

# The Quantity and Quality of Labor in China 1978-2000-2025

(May 2005)

Carsten A. Holz

The purpose of this manuscript is to construct the economy-wide number of laborers by “age times education” (at each age with a break-down by education) for each year of the reform period (1978-2003), as well as to forecast the number of laborers by age times education through 2025. These are the raw data for use in the construction of measures of the quantity and quality of labor.

The manuscript starts with an introduction to how data on population (including laborers) are collected and evaluates the age-specific data. The second and third section examine the changes to the classification into different education categories over time and the changes in the rate of completion of a particular level of schooling. The fourth section explains how missing education-specific data on *laborers* (by age) in key years are constructed, while the fifth section derives data on total employment. The sixth section combines the available data from censuses and 1% sample surveys to construct age- and education-specific data on laborers for each year of the reform period, from 1978-2003; the seventh section creates forecasts of the number of laborers by age times education through 2025.

The results, the aggregate quantity of laborers and various human capital measures for each year are reported in four tables: in Table 8 and Table 9 for the period 1978-2003, and in Table 12 and Table 13 for the period 2000-2025.

The term “population” in the following refers to all persons, the term “laborers” to those defined as laborers in the different sources. “Census” refers to a survey of the total population or of all laborers, and “sample” to a survey of a subset of the total population or a subset of all laborers. The term “population” is not used in the sense of “sample vs. population” or in the sense of “sampling population.”

## **1. Introduction to the Population Censuses, 1% sample surveys, and 1‰ sample surveys**

Since the beginning of the economic reform period in 1978, China has conducted three economy-wide population censuses, in 1982, 1990, and 2000, and two population 1% sample surveys, in 1987 and 1995. At least since 1988, data from a 1‰ sample survey on population change are published for those years when no census or 1% sample survey was done. Censuses and 1% sample surveys cover a wide range of information, including information on labor. The 1‰ sample surveys are likely to also collect information on labor but that information is not published. Table 1 summarizes the information available in the three census and two 1% sample surveys as relevant for labor and human capital data (and used in the following).

## *Census and 1% sample survey days*

The census/ survey day in 1982, 1987, and 1990 was 1 July (0:00 am); in 1995 and 2000 it was 1 November (0:00 am). (*Census 1982*, p. 4; *1990*, Vol. 1, p. 1; *2000*, Vol. 1, p. I; *Survey 1987*, p. 1; *1995*, p. 1) The survey day for the 1% sample surveys is probably end-year.

For comparisons across censuses/ surveys the 1995 survey and 2000 census data are transformed into midyear (1 July) values.<sup>1</sup> This is achieved, for each age group X (number of persons or laborers age X), through the following transformation:

$$X_{1 \text{ July}} = 2/3 * X_{1 \text{ Nov.}} + 1/3 (X+1)_{1 \text{ Nov.}} .$$

This assumes, at each age-level in the 1995 survey and the 2000 census, that the distribution of total birthdays (at that age level) occurring between 1 November of any year and 31 October of the following year is uniformly distributed across all days of this 365-day period.

In employment statistics, the age-specific data typically end with a summary value for all persons age 65 and over (“65+”). In this case, the group of 64-year olds at 1 July 1995 and 1 July 2000 is obtained as  $2/3 * 64\text{-year-olds}_{1 \text{ Nov.}} + \alpha * 1/3 * 64\text{-year-olds}_{1 \text{ Nov.}}$ , where  $\alpha$  is an adjustment factor obtained from the previous-age cohort. Specifically,

$$\alpha = 64\text{-year-olds}_{1 \text{ Nov.}} / 63\text{-year-olds}_{1 \text{ Nov.}} .^2$$

The group of 64-year olds at 1 July, thus, is obtained as

$$2/3 * 64\text{-year-olds}_{1 \text{ Nov.}} + 1/3 * ( [64\text{-year-olds}_{1 \text{ Nov.}}]^2 / 63\text{-year-olds}_{1 \text{ Nov.}} ) .$$

The age category  $65_{+1 \text{ July}}$  is set equal to  $65_{+1 \text{ Nov.}}$ . Ideally, when the 65+ category refers to the total population,  $\alpha * 1/3 * 64\text{-year-olds}_{1 \text{ Nov.}}$  should be subtracted out, and those who died between 1 July (0:01 am) and 31 Oct. 2000 (11:59 pm) be added in; the latter data could be approximated. When the 65+ category refers to laborers,  $\alpha * 1/3 * 64\text{-year-old}_{1 \text{ Nov.}}$  laborers should be subtracted out, and those who retired between 1 July (0:01am) and 31 Oct. 2000 (11:59 pm) be added in; the latter data are not available. The approximations which would have to be done in the first case, of the total population, and the guesses which would have to be made in the second case, of laborers, are unlikely to produce a gain in accuracy over simply setting the age category  $65_{+1 \text{ July}}$  equal to  $65_{+1 \text{ Nov.}}$ .<sup>3</sup>

---

<sup>1</sup> The ideal measure of laborers in production function estimations is the annual average. The best available estimator of the annual average number of laborers is the midyear value.

<sup>2</sup> A double-check with the *population* data for 58-, 59-, and 60-year-olds in 1990 (relevant for the comparison in 1995) and for 53-, 54-, and 55-year-olds in 1990 (relevant for the comparison in 2000) shows that the size of these cohorts within each of these two groups of years is rather similar, so that the correction factor is minimal.

Ideally, the correction would be based on the relationship between the year 2000 (and 1995) age 64 and 65 age cohorts 10 years (5 years) earlier, in the 1990 census. This is not done because the 1990 census does not have laborer data by age times education. This means different procedures would have to be used depending on the type of variable on which year 2000 (or 1995) age-64 values are sought. Instead, the one procedure discussed in the text here is used throughout.

<sup>3</sup> Due to the transformation of the youngest age group (age 15 in the case of laborers) and of the 64-year age group (when the data end with a 65+ group), the transformed total across age groups does not equal the original total across age groups.

## *Reliability of age-specific population data in surveys and censuses*

If the censuses, 1% sample surveys, and 1‰ sample surveys were perfectly reliable, the size of age cohort X in year Y should equal the size of (the same) age cohort X+1 in year Y+1, less the number of those who died during the year. However, the 1‰ sample survey data do not appear reliable and are therefore not used in the construction of laborers below. The population age distribution in the 1% sample surveys of 1987 and 1995 have shortcomings. The census data appear the most consistent.

Across the 1‰ sample surveys several patterns persist: (i) of those aged 0 in year Y, only 79% to 93% survive to age 1; (ii) around age 4-6, and then around 12, the survival rates exceed one (new persons appear); (iii) around age 17-19, several percentage points of a cohort disappear each year, to reappear in the early 20s; (iv) the 49-year cohort tends to grow by 3-15% when it turns 50, and then shrinks correspondingly when it turns 51; (v) the same pattern is repeated at age 59-61.<sup>4</sup> The first two patterns reflect the underreporting of births, with unreported births later surfacing at school age. The third pattern could be related to entry into / exit from the military if military personnel were not included in the population data (as they are unlikely to be). The fourth and fifth patterns, given the age groups, may be related to retirement, but it is not clear how.

The 1‰ sample surveys can also be contrasted with the 1% sample surveys and the censuses. Figure 1 shows the cohort growth rates of 1989 (compared to the 1987 1% sample survey, since data for annual age groups are not available in the 1988 1‰ sample survey), 1990 (vs. 1989), and 1991 (vs. 1990). The figure suggests that if the 1990 census data were correct, then the 1991 1‰ sample survey over-sampled the older age groups in comparison to the 1990 census. The data also indicate a problem around the 28-year olds in 1990, further examined below. While the 1987 1% sample survey and the 1990 census were conducted on 1 July, it is unclear when the 1‰ sample surveys of 1989 and 1991 were conducted; there could be a problem of time inconsistency if the latter were conducted at end-year.<sup>5</sup> Figure 2 makes the same comparisons a decade later, for 1999 (vs. 1998), 2000 (vs. 1999), and 2001 (vs. 2000). The 1‰ sample surveys, in comparison to the (1 Nov.) 2000 census, under-sample the age groups 16 through 28, and then over-sample the age groups above 40, and increasingly so, except at the octogenarian stage.

The comparisons in Figure 1 and Figure 2 raise questions about the age distribution of the population in the 1‰ sample surveys, as do the patterns noted above. Given these problems, and given that the 1‰ sample surveys have nothing more to offer than population data by age, they are not further used here.

Figure 3 compares the 1990 census data to the 1987 1% sample survey data (both are 1 July data) and the 2000 census data to the 1995 1% sample survey data (both are 1 Nov. data). If the year 1990 census data were correct, then Figure 3 implies under-sampling in the 1987 1% sample survey of the age groups 3-6 and 22-25. If the year 2000 census data were correct,

---

<sup>4</sup> The 1‰ sample surveys are compared under the assumption that they are conducted on the same day of each year.

<sup>5</sup> All 1% and 1‰ sample survey data are augmented to nationwide levels (multiplied by a factor of approximately 100, and 1000), following the end-year population data from the *Statistical Yearbook 2001*, p. 91, and *2004*, p. 95 (for years since 1989); the augmentation in the case of the 1987 1% sample survey is based on a calculated mid-year value (linear interpolation).

then Figure 3 also implies under-sampling in the 1995 1% sample survey, on a much larger scale than in 1987, of the age groups 6, 16, 22-30, 32, and 38.

Comparing the presumably accurate census data, for 1982, 1990, and 2000, also reveals some inconsistencies (Figure 4). One inconsistency is the well-known underreporting of births. The 1990 census reveals significant positive growth in the 0-2 and 4 year age categories in 1982 (age 8-10, 12 in 1990), and the 2000 census, on an even larger scale, in the 0-7 and 9 year age categories of 1990 (age 10-17, 19 in 2000). Between 5 and 10% of births are not initially reported. Apart from this problem, the comparison of the 1990 and the 1982 censuses only shows abnormal (positive) growth for the age categories 28 and 29 in 1990; the comparison of the 2000 and the 1990 censuses only shows abnormal (positive) growth for the age categories 30, 32, 38 and 39 in 2000 (with two slight outliers at age 50 and 51). Overall, the transition between censuses is smoother than between 1% sample surveys and censuses (Figure 3 vs. Figure 4); the variation around the trend is also smaller. If the censuses provided the most reliable data, this suggests that the population age distribution in the 1% sample surveys has some shortcomings.

The age at which the data on laborers start is 15. The only significant problem across the censuses in age groups above age 15 is the cohorts age 20 and 21 in 1982, i.e., 28/29 in 1990, and 38/39 in 2000. As Table 2 shows, the cohort born between 1 July 1961 and 30 June 1962 keeps growing over time, especially in the year 2000 (then age 38). The cohort born a year later also grew strongly, by about 10%, in both 1990 and 2000. The surrounding age cohorts also occasionally grow, but not on that scale.<sup>6</sup> Without any clue as to which data are correct, the census population data, and therefore also the census labor data, cannot be adjusted.

#### *Reliability of age-specific labor data in 1% sample surveys and censuses*

Census and 1% sample survey data appear reliable when labor data are compared to population data, and perhaps even when age distributions across laborers are compared.

Figure 5 reports, at each age, the ratio of ‘the share of laborers at the particular age in all laborers’ divided by the ‘share of the population at the particular age in the total population,’ in the following labeled “participation share ratio.” The overall pattern is appropriate: the young constitute a smaller share of laborers than of the population, the middle-aged a larger share of laborers than of the population, and the old again a smaller share of laborers than of the population. Across the 1% sample surveys and the censuses, a clear trend towards lower participation share ratios at young age and higher ratios at old age emerges. This shift over time is consistent across the mix of censuses and 1% sample surveys. The participation structure across age for each individual census and each individual 1% sample survey appears plausible.

Figure 6 reports, at each age, the employment rate, i.e., the number of laborers divided by the number of persons (population). The overall pattern and the shifts in the individual patterns over time are identical to the case of the participation share ratio. The smooth shifts and age structures in Figure 5 and Figure 6 also imply that the participation share ratios and employment rates based on the 1982 official 1% sample of laborers—when no data on

---

<sup>6</sup> Movements in and out of the military are too small to significantly affect the data in the table. Age 28 military personnel in 1990 numbered 67,158, age 38 military personnel in 2000 17,586; for the 29/39 groups the figures are 61,521 and 14,539.

laborers in the complete census are available—are reliable. Figure 7, for 2000, when the data are available, reveals plausible patterns for the unemployed and the residual population (neither employed nor unemployed): while formal unemployment is low throughout all age categories, the residual is large at the young age categories (schooling) as well as at the old age categories (retirement).

Figure 8 shows the age-specific share of laborers (in all laborers), for all censuses and 1% sample surveys. Figure 9 moves the data from all three censuses and two 1% sample surveys to one age standard, namely age in year 2000 (and, consequently, in 1995 5 years less than shown on the horizontal axis, etc.). The charts clearly show the severe impact of the Great Leap Forward (very low birth rates) on the shares of different age groups in total laborers. The sharp drop and then recovery of birth rates (number of laborers) gradually shifts through the figure as the censuses and surveys progress in time; the shifts occur consistently across the censuses and sample surveys, again confirming the quality of these data. The 1987 and 1995 1% sample survey data are perhaps slightly erratic in comparison to the census data.

If one assumes that the census data are (the most) accurate, these results suggest that the age-specific data on laborers, relative to the population, i.e., the employment rates, in the 1% sample surveys of 1987 and 1995 are also reliable. The age distribution of laborers in the 1% sample surveys, when compared to the censuses, are not as reliable but still appear acceptable. (The population age distribution in these surveys, examined in Figure 3, was also less than perfect.)

## 2. Education categories

Table 3 shows the education categories on which data were collected in the three censuses and two 1% sample surveys. (The translation of Chinese terms is appended at the end of the manuscript.) The education level always refers to the highest education level achieved. While there is a good equivalence at the secondary school level across censuses/ surveys (folding the special middle school into the upper middle school), it is not as good below and above the secondary level.

### *Below secondary school level*

Below the primary school level, the practice differs between 1982/90 and 2000. The 1982 and 1990 censuses included a category “knows no or few characters,” defined as knowing less than 1500 characters. In 1982, a person age 6-11 cannot “know no or few characters;” presumably persons aged 6-11 are considered primary school educated. In 1990, a person currently in primary school cannot “know no or few characters,” while a person who has no schooling but knows 1500 or more characters, can read popular books and newspapers, can write brief notes, and has reached literacy level, is considered “primary school” educated.

The published 1982 and 1990 census data and the 1987 sample survey data do not include data on the category “knows no or few characters.”<sup>7</sup> What they provide is data on the number of “illiterate or semi-illiterate” persons, even though none of the three censuses and two

---

<sup>7</sup> The *Population Statistical Yearbook 1993*, as an exceptional event, reports the 1990 census data on the number of persons who “know no or few characters” at all ages. These data are identical to the data on the number of illiterate and semi-illiterate persons in the 1990 census for the ages 15 and higher.

surveys includes a question on “illiterate or semi-illiterate” persons. By explicit definition in the 1982 census, persons who know no or few characters and are age 12 or above are illiterate or semi-illiterate. I.e., at age 12 and above the two categories, knowing no or few characters and being illiterate or semi-illiterate, are identical.<sup>8</sup> Similarly, in 1990, persons who know no or few characters and are *age 15* or above are tautologically illiterate or semi-illiterate.<sup>9</sup>

The 2000 census and the 1995 sample survey use the question “Knows characters?” (*shifou shizi*) as a separate question with partial filter character; if the answer is no, the question on education level can only be answered as “no schooling,” “literacy class,” or “primary school.” On the other hand, if a person is currently in primary school, the answer to the “Knows characters?” question is automatically “yes.”

For the 1995 1% sample survey, the published data cover the six education categories (for total population and total laborers separately): no schooling, literacy class, primary school, lower middle school, upper middle school, and university. These education categories are exhaustive, i.e., the sum of the number of persons (or laborers) across education categories, at each age, adds up to the population total (or total laborers). For the population, a separate table labeled “illiterate and semi-illiterate persons” age 15 through 65+ offers data on such persons in the categories no schooling, literacy school, and primary school. These data presumably reflect all those persons who answered the question “Knows characters?” to the negative.<sup>10</sup> The category illiterate and semi-illiterate persons, thus, no longer constitutes an independent “education” category.

The category illiterate and semi-illiterate persons could be re-integrated into the education scale by subtracting the number of persons who are illiterate, listed as having no schooling, literacy school, and primary school, from the corresponding categories in the education classification. One would then end up with the new education categories illiterate and semi-illiterate, no schooling (no longer including those without schooling who are illiterate), literacy class (with the same exclusion), primary school (with the same exclusion), and lower middle school etc. as before. Figure 10 shows that the persons listed in the population-wide education tables of 1995 as having no schooling are virtually all illiterate (or semi-illiterate), that around 12% of those with literacy school are illiterate (right scale in the figure), and that around 1% of those with primary school are illiterate (right scale in the figure).

The published 2000 census data also include data on “no schooling” and on “literacy class,” as well as, separately, on illiterate (not mentioning “semi-illiterate”) persons. The

---

<sup>8</sup> In the *Census 1982*, the total population less the population below age 6 is 889,788,266 (calculated from p. 272). This compares to the population age 6 and above in the education table (the education table is limited to the population starting age 6) of 889,787,495 (p. 360), i.e., 771 persons less (presumably of no relevance); the education table lists the exhaustive 6 education categories from university graduates to primary school and “illiterates and semi-illiterates.” A separate table on illiterates and semi-illiterates starts at age 12 and excludes the illiterates and semi-illiterates age 6-11 listed in the education table (pp. 362f.).

<sup>9</sup> In the *Census 1990*, the total population at age 15 and above is 817,508,784 (Vol. 2, p. 2). An identical total number appears in a separate table on illiterate and semi-illiterate persons which starts at age 15 only (Vol. 2 pp. 278f.). Of these, 181,609,097 persons are illiterate and semi-illiterate. The difference of 635,899,687 equals the population at age 15 and above in the education table, a table which comprises the education categories primary school to university only (Vol. 2, p. 112). The total figures in the education table, across primary school through university, for the ages *6 through 14* do not add up to the total population, which suggests that, as in 1982, illiterate and semi-illiterate persons at ages 6-14 exist, but these data are not available in 1990 except as residual of the total population and the population classified by education categories.

<sup>10</sup> The explanations to the survey also suggest so. *Survey 1995*, p. 645, further explains the question with, among others, “not having attained the above-illiteracy level” set by the State Council.

category “illiterate” again does not appear in the census form. One possibility is that it covers all those who answered the question “Knows characters?” to the negative.<sup>11</sup> The published data only come as a total, unlike in 1995, with no breakdown into no schooling, literacy school, and primary school; the illiterate persons can thus not be backed out of the exhaustive education categories covering the total population. In the exhaustive education classification in 2000, the number of persons with no schooling, at age 15 and above, is equivalent to 96.9027% of the number of persons who are illiterate at age 15 and above (the data on illiterate persons start only at age 15); the number of persons with no schooling *or* only literacy class, at age 15 and above, is equivalent to 120.7754% of the number of persons who are illiterate at age 15 and above. (Also see Figure 11 for the age-specific, reverse ratio.) It is possible that in 2000 the persons who have no schooling or only literacy class, while just exceeding the number of illiterate persons, are approximately equal to the number of illiterate *and* semi-illiterate persons, if this category were still in use.

These changes between the censuses have two overall effects. First, the category “primary school” is more comprehensive in 1982 and 1990 than in 2000. In 1982 and 1990, primary school appears to include all persons aged 6-11; furthermore, at least in 1990, primary school includes those who have never attended primary school, are older than age 11, but have achieved basic literacy. In 2000, primary school appears limited to those actually attending (or having attended) primary school. In 2000, the categories no schooling and literacy class capture some of those types of persons who in the earlier census would have been included in the primary school category.

Second, while the 1982 and 1990 census data include data on “illiterate and semi-illiterate” persons, the 2000 census data only include data on “illiterate” persons, and furthermore no longer as one of the exhaustive education categories. However, there could be a close correspondence between the “illiterate and semi-illiterate” category in 1982 and 1990 (as well as in 1987 and 1995) and the sum of “no schooling” and “literacy class” in 2000, with two limitations. First, some persons who have never been to primary school but are considered literate in 1990, or were 6-11 year olds in 1982, were added to the primary school category in those years, while in 2000 these persons would presumably have entered the “no schooling” or “literacy class” category. Second, a very few primary school persons in 2000 are likely to be illiterate (are reported illiterate in 1995) and might have entered the category illiterate and semi-illiterate persons in 1982 and 1990.

In the following, when linking the three censuses and the two surveys, the primary school education level is assumed comparable across all three censuses; “illiterate and semi-illiterate” in 1982 and 1990 is assumed to be equivalent to “no schooling” and “literacy class” in 2000 (and in 1995). For laborers, on which data start only at age 15, some of the classification issues over the years are not relevant.

### *Above secondary school level*

The number of categories at the tertiary level expanded from one in 1982 (and in the sample surveys of 1987 and 1995) to two in 1990 and three in 2000. In 1982, all tertiary education levels were grouped into one category; a distinction was only made between

---

<sup>11</sup> The explanations to the census, *Census 2000*, Vol. 3, p. 1895, are again not explicit but mention the “above-illiteracy level” criterion set by the State Council.

whether a person had already graduated or had not yet graduated (was still studying, had completed the course of study but did not graduate, or had discontinued the studies).<sup>12</sup> In 1990, a distinction was made between a regular university education (at all levels, from undergraduate to PhD studies) and a college-level associate degree education. In 2000, the first category was further divided into an undergraduate education vs. graduate education, but the data published on the latter category (graduate education) only covered those who had already graduated. This may bias comparisons of the 1982 and 1990 census data on all tertiary level education with 2000 data on all tertiary level education, in that the graduate education in 2000 only covers those who already graduated. A direct comparison between 1990 and 2000 is not possible because while the 1990 data are published according to 4 completion levels, including “graduated,” the 1990 data do not distinguish between undergraduate and graduate studies.

A second complication is that while accredited university-level distance learning, employee universities and evening universities, etc., are always acknowledged in the census, in 2000 this is only the case if the person has already graduated. Since the census data provide no details on distance learning etc. cases, no adjustments for this inconsistency can be made. (If year 2000 data on graduates at *each tertiary* education level were available, which are not, these data could be compared to the corresponding 1990 data.)

In the following, when linking the three censuses (relevant in the case of laborers), the 1990 university undergraduate/ graduate education is split into undergraduate vs. graduate level using the age-specific ratios of 2000, transferred back ten years, in as far as possible.<sup>13</sup> This ignores that the 2000 *ratio* may, in comparison to the 1990 ratio, have become inflated in the undergraduate category only, due to adult education, which is unlikely to occur to the same extent at the graduate level. However, adult education at the tertiary level may well be concentrated in the college category, in which case the 2000 ratio could be perfectly appropriate to split university/ graduate education into the two components.<sup>14</sup> It is also ignored that the 2000 census *graduate level* data only cover persons who already graduated, while the 1990 census *university education* data cover the total (graduated, currently enrolled, completed/discontinued studies, others).<sup>15</sup>

The 1982 university education is split into college, undergraduate level, and graduate level using the age-specific ratios of 1990, transferred back eight years, in as far as this is

---

<sup>12</sup> This distinction was also made in the 1987 sample survey but not in the 1995 sample survey.

<sup>13</sup> Age-specific *ratios* of undergraduate to graduate education are transferred back for the 1990 age groups 25-54 (2000 age groups 35-64). Because university education is only just happening at the young ages (15-24), with graduate studies *following* undergraduate studies, the 2000 ratios of the age groups 15-24 are directly applied to the 1990 age groups 15-24 (rather than transferring back ratios from those age 25-34 in 2000, who had a chance to complete both, undergraduate and graduate studies); in these age groups (15-24), in 2000, laborers with graduate education never exceed a number equivalent to 3% of laborers with undergraduate education. For those age 55-64 and 65+ in 1990, the 2000 ratio of those in the category 65+ is used (in 2000, the number of those age 65+ with a university graduate education is equivalent to 5.84% of the number of those with a university undergraduate education).

<sup>14</sup> The *Census 2000* (Vol. 2, pp. 874ff.) has limited adult education data by age and education level from the long-form sample, which shows adult education at the university-college level to be five times the size of adult education at the university-undergraduate level (in the regular education system, the two are of about equal size).

<sup>15</sup> It would have been possible to only consider persons who graduated with a university level education in 1990, but this biases the 1990 data downward in comparison to 2000, where for the undergraduate level only total data are available.

possible.<sup>16</sup> The same procedures are applied to the category university education in the 1987 and 1995 1% sample surveys.<sup>17</sup>

### 3. Education completion levels

The censuses ask for the highest education level achieved but achievement does not necessarily imply graduation. If a large share of those who state a given education level as their highest education level did not graduate at that level, using the official data exaggerates the actual graduation level.

The 1982 census data do not distinguish between completion levels except in the case of a university education, where a distinction is made between whether a person has already graduated or has not yet graduated (i.e., is still studying, has completed the course of study but did not graduate, or has discontinued the studies).

The 1990 census data, across all education categories (starting at primary school), report a total and then break this total down into four completion levels: graduated, currently attending, completed (including discontinued studies), and “others.”

The 2000 census data, for the census population, report total figures only (with no distinction according to completion levels); completion level data are not collected in the short form. For the sample of approximately 9.5% of the population who filled in the long form, completion data are collected at five levels for each education category from primary school up: graduated, currently attending, completed (but not graduated), discontinued, and others; these data are not published except for the primary and lower middle school. For the two sample surveys of 1987 and 1995, only total data are available, without distinction according to completion level (although such data were collected in 1995).

#### *Education levels in 1990 and 2000*

Figure 12 shows the education level of each age group in 1990, as share of the total population at this age (in percent). Obviously, most persons between age 6 and 12 in the 1990 census stated “primary school” as education level. Starting at age 15, those with an education level below primary school are considered “illiterate/ semi-illiterate,” and the different education levels from age 15 on, together with the category of illiterate and semi-illiterate persons, add up to 100%. For the generations below age 35 in 1990, Figure 12 shows a cycle of primary school, for half of all persons followed by lower middle school. Around one-quarter of those who attended lower middle school (the exact share is highly dependent on

---

<sup>16</sup> The exact procedure is as in the construction of the relative shares of undergraduate and graduate education in 1990, except that the 1982 age groups for who the cohort-specific eight-year transfer of the ratio is feasible and meaningful are now the age groups 33-64.

<sup>17</sup> In 1987, the age groups for who the cohort-specific three-year transfer of the ratio is feasible and meaningful are the age groups 25-61, in 1995, with a five-year transfer, the age groups 25-59. (In 1987, the age group 16 has 3 laborers at university level, while in 1990 the same age group has none; the 1990 ratios of the age group 17 are therefore used to break down the 1987 university category into its three sub-categories.)

In sum, in each of the four census/survey years 1990, 1982, 1987, and 1995, three different procedures are applied to three different age levels (the young, the old, and the three decades in between). The transitions between any two procedures in any of these years tends to be rather smooth, often as if a trend continues uninterrupted.

the age group) later attended upper middle school. The upper middle school reaches a narrow peak of just above 20% of the population age group 29/30. Those around age 29/30 in 1990 who cited upper middle school as their highest education level must have attended upper middle school in the early and mid-1970s. It is possible that the upper middle school then lasted only 2 years; it is also possible that completion levels became inflated during the Cultural Revolution.

In the older generations the share of illiterate/ semi-illiterate persons increases drastically to reach more than 70% in the case of those aged 65 and above. College- or university-level education, as well as special middle school education, play a very minor role throughout.

By 2000 (Figure 13) education levels at the younger ages had clearly risen. Lower middle school education started to reach 70% of the younger age groups, compared to 50% in 1990, and upper middle school education for those in their late teens 20%, up from 10% in 1990, with an extra 10% now in special middle school (Figure 13, based on the 1 November 2000 census date). About 5% of the younger age cohorts in 2000 gained a college-level education, and another 5% a regular undergraduate university education.

#### *Completion levels in 1990 and 2000*

A person stating a particular education level in the 1990 census could either, in 1990, be attending a school at that level, or could already have graduated, or could have completed the education but not yet have graduated (or could have discontinued the studies), or could fall into a residual category “others.” Figure 14 shows the shares of persons (in percent) at a given age who in the 1990 census stated a particular education level and were currently *attending* the institution that would provide the stated education. The figure reveals a natural sequence of school/ university attendance that peters out around age 26.

Figure 15, in contrast, shows the share of persons at a given age and stated education level who at the time of the census had already *graduated*, i.e. had the degree in their hands. The natural sequence of graduating first from primary school, then from lower middle school etc. again is clearly visible. But the figure also reveals that of those age 18 and over, roughly 30% of those who checked primary school education in the 1990 census as their highest education level actually did not graduate from primary school, a share that rises to about 40% for the oldest age groups. At the lower middle school level, that share is 10%, rising to 20% for the oldest age groups. Above the lower middle school level, the share is between 0 and 15%.

Figure 16 shows the shares of persons at a given age and stated education level who had completed the particular course of study but not graduated (or had discontinued their studies at this education level), and Figure 17 does so for the explicit residual “others.” The counterpart to the low graduation rate for primary school is completed (but not graduated) or discontinued studies. For lower and upper middle school, the discrepancy between the total vs. currently attending and graduated is also largely made up of completed studies. Only in the case of a university (college) education, for middle-aged persons, does the completion level “others” clearly rise above the “completed” level. Still, the residual others is a small

share throughout, at between 0 and 4% of the total, rising higher only for the age groups above 50.<sup>18</sup>

For the 2000 census, data on completion levels is only available for the sample using the long form, and only for primary and lower middle school. These 2000 data reveal much higher graduation rates, at around 85-90% for primary school and around 95% for the lower middle school (Figure 18 and Figure 19, based on the 1 November 2000 census date, and excluding military personnel).<sup>19</sup> Even older generations exhibit much higher graduation rates in 2000 than in 1990. Thus, for example, the group of 50-year olds in 1990 exhibited a graduation rate at the primary school level (for those who stated primary school as highest education level) of around 70%; by 2000, the group of 60-year olds exhibited a graduation rate, at the primary school level, of around 85%. At the lower middle school level, the graduation rate moved from around 90% for those in their 30s in 1990 to around 95% for those in their 40s in 2000. Unless the categorization practice changed between the two censuses, for which the census regulations appended to the data collections give no evidence, these improvements in graduation levels either suggest adult education or wrong self-reporting of education levels in one of the two censuses.<sup>20</sup>

With no graduation data for other education levels available in 2000, and no graduation data for any education level even collected in the 1982 census or in the 1% sample surveys of 1987 and 1995,<sup>21</sup> the question arises of what to do about those who stated a particular education level in the census but did not graduate at that level. Non-graduation from primary school is probably not a big issue affecting the quality of labor in that these persons can at least read and write brief notes, otherwise they would be considered illiterates or semi-illiterates in 1990; in 2000 they would enter the categories “no schooling” or “literacy class,” with possible a very few illiterates in the primary school category. In 1990, starting from the lower middle school level and clearly at the higher education levels, the graduation rates appear quite close to 100%

If one wanted to make corrections for non-graduation one could, for example, “rank” a person who in the 1990 census stated upper middle school as highest education level but had not yet graduated from upper middle school as above somebody who had graduated from lower middle school, but below somebody who had graduated from upper middle school, say half-way or three-quarters in between the two. Such a correction could be done for 1990, but for 2000 and 1982 an additional assumption would have to be made about some correspondence to the 1990 census, be it that persons are aged or “un-aged” by ten years,

---

<sup>18</sup> One complication about the completion level is that the census date of 1 July 1990 may mean that students who had completed a course of study and had taken their final exams may simply not yet have received their graduation certificate. It could thus be the case that one entire cohort of students entered the category “complete (or discontinued)” studies even though it would graduate a few weeks after the census date. However, this means that the completion share just before the school-level-specific graduation age should then show a clear peak, which is not the case, at any education level.

<sup>19</sup> The 2000 census was conducted with 1 November 2000 as census date, which avoids the completion vs. graduation problem (previous note), should it have been a problem in 1990.

<sup>20</sup> There appears to be no reason why the NBS’ practice of making no distinction between regular and adult education in the statistics by education level should not be appropriate. Only adult education following the official government requirements is accepted as equivalent to the corresponding regular education in the censuses. If a person age 50 can write, what difference should it make if this person learnt to write at age 10 or at age 30? At the higher education levels, such as the university-college level, adult education, occurring after work experience has been gained, may even yield better results in terms of contribution to the production process than a university-college level education obtained around age 20.

<sup>21</sup> For the university level, the 1987 survey collected and provides data in two categories, graduated vs. all else.

plus/minus the (largely unknown) extent of adult education, or that age-specific graduation rates are maintained over time. The overall impact of such a correction would appear to be below 5% of whatever measure is constructed, since approximately 90% of those stating a particular education level above the primary school level in 1990 had actually graduated (and more than 90% at the middle school level in 2000).

Given the assumptions needed to approximate graduation rates in 1982 and in 2000 (in 2000 above the lower-middle school level), the over time most reliable measure of education levels may simply be to take the totals reported and not try to adjust for graduation level. These totals slightly exaggerate the actual education levels in that they ignore that some persons did not graduate at the level at which they report themselves in the census to have been educated. In as far as the degree of exaggeration is constant over time, this may be near-irrelevant for the growth rate of any human capital measure constructed based on the totals. The few possible 1990-2000 comparisons suggest that the graduation levels are rising over time. The growth rates of human capital measures, thus, are likely to be slightly biased downward.<sup>22</sup>

### *Statistical Yearbook graduation data*

In order to project education levels into the future, it would be good if the most recent enrollment and graduation numbers provided in the *Statistical Yearbook* could be used. How do the year-specific graduation numbers in the *Statistical Yearbook* compare to the age-specific graduation numbers in the censuses? The 1990 census has graduation data at all education levels, and the 2000 census for primary and lower middle school. Table 4 has the comparison to the *Statistical Yearbook* data.

The first step is to identify at which age, for a particular education level, most students have (just) graduated.<sup>23</sup> For the primary school, in 1990, this is age 17 (or 16); after age 17, graduation rates do not change much any more. For the lower middle school it is age 19 (or 18), for the upper middle school age 21 (or 20), for the special middle school age 23 (or 22), and for the university age 25 (or 24).<sup>24</sup> Comparing the education-specific *Statistical Yearbook* graduation numbers in 1990 to the graduation numbers at these age groups (17, 19, 21, 23, 25) shows a good match for the lower middle school (the first is 1.22 times the latter), the upper middle school (1.03), and at the university level (1.01).

---

<sup>22</sup> In the data on laborers by education, the completion level “currently attending school” is not relevant, as these persons would then not be regarded as laborers (with minor complications explained in note 43).

<sup>23</sup> The problem for the comparison is that different persons graduate at a different age. If the number of students enrolled at a particular education level were identical across the years (time), one could simply take the maximum number of graduates at a particular education level from the census, regardless of exact age. But student cohorts are likely to change in size over the years. The attempt here then is to take the age at which there is a big jump in graduation rates compared to the age group one year younger and at the same time where the cumulative rate of graduation is high.

<sup>24</sup> Lower middle school and upper middle school each last three years, but the typical graduation age only goes up by two years in each case (from 17 in the case of primary school to 19 in the case of the lower middle school, to 21 in the case of the upper middle school). Presumably, those who go on to the next education level tend to graduate earlier rather than later, for example, at age 14 or 15 from primary school, at age 17 from lower middle school, and age 19 or 20 from upper middle school. (There could also be some ambiguity about some upper middle schools in earlier years possibly lasting for two years only).

The match is not good at the primary school level, presumably because a fair share of those who graduated from primary school in 1990 immediately enrolled in the lower middle school, and thus in the census did not enter the primary school category but the lower middle school category (as “attending school” rather than “graduated”). The match is mediocre at the special middle school level; the number of graduates might include a fair number of students from adult education at higher age. The 2000 census data has no further insights, except that the age at which a large share of students has graduated has moved back, at the primary school to 16 (or 15), and at the lower middle school to 18 (or 17).

The good match for upper middle school, university, and possibly lower middle school suggests that at these levels the recent *Statistical Yearbook* graduation data are likely to be good predictors of graduation reported in future population censuses and surveys. Below, *Statistical Yearbook* data on *enrollment*, reported in the same section and in some instances same tables as the graduation data, will be used in projections of future education levels; the censuses do not provide enrollment data for comparison, only graduation data.

#### 4. Education level of laborers vs. the population

The relevant labor variables in production function estimations are total laborers and their education levels. All three censuses, and both 1% sample surveys, contain data on the age-specific number of employed (the 1982 census only in 5-year age groups, with the 1982 official 1% sample in 1-year age groups). Data on education levels of laborers by age, however, are not always available. These data, laborers by age times education, are readily available only in the 2000 census publications for the long-form sample, and from the 1987 and 1995 sample surveys. For 1982 and 1990 they need to be constructed.

For 1982 and 1990 Margaret Maurer-Fazio has kindly provided education data on population and laborers at each age level 15 through 99 (four tables by age times education, for the population and for laborers, in 1982 and in 1990). The source is a 1% sample provided by the National Bureau of Statistics (NBS). The data I obtained from Margaret Maurer-Fazio is a 10% random sample of this NBS 1% sample in each of the two years, i.e., a 1‰ sample (665,783 persons in 1982, and 868,916 persons in 1990).<sup>25</sup>

For a simple check on the 1‰ sample data, Figure 20 through Figure 24 report the age- and education specific employment rates in the 1‰ (10% of 1%) sample of 1982, the 1% sample of 1987, the 1‰ sample of 1990, the 1% sample of 1995, and the long-form (approximately 9.5%) sample of 2000. The 1‰ sample data show quite some variation across age cohorts starting from about age 55 (the 1‰ sample figures are drawn through age 75, in contrast to the 1% sample and 2000 census sample figures where the data end with the category 65+). Otherwise, the patterns look similar across the censuses and surveys, and appear credible throughout.

The following sections examine the data in more detail and explain how the final data on laborers by age and education are derived.

---

<sup>25</sup> The rationale for the 10% random sample of the 1% sample is that the number of observations in the underlying 1% sample is so large that the demand on computing power to create the summary tables (education by age for population and, separately, laborers) is high; the problem is exacerbated by technical problems with the 1% samples.

1982

How reliable is the 1‰ sample of 1982 data? Figure 25 compares the complete 1982 population data (from the *Census 1982*) to the data in the 1‰ sample, with the latter, annual data summed to match the 5-year data in the *Census 1982*. For each age group, the percentage share of a particular education level according to the *Census 1982* is divided by the percentage share of the same education level in the 1‰ sample. If the 1‰ sample were a random sample of the population, the resulting ratio should be unity. As Figure 25 shows, the ratio is indeed very close to unity across all age groups (note the vertical range of the figure), except perhaps for the category university, presumably due to the relatively small number of observations in the sample. But even at the university level, the difference from unity is only a few percentage points.<sup>26</sup> Figure 26 conducts the same comparison by annual age, using the 1982 census official 1% sample (vs. the 1‰ sample data). The ratios for the categories illiterate/ semi-illiterate, primary school, and lower middle school are again very close to unity; these three categories cover 99% of the population. For the categories upper middle school and university the match is less perfect, especially at the higher ages; for these two categories the number of persons at each age dwindles to the lower hundreds in the census 1% sample and to the lower teens in the 1‰ sample.

Even if the 1982 1‰ sample were regarded as misrepresenting the shares of different education levels in a particular age group, it should do so equally for the population and the laborers, since the laborers are simply a sub-category of the 10% population sample of the original NBS 1% sample. Relationships between education measures of sample-laborers (in the 1‰ sample) vs. the sample-population (in the 1‰ sample) then should also hold for all (census) laborers vs. the total (census) population.

What is done in the following in order to obtain education data for the 1982 economy-wide (census) laborers, at each age level, is to take the 1‰ sample age- and education-specific characteristics of laborers relative to the population at this age (Table 5), and to apply them to the census population of this age. The sample characteristics consist of the number of sample laborers at a particular education level and age relative to the number of sample persons (population) at this age. Multiplying by the number of census persons (population) at this particular age then yields the total number of (implicit: census) laborers at the particular age and education level. That is the value needed for the calculation of measures of human capital. The remainder of this section discusses an alternative procedure and further explores the justifications for using the procedure chosen here.

A more detailed procedure which makes use of all sample data would have been to use the sample age- and education-specific ratios of laborers to the population but census population data by (annual) age and education are lacking, except for the 1% sample of the 1982 census. Census population data by 5-year age groups times education, however, are available. An alternative procedure then is to regroup the 1‰ sample data into 5-year data starting from age 15-19 through 55-59 and “60 and above,” to divide the resulting labor data (age by education) by the resulting population data (age by education), and to apply these age- and education-specific employment rates to the (5-year) age- and education-specific

---

<sup>26</sup> In 1982, the number of observations in the 1‰ sample tends to be several thousand at each combination of education and 5-year age group. Exceptions are the youngest and oldest age groups, and university level education throughout (which is in the hundreds except for the 40-44 age group, where it just exceeds one thousand); the upper middle school falls below the thousand threshold starting at age 50-54.

population data in the 1982 census. Comparing the total number of illiterate/ semi-illiterate laborers derived following this alternative procedure to the number obtained following the procedure explained in the previous paragraph, the 5-year alternative procedure yields a number that is 3.32% larger; for primary school, lower middle school, upper middle school, and university, the 5-year alternative procedure yields values that are 1.14%, 0.20%, 0.40%, and 0.09% larger. The size of these differences appears negligible. Given the advantage of having annual age-specific 1982 data for comparisons with other censuses, the first procedure is adopted in the following.<sup>27</sup>

Using the share of 1‰ *sample* laborers at a particular age and education level in the 1‰ sample population at that age, combined with the age-specific census population, to obtain economy-wide age- and education specific labor data assumes that the age- and education-specific share of laborers in the population at that age is the same in the sample as in the census. This is plausible because (i) the age-specific employment rates are the same in the sample as in the census, and (ii) the age-specific education distribution of laborers is the same in the sample as in the census.

Ad (i), this can be shown to hold for 5-year age groups in the census as well as for annual age groups in the official 1% sample. Figure 27 shows the 1‰ age-specific ratios of laborers to population, in 5-year intervals and by annual age group, as well as the census 5-year age-specific ratios and the census 1% sample annual age-specific ratios. The 1‰ sample and official census values are so close that they cannot be distinguished.

Ad (ii), laborers are a subsample of the population. If the education structure of the sample population were the same as the education structure of the census population, one could reasonably expect that the similitude also extends to the laborers. Figure 25, for 5-year age groups, showed that the population education structure of census and 1‰ sample are near-identical. The ratio, at a particular age, of the education-specific share in the population in the sample to that in the census is usually 1.00 or very close to 1.00, ranging at most between 0.98 and 1.03, except in the university case, where the range is 0.92-1.07.

With near-identical age-specific employment rates in census and sample, and near-identical age-specific population education distributions (and therefore in all likelihood also laborer education distributions), the only scope for mistakes in applying sample characteristics to the census population data arises from the possibility that employment rates, at each age, differ between sample and population depending on education level. I.e., a reason would have to be found why a specific age-education group in the sample should show different work tendencies than the same age-education group in the census, when the age category in total shows exactly the same work tendencies in the sample as in the census, and when the education distribution at each age (at least in the population data, and thereby presumably also in the labor data) is the same in the sample and in the census. Unless such a reason can be found and it is plausible, the procedure used here is justified.

---

<sup>27</sup> Yet another way to derive age by education labor data would have been to simply, at each age, take the sample share of laborers at a specific education level in all laborers and to apply this share to the official 1% sample laborer numbers, times 100. But this ignores that the distribution of laborers across age in the official 1% sample could be biased (the considerations at the beginning of this manuscript suggest not), or that the sample data I obtained contain some bias in the education distribution at each age (which turns out to not be the case in the 5-year groups, and to be unlikely given the population-based sample vs. census comparison of age-specific education shares).

For example, if the sample over-sampled highly educated people, this creates no bias in the calculations. If the sample over-sampled elderly people, this also creates no bias in the calculations. What is needed to create a bias is a reason for having *employment rates* at some combination of age and education systematically differ between sample and census. This could be the case, for example, if the sample focused on unemployed persons at certain combinations of age and education, and happened to compensate exactly with a focus on employed persons at the same age but different education level. (The comparison of employment rates between sample and census by *age* groups (item (i) above) revealed no bias. By education, the data are not available for a comparison.)

The procedure used here is further justified through the near-identical aggregate results obtained using the alternative procedure, which simply assumes that whatever bias the sample may have, it equally affects labor and population data in the sample.

The constructed data on age- and education-specific laborers for 1982 are capped at age 65 with a final category of age “65 and above.” This is done to match the laborer data in the other censuses.<sup>28</sup> In the 1‰ sample in 1982, the laborers age 65 and above accounted for only 1.53% of all laborers. (In 1990, 1.88% of all laborers were age 65 and above; in 2000, 3.25% of all laborers in the official long-form sample were age 65 and above.)

## 1990

For 1990, two approaches to constructing economy-wide laborers by age times education are equally plausible. One is to apply the age- and education-specific employment rates of the sample to the (age- and education specific) census population. The other is to, at each age, apply the education shares of laborers in the sample to the number of laborers in the census. The first approach has the advantage that if the sample were in any way biased, the sample laborers and sample population are likely to be equally biased, and the bias should thus cancel out; it has the disadvantage that the resulting age-specific (as well as total) number of laborers does not perfectly match the number in the census. The second approach has the advantage that the census numbers on laborers can be used (at each age, and thus also in total); it has the disadvantage that if the education structure of laborers at a given age in the sample were biased, this bias is not corrected in the derivation of economy-wide laborers by age times education.

The census also provides data on total laborers by education (independent of age). The results of the two approaches, the number of laborers by age times education, can be aggregated across age to yield two sets of data on total laborers by education. Dividing the share of laborers in each education category in the census by the share of laborers in each education category following the first (second) approach yields a ratio of 0.9999 (1.0271) for the education category illiterate and semi-illiterate persons, a ratio of 1.0143 (1.0020) for primary school, 1.0108 (0.9892) for lower middle school, 0.9993 (0.9855) for upper middle school, 0.9961 (0.9874) for special middle school, 1.0244 (1.0093) for university-college, and 1.0352 (0.9689) for university undergraduate and graduate studies. The first approach yields values closer to unity than in the second approach in three cases, and values further away also in three cases; in other words, the two approaches are equivalent in terms of aggregate

---

<sup>28</sup> In the 1‰ sample, the data go through age 99 in annual age levels. A residual category captures 171 persons in the population, of which 71 are working (laborers); these are ignored in the following. The age group “65+” in the 1‰ sample is taken to consist of those between the ages of 65 and 99 (both inclusive).

results.<sup>29</sup> The second approach, in comparison to the first approach, has the advantage that it uses the actual census laborer data at each age. Overall, therefore, the second approach appears more desirable and is used here.

The risk involved in using the second approach is that the sample age-specific education distribution of laborers need not perfectly match the (unknown) age-specific education distribution of laborers in the census. In the aggregate, as total across all age groups, the education distribution of laborers in census and sample match well, as seen above.

The risk involved in the first approach, not used here, is that the age- and education-specific employment rates of the sample need not be the same for the census. Figure 28 shows that the age-specific employment rates are near-identical (although perhaps not as good as in 1982, Figure 27, but then in 1982 the census data are 1% census sample data). The education-specific employment rates are also rather similar, with a ratio of census employment rates to sample employment rates across the seven education categories of 1.0104, 1.0182, 1.0055, 0.9764, 0.9531, 1.0298, and 1.0536. In the population, the age-specific education characteristics are also rather similar. Figure 29 compares the *population* education shares in a particular age group from the census to the population education shares in the same age group from the 1% sample. Apart from one sharp outlier for the university category at age 43, and some inconsistencies at the very youngest ages, the sample population data match the census data well; since the sample laborer data are a subset of the population data, the same should hold for laborers across age and education.<sup>30</sup> (The outliers / inconsistencies are based on a very small number of persons, at the very youngest ages on single-digit numbers of persons.)

The fact that age-specific laborer data are available in the census and can be used in the second approach tilts the balance towards the second approach. The results again go from age 15 through age “65 and above,” with age-specific data on laborers from the census only available for these age groups.<sup>31</sup> Table 6 reports the 1% sample laborer data as well as the employment rates for use in the first approach.

## 2000

The 2000 census contains data on laborers by age times education in the long-form sample. Data on the population in the long-form sample are also available, however, only by age *or* education, not by age times education. In total, the population in the long-form sample equals approximately 9.5% of the census population.

Economy-wide data on laborers by age times education are obtained by augmenting the laborer data in the long-form sample by the age-specific ratio of the census population to the long-form sample population.

---

<sup>29</sup> In terms of the sum of squared differences from unity, the value for the first approach is 95.33% of the value of the second approach, i.e., the first approach has a slightly lower value.

<sup>30</sup> The outlier at age 43 is due to a different transition in census and 1% sample from a low to a higher percentage plateau. The university shares in the census at ages 42, 43, and 44 are 0.41, 0.71, and 1.16; in the 1% sample they are 0.42, 0.41, and 1.10.

<sup>31</sup> The sample data again go through age 99 (with 7 persons age 99). There is no residual category.

## 1987 and 1995

In the absence of any economy-wide data, the characteristics of the 1% sample survey data cannot be directly combined with other data to obtain implicit economy-wide/nationwide laborer data. Furthermore, as seen above, the 1987 and 1995 1% sample over/under-sample certain age groups.

The procedure adopted here is to obtain the number of economy-wide laborers at each age via interpolation from the censuses (explained further below). At each age, the number of total laborers is then allocated to the individual education categories using the sample distributions (same procedure as in 1990). An alternative would have been to make use of the sample characteristics of laborers vs. population, and to apply these to nationwide population data, but the nationwide population data would also have to be interpolated from the censuses.

One would expect the education shares of a specific age group to gradually change over time, as the cohort ages. Figure 30, for the share of illiterate and semi-illiterate laborers (in all laborers of a given age), shows this succession. The horizontal axis shows the age in 2000, independent of the date of the census or survey. Taking the age group 45 in 2000 as example, of all laborers age 27 in 1982 (45 in 2000), 21.79% were illiterate or semi-illiterate. By 1987, of all laborers age 32 in 1987 (27 in 1982, 45 in 2000), 20.00% were illiterate or semi-illiterate (which suggests that illiterate and semi-illiterate persons at this particular age are gradually withdrawing from the labor force or undergoing adult education). By 1990, of all laborers age 35 in 1990, 14.13% were illiterate; by 1995, of all laborers age 40 in 1995, 12.90% were illiterate; and by 2000, of all laborers age 45 in 2000, 8.52% were illiterate.

Figure 31 through For notes and sources see Figure 30.

Figure 34 proceed similarly with the other education levels (primary school, lower middle school, upper middle school, and university). The succession is rather smooth throughout. The succession is not always perfectly smooth, but usually works acceptably well to justify the use of the 1987 and 1995 sample survey education structure of laborers.<sup>32</sup>

## 5. Total Employment

### *Census employment definitions*

Across the three censuses and two 1% sample surveys, the category laborers (*zai ye renkou*) covers persons age 15 and above. In 1982 and 1990, laborers comprise two groups of person, (i) those with a fixed occupation (*gudingxing zhiye*), except those who are on leave for further, tertiary level education,<sup>33</sup> and (ii) those without a fixed occupation. Those without a fixed occupation are defined as persons who on 30 June 1982/1990 held a temporary job and had worked on a cumulative 16 days or more in June 1982/1990. (*Census 1982*, p. 607; *Census 1990*, Vol. 4, p. 515 ) The 1987 1% sample survey data do not come with detailed definitions but presumably the category of laborers (in 1987 also *zai ye renkou*) is defined as in the 1982 and 1990 censuses.

---

<sup>32</sup> For example, the 1995 shares of university educated laborers in the age groups 55 and higher (in 2000) seem to be too high, but at these age levels the numbers of laborers are relatively small.

<sup>33</sup> Retirees who still worked in the month before the census and drew a salary were to be counted as laborers.

In 2000, laborers comprise those persons who worked for monetary or non-monetary compensation (including profit and family gain) for at least one hour in the week preceding the census (25 through 31 October 2000); the same definition holds for the 1995 1% sample survey.<sup>34</sup> In 2000, the labor question is only asked in the long form of the census, and, thus, was only answered by about 9.5% of the population. (*Census 2000*, Vol. 3, p. 1899) In the published data, laborers starting in 2000 (not yet in 1995) are no longer explicitly labeled “laborers” (*zai ye renkou*); they are now the “population” in the various economic sectors, or “employed persons” (*jiuye renkou*).

The new definition of employment since 1995/2000 corresponds to international practice.<sup>35</sup> Unfortunately, the switch in definition is likely to create time inconsistencies in the Chinese employment series.

### *Statistical Yearbook employment data*

An alternative source of economy-wide labor data is the *Statistical Yearbook*. The *Statistical Yearbook* traditionally reports employment according to sector, labeled “social laborers” (*shehui laodongzhe*) prior to 1993, and labeled “employed” since 1993, with the Chinese term *congye renyuan* in 1993-00, and the term *jiuye renyuan* since 2001. The switch in the Chinese term in 2001 also comes with a slight redefinition of who is included: the staff and worker category in 2001 turns into an “on-post” (*zaigang*) staff and worker category, i.e., now only counts those staff and workers actually working rather than those listed as staff and workers of a work unit. In practice, the re-definition appears to have been implemented already starting 1998, under the old definition and previous Chinese term of *congye renyuan*.<sup>36</sup>

The sectoral data in the *Statistical Yearbook* presumably reflect data collected through the traditional reporting system. The introductory passage to the employment section in the *Statistical Yearbook* (for example, 2004, p. 117) mentions that the employment data are based on the “labor statistics report form system” covering all units with independent accounting system, the “population change sample surveys,” and the “township/village social and

---

<sup>34</sup> In the 2000 census, the work is further qualified as (undefined) “social labor” (*shehui laodongli*). Compensation can be in cash or in kind. The *Survey 1995*, pp. 646ff., provides instructions on how to handle a list of special cases.

<sup>35</sup> See the “Resolution concerning statistics of the economically active population, employment, unemployment and underemployment, adopted by the Thirteenth International Conference of Labour Statisticians (October 1982),” at <http://www.ilo.org/public/english/bureau/stat/download/res/ecacpop.pdf>, item 9, as kindly pointed out to me by Sara Elder of the ILO Key Indicators of the Labour Market Team on 26 April 2005. Item 9 distinguishes between paid employment (for wage or salary, in cash or in kind) and self-employment (for profit or family gain, in cash or in kind), and requires “some work” within a specified brief period of either one week or one day, for operational purposes defined as work for at least one hour. China’s census 2000 (and 1995 1% sample) definition of labor matches the definition in this resolution. The U.S. uses the same criterion of one hour of work for compensation in the previous week. For the U.S. and China’s definition of employment see, for example, the International Labor Organization website (<http://laborsta.ilo.org/applv8/data/ssm3/e/US.html> vs. .../CN.html).

<sup>36</sup> The term “employees” by definition covers: staff and workers (in the definition, since 2000, on-post staff and workers only), re-employed retirees, private entrepreneurs, the self-employed (*getihu*), employees in private and individual-owned (*getihu*) enterprises, employees of township (and village) enterprises, village laborers, and other laborers (including teachers in unofficial schools, religious workers, and military personnel). See, for example, *Statistical Yearbook 2001*, p. 152, 2002, p. 170. On the re-definition in the data since 1998 also see Carsten Holz and Yi-min Lin (2001, p. 52), or a brief note in the *Statistical Yearbook 1999*, p. 140.

economic surveys,” without further specifying which reported data were collected using which method; the sectoral data are likely to rely primarily, if not completely, on the report form system.

Prior to 1990, the employment sum across sectors equals the reported total employment in the same table in the *Statistical Yearbook*; for the years since 1990, in the *Statistical Yearbook* editions starting 1997, it does not. In the *Statistical Yearbook 1997* through 2001, the sum across sectors for the years 1990 through the most recent year is adjusted upward to obtain the total reported in the same table (creating an implicit residual). This adjustment is done in accordance with the 1990 population census data (and presumably also in accordance with the annual population change sample survey data).<sup>37</sup> The *Statistical Yearbook 2002* revises all of these total employment data since 1990 (the previously published sectoral data are not revised); the introductory passages to the employment section in the *Statistical Yearbook 2003* and 2004 explicitly say that the employment data since 1990 are “adjusted according to the 5<sup>th</sup> population census [of 2000]” (*Statistical Yearbook 2004*, p. 117). (See Table 7.)

This means that the *Statistical Yearbook* series contains two total employment series since 1990. One is the sum across sectors, explicitly reported through, at the latest, 1995, in a tabulation that ends with the 2002 data; the other is the adjusted total. The adjusted total comes in two variations, as published for the years since 1990 in the *Statistical Yearbook 1997* through 2001, and as published in the *Statistical Yearbook* since then. Figure 35 shows the differences. The adjusted total value for 1990 and the years thereafter is significantly above the sum across sectors. The sum across sectors, furthermore, shows a discontinuity in 1998; this presumably reflects the de facto redefinition in 1998 (rather than, as the formal definitions suggest, in 2001) to only include those laborers on their post, i.e., to exclude those laborers laid off but still associated with a work unit. All *Statistical Yearbook* employment values are end-year values.<sup>38</sup>

The most recent, adjusted total employment value in the *Statistical Yearbook* for end-1990 almost coincides with the mid-1990 census value (647.49m vs. 647.24, or 650.44m including military), and for end-1995 almost coincides with the 1 Nov. 1995 survey value (680.65m vs. 680.03m, the latter presumably excluding military).<sup>39</sup> In 2000, the *Statistical Yearbook* value is larger than the census value (a for 1 Nov. linearly interpolated *Statistical Yearbook* value of 719.70m vs. a census value of 703.83m, or of 706.33 including military). It is not clear what the reason for the discrepancies in the different years is.<sup>40</sup>

---

<sup>37</sup> The *Labor Yearbook 1999*, p. 8, explicitly states—in a note to a table carrying the same data as in the *Statistical Yearbook*—that the data are calculated according to the 1990 population census.

<sup>38</sup> One further complication is the age aspect in the definition of laborers. Liu Chengxiang et al., 2000, p. 70, in a book explaining the variables used in the *Statistical Yearbook*, define employment as the *Statistical Yearbook* does (see note 36), but then follow this definition with a statement on what this “concretely” (*juti*) means in population surveys (not further specified): laborers include those age 16 and above with the capability to work, who have worked at least one hour in the survey week for compensation or profit, or who are currently not working because they are studying or on vacation but who have a work unit or work place. In contrast, the age limit in the censuses and in the 1987 and 1995 surveys is 15. Presumably, the employment data in the *Statistical Yearbook* also cover all laborers age 15 and above, and the annual 0.1% survey data that Liu Chengxiang et al. presumably refer to, if used, are adjusted.

<sup>39</sup> The *Statistical Yearbook* employment data comprise military personnel. (Liu Chengxiang et al., 2000, p. 70)

<sup>40</sup> In the years prior to 1998, the *Statistical Yearbook* employment data possible include a small number of persons who are officially part of an urban work unit but are no longer “on their post;” in the years since 1998, those not on their post are not included in the employment data (compare *Statistical Yearbook* data with those in individual volumes of the *Labor Yearbook*). The *Statistical Yearbook* economy-wide employment data used

## Military

All three censuses list military personnel separately (4,238,210 in 1982, 3,199,100 in 1990, and 2,498,600 in 2000). Military personnel refers to military employment only; family members of military personnel, for example, are included in the regular census data. Military personnel are not part of any of the published census tables such as the published population or laborer data.<sup>41</sup>

Total numbers of military personnel are available in all three censuses publications with a breakdown by age. In 1982, the breakdown is into “below age 20” and then five-year age groups starting with age 20. In 1990 and 2000, the breakdown is by annual age except for the first and last groups, which are “below 18,” and “60 and above.” Education data are also available, but only in the aggregate, for the total of all military personnel, in each of the three censuses.<sup>42</sup> Since the military is supposedly included in GDP (as part of government), military personnel needs to be included in labor.

The 1987 and 1995 sample survey data do not come with separate military data. The instructions on the 1995 sample survey state explicitly that the military personnel data are to be collected by the military and to be passed on to the national 1% sample survey joint meeting (*lianxi huiyi*). They are to be publicized “lumped into” the national data (*yu quanguo shuju yibing gongbu*). (*Survey 1995*, p. 633) Presumably the military personnel data are complete data rather than 1% sample survey data. My interpretation is that these complete data are included in the official estimate of the nationwide total population which is based on the 1% sample survey, and are otherwise not published. It does not seem plausible that the

---

here match the total employment data in a table listing laborers by ownership form. For the urban sector, laborers in state-owned units, collective-owned units, and “other” units (joint cooperatives, joint enterprises, limited liability companies, shareholding companies, overseas Chinese and foreign-funded enterprises, and “others,” but not including private enterprises and the self-employed) can be divided into formal laborers (*zhigong*), which typically covers the vast majority (more than 95%), and “other” laborers. The formal laborers, in turn, can be divided into those on their post and those not on their post. The issue of not being on one’s post is likely to have become relevant only in the early to mid-1990s, thus affecting no more than the years since perhaps 1994, up through 1997; a correction is not needed if those not on their post in these years actually became self-employed (and were not counted in the category self-employed, perhaps because of their special status). For 1996, data on those laid off from their post are available; their number is equivalent to 1.29% of total employment in China, but having been laid off does not mean that these persons had no employment (or had not retired), so that the true share of those not on their post and not employed in another category is possibly significantly smaller. No correction is attempted here to take into account those possibly counted as employed but de facto unemployed. (I learnt about the issue of laborers who are not on their post prior to 1998 being included in the total from Tsui Kai-yuen.)

The fact that the *Statistical Yearbook* total in 2000 well exceeds the census value also suggests that the *Statistical Yearbook* value is unlikely to cover only laborers at age 16 and above, rather than 15 and above, as the census does (see discussion in note 38).

<sup>41</sup> The *Census 2000*, Vol. 3, p. 1, makes this explicit: “this material [these 3 books with data on the 2000 census] does not include the 2.50m persons in the Chinese People’s Liberation Army.” The third volume, at the very end, in an appendix, has two pages with data on military personnel. The previous two censuses have similar brief tables at the end, but do not come with an explicit statement that military personnel are not included in the other tables. Given the same practice in the 1982 and 1990 censuses of listing the military personnel separately in an appendix at the end, as in the 2000 census, I assume the published population and laborer data in the 1982 and 1990 censuses also exclude military personnel, as in 2000.

<sup>42</sup> Military personnel are vastly higher educated than non-military laborers, as a comparison of military personnel qualifications and that of an artificial aggregation of laborers with *same age distribution* (as the military) shows.

NBS takes a 1% sample of the complete military personnel data and then adds these military personnel to the laborer and population tables in which it reports the 1% sample survey data. First, it does not mix military personnel data with non-military data in the published data from the censuses. Second, in the published census data on military personnel, education data are only reported for the total of all military personnel, not by age; it appears far-fetched for the NBS to have military personnel data in 1987 and 1995 by age times education, to create a 1% sample of these data, and to then add the sample into the population and laborer tables by age and education.

The number of military personnel and the education level of military personnel in the years other than the census years can be obtained by assuming geometric growth between the censuses, both in the number of military personnel, and in the relative education shares. Military personnel values can thus be obtained separately from non-military laborers.

### *Final employment data*

Which data provide the best approximation of the actual *total* number of laborers? The economic sector data are likely to be too restrictive if they are based on report forms. The fact that the time series on sectoral data ends in 2002 may also signal that it has severe shortcomings. For the years since 1991, then, the best guess is probably the revised total employment data in the *Statistical Yearbook*. Unfortunately, these data do not come with a clear definition, and it can only be hoped that they are defined by the NBS, internally, in consistent fashion. It is not clear how the revised series handles the changes in the definition of laborers between the 1990 and 2000 censuses, including the switch in definition to a one-hour work minimum in the week prior to the 1995 1% sample survey and the 2000 census.

The total employment data can be turned into midyear values by taking the arithmetic mean of previous year end-year and current year end-year values, starting with the 1991 value (using 1990 and 1991 employment data from the *Statistical Yearbook*).

Prior to 1991, the cornerstones are the 1982 and 1990 census values; these values are identically defined, although slightly differently from the 2000 census. In 1990, the end-year 1990 total employment value in the *Statistical Yearbook* of 647.49m compares to a mid-year census value, including military, of 650.44m. Given that employment in China is growing over time, the *Statistical Yearbook* value should be slightly higher, or the census value slightly lower, for a smooth transition. Lacking a plausible method to bridge the change in definition, the 1990 (small) statistical break is accepted in the following. One could have adjusted the 1990 census value along with the 1983-1989 values (as explained below), to link up better to the years after 1990, but then the 1990 aggregate labor data no longer match with the detailed education and age data in the 1990 census. Since these detailed census data are a cornerstone in deriving education values for other years, it is important to retain the 1990 value of total laborers from the census. A choice could also have been made to revise the data of 1991 or of several years after 1990, but this would only distribute the problem to other years.

For the years 1983-89, the annual employment values are obtained through interpolation, each year gaining the same share of the gap between 1982 and 1990 as the sum-across-sector employment does. This procedure appears acceptable in that the 1982 census number of laborers is 1.1433 times the sum-across-sectors value, and in 1990 an almost identical 1.1464

times. The 1987 1% sample survey value, on the other hand, augmented by a multiplication factor to take into account that the sample population is supposedly equal to 0.999% of the actual population, and augmented by a linearly interpolated military personnel value, would have been equal to 1.1154 times the sum-across-sector value (2.62% lower than the interpolated value). For the years 1978-81 the backward growth rates in the sum-across-sector employment are applied to the 1982 census value.

The only known reason for the statistical break in 1990 is the change in the coverage of laborers. While one might expect that the switch to counting everybody who worked for compensation for one hour or more in the week prior to the census/survey date to increase the employment number, the close match between the 1990 midyear census value and the 1990 end-year total employment value at a time of rising employment suggests that the statistical break in fact implies a reduction in the number of laborers counted. No matter what exact (unknown) change in definition happened in 1990, the one-hour new limit may have little practical impact. In the 1995 1% sample survey, only 0.0308% of all those counted as laborers worked 1-8 hours in the week before the survey, a very small share. Of all laborers, 0.24% worked 9-16 hours, 1.16% 17-24 hours, 2.65% 25-32 hours, 53.06% 33-40 hours, 8.56% 41-48 hours, and 34.31% more than 48 hours. Using mid-interval values and 48 hours for the last category, yields an average work week of 40.67 hours.<sup>43</sup> Such data are not available for other years.<sup>44</sup>

## 6. Derivation of the year-, age-, and education-specific number of laborers 1978-03

To derive the year-specific, age-specific, and education-specific number of laborers requires two steps. The first step is to, based on the final employment data, create age-specific labor data including military for all years. This is done first for the census years, which are subsequently linked to derive age-specific labor data (including military) for all other year by aging certain age cohorts and interpolating other age cohorts.

The second step is to break the yearly age-specific number of laborers down by education level. This requires first that military personnel is taken out of the year-specific, age-specific labor data because data on military personnel by age times education are not available, only by education *or* by age.<sup>45</sup> For non-census years, the number of military personnel and their

---

<sup>43</sup> For the absolute data see *Population Statistical Yearbook 1999*, pp. 84f. The *Survey 1995*, pp. 646f., provides details on who is to be counted as laborer; it also lists a default work time of 40 hours in a variety of special circumstances: sickness/ machinery breakdown that temporarily prevents work, temporarily not working during a job switch (left previous job, not yet assumed new job), seasonal labor not working at the time of the survey, and temporary leave from work for studies (less than one semester). The average reported hours, thus, appear biased upward. On the other hand, those persons who are on leave from work to study for more than one semester, by definition, work 0 hours.

<sup>44</sup> For later years the *Labor Yearbook 2004*, p. 110, for the urban sector only, has some work hour data. In as far as average weekly work hours in agriculture in 1995 were 41.2 hours, not much above the economy-wide average of 40.7 hours, a limited comparison might be possible. Thus, the *Labor Yearbook 2004* reports average weekly work hours in the urban sector for Oct. 2001 of 44.9 hours, for Nov. 2002 of 45.2 hours, and for Nov. 03 of 45.4 hours. This is in stark contrast to the reduction in weekly work hours of staff and workers to 40 through government regulation in the mid-1990s. But not all urban laborers are "staff and workers," and the default assumption of 40 hours, at least practiced in the 1995 survey, for a number of special categories may be too high.

<sup>45</sup> It may seem that total military personnel could simply have been subtracted from final employment data in the first step to facilitate the calculations. However, the age distribution of military personnel is different from that of other laborers, which implies that the interpolation of laborers in the first step, if not taking into consideration military personnel, is biased due to the omitted entry and exit of military personnel in/from the

distribution across age groups needs to be interpolated (using the average annual geometric growth rates between censuses). After subtracting, in each year, age-specific military personnel from the age-specific numbers of laborers, the remaining numbers of laborers are broken down, in each year at each age, into the different education categories. In order to obtain the age-specific numbers of laborers in 1982, 1987, 1990, 1995 and 2000, the census and 1% sample survey age-specific proportions are applied. For other years, the age-specific proportions are interpolated from the proportions of these five years, and then applied to the year- and age-specific numbers of laborers of the other years. The result is the year-specific, age-specific, and education-specific number of laborers (excluding military).

In addition, when interested in economy-wide labor qualification, a further step is to aggregate the year-, age-, and education-specific numbers of laborers across age, and to add in the military personnel with their aggregate education structure (interpolated for non-census years).

### *Step 1: Derivation of age-specific number of laborers (including military personnel)*

Age-specific numbers of laborers are available from the 1982, 1990, and 2000 censuses. Age-specific numbers of laborers are also available from the 1987 and 1995 1% sample surveys, but these surveys, as observed above, may over-/under-sample certain age groups; these data, therefore are not used in the derivation of the *age-specific* numbers of laborers. All data from the 2000 census (and when used below, from the 1995 1% sample survey) are adjusted to 1 July 2000, so that all data in the following are midyear data.

In order to obtain the economy-wide numbers of census laborers, military personnel need to be added at each age. The 1990 and 2000 census provide annual age-specific numbers of military personnel.<sup>46</sup> The age group “below 18,” in 1990 (2000) equivalent to approximately 10% (5%) of the age group 18, is assumed to be age 17 in total; the age group “60 and above,” in 1990 (2000) accounting for 1.98% (0.08%) of total military personnel, is split evenly among ages 60 through 64. The 1982 census provides age-specific number of military personnel only in 5-year age groups. The age group “below 20” is split into the age groups 17, 18, and 19 following the proportions in the 1990 census, the age groups 20-24, 25-29 etc. are each split among the five individual (one-year) age groups they cover according to the 1990 proportions of the five individual age groups.

The final employment data (including military), for all years 1978 through 2003, were derived above. In 1990, final employment equals total laborers in the census (including military). In 1982 and 2000 (when only sample data on laborers are available from the censuses and when economy-wide *age-specific* laborer data are obtained through the procedures explained above), the sum of constructed economy-wide laborers across *age groups* is compared to the final number of laborers, and each age group is adjusted proportionally for the two to match. (In 1982, the sum of constructed age-specific laborer data including military personnel is 1.55% larger than the final number of laborers, in 2000 it is 0.40% smaller.)

---

number of laborers. Military personnel is also vastly better educated than an artificial laborer aggregate with the same age characteristics.

<sup>46</sup> The military personnel data in the 2000 census are of 1 November, whereas those in the 1982 and 1990 censuses are of 1 July. It is assumed the total number of the military personnel on 1 November 2000 is the same as the (unknown) number on 1 July, and that the age and education structure are also the same.

In order to derive age-specific laborer data for all other years, the just derived 1982, 1990, and 2000 census data (including military personnel) are linked. How they are linked depends on the age group. The individual age groups 19-56 in 1982 are aged by cohort through 1983, 1984, ... 1989. This is done by applying the average annual geometric growth rate implicit in a comparison of the age group X in 1982 with the age group X+8 in 1990; the result is the age-specific number of laborers in 1983 for those in age groups 20-57, in 1984 for those in the age groups 21-58, etc. The same procedure is applied to the individual age groups 19-54 in 1990 (now with X+10 in 2000) to obtain the age-specific number of laborers in 1991 through 1999; this procedure is also continued for 2001-03, using the 2000/1990 average annual geometric cohort-specific growth rate. For 1978-1981, the 1990/1982 average annual geometric cohort-specific growth rates are applied backwards.

For the young and the old, special aging procedures are needed. For the age groups 15-19, the cohort-specific aging procedure is not appropriate because of drastically changing employment rates between age 15 and 19 within one year—rather than gradually, for example between age 15 in 1982 and age 23 in 1990. Above age 19, employment rates do not change much. The age cohort 15 turns into age 25 ten years later, therefore data on the originally even younger age groups, for example those who turn age 15 through 23 after nine years (and who are age 6 through 14 in the original year), are also needed. This is best done backwards.

Starting with the census 1990, the numbers of laborers in the age groups 15 through 25 in 1989 are obtained by adjusting for different age- and year-specific employment rates and different age- and year-specific death rates. The death rate is defined as the number of deaths at a particular age in the 12 months before the census, divided by the number of persons (population) at the same age at the time of the census (the survivors).<sup>47</sup> Age-specific employment rates and death rates can be calculated from the three censuses for 1982, 1990, and 2000.<sup>48</sup> For all other years, employment and death rates are obtained by using geometric growth rates between age group X in 1982 and age group X in 1990, and similarly between age group X in 1990 and age group X in 2000 (i.e., these rates are not cohort-specific but age-specific); 2001-3 rates are assumed to change annually as those between 1990 and 2000, and 1978-81 rates as those between 1982 and 1990. Obviously, employment rates and death rates at a given age do not change much over time (but they do change).

For example, the number of laborers age 16 in 1989 is obtained by multiplying the number of laborers age 17 in 1990 by the employment rate of 16-year olds in 1989, dividing by the employment rate of 17-year olds in 1990, and by multiplying with 'one plus the death rate of 17-year olds in 1990.'

These backward calculations are applied to the young in 1990 for who the cohort-specific aging process is not applicable (to age groups 16-26 in 1990), for all years 1989 back to 1983, and similarly to all young in 2000 (age groups 15-28 in 2000), for all years 1999 back to

---

<sup>47</sup> The death rates, thus, are derived based on population rather than laborer data. Death rates of the population and the subgroup of laborers need not be identical if persons who are more likely to die, perhaps because they are ill, are less likely to work; on the other hand, the death rate among laborers could be higher than in the population due to work accidents.

<sup>48</sup> The 1 July 1982 census data come with death data of the calendar year 1981. Age-specific deaths for the 12 months prior to 1 July 1982 are obtained by taking the arithmetic average of the corresponding two age groups in 1981.

1991. It is further applied to the age groups 15-19 in each of the years 1978-81. For 2001 through 2003, the procedure is applied forward for the age groups 15-27 in 2000; the new age 15 cohorts in 2001, 2002, and 2003 are obtained by assuming their laborer size to be in the same proportion to their population size in 2000 as is the case for the 15-year olds in 2000, less deaths in the following years, and considering the year- and age-specific employment rates.<sup>49</sup>

For the oldest age groups, a similar procedure is needed, since, for example, the 58-year olds in 1982 do not have a comparable age group in the 1990 census. Thus, the individual age groups 57 through 63 in 1982 turn into the 1983 age groups 58 through 64 by taking into consideration the age- and year-specific employment rates and death rates. For example, to obtain the number of laborers age 58 in 1983, the age group 57 in 1982 is divided by the employment rate of the age group 57 in 1982 and multiplied by the employment rate of the age group 58 in 1983, and then divided by 'one plus the death rate of 58-year olds' in 1983. The same procedure is used for the years between 1990 and 2000.

A further complication is that the age group "65 and above" is a cumulative age group; the age group 57 in 1982, aged through 1990, does not match the cumulative "65 and above" category in 1990. Therefore, the age group "65 and above" across all years is obtained differently. The average annual geometric growth rate of the age group "65 and above" between 1982 and 1990 is applied to the 1982 value in order to obtain the 1983 value, etc. I.e., cohorts which are 8 years apart (while the age-level remains constant) are directly linked. The same procedure is used for the years between 1990 and 2000, with the average annual growth rate also applied to the years 2001-03.

In the years 1978-81, both the age group "65 and above" and 64 need to be constructed using average annual age-specific (rather than cohort-specific) growth rates of the period 1982 through 1990. The age group 64 also follows this procedure since it cannot be reconstructed, for example in 1981, from the age group "65 and above" in 1982. In 1981, the age groups 56-63 are reconstructed from the age groups 57-64 in 1982, backwards, by correcting for the different employment rates and including those who died between mid-year 1981 and mid-year 1982 at the corresponding 1982 age. In 1980, the age groups 55-63 are reconstructed etc.

The resulting age- and year-specific data on laborers do not necessarily add up in each year to the final employment data. The final (annual) employment data are up to one percent different from than the sum of age- and year-specific data, except in 1978 through 1980, when the difference is between one and two percent. The discrepancy is resolved by multiplying all age groups in a particular year with one and the same factor that makes the sum across age groups equal to the final employment value.

### *Step 2: Derivation of age- and education-specific numbers of laborers (excluding military personnel)*

In the second step, the age-specific data on laborers for all years are broken down into the educational categories illiterate/ semi-illiterate, primary school, lower middle school, upper

---

<sup>49</sup> The age-specific, year-specific death rates of 2001-2003 are interpolated using the age-specific average annual geometric growth rates between 1990 and 2000.

middle school (including special middle school), university-college, university-undergraduate studies, and university-graduate studies. Starting point is the education data by age from the three censuses and two 1% sample surveys. Since these data do not cover military personnel, a first step is to subtract the age-specific military personnel from the final age-specific laborer data in each year.

With age-specific military personnel data only available for 1982 (with the manipulations described above to obtain data in annual age categories), 1990 and 2000, age-specific military personnel data for other years are obtained using age-specific average annual geometric growth rates derived from the periods 1982-1990 and 1990-2000 to apply to the years within these two periods, as well as to 1978-81 and 2001-03. (Since, in each year apart from 1982, 1990, and 2000, the sum across age categories need not equal the interpolated total, each age category's value is slightly adjusted by one and the same year-specific factor to make the sum across age categories equal to the total.) The age-specific military personnel data in each year are then subtracted from the final age-specific laborer data in each year to yield age-specific data on laborers, excluding military personnel, in each year.

The age-specific education shares for laborers in the years 1982, 1987, 1990, 1995, and 2000 were calculated or constructed above. For all other years, the age-specific education shares are interpolated using the average annual geometric growth rates between age levels in any two adjacent censuses/surveys (with the first period's and last period's growth rates also applied to the years before 1982 and after 2000).<sup>50</sup>

In each year, these age-specific education shares for laborers can be applied to the age-specific number of laborers in order to obtain the *age- and education-specific* number of laborers. The year-specific final employment number, less military personnel, need not necessarily equal the year-specific *sum of laborers across age and education*; a small adjustment needs to be made to the number of laborers at each age-education combination using that year-specific factor which makes the sum of laborers across age and education equal to the final employment number less military personnel.<sup>51</sup>

*Additional step: Aggregation of year-specific, age-specific, and education specific number of laborers by age, and inclusion of military personnel*

For the purpose of obtaining aggregate human capital measures, the number of laborers need to be aggregated across age or education. To reflect the total of all laborers, military personnel can be added to the aggregate data on laborers. For the military personnel, a breakdown of the total by education is available in all three censuses. Values for other years can be interpolated using average annual education-specific geometric growth rates.<sup>52</sup>

---

<sup>50</sup> Cohort-specific average annual geometric growth rates could have been used, again raising questions, now in each education category, on how to handle the youngest and oldest age groups. A further issue is that cohort-specific growth rates would reflect changes in educational distribution as well as moves in and out of the military, i.e., would have no clear-cut meaning. With data on five benchmark years available, the best procedure now seemed to be interpolation between specific age levels.

<sup>51</sup> The adjustment is always below two percent except in some of the years before 1982 and after 2000.

<sup>52</sup> The sum across education categories in each year other than the census years need not perfectly equal the interpolated total; a small correction factor is needed, applied equally to each education category in a given year. In the 1982 census, the highest education category is tertiary level education; this category is split into the three different categories of tertiary level education (university-college, university-undergraduate studies, university-graduate studies) using the 1990 proportions.

Table 8 for 1978-90 and Table 9 for 1990-2003 report summary labor and human capital measures on all laborers, excluding or including military personnel. For example, the average number of years of schooling of the total of all laborers fell in the early reform years from 5.83 years or 5.87 years (excluding or including military personnel) in 1978 before rising gradually to 8.37 or 8.38 years in 2003. This calculation attributes 0 years of schooling to the illiterate and semi-illiterate laborers, 6 years to those with primary school education, 9 years to those with lower middle school education, 12 years to those with upper middle school (or special middle school) education, and 15/16/19 years to those with a university – college/ university – undergraduate/ university – graduate education.

The share of laborers with a tertiary level education rose from 0.83% in 1978 to 5.94% in 2003 if military personnel are excluded, and from 0.85% to 6.07% if military personnel are included. Although military personnel are vastly higher educated than other laborers, their numbers are so small in comparison that they make little difference overall. The share of laborers age 25-40 (excluding military personnel) rose from 37.75% in 1978 to 45.68% in 2003. The average age of laborers (excluding military personnel) rose from 32.18 years in 1978 to 37.71 years in 2003.<sup>53</sup>

Between 1960 and 1978, in some parts of China, primary and middle school education each lasted only for five years (rather than six). During the Cultural Revolution, furthermore, school degrees may have become inflated either because the duration of schooling was further reduced or because the content/ quality of schooling deteriorated. In as far as the education level of older age cohorts is exaggerated in comparison to the younger age cohorts, due to shorter periods of schooling at individual education levels in the pre-reform period or due to education inflation in the Cultural Revolution, growth in human capital over time is underestimated. Similarly, in as far as the quality of schooling rises over time, growth in human capital over time is underestimated. The rise in average years of schooling from 5.83 years in 1978 to 8.37 years in 2003, a rise of 43.57%, could well be an underestimate of the increase in the average education level of, very subjectively, perhaps 60-80%.

## **7. Derivation of the year-, age-, and education-specific number of laborers 2000-25**

The derivation of the year-, age-, and education-specific numbers of laborers in the years 2001-25 is achieved in three steps. First, the age-specific population is derived for all years. Second, the age-specific education characteristics of the population (in each year) are derived. Third, age-, education-, and year-specific employment rates are applied to the population data.

In contrast to the years 1978-2003, population rather than laborer data serve as basic building block. The starting point is population data because data on future laborers at the younger ages cannot be derived through any aging procedure but need to be constructed anew (age groups 15-40 in 2025). The population age 0-15 in 2000, on which data are available, turns into potential laborers, age 25-40, in 2025.

---

<sup>53</sup> The average age of the age group 65+ is taken to be 69 years. The 1% sample data of 1982 and 1990 have data on laborers through age 99; the average age of those age 65 or above in 1982 was 68.71 years, and in 1990 68.97 years.

## *Population by age, 2001-25*

Deriving the future age-specific population is complicated by the fact that births are routinely underreported. The number of youngest persons in the year 2000 benchmark population, thus, needs to be adjusted upward.

This adjustment is done by comparing the 1990 population with the 2000 population, shifting each age cohort (age X) in 1990 through to 2000 (then age X+10) using year- and age-specific death rates. (The year- and age-specific death rates are based on the 1990 and the 2000 age-specific death rates with age-specific interpolations for the years 1991-99 using the average annual geometric growth rates in the age-specific death rates between 1990 and 2000.) After taking into consideration death rates, the cohorts age 0-7 in 1990 grow significantly between the 1990 census and the 2000 census, by 12.68%, 7.90%, 4.73%, 4.59%, 4.60%, 2.60%, 6.73%, and 4.92% (also see Figure 4, which does not incorporate death rates).<sup>54</sup> Assuming that the 2000 population census data exhibit exactly the same downward bias at ages 0-7 as the 1990 population census data do, the 2000 population census data for the age groups 0-7 are adjusted upwards by these percentages. The result is an adjusted year 2000 benchmark population to be used in the forecasts of the age-specific population.

A second complication is that military personnel enters and exits the population data over time (military personnel data in the 2000 census are not included in the population data). Consequently, the year 2000 military personnel data, by age, are included with the 2000 population for the purpose of constructing post-2000 population data by age (including military personnel). Once the populations of the years 2001-25, by age and including military personnel, have been constructed, the military needs to again be excluded since all age times education information available excludes military personnel. It is assumed that the size and age structure of the military in the future is identical to that in 2000, and the 2000 military personnel, by age, are then subtracted from the constructed 'population plus military' in each year.<sup>55</sup>

The 2000 benchmark population is aged year for year using the age-specific death rates of 2000.<sup>56</sup> For example, the death rate of the age group X in 2000 is applied to the age group X-1 in 2000 to derive the population age X in 2001, to the age groups X-1 in 2001 to derive the population age X in 2002, etc. No adjustments to the year 2000 death rates are made; death rates between 1990 and 2000 fell across almost all age groups. But if death rates were to continue to fall at the rate at which they did between 1990 and 2000, they would, for some age groups, quickly approach zero. Environmental pollution, deteriorating health care, and an increasing number of work accidents could in the future counterbalance the downward trend

---

<sup>54</sup> The 1990 age 9 cohort also grows by 5.02%, but the age 8 cohort does not; neither do the cohorts at the ages above 9. The year 2000 number of persons age 9 is not corrected below.

<sup>55</sup> Not correcting for entry into/ exit from the military would imply that the 2000 age 18-21 cohorts, which provided half of all military personnel, remain of too small a size in the population/ laborers in future years (when military personnel leaves the military to join the population/ laborers). In practice, given the small size of the military personnel, the difference in final results is so small that ignoring entry into and exit from the military make no significant difference in the final results on human capital measures.

<sup>56</sup> The death rates used to age those in the age groups 0-7 each year may slightly overestimate the true death rate since the denominator in the calculation of the death rate is the unadjusted number of persons surviving at the specific age, rather than the adjusted number (which would take into account those persons whose birth was not reported). It is also unclear if deaths among those whose birth was not recorded are included in the census data. No correction was attempted.

in death rates. Ageing the 2000 benchmark population raises no further complications except that each year a new cohort is born; the number of births (age 0 on 1 July each year) is constructed using the 2000 ratio of newly born persons to the population age 20-40. It is assumed that this ratio is constant over time.<sup>57</sup>

### *Age-specific education characteristics*

One possible approach to obtain the age- and education-specific number of persons is based on extending year 2000 age-specific education characteristics into the future. At each age, the share of a particular education category in the benchmark year 2000 can be extended into the future based on the average annual change in this share between 1990 and 2000. In detail, in each year, at each age, the age-specific population number is divided into the category illiterate/ semi-illiterate persons and the six education categories using seven share values. The seven shares are the previous year's shares at the same, particular age (starting with 2001), times the average annual growth rate in each of the seven shares between 1990 and 2000. This procedure yields the number of persons by age and education in each year. But due to drastic changes in education characteristics between 1990 and 2000, this approach after a dozen years yields shares which in sum begin to well exceed one.<sup>58</sup> Past trends of changes in age-specific education shares, thus, cannot be easily extended into the future.

Therefore, this approach is used in the following only for the category of "illiterate/ semi-illiterate" persons and the education category "lower middle school," and only for the age groups 16-27 in the years 2001-11, reducing the coverage each year by one age group, from 16-27 to 17-27, 18-27, etc.<sup>59</sup> All other age times education numbers in the age groups 15-27 are derived in a more elaborate procedure which involves the imposition of official enrollment numbers on the youngest age cohorts each year. The age groups 28-65+, finally, are subject to a yet different method.

### Age groups 15-27

The main approach for the age groups 15-27 adopted here is to consider the most recent official enrollment data and to project these enrollment data into the future. The most recent years' official enrollment data can be compared to the official population data. This requires assumptions about the age at which persons enroll in a particular school. Table 10 has the assumed average enrollment age and the number of new students enrolled in a particular year

---

<sup>57</sup> The population age 20-40 includes military personnel, both in the construction and in the application of the ratio. The number of newly born persons in 2000 is the adjusted number (adjusted upwards by 12.68%, to reflect the between 1990 and 2000 revealed underreporting of births in 1990). The projected births increase slightly from 2000 through 2003, before declining gradually over time. The number of persons age 0 in 2020 is projected to be 85.19% of the number of persons age 0 in 2003, and in 2025 79.44%. In 2003, the constructed number of persons age 0 exceeds the official number from the 2003 1% sample survey by 0.70% (taking into consideration the exact sample size, independent of age, and the adjustment factor for the age group 0). (*Population Statistical Yearbook 2004*, preface and p. 4)

<sup>58</sup> The number of persons in all education categories could be adjusted proportionally, by dividing by the ratio of the sum of persons across education categories to the age-total, but at the younger ages the sum of education shares after two dozen years exceeds unity by so much that this approach becomes groundless.

<sup>59</sup> In as far as the sum of shares at a particular age level exceeds unity, the individual education categories at this age level are adjusted proportionally. The adjustments are on the scale of a few percentage points, with the highest value 15.61%.

relative to the population at the enrollment age. Table 11 reports the absolute enrollment numbers. Both tables also report data on graduates, which are smaller than the enrollment data throughout, a testimony primarily to the rapidly rising enrollment levels over time.

A standard education progression is assumed that all young persons follow. Students enter primary school at age 7 and leave primary school at age 13, they (most/ some) enter and exit lower middle school at the ages 13 and 16, upper middle school at the ages 16 and 19, college at the ages 19 and 22, undergraduate studies at the ages 19 and 23, and graduate studies at the ages 23 and 26.

According to Table 10, in 2000 practically every child age 7 enrolled in primary school.<sup>60</sup> In the following, primary school enrollment is taken to be 100% every year starting 2000. For simplicity, across all education levels, enrollment is meant to imply graduation at this education level after the typical duration of the particular school. With the minimum age for laborers being 15, this implies that starting 2000 all laborers at the youngest ages (15 in 2000, 15 and 16 in 2001, etc.) have at least a primary school education.<sup>61</sup>

New enrollment in lower middle school, even though supposedly compulsory, was equal to only 89.63% of the size of the age group 13 (Table 10). This ratio is higher than the self-reported number of “lower middle school” education in the 2000 census would suggest.<sup>62</sup> Possibly not all those newly enrolled in lower middle school in 2000 stated lower middle school as their highest educational level in the census. Alternatively, the official new enrollment data could be overestimates.

The ratio of lower middle school enrollment to the size of the age group 13 rises further, to 94.09% in 2004. Figure 36 shows how this drastic increase may have come about. The figure shows the size of each age group in the population age 0 through 28 in 2000 (with adjusted number for the age groups 0-7); year 2000 census data here and below are adjusted to 1 July. If children enter lower middle school at age 13, then those age 9 in 2000 entered lower middle school in 2004. But the age group 9 in 2000 (Figure 36) is about twenty percent smaller than the age group 10 in 2000. This suggests that lower middle schools faced a drastic decrease in new enrollment in 2004 unless they could reach those who in previous years would not have gone on to lower middle school. What is done in the following, in the forecast of future lower middle school enrollment, is to use the actual ratios in 2000 through 2004, and then to increase the 2004 share every year by a one-tenth share of the remaining gap to 100%.<sup>63</sup>

Ratios of new enrollment at the upper middle school level rose drastically in recent years even at a time of increasing age 16 cohorts. (See Figure 36; age group 16 in 2000 entered

---

<sup>60</sup> The ratio of new enrollment to population age 7 exceeds unity, presumably due to underreporting of the population age 7 in the census (no adjustments are made to the population data on the young in these tables). If the cohort age 6 had been used, the ratio would even be higher, if the cohort age 8 had been used, it would be about five percent smaller. The population age 6, 7, and 8 in the 2000 census is 16,951,679m, 18,193,873m, and 19,195,413m.

<sup>61</sup> In the 2000 census, 0.97% of the age group 15 either had no schooling or only literacy class.

<sup>62</sup> Figure 13 shows that only about 75% of recent, young age groups went to lower middle school.

<sup>63</sup> The actual ratios show a sharp one-period drop in 2003 which matches a population peak in the size of the age 13 group in that year (age 10 in 2000); the ratios of 2000 through 2004 are 0.8963, 0.8983, 0.9034, 0.8502, 0.9409. The drop seems excessive and could be smoothed. But in the following, the 2003 value is taken as is. This implies a spike in primary education as highest education level for this age cohort that remains visible through 2025.

upper middle school in 2000, age group 15 in 2000 entered upper middle school in 2001, etc.) In the following it is assumed that the year 2005 enrollment ratio equals the previous year's ratio times 1.1, and similarly for 2006; after 2006, the distance to an enrollment ratio of 75% is decreased by a one-tenth share of the gap every year. (For the years 2000-2004, actual new enrollment numbers are used; similarly below.)

New enrollment in university-level college or undergraduate education approximately doubled between 2000 and 2004, although the pool of potential applicants rose only in 2001, then fell back in 2002 and 2003, and stayed constant in 2004. New enrollment between 2003 and 2004, similar to previous years, still went up by about twenty percent. Between 2004 and 2006 the size of the age cohort 19 rises by about 30%. (See age groups 14 and 13 in 2000 for the years 2005 and 2006, in Figure 36.) In the following, it is assumed that the absolute number of new enrollment at these two levels increases by 20% in each year 2005 and 2006, stays constant in 2007-09, and then falls by 5% in each year 2010 and 2011 and by 2% in each year 2012-19 (parallel to the decrease in the size of the new age cohorts), before leveling out and remaining constant. In contrast to the middle school level, thus, the forecast is done in terms of absolute numbers (rather than in percentages of an entering age cohort). The reason is that universities are likely to exhibit considerable resistance to cut-backs in new enrollment numbers. Viewed differently, there is considerable momentum to maintain a status quo in absolute new enrollment numbers at the tertiary level once it has been reached.

At the university graduate level the growth in new enrollment over the recent years has been dramatic, a 2.55-fold increase between 2000 and 2004, even though the size of the relevant age groups was constant between 2000 and 2004. (See age groups 23, 22, 21, 20, and 19 in 2000 for the years 2000 through 2004, in Figure 36.) In the following it is assumed that the absolute number of new enrollment rises by 10% in each year 2005-08, a much lower growth rate than in the past, and then by 20% in each 2009 and 2010 to match the upcoming increase in the size of the relevant age cohorts; new enrollment is taken to then remain constant in 2011-13, to fall by 5% in each year 2014-15, by 2% in each year 2016-22, and then to remain constant at the 2022 level in 2023-25.

In the following, the standard education progression is imposed on the age groups 15-27, starting in 2000. No deviation from this pattern is allowed except for the illiterate and semi-illiterate persons and the lower middle school persons through 2011, until the age 15 cohort of year 2000 has moved through all age groups 16-27; this is done to allow some adult education at early ages to occur through the lower middle school level in the next few years. After 2011, the standard education progression applies fully to all persons age 15-27; education outside the standard government education system by assumption cannot occur below age 28. The imposition of the standard education progression on the age groups 15-27 involves a number of detailed rules on how to proceed. These are explained in the remainder of this section.

In each year the age cohort 15 faces the following manipulations. The number of lower middle school students is imposed via the ratio of new enrollment in lower middle school two years earlier divided by the number of persons age 13 two years earlier, times the number of persons age 15 in the current year.<sup>64</sup> No person is allowed in the categories upper middle

---

<sup>64</sup> In 2000, 2001, and 2002 the year 2000 ratio is used, in 2003 the year 2001 ratio, in 2004 the year 2002 ratio, etc. The logic behind the procedure is that two years earlier, at age 13, a certain share of the age cohort enrolled in lower middle school. The number of persons who enrolled stays the same in the following two years (nobody

school and above. Due to the assumption that 100% of each new age 15 cohort goes through primary school, there are no illiterate or semi-illiterate persons. The number of persons with primary school education is obtained as residual.

In each year the age cohort 16 faces the following manipulations. The ratio of upper middle school new enrollment is applied to the number of persons at this age; those age 15 and in lower middle school in the previous year are aged by one year and those newly enrolled in upper middle school are subtracted (graduation from lower middle school is a requirement to be admitted to upper middle school).<sup>65</sup>

In each year the age cohort 19 is subjected to the new enrollment numbers for college and undergraduate studies; those age 18 and in upper middle school in the previous year are aged by one year and those newly enrolled at university level are subtracted. In each year the age cohort 23 is subjected to the new enrollment number for graduate studies; those age 22 and enrolled in undergraduate studies in the previous year are aged by one year and those newly enrolled in graduate studies are subtracted. At ages 16-18, the number of persons in the three university categories is set equal to zero; at ages 19-22, the number of persons enrolled in graduate studies is set equal to zero.

In 2000, the number of persons age 16 and with a primary education is obtained as residual; the number of persons age 17-18 in the upper middle school category is the sum of those enrolled at upper middle school and above (nobody age 17-18 is allowed to be in a category higher than upper middle school, with actual numbers pulled back into the upper middle school category); the number of persons age 20-22 in the undergraduate studies category is the sum of those enrolled in undergraduate or in graduate studies (the graduate studies category is set equal to zero for the age groups 20-22).

In 2001, the standard education progression is imposed on all those at upper middle school through university graduate studies (starting age 16, through age 27), be it through the imposition of new enrollment numbers on specific age groups as outlined above, or through aging of the corresponding age cohort in 2000.<sup>66</sup> For the age group 16 the persons age 15 in 2000 with primary school education are aged (yielding an identical result to calculating this number as residual). The primary school numbers for the age groups 17-27 are obtained as residuals. The number of illiterate and semi-illiterate persons age 17 and above and the number of persons in the lower middle school category age 18 and above are allowed to continue to change following the default procedure (based on trended shares in total population of a specific age); i.e., illiterate and semi-illiterate persons age 17 and above may gain a primary education outside the official school system, and primary school educated persons age 18 and above may gain a lower middle school education outside the official school system.

---

exits, nobody enters); deaths are taken into consideration by using the share in the cohort total two years ago, applied to the cohort total in the current year, where cohort totals take into consideration deaths.

<sup>65</sup> Aging by one year means applying the year 2000 age-specific death rate; this rate is already embodied in the number of persons age X this year vs. the number of persons age X-1 in the previous year. The total number of persons at each age level in all years after 2000 was constructed based on the year 2000 census population and the year 2000 death rates as explained above.

<sup>66</sup> "Aging" implies the application of the age-specific death rate. (This is done by multiplying with the current-year total number of persons age X and dividing by the previous-year total number of persons age X-1.)

In 2002, the age group 15 in 2000 turns into the age group 17; for illiterate and semi-illiterate persons, this number is zero (above age 17, the default procedure applies). The in 2001 age 17 lower middle school number of persons is aged by one year to become the age 18 number of lower middle school persons (above age 18, the default procedure applies). The numbers of persons with primary school education as highest education level in the age groups 18-27 are obtained as residuals. For the two categories illiterate and semi-illiterate persons and lower middle school persons the default procedure recedes by one age group each year, and so does the calculation of the primary school numbers as residuals.

#### Age groups 28-65+

Starting with age 28 a new procedure applies. The key is a comparison of the number of persons age X in 2000 in a particular education category to the number of persons age X-10 in 1990 (less deaths during 1990 and 2000) in this education category. In comparisons of persons age X-10 in 1990 and age X in 2000 below, the 1990 data are always aged, i.e., the 1990 number of persons at age X-10 is subjected to the death rates of persons age X-9 in 1991, age X-8 in 1992, etc., before the comparison is made.

A number of figures explore the relationships by age and education level. Figure 37 relates the 'number of persons age X in 2000 in a particular education category less the number of persons age X-10 in 1990 (aged to 2000) in this education category' to the number of persons age X-10 in 1990 (aged to 2000). I.e., it standardizes the absolute change in the number of persons in a specific education category for a specific age-cohort by the (aged) original size of the age cohort. At each age level in 1990, the number of illiterate and semi-illiterate persons who leave this category over the next 10 years is between 3% and 8% of the age cohort. The primary school category loses a substantial number of young persons over the 10-year period, but then gains middle-aged and older persons. The lower middle school gains substantial numbers of persons throughout. The category upper middle school flip-flops at the young age groups but then gains a small number of persons starting around age 45 in 2000. The category college gains persons throughout, although not as substantially as the lower middle school group. The category undergraduate studies gains persons only at the youngest ages, and the category graduate studies experiences no substantial in- or outflow.

But what is substantial relative to the size of the whole age cohort need not be substantial for the specific education category. Figure 38 reports the X-fold (percentage divided by 100) change in the number of persons of a specific age cohort at a specific education level between 1990 and 2000 (with 1990 numbers aged to 2000), Figure 39 the 2000 number relative to the 1990 (aged) number, and Figure 40 the average annual change. The picture changes above all for the highest three education levels. These show increases that may not be substantial when compared to the size of the whole age cohort but are significant when compared to the same age cohort's same education category 10 years earlier. An exception is the number of persons with a university graduate level education; for the older age cohorts this number is negative (a few thousand persons at each age). Presumably this is a Cultural Revolution effect; technically, the age cohort-wide death rate is too low for this education category. Presumably, the Cultural Revolution is also the cause of the significant increases in the number of middle-aged persons in the tertiary level categories between 1990 and 2000. The number of illiterate and semi-illiterate persons still falls significantly every year (for every age cohort) and the number in the lower middle school category still increases, although now by a smaller proportion, every year.

These figures provide an impression of the change in the number of persons by age and education level over the 10-year period 1990 and 2000. But it is unclear if this 10-year change, even if annualized, still holds for the transition between 2000 and 2001. A reliable comparison between 1999 and 2000 would have been more informative for the future than the comparison between 1990 and 2000.<sup>67</sup> A compromise is to assume that these measures of change apply most accurately sometime *between* 1990 and 2000. For convenience below, they are assumed to be accurate measures for 1997. This implies that a measure of change for the age cohort X in 2000 presented in the figures is really a measure of change for this age cohort in 1997, when it was X-3 years old. Thus a measure of change for the age cohort 31 in 2000 in the figures is assumed to be the accurate measure of change for the age group 28 in 1997.

The next step is to smoothen the series of average *annual* changes in the individual education levels across age groups (Figure 40) and to apply them to the year 2000 benchmark population data in order to obtain year 2001 data, then to apply them to the year 2001 data in order to obtain year 2002 data, etc. Figure 41 depicts the smoothed series starting from age 28, corresponding to age 31 in 2000. Age 31 in 2000 in the previous figures was a turning point for transition rates at the tertiary level; starting age 31 in 2000, transition rates appear somewhat reasonable (reflecting adult education rather than regular education). At the same time, transition rates in the creation of post-2000 populations by age and education are needed only for the age groups 28 and above (age groups 15-27 are subjected to a different approach as explained above).<sup>68</sup>

For the category illiterate/semi-illiterate persons, the first and last values (0.9225 at age 28, i.e., age 31 in 2000, and 0.9855 at age 61, i.e., age 64 in 2000) are connected linearly. The primary school category is depicted using the actual (1990-2000) average annual growth rates, but in the calculations below the number of persons in the primary school category is obtained as residual. For the lower middle school, a 0.5% growth rate is imposed across all age groups. For the upper middle school, zero growth is assumed through age 45 in 2000 (some persons enter, but others exit this category) and a 0.7% average annual growth rate thereafter (corresponding to an approximately 10% 10-year growth rate). For the university-college category, the actual average annual growth rate is approximately halved (arithmetically halved) by imposing a 4% growth rate on the age group 31 (in 2000) and a 1.5% growth rate on the age group 50 (in 2000), and by connecting these two values linearly; above age 50 the growth rate is held constant at 1.5%. University undergraduate and graduate studies are assumed to not be attainable through adult education (zero growth rate).<sup>69</sup>

---

<sup>67</sup> The 1995 1% sample survey, as noted above, suffers from biases in the distribution of the population across age groups, and can thus not be used for the purpose here.

<sup>68</sup> In the comparisons above and in the construction of Figure 41 the military had to be ignored since military personnel data by age times education are not available. This introduces inconsistencies in the rates of change between 1990 and 2000. Since the size of the military is very small relative to the population, these inconsistencies should be very minor and should make no difference in the choice of smoothed series.

Education while in the military is likely to contribute to educational upgrading outside the regular education system over time. Higher-level education of military personnel is likely to occur through educational institutions within the military, enrollment numbers at which are presumably not included in the official data on enrollment numbers at regular institutions of education.

<sup>69</sup> The *Census 2000*, Vol. 2, pp. 874-7, has limited data on adult education covering the long-form sample. Of all persons age 15 through 65+, 242,380 gave as highest education level university undergraduate studies through adult education, but the numbers are concentrated in the age groups 29-37, which could suggest a catching up process in the reform period. For comparison, 1,278,851 persons had studied at the university-

This smoothed transition pattern implies significantly lower transition (growth) rates for tertiary level education than the actual data of 1990 and 2000 suggest. However, the 1990s may have seen a large volume of “catch-up” following the Cultural Revolution that may have run its course by the first decade of the twentieth century. Furthermore, the absolute increase in numbers was related to a rather low base. Even if the imposed transition structure were too conservative for the tertiary level, this may not matter much in the aggregate (across all age cohorts in a particular year) because in as far as late graduation from regular tertiary level institutions is underestimated, the official, regular (not adult) enrollment numbers are imposed in full on the youngest age cohorts. The calculations here then impose tertiary level education slightly earlier, on average, than it may occur in reality. Over a 25-year horizon, this does not matter much as cohorts age 40 and below in 2000 (where most adult education may occur) will all be included in the aggregate data, in all years.

Lacking a sound basis on which to age the transition pattern over the years, one and the same transition pattern (Figure 41), age 28-61 by education, is applied equally in each year.<sup>70</sup> In 2001, the transition pattern is applied to the age groups 27-60 in 2000 to obtain age times education data of the age groups 28-61 in 2000. In 2002, the same transition pattern is applied to the age groups 27-60 in 2001, where the age group 27 in 2001 was derived in the previous section (together with the age groups 15-26). In each year, the primary school category is obtained as residual.

The number of persons, by education level, at age 62, 63, and 64 in each year is obtained by aging the age groups 61, 62, and 63 of the previous year (by education level); the transition rate is uniformly one. To obtain the number of persons age 65 and above in each year, first, for each education category in 2000 the number of persons age 65 and above is divided by the number of persons age 60-64. These education-specific ratios of 2000 are then applied to the education-specific numbers of persons age 60-64 in all other years.<sup>71</sup>

An alternative procedure for the age groups 28-65+ would have been to use the absolute change in the age- and education-specific number of persons between 1990 (aged) and 2000, and to relate it to the absolute number of persons of the same age cohort in the education category one degree lower (in 1990) from which they moved up. This would have been a measure of what share of an age- and education-specific group of people “upgraded” to the next higher education level during 1990 and 2000. Ideally, the denominator would be this age cohort at the lower education level some years earlier (before the upgrading began), but it is unclear how many years earlier would be appropriate (for example in the presence of part-time studies).

Figure 42 depicts these series. The patterns are not much different from those obtained in the procedure used here (compare Figure 37); prominent developments are perhaps slightly sharpened, and variability increases in size. Thus, a large share of those illiterate at age 28/29 in 1990 (8/9 in 1970) gained a primary school education between 1990 and 2000, and of those with a primary school education gained a lower middle school education; presumably

---

college level through adult education, 761,725 at the special middle school level, and 512,821 at the upper middle school level. (Data on adult education at other education levels were not collected in the census.)

<sup>70</sup> Age-specific death rates are considered through adjustment in each year by applying the ratio of total number of persons age X this year divided by total number of persons age X-1 in the previous year.

<sup>71</sup> The death numbers in the population census 1990 only go through age 90, which means that a complete number of persons age 65 and above in years after 2000 cannot be derived.

this is an aftereffect of the Cultural Revolution which will not be repeated in the future. Ignoring such special effects, the lower middle school and college education categories tend to gain in size (with large gains for the lower middle school at young ages, falling over time), and the upper middle school only gains at the higher age levels (above 45), patterns which the procedure used here incorporates.

### *From population to laborers*

The final step is to apply age- and education-specific employment rates to the age- and education-specific population data in each year. The age- and education-specific employment rates are those of 2000.

The year 2000 census collected age- and education-specific data on laborers in the long-form sample. Age- and education-specific data on the population were collected from everybody (those who filled in the short form, and those who filled in the long form), but data on the age- and education-specific population are only reported in full for the complete population (not for the long-form sample separately). Age- and education-specific employment rates then are obtained in two steps. First, all laborers at a particular age, independent of education category, are augmented by the ratio of the census population at this age to the long-form sample population at this age; second, these economy-wide age- and education-specific numbers of laborers are divided by the age- and education-specific population to obtain the age- and education-specific employment rates.

Comparing the 2000 age- and education-specific employment rates to those in 1990, at most age-education combinations there is little difference. But employment rates in a few age-education combinations are markedly different, such as higher employment rates for the young with tertiary education in 2000 (presumably because they obtained their tertiary education a year or two earlier than in 1990). If post-2000 employment rates are obtained by applying the average annual change in age- and education-specific employment rates between 1990 and 2000 to the 2000 employment rates, some age-education combinations end up with employment rates a few years after 2000 which do not make sense. Therefore, the year 2000 employment rates are used below for all years 2000-25.

While Figure 6 above suggested a lowering of employment rates over time, Figure 22 and Figure 24, when examined carefully, show little change (note the different vertical scale). What causes the lowering of aggregate employment rates (age-specific, across all education categories) in Figure 6 is a shift in the relative number of laborers between different education categories, with different education categories exhibiting different employment rates. The education categories lower and upper middle school tend to have lower employment rates at all ages than the primary school category; the tertiary level categories tend to have the highest employment rates, but the number of laborers in these latter categories tends to be small. As an increasing number of laborers moves to a lower or upper middle school education as highest education level, aggregate employment rates fall.

Consequently, by 2025, the sum of employment across education categories (where the age- and education-specific population have been subjected to the age- and education-specific 2000 employment rates) is only 92.76% of a total employment number derived by applying the 2000 age-specific (aggregate across education categories at that age) employment rates to the 2025 age-specific population (and summing up across age). The sum across education

categories (and age groups) appears the more reliable measure of laborers than the total (across age groups), and in the following, therefore, no adjustments are made for the discrepancy between the sum across age- and education categories and the total (sum of total across age groups).<sup>72</sup>

The resulting data on age- and education-specific numbers of laborers exclude military personnel. Assuming that the size of the year 2000 military personnel and their education characteristics stay constant in the future, military personnel can be added in every year. At education levels up through lower middle school, this may imply some double-counting (military personnel presumably received primary and lower middle school education in the regular education system). The fact that about half of the military personnel in 2000 was age 21 or below suggests that above lower middle school these persons received further education while in the military, and presumably in military-run schools which are unlikely to be included in the data on regular institutions of education. But because the military personnel numbers are so small relative to the population and laborers, even at the ages 18-21 (which account for half of all military personnel) equivalent to at most three percent of the population in these age groups, it makes no big difference how the military is treated. The military, on average, is better educated than the population, and their inclusion therefore slightly raises human capital measures.

Table 12 and Table 13 report the aggregate human capital measures for each year between 2000 and 2025, with or without military. The underlying year 2000 population at the youngest ages already has the official enrollment numbers imposed, which implies that the year 2000 human capital measures in Table 12 are slightly different from those in Table 9 (census values).

Figure 43, Figure 44, and Figure 45 chart the education characteristics of the age groups 15 through 65+ in 2000, 2015, and 2025. No attempt is made to correct for historical anomalies, such as the large number of persons and therefore laborers in the upper middle school category around age 35-37 in 2000, or the size of the age 25 group in 2015 (unusually low enrollment rate in lower middle school in 2003, combined with a large age group 13). In the aggregate, in human capital measures across the economy, these small aberrations—if they are aberrations—play no significant role.

---

<sup>72</sup> The impact of adjusting (dividing each age-education combination by the ratio of the sum across education categories at that age to the total at that age) on the final human capital measures of interest here is very minor.

## References

- Census 1982. Zhongguo 1982 nian renkou pucha ziliao* (Tabulation of the 1982 Population Census of the People's Republic of China). Compiled by the State Council Population Census Office and the NBS Population Division. Beijing: Zhongguo tongji chubanshe, 1985.
- Census 1990. Zhongguo 1990 nian renkou pucha ziliao* (Tabulation of the 1990 Population Census of the People's Republic of China). Compiled by the State Council Population Census Office and the NBS Population Division. Four volumes. Beijing: Zhongguo tongji chubanshe, 1993.
- Census 2000. Zhongguo 2000 nian renkou pucha ziliao* (Tabulation of the 2000 Population Census of the People's Republic of China). Compiled by the State Council Population Census Office and the NBS Population, Society, and Technology Division. Three volumes. Beijing: Zhongguo tongji chubanshe, 2002.
- Holz, Carsten A., and Yi-min Lin. "Pitfalls of China's Industrial Statistics: Inconsistencies and Specification Problems." *The China Review* 1, no. 1 (Fall 2001): 29-71.
- Labor Yearbook. Zhongguo laodong tongji nianjian* (China Labor Statistical Yearbook). Beijing: Zhongguo tongji chubanshe, various years.
- Liu Chengxiang, Liu Ke, Jin Zhaofeng. *Ruhe shiyong tongji nianjian* (How to use the Statistical Yearbook). Beijing: Zhongguo tongji chubanshe, 2000.
- Population Statistical Material 1949-1985. Zhonghua renmin gongheguo renkou tongji ziliao huibian* (Collection of PRC population statistics materials). Beijing: Zhongguo caizheng jingji chubanshe, 1988.
- Population Statistical Yearbook. Zhongguo renkou tongji nian jian* (China Population Statistical Yearbook). Beijing: Zhongguo tongji chubanshe, various years.
- Statistical Yearbook. Zhongguo tongji nianjian* (China Statistical Yearbook). Beijing: Zhongguo tongji chubanshe, various years.
- Survey 1987. Zhongguo 1987 nian 1% renkou chouyang diaocha ziliao*. (Tabulation of the 1987 1% Population Sample Survey of the People's Republic of China). Compiled by the NBS Population Division. Beijing: Zhongguo tongji chubanshe, 1988.
- Survey 1995. Zhongguo 1995 nian 1% renkou chouyang diaocha ziliao*. (Tabulation of the 1995 1% Population Sample Survey of the People's Republic of China). Beijing: Zhongguo tongji chubanshe, 1997.

**Table 1. Data Availability in Censuses and 1% Sample Surveys**

	Population		Laborers		Military	
	by age	by age * education	by age	by age * education	by age	by education
1 July 1982	v	5-year; ann. for 1%	5-year; ann. for 1%	NA	5-year	for total
1 July 1987 (survey)	v	v	v	v		
1 July 1990	v	v	v	educ. for total	v	for total
1 Nov. 1995 (survey)	v	v	v	v		
1 Nov. 2000	v	v	v (long form)	v (long form)	v	for total
<i>Sources (Volume / page numbers)</i>						
<i>Census 1982</i>	273-81	360f.	468f. [ <i>Pop. St. Y. 1988</i> , pp. 742-4]		505	505
<i>Survey 1987</i>	135f.	164f., 172	224f.	230-3, 468f.		
<i>Census 1990</i>	II/2-5	II/112-51, 278-81	II/476-8	[ <i>Pop. St. Y. 1993</i> , pp. 66f., 88f.]	IV/495	IV/496
<i>Survey 1995</i>	11-3	26-33	124-83	124-83		
<i>Census 2000</i>	I/215-7	I/593-602	II/1241f	II/1249-68	III/1883	III/1884

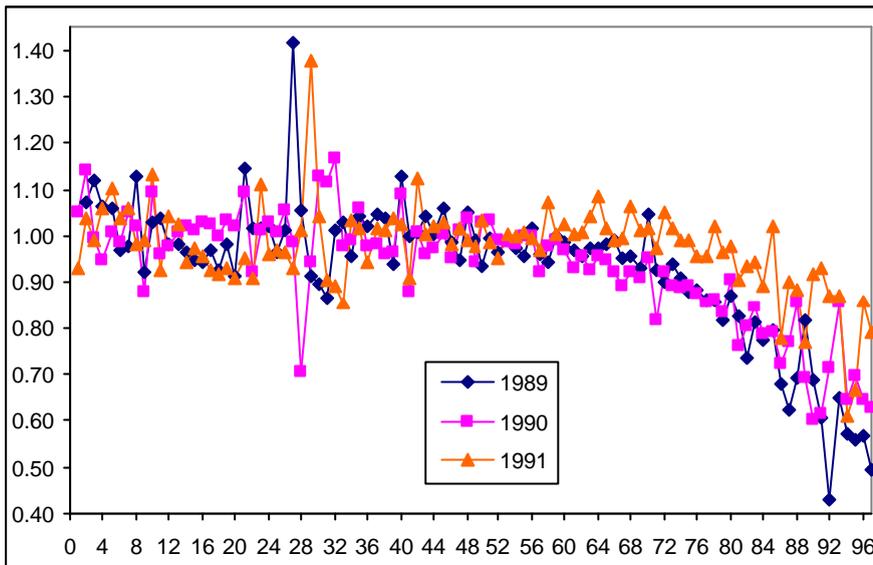
For the 1990 population by age times education, data are available according to education *completion level* for each of the education categories (rather than as totals across all completion levels only). For the 2000 population by age times education, data are available according to education completion level for each of the two education categories primary school and lower middle school only.

Labor data cover the age range 15-65+, i.e., data are available in the age categories, 15, 16, ..., 64, "65 and above," except in 1982, when the data are in 5-year age categories and end with a category "60 and above" (15-19, 20-24, ..., 60+); 1982 labor data by (annual) age are available for a 1% sample of the census in the *Population Statistical Yearbook 1988*, pp. 742-4. Year 1990 labor data by education (but not age, i.e., only the education breakdown of total laborers) are available in the *Population Statistical Yearbook 1993*, pp. 66f., 88f. Year 2000 labor data are from the long form of the census, covering approximately 9.5% of the population. Military data by age in 1982 start with a category "19 and below," then go in 5-year intervals to "60 and above;" in 1990 and in 2000 military data start with a category "below 18," then go in 1-year steps to "60 and above."

In 2000, the military is explicitly not included in the data on population and laborers; presumably the same is the case in the other censuses/ surveys.

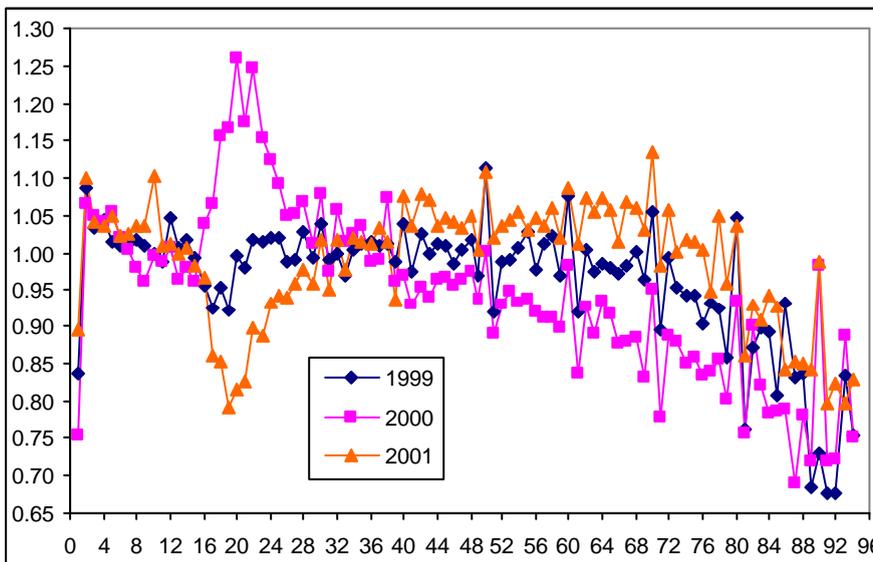
Annual population change 1% sample surveys cover population by age; these are reported for 1988 (in 5-year age groups), 1989, 1991, 1992, 1993, 1994, 1996, 1997, 1998, 1999, 2001, 2002, and 2003 in the *Population Statistical Yearbook*. The 1994 issue, pp. 30-5, reports the 1993 population by age times education.

Sources: *Census 1982*, *Survey 1987*, *Census 1990*, *Survey 1995*, *Census 2000*, and as otherwise stated.



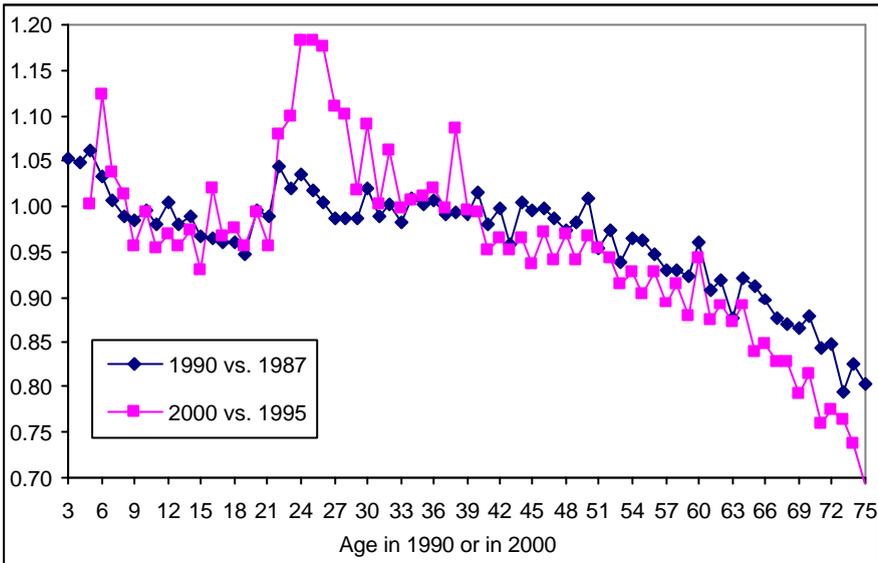
Each data point reflects the number of persons age X in this year relative to the number of persons age X-1 in the previous year, except those data of 1989, which compare the number of persons age X in 1989 to the number of persons age X-2 in 1987. 1987 population data are 1% sample survey data, 1989 and 1991 population data are 1‰ sample survey data from the *Population Statistical Yearbook* (1990, pp. 316-8; 1992; pp. 8-10), 1990 population data are census data. For comparability, 1987, 1989 and 1991 sample survey data are augmented (multiplied by a factor of approximately 100, 1000, and 1000) using total annual end-year (midyear for 1987, obtained by linear interpolation) population data from the *Statistical Yearbook 2001*, p. 91, for 1986 and 1987, and *2004*, p. 95 for 1990 and 1991. Year 1987 sample survey and 1990 census population data are of 1 July of that year; the dates of the 1‰ sample survey data are unknown (possibly end-year).

**Figure 1. Cohort Growth Rates 1989, 1990, 1991**



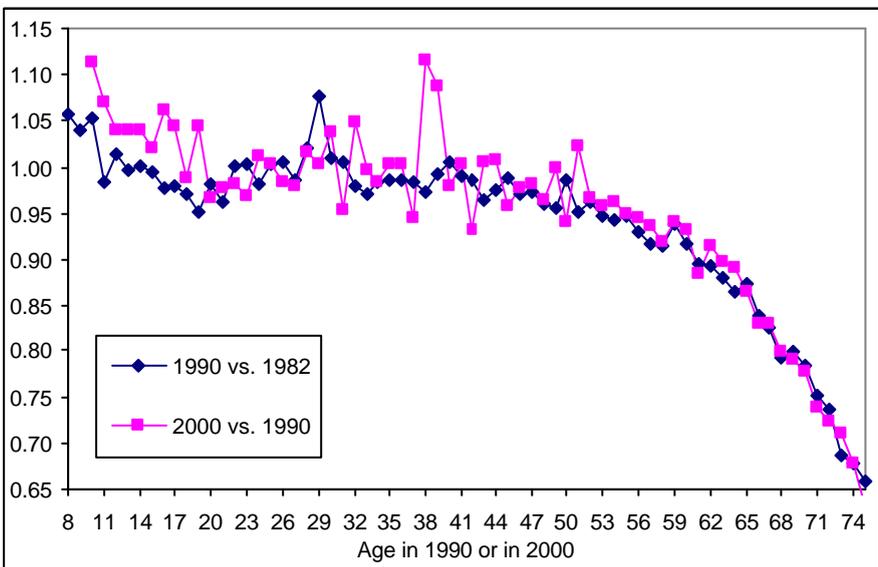
Each data point reflects the number of persons age X in this year relative to the number of persons age X-1 in the previous year. 1998, 1999 and 2001 population data are 1‰ sample survey data from the *Population Statistical Yearbook* (1999, pp. 4-6; 2000, pp. 4-6; 2002, pp. 4-6), 2000 population data are census data. For comparability, 1998, 1999 and 2001 sample survey data are augmented (multiplied by a factor of approximately 1000) using total annual end-year population data from the *Statistical Yearbook 2004*, p. 95. Year 2000 census population data are of 1 November 2000; the dates of the 1‰ sample survey data are unknown (possibly end-year).

**Figure 2. Cohort Growth Rates 1999, 2000, 2001**



Each data point in the 1990 vs. 1987 comparison reflects the number of persons age X in 1990 relative to the number of persons age X-3 in 1987, and similarly for the 2000 vs. 1995 (X-5) comparison. 1987 and 1995 population data are 1% sample survey data, 1990 and 2000 population data are census data. For comparability, 1987 and 1995 sample survey data are augmented (multiplied by a factor of approximately 100) using total annual end-year (midyear for 1987, obtained by linear interpolation) population data from the *Statistical Yearbook 2001*, p. 91, for 1986 and 1987, and 2004, p. 95, for 1995. Year 1987 sample survey and year 1990 census population data are of 1 July of that year, year 1995 sample survey and year 2000 census population data are of 1 November of that year. The final year in the chart, of 75, is chosen to keep the vertical range of the chart relatively large.

**Figure 3. Cohort Growth Rates 1990 vs. 1987, 2000 vs. 1995**



Each data point in the 1990 vs. 1982 comparison reflects the number of persons age X in 1990 relative to the number of persons age X-8 in 1982, and similarly, but with X-10, for the 2000 vs. 1990 comparison. All data are population census data; year 2000 population census data are adjusted to 1 July, to match the census day of the other two censuses. The final year in the chart, of 75, is chosen to keep the vertical range of the chart relatively large.

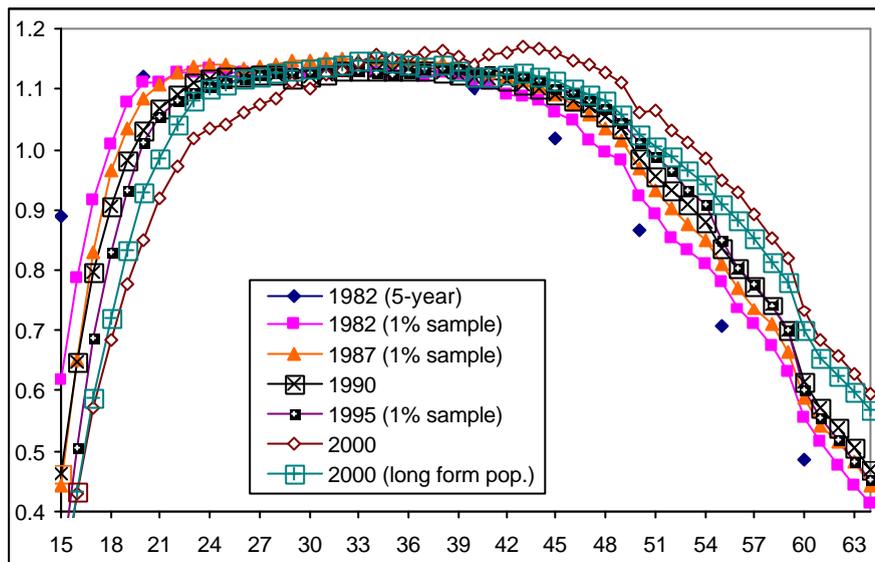
**Figure 4. Cohort Growth Rates 1990 vs. 1982, 2000 vs. 1990**

**Table 2. Inconsistencies Between Censuses**

Date of birth	Age	1964 Population	Age	1982 Population	Age	1990 Population	Age	2000 Population
64/65			17	24,417,934	25	24,520,614	35	24,581,035
63/64	0	28,483,827	18	25,131,408	26	25,287,489	36	25,385,295
62/63	1	30,248,104	19	27,379,596	27	27,026,864	37	25,551,830
61/62	2	<i>15,569,154</i>	20	<i>15,620,970</i>	28	<i>15,928,062</i>	38	<i>17,753,414</i>
60/61	3	<i>11,535,264</i>	21	<i>10,690,834</i>	29	<i>11,504,496</i>	39	<i>12,504,253</i>
59/60	4	14,305,607	22	14,307,196	30	14,443,119	40	14,143,080
58/59	5	14,883,857	23	14,284,240	31	14,378,626	41	14,418,518
57/58	6	20,396,314	24	19,459,780	32	19,060,366	42	17,756,795
56/57	7	19,716,266	25	18,876,181	33	18,350,283	43	18,453,625

Date of birth refers to between 1 July in the first year and 30 June in the second year. Italics are for emphasis.

Sources: see Table 1; for 1964 census data see *Population Statistical Material 1949-1985*, pp. 602-5.



Each point in the chart shows the ratio of (a) to (b), where (a) is the share of laborers at the particular age in all laborers age 15-64, and (b) is the share of persons at the particular age in the total population age 15-64. Age 64 is taken as cut-off point, since labor data for age 65 comprise age “65 and above;” a comparable population value could have been constructed, but appears meaningless since laborers increasingly retire as age exceeds 65.

1982 (5-year): data on laborers are only available in 5-year groups, with as last group “age 60 and above;” these data are compared to population data aggregated into similar groups, with the last group consisting of those aged 60-64. Data are census data.

1982 (1% sample): data on laborers are from an official 1% sample of the census; population data are the complete census data.

1987 (total laborers): data covers all laborers (*zaiye renkou*) in the 1% survey sample. (Population data are 1% sample survey data.)

1990: census data.

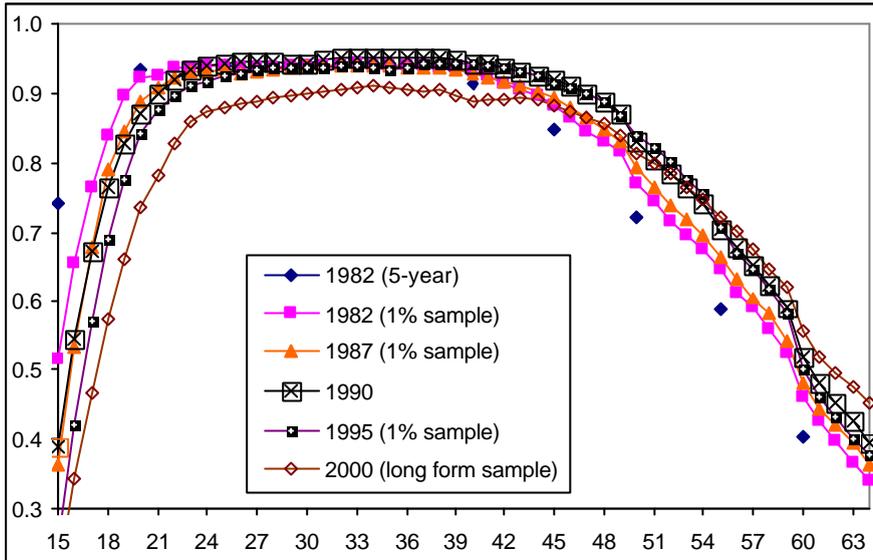
Census/ sample survey day for all of the above is 1 July.

1995: 1% sample survey data. (Sample survey date is 1 Nov. 1995.)

2000: laborers are from the long-form (approximately) 9.5% sample of the census, population data from the complete census. (Census date is 1 Nov. 2000.)

2000 (long-form pop.); both laborers and population are from the long-form (approximately) 9.5% sample of the census. (Census date is 1 Nov. 2000.)

**Figure 5. Participation Share Ratios**



Age 64 is taken as cut-off point, since labor data for age 65 comprise age “65 and above;” a comparable population value could have been constructed, but appears meaningless since laborers increasingly retire as age exceeds 65.

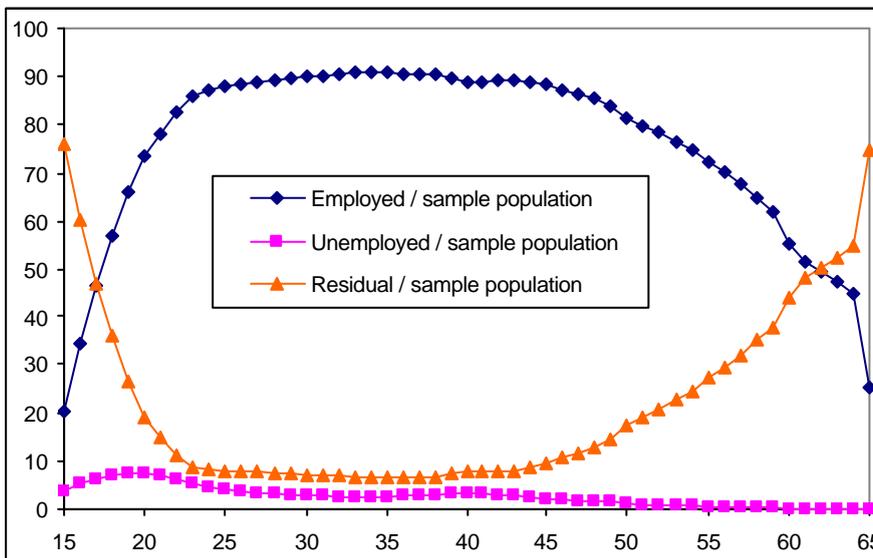
1982 (5-year): data on laborers are only available in 5-year groups, with as last group “age 60 and above;” these data are compared to population data aggregated into similar groups, with the last group consisting of those aged 60-64.

In 2000, data on laborers and population are from the long form sample.

Census/ sample survey day through 1990 is 1 July, and 1 November since then.

Sources: 1% samples in 1982 and 1990, *Census 2000*, Vol. 2, pp. 1241f. for 2000.

**Figure 6. Employment Rates (in %), Given Age, 1982, 1987, 1990, 1995, 2000**

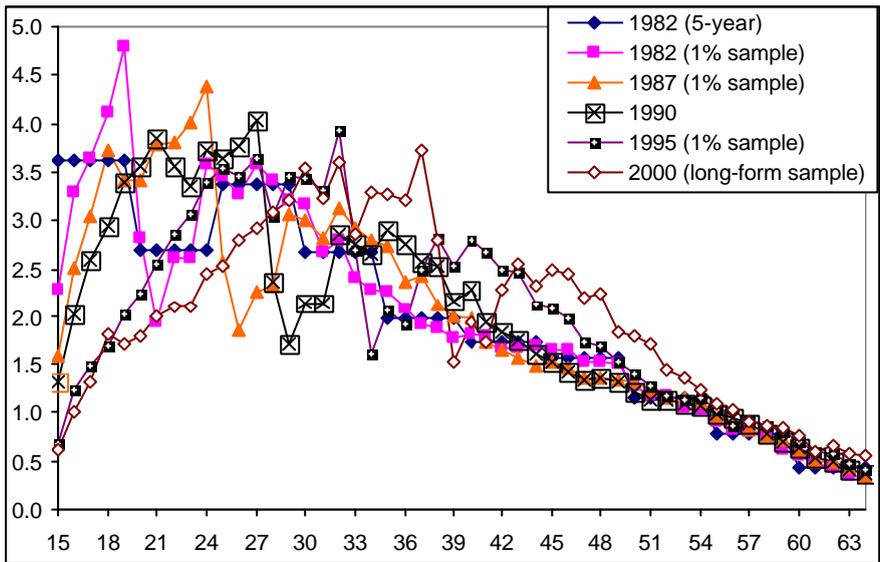


Sample refers to the persons who filled in the long form. Employment data are only collected in the long form.

Age 65 refers to “age 65 and above.”

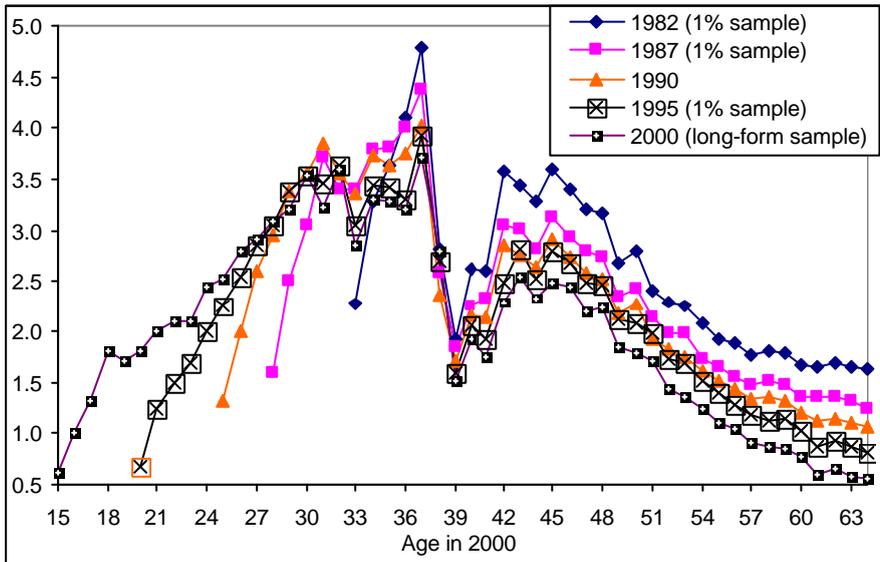
Sources: *Census 2000*, pp. 1241f.

**Figure 7. Laborers, Unemployed, and Residual Population (in %), 2000**



1995 and 2000 data are of 1 Nov., all other data of 1 July.  
Sources: see Table 1.

**Figure 8. Age-Specific Shares in Labor (in %)**



Age reflects the age in 2000, across all censuses and surveys. Thus, for example, the age category 40 (all data points in the figure where the horizontal value is 40) is age 40 in the census 2000, age 35 in the 1995 survey, age 30 in the 1990 census, etc.

1995 and 2000 data are of 1 Nov., all other data of 1 July.  
Sources: see Table 1.

**Figure 9. Age-Specific Shares in Labor, at Year 2000 Age (in %)**

**Table 3. Education Level Correspondence between Censuses**

	Census 1982 Whole population age 6-65+ (reported by 5-year age groups)	Census 1990 / survey 1987* Whole population age 6-65+ (reported by age)	Census 2000 / survey 1995* Whole population age 6-65+ (reported by age)
A	Knows no or few characters (age 12+: identical to illiterate/ semi-illiterate) -> available data on illiterate/ semi-illiterate only	Knows no or few characters (age 15+: identical to illiterate/ semi-illiterate) -> available data on illiterate/ semi-illiterate only	No schooling Literacy class
B	Primary school	Primary school	Primary school
C	Lower middle school	Lower middle school	Lower middle school
D	Upper middle school (incl. special middle sch.)	Upper middle school	Upper middle school
		Special middle school	Special middle school
E	University (graduated, vs. all else: currently attending/ completed but not graduated or discontinued studies)	University – college	University – college
		University – undergraduate/ graduate studies	University – undergraduate studies
			University – graduate studies (graduated)
	B-D: data not available according to completion level; E: reported at two completion levels	B-E: data available at all completion levels (and total)	No data available according to completion level (only total), except for primary and lower middle school in sample

\* The survey data (1987 and 1995) do not distinguish further at the university level (provide data only in one category), and do not have data on the special middle school (presumably it is included in the upper middle school, as it explicitly is in 1982).

65+: age 65 and above (in one group). The education question in all three censuses applies only to persons age 6 and older.

Military personnel are excluded (in the reported education data) throughout. These number 4,238,210 in 1982 (Census 1982, p. 505), 3,199,100 in 1990 (Census 1990, Vol. 4, p. 495), and 2,498,600 in 2000 (Census 2000, Vol. 3, p. 1883).

Availability of data according to completion level refers to availability on population in census and survey books published by the NBS' publishing house.

The Census 1982 and Census 1990 consist of one standard form, the Census 2000 of one long form applicable to an approximately 9.5% sample of the population, and one short form applicable to the rest of the population. The 1982 and 1990 census both contain the standard education question (with more categories in 1990, see the table); the 1990 census also asks for completion levels (4 possible answers). In the 2000 census, the education question, with yet again slightly different categories, is included in both the short form and the long form (in the long form with the additional question regarding 5 completion levels, and for the categories upper middle school through university undergraduate studies with an additional question as to if it is adult education); in both forms, the education question is preceded by a general literacy question ("Knows characters?").

A: This category comes with identical definitions in the 1982 and 1990 censuses, with as cut-off point 1500 characters; the 2000 census uses 1500 characters for the rural population and 2000 characters for the urban population. In 1982, a person age 6-11 cannot “know no or few characters;” in 1992, a person currently in primary school cannot “know no or few characters.” The 2000 census uses the question “Knows characters?” (*shifou shizi*) as a separate question with partial filter character; if the answer is no, the question on education level can only be answered as “no schooling,” “literacy class,” or “primary school.” On the other hand, if a person is currently in primary school, the answer to the “Knows characters?” question is automatically “yes.”

No data are available on the number of persons who know no or few characters, or who know characters or not (except in the *Population Statistical Yearbook 1993* for 1990). What is reported in the 1982 and 1990 data is the number of persons who are illiterate or semi-illiterate; by explicit definition in the 1982 census, persons who know no or few characters and are age 12 or above fall into the category illiterate or semi-illiterate (at age 12 and above the two categories, knowing no or few characters and being illiterate or semi-illiterate, are identical). In 1990, the reported data on illiterate and semi-illiterate persons starts at age 15; these data are identical to those of persons who “know no or few characters” published in an exceptional instance in the *Population Statistical Yearbook 1993*, pp. 54ff. The 2000 books report data on “no schooling” and on “literacy class,” as well as, separately, on illiterate (not mentioning “illiterate”) persons. The number of persons with no schooling, at age 15 and above, is equivalent to 96.9027% of the number of persons who are illiterate at age 15 and above (the data on illiterate persons start only at age 15); the number of persons with no schooling or only literacy class, at age 15 and above, is equivalent to 120.7754% of the number of persons who are illiterate at age 15 and above.

B: In 1992 primary school includes persons who have no schooling but know 1500 or more characters, can read popular books and newspapers, can write brief notes, and have reached the literacy level. The number of persons in the primary school category, compared to 1982 and 2000, may thus be slightly overestimated.

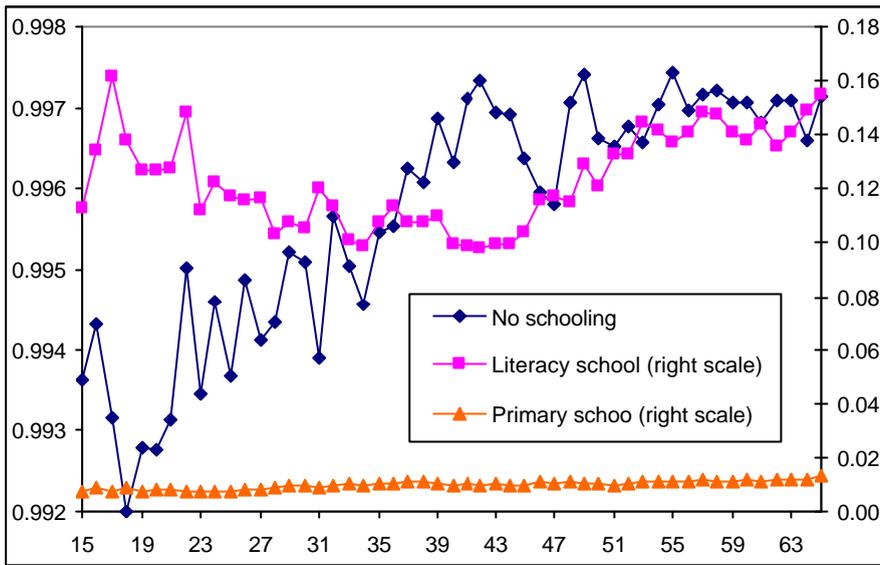
B-E: The completion level categories in 1990 are graduated, currently attending, completed (including incomplete, discontinued studies), and others, with data available according to all categories education categories. (The total, i.e., the sum of these four completion levels, is also published.) The completion level categories in 2000 are graduated, currently attending, completed, not completed (discontinued studies), others (apart from the total), with completion level data collected only for the sample, and then reported in the census books only for primary school and lower middle school.

C-E: The explanations to the 1982, 1992 and the 2000 census state explicitly that distance learning (broadcast and television university, or organized by tertiary level institutions of education) and persons studying at factory/ employee universities or evening universities etc., if these institutions meet the formal (1992: State Education Commission, 2000: Education Ministry) education criteria, can lead to an in the census recognized degree at the corresponding level.

C-D: In 1982 and 2000, the technical schools (*jigong xuexiao*) are equivalent to either lower or upper middle school, depending on the specific technical school. In 1982, special middle schools are included in the upper middle school category. In 1990 and 2000, vocational schools are included in the upper middle school category.

E: In 2000, distance learning, employee universities and evening universities etc. at the university college and university undergraduate studies level is only acknowledged in the census if the person has graduated already; this causes a slight inconsistency in comparison to the 1990 and 1982 census. In 2000, university graduate studies (distance learning seems unlikely in this case) are also only acknowledged in the census if graduation has already occurred.

Sources: *Census 1982*, pp. 605f. (and census form between pp. 589 and 590); *Census 1990*, Vol. 4, pp. 514f. (and census form between pp. 520 and 521); *Census 2000*, Vol. 3, pp. 1895f. and 1898f. (and census forms between pp. 1892 and 1893).

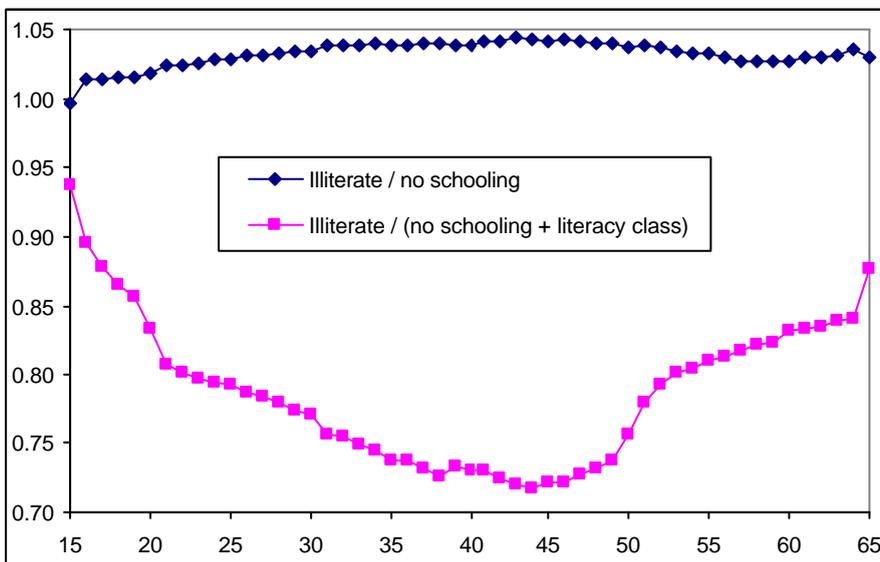


The series “no schooling” reflects, at each age, the number of persons classified in the table on illiterate and semi-illiterate persons as having no schooling, divided by the number of persons classified in the population-wide table on different education levels as having no schooling; similarly for the other two categories.

Age 65 refers to “age 65 and above.”

Sources: *Survey 1995*, pp. 26-33, 82-5.

**Figure 10. Persons in Illiteracy Categories Relative to Persons in Corresponding Education Categories, 1995 1% Sample Survey**

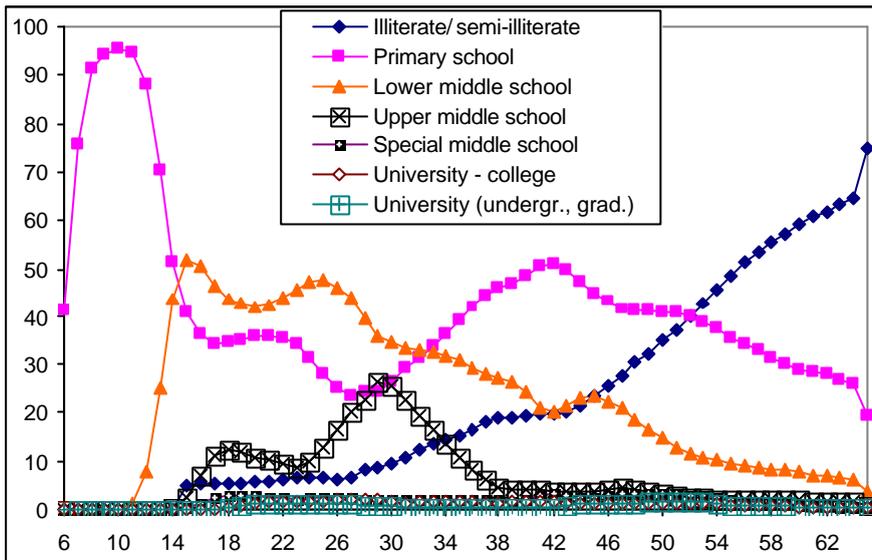


“No schooling” and “literacy class” are part of the exhaustive education categories (covering the total population). The category “illiterate” (persons age 15 and above) is not part of the education categories.

Age 65 refers to “age 65 and above.”

Sources: *Census 2000*, Vol. 1, pp. 593-602, 633f.

**Figure 11. Category Illiterate Vs. Education Levels, 2000 Census**



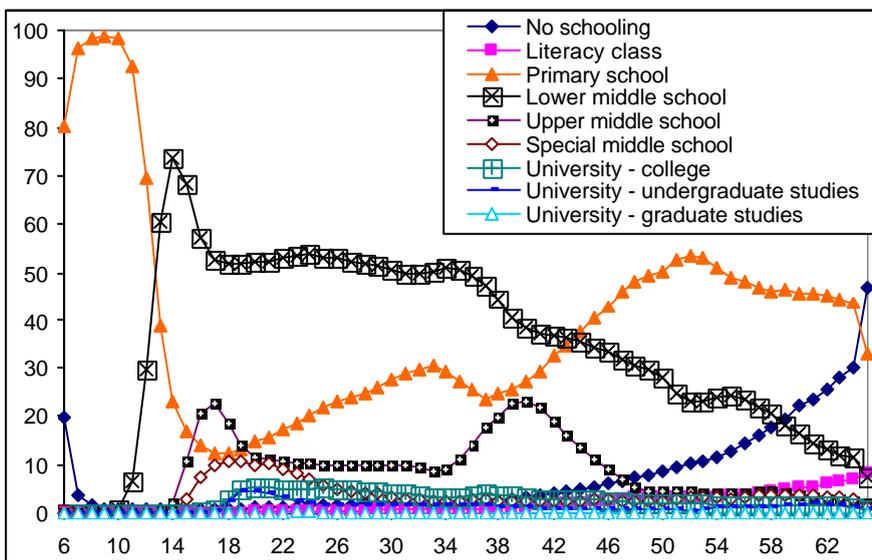
“Education level” here includes “illiterate/ semi-illiterate.”

Starting at age 15, at each specific age the shares of the different education levels add up to 100%, below age 15 they do not as some persons enter the census without any education level (only persons age 15 and above can be “illiterate/ semi-illiterate”).

Age 65 refers to “age 65 and above.”

Sources: *Census 1990*, Vol. 2, pp. 2-5 (total population), pp. 278-81 (total population at age 15 and above, illiterate/ semi-illiterate persons), and pp. 112-51 (education level, excluding illiterate/ semi-illiterate persons).

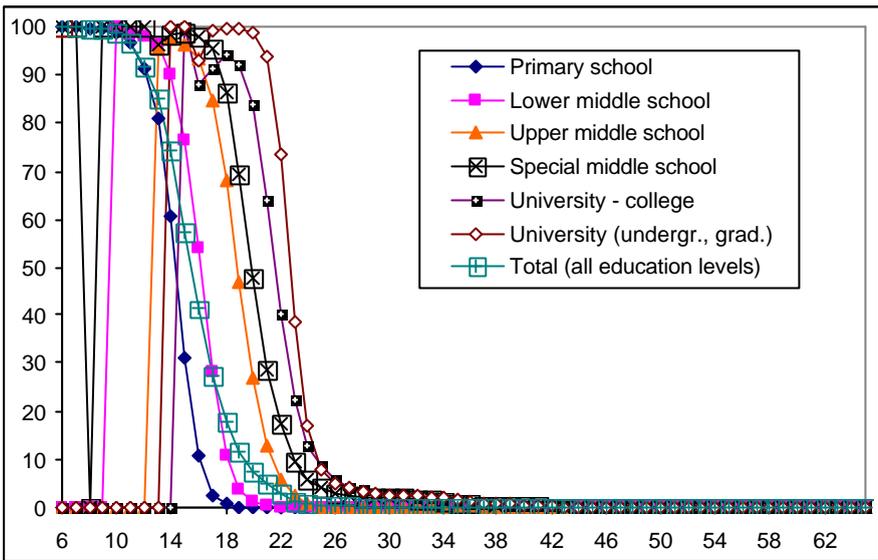
**Figure 12. Share of Education Level (in %), Population at Each Age, 1990 Census**



Data are available from age 6 to “65 and above” (“65” in the chart). At each specific age the shares of the different education levels add up to 100%.

Sources: *Census 2000*, Vol. 1, pp. 593-602 (total population).

**Figure 13. Share of Education Level (in %), Population at Each Age, 2000 Census**

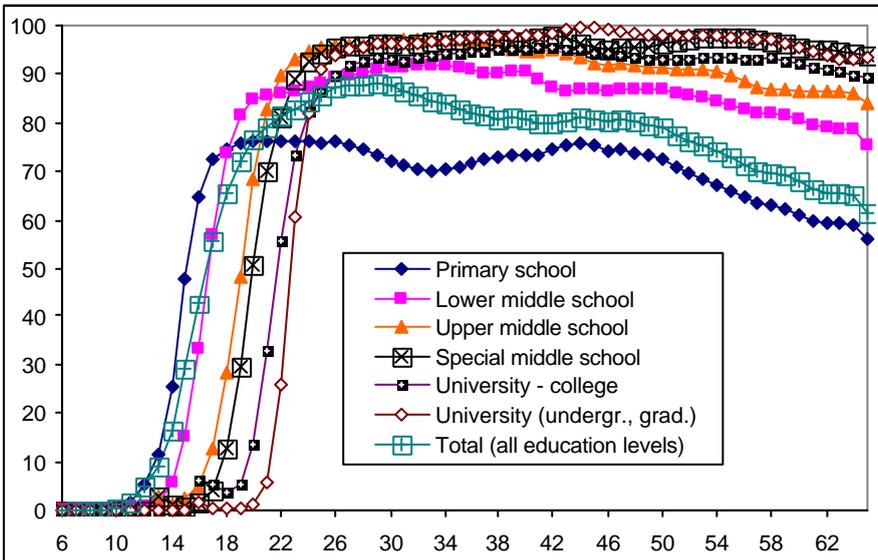


The figure shows the share (in %) of persons, at each age and education level stated in the census, which at the time of the census was attending school to complete that particular education level. For example, approximately 10% of those who were age 18 and stated lower middle school as their highest education level were in 1990 attending lower middle school.

Age 65 refers to “age 65 and above.”

Sources: see Figure 12.

**Figure 14. Share of Age-Specific Population Attending School at Time of 1990 Census (in %), By Education Level**

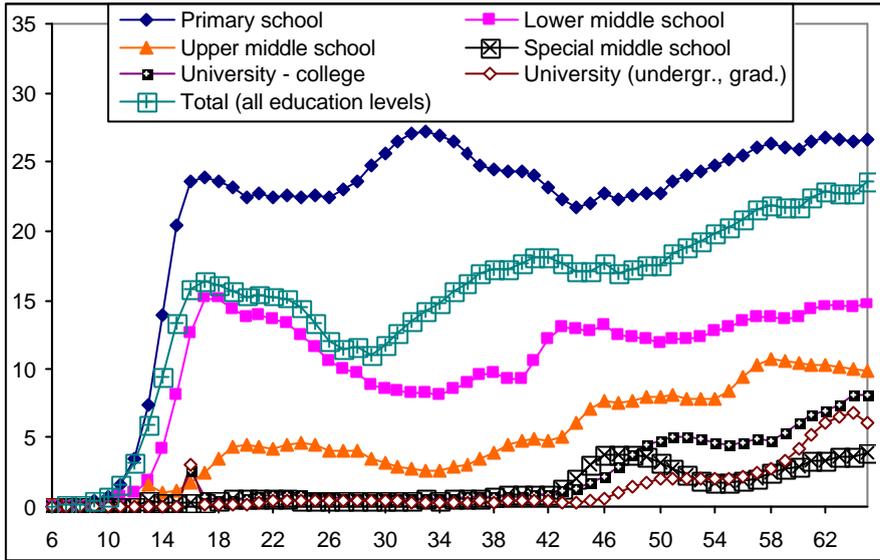


The figure shows the share (in %) of persons, at each age and education level stated in the census, which at the time of the census had graduated at that particular education level. For example, approximately 25% of those who were age 22 and stated university (undergr., grad.) as their highest education level in 1990 had already graduated from university.

Age 65 refers to “age 65 and above.”

Sources: see Figure 12.

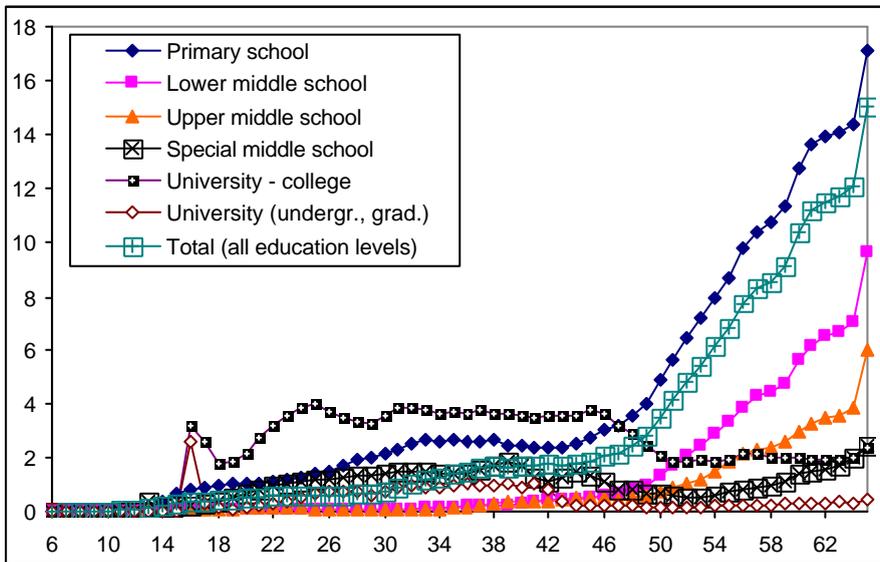
**Figure 15. Share of Age-Specific Population Graduated at Time of 1990 Census (in %), By Education Level**



The figure shows the share (in %) of persons, at each age and education level stated in the census, which at the time of the census had completed their studies at that particular education level but had not graduated, or had discontinued their studies. For example, approximately 10% of those who were age 26 and stated lower middle school as their highest education level in 1990 had completed or discontinued their studies. Age 65 refers to “age 65 and above.”

Sources: see Figure 12.

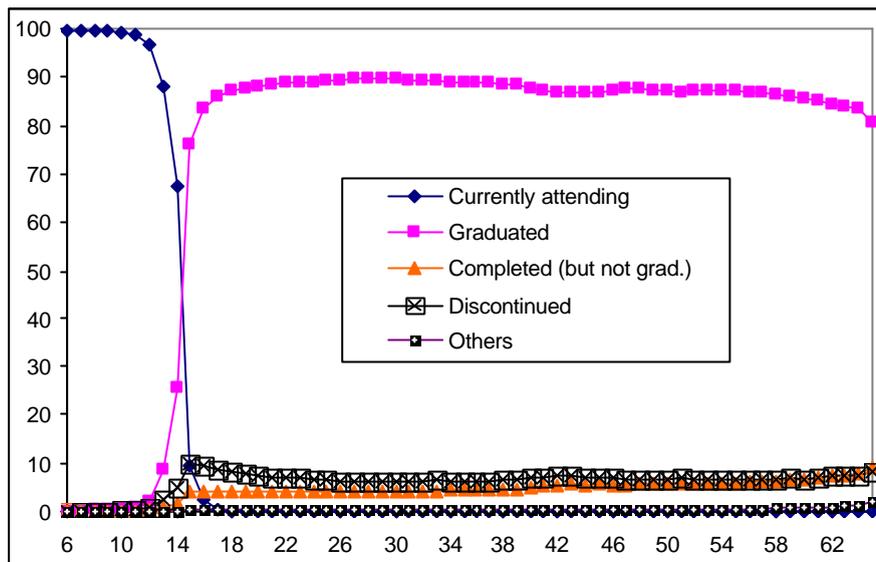
**Figure 16. Share of Age-Specific Population Which Had Completed (But Not Graduated) Or Discontinued Studies at Time of 1990 Census (in %), By Education Level**



The figure shows the share (in %) of persons, at each age and education level stated in the census, which at the time of the census were not attending school, had not yet graduated, and were not in the category “completed studies” (including discontinued studies), but still stated the particular education level as their highest education level at the census. These persons constitute the official, explicit category “others.” For example, approximately 4% of those who were age 38 and stated university-college as their highest education level in 1990 were neither attending school, nor had they graduated, completed or discontinued their studies. Age 65 refers to “age 65 and above.”

Sources: see Figure 12.

**Figure 17. Residual Share of Age-Specific Population at Time of 1990 Census (in %), By Education Level**

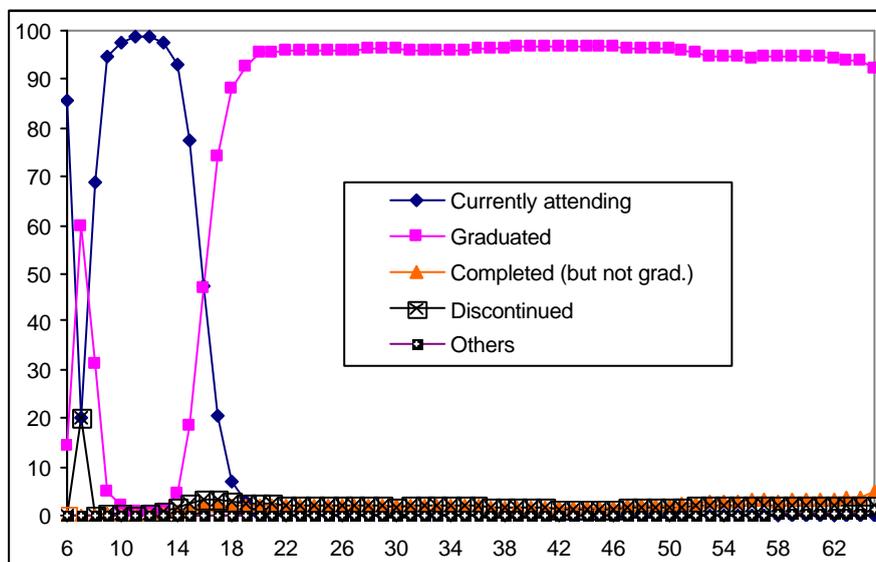


Age 65 refers to “age 65 and above.”

Data are from the census long form, i.e., cover 9.501550% of China’s population.

Sources: *Census 2000*, Vol. 2, pp. 842f., 846f.

**Figure 18. Age-Specific Completion Levels of Persons Who Stated Primary School As Their Highest Education Level at Time of 2000 Census (in %)**



Data in the age bracket 6-10, except for “currently attending” starting age 9, reflect an extremely small number of observations, from single-digit numbers rising to a maximum in the age bracket 6-10, (apart from “currently attending” age 9 and above) of 375 graduated at age 10.

Age 65 refers to “age 65 and above.”

Data are from the census long form, i.e., cover 9.501550% of China’s population.

Sources: *Census 2000*, Vol. 2, pp. 844f., 848f.

**Figure 19. Age-Specific Completion Levels of Persons Who Stated Lower Middle School As Their Highest Education Level at Time of 2000 Census (in %)**

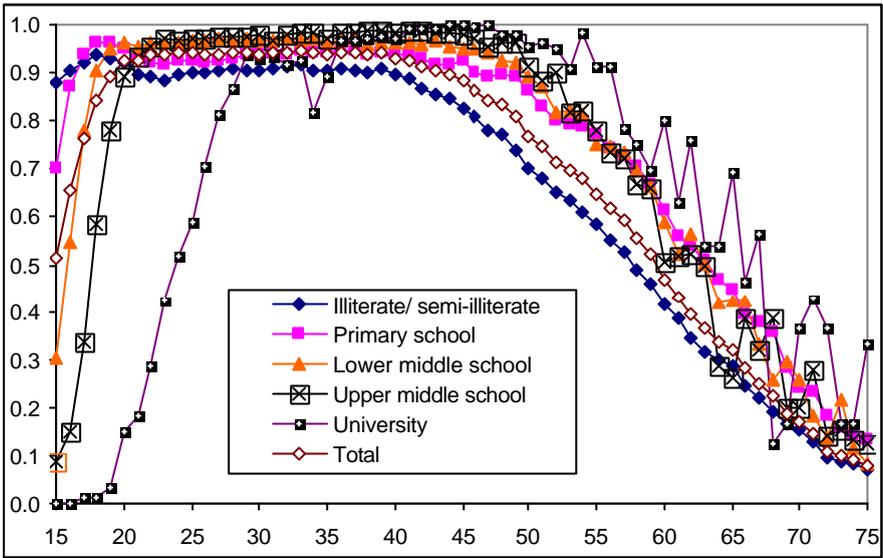
**Table 4. Census 1990 and 2000 Graduation Numbers and Graduation Rates**

Age	Primary school Grad. %	Lower middle school Grad. %	Upper middle school Grad. %	Special middle school Grad. %	University -- college Grad. %	University -- undergr./ grad. Grad. %
<b>Census 1990</b>						
13	1550691 11.40					
14	2659275 25.30					
15	4216177 47.99					
16	5496861 64.80	3914610 33.04				
17	6115730 72.54	6481088 56.78				
18	6303962 74.78	7923975 73.82	867013 28.31			
19	6888564 75.66	9122542 81.70	1494807 48.33	228513 29.46		
20		9325044 84.93	1946322 68.62	366781 50.70	49264 13.34	
21		9916358 85.79	2270407 82.84	471953 69.98	152529 32.62	20578 5.65
22		9301326 86.25	2045679 89.65	453692 81.26	233780 55.82	84823 25.77
23			1844363 93.05	447930 89.05	261360 73.32	162738 60.53
24			2328502 94.57	532473 92.82	334618 82.42	246939 82.05
25				572155 94.22	373680 86.52	269160 91.04
26				634276 95.37	476900 89.84	285082 93.84
27					525010 91.80	271674 94.88
28					282969 92.54	139423 95.42
29						100309 96.11
<b>Graduation data 1990 (Statistical Yearbook)</b>						
	18.631m	<sup>a</sup> 11.091m	<sup>b</sup> 2.330m	<sup>c</sup> 661000	<sup>d</sup> 649440	incl. in left
<i>Ratio graduation data to census data (at italics-age)</i>						
	3.05	1.22	1.03	1.48	1.01	
<b>Census 2000</b>						
13	818834 8.52					
14	1264535 25.68					
15	1760112 76.26	2727714 18.34				
16	1900227 83.47	5504489 47.19				
17	1908666 86.20	7597121 74.11				
18	2287538 87.14	10145971 88.01				
19	2096009 87.64	8896838 92.43				
20		8850600 95.49				
21		9230419 95.68				
<b>Graduation data 2000 (Statistical Yearbook)</b>						
	24.192m	<sup>a</sup> 16.071m				
<i>Ratio graduation data to census data (at italics-age)</i>						
	12.73	1.58				

“Grad.” refers to the absolute number of students at the age-specific level who reported themselves to have graduated at this education level (as their highest education level). “%” refers to the share of persons who had *graduated* at the particular education level which they list as their highest level

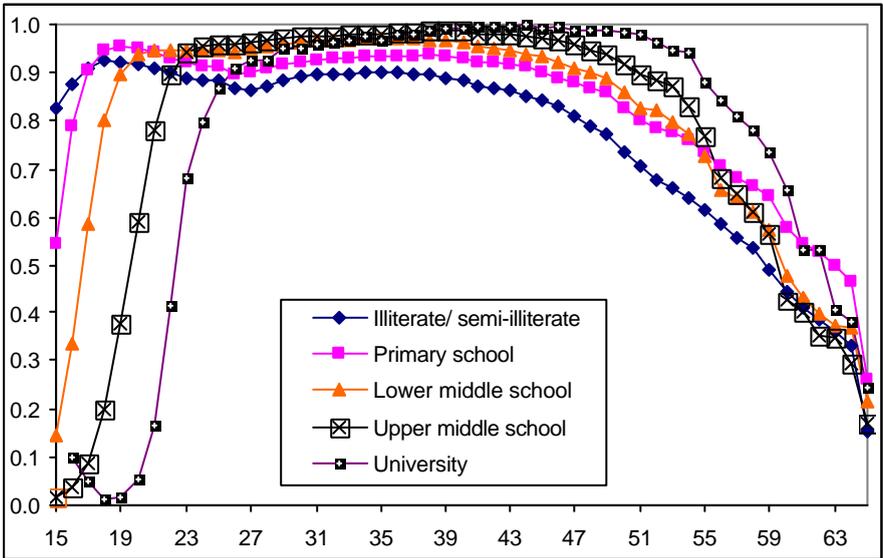
- a Regular lower middle school. “Regular” refers to the state education system. It excludes vocational schools and presumably also adult education.
- b Regular upper middle school. It excludes vocational schools, special schools, and presumably also adult education.
- c Special middle school. The excluded vocational schools, which provide an education equivalent to either lower or upper middle school, in 1990 had a further 893,000 graduates.
- d Regular institutions of higher education (0.614m), plus graduates of research programs (“postgraduates,” 35440).

Sources: *Census 1990*, Vol. 2, pp. 112-151; *Census 2000*, Vol. 1, p. 215, and Vol. 2, pp. 800, 842-9 (including age-specific ratios of census to long-form population to augment long-form education data); graduation data: *Population Statistical Yearbook 2002*, pp. 220f.; *Statistical Yearbook 2004*, pp. 780f.



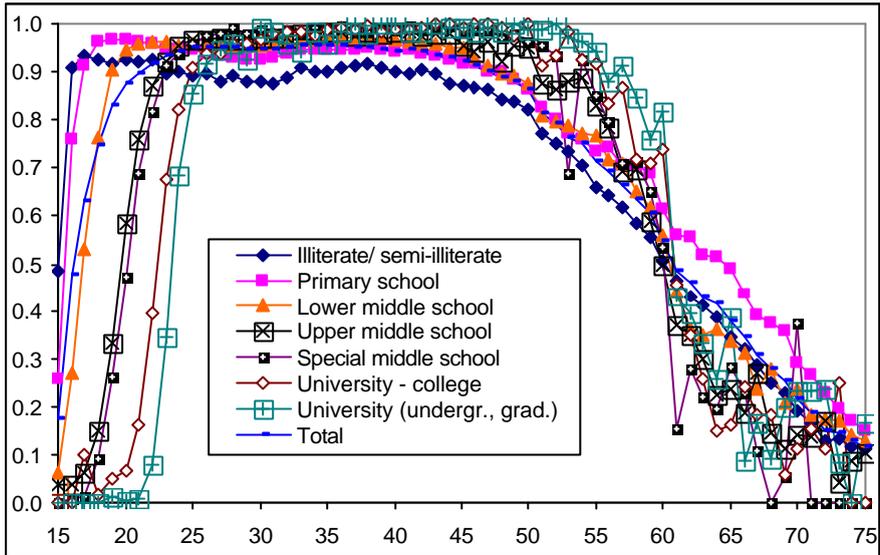
Each data point is the ratio of laborers at this age and education level to total persons (population) in the sample at this age and education level.  
Sources: 1‰ sample.

**Figure 20. Age- and Education Specific Employment Rates (1‰ Sample Data), 1982**



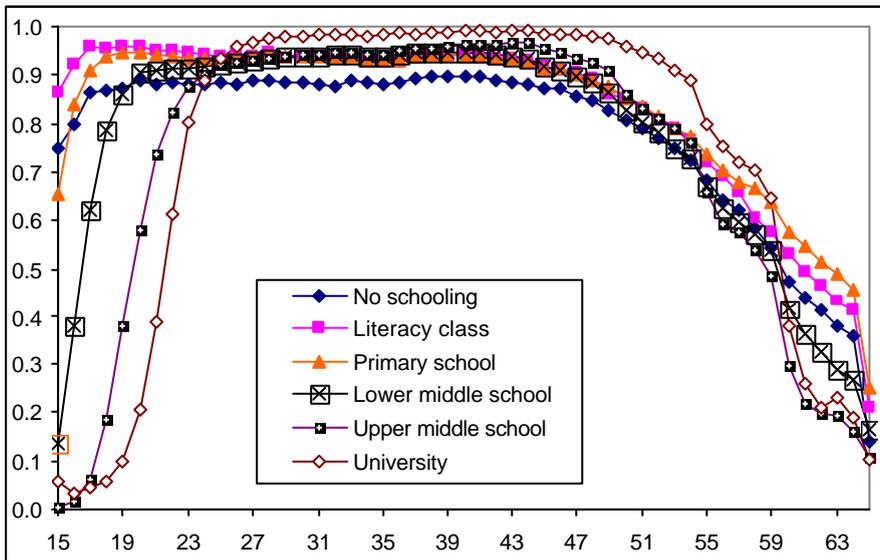
Each data point is the ratio of laborers at this age and education level to total persons (population) in the sample at this age and education level.  
Age 65 refers to “age 65 and above.”  
Sources: *Survey 1987*, pp. 164f., 172, 230-3, 468f.

**Figure 21. Age- and Education Specific Employment Rates (1‰ Sample), 1987**



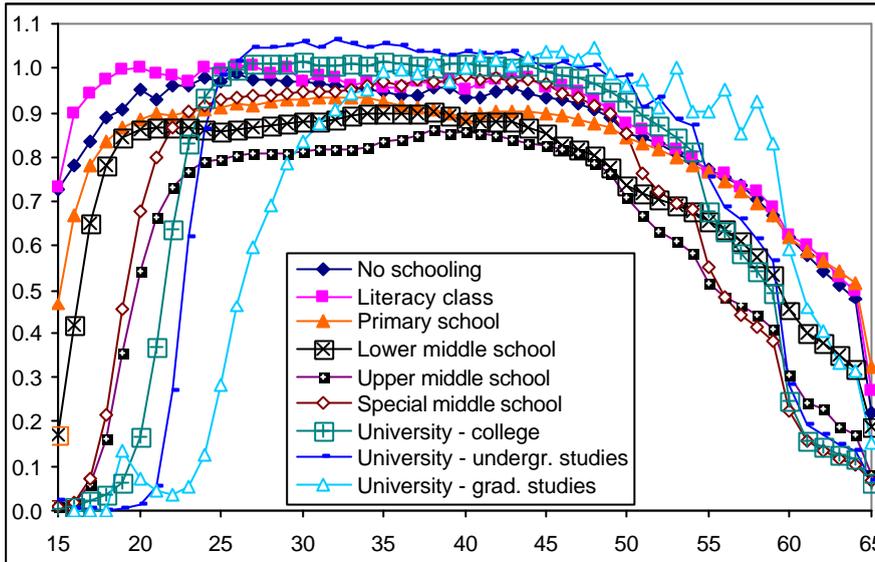
Each data point is the ratio of laborers at this age and education level to total persons (population) in the sample at this age and education level.  
Sources: 1‰ sample.

**Figure 22. Age- and Education Specific Employment Rates (1‰ Sample Data), 1990**



Each data point is the ratio of laborers at this education level, given the specific age, to total persons (population) at this education level (at the same specific age). Survey day is 1 November.  
Age 65 refers to “age 65 and above.”  
Sources: *Survey 1995*, pp. 26-33, 124-83.

**Figure 23. Age- and Education Specific Employment Rates (1‰ Sample), 1995**

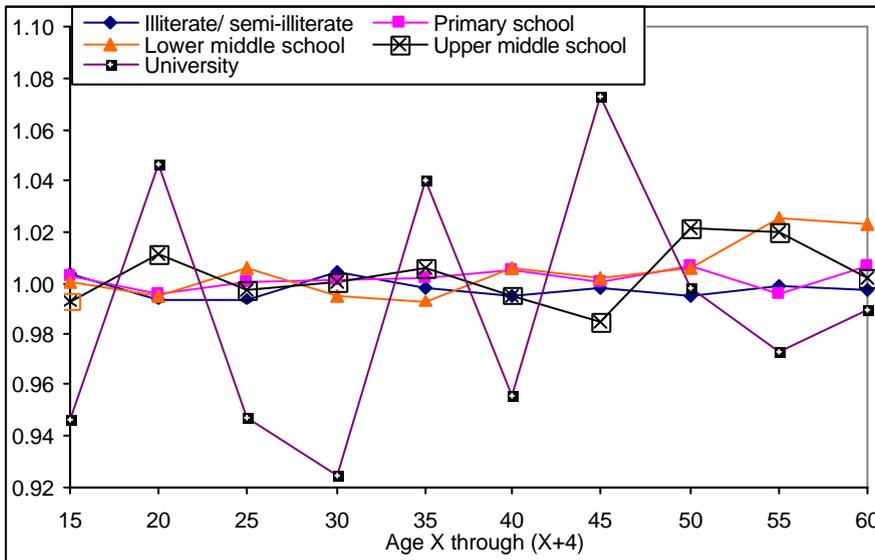


Each data point is the ratio of laborers at this age and education level to total persons (population) in the sample at this age and education level. Data on laborers are from the long form sample, and therefore augmented by age-specific sampling rates in order to be comparable with the population data. Census day is 1 November.

Age 65 refers to “age 65 and above.”

Sources: *Census 2000*, Vol. 1, pp. 593-602; Vol. 2, pp. 1249-68.

**Figure 24. Age- and Education Specific Employment Rates (Long-form Sample Data), 2000**

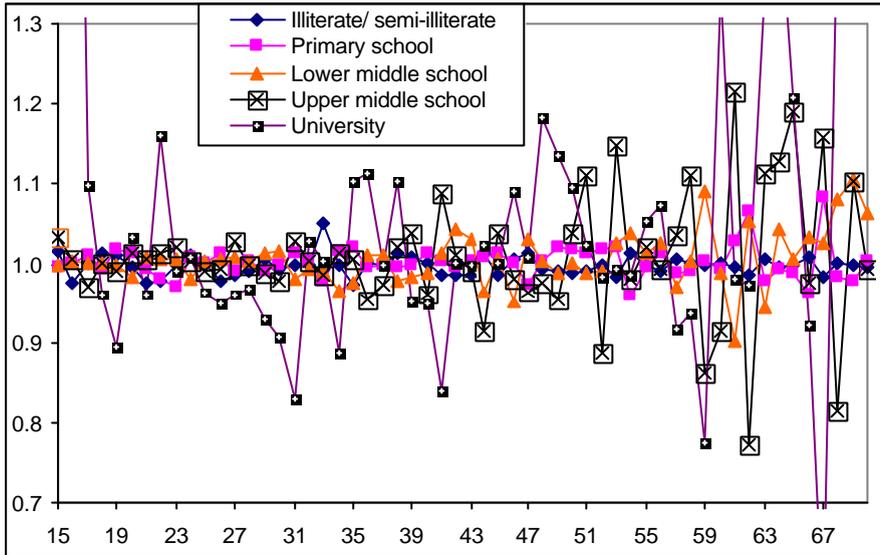


The data points are the ratios of census values in the census publication to 1% sample values. The individual values are, for each age group, the share of persons (population) at the particular education level.

Data in the census publication are only available in 5-year age groups. Age 60 refers to “age 60 and above.”

Sources: 1% sample, *Census 1982*, pp. 360f.

**Figure 25. Census Vs. 1% Sample Population Data, 1982**



The data points are the ratios of census 1% sample values in the census publication to 1% sample values. The individual values are, for each age group, the share of persons (population) at the particular education level.

Age 70 refers to “age 70 and above.”

Sources: 1% sample, *Population Statistical Yearbook 1988*, pp. 820-42.

**Figure 26. Census 1% Sample Vs. 1% Sample Population Data, 1982**

**Table 5. 1982 1% Sample Data**

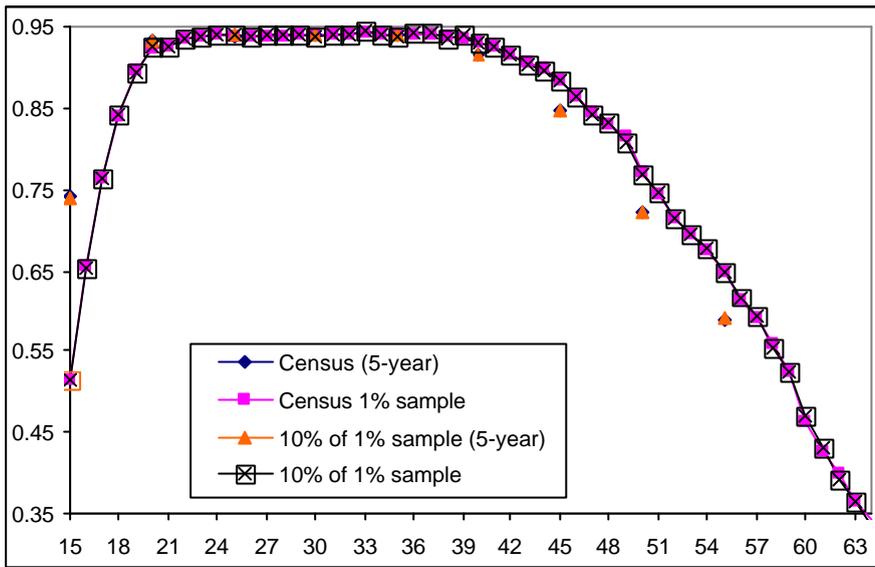
Age	Popu- lation	Illit.	Number of laborers				Share of laborers in population				
			Prim.	Low. m.	Upp. m.	Univ.	Illit.	Prim.	Low. m.	Upp. m.	Univ.
15	22869	2004	6432	3297	51	0	0.088	0.281	0.144	0.002	0.000
16	25701	2283	7285	6937	326	0	0.089	0.283	0.270	0.013	0.000
17	24484	2007	6171	9190	1313	1	0.082	0.252	0.375	0.054	0.000
18	24886	2043	5500	10304	3128	3	0.082	0.221	0.414	0.126	0.000
19	27737	2413	5352	11390	5628	11	0.087	0.193	0.411	0.203	0.000
20	15412	1700	2909	5775	3862	26	0.110	0.189	0.375	0.251	0.002
21	10722	1157	2011	3427	3301	26	0.108	0.188	0.320	0.308	0.002
22	14406	1713	3010	4445	4292	31	0.119	0.209	0.309	0.298	0.002
23	14166	1964	3314	4262	3690	39	0.139	0.234	0.301	0.260	0.003
24	19704	3098	4758	6142	4466	66	0.157	0.241	0.312	0.227	0.003
25	18802	3303	4889	5713	3692	89	0.176	0.260	0.304	0.196	0.005
26	17653	3353	4913	5322	2864	112	0.190	0.278	0.301	0.162	0.006
27	19514	3992	6114	5712	2366	140	0.205	0.313	0.293	0.121	0.007
28	18655	4065	6316	5235	1779	138	0.218	0.339	0.281	0.095	0.007
29	17382	4048	6399	4532	1240	135	0.233	0.368	0.261	0.071	0.008
30	17555	4370	6679	4365	929	133	0.249	0.380	0.249	0.053	0.008
31	14700	3555	5741	3755	641	125	0.242	0.391	0.255	0.044	0.009
32	15350	3613	6434	3555	727	113	0.235	0.419	0.232	0.047	0.007
33	13118	2970	5982	2539	805	100	0.226	0.456	0.194	0.061	0.008
34	12533	2865	5675	2365	785	93	0.229	0.453	0.189	0.063	0.007
35	12280	2876	5423	2414	698	106	0.234	0.442	0.197	0.057	0.009
36	11266	2667	4836	2380	594	153	0.237	0.429	0.211	0.053	0.014
37	10396	2604	4154	2313	562	155	0.250	0.400	0.222	0.054	0.015
38	10125	2711	3847	2202	581	133	0.268	0.380	0.217	0.057	0.013
39	9699	2789	3568	1960	645	155	0.288	0.368	0.202	0.067	0.016
40	10064	3154	3568	1774	695	179	0.313	0.355	0.176	0.069	0.018
41	9849	3219	3563	1521	588	236	0.327	0.362	0.154	0.060	0.024
42	9440	3169	3421	1244	578	222	0.336	0.362	0.132	0.061	0.024
43	9363	3293	3323	1098	515	231	0.352	0.355	0.117	0.055	0.025
44	9692	3532	3305	1091	538	229	0.364	0.341	0.113	0.056	0.024
45	9694	3767	3218	965	420	206	0.389	0.332	0.100	0.043	0.021
46	9585	3795	3023	933	377	162	0.396	0.315	0.097	0.039	0.017
47	9311	3775	2807	799	322	148	0.405	0.301	0.086	0.035	0.016
48	9289	3991	2601	779	273	97	0.430	0.280	0.084	0.029	0.010
49	9356	3995	2454	756	255	93	0.427	0.262	0.081	0.027	0.010
50	8490	3594	2033	628	189	82	0.423	0.239	0.074	0.022	0.010
51	8180	3391	1879	592	165	79	0.415	0.230	0.072	0.020	0.010
52	8245	3318	1755	541	203	76	0.402	0.213	0.066	0.025	0.009
53	7541	3068	1554	442	120	61	0.407	0.206	0.059	0.016	0.008
54	7915	3039	1691	434	142	67	0.384	0.214	0.055	0.018	0.008
55	6985	2665	1385	342	100	41	0.382	0.198	0.049	0.014	0.006
56	6876	2558	1240	304	95	42	0.372	0.180	0.044	0.014	0.006
57	7162	2552	1258	312	83	36	0.356	0.176	0.044	0.012	0.005
58	6505	2220	1045	245	64	30	0.341	0.161	0.038	0.010	0.005
59	6054	1970	910	187	70	30	0.325	0.150	0.031	0.012	0.005
60	6117	1818	807	182	48	20	0.297	0.132	0.030	0.008	0.003
61	5853	1651	671	152	31	17	0.282	0.115	0.026	0.005	0.003
62	5540	1434	556	121	44	19	0.259	0.100	0.022	0.008	0.003

63	5002	1185	504	101	22	7	0.237	0.101	0.020	0.004	0.001
64	4794	1096	427	69	12	7	0.229	0.089	0.014	0.003	0.001
65+	49595	5643	1887	297	81	49	0.114	0.038	0.006	0.002	0.001

“Share of laborers in population” refers to, at each age, the number of laborers at a particular education level divided by the population number (at the same age, independent of education level). The age group 65+ covers persons between the age of 65 and 99.

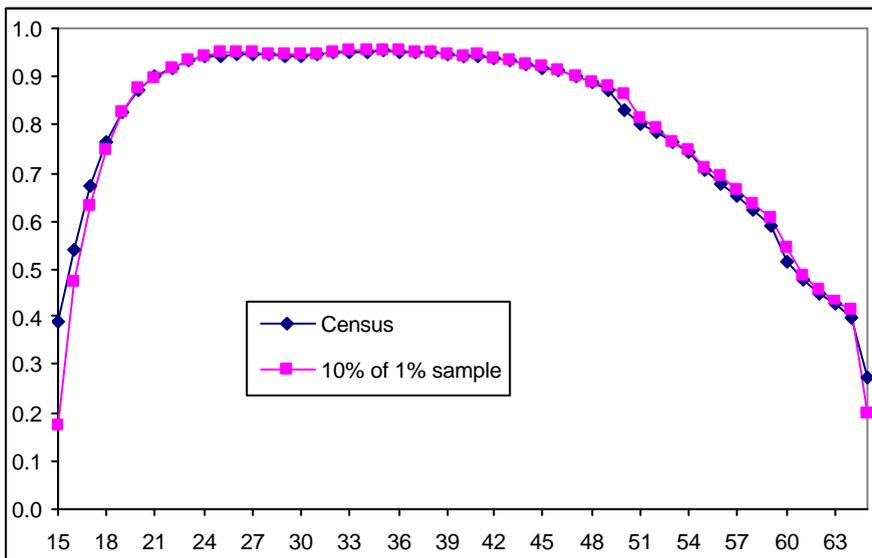
“Illit.” refers to “illiterate/ semi-illiterate,” “Prim.” to “primary school,” “Low. m.” to “lower middle school,” “Upp. m.” to “upper middle school,” and “Univ.” to “university.”

Sources: The 1‰ sample is a 10% random sample drawn from the official NBS 1982 1% sample.



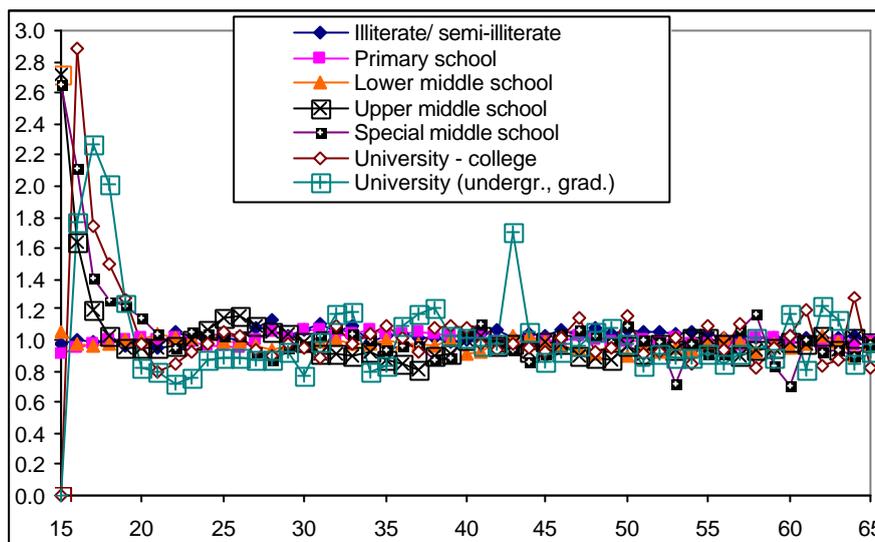
Each data point is the ratio of laborers to population at the specific age.  
 Sources: 1% sample, *Census 1982*, pp. 273-81, 468f.; *Population Statistical Yearbook 1988*, pp. 742-4, 820, 822.

**Figure 27. Age-Specific Employment Rates 1982, Census Vs. 1% Sample**



Each data point is the ratio of laborers to population at the specific age.  
 Sources: 1% sample, *Census 1990*, Vol. 2, pp. 112-51, 278-81.

**Figure 28. Age-Specific Employment Rates 1990, Census Vs. 1% Sample**



The data points are the ratios of census values in the census publication to 1‰ sample values. The individual values are, for each age group, the share of persons (population) at the particular education level.

Data in the census publication are only available in 5-year age groups. Age 65 refers to “age 65 and above.”

Sources: 1‰ sample, *Census 1990*, pp. 112-51, 278-81.

**Figure 29. Census Vs. 1‰ Sample Population Data, 1990**

**Table 6. 1990 1% Sample Data**

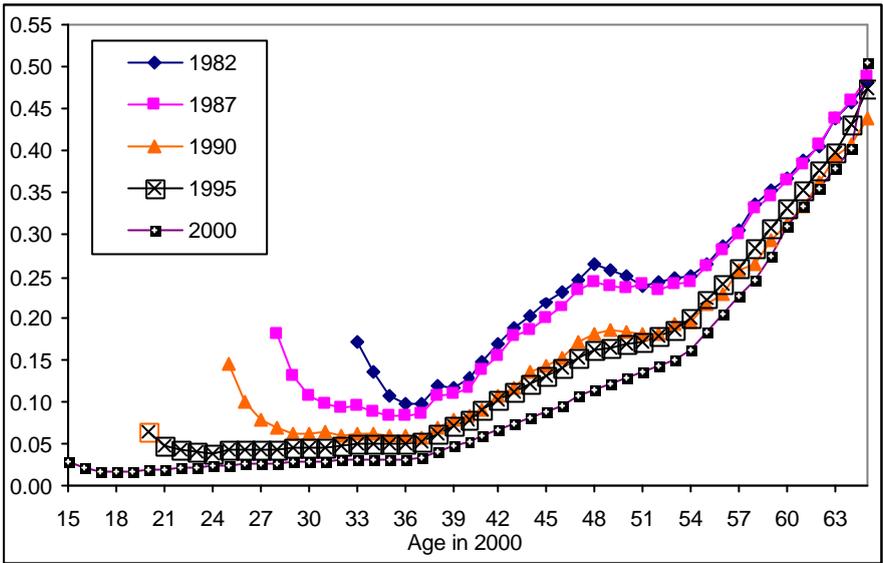
Age	Popu- lation	Illit.	Prim.	Number of laborers				Laborers / population (by age and educ.)							
				Low. m.	Upp. m.	Spec. m.	Univ.- coll.	Univ.	Illit.	Prim.	Low. m.	Upp. m.	Spec. m.	Univ. -coll.	
15	22196	550	2579	697	7	0	0	0	0.487	0.259	0.064	0.038	0.000		
16	23765	1116	6846	3311	37	1	0	0	0.908	0.759	0.268	0.037	0.006	0.000	0.000
17	24939	1240	7991	6366	142	6	1	0	0.932	0.911	0.529	0.061	0.013	0.100	0.000
18	25787	1311	8580	8868	471	56	1	0	0.927	0.963	0.765	0.151	0.089	0.017	0.000
19	26412	1366	8984	10236	1102	170	9	2	0.916	0.967	0.904	0.333	0.262	0.049	0.012
20	27465	1452	9363	11006	1839	318	26	2	0.924	0.968	0.948	0.583	0.473	0.065	0.005
21	26478	1508	9157	10414	2138	434	94	3	0.921	0.961	0.960	0.758	0.688	0.163	0.007
22	29374	1602	9844	12197	2478	577	233	43	0.926	0.959	0.962	0.872	0.816	0.392	0.078
23	23481	1355	7704	10131	1884	448	270	127	0.896	0.947	0.962	0.920	0.912	0.675	0.345
24	26317	1502	8060	11804	2298	543	358	245	0.902	0.946	0.956	0.956	0.936	0.821	0.684
25	25766	1474	7128	11817	2745	595	394	299	0.893	0.946	0.958	0.966	0.958	0.910	0.857
26	25436	1455	6300	11397	3539	635	491	317	0.894	0.939	0.956	0.966	0.977	0.939	0.919
27	29502	1604	6473	12845	5323	777	622	337	0.881	0.934	0.955	0.972	0.979	0.938	0.947
28	23769	1559	5136	9561	4960	591	490	241	0.893	0.928	0.956	0.969	0.990	0.957	0.960
29	12093	885	2668	4289	2971	272	231	110	0.880	0.926	0.953	0.968	0.978	0.959	0.924
30	14250	1107	3273	4854	3609	291	226	119	0.881	0.925	0.956	0.969	0.986	0.966	0.992
31	14131	1218	3585	4584	3431	257	219	66	0.878	0.928	0.958	0.969	0.973	0.978	0.985
32	17708	1804	5058	5667	3690	321	232	67	0.886	0.938	0.968	0.973	0.988	0.983	0.957
33	20122	2248	6115	6494	3594	395	313	80	0.909	0.945	0.968	0.980	0.978	0.984	0.941
34	18787	2442	6032	5914	2721	412	283	111	0.901	0.947	0.967	0.980	0.993	0.976	0.982
35	19554	2637	7016	5829	2354	463	277	91	0.902	0.946	0.975	0.983	0.994	0.986	0.958
36	20068	2898	7637	5888	1899	454	300	65	0.911	0.947	0.971	0.981	0.996	0.990	0.985
37	18343	2983	7389	5059	1317	375	275	49	0.914	0.947	0.970	0.973	0.984	0.989	1.000
38	18853	3224	7954	5145	941	402	240	45	0.916	0.951	0.968	0.983	0.983	0.996	0.978
39	15536	2719	6747	4011	670	321	213	49	0.908	0.945	0.972	0.981	0.991	0.982	1.000
40	15170	2617	6638	3903	585	298	219	53	0.902	0.940	0.966	0.983	0.987	0.986	1.000
41	15654	2655	7392	3435	599	410	236	58	0.896	0.947	0.962	0.982	0.998	0.987	0.983
42	13075	2218	6321	2557	519	421	200	54	0.906	0.937	0.959	0.974	0.988	0.990	0.982
43	13299	2376	6295	2628	490	416	180	55	0.895	0.933	0.959	0.978	0.990	1.000	1.000
44	12091	2179	5354	2582	455	359	155	132	0.876	0.927	0.951	0.981	0.994	1.000	0.992
45	11060	2187	4530	2484	390	301	132	152	0.872	0.918	0.949	0.954	0.984	0.992	0.993
46	10831	2271	4320	2331	420	290	115	132	0.866	0.917	0.930	0.944	0.986	1.000	0.985
47	9676	2261	3530	1950	456	302	96	120	0.865	0.899	0.911	0.962	0.987	1.000	0.992
48	9875	2336	3624	1825	435	322	121	119	0.842	0.901	0.898	0.922	0.973	0.992	0.983
49	10171	2644	3718	1550	438	349	124	140	0.840	0.884	0.894	0.956	0.980	0.976	1.000
50	10002	2702	3579	1423	341	296	108	170	0.822	0.864	0.875	0.950	0.964	1.000	0.988
51	9060	2472	3059	1003	258	283	116	186	0.771	0.825	0.809	0.878	0.956	0.913	0.989
52	9884	2832	3185	996	230	289	123	178	0.752	0.800	0.797	0.861	0.932	0.932	1.000
53	9808	2951	2908	875	216	269	111	158	0.733	0.773	0.788	0.878	0.690	0.982	0.969
54	9603	2936	2809	802	182	227	114	130	0.706	0.760	0.773	0.888	0.927	0.927	0.963
55	9337	2912	2514	707	158	184	78	100	0.661	0.736	0.769	0.832	0.852	0.918	0.943
56	9113	2963	2313	601	146	140	71	81	0.642	0.741	0.720	0.785	0.795	0.835	0.880
57	9497	3044	2289	604	143	91	58	72	0.618	0.705	0.711	0.694	0.711	0.866	0.911
58	8359	2697	1832	497	122	62	53	50	0.585	0.707	0.651	0.697	0.713	0.716	0.847
59	7772	2461	1604	404	96	63	39	44	0.557	0.689	0.617	0.589	0.649	0.709	0.759
60	8504	2511	1537	383	83	56	37	36	0.507	0.613	0.561	0.497	0.533	0.740	0.818
61	7200	2020	1154	234	50	8	15	22	0.466	0.559	0.442	0.368	0.154	0.455	0.431
62	7709	2031	1237	187	47	15	16	13	0.432	0.557	0.361	0.348	0.278	0.348	0.394

63	6749	1744	959	156	37	9	9	10	0.412	0.520	0.351	0.298	0.220	0.257	0.333
64	6321	1531	884	158	23	7	3	9	0.385	0.516	0.360	0.225	0.194	0.150	0.257
65+	68964	8876	4021	583	107	24	26	36	0.172	0.298	0.218	0.156	0.134	0.138	0.164

“Laborers / population (by age and educ.)” refers to, at each age and education level, the number of laborers divided by the population number. The age group 65+ covers persons between the age of 65 and 99.

“Illit.” refers to “illiterate/ semi-illiterate,” “Prim.” to “primary school,” “Low. m.” to “lower middle school,” “Upp. m.” to “upper middle school,” “Spec. m.” to “special middle school,” “Univ.-coll.” to “University – college,” and “Univ.” to “University (undergraduates and graduates).”

Sources: The 1‰ sample is a 10% random sample drawn from the official NBS 1990 1% sample.

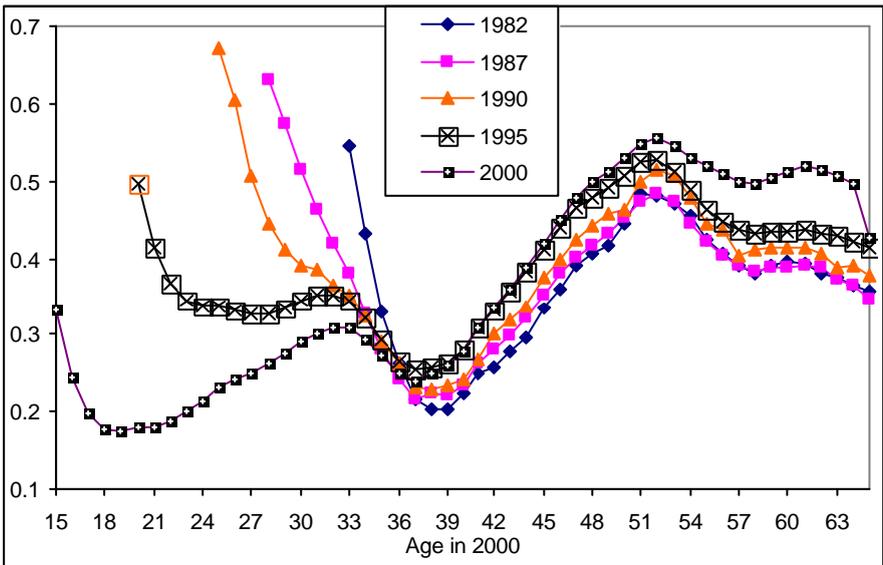


At a given age, a data point shows the share of illiterate and semi-illiterate laborers in all laborers that age.

1982 and 1990 data are from the 1% sample surveys, 2000 data from the census long-form sample, and 1987 and 1995 data from the 1% sample surveys. All data are midyear data (1995 and 2000 data are adjusted to midyear values). “Illiterate and semi-illiterate” laborers in 1995 and 2000 are laborers with no schooling or literacy school only. The 5 data points at a particular age level, for example age 42 in 2000, cover the shares of those age 24 in 1982, those age 29 in 1987, those age 32 in 1990, those age 37 in 1995, and those age 42 in 2000.

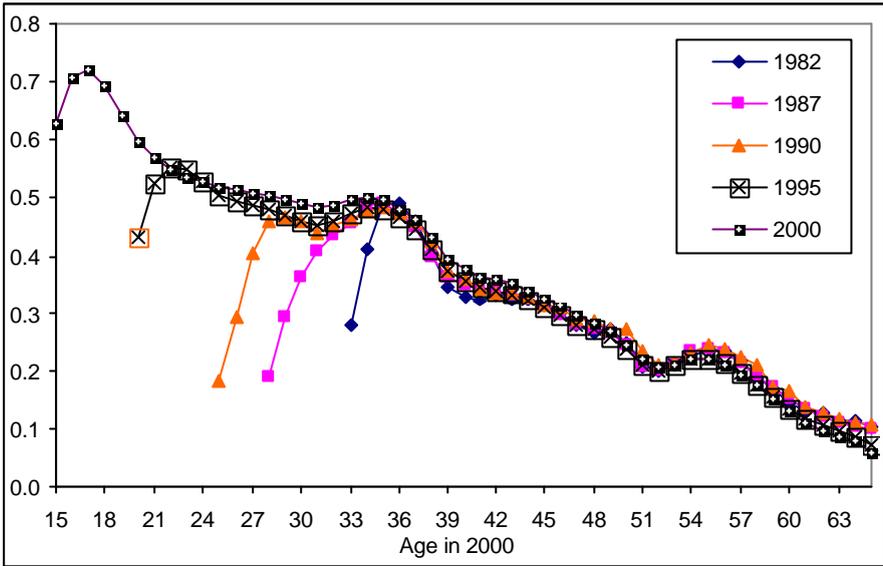
Sources: 1% samples of 1982 and 1990; for 1987, 1995, and 2000 see Table 1.

**Figure 30. Age-Specific Shares of Illiterate/ Semi-Illiterate Laborers**



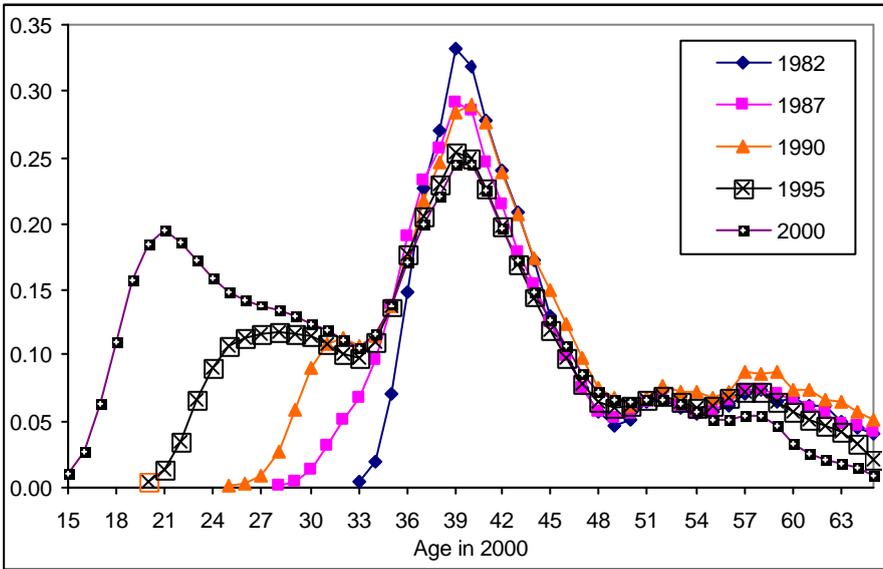
For notes and sources see Figure 30.

**Figure 31. Age-Specific Shares of Laborers with Primary School Education**



For notes and sources see Figure 30.

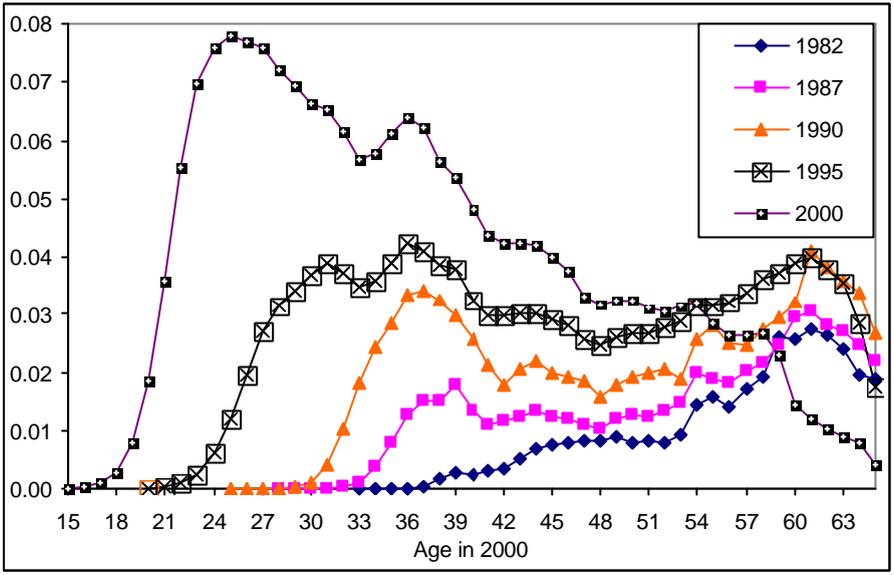
**Figure 32. Age-Specific Shares of Laborers with Lower Middle School Education**



Separate data on the special middle school in 1990 and 2000 were folded in with the upper middle school; the upper middle school data in other years, when the available education categories without special middle school are exhaustive, in all likelihood include the special middle school.

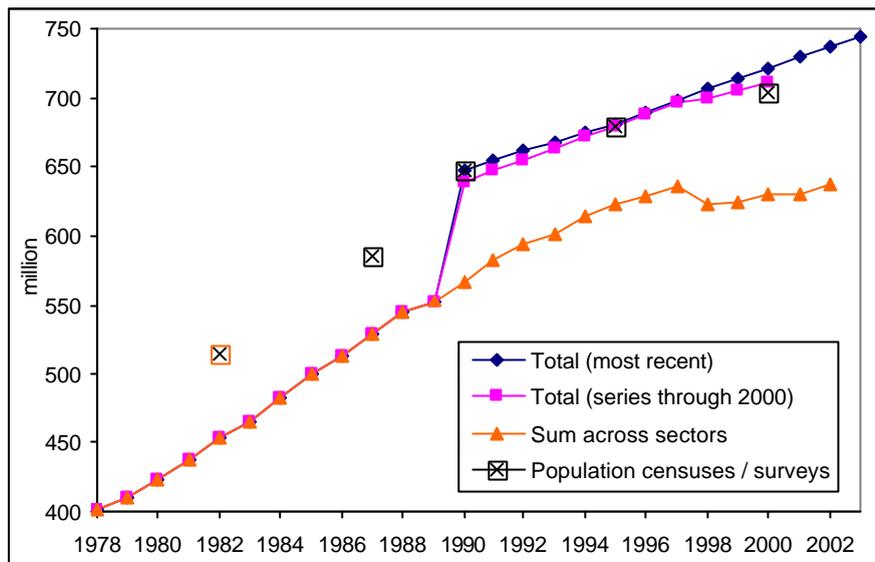
For further notes and sources see Figure 30.

**Figure 33. Age-Specific Shares of Laborers with Upper Middle School Education**



For notes and sources see Figure 30.

**Figure 34. Age-Specific Shares of Laborers with University Education**



Prior to 1990, the sum across sectors equals the total employment reported in the *Statistical Yearbook*. (The *Statistical Yearbook* series reports a total that equals the sum across sectors through the 1996 issue; the 1997 issue is the first to switch to new totals for all years since 1990.) The sectoral data end in 2002; the *Statistical Yearbook 2004* should have reported 2003 values, but didn't. Neither did the *Labor Yearbook 2004*.

*Statistical Yearbook* data are end-year data. Census and survey data through 1990 are midyear data, survey and census data in 1995 and 2000 are 1 Nov. data.

Sources: total (most recent): *Statistical Yearbook 2001*, p. 112 (for years through 1988), *2004*, p. 127; total (series through 2000): *Statistical Yearbook 2001*, p. 112; sum across sectors: reported as total (and equals the sum across sectors except for what appear rounding differences) in *Statistical Yearbook 1994*, p. 86, *1996*, p. 92, and calculated as sum across sectors since 1996 from data in *Statistical Yearbook 2004*, p. 127; population censuses / surveys: *Census 1982*, p. 468, *Census 1990*, Vol. 2, p. 476, *Census 2000*, Vol. 1, p. 215, Vol. 2, pp. 800, 1241 (with adjustments of the long-form labor data according to the share of long-form respondents in the total population), *Survey 1987*, pp. 1, 224 (p. 1 reports the sample size relative to the economy-wide population, used to augment the sample number of laborers to the economy-wide number of laborers), *Survey 1995*, pp. 1, 124 (p. 1 reports the sample size relative to the economy-wide population, used to augment the sample number of laborers to the economy-wide number of laborers).

**Figure 35. Employment Data**

**Table 7. Total Employment Data, million**

	<i>Statistical Yearbook</i>		Censuses/ surveys		Final mid-year series	
	Sum sectors	2001	2004	Total		Military
1978	401.52	401.52	401.52		459.06	
1979	410.24	410.24	410.24		469.03	
1980	423.61	423.61	423.61		484.32	
1981	437.25	437.25	437.25		499.91	
1982	452.95	452.95	452.95	513.63	4.24	517.86
1983	464.36	464.36	464.36			531.08
1984	481.97	481.97	481.97			551.48
1985	498.73	498.73	498.73			570.90
1986	512.82	512.82	512.82			587.22
1987	527.83	527.83	527.83	(585.15)		604.61
1988	543.34	543.34	543.34			622.57
1989	553.29	553.29	553.29			634.10
1990	567.40	639.09	647.49	647.24	3.20	650.44
1991	583.60	647.99	654.91			651.20
1992	594.32	655.54	661.52			658.22
1993	602.20	663.73	668.08			664.80
1994	614.70	671.99	674.55			671.32
1995	623.88	679.47	680.65	(680.03)		677.60
1996	628.42	688.50	689.50			685.08
1997	636.67	696.00	698.20			693.85
1998	623.63	699.57	706.37			702.29
1999	624.91	705.86	713.94			710.16
2000	629.78	711.50	720.85	703.83	2.50	717.40
2001	630.52	401.52	730.25			725.55
2002	637.79	410.24	737.40			733.83
2003			744.32			740.86

Final midyear series: 1991-2003: midyear values of *Statistical Yearbook 2004* series; 1982 and 1990: census values including military; 1983-89: interpolated values with annual change proportional to annual change in sum-sectors series; 1978-1981: 1982 census value (including military) combined with annual growth rates of *Statistical Yearbook 2004* series.

An alternative 1987 value can be obtained by using the information that the 1% sample population is supposedly equal to 0.999% of China's total population, yielding a number of laborers of 585.15m (*Survey 1987*, pp. 1, 224), and by adding in a linearly interpolated military personnel value of 3.59m (from the 1982 and 1990 censuses). In 1995, the 1% sample survey supposedly covered 1.02666% of the population, implying an economy-wide number of laborers of 680.03m (*Survey 1995*, pp. 1, 124); the linearly interpolated number of military personnel is 2.85m.

Sources:

*Statistical Yearbook* data: sum sectors: *Statistical Yearbook 2001*, p. 112, *2004*, p. 127; columns 2001 and 2004 from *Statistical Yearbook 2001*, p. 112, *2004*, p. 120.

Censuses: *Census 1982*, p. 468; *Census 1990*, Vol. 2, p. 476; *Census 2000*, Vol. 1, p. 215, Vol. 2, pp. 800, 1241.

Military: *Census 1982*, p. 505; *Census 1990*, Vol. 4, p. 496; *Census 2000*, Vol. 3, p. 1883.

**Table 8. Human Capital Measures 1978-1990**

	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
<i>Without military</i>													
Laborers (m)	454.18	464.32	479.77	495.52	513.63	526.99	547.53	567.08	583.54	601.05	619.14	630.79	647.24
Educ. ave.	5.83	5.74	5.72	5.75	5.80	5.86	5.94	6.03	6.13	6.24	6.44	6.64	6.84
Share primary	0.3041	0.3192	0.3297	0.3375	0.3430	0.3490	0.3548	0.3596	0.3627	0.3641	0.3706	0.3752	0.3775
Share lower m.	0.2071	0.2225	0.2373	0.2496	0.2599	0.2684	0.2754	0.2826	0.2898	0.2965	0.3079	0.3179	0.3266
Share upper m.	0.1673	0.1409	0.1227	0.1120	0.1055	0.1011	0.0986	0.0975	0.0982	0.1008	0.1033	0.1071	0.1123
Share univ.-coll.	0.0036	0.0039	0.0042	0.0045	0.0049	0.0053	0.0057	0.0062	0.0067	0.0073	0.0086	0.0101	0.0120
Share univ.-UG	0.0046	0.0042	0.0040	0.0038	0.0037	0.0037	0.0037	0.0038	0.0040	0.0043	0.0048	0.0055	0.0065
Share univ.-Grad	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003	0.0004
Share age 25-40	0.3775	0.3897	0.4014	0.4074	0.4140	0.4189	0.4152	0.4106	0.3995	0.3957	0.4060	0.4153	0.4244
Share age 30-45	0.3126	0.3159	0.3209	0.3235	0.3305	0.3356	0.3419	0.3494	0.3524	0.3554	0.3593	0.3577	0.3570
Share age 35-50	0.2722	0.2719	0.2715	0.2705	0.2723	0.2720	0.2711	0.2732	0.2734	0.2767	0.2824	0.2897	0.2991
Share age 40-55	0.2295	0.2313	0.2302	0.2284	0.2272	0.2243	0.2214	0.2195	0.2176	0.2170	0.2184	0.2202	0.2253
Share age 25-50	0.5397	0.5510	0.5614	0.5646	0.5706	0.5735	0.5660	0.5584	0.5452	0.5398	0.5511	0.5620	0.5743
Share age 50+	0.1164	0.1221	0.1265	0.1287	0.1310	0.1334	0.1349	0.1352	0.1356	0.1355	0.1376	0.1388	0.1402
Average age	32.18	32.47	32.67	32.80	32.96	33.02	33.06	33.09	33.12	33.16	33.33	33.53	33.77
<i>With military</i>													
Laborers (m)	459.06	469.03	484.32	499.91	517.86	531.08	551.48	570.90	587.22	604.61	622.57	634.10	650.44
Educ. ave.	5.87	5.78	5.76	5.79	5.84	5.90	5.97	6.06	6.16	6.27	6.47	6.67	6.86
Share primary	0.3011	0.3162	0.3269	0.3347	0.3404	0.3464	0.3524	0.3573	0.3605	0.3621	0.3687	0.3733	0.3757
Share lower m.	0.2104	0.2254	0.2398	0.2519	0.2619	0.2701	0.2769	0.2839	0.2909	0.2975	0.3087	0.3186	0.3272
Share upper m.	0.1702	0.1438	0.1257	0.1148	0.1082	0.1038	0.1011	0.0998	0.1003	0.1028	0.1051	0.1088	0.1139
Share univ.-coll.	0.0037	0.0040	0.0044	0.0047	0.0051	0.0055	0.0059	0.0064	0.0069	0.0075	0.0089	0.0104	0.0123
Share univ.-UG	0.0046	0.0043	0.0040	0.0039	0.0038	0.0038	0.0038	0.0039	0.0041	0.0044	0.0049	0.0057	0.0066
Share univ.-Grad	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003	0.0004

Educ. ave.: average years of education. (Illiterate/ semi-illiterate: 0, primary school: 6, lower middle school: 9, upper middle school (including special middle school): 12, university-college/ university-undergraduate studies/ university-graduate studies 15/16/19 years.

Share primary, lower m., upper m., univ.-coll., univ-UG, univ-Grad: share of laborers with highest level of education at the primary school, lower middle school, upper middle school, university-college, university-undergraduate, and university-graduate level.

Share age 25-40 (etc.): share of laborers age 25-40 (inclusive).

The average age of the age group 65+ is taken to be 69 years. The 1% sample data of 1982 and 1990 have data on laborers through age 99; the average age of those age 65 or above in 1982 was 68.71 years, and in 1990 68.97 years.

**Table 9. Human Capital Measures 1990-2003**

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
<i>Without military</i>														
Laborers (m)	647.24	648.08	655.17	661.83	668.42	674.77	682.32	691.16	699.66	707.59	714.90	723.11	731.45	738.54
Educ. ave.	6.84	6.93	7.02	7.10	7.19	7.27	7.43	7.58	7.72	7.86	7.99	8.12	8.24	8.37
Share primary	0.3775	0.3751	0.3731	0.3716	0.3703	0.3689	0.3591	0.3501	0.3415	0.3330	0.3247	0.3168	0.3091	0.3012
Share lower m.	0.3266	0.3366	0.3452	0.3530	0.3597	0.3657	0.3794	0.3916	0.4029	0.4129	0.4213	0.4279	0.4325	0.4353
Share upper m.	0.1123	0.1115	0.1110	0.1106	0.1103	0.1105	0.1142	0.1174	0.1204	0.1237	0.1274	0.1319	0.1373	0.1441
Share univ.-coll.	0.0120	0.0133	0.0147	0.0162	0.0178	0.0196	0.0221	0.0247	0.0274	0.0302	0.0331	0.0361	0.0391	0.0420
Share univ.-UG	0.0065	0.0067	0.0071	0.0075	0.0080	0.0086	0.0093	0.0101	0.0110	0.0119	0.0129	0.0140	0.0151	0.0162
Share univ.-Grad	0.0004	0.0004	0.0004	0.0005	0.0005	0.0006	0.0006	0.0007	0.0007	0.0008	0.0009	0.0010	0.0011	0.0012
Share age 25-40	0.4244	0.4303	0.4351	0.4403	0.4477	0.4535	0.4561	0.4593	0.4601	0.4600	0.4623	0.4637	0.4640	0.4568
Share age 30-45	0.3570	0.3510	0.3523	0.3677	0.3805	0.3910	0.3984	0.4028	0.4085	0.4147	0.4209	0.4257	0.4308	0.4338
Share age 35-50	0.2991	0.3044	0.3109	0.3180	0.3191	0.3196	0.3156	0.3184	0.3323	0.3451	0.3556	0.3651	0.3714	0.3801
Share age 40-55	0.2253	0.2279	0.2338	0.2404	0.2481	0.2572	0.2641	0.2713	0.2783	0.2808	0.2825	0.2818	0.2875	0.3013
Share age 25-50	0.5743	0.5854	0.5939	0.6058	0.6204	0.6338	0.6446	0.6536	0.6602	0.6650	0.6661	0.6647	0.6600	0.6533
Share age 50+	0.1402	0.1419	0.1434	0.1439	0.1449	0.1460	0.1494	0.1535	0.1577	0.1632	0.1700	0.1764	0.1846	0.1930
Average age	33.77	34.01	34.29	34.56	34.83	35.11	35.44	35.75	36.07	36.40	36.74	37.07	37.39	37.71
<i>With military</i>														
Laborers (m)	650.44	651.20	658.21	664.80	671.32	677.60	685.08	693.85	702.29	710.16	717.39	725.55	733.82	740.86
Educ. ave.	6.86	6.95	7.04	7.12	7.21	7.29	7.45	7.60	7.74	7.87	8.01	8.13	8.26	8.38
Share primary	0.3757	0.3733	0.3714	0.3700	0.3687	0.3674	0.3577	0.3488	0.3402	0.3318	0.3236	0.3157	0.3081	0.3003
Share lower m.	0.3272	0.3370	0.3455	0.3532	0.3598	0.3657	0.3792	0.3914	0.4026	0.4125	0.4208	0.4273	0.4319	0.4346
Share upper m.	0.1139	0.1131	0.1124	0.1120	0.1117	0.1117	0.1153	0.1185	0.1214	0.1246	0.1283	0.1326	0.1379	0.1446
Share univ.-coll.	0.0123	0.0136	0.0151	0.0166	0.0182	0.0201	0.0226	0.0253	0.0280	0.0309	0.0338	0.0368	0.0398	0.0428
Share univ.-UG	0.0066	0.0069	0.0073	0.0077	0.0082	0.0088	0.0096	0.0104	0.0113	0.0123	0.0133	0.0144	0.0155	0.0166
Share univ.-Grad	0.0004	0.0004	0.0004	0.0005	0.0005	0.0006	0.0006	0.0007	0.0008	0.0009	0.0009	0.0010	0.0012	0.0013

Also see notes to previous table.

**Table 10. New Enrollment and Graduation Rates, 1990, 2000, 2003, 2004**

	New enrollment in year X / population at age Y					Graduated in year X / population at age Y				
	Age	1990	2000	2003	2004	Age	1990	2000	2003	2004
Primary school	7	1.0220	1.0699	1.1093	1.1385	13	0.9620	0.9580	0.8783	0.9667
Lower middle school	13	0.7073	0.8963	0.8502	0.9409	16	0.4722	0.7944	0.7915	0.8250
Upper middle school	16	0.1899	0.3673	0.4593	0.5457	19	0.1494	0.3329	0.3676	0.4416
University – college	19	0.0117	0.0554	0.0989	0.1149	22	0.0125	0.0245	0.0503	0.0556
University – undergrad.	19	0.0117	0.0615	0.0904	0.1051	23	0.0135	0.0264	0.0502	0.0629
University – graduate	23	0.0013	0.0068	0.0145	0.0173	26	0.0014	0.0025	0.0059	0.0082

New enrollment and graduation data from the *Statistical Yearbook* for the three years (1990, 2000, 2003) and from the *Statistical Communiqué 2004* for 2004 are related to the population data in the 1990 census, the 2000 census, and the 2000 census with each age group aged by 3 and 4 years (taking into consideration the constant, year 2000 age-specific death rates).

At the primary school level, new enrollment exceeds the population number presumably because the population number at age 7 is an underestimate (due to underreporting in the censuses).

The upper middle school category includes the regular upper middle schools, special middle schools, and vocational schools.

In 1990, the available number for university students (excluding graduate students) is split evenly into college and undergraduate students, as the data of 1994, the first which come with a breakdown, suggest. The in 2004 again only available number for all university students (excluding graduate students) is split in the proportions of the 2003 data (both for new enrollment and graduates).

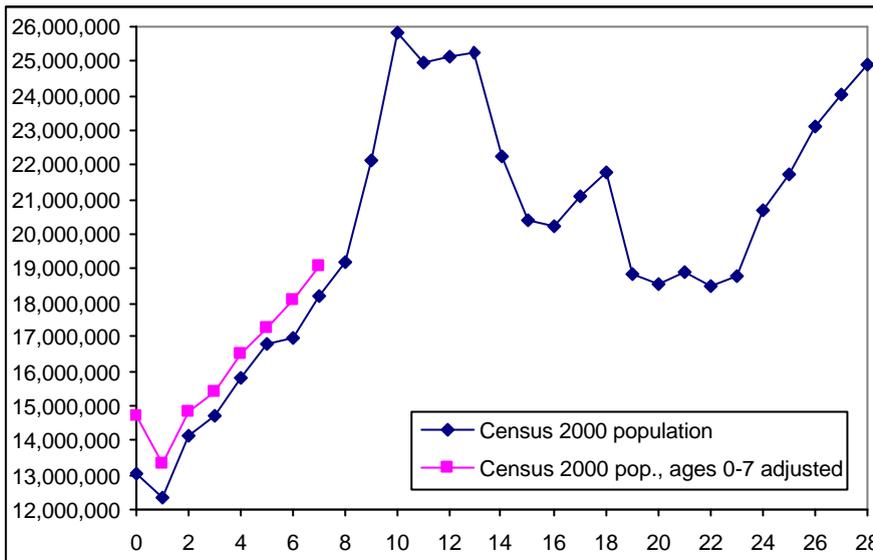
Students at “special schools” are ignored throughout; this is a very small number of around 50,000 every year. Special schools are presumably a substitute for primary school.

Sources: *Statistical Yearbook 2004*, pp. 780-2; 2002, p. 676; *Census 2000*, Vol. 1, pp. 570, 713; *Census 1990*, Vol. 2, p. 2f.; 2004 enrollment and graduation data from the 2004 *Statistical Communiqué* at [http://www.stats.gov.cn/tjgb/ndtjgb/qgndtjgb/t20050228\\_402231854.htm](http://www.stats.gov.cn/tjgb/ndtjgb/qgndtjgb/t20050228_402231854.htm).

**Table 11. New Enrollment and Graduation Numbers (million), 1990, 2000, 2003, 2004**

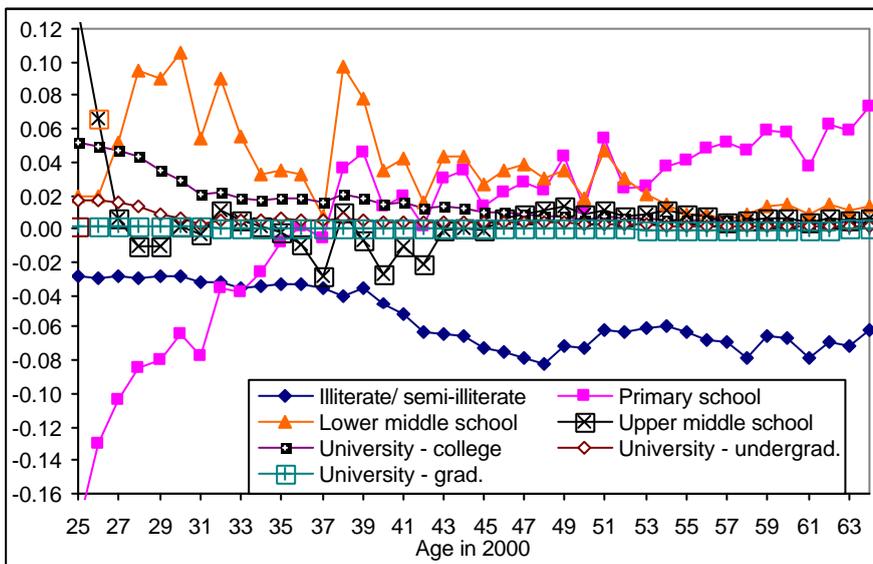
	Newly enrolled				Graduates			
	1990	2000	2003	2004	1990	2000	2003	2004
Primary school	20.640	19.465	18.294	17.470	18.631	24.192	22.679	21.352
Lower middle school	13.699	22.633	21.953	20.782	11.091	16.071	19.956	20.704
Upper middle school	4.460	7.430	11.581	13.696	3.884	6.285	7.420	8.979
University – college	0.305	1.046	1.996	2.337	0.307	0.454	0.948	1.207
University – undergraduates	0.305	1.160	1.825	2.136	0.307	0.496	0.930	1.184
University – graduate	0.030	0.128	0.269	0.326	0.035	0.059	0.111	0.151

For notes and sources see previous table.



Adjusted values assume that age groups 0-7 are underreported in the 2000 census to the same degree as the age groups 0-7 in the 1990 census were (as revealed by the 2000 census for the age groups 10-17, taking into consideration interpolated year- and age-specific deaths between 1990 and 2000).  
 Sources: *Census 2000*, Vol. 1, pp. 570, 713; *Census 1990*, Vol. 2, p. 2f., Vol. 4, p. 9.

**Figure 36. Census 2000 Population Age 0-28, Absolute Numbers**

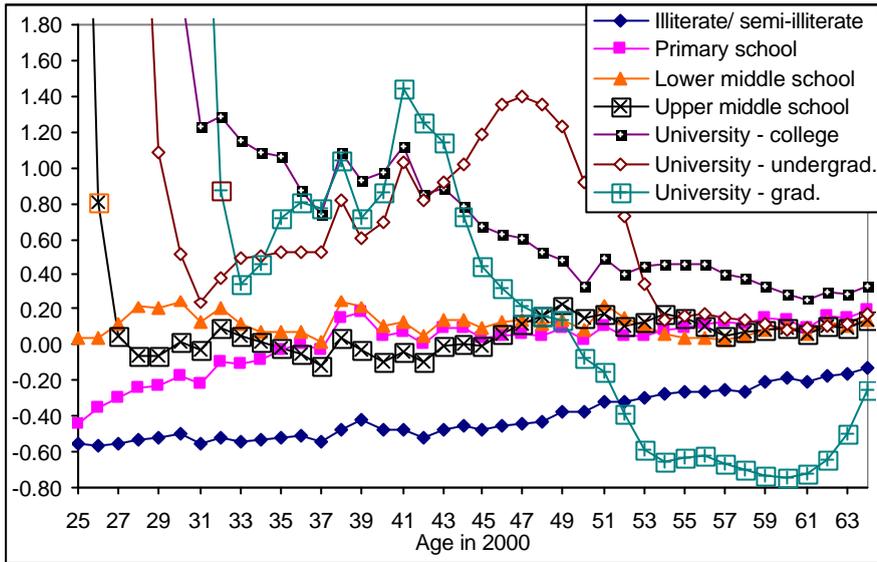


The figure depicts the number of persons with education X at age Y in 2000 minus the number of persons with education X at age Y-10 in 1990, and this absolute difference is divided by the number of all persons age Y-10 in 1990. The figure is limited to the working age population, starting age 15 in 1990 (corresponding to age 25 in 2000), and ending with age 54 in 1990 (corresponding to age 64 in 2000); the age group 65+ in 2000 is not included since there is no comparison group in 1990.

Each cohort in 1990 is aged from 1990 to 2000, to take into account deaths, before making the comparison. (Age-specific death rates of 1991 through 1999 are obtained via interpolation between the 1990 and 2000 age-specific death rates, using average annual geometric growth rates.)

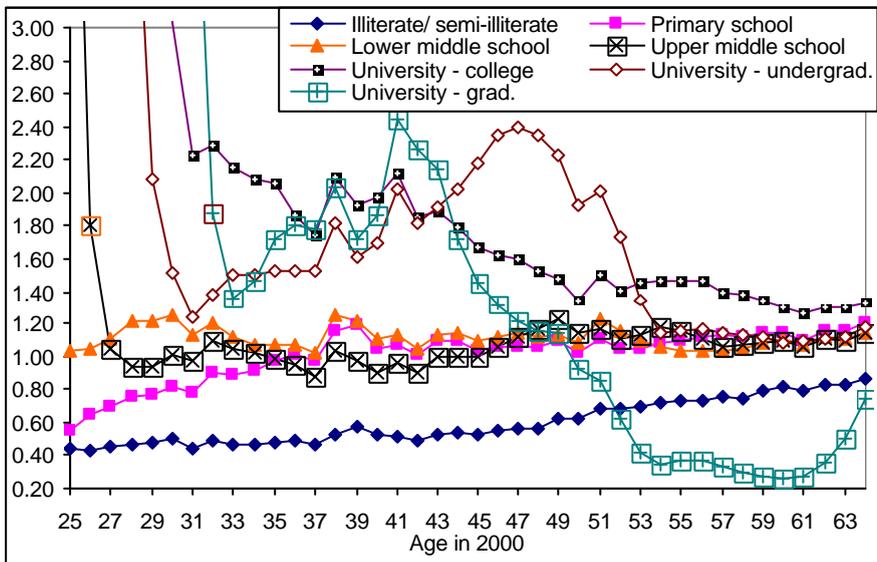
Sources: *Census 1990*, Vol. 2, pp. 2-5, Vol. 4, pp. 9-13; *Census 2000*, Vol. 1, pp. 215-7, 713-5.

**Figure 37. Cohort-Specific Absolute Changes in Education Level, Relative to Age-Specific Population 1990 to 2000**



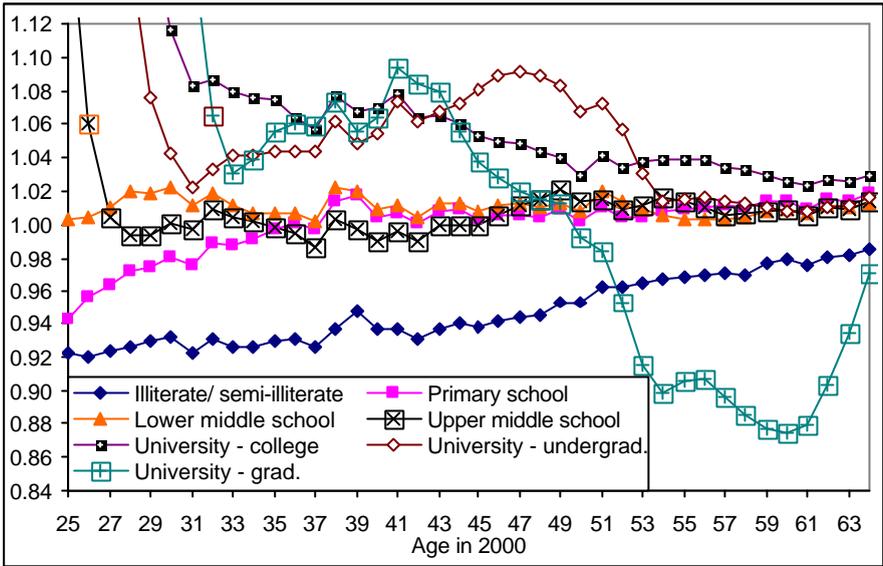
The figure depicts the number of persons with education X at age Y in 2000 minus the number of persons with education X at age Y-10 in 1990, with this difference divided by the number of persons with education X at age Y-10 in 1990. I.e., the figure shows the X-fold growth (percentage divided by 100) in the number of persons with education X and age Y-10 in 1990 over the next 10 years. For further notes and sources see Figure 37.

**Figure 38. Cohort-Specific X-Fold Changes in Education Level, 1990 to 2000**



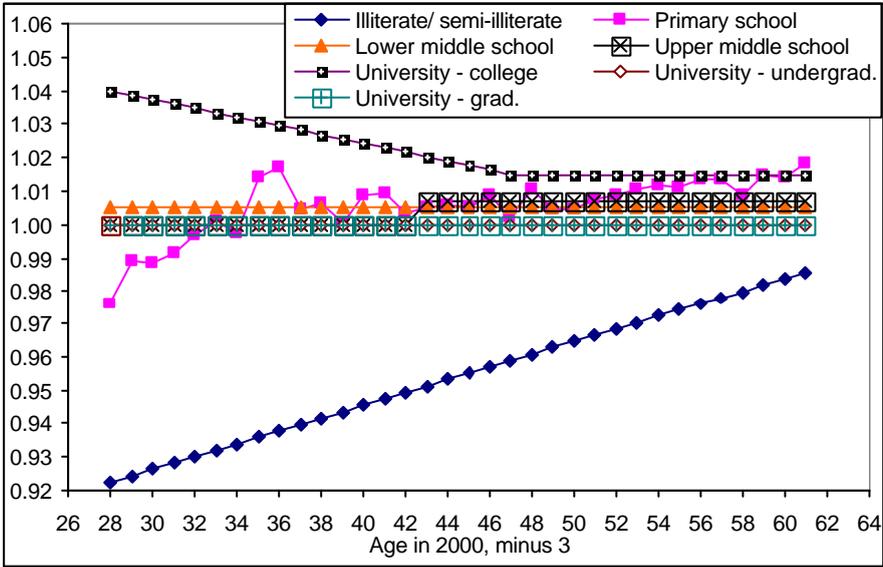
The figure depicts the number of persons with education X at age Y in 2000 divided by the number of persons with education X at age Y-10 in 1990. For further notes and sources see Figure 37.

**Figure 39. Cohort-Specific Relative Changes in Education Level, 1990 to 2000**



The figure depicts the number of persons with education X at age Y in 2000 divided by the number of persons with education X at age Y-10 in 1990, raised to the power of 1/10. For further notes and sources see Figure 37.

**Figure 40. Cohort-Specific Average Annual Relative Changes in Education Level, 1990 to 2000**



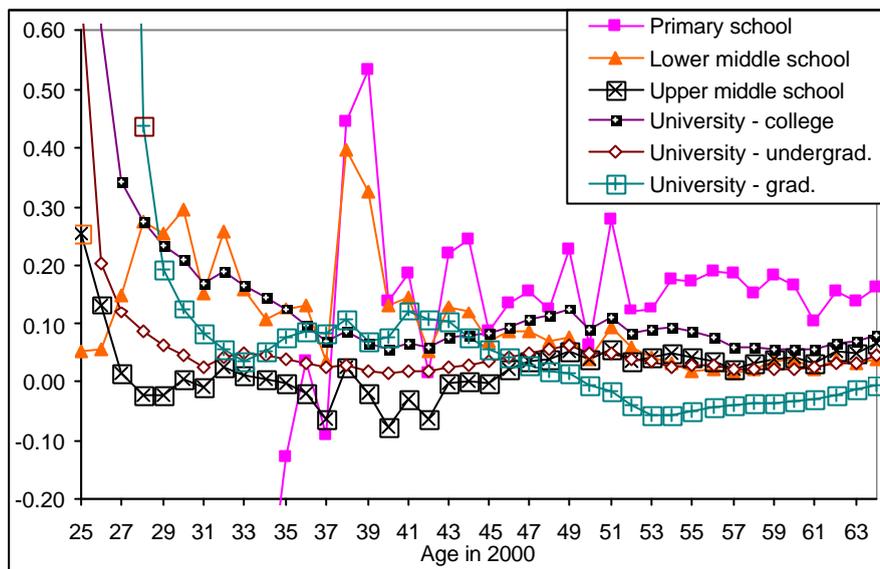
The figure depicts the number of persons with education X at age Y in 2000 divided by the number of persons with education X at age Y-10 in 1990, raised to the power of 0.1; the growth structure is imposed, and the growth rates are shifted backwards from 2000 by three years.

For the category illiterate/semi-illiterate persons, the first and last values (0.9225 and 0.9855) are connected linearly. The primary school category is depicted using the actual (1990-2000) average annual growth rates. For the lower middle school, a 0.5% growth rate is imposed. For the upper middle school, zero growth is assumed through age 45 in 2000 (some persons enter, but others exit this category) and a 0.7% average annual growth rate thereafter (corresponding approximately to a 10% 10-year growth rate). For the university-college category, the actual average annual growth rate is approximately halved (arithmetically halved) by imposing a 4% growth rate on the age group 31 (in 2000) and a 1.5% growth rate on the age group 50 (in 2000), and by connecting these two values

linearly; above age 50 the growth rate is held constant at 1.5%. University undergraduate and graduate studies are assumed to not be attainable through adult education (zero growth rate).

For further notes and sources see Figure 37.

**Figure 41. Imposed Average Annual Cohort-Specific Relative Changes in Education Level**



The figure depicts the number of persons with education X at age Y in 2000 minus the number of persons with education X at age Y-10 in 1990, with this difference divided by the number of persons with education X-1 (next lower education level for college and undergraduate studies is the upper middle school, for graduate studies it is undergraduate studies.) For further notes and sources see Figure 37.

**Figure 42. Cohort-Specific Changes in Education Level Relative to Original Education Level, 1990 to 2000**

**Table 12. Human Capital Measures 2000-2012**

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
<b><i>Without military</i></b>													
Laborers (m)	711.91	717.96	723.98	729.71	735.78	741.92	745.94	750.65	756.61	762.80	768.73	773.20	776.43
Educ. ave.	8.00	8.11	8.21	8.31	8.42	8.51	8.62	8.73	8.84	8.96	9.08	9.19	9.31
Share primary	0.3231	0.3176	0.3112	0.3043	0.2964	0.2915	0.2848	0.2773	0.2690	0.2604	0.2518	0.2428	0.2338
Share lower m.	0.4228	0.4307	0.4385	0.4462	0.4534	0.4565	0.4591	0.4602	0.4605	0.4595	0.4579	0.4563	0.4549
Share upper m.	0.1278	0.1312	0.1343	0.1373	0.1402	0.1433	0.1472	0.1517	0.1567	0.1622	0.1675	0.1724	0.1768
Share univ.-coll.	0.0331	0.0347	0.0363	0.0382	0.0404	0.0429	0.0460	0.0495	0.0534	0.0574	0.0616	0.0659	0.0703
Share univ.-UG	0.0129	0.0132	0.0137	0.0143	0.0152	0.0163	0.0178	0.0197	0.0221	0.0248	0.0278	0.0309	0.0342
Share univ.-Grad	0.0009	0.0010	0.0011	0.0012	0.0014	0.0016	0.0019	0.0022	0.0026	0.0030	0.0035	0.0041	0.0047
Share age 25-40	0.4624	0.4663	0.4695	0.4660	0.4535	0.4410	0.4315	0.4248	0.4188	0.4086	0.3988	0.3909	0.3883
Share age 30-45	0.4210	0.4262	0.4307	0.4335	0.4358	0.4402	0.4449	0.4485	0.4449	0.4331	0.4215	0.4124	0.4071
Share age 35-50	0.3557	0.3643	0.3699	0.3772	0.3839	0.3907	0.3969	0.4021	0.4047	0.4069	0.4107	0.4141	0.4171
Share age 40-55	0.2826	0.2807	0.2845	0.2969	0.3081	0.3167	0.3252	0.3306	0.3372	0.3431	0.3494	0.3547	0.3596
Share age 25-50	0.6662	0.6667	0.6644	0.6601	0.6554	0.6496	0.6455	0.6452	0.6420	0.6386	0.6390	0.6429	0.6524
Share age 50+	0.1701	0.1734	0.1785	0.1837	0.1905	0.1971	0.2030	0.2096	0.2149	0.2164	0.2172	0.2158	0.2194
Average age	36.75	37.04	37.35	37.66	38.04	38.38	38.83	39.26	39.62	39.96	40.28	40.59	40.94
<b><i>With military</i></b>													
Laborers (m)	714.41	720.46	726.48	732.21	738.28	744.42	748.44	753.15	759.11	765.30	771.23	775.70	778.93
Educ. ave.	8.01	8.12	8.23	8.33	8.43	8.53	8.63	8.74	8.85	8.97	9.09	9.20	9.32
Share primary	0.3220	0.3165	0.3102	0.3033	0.2954	0.2906	0.2839	0.2764	0.2681	0.2595	0.2510	0.2421	0.2331
Share lower m.	0.4222	0.4302	0.4379	0.4456	0.4528	0.4559	0.4585	0.4596	0.4599	0.4589	0.4573	0.4557	0.4543
Share upper m.	0.1287	0.1320	0.1351	0.1380	0.1409	0.1441	0.1479	0.1524	0.1574	0.1628	0.1681	0.1730	0.1774
Share univ.-coll.	0.0338	0.0353	0.0370	0.0389	0.0410	0.0435	0.0466	0.0501	0.0539	0.0580	0.0622	0.0665	0.0708
Share univ.-UG	0.0132	0.0136	0.0141	0.0147	0.0155	0.0166	0.0181	0.0201	0.0224	0.0251	0.0281	0.0312	0.0345
Share univ.-Grad	0.0010	0.0010	0.0011	0.0013	0.0014	0.0017	0.0019	0.0023	0.0026	0.0031	0.0036	0.0041	0.0047

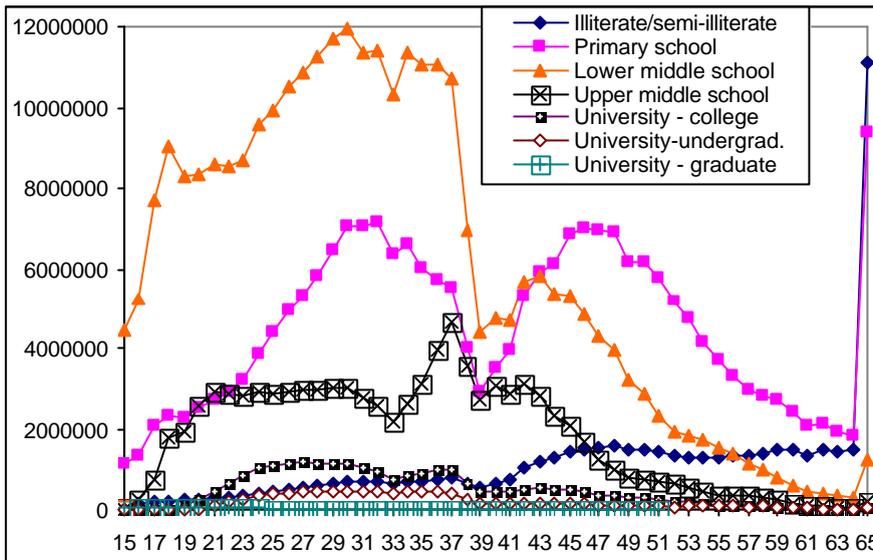
Human capital measures are unadjusted measures. I.e., the sum of projected laborers in different education categories at a particular age need not equal projected total laborers at that age; the first rather than the latter is used in the calculation of the aggregate values. (The difference is due to the application of age- and education-specific employment rates to population data. Employment rates differ across education categories and age. As new age cohorts become more highly educated, the aggregate age-specific employment rate of the total across all education categories at a particular age no longer equals the sum of laborers across education categories at that age.)

Also see notes to Table 8.

**Table 13. Human Capital Measures 2013-2025**

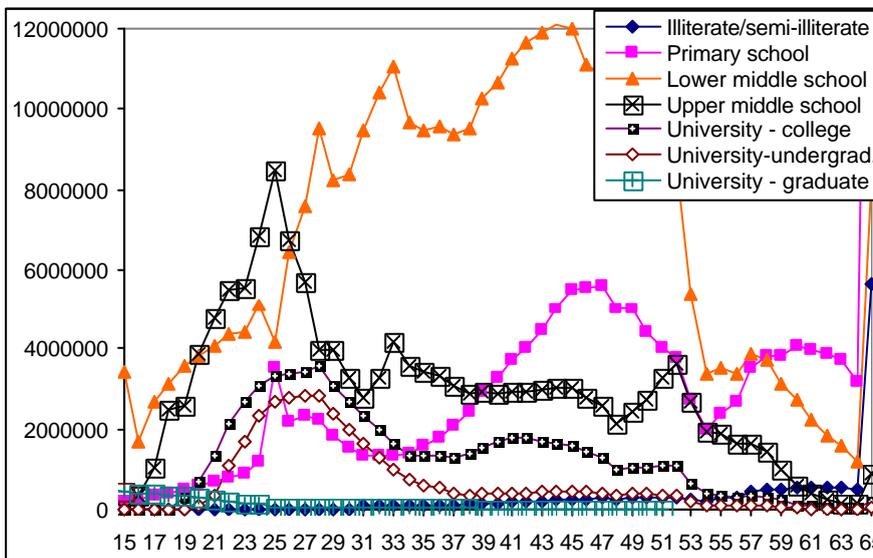
	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
<b><i>Without military</i></b>													
Laborers (m)	778.29	778.02	776.62	773.76	768.97	762.84	754.67	746.11	737.67	730.61	726.46	724.63	723.21
Educ. ave.	9.42	9.54	9.66	9.77	9.89	10.02	10.15	10.28	10.42	10.54	10.65	10.74	10.83
Share primary	0.2246	0.2149	0.2047	0.1942	0.1828	0.1710	0.1582	0.1452	0.1325	0.1207	0.1102	0.1014	0.0938
Share lower m.	0.4538	0.4530	0.4529	0.4527	0.4532	0.4539	0.4546	0.4549	0.4546	0.4538	0.4530	0.4514	0.4490
Share upper m.	0.1808	0.1848	0.1886	0.1922	0.1959	0.1995	0.2035	0.2079	0.2125	0.2167	0.2202	0.2230	0.2256
Share univ.-coll.	0.0747	0.0792	0.0838	0.0886	0.0936	0.0988	0.1043	0.1100	0.1158	0.1214	0.1265	0.1312	0.1359
Share univ.-UG	0.0374	0.0406	0.0438	0.0469	0.0501	0.0534	0.0567	0.0601	0.0634	0.0666	0.0694	0.0720	0.0746
Share univ.-Grad	0.0054	0.0061	0.0069	0.0077	0.0085	0.0094	0.0103	0.0111	0.0120	0.0129	0.0138	0.0145	0.0153
Share age 25-40	0.3877	0.3890	0.3931	0.3953	0.3965	0.4004	0.4046	0.4075	0.4098	0.4095	0.4033	0.3952	0.3892
Share age 30-45	0.4038	0.3976	0.3922	0.3886	0.3906	0.3944	0.4002	0.4084	0.4139	0.4164	0.4194	0.4202	0.4192
Share age 35-50	0.4152	0.4075	0.4005	0.3955	0.3946	0.3953	0.3933	0.3916	0.3913	0.3952	0.3988	0.4019	0.4069
Share age 40-55	0.3634	0.3678	0.3737	0.3793	0.3846	0.3859	0.3826	0.3794	0.3773	0.3776	0.3770	0.3720	0.3671
Share age 25-50	0.6577	0.6573	0.6592	0.6589	0.6568	0.6575	0.6560	0.6537	0.6494	0.6436	0.6358	0.6253	0.6169
Share age 50+	0.2301	0.2396	0.2472	0.2547	0.2588	0.2651	0.2697	0.2751	0.2796	0.2842	0.2907	0.2984	0.3051
Average age	41.30	41.69	42.06	42.44	42.82	43.19	43.45	43.67	43.85	44.11	44.56	45.09	45.55
<b><i>With military</i></b>													
Laborers (m)	780.79	780.52	779.12	776.26	771.47	765.34	757.17	748.61	740.17	733.11	728.96	727.12	725.70
Educ. ave.	9.43	9.55	9.66	9.78	9.90	10.03	10.16	10.29	10.42	10.55	10.66	10.75	10.84
Share primary	0.2239	0.2142	0.2040	0.1936	0.1822	0.1704	0.1577	0.1447	0.1321	0.1203	0.1098	0.1011	0.0935
Share lower m.	0.4532	0.4524	0.4523	0.4521	0.4526	0.4533	0.4540	0.4543	0.4539	0.4532	0.4523	0.4508	0.4484
Share upper m.	0.1814	0.1854	0.1891	0.1928	0.1964	0.2000	0.2040	0.2084	0.2130	0.2172	0.2207	0.2235	0.2260
Share univ.-coll.	0.0752	0.0797	0.0843	0.0890	0.0940	0.0992	0.1047	0.1104	0.1162	0.1217	0.1268	0.1316	0.1362
Share univ.-UG	0.0377	0.0409	0.0440	0.0472	0.0504	0.0536	0.0569	0.0603	0.0636	0.0667	0.0696	0.0722	0.0747
Share univ.-Grad	0.0054	0.0062	0.0069	0.0077	0.0085	0.0094	0.0103	0.0112	0.0121	0.0129	0.0138	0.0146	0.0153

Also see notes to previous table and to Table 8.



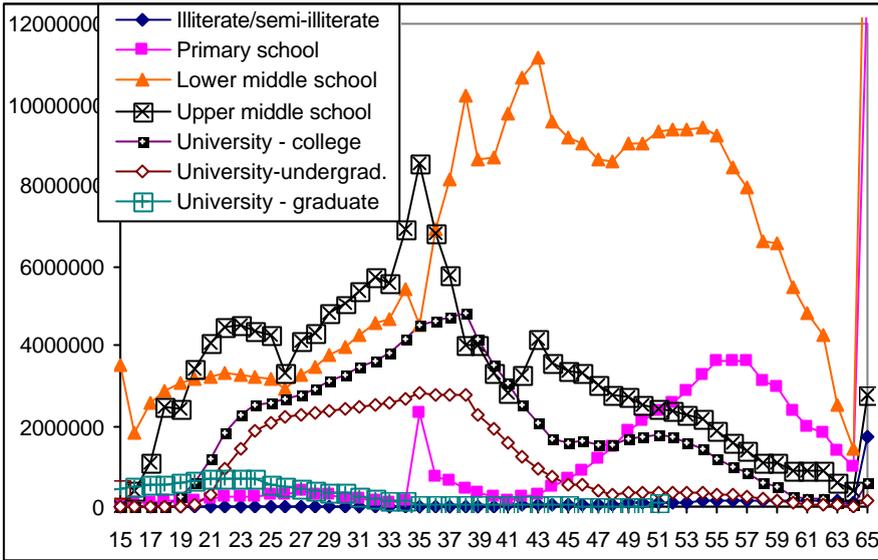
Age group 65 is 65+. Data are based on age- and education-specific employment rates in the long-form sample of the 2000 census, applied to the year 2000 age- and education-specific population (with enrollment figures imposed at youngest ages).

**Figure 43. Age- and education-specific laborers, 2000**



Age group 65 is 65+. Data are based on age- and education-specific employment rates in the long-form sample of the 2000 census, applied to the current-year age- and education-specific population.

**Figure 44. Age- and education-specific laborers, 2015**



Age group 65 is 65+. Data are based on age- and education-specific employment rates in the long-form sample of the 2000 census, applied to the current-year age- and education-specific population.

**Figure 45. Age- and education-specific laborers, 2025**

## Translations

Knows no or few characters *bu shizi huo hen shao*

No schooling *wei shangguo xue*

Literacy class *saomang ban*

Primary school *xiaoxue*

Middle school

Lower middle school *chuzhong*

Technical middle school *jigong* (corresponds to either lower or upper middle school)

Upper middle school *gaozhong*

Special middle school *zhongzhuan*

Vocation school *zhiye gaozhong*

University

College (college-level associate degree) *daxue zhuanke*

Undergraduate studies *daxue benke*

Graduate studies (Masters degree, PhD) *yanjiusheng*

Degrees of completion (1992, 2000 censuses)

Graduated *biye*

Currently attending *zaixiao*

Completed *yiye*

in 1992 census: completed course of study but did not graduate, or attended but did not complete course of study

in 2000 census: completed course of study, but not graduated

Not completed (discontinued) *chuoye* (this category is not included in 1992 census)

Others *qita*

Illiterate or semi-illiterate *wenmang, ban wenmang*