification.

**Table 2. Sub-Categories of “Residence” and Housing -related Categories in Household Surveys**

Rural	Urban
<i>Recent classification</i>	
Residence ( <i>juzhu</i> )	Residence ( <i>juzhu</i> )
1. Housing ( <i>zhufang</i> ) <ul style="list-style-type: none"> <li>(1) Construction materials (<i>jianzhu cailiao</i>) of which: building repairs (<i>weixiu fangwu feiyong</i>)</li> <li>(2) Rent (<i>fangzu</i>)</li> <li>(3) Others (<i>qita</i>)</li> </ul>	1. Housing ( <i>zhufang</i> ) <ul style="list-style-type: none"> <li>(1) Construction materials (<i>jianzhu cailiao</i>)</li> <li>(2) Rent (<i>fangzu</i>)</li> <li>(3) Repair services (<i>weixiu fuwu</i>)</li> </ul>
2. Electricity ( <i>dianfei</i> )	2. Water, electricity, fuel, and others ( <i>shui, dian, ranliao ji qita</i> )
3. Fuels ( <i>ranliao</i> )	
<i>Early classification</i>	
Within living expenditures for goods (no relevant categories within living expenditures for non-goods):	A. Within living expenditures for goods:
A. Housing ( <i>zhufang</i> ) <ul style="list-style-type: none"> <li>1. Rent (<i>fangzu</i>)</li> <li>2. Electricity (<i>dianfei</i>)</li> <li>3. Building repairs (<i>fangwu weixiu feiyong</i>)</li> <li>4. Expenses on newly built houses (<i>xinjian zhufang kaizhi</i>)</li> <li>5. Materials purchased for newly constructed housing, renovation, or repairs (<i>wei xinjian, weixiu zhufang er goumai de jianzhu cailiao</i>)</li> <li>6. Self-produced bamboo and wood (<i>zhumu</i>), but only if they are also included on the income side</li> </ul>	1. Construction materials ( <i>jianzhu cailiao</i> )
B. Fuels ( <i>ranliao</i> ): coal ( <i>mei</i> ), charcoal ( <i>mutan</i> ), firewood ( <i>chai</i> ), straw ( <i>cao</i> ), all only as far as used for cooking, heating water, and space heating	2. Fuels ( <i>ranliao</i> ): coal ( <i>mei</i> ), liquefied petroleum gas ( <i>yehua shiyouqi</i> ), and other fuels (including firewood [ <i>chai</i> ], straw [ <i>cao</i> ], and kerosene [ <i>meiyou</i> ]))
	B. Within living expenditures for non-goods:
	1. Rent ( <i>fangzu</i> )
	2. Water ( <i>shuifei</i> )
	3. Electricity ( <i>dianfei</i> )
	4. Gas ( <i>meiqfei</i> )

Sources: Xu Xianchun (2000), pp. 58f., for the recent classification; *Statistics Manual*, pp. 1246, 1255, for the early classification.

**Table 3. Ratio of NIA Rural Own Account Consumption to Household Survey Data**

	Household survey, recent classification			Household survey, early classification			
	Foods, clothing, equipment, approximated fuels <sup>a</sup>	Foods, clothing, equipment, residence	All self- provided goods and services (8 categories)	Foods, clothing, equipment	Foods, clothing, articles for use, fuels	Foods, clothing, articles for use, fuels, housing	Foods, clothing, articles for use
1978	0.98	0.97	0.95	1.06	0.96	0.96	1.05
1979			0.98				
1980	0.90	0.90	0.90	1.00	0.92	0.90	1.01
1981	0.94	0.94	0.95	1.07	0.97	0.95	1.07
1982	0.89	0.88	0.89	0.99	0.90	0.89	1.00
1983	0.91	0.91	0.90	1.02	0.91	0.90	1.01
1984	0.86	0.86	0.86	0.96	0.87	0.86	0.97
1985	0.84	0.83	0.83	0.95	0.84	0.83	0.95
1986	0.83	0.83	0.83	0.94	0.83	0.83	0.94
1987	0.85	0.85	0.85	0.96	0.85	0.85	0.96
1988	0.88	0.88	0.88	0.99	0.88	0.88	0.99
1989	0.90	0.90	0.90	1.01	0.90	0.90	1.01
1990 <sup>b</sup>			0.95		0.95	0.95	1.07
1990 <sup>b</sup>	0.74	0.74	0.74	0.82	0.74	0.74	0.82
1991	0.75	0.75	0.75	0.83	0.75	0.75	0.83
1992	0.78	0.78	0.78	0.86	0.78	0.78	0.86
1993	0.76	0.76	0.75	0.84			
1994	0.79	0.79	0.79	0.87			
1995	1.00	1.00	1.00	1.08			
1996	1.00	1.00	1.00	1.07			

a The value of fuels is obtained as follows. In the earlier classification, data on housing and fuel are separately available. Table 2 with the definitions shows that in the early classification housing and fuel together correspond either perfectly, or very closely, to the item residence in the recent classification. In terms of actual (non-monetary) values, dividing the residence value in the recent classification by the sum of housing and fuel in the early classification yields a ratio of unity in all years 1978 through 1992 except 1981 (1.20), 1984 (0.99), and 1991 (0.98); no data are available for 1979. Taking the average ratio of fuels to 'housing plus fuel' in the early classification for 1985 through 1992 (0.9469, standard deviation 0.0065) and applying this ratio to the residence values in the recent classification yields an approximate value of fuel in the recent classification. (In calculating the share in the early classification, 1985 was used as the starting year because data for 1980 and 1981 appear unusually low, at 0.83 and 0.88; 1992 is the last year for which data according to the early classification are available.)

b The NBS obtains the values of self-provided goods and services in its survey of total expenditures by multiplying survey quantities with imputed prices. The choice of imputed prices changed in 1990, with data available both according to the previous practice (in the table in italics) and the new practice. (Further revisions occurred in 1998 and 1999; see note 12 in the text.)

Sources: *Statistical Yearbook* (several issues; national income accounts and livelihood sections); *GDP 1952-96*, pp. 56, 60, 62; *Rural Survey Yearbook 2002*, pp. 37, 37.

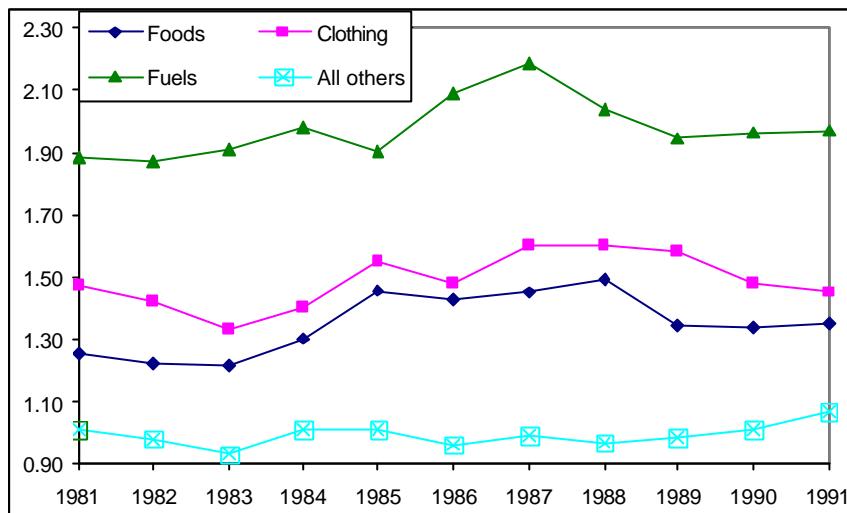
**Table 4. Ratio of NIA Commodity Consumption to Underlying Data**

	Retail sales of consumer goods to residents	Retail sales of consumer goods * 0.85	Household survey: recent classification		Household survey: early classification			
			Urban	Rural	Version I		Version II	
					Urban	Rural	Urban	Rural
1978	0.97	1.01				1.41		1.53
1979	0.97	1.01						
1980	0.97	1.02		1.39		1.21		1.41
1981	0.98	1.03		1.23	1.10	1.06	1.10	1.28
1982	1.00	1.05		1.16	1.11	1.03	1.11	1.26
1983	1.00	1.05		1.14	1.08	1.01	1.08	1.24
1984	1.00	1.04		1.27	1.08	1.11	1.08	1.38
1985	0.99	1.04	1.19	1.40	1.14	1.21	1.16	1.53
1986	0.98	1.03		1.31	1.13	1.13	1.16	1.47
1987	0.97	1.02		1.29	1.19	1.12	1.22	1.46
1988	0.98	1.04		1.28	1.22	1.12	1.25	1.46
1989	0.99	1.05		1.20	1.21	1.07	1.24	1.37
1990	1.00	1.06	1.23	1.19	1.19	1.09	1.22	1.37
1991	1.00	1.05		1.21	1.20	1.11	1.22	1.37
1992	1.03	1.07	1.27	1.33		1.23		1.50
1993	1.01	1.04	1.32	1.35				
1994		1.07	1.28	1.35				
1995		1.05	1.30	1.19				
1996		1.04	1.31	1.21				

Household survey, recent classification: the value of the consumption of goods in the two household surveys (rural monetary, urban (only monetary available to begin with)) are obtained by adjusting the values of each of the eight survey categories in each of the two surveys in each year, subtracting out the service component in each category to obtain the commodity component. The adjustment is based on urban household survey data for the years 1993, 1995, and 1996 (*Urban Survey 1994*, pp. 17, 19; *1996*, pp. 119, 121; *1997*, pp. 123, 125); for these years, some detailed sub-category data are available that allow a sometimes perfectly accurate and sometimes approximate adjustment of urban survey data; the adjustment rates of each of the eight survey categories do not differ much across years, and therefore for each category the average adjustment rate of the three years is calculated and applied to all data (rural and urban, across all years). The average downward adjustment rate of the three years for the category foods is 0.0412% (exact rate, since data on the sub-category food processing costs are available), for clothing 2.7842% (exact rate), for equipment 7.6533% (exact rate), and for medicine and health care 20% (guesstimate); only 20.7868% of the value of transport, post and telecommunications is used (a subjectively chosen 30% of the sub-category transport, and 10% of the sub-category telecommunications), 26.0170% of the category culture, education and recreation (exact rate), 59.8381% of the category residence (to cover the sub-category “water, electricity, fuels, and others,” thus mistakenly including water, electricity and others as commodities, but on the other hand mistakenly lacking construction materials used to repair houses), and 91.8683% of the category “others” (exactly covering “individuals’ consumption” and “other goods”).

Household survey, early classification: in the urban case, the first version covers all goods, namely the categories foods; clothing; articles for daily use; cultural and recreational goods; books, newspapers and magazines; medical and health goods; housing and construction materials (only construction materials used to repair houses should be included); fuels; and others; in the second version, the category housing and construction materials is omitted. In the rural case, the first version covers all goods, namely the categories foods, clothing, housing, fuels, and articles for use; in the second version, the category housing is omitted.

Sources: *Statistical Yearbook* (several issues; national income accounts and livelihood sections); *GDP 1952-96*, pp. 56, 60, 62. Retail sales data are from the *Statistical Yearbook 1992*, p. 605; *1994*, p. 496; *1998*, p. 594; *2002*, p. 578; *Statistical Abstract 2003*, p. 152. (Data on the sale of consumer goods to residents end in 1993.)



An economy-wide breakdown of retail sales of consumer goods into seven categories is available until 1994 in *Statistical Yearbook 1992*, p. 610; 1995, p. 525. Urban and rural household survey data on consumer goods consumption following the early classification are available in nine and in five categories, simultaneously, for the years reported here. (The source is the livelihood section of various years' *Statistical Yearbook*.) Urban, and separately, rural household survey data are multiplied by the population numbers implicit in the aggregate and per capita NIA consumption data to obtain aggregate urban and aggregate rural data for each category. To maintain unambiguous comparability, the different classifications were reduced to four categories. Household survey data following the new classification would have been available for the years 1992 through 1994, but are not immediately suitable for comparison since in the recent classification each category contains both goods and services.

In 1991, foods accounted for 55.03 of total retail sales of consumer goods, clothing for 16.45%, fuels for 3.07%, and all others for 25.45%.

**Figure 3. Ratio of Retail Sales of Consumer Goods to Household Survey Consumer Goods Consumption**

**Table 5. Ratio of NIA Cultural and Personal Service Consumption to Underlying Data**

	Urban		Rural	
	Recent classification	Early classification	Recent classification	Early classification
1978				1.24
1979				1.07
1980			0.53	1.05
1981		1.94	0.47	1.02
1982		1.94	0.62	1.10
1983		1.95	0.61	1.13
1984		2.06	0.63	1.16
1985	1.12	1.95	0.55	1.06
1986		1.71	0.53	0.95
1987		1.76	0.56	0.70
1988		2.25	0.61	0.74
1989		2.27	0.69	0.72
1990	1.89	2.66	0.79	0.74
1991		2.52	0.77	0.70
1992	2.37		0.91	0.82
1993	2.09		1.04	
1994	2.02		1.12	
1995	2.14		1.26	
1996	2.01		1.31	

Recent classification: the value of the consumption of services in the two household surveys are obtained by adjusting the values of each of the eight survey categories in each of the two surveys in each year. The adjustment factors are the converse factors to those listed for the case of goods in the notes to Table 4, except that the category residence is omitted altogether here. Two, in terms of values very likely negligible complications are the following: the subcategory education within the category culture, education and recreation is fully counted as services, even though it contains (an unknown) value of textbooks and reference books; in contrast, the subcategory individual consumption in the category others is fully counted as goods consumption, even though it contains (an unknown) value of travel and services.

Early classification: in the urban case, the official service value is used, less rent, and less water and electricity (these belong into NIA housing consumption); in the rural case, lacking further details, the official service value (cultural and living services) is used.

Sources: *Statistical Yearbook* (several issues; national income accounts and livelihood sections); *GDP 1952-96*, pp. 56, 60, 62; *Urban Survey 1994*, pp. 17, 19; *1996*, pp. 119, 121; *1997*, pp. 123, 125.

**Table 6. Coverage of NIA Housing, Water and Electricity Consumption**

	NBS (1997, pp. 143f.)	Xu Xianchun (2000, pp. 62-64)	GDP Manual (pp. 79f., 84-86)
Rural	<p>1. Rent = item rent in cat. residence</p> <p>2. Water, electricity and gas = electricity costs, plus other expenses (in category residence)</p> <p>3. Owner-occupied housing: implicit depreciation = year-end value per room of housing * year-end average number of rooms per household * number of rural households * depreciation rate (less depreciation on those houses which are rented out)</p>	<p>1. Rent = item rent in category residence</p> <p>2. Water, electricity, and fuels = electricity costs, fuels, and water costs (within "others"), all in catg. residence</p> <p>3. Owner-occupied housing: implicit depreciation = year-end value per room of housing * year-end average number of rooms per household * number of rural households * 2% (depreciation rate)</p>	<p>1. Rent</p> <p>2. Water, electricity and gas = electricity costs, plus water costs (item "other expenses"), both in category residence</p> <p>3. Owner-occupied housing:<sup>a</sup> implicit depreciation = [(year-end average per capita housing area * year-end population + beginning-of-year per capita housing area * beginning-of-year population)/2] * construction costs of rural housing units per area constructed * depreciation rate (experience value, approx. 2%) * share of owner-occupied housing in all housing (experience value according to actual situation, can equal unity if data are difficult to obtain)</p>
Urban	<p>1. as above</p> <p>2. as above</p> <p>3. Based on residential housing construction area, construction costs per square meter of private completed housing, and maximum usage period for housing (with data supplied by Construction Ministry). Depreciation rate is between 2 and 4%. (No further details are provided.)</p>	<p>1. Rent = items rent and repair services in category residence</p> <p>2. Water, electricity, and fuels = cost items "water, electricity, and fuels" and "others," both in category residence</p> <p>3. Owner-occupied housing: implicit depreciation = average square meter construction costs of completed urban private housing buildings (data from Construction Ministry) * area of owner-occupied urban housing buildings (data in household survey) * 4% (depreciation rate)</p>	<p>1. Rent</p> <p>2. Water, electricity and gas = electricity costs, water costs, liquefied petroleum gas, pipeline (coal) gas, water-electricity-fuels, "others" in the category "others"</p> <p>3. Owner-occupied housing:<sup>a</sup> implicit depreciation = [(year-end average per capita housing area * year-end population + beginning-of-year per capita housing area * beginning-of-year population)/2] * construction costs of urban (and industrial/ mining area) housing units per area constructed * depreciation rate (4%, adjustable according to actual situation) (all with data from Construction Ministry)</p>

a The GDP Manual covers only "self-owned" housing (*ziyou zhufang*).

Throughout the table, the original phrasing has been retained as much as possible, at the cost of stylistic perfection. Only the GDP Manual explicitly multiplies the per capita household categories (or sub-categories) in 1. and 2. by the relevant population numbers. The Chinese terms used for per capita housing area are, in the rural case, in the GDP Manual, *juzhu mianji* ((net) living space), and in the urban case, in NBS (1997), *sifang shiyou zhuzhai jianzhu mianji* (private actual residence construction area), in Xu Xianchun (2000), *zhufang jianzhu mianji* (housing area), and in the GDP Manual, *jumin zizhi zhufang mianji* (citizens' own housing area).

**Table 7. Per Capita Housing, Water and Electricity Consumption (RMB)**

	NIA	Urban			NIA housing / w-e-g + housing construction	NIA	Rural			NIA housing / w-e-g + housing construction		
		Housing cons. based on construction costs	Household survey	Rent			househ. survey value	construct. costs	current- year	survey		
Housing, water + electricity	cumulative	current-year			cumulative	current-year	Housing, water + electricity			cumulative	current-year	
1978	10.08						5.76					
1979	10.77						6.02					
1980	12.04						6.61					
1981	14.03		6.36	4.44			6.63	3.42				
1982	14.27	10.16	10.16	7.08	5.04	0.94	0.94	7.38	3.88	6.44	1.07	
1983	14.23	10.65	11.43	7.68	5.76	0.87	0.83	8.07	4.58	7.00	7.21	
1984	17.88	11.46	13.16	7.80	6.24	1.01	0.92	8.98	5.76	8.64	11.18	
1985	19.33	12.40	16.29	6.48	6.84	0.60	0.97	11.14	7.18	9.60	13.23	
1986	22.29	15.29	21.70	7.20	8.88	0.96	0.89	13.36	8.40	10.08	12.54	
1987	26.63	15.72	26.46	7.74	10.32	1.35	0.97	15.34	9.34	10.89	18.24	
1988	31.61	16.91	37.43	7.83	12.32	1.57	1.03	18.45	10.42	11.73	23.88	
1989	41.99	18.85	42.53	8.82	16.20	2.07	1.13	20.55	11.97	12.95	33.39	
1990	48.03	19.53	45.54	9.43	19.81	2.62	1.14	23.04	14.37	14.14	34.23	
1991	61.54	21.00	50.71	10.66	24.16	4.27	1.24	31.07	18.41	15.49	37.72	
1992	88.83	22.76	62.77				1.70	0.96	32.55	21.80	16.36	42.29
1993	123.87	27.34	85.73				1.84	0.98	45.76	26.54	22.40	68.34
1994	184.52	31.12	98.42				2.39	1.28	51.32	32.01	20.69	70.37
1995	222.18	35.41	115.91				2.48	1.31	71.97	38.17	23.92	85.72
1996	289.88	41.62	131.82				2.81	1.50	89.68	50.39	26.94	96.30
1997		46.32	137.84						62.64	30.66	110.01	
1998		54.30	148.42						69.26	34.93	115.62	
1999		62.42	159.24						73.83	39.23	113.40	
2000		71.00	176.74						84.69	42.12	121.62	
2001		78.79	184.60						96.98	46.82	132.77	

W-e-g: water, electricity and gas. In all ratios of NIA housing consumption to w-e-g and imputed housing consumption, except in the years 1982 through 1991 in the urban case, w-e-g is taken to amount to 46.79% of imputed housing consumption based on current-year construction costs. The percentage was obtained as the mean value of the years 1982 through 1991 in the urban case, with a standard deviation of 5.08 percentage points. W-e-g was also related to other measures of housing consumption; the specific method chosen yielded an average value with the lowest coefficient of variation (and standard deviation.)

In the urban case, imputations of housing consumption use the official construction costs per square meter (sqm) of residential buildings from the construction section in the *Statistical Yearbook* and the year-end per capita living space (*juzhu mianji*) from the household surveys in the

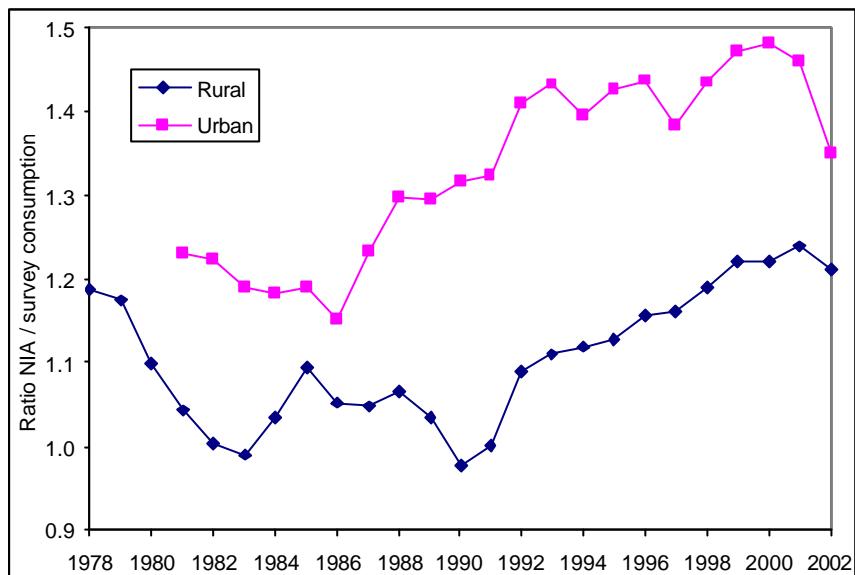
*Statistical Yearbook*. Year 1982 construction costs are set equal to year-end 1982 area per capita times 1982 construction costs per sqm. To obtain cumulative construction costs, in 1983 the additional area per person is priced at current-year construction costs and added to the 1982 value; the same procedure is applied in all later years. To obtain the value of per capita housing at current costs, in every year the total per capita housing area is multiplied by current-year construction costs per sqm. A depreciation rate of 4% is used.

Per capita living space (*juzhu mianji*) data are used in the calculations since both Xu Xianchun (2000) and the *GDP Manual* report the use of *zhufang mianji* (housing area) data, which the rural statistics (see below) reveal as identical to *juzhu mianji*. Only NBS (1997) reports the use of “construction area of private actual residence” (*sifang shiyou zhuzhai jianzhu mianji*; also see notes to previous table). The differences in the data are significant. In 1990 (2000), the official data show a (per capita) constructed (built-up) area (*jianzhu mianji*) of 13.7 sqm (20.3 sqm), used area (*shiyong mianji*) of 9.9 sqm (14.9 sqm), and living space (or “housing area,” *juzhu mianji*) of 6.7 sqm (10.3 sqm). If the data for per capita built-up area were used in calculating the ratio of NIA housing to w-e-g and housing construction, the urban ratios reported in the table would be consistently smaller by close to one half. A further complication is that *Statistical Yearbook* issues of the 1980s report a second urban “living space” time series for the years 1978 through 1988 with values between 17% and 43% higher than those reported in later issues of the *Statistical Yearbook*; the same is true for used space, with a separate time series for the years 1983 through 1988. Here, the more recently published data which cover both the 1980s and the 1990s with what appears a consistent time series are used.

In the rural case, the value of imputed household survey per capita housing consumption is the mid-year value—this is the only imputation using mid-year values. It amounts to 2% of the arithmetic mean of this and the previous year’s “year-end housing area (*zhufang mianji*) per person (as published in the *Rural Survey 2002*, p. 12) times the year-end per sqm value of housing (*zhufang jiazhi*).” The *Statistical Yearbook 2001*, p. 333, contains a time series “net living space” (*juzhu mianji*) which is identical to “housing area” (*zhufang mianji*).

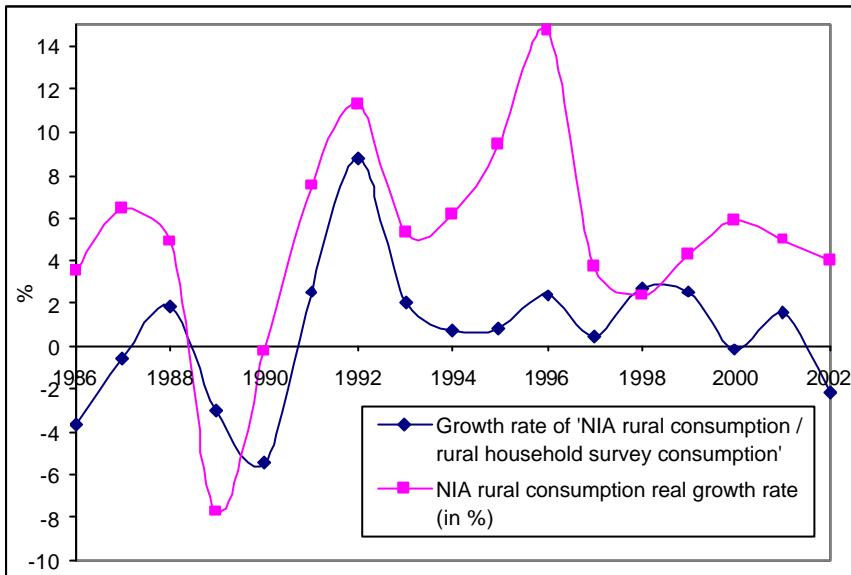
Virtually identical *year-end* housing consumption data for the years 1993 through 1998 can be obtained from a calculation based on the number of rooms per household, the value per room, and the number of persons per household. In the years 1980 and 1985 through 1992 (the other years for which data are available) the year-end area-based values are smaller than the year-end room-based values (the first are equivalent to 0.81 of the latter in 1980, and to 0.85 or 0.84 in 1985 through 1992). However, if the 1980 and 1985-92 values of per sqm housing area are multiplied by per capita “*used*” space (*shiyong mianji*; a gross measure) instead of by housing area, the area-based imputed housing value is identical to the room-based imputed housing value. This has the following implications. (i) Assume the NBS in the countryside collects data only on the average value of a room. Then the area-based imputed housing values used here experience a statistical break in 1992/93; the published value of one sqm of housing in the years prior to 1993 would be for a sqm of *used* space (and should be multiplied by the larger value of used space rather than living space as is done here), while the sqm value after 1992 would be for a sqm of *living* space. But the actual reported value per sqm rises only from 60.11 RMB in 1992 to 73.36 RMB in 1993, an increase of 22.05%, which, given the trend of the time series, appears a normal annual change rather than a statistical break (an approximate 20% rise) plus normal annual change. The assumption, thus, is unlikely to be correct. (ii) If, on the other hand, the NBS in the countryside collects data only on the average value of one sqm of living space, and if this value is consistently based on living space, then the published value of one room is overestimated in the years prior to 1993 (these data are only published for 1980 and 1985 through 1992); the area-based values used here then yield correct values for all years.

Sources: Urban construction costs: *Statistical Yearbook 1990*, p. 220; *1995*, p. 182; *2002*, p. 222; per capita living space (*juzhu mianji*): *Statistical Yearbook 1990*, p. 324; *1995*, p. 290; *1998*, p. 356; *1999*, p. 349; *2000*, p. 324; *2001*, p. 333; year-2001 living space data were obtained by regressing the per capita living space area of the years 1985 through 2000 on a constant and the per capita used space (*shiyong mianji*), a gross measure available for these years in *Statistical Yearbook 2002*, p. 353 (with the two coefficients significant at the 5% and the 0.01% level), and then applying the coefficients to the year 2001 per capita used space. Urban household survey items rent, water, electricity, and gas: *Statistical Yearbook 1988*, p. 807; *1989*, p. 727; *1990*, p. 300; *1992*, p. 284. Rural imputed housing consumption: (i) household surveys: *Rural Survey Yearbook 2002*, p. 12; (ii) construction costs: *Statistical Yearbook 1990*, p. 220; *1995*, p. 183; *2002*, p. 223. NIA housing, water, and electricity consumption: *GDP 1952-96*, pp. 60, 62.



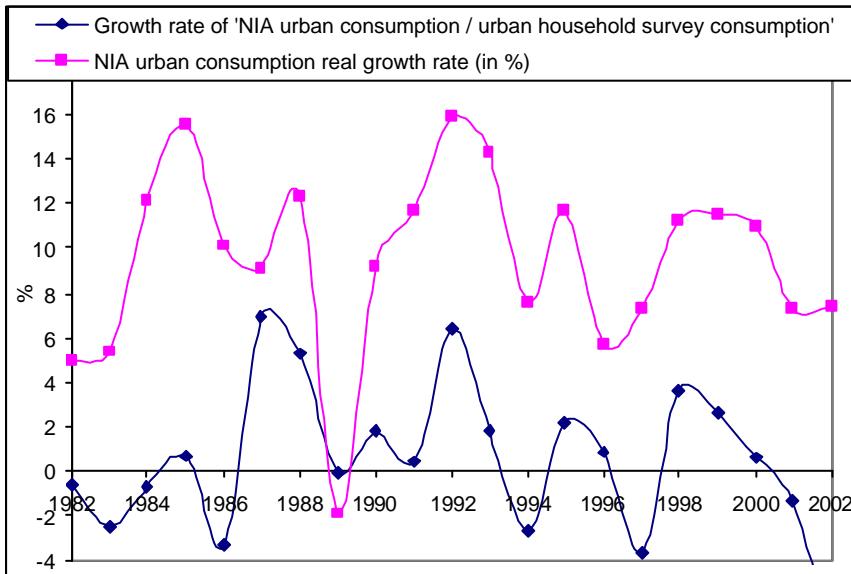
NIA consumption data for all years are taken from the *Statistical Yearbook 2000* and from the *Statistical Abstract 2003*. GDP 1952-96 would have had slightly larger data for 1994 through 1996. Sources: *Statistical Yearbook 2000*, p. 64; *Statistical Abstract 2003*, p.p. 31f.; livelihood section in several issues of the *Statistical Yearbook*.

**Figure 4. Ratio of Consumption in National Income Accounts to Household Surveys**



Sources: See sources to Figure 4; *Urban Survey Yearbook 2002*, p. 16; *Statistical Abstract 2003*, p. 90.

**Figure 5. Relative Change in the Ratio of NIA Rural Consumption to Rural Household Survey Consumption Versus the Real Growth Rate of NIA Rural Consumption**



The 2002 value of the growth rate of 'NIA urban consumption / urban household survey consumption' is negative 7.68%.

Sources: See sources to Figure 4 and Figure 5.

**Figure 6. Relative Change in the Ratio of NIA Urban Consumption to Urban Household Survey Consumption Versus the Real Growth Rate of NIA Urban Consumption**

**Table 8. Population Statistics Implicit in Household Consumption Vs. Official Population Data (absolute data in mio. people)**

	A <i>Statistical Yearbook, NIA section (implicit)</i>			B <i>GDP 1952-96 (implicit)</i>			C <i>Statistical Yearbook, population section</i>			A / B			A / C			Rural share (%)		
	Total	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	A	B	C
1978	956.03	791.59	164.62	956.03	791.59	164.62	962.59	790.14	172.45	1.000	1.000	1.000	0.993	1.002	0.955	82.80	82.80	82.08
1979	969.51			968.79	797.28	171.82	975.42	790.47	184.95	1.001			0.994			82.30	81.04	
1980	981.82	801.85	179.40	981.82	801.85	179.40	987.05	795.65	191.40	1.000	1.000	1.000	0.995	1.008	0.937	81.67	81.67	80.61
1981	994.89			993.93	819.50	173.19	1000.72	799.01	201.71	1.001			0.994			82.45	79.84	
1982	1010.08			1009.82	826.47	180.80	1016.54	801.74	214.80	1.000			0.994			81.84	78.87	
1983	1023.00			1023.31	838.78	185.59	1030.08	807.34	222.74	1.000			0.993			81.97	78.38	
1984	1035.84			1037.99	843.00	194.68	1043.57	803.40	240.17	0.998			0.993			81.21	76.99	
1985	1050.11	841.93	207.92	1050.11	841.93	207.92	1058.51	807.57	250.94	1.000	1.000	1.000	0.992	1.043	0.829	80.18	80.18	76.29
1986	1066.31			1067.01	853.72	213.59	1075.07	811.41	263.66	0.999			0.992			80.01	75.48	
1987	1083.85	870.53	214.06	1083.85	870.53	214.06	1093.00	816.26	276.74	1.000	1.000	1.000	0.992	1.066	0.774	80.32	80.32	74.68
1988	1101.46	880.55	220.82	1101.46	880.55	220.82	1110.26	823.65	286.61	1.000	1.000	1.000	0.992	1.069	0.770	79.94	79.94	74.19
1989	1118.57	889.66	229.83	1118.57	889.66	229.83	1127.04	831.64	295.40	1.000	1.000	1.000	0.992	1.070	0.778	79.54	79.54	73.79
1990	1134.89	898.27	236.30	1134.89	898.27	236.30	1143.33	841.38	301.95	1.000	1.000	1.000	0.993	1.068	0.783	79.15	79.15	73.59
1991	1151.33	908.18	242.91	1151.33	908.18	242.91	1158.23	846.20	312.03	1.000	1.000	1.000	0.994	1.073	0.778	78.88	78.88	73.06
1992	1164.47	915.26	249.92	1164.47	915.26	249.92	1171.71	849.96	321.75	1.000	1.000	1.000	0.994	1.077	0.777	78.60	78.60	72.54
1993	1178.24	920.14	258.18	1178.24	920.14	258.18	1185.17	853.44	331.73	1.000	1.000	1.000	0.994	1.078	0.778	78.09	78.09	72.01
1994	1191.86	922.03	269.89	1192.03	921.91	269.88	1198.50	856.81	341.69	1.000	1.000	1.000	0.994	1.076	0.790	77.36	77.34	71.49
1995	1205.03	923.79	281.03	1204.63	923.84	281.03	1211.21	859.47	351.74	1.000	1.000	1.000	0.995	1.075	0.799	76.66	76.69	70.96
1996	1217.43	927.49	290.13	1217.46	927.55	290.09	1223.89	850.85	373.04	1.000	1.000	1.000	0.995	1.090	0.778	76.18	76.19	69.52
1997	1229.87	929.47	300.51				1236.26	841.77	394.49				0.995	1.104	0.762	75.57		68.09
1998	1242.30	932.31	309.70				1247.61	831.53	416.08				0.996	1.121	0.744	75.05		66.65
1999	1253.49	941.75	311.75				1257.86	820.38	437.48				0.997	1.148	0.713	75.13		65.22
2000	1262.75	942.41	320.17				1267.43	808.37	459.06				0.996	1.166	0.697	74.63		63.78
2001	1271.77	941.90	329.73				1276.27	795.63	480.64				0.996	1.184	0.686	74.06		62.34
2002	1280.25	946.31	334.23				1284.53	782.41	502.12				0.997	1.209	0.666	73.92		60.91

The total, rural, and urban implicit population in the *Statistical Yearbook NIA* section, the *Statistical Abstract*, and in *GDP 1952-96* is derived as aggregate consumption in each of the three categories divided by the corresponding per capita consumption. The *Statistical Yearbook NIA* section calls “rural” in Chinese “nongcun,” and urban “chengzhen,” while the population section uses “xiangcun” and “chengzhen;” *GDP 1952-96* talks of *nongye jumin* and *fei nongye jumin*.

Sources: NIA consumption (*Statistical Yearbook*): aggregate: *Statistical Yearbook 2002*, p. 64; per capita: *Statistical Yearbook 2000*, p. 70 (for years 1987-89); 2002, p. 68; aggregate and per capita for years 2000-2002: *Statistical Abstract 2003*, pp. 31f., 37. NIA consumption (*GDP 1952-96*): *GDP 1952-96*, pp. 56, 60, 62. Official population statistics: *Statistical Yearbook 1992*, p. 77 (for years 1979, 1981-84); *2000*, p. 95 (for years 1986-89); *2002*, p. 93; *Statistical Abstract 2003*, p. 39 (for 2002).

## Notes

<sup>1</sup> For example, Cai Yongshun (2000) reports on data falsification in the countryside. Thomas Rawski (2001a, 2001b, 2002), Gerard Adams and Chen Yimin (1996), and Meng Lian and Wang Xiaolu (2000) question the reliability of GDP statistics in the most recent years, with Thomas Rawski focusing especially on the years since 1998. Albert Keidel (2001) documents the non-systematic differences between the official real GDP growth rates based on the production approach and the real GDP growth rates following the expenditure approach. Carsten Holz (2003), on the other hand, questions the evidence that purports to show data falsification in the nationwide aggregate official statistics.

<sup>2</sup> In 1978, household consumption accounted for 48.79% of expenditure approach GDP and in 2002 for 45.26% (*Statistical Yearbook 2002*, pp. 63f., *Statistical Abstract 2003*, pp. 30f.).

<sup>3</sup> For all ten fundamental principles of official statistics see Tim Holt (1998, pp. 211f.).

<sup>4</sup> Hsueh Tien-tung and Li Qiang (1999) calculate detailed GDP data at the national and provincial level for the years 1952-1995; they provide explanations on how they proceed, but these explanations are too brief to be useful in either reconstructing these authors' calculations or in double-checking on the official data.

<sup>5</sup> Per capita rural as well as urban NIA household consumption data were first published in the *Statistical Yearbook 1998*, with retrospective data for some but not all years since 1978. Per capita rural and urban data are no longer published beginning with the *Statistical Yearbook 2001* (data for the year 2000). However, year 2000-2002 data were later published in the *Statistical Abstract 2003*; the *Statistical Abstract 2003* reports data on all years since 1978, thus also filling earlier data gaps. Aggregate rural as well as urban NIA household consumption data are no longer published beginning with the *Statistical Yearbook 2002* (data for the year 2001), but, again, the *Statistical Abstract 2003* provides these data.

<sup>6</sup> *GDP 1952-96* is an abbreviated version of *GDP 1952-95*; the latter comprises the most extensive official retrospective GDP calculations. *GDP 1952-96* has one more year of data, 1996.

<sup>7</sup> The data in the *Rural Survey Yearbook 2002* following the new classification in all other years are also provided in the *Statistical Yearbook*; these data are identical. The 1979 data reported in the *Rural Survey Yearbook 2002* are incomplete. A further potential data source is the *Rural Statistical Yearbook*, but it contains no data beyond those reported in the *Statistical Yearbook*, and therefore served only as a double-check on data.

<sup>8</sup> This definition conflicts with the list of items in the rural household survey provided by Xu Xianchun (2000, p. 58), who lists housing repair services as a sub-sub-category of the sub-category construction materials, not a separate sub-category. In addition, Xu Xianchun lists no sub-category "others" for residence, but only a sub-sub-category "others" within housing, which in turn is a separate sub-category of residence. (Also see Table 2.)

<sup>9</sup> The household survey category "residence" should never be part of own account consumption in full if only because it includes an investment sub-category "construction materials" (Table 2). Nevertheless, Xu Xianchun and the *GDP Manual* both include the item residence from the household surveys in full in the calculation of NIA own account consumption.

<sup>10</sup> The percentage is obtained by multiplying aggregate NIA rural 1995 own account consumption of 414.820b RMB by 0.21 (1 - 0.79) and dividing the thus obtained 87.112b RMB by total 1995 expenditure approach GDP of 5851.05b RMB (*GDP 1952-96*, p. 60; *Statistical Yearbook 2002*, p. 63).

<sup>11</sup> The data on self-provided goods and services in the recent classification, which does not contain a separate service category, are virtually identical to total self-provided goods in the early classification, which has a separate service category (with comparison data available in 1978 and 1980 through 1992).

<sup>12</sup> Since 1998 (*1997 nian hou*), the average annual price at provincial or county level obtained by farmers when selling a particular good on the market is used to impute the value of the self-provided quantity of that good. Since 1999 (*1998 nian hou*) a multiplier of 0.9 is applied to a such obtained market price in the case of grain and meat, and a multiplier of 0.85 in the case of all other agricultural goods. The 1998 and 1999 changes were implemented without providing a second set of data following the old procedure; the effect of the innovations on the data can thus not be isolated. For details on the changing imputation methods see Liu Chengxiang, Liu Ke, and Jin Zhaofeng, 2000, pp. 129f. The impact on rural own account consumption is unclear as these data are no longer published.

<sup>13</sup> Subtraction of the construction materials is correctly justified by the non-consumption nature of these materials.

<sup>14</sup> For each category, the value of consumer goods can be obtained by adding up the relevant sub-categories or by subtracting the irrelevant sub-categories from the total. The text here reports the most practical method. In the case of residence, the subtraction method involves numerous sub-categories, with the classification of these not matching the household survey table provided in Xu Xianchun (2000, p. 58).

<sup>15</sup> Xu Xianchun only provides the subtraction method. He implicitly retains the construction materials used to repair housing; his explanations on how to obtain commodity consumption uses residence sub-categories that do not fully match his survey table, but the exact wording implicitly suggest this one residual.

<sup>16</sup> Comparing the available urban (rural) NIA household commodity consumption data to the retail sales of consumer goods to urban (rural) residents (with the latter obtained by splitting total such retail sales into an urban and a rural component based on household survey commodity consumption) does not reveal any significant bias in favor of either rural or urban areas (not reported in the table), but larger variation than in the nationwide data. Starting with the recent classification, in the urban case, NIA urban commodity consumption was approximately equal to the underlying value of retail sales of consumer goods in all years, with a ratio close to unity, except in the most recent years, 1992 through 1996, when the ratio varied between 1.03 and 1.09. In the rural case the ratios vary slightly more, but are still close to unity. A drastic fall in the ratio in 1995 from previously 1.10 to 0.99 does not appear credible; it goes hand in hand with an increase in the urban ratio. The ratios following the early classification(s) are all close to unity without significant trend.

<sup>17</sup> Urban in-kind income consumption could theoretically account for some of the excess in the urban ratio, but, as reported in section four below, urban in-kind income consumption is part of a separate (not yet discussed) NIA household consumption component.

<sup>18</sup> The percentage is obtained by multiplying aggregate NIA rural 1992 commodity consumption of 422.30b RMB by 0.0992 ( $1.33/1.21 - 1$ ) and dividing the thus obtained 41.88b RMB by total 1992 expenditure approach GDP of 2586.37b RMB (*GDP 1952-96*, p. 60; *Statistical Yearbook 2002*, p. 63).

<sup>19</sup> For the data see *Statistical Yearbook 1993*, pp. 31f.; *1994*, p. 32; *1995*, p. 32; *2002*, p. 51.

<sup>20</sup> For the data see *Statistical Yearbook 1995*, p. 525; *1998*, p. 595; *1999*, p. 547.

<sup>21</sup> In the category foods, this item covers only processing, in the category clothing it covers processing and repairs, and in the equipment, medical, transport, and culture categories only repairs. For the category residence, only the *GDP Manual*, pp. 79 and 85, mentions a housing repair service cost. Xu Xianchun (2000, p. 62), does not mention the repair service item for the category culture, education and recreation in the rural case. The *GDP Manual* in the urban case does not mention the repair services in the category medicine and health care.

<sup>22</sup> The coverage follows Xu Xianchun (2000, p. 63). The *GDP Manual* (pp. 78f.) neglects to list services for the category culture, education and recreation, except repair services. NBS (1997, pp. 142f.) covers only one out of four relevant items in the sub-category education (following the survey classification in Xu Xianchun, 2000, p. 59); it then adds an item technological education costs as part of the category culture, education, and recreation, which is not part of the urban household survey classification presented by Xu Xianchun (2000, p. 59).

<sup>23</sup> NBS (1997, p 143), and the *GDP Manual* (pp. 78f., 85), but not Xu Xianchun (2000) also include an item child care in the equipment (“household equipment, articles, and services”) category; this item is part of the sub-category others in the equipment category.

<sup>24</sup> NBS (1997, pp. 142f.) does not distinguish between the rural and urban classifications, with the items listed for the category culture, education and recreation suggesting that it focuses on the urban classification. Xu Xianchun (2000, pp. 62f.) and the *GDP Manual* (pp. 78f., 85) distinguish between the rural and urban classifications. The various minor differences in the definition of NIA consumption of cultural and personal services are listed exhaustively in the notes above.

One further inconsistency is that Xu Xianchun (2000) in the calculation of NIA own account consumption does not subtract out the processing/repair services from the value of self-provided consumption in the first four household survey categories. Since he uses the total household survey values (monetary and non-monetary) in the calculation of services, the self-provided processing/repair services in the first four categories are double-counted (in the calculation of NIA own account consumption plus in the calculation of NIA cultural and personal consumption). The *GDP Manual* is double-counting all self-provided services. In practice, these effects are likely to be negligible, since the value of services included in self-provided goods and services is likely to be close to zero.

<sup>25</sup> The percentage is obtained by multiplying aggregate NIA urban 1993 cultural and personal services of 134.69b RMB by 0.1340 ( $2.37/2.09 - 1$ ) and dividing the thus obtained 18.05b RMB by total 1993 expenditure approach GDP of 3450.07b RMB (*GDP 1952-96*, p. 62; *Statistical Yearbook 2002*, p. 63).

<sup>26</sup> The percentage is obtained by multiplying aggregate NIA urban 1990 cultural and personal services of 62.24b RMB by negative 0.1718 ( $1 - 2.66/2.27$ ) and dividing the thus obtained 10.69b RMB by total 1990 expenditure approach GDP of 1831.95b RMB (*GDP 1952-96*, p. 62; *Statistical Yearbook 2002*, p. 63).

<sup>27</sup> According to the definition of retail sales in the *Statistical Yearbook 2002*, p. 606, fuels are included.

<sup>28</sup> The available data on rent are reported in the table for the years (and urban case) when they are available; compared to the imputed housing values, rent is small, with a value of perhaps little more than 10% of the former. Its value may be larger than the housing consumption imputed to the rented housing (by imputing the value of *all* housing), but the difference may not be large enough to cause a significant difference in the ratios.

<sup>29</sup> Imputed housing calculated from square meter values (data in the household surveys) was constructed using mid-year values. Construction data use year-end data; they comprise such a large margin of error that using mid-year values seemed a misplaced attempt at accuracy. The use of year-end data also avoids the loss of one data point.

<sup>30</sup> The gross output value of the banking sector is calculated by adding up the relevant items, as far as the data are available, listed by Xu Xianchun (2000, pp. 43f.) for the production approach to the calculation of value-added in the financial sector. The data are taken from the profit and loss account of the Industrial and Commercial Bank of China and then multiplied by three to approximate the whole banking sector; the Industrial and Commercial Bank of China's volume of loans accounts for one third of all lending by China's financial sector (*Financial Yearbook 2001*, pp. 375, 424, 438). The imputed service share of rural households in this gross output value is approximately one-eighth. One-eighth is obtained by realizing that rural households account for approximately half of all household deposits, which in turn account for approximately one-quarter of financial institutions' total deposits and loans. The calculation of rural households' share in all households' deposits (one half) is based on rural vs. urban per capita net financial property income—interest income and interest expense data, which should be added up, are not separately available—times the rural vs. urban population data underlying the NIA. (*Statistical Yearbook 2001*, pp. 305, 322, 637)

<sup>31</sup> Regarding collective welfare consumption, Xu Xianchun does not list the recreation, sports and propaganda costs, while NBS (1997) lists numerous items in this category, such as showers or barbershops provided by the collective. The three compendia show minor differences in phrasing on what is included.

<sup>32</sup> In the three years, the *GDP 1952-96* total NIA household consumption exceeds the corresponding value in the *Statistical Yearbook* by 2.02%, 3.32%, and 3.22%; for the rural data, the percentages are 1.78%, 3.14%, and 2.72%, and for the urban data the percentages are 2.26%, 3.49%, 3.74%. (*Statistical Yearbook 2002*, p. 64; *GDP 1952-96*, pp. 58, 60, 62)

<sup>33</sup> The ratio of NIA urban commodity consumption to retail sales of consumer goods to urban “residents” (with total retail sales split into an urban and rural component based on household survey data, as explained in the three compendia and reviewed in the text above), is 0.91 in 1985, 1.02 in 1990, and 1.00 in both 1992 and 1993, the only years for which data are available (following the recent classification, with ratios based on the early classification more problematic and ending in 1991). A perfect ratio of unity does not allow for the inclusion of any additional items in NIA urban commodity consumption.

<sup>34</sup> The percentage is obtained by multiplying aggregate NIA rural year 2000 household consumption of 1919.69b RMB by (negative) 0.1803 ( $1 - 1/1.22$ ) and dividing the thus obtained (negative) 346.12b RMB by total 2000 expenditure approach GDP of 8934.09b RMB (*Statistical Abstract 2003*, pp. 30, 31).

<sup>35</sup> The percentage is obtained by multiplying aggregate NIA urban year 2000 household consumption of 2369.87b RMB by (negative) 0.2230 ( $1 - 1.15/1.48$ ) and dividing the thus obtained (negative) 528.48b RMB by total 2000 expenditure approach GDP of 8934.09b RMB (*Statistical Abstract 2003*, pp. 30, 32).

<sup>36</sup> The percentage is obtained by multiplying aggregate NIA rural year 2000 household consumption of 1919.69b RMB by (negative) 0.1475 ( $1 - 1.04/1.22$ ) and aggregate urban year 2000 consumption of 2369.87b RMB by (negative) 0.1689 ( $1 - 1.23/1.48$ ) and dividing the thus obtained (negative) 283.15b RMB and 400.32b RMB by total 2000 household consumption of 4289.56b RMB and total 2000 expenditure approach GDP of 8934.09b RMB (*Statistical Abstract 2003*, pp. 30-32).

<sup>37</sup> The theoretical possibility that the aggregate NIA consumption data are the starting point from which per capita data are obtained would render meaningless the definitions provided in the three compendia on how NIA consumption data are obtained (except for the section on NIA commodity consumption).

<sup>38</sup> Three different definitions have been in use over the years (*Statistical Yearbook 2002*, p. 112). Prior to 1982, and between 1982 and 1999, the identification of urban areas was based on (two slightly differing) administrative area criteria. Since 2000 it follows a 1999 NBS regulation on which the *Shaanxi Statistical Yearbook 2002*, p. 79, reveals more details; the underlying criterion is a population density criterion.

<sup>39</sup> The *Shaanxi Statistical Yearbook 2002*, p. 64, in contrast to the *Statistical Yearbook*, carries not only data on rural (*xiangcun*) and urban (*chengzhen*) population, but also on the farming population (*nongye renkou*) and the non-farming population (*fei nongye renkou*). These latter data almost perfectly match the population data implicit in the

Shaanxi Province NIA household consumption data in all four years which were checked (1990, 1995, 2000, and 2001). *GDP 1952-96* also uses the terms farming and non-farming population. In Shaanxi Province, the farming population in 1978 was approximately of the same size as the rural population; in 1990, the *farming* population was 1.49 times the *rural* population, in 2000 2.78 times, and in 2001 3.77 times. By 2001, Shaanxi had a farming population of 28.07m persons and a non-farming population of 8.52m. These data contrast with a primary sector labor force of 9.94m and a non-primary sector labor force of 7.91m, implying a dependency ratio among the farming population of 2.83 persons per laborer (28.07/9.94) and in the non-farming population of 1.08 (8.52/7.91), which is not credible. In other words, the farming population (at least in the provincial NIA statistics, but, as *GDP 1952-96* suggests, also at the nationwide level) contains more people than just farmers and their families. If “farmers” were a synonym for “villagers,” the NIA statistics would indeed make sense (the rural household survey covers only villages), but then almost three-quarters of the villages in Shaanxi Province are as of 2001 regarded as “urban” by the statistical authority, which does not seem feasible. The problem that registered villagers living in non-villages are not counted as urban in the NIA statistics continues to be relevant. (For the data see *Shaanxi Statistical Yearbook 2002*, pp. 64, 84.)

<sup>40</sup> For the data see the *Statistical Abstract 2003*, pp. 30-32, 36.

<sup>41</sup> In the U.S., in 2002, housing consumption (labeled “housing services”) accounted for 14.67% of total personal consumption expenditures. (See <http://www.bea.doc.gov/bea/newsreel/gdpnewsrelease.htm>, accessed on 25 Oct. 2003.) If the housing consumption share in China were the same as in the U.S., NIA household consumption in China would be about ten percent larger. (For the recent shares in China see Figure 1 and Figure 2.)

<sup>42</sup> I have also considered the possibility that aggregate expenditure approach GDP with the NIA household consumption data could simply be the sum of these data across provinces. But even then, if statistics were compiled conscientiously (for example, nationwide average household survey data should be properly weighted by provincial population numbers), the nationwide calculations presented here should hold. To be on the safe side, I have run several double-checks for one province, Shaanxi Province, for the years 1990, 1995, 2000, and 2001 only to find similarly large if not even larger inconsistencies than in the nationwide data. For example, NIA rural own account consumption in 1990 was equivalent to 27% of rural household survey self-provided consumption, but then to 100% in 2000. NIA commodity consumption was equivalent to 91% of total retail sales of consumer goods (to residents and social entities) in 1990, and to 75% in 2000. NIA urban commodity consumption plus services were equivalent to 130% of the urban household survey total in 1990, and to 107% in 2001; the corresponding rural values (based on the monetary survey) are 73% and 49%. Total NIA urban household consumption finally was equivalent to 137% of the urban household survey total in 1990, and to 114% in 2000; the corresponding rural values (based on monetary and non-monetary survey) are 52% and 95%.

<sup>43</sup> For a few provinces, data on NIA household consumption components are available; applying the relative size of the components at provincial level to nationwide NIA household consumption yields absolute values for the nationwide components which can then be compared to the underlying data as in the years up through 1996. Figure 1, Figure 2, Table 3, Table 4, Table 5, and Table 7 can then be updated for all years 1997 through 2001. The estimates are not reported here because they appear rather unreliable. (i) The set of provinces for which NIA household consumption component data is available in the years 1995 (or 1996) through 2001 accounted for only 24% of nationwide NIA household consumption. (Including all available provinces in any one year, independent of whether data on a particular province are available for all years, in some years leads to a one-third share.) (ii) Comparing the 1995 and 1996 component shares based on provincial-level estimates to those based on the actual nationwide data reveals quite some variation in a ratio of estimated to actual NIA household consumption component values between 1995 and 1996, especially in services. Housing values differ significantly between the two data sources. (iii) The ratio of NIA own account consumption (derived by multiplying the provincial own account consumption share by nationwide rural household consumption) to rural household self-produced-self-consumed goods was 1.00 in 1997, but then rose to 1.05 in 1998 and 1999, to 1.06 in 2000, and to 1.07 in 2001. Among all ratios, this is the one which should most reliably be equal to unity (as the practice nationwide, of 1995 and 1996, suggests), but it is not. Overall, comparing the 1996 nationwide ratios of NIA household consumption components to the underlying data to the 1997 estimated ratios (based on provincial-level component share data), reveals statistical breaks in the tens or twenties of percentage points in 1997, but a relatively stable level afterwards.