

## Appendix I

### **The 2007 Household Surveys: Sampling Methods and Data Description**

LUO Chuliang, LI Shi, Terry Sicular, DENG Quheng, and YUE Ximing

To track the dynamics of income distribution in China, the Chinese Household Income Project (CHIP) has conducted four waves of household surveys, in 1988, 1995, 2002, and lastly 2007. These surveys were carried out as part of a collaborative research project on incomes and inequality in China organized by Chinese and international researchers, with assistance from the National Bureau of Statistics (NBS). The CHIP project participants and other researchers have analyzed the data from the first three waves and published a wide range of articles, reports, and books. Descriptions of the CHIP surveys and key findings can be found in Griffin and Zhao (1993), Riskin, Zhao, and Li (2001), and Gustafsson, Li, and Sicular (2008). This volume contains analyses based on the data from the fourth wave, 2007, but also uses data from the earlier waves so as to understand trends over time.

Eichen and Zhang (1993) describe the 1988 survey, and Li, Luo, Wei, and Yue (2008) describe the 1995 and 2002 surveys. This Appendix provides basic information about the 2007 survey. The CHIP surveys are closely related to the NBS household survey. Li et al. (2008) discuss how the NBS household survey samples were selected. Additional details about the NBS household surveys can be found in recent NBS statistical reports and publications.

All the CHIP waves contain surveys of urban and rural households. In view of the increased importance of rural-to-urban migration, and because the urban and rural household subsamples do not adequately cover migrants, the 2002 survey added a survey of rural-to-urban migrants. Thus, the 2002 CHIP survey includes three subsamples. The same procedure was adopted for the 2007 survey, which is also composed of three parts: the urban household survey, the rural household survey, and the rural-to-urban migrant household survey.<sup>1</sup> This structure reflects China's urban-rural division and the increased number of rural individuals who have migrated into the urban areas, especially during the last two decades. There are both similarities and differences between the data collection in 2007 and that in the previous three waves, as we will describe below.

In the CHIP surveys the sample sizes for the urban, rural, and rural-urban migrant surveys are not proportional to their shares in the national population. For this reason, for many analyses population-based weights are needed in order to obtain representative results. Moreover, the regional and provincial sample sizes are not proportional to their regional and provincial shares in the national population, thus multi-level weights are needed. Appendix II provides a detailed discussion of weights for the 2002 and 2007 CHIP survey samples. In this Appendix, however, all information is reported without reweighting. The statistics reported in this Appendix are intended to describe the original survey data and may not be representative of China as a whole.

The 2002 and 2007 migrant surveys include various kinds of rural-urban migrants,

some of whom may also be covered in the rural and urban household surveys. The migrant samples include residents of cities with local agricultural residence registrations (*hukou*); this group is also found in the urban samples. The migrant samples also contain temporary and short-term rural migrants with non-local *hukou*; this group is also included in the rural samples. Due to such overlap, analyses combining the migrant sample survey data with the urban or rural survey data may require adjustments to the samples so as to avoid double-counting. Appendix II discusses some ways to address such double-counting. In this Appendix we discuss the entire survey samples, including the types of households that may be double-counted. Again, the statistics in this Appendix are intended to describe the original survey data, without modifications or adjustments.

### **I. Sampling and Sample Sizes**

Table AI.1 presents the sample sizes for the rural, urban, and migrant subsamples of the 2007 survey. The urban survey covered 10,000 households containing 29,262 individuals selected from 302 cities in sixteen provinces, whereas the rural survey covered 13,000 households containing 51,847 individuals selected from 287 counties in sixteen provinces. The migrant survey covered nearly 5,000 households containing 8,404 individuals selected from fifteen cities in nine provinces.

To obtain a nationally representative sample, the provinces were selected from four distinct regions to reflect variations in economic development and geography. Beijing and Shanghai were selected to represent China's large metropolitan cities;

Liaoning, Jiangsu, Zhejiang, Fujian, and Guangdong to represent the eastern region; Shanxi, Anhui, Hebei, Henan, Hubei, and Hunan to represent the central region; and Chongqing, Sichuan, Yunnan, and Gansu to represent the western region. The provinces covered in the urban and rural surveys are almost identical, with the exception that Shanghai is only included in the urban survey and Hebei is only included in the rural survey.

The migrant household survey was conducted in fifteen cities in nine provinces that are also represented in the urban and rural surveys, including Shanghai (a large metropolitan city region); Guangzhou, Shenzhen and Dongguan in Guangdong (eastern region); Nanjing and Wuxi in Jiangsu (eastern region); Hangzhou and Ningbo in Zhejiang (eastern region); Wuhan in Hubei (central region); Hefei and Bengbu in Anhui (central region); Zhengzhou and Luoyang in Henan (central region), Chongqing (western region); and Chengdu in Sichuan (western region). The majority of migrants in China are concentrated in the aforementioned cities.

The 2007 sample size is larger than the 2002 survey, both in terms of numbers of individuals and numbers of households. Compared with the 2002 survey, the 2007 survey covers more provinces in the urban survey but fewer provinces in the rural survey. In the 2007 urban survey Shanghai, Zhejiang, Fujian, and Hunan were added and twelve provinces in the 2002 urban survey were retained—Beijing, Shanxi, Liaoning, Jiangsu, Anhui, Henan, Hubei, Guangdong, Chongqing, Sichuan, Yunnan, and Gansu. Hebei appears in the urban survey in 2007 but not in 2002.

The provinces covered in the rural survey in both 2002 and 2007 are Beijing,

Hebei, Shanxi, Liaoning, Jiangsu, Zhejiang, Anhui, Henan, Hubei, Hunan, Guangdong, Chongqing, Sichuan, Yunnan, and Gansu. Seven provinces, Jilin, Jiangxi, Shandong, Shaanxi, Guangxi, Guizhou, and Xinjiang are excluded from the 2007 survey, but Fujian is added to the 2007 rural survey.

In the 2002 migrant household survey, the sample contains 2,000 rural-urban migrant households from provincial capitals and one or two medium-sized cities in the same provinces that are also included in the 2002 urban survey. The migrant household sample in 2007, as opposed to in the 2002 survey, is drawn from cities where migrants are more concentrated nationally.

<Table AI.1 about here>

In addition to the household and individual surveys outlined above, in some areas surveys were also conducted at the village level to obtain relevant information about the communities where the rural households were located. Village-level variables are available for the 8,000 rural households about which we have information from the CHIP questionnaires (see Section II). Information about the village-level survey data is available on request.

## **II. Data from the CHIP Questionnaires versus Data Provided by the NBS**

The CHIP dataset contains two types of data. One type was collected by the NBS as part of its annual urban and rural household surveys and then provided to the CHIP. The other type was collected through household interviews conducted using independent CHIP questionnaires. The CHIP questionnaires were designed to

supplement the NBS survey data. They contain questions asking for some information that was also collected in the NBS surveys, as well as for some information on variables that are unavailable in the NBS household surveys. With respect to the migrant surveys, because the NBS does not conduct a survey of rural-urban migrants, information in the migrant dataset is based entirely on interviews using the independent CHIP questionnaires.

Unfortunately, not all types of information are available for all households and individuals. The types of data collected for each subsample are summarized in Table AI.2. For the 2007 urban survey, the NBS provides comprehensive data for 10,000 households. These 10,000 households, however, did not answer the CHIP questionnaires.

We note that the CHIP questionnaires were used for an additional sample of 5,000 urban households that we do not include in our description of the 2007 survey sample here. The additional 5,000 households are in nine provinces (Shanghai, Jiangsu, Zhejiang, Anhui, Henan, Hubei, Guangdong, Chongqing, and Sichuan). They had been part of the NBS urban household survey sample in 2006, but due to sample rotation they were not retained in the 2007 survey. For these additional 5,000 households NBS data are not available. As the CHIP urban questionnaire was designed to be matched with the data provided by the NBS, information on these 5,000 households is incomplete and of limited use for analysis of incomes, inequality, and poverty. Consequently, in this Appendix we limit our discussion to the 10,000 households for which NBS data are available. Most of the chapters in this volume

that analyze the urban survey also use only the 10,000 households for which we have NBS data.

For the 2007 rural survey, data from the CHIP rural questionnaire are available for 8,000 households. For these 8,000 households income and expenditure data from the NBS household survey are also available. The NBS also provided its household survey data for an additional 5,000 households, but these additional households were not interviewed using the CHIP questionnaires. In other words, CHIP data and partial NBS data are available for 8,000 households, and NBS data but no CHIP data are available for an additional 5,000 households. However, because comparable data on key household characteristics as well as on household incomes and expenditures are available for all 13,000 households, these two rural subsamples can be combined for analyses of incomes, inequality, and poverty. In this Appendix we discuss the full sample of 13,000 households.

For the migrant sample only CHIP data and no NBS data are available. The CHIP migrant questionnaire was designed accordingly, with questions that collect information comparable to that in the urban and rural household surveys as well as additional information on topics of special relevance to migrants and migration. We note, however, that all the data in the CHIP migrant survey, including data on incomes and expenditures, are based on recall questions. For the urban and rural datasets income and expenditure data are diary-based. In principle, diary-based income data are more accurate than recall data, but recall data can still be useful. Analysts who use the migrant data should be aware of this difference between the data for the

migrant sample and those for the urban and rural samples.

<Table AI.2 about here>

### **III. Characteristics of the 2007 Urban Survey**

The distribution of households and individuals by province in the 2007 urban survey sample of 10,000 households for which NBS data are available is reported in Table A1.3. The sample was designed to draw more households from the more populous provinces. The sample size for each province or region, however, was not strictly proportional to its actual population, so that for some analyses it may be necessary to reweight the sample in order to obtain results that are representative (see Appendix II).

<Table AI.3 about here>

The gender composition in the 2007 urban data, by province, is presented in Table AI.4. Slightly more female individuals than male individuals are included in the overall urban sample. The gender compositions are almost the identical in the 2002 and 2007 surveys, 100:102.6 (male:female) in the 2002 data (Li et al. 2008, p. 348), and 100:102.1 in the 2007 data. Within provinces, however, there are some differences between the two years. For example, in Beijing the percentage of females relative to males is 101 percent in the 2002 survey but 95.5 percent in the 2007 survey. Although we mention these differences, we do not attempt to explain them in detail.

<Table AI.4 about here>



Table AI.5 shows the distribution of households by household size in the 2007 urban survey. More than 57 percent of the sampled households were composed of three members, reflecting the enforcement of the one-child policy in urban China. The proportion of three-member households decreased by 4 percentage points, from 61.6 percent in 2002 to 57.7 percent, in 2007, whereas the proportion of small-size households (fewer than three members) increased slightly. More than 80 percent of the urban sample in 2007 consists of households with two or three members, suggesting that “nuclear households” dominate in the urban sample. The proportion of two- and three-member households is generally higher in the more developed provinces. The average household size in the NBS national urban household survey was 2.91 in 2007, very close to 2.93, the average (unweighted) household size calculated from our urban survey.

<Table AI.5 about here>

Table AI.6 reports the distribution of individuals among different age groups in the 2007 urban survey. Figure A1.1 compares the age-gender profiles in the 2002 and 2007 urban surveys. In 2007 individuals between the ages of 20 and 60 account for 68.8 percent of all sampled individuals. The percentages vary among provinces from 65 percent to 76 percent. In Beijing and Shanghai, for example, the percentages of individuals between the ages of 20 to 60 are higher than that in other provinces. Compared with the 2002 urban data, in 2007 a higher percentage of individuals are in older cohorts and a lower percentage in younger cohorts. This reflects the aging of China’s urban population.

<Table AI.6 about here>

<Figure AI.1 about here>

<Table AI.7 about here>

Table A1.7 reports the distribution of individuals over the age of 15 in the 2007 urban survey by educational attainment. The ratio of those with “junior college” and “college and above” educations increased, from 16.7 percent and 8.7 percent in 2002 to 18.9 percent and 12.1 percent in 2007, respectively. This increase in post-secondary attainment reflects the expansion in the availability of college education since 1999. In Beijing, individuals over the age of 15 with a post-secondary education accounted for more than 43 percent of the individuals in the subgroup; this is much higher than the percentages in the other provinces. In general, educational attainment is higher in the more developed provinces.

#### **IV. Characteristics of the Rural Survey**

Table AI.8 presents the distribution of households and individuals in the 2007 rural sample by province. In order to capture the fact that the population in the rural areas is larger than that in the urban areas and also that rural China is more heterogeneous, the rural survey sampled more households and individuals than the urban survey. The number of households selected in the survey design is based on the population in each province, with the more populous provinces assigned more households. Similar to the urban survey, however, the provincial sample sizes are not consistent with the actual provincial distributions of the population. Consequently, depending on the question

being analyzed, reweighting may be required to obtain representative results (see Appendix II).

Although fewer provinces were surveyed in 2007 than in 2002, in 2007 the sample sizes in terms of number of households and individuals increased and more counties were drawn from within the provinces.

<Table AI.8 about here>

Table AI.9 provides the gender composition of the 2007 rural survey sample, both overall and by province. In contrast to the urban survey (see Table AI.4), the rural survey contains fewer females than males, both overall and in each individual province. This difference reflects the strong preference for males in rural China. The gender composition changed only slightly between the 2002 and 2007 rural surveys. The ratio of females to males was 92.2 percent in 2002 and 93.2 percent in 2007.

<Table AI.9 about here>

Table AI.10 shows the distribution of households by household size in the 2007 rural survey. Households with four members account for 30 percent of all the sampled households, and three- and four-member households together account for 56 percent of all sampled households. The larger household sizes in rural China reflect rural-urban differences in implementation of population planning policies. In most counties rural couples are allowed to give birth to a second child. The average (unweighted) household size in the 2007 rural survey was 4.0 members (as calculated from Table AI.1), which is close to the 4.03 members per household officially reported by the NBS in its annual rural household survey.

<Table AI.10 about here>

The distribution of individuals among different age groups in the 2007 rural survey is presented in Table AI.11 and Figure AI.2. The rural sample is younger than the urban sample. The proportions of individuals between the ages of 0 and 5 and between the ages of 11 and 20 are 4.6 percent and 18.6 percent, respectively, which is much higher than the proportions in the urban data. However, there are relatively fewer working-age individuals in the rural sample than in the urban sample. Individuals between the ages of 20 and 60 account for 63.4 percent of the rural sample, 5 percentage points less than in the urban sample, possibly reflecting that some rural laborers who had migrated to the urban areas were excluded from the rural survey. The treatment of individuals who were away from their homes at the time of the survey is explained in more detail in Appendix II. Similar to the urban sample, between 2002 and 2007 the age distribution of the rural sample shifted upward, reflecting the aging of the population (Figure AI.2).

<Table AI.11 about here>

<Figure AI.2 about here

<Table AI.12 about here>

Table AI.12 gives the distribution of individuals over the age of 15 in the 2007 rural sample by level of educational attainment. As expected, educational attainment in the rural areas is generally lower than that in the urban areas. The majority of rural adults have attended primary or lower-middle school; these two groups account for 37.5 percent and 43.0 percent of all the rural sampled adults, respectively. Compared

with the 2002 rural survey, the share of adults who attended primary school increased by nearly 8 percentage points, whereas the proportion who attended lower-middle school decreased by 2 percentage points.

## **V. Characteristics of the Rural-Urban Migrant Survey**

Since the mid-1990s rural laborers have increasingly migrated to the urban areas to seek employment. Neither the rural nor urban household survey regularly conducted by the NBS, however, adequately captures the rural-to-urban migrants. The NBS rural household survey includes some migrants, as it counts as household members individuals who are present in the household for up to six months or who are away from the household but still maintain strong economic ties with it (see Appendix II). Historically, the NBS urban household survey only covered households with a local *hukou*, but in recent years the NBS has deliberately expanded its urban sample to include households without an urban *hukou*; still, rural-urban migrants are generally underrepresented in the NBS data.

Our 2002 survey includes a sample of rural-to-urban migrant households within the same provinces as the urban sample. In the absence of a sampling frame specifically designed to capture migrants, we relied on neighborhood resident committees, grassroots organizations in urban China, to identify and select the migrant households for our survey. Details about selection of the migrant sample for the 2002 are discussed in Li et al. (2008).

The 2007 migrant survey covers nine of the sixteen provinces covered in the

urban survey. Similar to the urban survey, the provincial sample sizes are not consistent with the actual distribution of the migrant population among provinces. Therefore, depending on the question being analyzed, reweighting may be required to obtain representative results (see Appendix II).

In 2007 the migrant survey was conducted under the auspices of the Rural-Urban Migration in China (RUMiC) project, with assistance from Datasea, a company specializing in market research. The researchers made great efforts to construct the sampling frame, as explained at length in Kong (2010). Here, we will only briefly describe the process. First, each sample city selected was divided into equal-sized blocks averaging 0.25 square kilometers in size, based on up-to-date, equal-scale maps. Second, 10 percent of the blocks were randomly selected. The enumerators listed in a given order all the workplaces within each block and maintained a record of the number of migrants working in each workplace. These records were considered to be the sampling frame, and from this frame the researchers randomly drew the individuals for the migrant survey. The survey questionnaire then asked the migrant respondents for information about their living situations in the cities, including their household situation in the cities, i.e., whether they lived in the city with other people from the household and shared income and expenditures.

The migrant sample includes individuals with agricultural *hukou* who were not living in the location of their *hukou* registrations, including both temporary and long-term migrants.<sup>2</sup> Because temporary migrants are also included in the rural sample as household members, there is a potential for double-counting.

Consequently, several of the analyses in this volume drop the temporary migrants and only use the subsample of long-term, stable migrants. The migrant sample also includes some individuals with agricultural *hukou* in the local urban area. Such individuals are included in the urban sample as well, also creating the potential for double-counting. Appendix II provides a detailed discussion of how to handle potential double-counting between the migrant and the rural and urban surveys..

Tables AI.13 through AI.16 report information from the rural-urban migrant sample, similar to the information reported above for the urban and rural samples. The statistics in these tables are calculated from the full migrant sample, including both temporary and long-term migrants and including individuals with local agricultural *hukou*. Table AI.13 shows the distribution of migrant households and individuals among cities. In the table one can see that the average household size in the rural-urban migrant sample is much smaller than that in the urban and rural household samples. The majority of migrant households are composed of a single person.

Looking at the gender composition in Table AI.14, we find that the migrant sample contains more males than the other two samples, since married women typically return to their places of origin. Not surprisingly, working-age individuals comprise a higher share of the migrant sample.

The age distribution in the 2002 migrant sample is dominated by individuals between the ages of 25 and 40 (Figure AI.3). This probably reflects the sampling method, which in 2002 disproportionately captured longer-term migrants who had settled in urban neighborhoods. In 2007 the migrant sampling method captured

more young migrants between the ages of 17 and 25. Adjustments of the migrant samples, as discussed in Appendix II through weighting and eliminating the short-term migrants, will reduce this discrepancy.

Educational attainment in the 2007 sample of migrants over the age of 15 is shown in Table AI.16. Most of the migrant adults have only a middle-school education, which is similar to the findings from the rural sample and reflects the lower educational attainment among migrants compared to their urban counterparts.

<Table AI.13 about here>

<Table AI.14 about here>

<Table AI.15 about here>

<Figure AI.3 about here>

<Table AI.16 about here>



## References

Eichen, M. and M. Zhang (1993), "Annex: The 1988 Household Sample Survey-Data Description and Availability," in K. Griffin and R. Zhao, eds., *The Distribution of Income in China*, 331-346, New York: St. Martin's Press.

Griffen, K. and R. Zhao, eds. (1993), *The Distribution of Income in China*, Basingstoke: Macmillan.

Gustafsson, B., S. Li, and T. Sicular, eds. (2008), *Inequality and Public Policy in China*, New York: Cambridge University Press.

Kong, Sherry Tao (2010), "Rural-Urban Migration in China: Survey Design and Implementation," in X. Meng, C. Manning, S. Li, and T. N. Effendi, eds., *The Great Migration: Rural-Urban Migration in China and Indonesia*, 135-150, Northampton, MA: Edward Elgar.

Li, S., Luo, C., Wei, Z., and Yue, X. (2008), "Appendix: The 1995 and 2002 Household Surveys: Sampling Methods and Data Description," in B. Gustafsson, S. Li, and T. Sicular, eds., *Inequality and Public Policy in China*, 337-353, New York: Cambridge University Press.

Riskin, C., R. Zhao, and S. Li, eds. (2001), *China's Retreat from Equality: Income Distribution and Economic Transition*, Armonk, NY: M. E. Sharpe.

Table AI.1. *CHIP sample size for each subgroup, 2007*

	Urban	Rural	Rural-to-urban migrants
Individuals	29,262	51,847	8,404
Households	10,000	13,000	4,978
Provinces	16	16	9
Counties/districts/cities*	302	287	15
Regions	4	4	4

*Notes:* For the urban and rural areas, the row marked “region” provides the number of counties and districts covered by the survey. For migrants, this row provides the number of cities covered by the survey. Note that the sample sizes for the urban and migrants reported in this table are slightly different than the sample sizes reported in Appendix II. This is due to different treatment of some duplicates found in the urban dataset and also because here we have dropped the small number of migrant observations for which the personal and household datasets could not be merged.

Table AI.2. *Samples covered by the CHIP and NBS data (number of households)*

	CHIP data only	CHIP and NBS data	NBS data only
Urban	(5,000) <sup>1</sup>	0	10,000
Rural	0	8,000 <sup>2</sup>	5,000
Migrant	4,978	0	0

*Notes:*

1. As discussed in the text, data are available for 5,000 urban households from the CHIP questionnaire but there are no matching NBS data. Because the data for these 5,000 households are incomplete, we do not include them in the data descriptions and tables that follow.
2. Partial NBS data are available for these 8,000 households. See the text for further discussion.

Table AI.3. *Distribution of households in the 2007 urban sample, by province*

Province	Number of counties/districts	Number of households	Number of individuals
Beijing	18	800	2,289
Shanxi	24	600	1,771
Liaoning	52	800	2,244
Shanghai	12	500	1,519
Jiangsu	12	600	1,669
Zhejiang	16	600	1,653
Anhui	11	550	1,572
Fujian	32	800	2,443
Henan	19	650	1,893
Hunan	23	800	1,160
Hubei	8	400	2,327
Guangdong	23	700	2,268
Chongqing	13	400	1,186
Sichuan	9	600	1,740
Yunnan	17	600	1,794
Gansu	13	600	1,734

*Note:* Here, and below, statistics on the urban sample are for the 10,000 urban households for which NBS data are available. The additional 5,000 households for which we only have data from the CHIP questionnaire are not included.

Table AI.4. *Gender composition of individuals in the 2007 urban sample, by province*

Province	Male	Female	Females as a percentage of males
Total	14,478	14,784	102.1
Beijing	1,171	1,118	95.5
Shanxi	902	869	96.3
Liaoning	1,089	1,155	106.1
Shanghai	751	768	102.3
Jiangsu	828	841	101.6
Zhejiang	815	838	102.8
Anhui	782	790	101.0
Fujian	1,195	1,248	104.4
Henan	925	968	104.6
Hubei	579	581	100.3
Hunan	1,160	1,167	100.6
Guangdong	1,110	1,158	104.3
Chongqing	589	597	101.4
Sichuan	849	891	104.9
Yunnan	876	918	104.8
Gansu	857	877	102.3

Table AI.5. *Distribution of households in the 2007 urban sample, by household size and province (%)*

Province	Number of household members							
	1	2	3	4	5	6	≥7	All
Total	1.7	25.1	57.7	10.3	4.6	0.5	0.1	100.0
Beijing	0.0	22.5	70.6	5.1	1.8	0.0	0.0	100.0
Shanxi	1.7	25.7	53.7	14.5	3.8	0.7	0.0	100.0
Liaoning	2.1	34.1	50.5	8.0	4.9	0.4	0.0	100.0
Shanghai	0.4	16.4	69.8	6.4	6.4	0.6	0.0	100.0
Jiangsu	3.2	35.5	47.2	9.2	4.2	0.8	0.0	100.0
Zhejiang	3.5	28.3	60.5	5.0	2.2	0.5	0.0	100.0
Anhui	0.6	25.6	63.6	7.8	2.4	0.0	0.0	100.0
Fujian	1.1	21.3	57.8	11.8	7.5	0.4	0.3	100.0
Henan	1.7	28.2	53.4	11.4	4.8	0.6	0.0	100.0
Hubei	2.3	23.5	61.0	8.8	4.3	0.3	0.0	100.0
Hunan	2.5	26.1	55.4	10.6	4.8	0.6	0.0	100.0
Guangdong	0.3	10.1	65.0	15.9	7.6	0.9	0.3	100.0
Chongqing	1.0	25.8	54.8	13.0	5.3	0.3	0.0	100.0
Sichuan	1.2	29.8	52.3	11.7	4.5	0.5	0.0	100.0
Yunnan	5.0	24.7	47.0	15.3	6.3	1.2	0.5	100.0
Gansu	1.5	24.0	61.5	10.2	2.7	0.2	0.0	100.0

Table AI.6. *Distribution of individuals in the 2007 urban sample, by age group and province (%)*

Province	0-5	6-10	11-20	21-30	31-40	41-50	51-60	61-70	>70	Total
Total	2.7	4.2	11.9	10.7	17.8	21.8	18.5	8.1	4.3	100.0
Beijing	1.2	2.1	10.6	14.2	11.3	24.4	25.7	7.7	2.9	100.0
Shanxi	2.1	5.9	16.0	8.4	23.6	20.4	13.5	6.8	3.2	100.0
Liaoning	1.7	3.2	10.3	10.7	14.7	21.2	22.5	11.1	4.7	100.0
Shanghai	3.1	2.3	9.6	15.0	12.1	20.4	25.9	7.6	4.2	100.0
Jiangsu	2.5	3.2	9.4	10.3	14.9	18.6	21.0	11.2	8.9	100.0
Zhejiang	2.6	4.4	11.4	11.1	17.9	23.4	17.4	7.7	4.1	100.0
Anhui	2.0	3.4	14.1	10.0	15.0	28.9	15.7	6.4	4.6	100.0
Fujian	3.2	4.2	13.1	8.2	20.0	22.3	16.0	8.5	4.5	100.0
Henan	3.0	5.2	12.4	9.6	21.5	18.3	16.4	9.1	4.5	100.0
Hubei	2.8	2.1	10.3	13.9	12.9	22.9	22.2	7.8	5.2	100.0
Hunan	3.4	5.8	10.7	9.6	19.3	20.1	17.9	8.3	5.0	100.0
Guangdong	4.7	6.5	13.1	11.2	23.6	21.6	12.8	4.4	2.2	100.0
Chongqing	1.9	4.1	10.5	11.2	18.0	19.1	22.4	7.9	4.8	100.0
Sichuan	2.2	4.4	10.9	9.2	17.2	22.0	19.3	9.6	5.2	100.0
Yunnan	4.2	3.7	12.4	11.4	18.7	20.8	16.0	9.1	3.6	100.0
Gansu	2.8	5.8	14.0	8.9	21.2	23.5	14.9	5.9	2.9	100.0

Table AI.7. *Educational attainment of individuals over the age of 15 in the 2007 urban sample, by province (%)*

Province	Less than primary school	Primary school	Lower-middle school	Higher-middle school	Professional School*	Junior college	College and above	Total
Total	1.8	6.0	24.5	25.7	11.0	18.9	12.1	100.0
Beijing	0.5	2.2	19.0	23.4	11.6	22.5	20.8	100.0
Shanxi	0.8	4.6	31.8	25.1	11.8	15.2	10.8	100.0
Liaoning	1.5	5.3	31.8	23.1	10.2	19.1	9.1	100.0
Shanghai	1.0	2.6	24.9	29.1	11.4	16.9	14.1	100.0
Jiangsu	2.9	6.7	25.7	26.3	9.5	14.9	14.0	100.0
Zhejiang	1.9	9.5	31.7	22.7	6.4	17.2	10.6	100.0
Anhui	1.3	4.6	21.8	32.0	10.0	20.1	10.2	100.0
Fujian	2.2	7.7	25.8	26.5	14.6	14.8	8.4	100.0
Henan	1.4	3.5	19.6	26.7	11.6	24.1	13.1	100.0
Hubei	1.3	3.3	19.3	26.6	12.0	21.7	15.7	100.0
Hunan	1.8	8.1	24.0	25.3	10.8	18.3	11.7	100.0
Guangdong	1.0	5.9	16.3	30.4	10.7	23.0	12.8	100.0
Chongqing	1.6	7.2	25.4	26.4	8.7	20.2	10.6	100.0
Sichuan	2.7	6.9	24.8	24.2	9.4	20.9	11.0	100.0
Yunnan	4.5	13.6	26.8	16.0	14.4	15.7	9.0	100.0
Gansu	2.6	4.6	23.8	30.1	11.3	17.1	10.4	100.0

\*Here, and elsewhere, “professional school” refers to *zhongzhuan*, a technical/occupational stream of high school.



Table AI.8. *Distribution of households in the 2007 rural sample, by province*

Province	Number of counties	Number of households	Number of individuals
Beijing	13	500	1,717
Hebei	5	500	1,826
Shanxi	35	700	2,777
Liaoning	24	800	2,694
Jiangsu	10	1,000	3,714
Zhejiang	10	1,000	3,426
Anhui	9	900	3,683
Fujian	26	800	3,435
Henan	10	1,000	4,089
Hunan	37	800	3,168
Hubei	10	1,000	4,026
Guangdong	12	1,000	5,082
Chongqing	5	500	1,782
Sichuan	11	1,100	4,163
Yunnan	40	700	3,015
Gansu	30	700	3,250

Table AI.9. *Gender composition of individuals in the 2007 rural sample, by province*

Province	Male	Female	Females as a percentage of males
Total	26,838	25,009	93.2
Beijing	859	858	99.9
Hebei	930	896	96.3
Shanxi	1,436	1,341	93.4
Liaoning	1,386	1,308	94.4
Jiangsu	1,901	1,813	95.4
Zhejiang	1,782	1,644	92.3
Anhui	1,932	1,751	90.6
Fujian	1,792	1,643	91.7
Henan	2,119	1,970	93.0
Hubei	2,080	1,946	93.6
Hunan	1,651	1,517	91.9
Guangdong	2,640	2,442	92.5
Chongqing	928	854	92.0
Sichuan	2,148	2,015	93.8
Yunnan	1,581	1,434	90.7
Gansu	1,673	1,577	94.3

Table AI.10. *Distribution of households in the 2007 rural sample, by household size and province (%)*

Province	Number of household members							
	1	2	3	4	5	6	≥7	All
Total	0.3	12.0	25.8	30.2	19.1	9.1	3.5	100.0
Beijing	0.6	12.2	48.4	24.2	11.8	2.4	0.4	100.0
Hebei	0.4	19.6	27.2	29.6	15.6	5.8	1.8	100.0
Shanxi	0.3	12.6	19.3	37.9	20.1	7.9	2.0	100.0
Liaoning	0.6	21.8	38.4	23.3	11.9	3.8	0.4	100.0
Jiangsu	0.2	15.3	34.1	23.5	19.2	6.3	1.4	100.0
Zhejiang	0.5	18.4	40.6	24.9	10.9	3.9	0.8	100.0
Anhui	0.7	9.7	21.6	33.4	22.0	9.4	3.2	100.0
Fujian	0.0	5.6	20.9	35.9	21.6	10.0	6.0	100.0
Henan	0.4	10.6	16.9	39.0	20.2	11.0	1.9	100.0
Hubei	0.2	9.9	25.1	32.8	20.2	9.0	2.8	100.0
Hunan	0.0	12.3	23.5	34.8	18.3	8.9	2.4	100.0
Guangdong	0.3	2.4	8.0	26.8	28.3	19.1	15.1	100.0
Chongqing	0.2	20.6	33.2	24.6	14.8	4.6	2.0	100.0
Sichuan	0.5	14.7	32.9	23.1	17.9	9.1	1.8	100.0
Yunnan	0.1	6.9	18.1	38.0	21.3	10.1	5.4	100.0
Gansu	0.0	4.1	12.6	31.4	27.1	18.3	6.4	100.0

Table AI.11. *Distribution of individuals in the 2007 rural sample, by age group and province (%)*

Province	0-5	6-10	11-20	21-30	31-40	41-50	51-60	61-70	>70	Total
Total	4.6	4.7	18.6	15.3	15.3	17.5	15.3	5.9	2.9	100.0
Beijing	1.9	2.6	17.3	13.3	15.8	24.3	16.2	6.2	2.6	100.0
Hebei	4.8	4.6	16.5	13.2	14.7	18.1	19.7	6.0	2.6	100.0
Shanxi	3.1	4.6	22.5	15.2	14.5	19.0	14.5	4.9	1.7	100.0
Liaoning	3.0	3.4	14.9	13.1	14.3	21.0	19.9	6.6	3.8	100.0
Jiangsu	4.2	4.3	17.2	12.1	18.6	18.4	15.5	6.7	3.1	100.0
Zhejiang	3.7	4.6	13.1	13.4	14.1	21.3	19.7	7.0	3.2	100.0
Anhui	5.7	4.2	20.6	15.5	15.4	16.1	14.0	5.7	2.7	100.0
Fujian	4.4	4.3	18.9	18.5	16.7	17.4	12.5	4.7	2.7	100.0
Henan	6.0	5.9	21.7	14.4	16.3	15.6	12.9	5.2	2.2	100.0
Hubei	5.0	2.8	17.9	19.2	12.4	19.5	16.9	4.2	2.2	100.0
Hunan	4.6	4.4	18.2	16.8	13.3	16.5	17.1	6.3	3.0	100.0
Guangdong	5.5	6.2	22.8	19.2	11.5	14.5	12.8	4.4	3.2	100.0
Chongqing	4.8	3.5	12.1	13.3	15.8	16.0	21.0	9.2	4.4	100.0
Sichuan	4.5	5.1	14.8	12.2	18.8	14.7	18.7	8.0	3.4	100.0
Yunnan	5.3	6.4	19.4	17.0	16.1	16.3	10.2	5.8	3.5	100.0
Gansu	4.4	5.6	23.5	13.1	18.0	17.0	9.8	6.1	2.6	100.0

Table AI.12. *Educational attainment of individuals over the age of 15 in the 2007 rural sample, by province (%)*

Province	Less than primary school	Primary School	Lower-middle school	Higher-middle school	Professional school	Junior college and above	Total
Total	3.8	37.5	43.0	11.3	2.4	2.1	100.0
Beijing	3.3	7.7	44.7	20.3	11.0	13.1	100.0
Hebei	0.0	40.4	48.5	9.0	1.1	1.0	100.0
Shanxi	3.7	20.2	54.7	16.1	2.5	2.8	100.0
Liaoning	3.9	21.4	55.7	10.5	2.9	5.6	100.0
Jiangsu	0.0	44.3	41.0	11.1	2.2	1.3	100.0
Zhejiang	0.0	49.3	35.7	11.0	2.3	1.7	100.0
Anhui	0.0	39.8	51.2	6.9	1.1	1.1	100.0
Fujian	9.7	23.9	40.8	16.5	4.5	4.5	100.0
Henan	0.0	43.8	44.0	9.8	2.1	0.4	100.0
Hubei	0.0	47.9	41.4	8.5	1.3	0.8	100.0
Hunan	4.8	29.5	45.5	15.5	2.8	1.9	100.0
Guangdong	0.0	42.1	46.0	10.1	1.4	0.4	100.0
Chongqing	0.0	57.2	33.9	7.6	0.9	0.3	100.0
Sichuan	0.0	55.3	36.1	6.9	0.9	0.9	100.0
Yunnan	18.4	35.0	35.7	7.7	2.1	1.1	100.0
Gansu	18.6	25.5	35.3	15.8	2.7	2.1	100.0

Table AI.13. *Distribution of households and individuals in the 2007 rural-urban migrant sample, by city*

City	Number of individuals	Number of households	Averaged household size	Number of household members (%)				
				1	2	3	≥4	Total
Total	8,404	4,978	1.7	60.5	18.9	13.7	6.9	100.0
Guangzhou	617	400	1.5	66.0	18.5	12.3	3.3	100.0
Dongguan	427	272	1.6	66.9	17.3	9.9	5.9	100.0
Shenzhen	365	302	1.2	84.4	12.3	1.7	1.7	100.0
Zhengzhou	658	350	1.9	56.9	14.9	14.9	13.4	100.0
Luoyang	366	200	1.8	59.5	12.5	17.5	10.5	100.0
Hefei	705	350	2.0	47.4	20.3	19.4	12.9	100.0
Bengbu	428	200	2.1	43.5	18.5	22.0	16.0	100.0
Chongqing	682	400	1.7	56.3	25.3	12.0	6.5	100.0
Shanghai	852	503	1.7	58.3	20.5	15.7	5.6	100.0
Nanjing	611	400	1.5	65.3	22.8	7.0	5.0	100.0
Wuxi	331	200	1.7	63.5	12.0	20.0	4.5	100.0
Hangzhou	639	400	1.6	58.0	28.5	10.5	3.0	100.0
Ningbo	331	200	1.7	61.0	20.5	11.5	7.0	100.0
Wuhan	692	400	1.7	61.5	12.8	18.8	7.0	100.0
Chengdu	700	401	1.7	57.9	18.7	16.7	6.7	100.0

Table AI.14. *Gender composition of individuals in the 2007 rural-urban migrant sample, by city*

Province	Male	Female	Females as a percentage of males
Total	4,777	3,627	75.9
Guangzhou	343	274	79.9
Dongguan	257	170	66.1
Shenzhen	235	130	55.3
Zhengzhou	386	272	70.5
Luoyang	220	146	66.4
Hefei	416	289	69.5
Bengbu	251	177	70.5
Chongqing	367	315	85.8
Shanghai	484	368	76.0
Nanjing	342	269	78.7
Wuxi	153	178	116.3
Hangzhou	381	258	67.7
Ningbo	184	147	79.9
Wuhan	378	314	83.1
Chengdu	380	320	84.2

Table AI.15. *Distribution of individuals in the 2007 rural-urban migrant sample, by age group and city (%)*

City	0-5	6-10	11-20	21-30	31-40	41-50	51-60	61-70	>70	Total
Total	4.5	3.3	17.2	32.5	24.9	12.8	3.8	0.8	0.2	100.0
Guangzhou	3.7	2.6	14.3	36.1	25.9	14.3	2.6	0.5	0.0	100.0
Dongguan	4.9	2.8	11.9	42.6	24.6	8.7	3.3	0.9	0.2	100.0
Shenzhen	1.9	0.3	20.6	51.0	15.6	7.4	3.3	0.0	0.0	100.0
Zhengzhou	8.5	4.3	19.6	29.5	20.5	9.4	5.5	2.4	0.3	100.0
Luoyang	7.7	4.4	24.0	25.4	22.1	12.0	2.7	0.8	0.8	100.0
Hefei	5.0	4.3	21.7	25.3	26.4	11.8	5.3	0.4	0.0	100.0
Bengbu	6.3	6.1	20.8	24.3	29.0	8.9	3.3	1.2	0.2	100.0
Chongqing	2.8	3.2	14.8	30.1	27.9	16.7	3.4	0.7	0.4	100.0
Shanghai	3.6	2.9	14.3	32.8	27.4	16.7	2.1	0.2	0.0	100.0
Nanjing	2.6	1.6	19.8	36.2	21.1	11.5	5.7	1.3	0.2	100.0
Wuxi	4.5	3.0	16.3	38.1	16.9	17.8	3.3	0.0	0.0	100.0
Hangzhou	3.1	2.7	13.0	36.8	26.8	13.3	3.8	0.6	0.0	100.0
Ningbo	5.1	2.7	15.1	27.8	26.9	15.4	5.7	1.2	0.0	100.0
Wuhan	5.2	4.1	17.5	34.5	21.7	14.0	2.5	0.6	0.0	100.0
Chengdu	4.0	4.0	17.4	25.3	32.4	10.9	4.1	1.1	0.7	100.0



Table AI.16. *Educational attainment of individuals over the age of 15 in the 2007 rural-urban migrant sample, by city (%)*

City	Less than Primary school	Primary school	Lower- Middle school	Higher- middle school	Professional school	Junior college	College and above	Total
Total	2.5	13.3	55.3	19.8	5.2	3.3	0.6	100.0
Guangzhou	1.6	8.4	60.0	22.2	3.6	3.2	0.9	100.0
Dongguan	1.1	10.2	57.3	23.3	4.2	3.4	0.5	100.0
Shenzhen	0.6	7.1	49.1	26.9	10.9	5.1	0.3	100.0
Zhengzhou	3.2	14.5	46.2	18.3	10.3	7.0	0.6	100.0
Luoyang	1.4	7.1	62.6	18.2	6.7	2.7	1.4	100.0
Hefei	6.0	18.0	54.1	14.5	3.7	3.4	0.4	100.0
Bengbu	6.6	18.2	57.6	11.9	3.6	1.5	0.6	100.0
Chongqing	1.5	17.0	54.1	20.6	3.6	2.9	0.3	100.0
Shanghai	2.2	13.9	57.6	18.0	4.9	2.2	1.2	100.0
Nanjing	3.4	9.5	54.2	23.5	5.7	3.2	0.5	100.0
Wuxi	0.7	6.9	75.9	12.1	2.8	0.7	1.0	100.0
Hangzhou	2.2	15.2	49.0	25.7	2.0	5.4	0.5	100.0
Ningbo	3.6	25.8	46.4	19.2	4.0	0.7	0.3	100.0
Wuhan	0.9	8.2	60.0	19.3	8.4	2.9	0.5	100.0
Chengdu	2.5	17.2	52.7	19.3	5.1	3.0	0.2	100.0

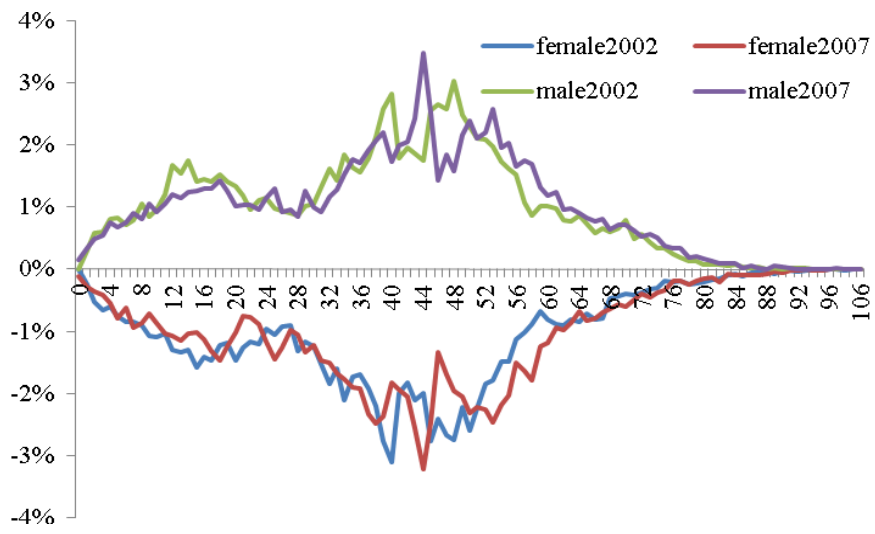


Figure A1.1 Age-Gender Profiles, Urban

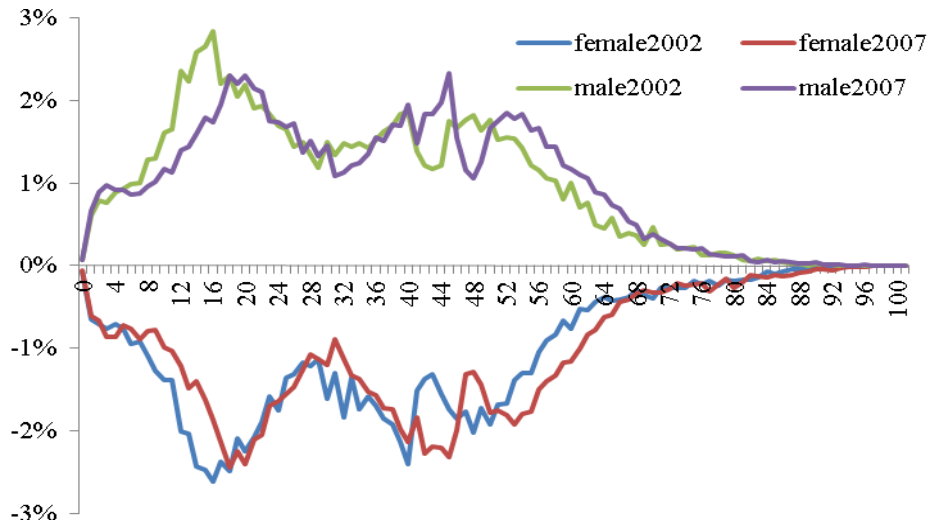


Figure AI.2 Age-Gender Profiles, Rural

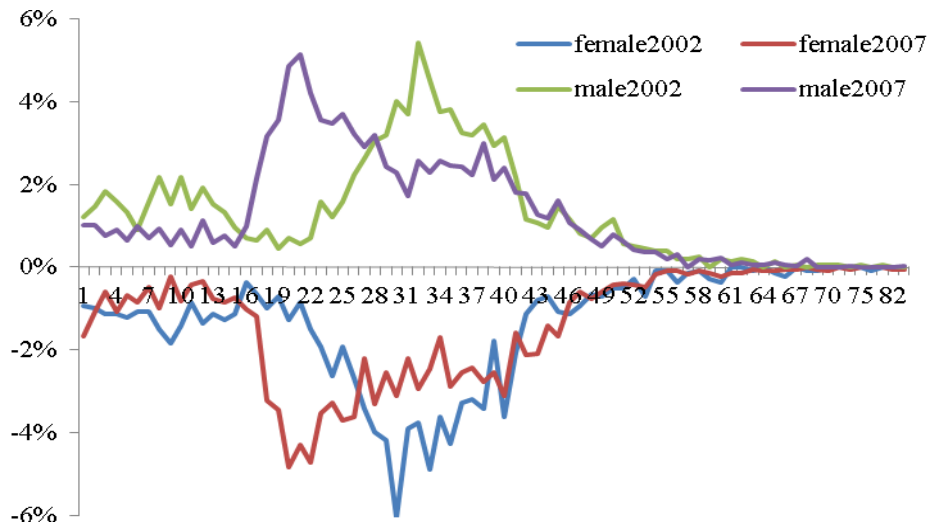


Figure AI.3 Age-Gender Profile, Migrants

---

<sup>1</sup> The 2002 surveys were carried out by the NBS. The 2007 urban and rural surveys were conducted by the NBS, but the rural-urban migrant survey was conducted by a survey company. The 2007 survey is also a part of the larger RUMiCI (Rural-Urban Migrants in Indonesia and China) survey project. The sampling procedure and survey method for the 2007 migrant survey are described in detail in the Rural-Urban Migration in China and Indonesia Project Survey Documentation. See <http://rse.anu.edu.au/rumici/documentation.php>, accessed September 3, 2011.

<sup>2</sup> In 2007 the *hukou* reform was only implemented in several provinces. For the most part, the *hukou* was still classified as either agricultural or nonagricultural.