The Size and Distribution of Hidden Household Income in China[*]

Xiaolu Wang
National Economic Research Institute
China Reform Foundation
C510 Guo Hong Building, No. A-11 Muxidi Beili
Xicheng District, Beijing, 100038, China
wangxiaolu@neri.org.cn

Wing Thye Woo
Economics Department
University of California
Davis, California 95616, USA
wtwoo@ucdavis.edu
and
School of Finance
Central University of Finance and Economics
Beijing, China

December 25, 2010 revision

The Size and Distribution of Hidden Household Income in China

Xiaolu Wang and Wing Thye Woo

Abstract

Official Chinese data on urban household income are seriously flawed because of significant underreporting of income by respondents and non-participation by the high income groups in official household surveys. We collected urban household income and expenditure data in a way that increased their reliability and the coverage of the rich. We utilized the well-known relationship between Engel's coefficient and income level through two different approaches to deduce the true level of household income for each of the seven Chinese income categories (lowest income, low income, lower middle income, middle income, upper middle income, high income, and highest income). We found that the ratio of our estimated income to official income increased from 1.12 for the lowest income group to 3.19 for the highest income group. Total household disposable income in 2008 is RMB 14.0 trillion according to the official data but RMB 23.2 trillion according to our estimate; and 63 percent of the unreported income went to the richest 10 percent of urban households. The income of the richest 10 percent of Chinese households is really 65 times that of the poorest 10 percent instead of the 23 times reported in the official data. The Gini coefficient is clearly much higher than the usually reported figure of 0.47.

In one of the estimations, we had to drop the 76 richest households (1.8 percent of our sample) from the analysis because there were no super-rich in the official data for us to match characteristics with. So, we still understate the income of the highest income households. As the amount of unreported income indicates the degree of corruption, it is troubling that it grew 91 percent in 2005-08 compared to the 71 percent growth in gross national income (GNI). Serious institutional reforms must be enacted if corruption is not to derail economic development and social harmony.

Key words: income distribution, unreported income, hidden income, grey income, simple-Engel, supplemented-Engel

JEL Code: C81, D31, E21, E26, O53, P24

The Size and Distribution of Hidden Household Income in China

Xiaolu Wang and Wing Thye Woo

1. Introduction

It is easy to sing the praises of China's economic performance. An average annual GDP growth rate of 10 percent during the 1978-2006 period has raised GDP per capita by almost nine-fold over the period. The prevailing expectation in 2006 was that China would continue to register impressive growth for some time to come.[1] It was therefore a big surprise at the end of 2006 when the Plenum of the Central Committee of the Communist Party of China (CPC) did not repeat what every Plenum had proclaimed since the famous 1978 Plenum that the chief task of CPC was economic construction. This 2006 Plenum proclaimed instead that the chief task of CPC was the establishment of a harmonious society by 2020.[2] The obvious implication from this new party line is that the present major social, economic and political trends within China might not lead to a harmonious society or, at least, not lead to a harmonious society fast enough.

We believe that this switch in emphasis occurred because the CPC has concluded that social stability requires not just a high economic growth rate to keep unemployment low but also a growth pattern that diffuses the additional income widely; and that there has been too a rapid increase in income inequality. In the 1985-1987 period, China's Gini coefficient was about 30%.[3] However, according to the Asian Development Bank (2007), China's Gini coefficient climbed from 40.74% in 1993 to 47.25% in 2004 and overtook the four Asian countries (Thailand, Philippines, Malaysia and Turkmenistan) that had higher Gini coefficients than China in 1993-1994. While Nepal has the highest Gini coefficient in Asia in 2001-2004, its value of 47.30% is statistically indistinguishable from China's value of 47.25%. If one combines this with the fact that China's income ratio of the richest 20 percent to the poorest 20 percent (11.37)

---

[1] See, for example, the Goldman-Sachs report by Wilson and Stupnytska (2007) which predicted that China's GDP would surpass that of the United States in 2027. For a review of the debate on how to interpret China's high growth in the post-1978 period, Woo (1999) and Woo (2001).

[2] The harmonious socialist society would be (1) a democratic society under the rule of law; (2) a society based on equality and justice; (3) an honest and caring society; (4) a stable, vigorous and orderly society; and (5) a society in which humans live in harmony with nature; see "CPC key plenum elevates social harmony to more prominent position," People's Daily Online, October 12, 2006. What is revealing is that the official descriptions of the harmonious society downplayed the prominence of achieving a prosperous society. Of the nine objectives listed in the Communique of the 2006 Plenum, "the objective of building a moderately prosperous society" was not only listed last, it was also qualified with the condition that the prosperity should be shared "all-around." And this qualifier is actually a repetition because the narrowing of income gaps had already been listed as the second objective.

[3] Wu and Perloff (2005) put the rural and urban Gini coefficients to be 27.2% and 19.1% respectively in 1985; and Benjamin, Brandt, Giles and Wang (2005) estimated them to be 32% and 22% respectively in 1987.

is the highest in Asia and is significantly higher than the next highest income ratio (9.47 for Nepal), China is probably the most unequal country in Asia today.

One severe difficulty with knowing the extent to which inclusive growth (*bao rong xing zeng zhang*) has not been achieved is the widespread phenomenon of hidden income. In 2007, one of us published a research report (Wang 2007) on unreported income in China in 2005 based on an urban income survey he had conducted in 2005-2006. Wang (2007) estimated that unreported Chinese urban household income -- which is the difference between his estimate of household income and the level of household income reported by the National Bureau of Statistics (NBS) -- totalled RMB 4.8 trillion in 2005, and that most of this unreported income belonged to the high income classes. He called this unreported income "hidden income." A correction of the income statistics by including the hidden income showed that the ratio of the income of the richest 10 percent of households to the income of the poorest 10 percent households in urban areas was 31 instead of the reported nine times, and that the same ratio on a nationwide basis, was 55 instead of the reported 21. In short, the 2007 report showed that China's income inequality problem is much more severe than what is usually reported.[4]

Wang (2007) also pointed out that the level of household income reported in the NBS Household Survey was also lower than the level reported in the NBS Economic Census. The latter was still lower than his estimate of household income, and he had called the difference between the estimates of the Economic Census and our survey "grey income".

What has happened to income distribution since 2005? We conducted a second survey on urban household incomes in 2009 to obtain the data for 2008, and this paper reports the findings of this second survey. In what that follows, Section 2 describes the survey method; Section 3 explains the analytical techniques and presents the results; and Section 4 estimates the levels of disposable incomes of urban residents. Section 5 discusses the sources of grey income; and Section 6 analyzes its impact on national income distribution. Section 7 makes some remarks on our findings.

2. Survey method and sample distribution

2.1 Gathering reliable data
The NBS survey samples of urban and rural residents are determined by random sampling that follows standard statistical procedures. We see two defects in the NBS approach:
1) NBS random sampling is based on the principle of voluntary participation. A considerable proportion of higher income residents, however, are unwilling to be included in the survey. The samples are therefore, deficient in high income residents.
2) Among higher income residents in the sampling, many are reluctant to provide true information about their income. They tend to report truthfully their regular salaries, but are relatively untruthful about other types of income, especially the "grey income" from unidentified sources.

---

[4] The 2007 report also verified the veracity of our estimated household income by using data like family ownership of cars and housing, number of overseas trips, and amount of private bank deposits.

In contrast, we obtained more reliable data about household income in our 2007 survey because we drew upon the methods of sociology. We asked our professional survey staff in different regions to interview only the people they are familiar with, i.e., their relatives, friends, neighbors and former schoolmates, whose family financial status they generally know. In the 2009 survey, we adopted the same method, but implemented even stricter quality control measures and expanded the sample size.[5] It is important to note that our method is different from that of random sampling and therefore our data cannot be used directly to extrapolate the general distribution of urban household income.

Before the survey, we trained our survey staff at various locations on questionnaire and survey methodology. To eliminate the interviewee's fears, the questionnaire had no information on the identity of the interviewee and the interviewee was assured of the research purpose of the survey as well as the confidentiality of his personal data. We also took measures to lower sensitivity to the survey (e.g. we emphasized that our main purpose was to study consumption structure rather than income levels). The questionnaire is designed to inquire about consumption details before the income details, and to inquire about different components of consumption and income before the total amount. As for the sources of incomes, the questionnaire only asked the interviewees to choose among a few simple categories, including wages and salary, part-time job and service payments, entity-business return, gains from capital and financial markets, property rents, intellectual property royalties, transfer income, and (unclassified) other incomes.

After the survey, the surveyors were required to report their relationship with each interviewee and their personal judgement about the creditability of the survey result (including their judgement on the direction and extent of possible deviations). Then, in addition to making a thorough check for completeness of the information and correctness of the survey locations, we also designed a set of screening procedures to examine the rationality of the logic between answers to different questions and the consistency between income, saving and expenditure data in each questionnaire. We then omitted the "suspicious" questionnaires.

2.2 Distribution of survey samples
This survey was conducted in 64 cities of different scale in 19 provinces (including cities under direct administration of the central government), as well as 14 county towns and administrative towns. The provinces are Beijing, Shanghai, Shandong, Jiangsu, Zhejiang, Guangdong, Shanxi, Henan, Hubei, Anhui, Jiangxi, Liaoning, Heilongjiang, Sichuan, Chongqing, Yunnan, Shaanxi, Gansu and Qinghai. This selection of the provinces achieves balance among the East, Northeast, Central and West regions, and between north and south China. The cities are Beijing, Shanghai, Jinan, Nanjing, Hangzhou, Guangzhou, Tai Yuan, Zhengzhou, Wuhan, Hefei, Nanchang, Shenyang, Harbin, Chengdu, Chongqing, Kunming, Xi'an, Lanzhou, Xining, Shenzhen, Qingdao, Suzhou, Datong, Anshan, Fushun, Qiqihar, Daqing, Xuzhou, Yangzhou, Fuyang (in Anhui province), Wuhu, Lu'an, Rizhao, Xiangfan, Yichang, Dongguan, Zhongshan, Mianyang, Xinzhou, Kaifeng, Sanmenxia, Zhumadian, Xiaogan, Yidu, Pizhou, Fuyang (in Zhejiang Province), Jinhua, Shaoxing, Shaoguan, Chaohu, Chuzhou, Ganzhou, Ji'an, Jingdezhen, Jiujiang, Dandong, Tieling, Mudanjiang, Xichang, Xianyang, Baiyin, Jiayuguan, Tianshui and Yuxi. Among these, 21 cities are either under direct administration of the central government, or

---

[5] The professional survey companies in the different regions employed a total of about 450 people to conduct the interviews.

provincial capitals, or 'sub-provincial' cities, 43 are smaller cities at the prefecture and county levels. In this way, a generally balanced distribution was kept among cities of different scales.

County towns and administrative towns include Fanzhi County in Shanxi Province, Pei County in Jiangsu Province, Xiangshan County in Zhejiang Province, Ping Yuan County and Qihe County in Shandong Province, Hua County in Henan Province, Dawu County in Hubei Province, Zhijiang County, Kai County and Zhong County in Chongqing City, Liquan County in Shaanxi Province, Gaolan County and Jingchuan County in Gansu Province and Minhe County in Qinghai Province. The geographical distribution of these towns was also generally balanced.

This survey has chosen a large number of cities with a relatively scattered distribution of observations in each city, mainly for two reasons. First, if we choose too many observations in a city, we cannot ensure that our surveyors are familiar with all the respondents, which is a prerequisite of the survey. Second, the wide geographical distribution achieved by the large number of cities ensures accurate representation of urban China.

Our methodology does have shortcomings. A major problem is that the survey is done in one interview, and all the income and consumption data of the interviewed families are provided by the interviewee according to his memory (we have excluded those family members who are unfamiliar with their family income and consumption). Compared with surveys that require a respondent to record his income and expenditure every day, our methodology has greater data error. However, requiring a respondent to record his information over an extended period of time is more prone to systematic distortion because of the respondent's sensitivity to some survey questions. As the data errors in our survey were caused by inaccurate memory, they are mostly random instead of systematic. When the desired value is calculated by taking group-averages, then the random errors should offset each other and cause limited bias, but systematic bias cannot be offset by averaging. Our adoption of this survey method is therefore rational.

Our survey covered 4,909 families. After strict inspection, 689 suspicious questionnaires were dropped and 25 negative income observations[6] were excluded to arrive at the final (effective) sample size of 4,195 observations. Table 1 shows the information about regional distribution of the total (collected) sample and the effective (actually used) sample, the scale of the cities, the age and household registration status of the respondents, and the education level and the profession of the family member with the highest income. The samples are generally evenly distributed across regions and cities of different size, and among various age group and education levels.

<insert Table 1 here>

However, it seems from Table 1 that our survey sample is skewed toward people living in larger cities, with better education, owning their business or working in white collar positions. For example, 27.3 percent of our sample population received tertiary level education while official data suggest that the proportion of urban population with tertiary education is less than 14.7

---

[6] Most of the negative income families are not normally low-income families. Their negative income was commonly due to temporary losses in their family business.

percent.[7] This skewing of our survey sample toward people living in large cities and having more education is actually an outcome that we had deliberately created. According to our 2007 report, the understatement of urban household income in the official data mainly occurs with higher income residents. To ensure a large enough sample of high income and very high income households, we intentionally increased the number of observations for this type of people. As will be explained later, the methods that we employ to analyze the data do not allow the sample distribution to influence our estimation of the income distribution of the total urban population.

3. Estimation methods and results

3.1 Engel's coefficient method

Economists call the proportion of food expenses in the total consumption expenditure of a family "Engel's coefficient", and they have long established that the value of Engel's coefficient declines with the rise in income. This happens because after the basic demand for food by the family has been met, its members start spending increasingly more on transportation and communication, luxury goods, higher level education and cultural entertainment. The growth rate of food expenditure becomes increasingly lower than the growth rate of consumption.

In presenting the household income data, the NBS divides the urban resident families into seven income groups according to their per-capita incomes:
  (1) lowest income,
  (2) low income,
  (3) lower middle income,
  (4) middle income,
  (5) upper middle income,
  (6) high income, and
  (7) highest income.

Groups (1), (2), (6) and (7) account for ten percent each of all urban families. Groups (3), (4) and (5) account for 20 percent each of all urban families. We calculated the average Engel's coefficients of the seven income groups from the published NBS data.

According to the information obtained from our 2005-2006 and 2009 surveys, higher income families in the NBS household survey usually under-report their incomes to a large extent. Some of them also underreport their food expenditure and total consumption expenditure (but to a much smaller extent on average) and the proportions of the underreporting in these two items are roughly the same. We therefore assume that each household group in the NBS survey underreports their food expenditure and total consumption by the same proportion, which implies that the Engel's coefficient for each income group that is calculated from the NBS survey is the true value for that group of households, even though the income level could be seriously underreported.

---

[7] Table 3-12 in the *China Statistical Yearbook 2009* (CSY2009) reports that 6.7 percent of Chinese population has at least a college-level education, and Table 3-4 in CSY2009 shows that 45.7 percent of the Chinese population are urban residents. The 14.7 percent is obtained if all such educated people live in urban areas.

The important implication is that if we can obtain an independent estimate of the true relationship between Engel's coefficient and income level in China, then we can use the Engel's coefficient of each NBS income group to deduce the true level of income in each NBS income group. The difference between the deduced income level and the NBS-reported income level is the "hidden income" of the average family in each income group.

As our samples were collected in a manner that encouraged respondents to report their true income and true expenditure, we can use the 2008 sample to calculate the true relationship between Engel's coefficient and income level. We calculated this relationship in two ways by assuming, in turn, that the size of Engel's coefficient depends:
1. only on the per capita income in the family. The use of this particular Engel's coefficient to estimate actual income is called the "simple-Engel approach".
2. not only on income but also on a number of other variables (that we will identify later.) We call this more general view of Engel's coefficient the "supplemented-Engel approach", and this is our preferred approach.

It is important to understand that
1. the estimation of the true multivariable Engel's coefficient equation can be done without our data sample to be representative of the national population; and
2. (as will be shown) when we use the estimated multivariable Engel coefficient equation in combination with the national average values[8] of the variables for each income group and the NBS value of the Engel's coefficient for each income group, we obtain the true level of the national average income for each income group.

3.2 The simple-Engel approach
First, we calculate the per-capita income and Engel's coefficients of all observations.[9]

Second, we sort all the valid study samples according to their households' per-capita disposable income from the lowest to the highest. To group the samples, we start at the lowest income and keep adding observations until we achieve an average Engel's coefficient that equals that of the "lowest income group" of the NBS survey. This chosen sample group is called "lowest income group". Then, we start with the next observation above the cut-off income of the lowest income group and use the same method as before to arrive at the upper cut-off income of the "low income group", i.e. this group of observations in our sample has the same average Engel's coefficient as the "low income group" of NBS data. This method is repeated for the next higher income group. Our procedure of grouping does not require consideration about the number of observations in each income group.

We had to leave out the 76 richest observations from the "highest income group" because the values of their Engel's coefficients are so low that their inclusion would render the average value of the Engel's coefficient to be far below the NBS value of the Engel's coefficient in the official "highest income group". This suggests that the NBS household sample does not contain the very-rich families in China. The 76 excluded observations accounted for 1.8 percent of our survey

[8] These national average values are not from our survey sample.
[9] The unit of observation is the family.

sample and have (a) a minimum annual per capita disposable income of more than RMB 400,000, (b) a maximum per capita income of RMB 1.76 million, and (c) an average per capita income of RMB 658,811.

Third, we calculate the average per-capita income of each income group of the study samples.

Fourth, we compare the per-capita income of the part of our sample in each income group with that of the corresponding official sample group, and discover the understatement of income in the official samples.

Table 2 shows the distribution of our survey sample and the NBS (official) sample by income groups. The "high income group" and "highest income group" together make up 28.6 percent of our sample, while they make up only 20 percent of the official sample. So we have achieved the desired skewing of our survey sample discussed earlier to encompass rich and super-rich households.

Table 3 shows the per-capita incomes between our samples and the official NBS samples. The per-capita income of each income group of our survey sample is always higher than that of the official samples. The gap expands for the higher income groups. In the highest income group, the NBS survey shows a per capita income of RMB 43,614 but our survey sample shows a per capita income of RMB 164,034, nearly 3.8 times larger. The unreported income gap of this "highest income group" accounted for about two-thirds of the total hidden income. These findings coincide with the findings reported in our 2007 study. The general consistency between the two studies is re-assuring about their credibility.

<insert Table 2 here>

<insert Table 3 here>

3.3 The supplemented-Engel approach
We see five other variables (beside income) to also be determinants of the size of Engel's coefficient. First, prices of consumption goods vary from city to city. For instance, food prices tend to be higher in large cities than in small cities, and so Engel's coefficient in large cities is likely to be higher. We use a city scale variable (hence "*city*") to catch this price effect. Extra large cities (population of more than 2 million), large cities (population between 1 and 2 million), small and medium cities (population of less than 1 million), and county towns are given the values of 1, 2, 3 and 4, respectively.

Second, residents in different places have different dietary habits. Since people in some regions may spend more on food than others, we insert *region-specific dummies* to the regression equations. An analysis of our survey data shows that, under the same circumstances, Engel's coefficients in Shanghai, Jiangxi and Sichuan are noticeably higher than the average level of all provinces. Dummy variable *H1* is used to represent these three provinces. Engel's coefficients in Beijing, Shandong, Hubei, Guangdong, Chongqing and Henan are moderately higher than the average, and they are presented by dummy variable *H2*. Engel's coefficients in Liaoning and Shanxi are lower than the average, and they are presented by dummy variable *L1*. Observations from the other

provinces (including Jiangsu, Zhejiang, Anhui, Heilongjiang, Yunnan, Shanxi, Gansu and Qinghai) constitute the reference sample.

Third, family size may have impact on Engel's coefficient, because bigger families tend to buy food in bulk to save on food expenses. A "*family*" variable is used to represent the number of family members.

Fourth, the education level may affect Engel's coefficient, because residents with a higher educational background may consume more communication, education, and cultural entertainment services, while residents with lower education backgrounds may consume more food, cigarettes, and drinks. A variable "*edu18*" is set to represent the average education level for family members at or above age of 18. The variable *edu18* is valued from 1 to 5 to refer to:
1. elementary school and below,
2. junior middle school,
3. senior middle school and vocational school,
4. college and university, and
5. post-graduate and doctoral studies.

Fifth, Engel's coefficient may be affected by the family's employment ratio (the proportion of employed family members in the whole family, the "*emp*" variable). On one hand, more family members being employed may mean less food expenses because they may eat at their workplaces and enjoy food subsidies there. On the other hand, however, they may prefer eating in regular restaurants and hence incur more food expenses.

To take non-linearity into account, we estimated the following four specifications of Engel's coefficient ("*eng*") equation:

$$eng = C_1 + a_1 lnY + a_2 city + a_3 family + a_4 edu18 + a_5 emp + a_6 H2 + a_7 H1 + a_8 L1 \tag{1}$$

$$eng = C_2 + b_1 lnY + b_2 city + b_3 family + b_4 edu18 + b_5 emp + b_6 H2 + b_7 H1 + b_8 L1 + b_9 (lnY)^2 \tag{2}$$

$$eng = C_3 + c_1 Y + c_2 city + c_3 family + c_4 edu18 + c_5 emp + c_6 H2 + c_7 H1 + c_8 L1 + c_9 Y^2 + c_{10} city^2 \\ + c_{11} family^2 + c_{12} edu18^2 + c_{13} emp^2 \tag{3}$$

$$eng = C_4 + d_1 Y + d_2 city + d_3 family + d_4 edu18 + d_5 emp + d_6 H2 + d_7 H1 + d_8 L1 + d_9 Y^2 + d_{10} city^2 \\ + d_{11} family^2 + d_{12} edu18^2 + d_{13} emp^2 + d_{14} Y^3 + d_{15} city^3 + d_{16} family^3 + d_{17} edu18^3 + d_{18} emp^3 \tag{4}$$

In preliminary regressions not reported here, the squared and cubic terms of some variables in equation (3) and (4) were found to be statistically insignificant at the 10 percent level. These variables were omitted from the final specifications. The regression results of the final specifications are shown in Table 4. We see that although the adjusted R^2 of the four models are not high, most of the variables found strong statistical support. As model (2) has the highest adjusted R^2, it is our preferred model, and we will use it in the subsequent grey income estimations.

<insert Table 4 here>

For each income group, we assign values to each control variable that equal to the national average values of those variables in that income group. Specifically, the national average value for all income groups of the:

- *city* variable is approximately 2.5. Because we know from our 2006 and 2009 surveys and from international experience that the richest households tend to live in the bigger cities, and the poorest household tend to live the smaller towns, we assume that *city* = 1.3 for the highest income group, and *city* = 3.3 for the lowest income group; and that the values for the *city* variable for the other five income groups lie proportionally within this range (e.g. *city* = 2.3 for the middle income group).
- education level of urban residents above 18 years of age is around 3. Again since *edu18* is closely related to income, we assume that *edu18* = 3.8 for the highest income group, and *edu18* = 2.6 for the lowest income group; and that the values for the *edu18* variable for the other five income groups lie proportionally within this range (e.g. *edu18* = 3.2 for the middle income group).
- family employment ratio is around 0.5. From the NBS household survey, we know the values for the *emp* variable in each income group, e.g. *emp* = 0.62 in the highest income group, and *emp* = 0.38 in the lowest income group.
- family size variable is 2.9. From the NBS household survey, we know the values for the *family* variable in each income group, e.g. *family* = 2.6 in the highest income group, and *family* = 3.3 in the lowest income group.
- regional dietary effect is about 0.01, which is obtained these by multiplying the national average value of each regional dummy variable with its estimated coefficient, and then adding them up.

Based on the estimated parameter values of regression equation (2), the above values for the five control variables for each income group, and the NBS value of Engel's coefficient (eng) for each income group, the equation (2) for each income group is reduced to a quadratic equation in lnY:

$$b_9(lnY)^2 + b_1 lnY + [(C_2 + b_2 city + b_3 family + b_4 edu18 + b_5 emp + b_6 H2 + b_7 H1 + b_8 L1) - eng] = 0 \quad (5)$$

From (5), it is straightforward to compute the value of lnY using the quadratic formula, and then of Y for each income group; see Table 5. This supplemented Engel approach allows us to estimate the national average income for each of the seven income groups without requiring that our survey sample be a representative national sample.

<insert Table 5 here>

4. Estimating the true income of urban residents

4.1 The estimated urban household income by group
Table 5 compares the estimated results from supplemented-Engel approach with the results from the official data and our simple-Engel approach. It shows that estimated incomes derived from the supplemented-Engel approach for the two low-income groups are only marginally higher than the official incomes. The gap between estimated income and the official income becomes

significantly greater for the middle-income groups and above. The greatest difference lay within the highest income group, with per-capita income at RMB 164,034 according to the simple-Engel approach and RMB 139,000 according to the supplemented-Engel approach, which are 3.76 and 3.19 times official income respectively. Driven by the high estimated income of the high income and highest income groups, the average per-capita income of all urban residents is nearly double the official income, i.e. RMB 35,462 according to the simple-Engel approach and RMB 32,154 according to the supplemented-Engel approach instead of RMB 16,885 according to the NBS survey.

Our judgement is that the supplemented-Engel approach to estimating the true income level is preferable to the simple-Engel approach because of the strong statistical significance of the other variables reported in Table 4. Hence, from this point onward, we will use the income estimates obtained from specification (2) of the supplemented-Engel approach in all calculations.

Table 6 shows the ratios between the estimated income and official data in 2005 and 2008. We find that in the high income group, the gap between the estimated income and official data has widened most significantly, from 138.7 percent in 2005 to 209.1 percent in 2008. The greatest deviation still occurs at the highest income group, 337.6 percent in 2005 and 318.7 percent in 2008.

<insert Table 6 here>

We want to reiterate that the fact that we had to exclude the 76 richest observations when we employed the simple-Engel approach means that the NBS household survey sample, most probably, has few (or even no) households with per capita disposable income greater than RMB 400,000. Therefore, strictly speaking, the "highest income group" category in Table 5 and 6 does not really capture the truly top-income people. Because we do not know the proportion and income level of the missing super-high income households from the official samples, we cannot correct this distortion in the official data. So, what we have done in this paper is to correct only the understatement of income in six of the seven income categories in the official data. While we still understate the income level of the highest income households, we are confident that our income estimates for this group is far closer to the true level than the official income statistics.

4.2 How large is the hidden income?
As indicated in Table 6, the hidden income of the highest income families accounts for 63 percent of all hidden income, and this makes the income gap between the top and bottom 10 percent of urban families 26 times rather than nine times according to the official data. Together with the hidden income of the high income group, the richest 20 percent of the urban population takes up more than 80 percent of total hidden income. Since hidden income occurs mostly in urban areas, if we use the richest 20 percent urban families and poorest 20 percent rural families to represent the nationwide top and bottom 10 percent families respectively[10], the income gap is 65 times instead of the 23 times in the official data.

---

[10] This is because half of the Chinese population are urban residents and that average rural per capita income on average is only one third of average urban per capita income.

Using our estimated urban income, we derive an approximate total household disposable income of RMB 23.2 trillion in 2008 compared to less than RMB 14 trillion in the official household statistics. This means that the total hidden income in China in 2008 is RMB 9.26 trillion, almost double the RMB 4.85 trillion in 2005 (up by 91 percent). As nominal GDP had increased by only 71.4 percent in the same period, hidden income had expanded at a much faster pace than GDP.

Table 7 reports the changes in some key indicators between 2005 and 2008 according to the official data and according to our corrected official data. After including the hidden income, total household disposable income increased by 69.3 percent from 2005 to 2008, similar to nominal GDP growth. According to the official statistics (excluding hidden income), total household disposable income had increased by only 57.4 percent over the 2005-08 period, causing it to decline from 48.4 percent of GDP in 2005 to 44.5 percent in 2008.

<insert Table 7 here>

4.3 Cross-checking the size of our estimated hidden income
There have been discrepancies between the official household income statistics and other official data series for a long time. However, as we shall see, some of these contradictions disappear once we include the estimated hidden income into the official household income. In this section, we check our estimate of RMB 9.3 trillion in hidden income in 2008 in three ways, by using (a) consumption and saving data, (b) property prices, and (c) private ownership of automobiles.[11]

Consumption and savings   According to the official data on urban and rural households, total savings (the difference between disposable income and consumption) in the nation should have been RMB 3.55 trillion in 2008. We can check the plausibility of this number by calculating the approximate total household savings from the amount of household savings that was put in each of the following six investment vehicles in 2008.

1. Household savings deposits in the banking system increased by RMB 4.54 trillion in 2008, which is more than the RMB 3.55 trillion in total household savings calculated from the official household statistics.

2. New (i.e. excluding second-hand) residential property sales in 2008 were RMB 2.12 trillion. After deducting the RMB 300 billion increase in mortgage loans, RMB 1.82 trillion of household savings was used in property purchases.

3. In 2008, RMB 371.1 billion was spent on private housing construction in rural areas. It is also common for urban residents to build their own houses, with self-built houses accounting for 15-16 percent of self-owned property. As most residents use personal savings instead of loans from banks, we estimate that RMB 700 billion of household savings was spent on private housing construction.

4. Equity in private industrial enterprises (excluding micro-businesses) increased by RMB 1.09 trillion in 2008, and this increase basically came from the owners' savings. Since equity increases in the services sector are estimated to be no less than that in the industrial

---

[11] All the data to be adjusted are from various issues of the NBS *China Statistical Yearbook*.

sector, private savings provided RMB 2.5 to 3.0 trillion of the overall private investment in industrial and service sectors.

5. The negotiable market value of A-shares shrunk by only 50.9 percent in 2008 when the Shanghai Composite Index and Shenzhen Composite Index dropped by 65.4 percent and 62.4 percent, respectively. This approximately 13 percentage point gap between the fall in market value and the fall in share prices means that there was a net investment of RMB 1.35 trillion in the stock market in that year. There was also a RMB 1.7 trillion net increase in treasury and corporate bonds during that year. If we make the conservative assumption that one-third of the investment in bonds and stocks came from household savings, then the amount was around RMB 1 trillion.

6. It is estimated that net private investment in commodity futures, gold, foreign exchange, financial derivatives, cash and deposits in overseas banks together amounted to RMB 500 billion in 2008.

All the above six estimations were based on official statistics, and together they imply total household savings of at least RMB 11-11.5 trillion in 2008, dwarfing the RMB 3.5 trillion computed from the official household survey data. Now, NBS household data put total household consumption to be RMB 9.46 trillion, which is RMB 1.4 trillion less than the household consumption from the GDP statistics. If we make the conservative assumption consumption in official household survey data was underreported by RMB 2 trillion, then the total hidden income in 2008 should be at least RMB 9.5-10 trillion.[12] This number coincides with the RMB 9.3 trillion in hidden income estimated from the supplemented-Engel approach.

Property price and income   International experience tells us that housing prices are usually three to five times that of annual household income to be affordable. In recent years, China's housing prices have been about 10 times the average urban household income, which is well above the affordability of urban residents. But then, the real estate market has been booming in the past years, total residential property sales reached RMB 3.8 trillion in 2009, sharply up from RMB 2.1 trillion in 2008. These events imply that the true average urban household income should be at least double the income level in the official data. This amount is about what we had found using the supplemental-Engel approach; see Table 6.

According to the official household survey, the richest 20 percent of urban households had an average household income of RMB 89,425 in 2008. Since the average property price in the primary market is about five times that amount, this means that high income families were barely capable of purchasing property. However, this is not consistent with what have we observed. During the 20 years between 1990 and 2009, more than 46 million apartments were sold in the open (commercial) market when the richest 20 percent urban families amounted to only 41 million households. This probably meant that some of these apartments were bought by middle-income families, and that some high-income families bought more than one apartment because our survey data show that more than one-third of high-income families did not buy property from the real estate market. This is because some of them live in properties provided either by their companies or by the government; and the others had bought the property at non-market prices

---

[12] This range is from (11.0 - 3.5 + 2.0) trillion, and from (11.5 - 3.5 + 2.0) trillion.

during the housing reform era in the late 1990s. The high-income families who did purchase property from the market usually paid prices that were much higher than the market average; and that around one-third of high-income families owned at least two residences. These facts show that the income of the high-income group must be much higher than reported in the official data.

Holding of private automobiles   According to the official data on car registration, individuals owned 28.14 million private sedans in 2008. Assuming that 90 percent of these belong to urban residents, then there are 12.1 cars for every 100 urban households, which means that a majority of the 20% richest urban families probably own a car. The catch is that the official household survey data in 2008 show only 8.8 private automobiles for every 100 households. This discrepancy may indicate that nearly one-third of high-income families are missing from the official survey data.

As the price of an average private sedan is about RMB 100,000, with RMB 20,000 related expenses each year (on fuel, maintenance, insurance, annual inspection, parking, and tolls), it is reasonable to expect that families who can afford to own a car would have an annual household income of not less than RMB 200,000. According to the official data, the annual disposable income of the richest 20 percent households was only RMB 89,425, which means that most of them cannot afford to buy a car. In contrast, our analysis indicates that the richest 20 percent of urban households have an actual annual income of RMB 248,192, which means that most of them can afford a car.

5. Grey income and its sources

5.1 What does this huge hidden income tell us?
Grey income (which equals the difference between our estimated income level and the income level from the NBS Economic Census) is income that cannot be clearly defined as legitimate or illegitimate. For instance, presents and gift money received during weddings are permitted by law, and some officials collect huge amounts of money at the weddings of their children and relatives. Some government organizations and state-owned enterprises (SOEs) also provide their staff with big bonuses and welfare benefits, far above normal market practices. Tax evasion is one of the major reasons for the grey income phenomenon.

Under the current circumstances, grey income is usually connected with the following four phenomena.

1. Abuse of power for personal gain: A survey in 2006 covering 4,000 enterprises in China included such a question: "How much did your company informally pay officials of government and regulatory agencies?" Only 19.8 percent of the managers replied "None", while 80.2 percent replied "a little", "quite a lot" and "a lot". Within which, those who answered "quite a lot" and "a lot" accounted for 18.1 percent. The situation is worse in industries related to natural resources and monopolies, and in industries under intensive supervision by state authorities. The proportion of managers saying that informal payment is "quite a lot" or "a lot" was 35.2 percent in the mining industry, 24.3 percent in power and gas supply, 23 percent in the real estate sector and 24.2 percent in the chemicals industry; Wang (2006).

The embezzlement of public resources is also common. According to the National Audit Office (2010) report on the central government budget in 2009, 5,170 fake invoices worth RMB 142 million were found in the 29,363 doubtful invoices already reimbursed by 56 central government departments. This finding is not surprising because it is common to see people selling fake invoices in the streets and to receive such advertisements in short message service (SMS) and emails.

Another "emerging industry" that reflects the fast growth in grey income is the gift purchase trade. In many cities, there are an increasing number of traders in the business of buying expensive cigarettes, wine, medicine, jewellery, and gift coupons from households at discount prices. It is certainly weird for households to be buying luxury products and consumer coupons at high prices from regular shops and then re-selling them to these traders at lower prices. There can only be one explanation for this strange phenomenon, i.e., many high-income households have received such items as gifts and were selling them for cash. A key reason for such a rampant gift-giving culture is that it is a safer form of corruption than receiving cash.

2. Public investment and corruption: Public investment is another source of grey income. Two recent examples are the Beijing-Shanghai express railway project and the western section of the West-to-East Natural Gas Transmission Project. When these projects were audited, overcharging of RMB 815 million in project construction was found, in addition to RMB 1.794 billion on irrelevant fees in construction and reimbursements of fake invoices. Furthermore, 80 percent of the work in the construction contract of RMB 3.6 billion for the West-to-East Natural Gas Transmission Project, was awarded without public bidding procedures (National Audit Office 2010).

3. Leaking of land revenue: Because many local governments do not have adequate budgetary resources under the current fiscal system for infrastructure construction and provision of public services, they rely heavily on the sale of land. In 2009, government revenue from land transfer fees reached RMB 1.5 trillion[13] (which is equal to 20 percent of government budgetary revenue) but this revenue is excluded from the formal government budget. Due to poor management of this revenue, it has become a major source of grey income for local government officials in some areas. In addition, the power of local authority to exempt land transfer fees could be another source of corruption. The National Audit Office revealed that in 2009, RMB 68.4 billion of land transfer fees in 11 provinces was not placed under the budgetary supervision system, and RMB 38.1 billion was not collected. One can only imagine how much of the fee exemption ended up in the pockets of local officials.

The government has the authority for land approval, expropriation and sales, and this conferred monopoly status on the supply of land and on the related real estate sector. It is estimated that in 2009, the profits of the real estate sector was RMB 1.7 trillion, which is more than half of total industrial profits, when the number of employees in the real estate sector was only 1.3 percent of those working in the industrial sector, and the assets of the real estate sector was worth only 6

---

[13] See Xinhuanet, 2010.

percent of industrial assets.[14] Of China's top 30 richest billionaires in *Forbes 2009*[15] list, 11 were in property; and of the 36 persons (some tied) listed as China's top 30 richest billionaires in *Hurun 2010*[16] list, 17 were real estate developers. Real estate is clearly a most profitable industry.

However, the RMB 1.7 trillion profit of the real estate industry does not all go to real estate developers. To acquire good pieces of land from local governments, real estate developers sometimes need to "contribute" significantly to people who have the authority to approve land development. The profit of the real estate sector is actually divided between property developers and those who have permit approval power.

4. Distribution of monopoly profits: The national wage statistics of 2008 and 2009 show that the average wage rate in highly monopolistic industries (such as oil, tobacco, power generation and supply, telecommunication, banking, and insurance) is about twice the national average. These data, however, fail to fully reflect the real gap between different industries. First, the actual per-capita income of workers and staff members in monopolistic industries is far more than reported income in official data. According to Bu Zhengfa, former Vice-Minister of Labor and Social Security, the actual per capita income gap between these industries and other sectors is between five to ten times.[17] Second, within some of the monopolistic industries, there is a much wider income gap between ordinary workers, staff and senior management than in the normal cases.

6. Recalculating the distribution of national income

Based on the NBS household income survey, household disposable income in 2008 is slightly below RMB 14 trillion, which is 44.2 percent of the Gross National Income (GNI). However, based on the NBS Flow of Funds (FOF) accounts, household disposable income was RMB 17.9 trillion[18] (roughly RMB 4 trillion higher). The household income from the FOF data is, however, still RMB 5.4 trillion lower than our estimate of RMB 23.2 trillion (see Table 7). We call this RMB 5.4 trillion gap "grey income".

According to the FOF data, disposable household income in 2008 was 56.5 percent of GNI. Of this, compensation of employees accounted for 46.7 percent of GNI, and non-labor income accounted for 9.9 percent. The disposable income of corporation sector (including both financial and non-financial corporations) and government sector accounted for 17.7 percent and 25.9 percent of total disposable income, respectively.

---

[14] This observation was first made by Chen Wanzhi, a member of National People's Congress; see *East Morning Paper* (4 March 2010). The figures here were calculated by the authors from updated NBS (2010) data.

[15] Available at http://www.forbes.com/lists/2009/74/china-billionaires-09_The-400-Richest-Chinese_Rank.html

[16] See "Hurun Rich List 2010 sponsored by Hainan Clearwater Bay" at http://www.hurun.net/listen186.aspx

[17] *Yangtse Evening Paper,* 15 May 2006.

[18] The FOF data are from the National Economic Census, which surveys the entire enterprise sector and not just a sample of households. The NBS has not published the 2008 FOF data. This number is derived from a linear projection of FOF data from 2005 to 2007. Data are from NBS (various years).

It is reasonable to expect that grey income does not generally come from wages (as the respondents have no reason to hide them). Thus the difference between estimated household income from this study (and Wang's (2007) study on 2005 data) and FOF accounts – RMB 5.4 trillion in 2008 (and RMB 2.7 trillion in 2005) – are treated as non-labor income in household income.

After the adjustment in household disposable income, there should also be some corresponding adjustment to GNI and GDP. For instance, it is common to see some companies report their irregular payment to various parties (e.g. bribes to people outside the company, non-reported payments to managers to evade taxes) as production costs. This practice understates the companies' value-added, and if such understatement is widespread then GDP will be understated substantially.

Another source of grey income is leakage of public funds and public assets, and transfer payments (e.g. bribe) among people. This part of grey income does not result in an understatement of GDP, but only increases household income (though only a few would benefit), at the expense of income distributed to government and state-owned enterprises. This type of grey income reduces the income of some groups but increases those of others.

We recalculated the gross national income in 2005 and 2008 by assuming that 60 percent of the RMB 5.4 trillion of grey income was an understatement of the value-added of enterprises and individual business; and that the remaining 40 percent was transfer from enterprise and government incomes to individuals. Tables 8 and 9 show the distribution of national income to households, enterprises, and government before and after these two adjustments.

\<insert Table 8\>

\<insert Table 9\>

Table 9 shows that before the adjustment, household income accounted for only 56.5 percent of national income in 2008. This share rises to 66.7 percent after the adjustment, an increase of 10 percentage points. Before the adjustment, the GNI share of household income dropped by 3.6 percentage points between 2005 and 2008; but with the adjustment, the decline was only 1.9 percentage points. This means that when grey income is taken into consideration, the share of household income in GDP is not that low, and has not been declining so rapidly.

Since most of the grey income is concentrated in the richest 10-20 percent of households, China's income inequality is much worse than shown in the official data. And as the bulk of grey income is likely to have come from the diversion of enterprise and government income and from the expropriation of the income (and, sometimes, property) of politically weak households, such embezzlements could cause social conflict and instability as well as economic inefficiency.

7. Final remarks

This paper analyzed the NBS statistics on household income alongside the income and expenditure data collected by us, and found that the richest 10 percent of urban households have a per-capita disposable income of RMB 139,000 in 2008 instead of the reported RMB 43,614, and that the second richest 10 percent households have a per-capita disposable income of RMB 54,900 instead of the reported RMB 26,250. The Gini coefficient is probably much higher than the 0.47 to 0.50 calculated by different experts.

This concentration of hidden income in the high-income groups is due to institutional defects. Grey income has its origins in the misuse of power and is closely connected with corruption. The widespread existence of grey income reveals that institutional reforms have lagged far behind economic reforms. Unless the government could stay largely uninfluenced by the rent-seeking lobbying of capital owners and other special interest groups, the free competition of the market economy would inevitably be replaced by the monopolistic practices of crony capitalism. Such a development would accentuate income inequality, economic inefficiency, and social conflict. To avoid these serious threats to economic development and social harmony, institutional reforms are essential, especially in the public finance system and the government administrative system.[19]

It might seem surprising to hear that China needs such fundamental and comprehensive reform when it has experienced high growth for almost 30 years. Why meddle with success? Why fix the economy if it is not broken? The frank answer is that "the economy in 1978 was a broken economy and the story of the last thirty years has been the story of successful repair." Post-1978 growth has stayed high because the government has continually changed policies to keep marketising the economy, deepening its integration into the international economy, and reducing the discrimination against the private sector. In short, policy changes and institutional reforms were the reason for keeping growth high in 1979-2009, and the reform process will have to continue if future growth is to remain high.

---

[19] The threat to economic development and social harmony in China comes from more than just the lag in reforming its administrative system. Increasing tensions in relations with other countries and deterioration of the natural environment are also becoming serious threats; see Woo (2007).

References

Asian Development Bank, 2007, <u>Key Indicators: Inequality in Asia</u>.

Benjamin, Dwayne, Loren Brandt, John Giles and Sangui Wang, 2005, "Income Inequality During China's Economic Transition," manuscript

National Audit Office. 2010. Annual Auditing Report on Implementation of 2009 Central Government Budget and Other Budgetary Revenue and Expenditure. Available at http://www.xinhuanet.com/.

National Bureau of Statistics. Various years. *China Statistical Yearbook.* China: China Statistics Press.

Wang, Xiaolu, 2006, A background paper for the China Entrepreneur Survey System 2006 report, unpublished.

Wang, Xiaolu. 2007. Grey Income and Income Inequality in China (in Chinese). *Comparative Studies* 31. China: CITIC Press.

Wilson, Dominic and Anna Stupnytska, 2007, "The N-11: More Than an Acronym," Global Economics Paper No: 153, Goldman Sachs, March 28, 2007.

Woo, Wing Thye, 1999, "The Real Reasons for China's High Economic Growth," <u>The China Journal</u>, Volume 41, January, pp115-137.

Woo, Wing Thye, 2001, "Recent Claims of China's Economic Exceptionalism: Reflections Inspired by WTO Accession," <u>China Economic Review</u>, Vol. 12, No. 2/3, pp. 107-136 2001.

Woo, Wing Thye. 2007. The Challenges of Governance Structure, Trade Disputes and Natural Environment to China's Growth. *Comparative Economic Studies* 40(4): 572-602.

Wu, Ximing and Jeffrey M. Perloff, 2005, "China's Income Distribution, 1985-2001," manuscript.

Xinhuanet. 2010. China Index Research Institute: countrywide land remise fee reached RMB 1.5 trillion in 2009. Available at http://www.xinhuanet.com/.

Table 1: Sample distribution in various classifications

	Total Sample (total collected observations)	Distribution (%)	Effective sample (observations actually used)	Distribution (%)
1. Geographical location				
Eastern region	1,863	37.95	1563	37.26
Central and Northeast regions	1,848	37.65	1605	38.26
Western region	1,198	24.40	1027	24.48
Total	**4,909**	**100.00**	**4195**	**100.00**
2. Distribution by the scale of the cities				
Cities with more than 2 million in population	2,495	50.83	2,083	49.65
Cities with 1 to 2 million in population	915	18.64	789	18.81
Cities with less than 1 million in population	995	20.27	889	21.19
County towns and administrative towns	504	10.27	434	10.35
Total	**4,909**	**100.00**	**4,195**	**100.00**
3. Age of the interviewees				
20-29	1,647	33.55	1,411	33.64
30-39	1,383	28.17	1,196	28.51
40-49	1,236	25.18	1,062	25.32
50-59	520	10.59	425	10.13
60 and above	123	2.51	101	2.41
Total	**4,909**	**100.00**	**4,195**	**100.00**
4. Household registration of the interviewees				
Local urban resident	4,457	90.79	3,808	90.77
Non-local urban resident	276	5.62	234	5.58
Non-local rural resident	156	3.18	138	3.29
Forget to answer	20	0.41	15	0.36
Total	**4,909**	**100.00**	**4,195**	**100.00**
5. Educational level of the family member with highest income				
Elementary school and below	165	3.36	136	3.24
Junior middle school	970	19.76	832	19.83
Senior middle school (including equivalency)	1,833	37.34	1,565	37.31
University and college	1,822	37.12	1,569	37.40
Post-graduate and PhD	82	1.67	74	1.76
Forget to answer or indefinable	37	0.75	19	0.45
Total	**4,909**	**100.00**	**4,195**	**100.00**
6. Profession of highest income member of the family				
General technical personnel	396	8.07	353	8.41
Middle and senior level technical personnel	262	5.34	227	5.41
Other professional (scientists, teachers, doctors, performers, etc.)	339	6.91	302	7.20
Low level officer of the Party, government, army, etc.	193	3.93	165	3.93
Middle and senior level official of the Party, government, army, etc.	52	1.06	47	1.12
Ordinary staff of enterprises and social organizations	561	11.43	483	11.51
Middle and senior manager of enterprises and social organisations	327	6.66	268	6.39
Service personnel	317	6.46	277	6.60
Worker	659	13.42	562	13.40
Family business or self employed	1,008	20.53	853	20.33
Owner, partner, shareholder of private enterprises	317	6.46	277	6.60
Other occupations	73	1.49	66	1.57
Students, post-graduates	20	0.41	17	0.41
Jobless (including retired)	349	7.11	278	6.63
Forget to answer or indefinable	36	0.73	20	0.48
Total	**4,909**	**100.00**	**4,195**	**100.00**

Note: The scale of the city is measured by its regular urban population. *Source:* Our 2009 Survey Sample Data.

Table 2: Distribution of effective sample and official sample by income group

Group	Study samples range (RMB)	Study samples (households)	Study samples proportion (%)	Official samples proportion (%)
Lowest income	1-7,000	365	8.7	10
Low income	7,001-10,000	622	14.8	10
Lower middle income	10,001-17,000	927	22.1	20
Middle income	17,001-26,500	650	15.5	20
Upper middle income	26,501-34,000	355	8.5	20
High income	34,001-75,000	635	15.1	10
Highest income	75,001-400,000	565	13.5	10
Excluded samples	>400,000	76	1.8	0
Total		**4,195**	**100.0**	**100.0**

Note: Altogether 65,000 urban households are included in the official samples.

Source: Our 2009 survey result and statistics (NBS 2009).

Table 3: The simple-Engle approach - Comparison between estimated and official per capita urban income in 2008 (RMB)

	Estimated data		Official data		Comparison between two samples	
Group	Engel's coefficient	Per-capita income (RMB)	Engel's coefficient	Per-capita income (RMB)	Gap (RMB)	Divergence (%)
Lowest income	0.4816	5,685	0.4814	4,754	931	19.6
Low income	0.4595	8,646	0.4594	7,363	1,283	17.4
Lower middle income	0.4297	13,392	0.4289	10,196	3,196	31.3
Middle income	0.4065	20,941	0.4042	13,984	6,957	49.7
Upper middle income	0.3790	29,910	0.3787	19,254	10,656	55.3
High income	0.3437	47,772	0.3403	26,250	21,500	82.0
Highest income	0.2908	164,034	0.2918	43,614	120,420	276.1
Excluded observations	0.2241	658,811				

Note 1: The tiny deviation between Engel's coefficients of study samples and correspondent official samples has little influence upon the analysis and therefore is treated as equal.

Note 2: "Gap" refers to the amount that estimated income exceeds the official income. "Divergence" refers to the proportion of income gap as percent of the official income.

Source: Our 2009 survey result and statistics (NBS 2009).

Table 4: Estimating four specifications of the supplemented-Engel approach

	(1) Semi logarithm function		(2) Semi logarithm quadratic function		(3) Quadratic function		(4) Cubic function	
	coefficient	t-ratio	coefficient	t-ratio	coefficient	t-ratio	coefficient	t-ratio
$\ln Y$	-0.05739	-28.66**	-0.12004	-4.63**				
$(\ln Y)^2$			0.00295	2.42*				
Y					-7.67E-07	-20.8**	-1.24E-06	-19.31**
Y^2					5.44E-13	13.88**	1.93E-12	12.15**
Y^3							-7.49E-19	-8.99**
city	-0.00664	-3.50**	-0.00677	-3.57**	-0.00385	-1.97*	-0.12508	-2.21*
$city^2$							0.05612	2.22*
$city^3$							-0.00774	-2.28*
edu18	-0.01116	-4.35**	-0.01066	-4.15**	-0.03194	-6.80**	-0.02741	-5.83**
$edu18^2$					0.00117	2.84**	0.00098	2.39*
family	-0.01427	-6.41**	-0.01423	-6.40**	-0.01559	-6.78**	-0.01498	-6.54**
emp	-0.01585	-1.95*	-0.01350	-1.65'	-0.03781	-4.53**	-0.03164	-3.82**
H1	0.07106	11.47**	0.07078	11.43**	0.07601	11.89**	0.07543	11.89**
H2	0.02557	5.66**	0.02544	5.62**	0.02615	5.58**	0.02858	6.12**
L1	-0.03938	-6.06**	-0.03979	-6.13**	-0.03298	-4.93**	-0.03149	-4.74**
C	1.06077	49.76**	1.38627	10.19**	0.5790	37.80**	0.64580	16.57**
Adj.R2	0.2463		0.2472		0.1973		0.2130	
Observations	4,195		4,195		4,195		4,195	

Note: t statistics with ' are significant at 10% level; those with * are at 5% level; those with ** are at 1% level. In every case, Prob.>F is 0.000

Table 5: The supplemented-Engel approach - Comparison between estimated and official per-capita urban income in 2008 (RMB)

Group	Engel's coefficient	Official income	Estimated income: simple-Engel	Estimated income: supplemented-Engel
Lowest income	0.481	4,754	5,685	5,350
Low income	0.459	7,363	8,646	7,430
Lower middle income	0.429	10,196	13,392	11,970
Middle income	0.404	13,984	20,941	17,900
Upper middle income	0.379	19,254	29,910	27,560
High income	0.340	26,250	47,772	54,900
Highest income	0.292	43,614	164,034	139,000
All urban residents	0.379	16,885	35,462	32,154
Left-out observations	0.224		658,811	

Note 1: The RMB 16,885 urban income is an weighted average from the official sample groups, while the published average by NBS is RMB 15,781.

Note 2: The estimated income of all urban residents does not include the left-out samples.

Source: NBS (2009), and authors' estimation.

Table 6: Ratio of estimated income to official incomes in 2005 and 2008

Group	Ratio of estimated income to official income in 2005 (%)	Ratio of estimated income to official income in 2008 (%)	Distribution of hidden income in 2008 (%)
Lowest income	99.1	112.5	0.4
Low income	101.8	100.9	0.0
Lower middle income	106.9	117.4	2.3
Middle income	114.0	128.0	5.1
Upper middle income	130.6	143.1	10.9
High income	138.7	209.1	18.8
Highest income	337.6	318.7	62.5
All urban residents	177.7	190.4	100.0

Note: The estimated incomes are based on the supplemented Engel approach.

Source: NBS (2006, 2009), and authors' estimation.

Table 7: Income changes between 2005 and 2008 from official data and supplemented-Engle approach

	2005	2008	Change %
Per-capita urban disposable income (RMB, official)	11,100	16,885	52.1
Per-capita urban disposable income (RMB, estimated)	19,730	32,154	63.0
Urban population (million)	562	607	7.9
Per-capita rural net income (RMB)	3,537	5,171	46.2
Rural population (million)	745	721	-3.2
Total household disposable income (RMB billion, official)	8,876	13,974	57.4
Total household disposable income (RMB billion, estimated)	13,727	23,237	69.3
Estimated hidden income (RMB billion)	4,851	9,263	91.0
GDP (RMB billion, official)	18,322	31,405	71.4

Note 1: The official data of urban and rural disposable income per capita are derived from the group statistics as weighted averages, which are slightly higher than the official average income data published by the NBS.
Note 2: GDP data is not adjusted.
Source: NBS (2006, 2009), Wang (2007), authors' estimations.

Table 8: GNI before and after two adjustments for hidden income in 2005 and 2008 (RMB trillion)

	Before adjustment		After adjustment	
	2005	2008	2005	2008
Household sector	11.06	17.87	13.73	23.24
Compensation of employees (wage income)	9.28	14.75	9.28	14.75
Non-wage income	1.78	3.12	4.45	8.49
Enterprise sector	3.73	5.61	3.20	4.74
Non-financial	3.60	5.20	3.09	4.39
Financial	0.13	0.41	0.11	0.35
Government sector	3.83	8.20	3.29	6.92
Gross National Income	**18.41**	**31.62**	**20.01**	**34.84**

Note: The national income component data in 2008 before adjustment are estimated by linear projection from previous Funds of Flows Accounts (NBS, 2005-07) with certain price adjustment. The data after adjustment are obtained by allocating the estimated hidden income into each sector.
Sources: NBS (2006, 2009), authors' estimations.

Table 9: Structure of GNI before and after two adjustments for hidden income in 2005 and 2008 (%)

	Before adjustment (%)		After adjustment (%)	
	2005	**2008**	**2005**	**2008**
Household sector	60.1	56.5	68.6	66.7
Compensation of employees (wage income)	50.4	46.7	46.4	42.3
Non-wage income	9.7	9.9	22.2	24.4
Enterprise sector	20.3	17.7	16.0	13.6
Non-financial	19.6	16.4	15.5	12.6
Financial	0.7	1.3	0.5	1.0
Government sector	20.8	25.9	16.4	19.9
Sum as Gross National Income	**101.1**	**100.2**	**101.0**	**100.2**

Note: Income by sectors is disposable income, and their sum differs slightly from GNI. This explains why the sum of these ratios differs slightly from 100%.

Sources: NBS (2006, 2009), this study.