経済研究
Vol. 67, No. 1, Jan. 2016

Changes in the Occupational Structure of Chinese Workforce in the Course of China's Industrialization and Urbanization, 1912–1952*

Harry X. Wu

This paper makes a preliminary attempt to fill a long-standing employment data gap in the study of China's economic growth in the Republican era. Instead of searching for other data than that used in the previous studies, it tries to organize the available estimates by various sources in a consistent occupational framework and reassess them against the macroeconomic background that might have impacts on the performance of the economy and hence its employment and occupational distribution. It focuses mainly on the construction of the 1915 benchmark and the revision of the existing 1933 and 1949–1952 benchmarks. It also provides time-series estimates for the employment by major occupation from the demand-side and for the total and working-age population from the supply-side of the system. The two sides are linked by estimated labor participation rate. Finally, it gauges the urbanization rate of the population implied by the change of the urban employment.
JEL classification: C82, E24, O47

1. Introduction

The essential driver of the modern economic growth is technological advancement with supportive institutions (Kuznets 1966). New technology induces a substantial increase in the division of labor within and across industries, which enhances what Adam Smith called "productive powers of labour" (Smith 1776). Pro-market institutions ensure the most efficient use of available technologies, hence helping maximize national income with given resources. The course of industrialization is driven by continuous technological progress as well as institutional changes, which affect not only the reallocation of physical resources, but also the division of labor with embodied human capital stock, both economy-wide and internationally. In this sense, changes in the occupational distribution and industrial structure of labor force are indispensible indicators for a better understanding of the modern economic growth.

China's traditional division of labor remained intact until the late nineteenth century. Although modern factories, transportation and telecommunication industries began to emerge in the "treaty port" cities after the two opium wars (1839–1842 and 1856–1860), the share of modern manufacturing in the Chinese economy was trivial. Based on the studies in Ou *et al.* (1947), Liu and Yeh (1965) and Rawski (1989) on China's growth between 1913 and 1933, Maddison (1998 and 2007) estimated that the industrial sector accounted for not more than 8 percent of the total GDP in 1890, of which 98 percent was attributable to handicraft industries. Nevertheless, a significant part of the handicrafts was not carried out by specialized workers but concentrated in rural households where family members engaged in non-farming activities in the form of by-employment (Skinner 1964–1965; Maddison 2007)[1]. Despite the government-engineered industrialization in the name of "self-strengthening movement" from the 1860s and the increasing influences of the foreign trade and direct investment alongside, the position of the

household-based handicrafts in rural China had not been significantly changed by the early 1920s (Feuerwerker 1983).

This is an important starting point for a better understanding of the occupational changes in China along with the early industrialization process in the first half of the twentieth century — how the household-based traditional division of labor first dissolved due to the incursion of the modern industries and then how increasingly market-driven occupational distribution evolved. This requires not only the construction of individual benchmarks for which better data are available, but also scholarly efforts to connect the benchmarks with conceptual, coverage and classification consistencies.

This paper takes two initiatives: constructing a new benchmark for China's occupational distribution in 1915 and establishing a time series between the available benchmarks for the period 1912–1952. The latter also involves revisions to the existing benchmark estimates for 1933 (Liu and Yeh 1965; Zhang 1947; Ou *et al.* 1947; Makino and Luo 2014) and for 1952 (Liu and Yeh 1965; Wu 2014a and 2014b).

This paper is organized as follows. Section 2 addresses the research problem by a brief but critical review of the literature. Section 3 presents a conceptual framework for a consistent data handling of the supply-side and demand-side data. Section 4 presents the main steps in constructing the 1915 benchmark. Section 5 reassesses and revises the available estimates for 1933 and 1949–1952. Section 6 makes a preliminary effort to construct time series estimates for the period 1912–1952. Section 7 briefly concludes this study and highlights future research priorities.

2. A Brief Literature Review

Research endeavors on China's occupational distribution in the first half of the twentieth century have so far mostly, if not only, focused on the early 1930s for two reasons. One is the availability of large-scale-survey-based occupation data and the other is the completion of China's first national accounts following the international standard. Two primary data sources played a fundamental role. They are land utilization survey led by John Lossing Buck for the period 1929–1933 (Buck 1937) and the National Resource Commission (NRC)'s factory survey for 1933 led by D. K. Lieu (Lieu 1937). Based on careful evaluations of these data Pao-san Ou led a team to construct China's first national income accounts for 1933 (Ou *et al.* 1947). In order to estimate the value-added on per worker basis, Ou gauged labor-unit-adjusted occupation estimates to deal with the by-employment problem.

Liu and Yeh (1965) made the first attempt to construct an economy-wide labor accounts with broad occupational classification for China in 1933 and coherently incorporated it in their new endeavor in constructing the Chinese national income accounts. Following Ou, they also took into accounts the by-employed and worked out full-time and single-occupation equivalent estimates. Besides, they assessed the "communist statistics" of the 1950s and reconciled them with their 1933 benchmark (1965: 184–212). Their estimates on both the labor and income accounts have been considered most carefully done, hence well accepted in studies on the modern Chinese economy (Perkins 1969; Feuerwerker 1983; Maddison 1998). Unfortunately, they made no attempt to establish an earlier yardstick, e.g. the 1910s and/or the 1920s, for a necessary backward judgment for their 1933 benchmark.

Almost at the same time when Ou *et al.* (1947) accomplished his work on the 1933 national accounts, Zhang (1947) produced a new set of employment estimates also for 1933. To measure income distribution among

Table 1. Various Estimates of Chinese Population, Employment and Occupational Distribution for 1933

(Thousand persons)

	Liu and Yeh (1965)	Zhang (1947)	Makino and Luo (2014)	This Study
Population	500,000	429,720	517,280	512,948
Employment	259,220	189,077	265,970	262,352
(Share of employment, %)	(51.8)	(44.0)	(51.4)	(51.1)
	PST Distribution (total=100)			
Primary	79.3	76.3	79.8	73.8
Secondary	7.1	8.9	8.7	11.0
Tertiary	13.5	14.9	11.5	15.2

Sources: Liu and Yeh (1965: Tables 54 and 55), Zhang (1947: 84), and Makino and Luo (2014: 60; Tables 2. 1. 2, 2. 2. 1); and Table 4 of this study.

Notes: Primary: Agriculture and mining. Secondary: Manufacturing, utility and construction. Tertiary: All services.

Chinese population, Zhang changed the employment estimates in Ou *et al.* from a *full-time equivalent* to a *natural number-based* measure. Interestingly, Zhang's results show a lower rather than a higher share of the primary sector compared to Liu and Yeh as presented in Table 1. There are however even much greater differences between the two studies in terms of population and employment estimates. That is, Liu and Yeh's results appear to be 16 and 37 percent higher than those of Zhang respectively. Therefore, a more careful examination of the Chinese demographics is called in order to settle these differences. Indeed, Zhang relied on Wang's population estimate for 1933 (Wang 1935: 202–204) which is criticized by Hou (2001: 258–259) for substantially underestimating the size of the population. However, any assessment of demographic data naturally involves an assessment over a certain time horizon; thus a time series appears to be indispensible. Revisiting these important benchmark works and considering how they may link with other time points motivate this study and eventually result in a new set of benchmark estimates, of which the 1933 benchmark is also included as shown in Table 1.

Table 1 shows that the estimates by Makino and Luo (2014) have further widened the differences. Their study is part of the *China Volume* that belongs to a long series of historical statistics for major Asian economies (Odaka, Saito and Fukao 2008)[2]. They also work on the 1933 benchmark following Ou *et al.* (1947) and Liu and Yeh (1965) in principle, but they show different results based on their reassessment of the historical data (2014: 57–62). They have made perhaps the first attempt to construct a time series for the Chinese population from 1912 to 1949 and for the employment by broad sector from 1931 to 1944 using econometrics (2014: Tables 2. 1. 2 and 2. 2. 1). Their empirical approach is nevertheless questionable. For example, in the case of population, they rely solely on the population and agricultural changes in 1952–1970 to predict the change of the population over the period 1912–1949 albeit anchored by the available benchmark estimates in the literature. It is hard to imagine why they ignored that the two periods suffered from very different shocks by nature. Like others, they made no attempt to construct an earlier and/or a later benchmark to link with their 1933 estimates, hence evaluating the latter.

Therefore, constructing an earlier benchmark for the mid-1910s and establishing a more or less coherent link between this new yardstick and the existing benchmarks of 1933 and the early 1950s, with revisions if necessary, are the primary goals of the present study.

3. A Conceptual Framework

In the early stages of development, when an economy shifts from traditional agricul-

ture to the modern industry and services, by-employment and irregular working hours are typically phenomenal. Although properly handling the problem is difficult with very limited historical data, ignoring it may mislead our understanding of labor input and labor productivity in China in the first half of the twentieth century. In this section, we propose a conceptual framework to categorize workers of different employment status and with different hours worked in any given time and across sectors or industries.

It is better to start with a simple accounting identity between the supply-side and the demand-side of our research problem as expressed in Equation 1. It shows how the demand for labor input, L, is met by the labor supply of an economy originated in its population, N, through the share of the working-age, α, and the participation rate of the working-age, ρ. The complete expression of this supply-demand identity also has to take into account the rate of the unemployment of the participated, μ. It is needless to say that all the parameters are time variant.

$$\frac{L_t}{(1-\mu_t)} = \alpha_t \rho_t N_t \qquad (1)$$

We now consider the concept of labor productivity. Equation 2 shows that output (value added) per unit of labor input[3], y, is defined as the total output, Y, divided by the total number of hours worked, H, which is a product of the numbers employed, L, and the average hours worked per employed person, h, in a given time period t:

$$y_t = \frac{Y_t}{H_t} = \frac{Y_t}{h_t L_t} = (h_t)^{-1}\frac{Y_t}{L_t} \qquad (2)$$

This means that the level of labor productivity is determined not only by the size of workforce but also by the number of hours worked by each worker on average. The industry or occupation origin of Equation 2 can be shown as:[4]

$$H_t = h_t L_t = \sum_{i=1}^{m} h_{i,t} L_{i,t} \quad (i = 1, 2, \cdots, m) \qquad (3)$$

Suppose that in any industry i there could be a formal sector A, which is regulated by labor laws that institutionalize standard working hours per week and is monitored through the regular statistical system, and an informal sector B in which workers are engaged in temporal or seasonal activities or subsidiary employment with irregular working hours, and may only be obserble in censuses or occasional surveys, as given in Equation 4:

$$H_t = h_t L_t = \sum_{i=1}^{m} h_{i,t} L_{i,t}^{A} + h_{i,t} L_{i,t}^{B}) \qquad (4)$$

However, if there is virtually no information for B at the industry level, the situation we face becomes:

$$H_t = h_t L_t = \sum_{i=1}^{m} h_{i,t} L_{i,t}^{A} + h_t L_t^{B} \qquad (5)$$

Now our key question is how to breakdown the employment in B by industry so that total hours worked at industry level can be fully measured. To this end, we have to rely on some assumptions so that occasionally available information from censuses and special surveys can be used to estimate industry-specific hours worked.

Consider that a temporary or seasonal worker does not work for full time on average and a full-time but by-employed worker is engaged in multi activities in a given time period. Assuming that all industries whthin the B sector are homegenous in employment, the problem can be simply solved if we have information on average hours worked per worker in this sector. As long as some workers are not fully engaged in gainful activities, no matter how many jobs in the form of by-employment, an average worker of the B sector is not be fully engaged. This means that a fraction of a working day must be idle which can be defined as a share of the total hours available for work (λ). Thus the total actual number of hours worked in the B sector can be defined by the following formula:

$$H_t^B = (1-\lambda) h_t L_t^B \qquad (6)$$

Besides, if we also consider that the

employment in numbers is not fully covered and it is only related to B as usually the case, Equation 6 can be revised to capture such an effect γ, then the total number of hours worked in B becomes:

$$H_t^B = \frac{(1-\lambda)}{(1-\gamma)} h_t L_t^B \tag{7}$$

Apparently, Equation 6 is a special case of Equation 7, when $\gamma=0$. Note that both λ and γ are time-variant.

Now turn to the breakdowns of industries. We can reasonably assume that workers who are engaged in temporal or seasonal activities or subsidiary employment are usually low-skilled and less-educated if not uneducated at all, and tend to work in more labor-intensive manufacturing and services. What we need is to establish industry-specific weights that take into account such industries through surveys. Since conceptually $H_t^B=\sum_{i=1}^m H_{i,t}^B$, and hence the share of hours worked in the ith industry can be defined as $\omega_{i,t}^B=H_{i,t}^B/\sum_{i=1}^m H_{i,t}^B$, the number of hours worked for the B sector can be estimated as:

$$H_{i,t}^B = \omega_{i,t}^B \left[\frac{(1-\lambda)}{(1-\gamma)} h_t L_t^B \right] \tag{8}$$

Finally, to entirely account for the economy-wide hours worked, considering both the A and B sectors of each industry, we have the following equation:

$$\begin{aligned} H_t &= \sum_{i=1}^m H_{i,t} \\ &= \sum_{i=1}^m \left[h_{i,t} L_{i,t}^A + \omega_{i,t}^B \left[\frac{(1-\lambda)}{(1-\gamma)} h_t L_t^B \right] \right] \end{aligned} \tag{9}$$

The above conceptual framework will guide the data work in this study focusing not only on the supply-side parameters, α and ρ, and the demand-side parameter, μ (Equation 1), but also on the consistency-adjustment parameters, h, λ, and r, as well as ω for gauging proper industry breakdowns. This is however by no means to say that we can always obtain such parameters in a satisfactory way. Nonetheless, when such pursuits are data constrained, the likely biases will be discussed to highlight the research priorities in this field.

4. Constructing the 1915 Benchmark

The 1910s is the beginning of the Republican period. It was the period in which China's modern statistical system also began as marked by China's two modern censuses, i.e. the 1909–1911 Qing census and the 1912 census under the Beiyang (Beijing) Government, and by the establishment of the official agricultural, industrial and commercial statistics in 1912 (see Data Source Appendix). Before the outbreak of the civil war in 1917, this was also a long-awaited stable and peaceful period, which encouraged private investment. There was no major economic shock during this time. The outburst of World War I (1914–1918) somehow turned out to be fairly positive to the Chinese economy because the substantial decline of the exports from the West provided a golden opportunity for the expansion of China's exports.

The quality of China's first two population censuses has been questioned for over one century (see a review in Hou, 2001: 56–62, 334–340). Although they provided a better quantitative base than all previous estimates based on the *hukou* records, they were by no means more reliable. Besides, they did not collect any meaningful data for the occupational distribution of the workforce. For this reason, we have to rely on industry employment data from the Ministry of Agriculture and Commerce (MAC) annual statistical tables (Data Source Appendix) and other less systematic data to fill sectoral gaps. The MAC statistical tables were compiled for ten consecutive years from 1912 to 1922. The year 1915 is chosen as the benchmark for this early Republican period because of the worsening coverage problem of the MAC tables from 1916 onwards

Table 2. Reconstructed Occupational Survey Data of the Chinese Rural Workforce, 1928–1933

(Thousand persons)

	Male			Female			Total
	National	North	South	National	North	South	National
Reported aged 7 and over	48,235	27,085	21,150	42,615	24,078	18,537	90,850
Adjusted aged 15 and over[1]	44,644	25,069	19,576	39,443	22,286	17,157	84,087
Adjusted aged 15 to 64[1]	36,349	20,411	15,938	32,114	18,145	13,969	68,464
Active Population[2]	40,179	22,339	17,840	15,158	8,218	6,940	55,337
(Participation rate %[3])	(110.5)	(109.4)	(111.9)	(47.2)	(45.3)	(49.7)	(80.8)
Agriculture	28,863	15,673	13,190	9,223	5,177	4,046	38,086
Mining	25	25	0	0	0	0	25
Manufacturing	1,392	682	711	1,219	397	822	2,611
–Factory	*1,143*	*524*	*619*	*22*	*0*	*22*	*1,165*
–Handicrafts	*250*	*158*	*92*	*1,196*	*397*	*800*	*1,446*
Transportation	438	211	227	17	0	17	455
Wholesales & retails	2,063	1,429	634	84	45	39	2,147
Professional services	3,965	2,346	1,619	316	178	137	4,281
Government services	334	231	104	0	0	0	334
Unidentified[4]	52	33	20	123	89	34	176
Total gainfully employed	37,133	20,629	16,504	10,981	5,886	5,096	48,115
Unemployed	3,046	1,710	1,335	4,177	2,333	1,844	7,222
(Unemployment rate %)	(7.6)	(7.7)	(7.5)	(27.6)	(28.4)	(26.6)	(13.1)
	PST Distribution (total=100)[4]						
Primary	77.7	76.0	79.9	84.0	88.0	79.4	79.2
Secondary	3.7	3.3	4.3	11.1	6.7	16.1	5.4
Tertiary	18.5	20.7	15.8	4.9	5.3	4.5	15.4

Sources: Reconstructed based on the sample survey data used in Notestein and Chiao (1937: Table 9) with adjustments in line with the principles of the framework proposed in this study. See text for explanations.

Notes: 1) Estimated using age structure from a life-table constructed by Chiao (1945) based on five surveys covering from 4 to 11 provinces in the early 1920s. 2) Economically active population is estimated based on the age group 15 and over rather than age 15–64 considering there were child laborers under age 15. 3) Participation rate is calculated as "active population" divided by the age 15–64 group. Rate greater than one indicate that there are laborers who are younger than age 15 and older than 64. 4) "Unidentified" category is included in the tertiary sector. See notes to Table 1 for PST definition.

especially in the case of mining and manufacturing industries. Also, none of the time points of 1912–1914 is selected because the decline in the industrial employment in that period looks unconvinced. A well established Chinese industrial output index based on factory products shows a strong growth at the same time (Chang 1969) suggesting a strong growth during that period. In what follows, we explain the data construction procedures for major sectors. The reader may also see data sources and problems in the Data Source Appendix.

Manufacturing

There are two sets of industry-specific employment statistics in the MAC annual tables, one for manufacturers with more than 7 employees, whereas the other for all producers. Logically, the former is included in the latter. In fact, the average number of employee per producer in 1912–1915 was 31 in the case of former and 5 in the case of latter. The range was from 15 to 65 and from 4 to 17 across industries, respectively. The two MAC tables are reconciled to identity workers in "factories" and "handicraft workshops". A note attached to the MAC tables explained that the broad MAC definition already included those working at home

(MAC 1919: 349). But we believe that there could be still a significant number of "handicraft workers" missing if judged by the average size of workshops and its range. The so-derived working place with 4 to 17 workers appears to be too big to be located within an averege household. We are then convinced that the MAC tables do not cover those engaged in handicraft as by-employment. We will get back to this matter when later working on the rural employment.

Mining

The MAC statistics for mining employment are less satisfied than the manufacturing data especially after 1916 when its coverage reduced and its measure of mining workers changed to "the unit of full labor/day" (MAC 1919: 487). Moreover, fluctuations in mining employment over the period 1912–1919 did not follow the output performance of coal and iron as suggested by the industrial production index constructed by Chang (1969: 60–61). We derive the annual change of mining employment with an assumption that 5 percent of mining output was attributable to the labor productivity growth (note that following Equation 2 the output growth can be defined as the sum of the growth of employment and the growth of labor productivity). We use 1914 as the yardstick which was a booming year for mining as confirmed by Chang's mining output index. This gives an estimate at 1.094 million for 1915 with a stable miner/engineer ratio of 232. Since the MAC statistics on mining came with regular reports on the number of engineers, technicians, clerks in both mining and prospecting, the estimates can be considered only for the modern mining industry. The results also imply that there were around 180–190 workers each operating mine[5] in 1914–1915. We can then safely judge that the MAC tables could have excluded a large number of native mining activities that were small in size and did not permanently hire their own engineers and technicians. This part of mining activities will be taken into account when we work on the rural employment below.

Hours worked

The MAC statistical tables report only annual days occupied without hours worked. It can be calculated that for the narrow category with an average size of 31 workers, the nation-wide annual average working days were 286.4 in the mid-1910s. This is almost the same as 285.6 days from Lieu's survey on factories across 16 industries in 1933 (Lieu 1937: 345–376). Lieu's survey data also show that the average number of hours worked was 12.8 per day. There were similar observations in the mid-1910s quoting 12–13 hours per day as the standard practice in textile and mining industries (Wang 1957: 1199–1203), as well as in handicrafts (Peng 1957: 728–733). We may hence argue that in term of hours worked per day and days occupied per year, our so-constructed MAC statistics for the mid 1910s are largely comparable with those constructed by Liu and Yeh (1965) for the early 1930s.

Transportation and post and telecommunication

China's employment for this sector is estimated using data from the first *China Labor Almanac* (CLA hereafter) published in 1928 compiled by Wang *et al.* (1928), supplemented by various transport indices compiled by Yan *et al.* (1955). The first CLA provides data on railway employees from 1916 to 1925 and post & telecommunication employees for 1925 only (1928: 27–32). The number of railway employees is extrapolated back to 1915 using the changes in freight and passenger service indices jointly with the changes in the number of equipment: locomo-

tives, freight cars and compartments. Besides, the number of seaman and inland shipping sailors, plus all kinds of supportive port workers, is estimated by the number of ships classified in tonnage assuming arbitrarily that 100 seamen (and supportive workers) per ship over 1000t, 50 per ship between 100t and 1000t, 30 per ship below 100t, and 20 per traditional ship without machine power (Yan *et al.* 1955: 228)[6]. Finally the number of post & telecommunication employees is assumed to move with that of the railway employees (Yan *et al.* 1955: 194, 207). For workers engaged in the traditional transport services, our estimation will be based on later estimates for the rural transport employment.

Other services

For the employment of other services, we rely on Liu and Yeh's (1965) study and Ou et al. 's (1947) national accounts, both for 1933, supplemented by small-scale surveys for the early 1920s. Modern professional services were still new and in small numbers even by 1933. Government services were however by no means small given China's long established bureaucratic system. Various locality surveys in the late 1920s also provided the number of public servants in rural China (e.g. Notestein and Chiao 1937: Table 9).

Military personnel cannot be ignored, but its size pre-1949 is not easy to gauge. Liu and Yeh (1965: 600) estimated it at a size of 3 million, which was adopted by Maddison as a constant over time (1998). Makino and Luo (2014: 60) estimated at 2.5 million for 1933 without discussing their assumption. Our estimation for Chinese military personnel as part of the public service employment follows the work by Ou *et al.* (1947) and Zhang (1947). Ou *et al.* showed that the value added of military service accounted for 54 percent of the national income attributable to government services (1947: 141). Based on an equal labor productivity assumption within the public sector, Zhang (1947: 89) estimated the number of employment in this sector at 5.4 million. Together with 0.71 million professional services (teachers, doctors, lawyers, journalists etc.), it accounts for 1.42 percent of the total population. Assuming this can be reasonably held, we can use it as a yardstick to estimate chaina's military personnel for 1915 at about 2.7 million. Considering the post-1949 military personnel estimates in Wu (2014b), this is about half of the size of the early 1950s.

Next, we assume that the urban construction industry grew by about 9 percent per annum in 1912-1933 as the factory sector given by the Chang index (1969: 60-61), of which 4 percentage points were attributed to the growth of labor productivity and 5 percentage points to the growth of employment considering the labor-intensive nature of this industry. We will further gauge the economy-wide employment in these services when dealing with the rural economy.

Occupational distribution of the rural workforce

In the absence of the national statistics on the Chinese rural employment, we use a top-down approach beginning with our reconstructed demographic data that are supported by available surveys of various sources. China's rural population for 1915 can be derived from the estimated total population and its urban share (see the section for time-series estimation). Its 15-64 working-age share of 0.640 is interpolated between two age structures, 0.652 and 0.634, from the reconstructed 1912 census (Hou 2001: 326) and Buck's land and population survey for 1929-1933 (Buck 1937, reconstructed by Chiao 1945: 58-59) respectively.

However, it is not easy to gauge the rural labor participation rate of the working-age because available surveys not only suffer

from sampling biases but also tend to count all who worked at the time of survey including working children whose effective or adult-equivalent hours worked were never clear. Table 2 is a reconstructed occupational distribution table based on Notestein and Chiao (1937: Table 9) using Buck's 1929–1933 land utilization survey data. It shows that China's rural participation rate could be 0.808 (the working-age population=1). Considering that farm households worked very hard in order to maintain their basic living standard, we assume that China's total participation rate could have been close to the maximum recorded in history, i.e. around 0.85[7]. We start with a hypothetical rate of 0.830 for the total economy and 0.800 for the urban area (less than national due to higher income and more education opportunity for the age 15–64 group). Assuming the urban to rural employment ratio as 1: 9, we then arrive at a rural participation rate of 0.833. This rate is given to estimate the economically active population in the rural China in 1915.

We gauge the rural occupational structure in 1915 based on Notestein and Chiao (Table 2), supplemented by Liu and Yeh's work using the same survey data. We first make three adjustments to the general accounts in Notestein and Chiao: 1) adjust the working-age population from aged 7 to aged 15 and above and estimate the 15–64 age group, 2) assume that the by-employed with farming as their major activity allocate 70 percent of their time to farming and 30 percent to other activities, and 3) redistribute the "unclassified" workers to services except "government", using the existing structure of these services. According to Buck (1937: 293–294), an average able-bodied man only engaged in farming worked for 85 percent of their time over a year. Nevertheless, we choose not to make any "full-time adjustment" because this phenomenon has been typically observed not only historically but also at presentl. With a much higher degree of mechanization than before, chinese farmers today still have different levels of time utilization over seasons. Perhaps, an easier way to "solve" the problem is to conceptually consider this type of underemployment as *underproductivity*.

With a careful examination of the reconstructed Notestein-Chiao estimates (Table 2), we can consider a proper occupational distribution that is plausible for 1915. For this purpose, we need to compare the estimates with those by Liu and Yeh (1965) based on the same data used in the former study. The comparison, however, has come out with a very distinct structural difference between the two, especially in the agricultural share of the rural employment (=1), that is, 0.792 in the case of Notestein and Chiao and 0.965 in the case of Liu and Yeh. It is never clear how the original data from Buck's land utilization survey was handled in these studies. Converting the by-employed to single occupation-based measures should not cause such a big difference. One possible reason is that aiming at constructing China's national accounts that were considered compatible to its labor accounts, Liu and Yeh had to adjust Buck's land survey data for a economy-wide full coverage and therefore take into account Ou's previous work that used labor compensation to gauge the sectoral value added (Ou *et al.* 1947).

It is difficult to be convinced by the Liu-Yeh estimates that there was still 97 percent of the rural workforce engaged in agriculture in 1933 given rich anecdotal accounts on how rural households worked in off-farm activities to support their basic, if not subsistence, living. However, the 79 percent provided by Notestein and Chiao appears to be too low to be true. If taking an average of the two, it will be close to 90 percent. Accepting the view that by the early 1930s the Chinese rural economy had basically remained unchanged

Table 3. Estimated Chinese Employment by Occupation, circa 1915

(Thousand persons)

	Male	Female	Total
Population	233,442	200,335	433,777
–Urban	*29,588*	*24,278*	*53,865*
Working-age 15–64	139,525	138,200	277,726
Participated[1)]	138,967	91,545	230,512
Agriculture	78,912	82,899	161,811
Mining	1,365	274	1,639
–Modern mining	*821*	*274*	*1,094*
–Native mining	*545*	*0*	*545*
Manufacturing	10,820	2,579	13,399
–Factory	*484*	*163*	*648*
–Handicrafts	*10,336*	*2,415*	*12,751*
Utilities	1	0	1
Construction	1,822	0	1,822
Transportation	4,602	44	4,647
–Modern	*515*	*7*	*522*
–Traditional	*4,087*	*38*	*4,124*
Wholesales & retails	5,435	1,095	6,530
Professionals	454	263	717
Government	4,875	567	5,442
–Military personnel	*2,585*	*136*	*2,721*
Other services	312	1,742	2,054
Total gainfully employed	108,598	89,463	198,061
Unemployed	30,369	2,082	32,451
(Rate of unemployed, %)	(21.9)	(2.3)	(14.1)
	PST Distribution (total=100)		
Primary	74.0	93.0	82.5
Secondary	11.6	2.9	7.7
Tertiary	14.4	4.1	9.8

Sources: Author's estimation. See text for explanations.

Note: 1) Including participated population who are younger than age 15 and older than age 64.

from the beginning of the twentieth century (Feuerwerker 1983), we assume 90 percent of the rural workforce was still engaged in agriculture in 1915. In the case of Liu and Yeh's estimates, this assumption is equivalent to relocating about 10 percent of the agricultural labor force to the non-agricultural sectors. The so-increased nonagricultural employment was redistributed by the industry structure as given in Notestein and Chiao (Table 2) with a slight downward adjustment to transport and commerce and an according upward adjustment to the rest of the services including government and professional services.

Final Adjustments for full coverage

To arrive at the full-coverage results in Table 3, we need further adjustments to fill the gaps left so far. First, we use our estimates for the rural area and the rural-urban employment distribution, 0.89: 0.11 for 1915, to gauge the employment of economy-wide commerce, road transport and other services. We assume that the so-estimated employment of the "road transportation" for the urban area can represent the employment of the highway transport industry that has not been covered in our estimation for the modern transport. Second, we treat mining employment estimated for the rural economy as part of the native mining industry and add it to the earlier estimate based on the MAC statistics tables. Third, we add the estimated handicraft employment of the rural economy to the same category of the MAC tables. However, we assume the estimated factory workers for the rural economy are already covered by the MAC tables. Besides, we add the so-estimated rural construction employment to what we have estimated for the urban economy. Lastly, the industry-specific occupational distribution between male and female follows Notesteign and Chiao in the case of the rural economy (Table 2) and Liu and Yeh (1965) in the case of the national economy.

We show that by 1915, there was still 82.5 percent of the Chinese labor force engaged in the primary sector, with the rest engaged in the secondary sector (7.7 percent) and the tertiary sector (9.8 percent). After more than a half century of China's industrialization, China's traditional industries had appeared to be still dominant as reflected by the occupational changes. The "modern share" accounted for merely 5 percent in manufacturing and 12 percent in transport. Studies have shown that many of

Table 4. Reconstructed Chinese Population and Employment by Occupation, 1933

(Thousand persons)

	Male	Female	Total
Total population	271,138	241,846	512,984
–Urban	*44,654*	*39,989*	*84,642*
Working-age 15–64	165,050	160,073	325,123
Participated[1)]	170,020	109,074	279,094
Agriculture	112,122	80,340	192,462
Mining	1,116	39	1,155
–Modern	*739*	*39*	*777*
–Native mining	*377*	*0*	*377*
Manufacturing	10,739	11,010	21,748
–Factory	*835*	*556*	*1,391*
–Handicrafts	*9,904*	*10,453*	*20,357*
Utilities	49	0	49
Construction	7,166	0	7,166
Transport	7,710	3,590	11,300
–Modern	*986*	*539*	*1,524*
–Traditional	*6,725*	*3,052*	*9,776*
Wholesales & retails	13,768	1,112	14,880
Professionals	538	311	849
Government	5,768	671	6,439
–Military personnel	*3,059*	*161*	*3,220*
Other services	5,301	1,003	6,303
Total gainfully employed	164,276	98,076	262,352
Unemployed	5,744	10,999	16,742
(Rate of unemployed, %)	(3.4)	(10.1)	(6.0)
	PST Distribution (total=100)		
Primary	68.9	82.0	73.8
Secondary	10.9	11.2	11.0
Tertiary	20.2	6.8	15.2

Sources: Author's estimation. See text for explanations.

Note: 1) Including participated population who are younger than age 15 and older than age 64.

the handicraft manufacturers had been much better equipped and managed in the 1910s–1920s than in the late of the nineteenth century (Shi 2005: 251–265), but their labor-intensive nature had not changed. The slower than expected pace of labor shift from the traditional to the modern sectors reflects China's technological choice when facing a tremendous pressure of surplus labor that suppressed the wages of the modern sectors.

5. Revisiting and Revising the 1933 and the 1952 Benchmarks

The 1933 benchmark

To ensure consistencies in concept, coverage and classification over the period in question, we revisit the estimates of occupational distribution by Liu and Yeh (1965) for the 1933 benchmark together with their source data in Buck (1937) as well as related studies in Ou *et al.* (1947) and Zhang (1947). As above discussed, one of our main concerns is the likely overestimation of the agricultural employment in Liu and Yeh (1965: 191) which accounted for 97 percent of the rural workforce and 79 percent of the national workforce.

Liu and Yeh followed Buck's man-labor adjustment to those engaged in by-employment but did not adjust the workforce engaged in single occupation. We make an arbitrary 10-percent downward adjustment to align the 1933 benchmark with that of 1915. We also make an upward estimate of Liu and Yeh's working-age population from age 7–64 and above to age 15–64. The grouping of age 7–64 in Liu and Yeh follows Buck. In fact, as pointed out by Liu and Yeh, workers in Buck's survey mostly fell in the category of 12–64 rather than 7–64 (1965: Table 55). We assume that only one fifth of those aged 12–14 were active in the labor market, and include them as participated laborers. This makes the resulted participation rate is higher than that using the standard measure of working-age population 15–64, especially in the case of male workers (171 against 165 million, Table 4).

For the across-industry occupational distribution we have basically maintained the existing structure of Liu and Yeh. For more detailed industry breakdowns we refer to those in the newly constructed 1915 benchmark allowing more laborers engaged in handicraft manufacturing, native mining and traditional transportation. We believe that this treatment has to some extent made the measured agricultural employment more "effective" while addressing the increasing

importance of the off-farm activities in sustaining farm household subsistence needs. This is an issue that is subject to further investigation with more convincing evidence.

We have arrived at a smaller but arguably more plausible share of the primary sector than that of Zhang (1947), Liu and Yeh (1965) and Makino and Luo (2014) as compared in Table 1. The 1933 occupational distribution across industries also looks better aligned with that of 1915. Our final results show a PST ratio around 74: 11: 15 (Table 4), compared to 79: 7: 14 estimated by Liu and Yeh (Table 1).

The 1952 benchmark

Two unavoidable problems when dealing with the 1952 benchmark are caused by the missing of military personnel and a serious structural break in 1990 of the official employment series. Our new adjustment in this study adds Wu's earlier estimate of 5.7 million armed forces for 1952 (Wu 2014b: 275–276). As discussed in Maddison and Wu (2008), the structural break is represented by an astonishing 17 percent or 94.2 million jump in 1990 in the official employment series. This new total is available with three broad-sector breakdowns (primary, secondary and tertiary) linking to the same breakdowns prior to 1990, but not with the estimates at industry level for 16 sectors. The existing industry level estimates, which follow the pre-1990 tradition, fall short of the new estimate of total employment in 1990 by 80.1 million. The post-1990 data series is built on this new level of total employment, hence sustaining a continuous gap with the underlying trend based on the pre-1990 data series. The choice of approaches to solving the problem inevitably affects the beginning time point of the series 1952.

My earlier attempts (Wu 2010, 2014a and 2014b) using China's 1982 census and 1987 one-percent sample survey confirmed that the 1990 structural break could have "appeared" earlier in 1982 if the 1982 census data were used to campare with the official regular annual estimates. It was caused by the insufficient coverage of the labor planning system that ignored informal activities outside the system. Assuming that this only began in the early 1970s when the government relaxed its control over small, collective-run enterprises, especially those in the rural area, to lessen the pressure of the rising surplus labor, I backward extrapolated the 1982 census-based adjustment to 1971, anchored the series on 1970. This absorbed the additional number of employment that was caused by the break.

In the present study, after the reexamining the official series of the period 1949–1952, which was untouched in Wu's eariler adjustment (2014a and 2014b) a strange PST ratio has appeared as 90: 5: 5. This is not only misaligned with what already observed in the prewar China, but also appears to be against the development experiences of low income countries suggesting that the share of the tertiary sector engagement should be somewhat twice that of the secondary sector. The estimates by Kuznets gave a PST ratio of 80: 6: 14 (1957: 23–27). We are thus convinced that forcing the 1982 census-based new series to fully absorb the positive break back to 1971 is unrealistic and ignored a significant amount of informal employment using the traditional or native technologies prior to the 1970s. In the present work, instead of absorbing the break along a new 1970–1982 trend, we allow the series to change along with the official series from 1982 back to 1949. This raises the total number of employment by 14.7% for 1949. To check it, we find that the result of such a backward adjustment is still slightly smaller than my estimate for 1948 (by 1.2 percent). Given the shock of the 1949 revolution that might substantially reduce employment, this new result looks

Table 5. Reconstructed Chinese Population and Employment by Occupation, 1952
(Thousand persons)

	Male	Female	Total
Total population	297,526	275,634	573,160
–Urban	*35,805*	*33,170*	*68,975*
Working-age 15–64	172,018	163,120	335,138
Participated[1)]	164,583	102,628	267,211
Agriculture	105,971	74,421	180,391
Mining	1,917	255	2,172
–Modern	*1,142*	*255*	*1,396*
–Native mining	*775*	*0*	*775*
Manufacturing	10,551	10,367	20,918
–Factory	*6,971*	*1,556*	*8,526*
–Handicrafts	*3,580*	*8,811*	*12,391*
Utilities	197	10	207
Construction	2,536	27	2,563
Transport	3,165	458	3,623
–Modern	*2,283*	*67*	*2,350*
–Traditional	*882*	*392*	*1,273*
Wholesales & retails	9,959	5,136	15,095
Professionals	2,535	1,075	3,610
Government	7,125	1,165	8,290
–Military personnel	*5,130*	*570*	*5,700*
Other services	4,157	2,144	6,300
Total gainfully employed	148,113	95,057	243,170
Unemployed	16,470	7,571	24,041
(Rate of unemployed, %)	(10.0)	(7.4)	(9.0)
	PST Distribution (total=100)		
Primary	72.8	78.6	75.1
Secondary	9.0	10.9	9.7
Tertiary	18.2	10.5	15.2

Sources: Author's estimation. See text for explanations.
Note: 1) Including participated population who are younger than age 15 and older than age 64.

much more plausible than my previous estimates. For the 1952 benchmark, the *additional* employment (=1) is allocated as follows: 0.368 to manufacturing, 0.265 to transportation and commerce, 0.179 to other services, and 0.188 to agriculture.

Our final results show a PST ratio 75: 10: 15 for 1952 (Table 5), which appears to be only slightly different from that of 74: 11: 15 for 1933 (Table 4). Although we cannot rule out errors in the estimation, this seemingly structural stagnation over the two decades, or even a slight reversal between the primary sector and the secondary sector, could be largely attributable to the impact of the wars and the revolution which restructured China's industrialization towards more capital-intensive development.

6. Interpolating between the Benchmarks

Working on these occupational benchmarks naturally makes one think of the connections between the benchmarks at both the sectoral and national levels that could be affected by the underlying macroeconomic changes, external shocks (market or political) and demographic transitions. Such connections cannot be addressed until one starts making an attempt to construct a time series between the constructed benchmarks. This is nevertheless a grave challenge because it is by no means a simple interpolation of individual indicators, but requires a systematic endeavor to explicitly establish coherent connections between the supply-side and the demand-side indicators.

Supply-side interpolations

The supply side indicators include basically the total population by gender and age, and hence the working-age population. It would be ideal if these data are also cross-classified by occupation and by industry, but none of the Chinese population censuses for the period in question could meet such a standard. For the 1915 benchmark, the 1909–1911 Qing census and the 1912 census cannot satisfy our needs. The responsible authorities not only failed to collect data from all localities at the same time of the census, but also introduced serious mistakes when later trying to fill the missing localities with data of other sources.

We first need to estimate the size of the Chinese population for the starting point 1912 or another time point that is close to it. There were two studies that tried to repair the 1912 census but came out with different estimates, i.e. 419.6 million from the Ministry of Internal

Affairs (MIA) (1936: 402–404) and 405.8 million from Chen (1934: 16–18). Some researchers accepted work by Chen but others proposed further adjustments (e.g. Zhao and Xie 1988: 480–501). Hou reasonably criticized MIA and Chen for obvious mistakes (2001: 248–254). Based on indices constructed by Nanking University, Perkins derived an estimate for 1913 at 430.4 million (1969: 212). However, the highest estimate is 436.0 million for 1910, made by Cao (2001: 831) after taking into account severe losses caused by famines and wars. Maddison gave a high appraisal to Perkin's estimate. His 423.0 million for 1910 appears to be a backward extrapolation from Perkin's 430.4 million for 1913 at an annual rate of 0.6 percent (1998: 167–169). This Perkins-Maddison hypothetical figure for 1910 is then used as our starting point. Our interpolation is conducted between the following benchmark estimates, 1910, 1919, 1928, 1936, 1946 and 1953, which anchor the pursued time series for the period 1912–1952.

For the age structure with a gender distribution of each age we rely on Hou's assessment of the available census and survey data (2001: 282–302). He convincingly points out that the data of various official sources before the war tended to significantly exaggerate the actual male-female ratio by at least 10 points or over 120 (female = 100) compared to around or below 110 using more closely monitored surveys after the war, and more importantly observed from the first post-revolution census in 1953. The exaggeration of the sex ratio was caused by the serious underestimation of females of all ages. Moreover, underreporting children below age 10 for both sexes is another serious problem. Normally, for a very young population as in the case of China in the first half of the twentieth century, this age group should be accounted for above 20 percent of the total population. The result of the 1912 census is however only 13.5 percent compared to 24.6 percent estimated by Chiao (1945) based on five large-sample rural surveys in the 1930s and 1940s, and 26.8 percent based on China's 1953 census (Hou 2001). We take into accounts these evaluations when establishing the benchmark working-age 15–64 group for 1912, 1930, 1947 and 1953.

Demand-side interpolations

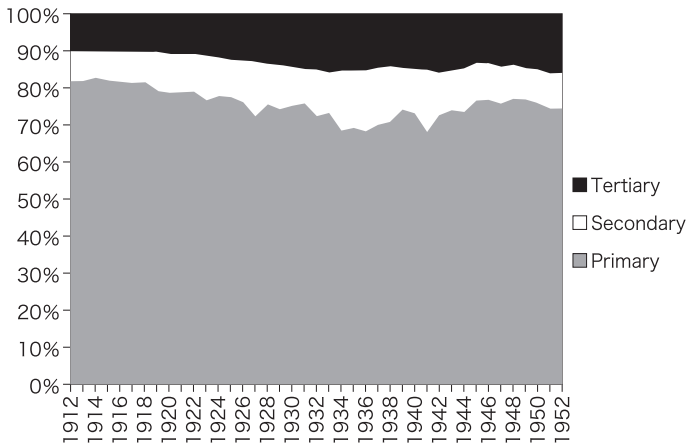
Our demand-side interpolations are conducted between the established benchmarks as presented in Tables 3, 4 and 5 with the following assumptions: 1) the change of the employment engaged in agriculture and handicrafts is assumed to move with that of the working-age population; 2) the change of the employment engaged in mining (both modern and native), utilities and factories is assumed to go with the change of the "modern" industrial output as measured by the Chang index (1969: 60–61); 3) the change of the employment engaged in construction is assumed to follow that of the railway development measured by the annual constructed length (Minami and Mokino 2014: CD Table 4. F. 1); 4) the change of the employment engaged in transportation (both modern and tradition) and commerce (both wholesale and retail trade) is assumed to move with that of the workforce engaged in non-service activities; and finally 5) the change of government and professional services is assumed to be in line with the development of the modern industry that is also given by the Chang index. The results are depicted as the PST shares of the total employment in Figure 1.

Measuring the urbanization of the population

Our time series construction of the Chinese urban population takes a supply-demand hybrid approach. That is, the source of the urban population is based on the total population estimated, but in the absence of

Figure 1. Estimated Occupational Distribution of Chinese Employment, 1912–1952

(Total employment=100 percent)



Sources: Author's estimation.

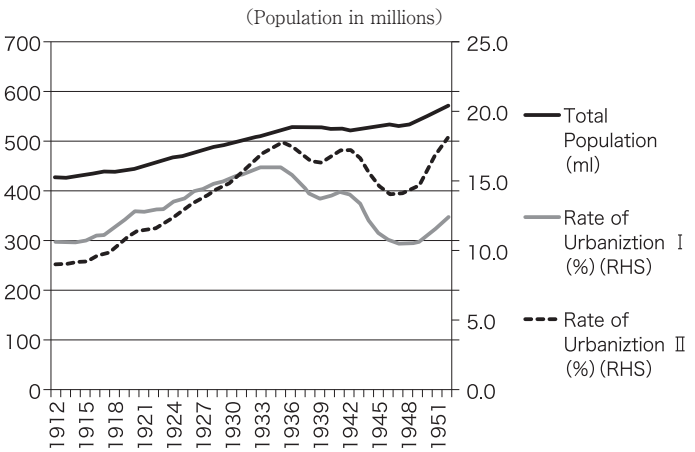
statistics on the urban population, we have to rely on the demand-side information, that is, the movement of the urban employment which has to be anchored by some exogenously given benchmarks for the urban share of the total population.

Most of the previous studies never seriously dealt with the quantitative measure of the urbanization in the light of China's industrialization. Attempts to estimate the size of Chinese urban population faced not only the problem of the lack of necessary data, but also the problem of the lack of agreement on how to define "urban population". The size threshold for "urban population" used by researchers ranged from 2,500 to 100,000 inhabitants. There were two estimates used the 100,000 threshold which arrived at similar estimates for the 1910s, i.e. 18.1 million by CCC (1918, quoted in Hou 2001: 477–479) and 17.1 million by Perkins (1969: 190–295), which was around 4 percent of the total population. Hou lowered the threshold to 25,000 to add the population of the cities with 25,000 to 99,999 inhabitants to the CCC's estimate. He arrived at 31.8 million, implying an urbanization rate of 7.3 percent for 1918. In Hou's view (2001: 483), the most reasonable and reliable estimate was made by Notestein and Chiao (1937: 365) for 1929–33 in three distinguished categories: cities, market towns and farm villages. Their estimate of the city population accounts for 10 percent and of the market-town population accounts for 11 percent of the total population.

We accept Hou's assessment of the Notestein-Chiao estimates with one adjustment. In their estimates, market towns also included agricultural population under the jurisdiction of towns. If assuming that only half of the town population can be classified as the real urban inhabitants, China's urbanization rate should be 15.5 percent for 1929–33. Taking 15.5 as the median number for 1931 and allowing for a maximum change of one percentage point within this period, the rate should be 15.0 for 1929 and 16.0 for 1933. Furthermore, the official estimates of the urban population for 1949–1952 should also be considered. Hou assesses two sources of official data for 1949 that suggest 10.5–11.5 percent of the urbanization rate (2001: 485–486). We find that based on the official population statistics, China's urbanization rate was 12.5 percent in 1952. If the official data are true, compared to 16.0 percent in 1933, this means that China underwent a significant de-urbanization process during the wartime and not yet fully recovered after the civil war.

In this study, we propose two scenarios to gauge the size of China's urban population, with or without taking into accounts the official estimates for 1949–1952. Both approaches require the time series of the "urban employment". It is constructed as the sum of the workforce engaged in modern mining, manufacturing, utilities and modern transportations, plus the half size of those engaged in native mining, handicrafts, and the rest of services, assuming the other half

Figure 2. Estimated Chinese Population and Urbanization Rate, 1912–1952



Sources: Author's estimation.

located in rural areas. The workforce engaged in construction is excluded in this estimation because of its high volatility. The idea is to gauge the change of the urban population by the change of the urban employment. To reduce the effect of the volatility of the latter on the formar which should not be very volatile, we use three-year moving average of the latter. Note that in both Scenarios I and II, the revised Notestein-Chiao median estimate of 15.5 percent is used to control for the level of 1931. However, in Scenario I the official urbanization rate of 10.6 percent is added to control for the level of 1949. There is another difference between the two scenarios. In Scenario I, the series was gauged with the relationship of the levels (i.e. the ratio of the total population to the urban employment), whereas in Scenario II, the series was gauged by the change of the urbanization rate as a result of the change of the urban employment. The results are depicted against the estimated size of the total population in Figure 2.

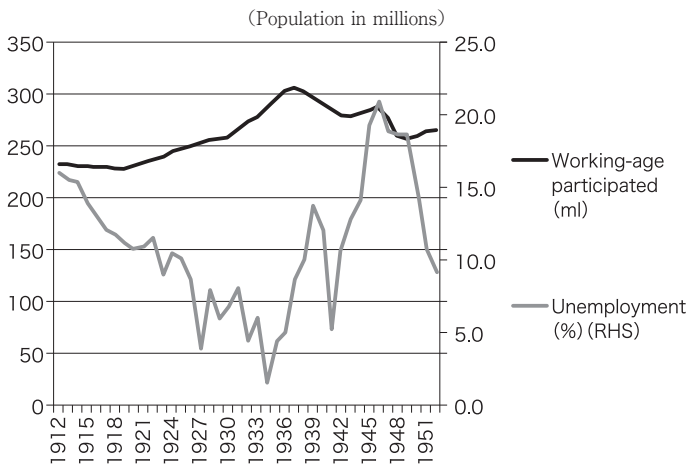
As shown in Figure 2, compared to Scenario I, Scenario II suggests a much faster urbanization process and a much quicker postwar recovery. If using the official estimate for 1949 to control this time-series exercise, we can see that it has a significant bearing on the results. However, assuming China's urbanization by 1952 had recovered to the prewar level as suggested by Scenario II may emphasize the power of the market despite the revolution. On the other hand, if arguing that China was still a half way to go to reach its prewar peak as suggested by Scenario I, it implies that the radical regime change in the wake of the Communist Revolution in 1949–52 had Signifcantly obstructed the recovery. In my view, since Scenario II is completely based on the change of the urban employment, it is hence more coherent with other variables within the system.

Integrating the whole system

We now face a critical issue of how to integrate the two sides of the estimates to explicitly establish the coherence between the population and the occupation accounts. The key is inevitably a proper measure of the economically active population. Logically, with our estimated total employment, what we need is just the statistics on the unemployed. Yet, this is difficult because there was simply no such statistics. Another practical problem is that there was no legal working-age implemented. If there are a large number of workers below age 15, the standard use of the age 15–64 category as the working-age population may exaggerate the labor participation rate (LPR) (see the case of 1933 in Table 4). Yet, this has to be ignored when we gauge the benchmark LPRs because of the lack of necessary data.

We first need to establish the LPRs for the benchmarks. For 1915, we exogenously set up a LPR using the available 1929–1933 land utilization survey data, which is 0.833 and implies an unemployment rate of 14.1% (Table 3; see explanations in the text). For

Figure 3. Participated Working-age Population and the Unemployment Rate, 1912–1952



Sources: Author's estimation.

1933, considering the macroeconomic situation, we assume that the unemployment rate was 50 percent above the natural rate of 4 percent, i.e. 6 percent which implies a plausible LPR of 0.858.

Next we search for two time points over the wartime and the early postwar period. They should be in distinct macroeconomic situations that make it easier to judge the estimated unemployment rate. Thus, 1949 and 1958 are chosen for this purpose. The year of 1949 was not only the last year of a three-year brutal civil war of an unprecedented scale in the twentieth-century Chinese history, but also the beginning of the communist regime that shocked the existing economic system. It should have experienced a very high unemployment rate. Although the year of 1958 is outside our time span, it is an indispensible yadstick for exploring the true labor participation rate in China. This year experienced an unprecedented full employment because of the Maoist feverish Great Leap Forward (1958–1959) that aimed to use a mass production campaign to boost the output of the economy in a very short period. We find such a full-employment situation implies a LPR of 0.800. If applying the same LPR to 1949–1952 and meanwhile applying a LPR of 0.830 to 1943, which is just the half way from the prewar 1933 (LPR=0.858) to 1952, then the so-derived unemployment rate for 1949 should be about 18 percent, which is consistent with the estimated unemployment rate of 19.1 percent on average over the civil war period 1946–1949.

The above explored LPRs are exogenously given for the benchmarks in the interpolations of LPR series from 1912 to 1952. The so-estimated China's working-age population participated in the labor market and the LPR series-implied unemployment rate series, given the estimated employment series, are presented in Figure 3.

7. Ending Remarks

In this paper we have made a preliminary attempt to fill a long-standing employment data gap in studies on China's long-run growth in the Republican era from 1912 to 1949. We focus on the construction of the 1915 benchmark and the revision of the existing 1933 and 1949–1952 benchmarks. From the demand-side, we provide the time-series estimates of the employment by occupation, and from the supply-side, we provide the time-series estimates of china's total and working-age population. The estimates of the two sides are then linked through the gauged labor participation rate to complete a conceptually coherent system of the chinese population and employment dynamics. Finally, we gauge the urbanization rate of the population implied by the change of the urban employment.

However, when writing down the end remarks here, I feel more like beginning rather than concluding an endeavor. There were important objectives that could not be

achieved due to limited time and resources. My grave challenge is how to go beyond the 1933 benchmark and establish new benchmarks, earlier and later. This is not only to explore the time series of the key population and employment variables in a coherent system, but more importantly to revisit the long-established dogma about the Chinese industrialization sololy based on the 1933 benchmark estimates. Researchers in this area have been stuck in the limited studies published in English using the easiest available data for 1933 or the 1930s.

Changes in the occupational distribution are not independent of changes in investment, income, as well as technology, and changes in one time point are not independent of changes in other times. I am convinced that tasking a systematic approach, considering the economic factors in time and space, is the key to filling data gaps. To advance the present study, the top priority is to systematically tap the rich historical materials in chinese compiled in a series of volumes and published in the 1950s that covered China's government-engineered industrialization from 1861 to 1949.

(Institute of Economic Research,
Hitotsubashi University)

Appendix: Sources of the Historical Data

Occupational data are usually found in the modern population censuses or surveys that covers economic attributes of the gainfully employed population. In the absence of such data, annual industry and labor statistics that are regularly produced by statistical authorities may be the sources of close substitutes. In the case of China, both qualified census data and regular statistics for occupational study are scant because of not only the late establishment of the modern statistical system but also the frequent interruptions of government administration due to wars and political chaos in our targeted period. Therefore, we have to explore implicit connections in the available unsystematic data of various sources. This study is mainly based on the available data introduced below.

National population censuses

China's first-ever population census was conducted in 1909–1911. Until then there was no any nation-wide survey on population. The widely appraised merit of the Chinese *hukou* (literally "number of mouths to be fed in a household") system in keeping historical records of household is overstated. The *hukou* system was developed mainly for taxation purpose. It lacked the needed demographic and occupational status of household members for today's occupational studies. Besides, it could be flawed due to the incentives of tax evasion and avoidance.

China's first census was conducted by the Ministry of Civil Affairs (MCA) of the Qing Government that prepare for the regime shift from the autocratic monarchy to the constitutional monarchy endorsed by Emperor Guangxu (of Aisin Gioro Zaitian). However, it aimed to construct quantitative accounts for the number of population and households rather than their basic social and economic status. Although it did not collect data on occupation, it provided some quantitative foundation for measuring employment. The census began in 1909 but was affected by the Republican Revolution in 1911. The Ministry of Internal Affairs (MIA) of the Republican Government (the Beijing Government, 1912–1928) continued the census and finally reported the census results in five volumes including five household-based tables and two population-based tables (Wang 1933 and 1935).

In addition to accomplishing the Qing census, the Beijing Government conducted a new census in 1912 to remedy the deficiencies in the Qing census. It ambitiously designed much more comprehensive questionnaires than the Qing census and included occupation questions as well (Lieu 1931), but ended up with almost half of the provinces failed to achieve the required tasks. Therefore, MIA had to use the Qing census data to fill the gaps (MIA 1936: 402–404). In 1928, the Nanjing Government (1927–1949) conducted a new population census that largely followed the questionnaire of the 1912 census. It was again unsuccessful because many localities were not cooperative (Hou 2001). Besides, there were

also two nation-wide "census-type" surveys in 1931 and 1936 respectively with the former concentrating on land and population at the city and county levels (DOS/MIA 1935) and the latter accounting for the potential eligible population for the forthcoming election (DOS/MIA 1936). They are considered better quality in terms of the coverage and timing (Hou 2001: 72–73).

Surveys and sub-national censuses

The most valuable sample survey in rural China was the one on land utilization and population carried out in 1929–1933 by the Department of Agricultural Economics, University of Nanking, under the leadership of professor John Lossing Buck. It was a five-year research project of the US-based Institute of the Pacific Relations (IPR), sponsored by the Rockefeller Foundation. It was the only comprehensive survey on the Chinese rural economy conducted in the first half of the twentieth century. It included occupational and industrial structures of employment and the time use of farm households covering 202,617 people from 38,256 families of 16,786 farms Living in 101 localities across 16 provinces. Its rich deta were published in three volumes (Buck 1937). Importantly, by-employment or subsidiary employment issues were paid attention and some useful information was collected.

In 1933, the National Resource Commission (NRC) of the Nationalist Government conducted a comprehensive industrial census, carried out by D.K. Lieu. It was the only survey of its kind before the war recording 2,435 Chinese-owned factories which employed 493,257 workers (Lieu 1937). However, the focus of the census was those "factories" using mechanical power and employing 30 or more workers as defined by China's first *Factory Law* in 1929. Although it also surveyed factories using mechanical power regardless the number of workers in each establishment, it still excluded a huge number of workers engaged in handicraft workshops not to mention many working at home with either full-time or part-time engagement (Liu and Yeh 1965: 429–431).

It should be noted that from the mid 1930s through the wartimes, there were at least ten sub-national censuses conducted in eight provinces that covered about four million people (Chen 1981: 12–23; BOS 1948: 50). The majority of the country-level censuses followed the design of the 1928 national census and collected employment data with occupational information[8].

Industrial and labor statistics

Regular industrial and labor statistics compiled by state agencies only began in 1912 but were interrupted by wars and radical changes in government. The employment data included in such statistics were categorized by industry rather than from occupation perspective as in the modern population census. The Ministry of Agriculture and Commerce (MAC) under the Beijing Government published annual "Statistical Tables for Agriculture and Commerce" for ten consecutive years from 1912 reporting the number of farm households, miners, industrial workers and members of commerce chambers[9]. These tables were regularly compiled in the same format although they suffered from geographic coverage problem after 1916[10]. Ministries in charge of transportation and telecommunication, healthcare and education also collected data on workers and professional personnel but lacked regular and consistent statistics. Often, there statistics were interrupted due to regime changes and wars. For example, statistics on railway workers began in 1916 but interrupted in 1926–1927 and stopped after 1937. Regular labor statistics appeared to be more difficult to be established. There were three *China Labor Almanacs* published in 1928, 1932 and 1933. However, these publications seldom constructed their own statistics but mainly collecting data from various sources on occupation, wages, working and living conditions regardless if they were contradoctory.

Statistics for 1949–1952

This is a "transition period" from the Republican to the Communist regime and to be served as an end-period anchor for this study. The new regime conducted its first population census in 1953, but unfortunately it mainly focused on a few demographic indicators rather than occupational and human capital information. The official industry-level labor statistics suffered from insufficient coverage problems and a serious break in 1990 in the employment series that affected the data series backward to 1949 (see the text for the solution to the

prololem). The data were compiled using local government reports through military-controlled administrative channels and offices in charge of different modern sectors of the economy, which tended to ignore employment in the traditional sectors and (deliberately) exclude military personnel (Maddison 1998). They were first released in 1987 by the Department of Social Statistics (DSS) of the National Bureau of Statistics (NBS) in a historical compilation for 1949–1985. Some of the raw data used in the compilation were later published in *Economic Archive Materials on Labor, Wage and Welfare* for the period 1949–1952 (CASS and Central Archives 1989).

For this "transition period", the present study is benefitted from estimates by Wu and his associates especially those on service employment adjusted for military personnel (Wu 2014a), sectoral employment estimates adjusted for the 1990 break (Wu 2014b) and hours worked based on various survey data (Wu and Yue 2012; Wu, Yue and Zhang 2015). The adjustment for the 1990 break takes into account information from China's 1982 and 1990 censuses as well as annual labor statistics from 1949 to 1990.

Notes

＊ I am indebted to constructive comments and suggestions by Osamu Saito as well as discussions at the IER regular faculty seminar, July 2015.

1) There are also a lot of anecdotal stories with piecemeal statistical information on subsidiary employment in rural China in the economic history literature (Peng 1962; Zhang 1957).

2) This project began in 1995 as Asian Historical Statistics Project (known as ASHSTAT) aiming to construct time series data for major Asian economies. It follows a model laid down by a 14-volume work to construct long-term economic statistics (LTES) for Japan since 1868 by Kazushi Ohkawa, Miyohei Shinohara and Mataji Umemura, also based in IER, Hitotsubashi University, from 1965 to 1988. The ASHSTAT series began its publication in 2008 and the China Volume (Minami and Makino 2014) is its third one in a row.

3) We assume the existence of homogeneous labor or the labor in constant quality. Such labor measure can be obtained by weighing all types of labor with their compensation, but beyond the scope of this study due to data constraint.

4) In this conceptual framework we use industry and occupation interchangeably assuming that this is also true in the real data albeit a strong assumption.

5) This is also termed as "mining area" or "mining district" in the MAC statistics without clear explanation (MAC 1919: 487).

6) This approach is subject to reassessment when there is more information available.

7) Using official population and labor statistics for 2000, and also officially defined "population of economically active", we obtain an economy-wide participation rate of 0.832. If we use a range of hypothetical unemployment rate because of likely underreporting, say from 5 to 7 percent, we can arrive at a rate from 0.853 to 0.872. Available records of major economies show that the postwar Japan once reached 0.740 in the mid 1950s and UK reached 0.783 in the late 1980s.

8) Include one county in Zhejiang in 1935, one in Yunnan in 1939, four in Yunnan in 1942, nine in Sichuan in 1943, and one in Xikang (now part of Sichuan) in 1944. Those in Yunnan were conducted by the Institute of National Conditions at Tsinghua University, and considered high standard in terms of the modern population census.

9) The Tenth (1921) Tables were published as a supplementary to the Ninth Tables (1920) (MAC 1924).

10) For details see notes in the Eighth (1919) Tables (MAC 1923: 223 and 456).

References

Bureau of Statistics (BOS), Republic of China (中華民國主計部) (1948)《中華民國統計年鑒》.

Buck, John Lossing (1937) "Farm Labor," in John Lossing Buck (ed.), *Land Utilization in China*: pp. 290–312, University of Chicago Press.

Buck, John Lossing (1937) *Land Utilization in China*. University of Chicago Press.

Cao, Shuji (曹樹基) (2001)《中國人口史，第五卷：清時 期》(*Chinese Population History*, Vol. 5: Qing Period), Fudan University Press.

China Academy of Social Sciences (CASS) (中國社會科學院) and Central Archives (中央檔案館) (1989)《中華人民共和國經濟檔案資料選編：勞動工資和職工福利卷》(Selection of Economic Archive Materials of PRC: Labor, Wage and Welfare)，北京：中國社會科學出版社.

Chang, John K. (1969) *Industrial Development in Pre-Communist China, 1912–1949*, Edinburgh University Press.

Chen, Changheng (陈长蘅) (1934) "人口," in Ministry of Industry and Commerce (MIC) (中華民國實業部) (ed.)《中國經濟年鑒》(*China Economic Yearbook*).

Chen, Da. (1981)《現代中國人口》(*Contemporary Chinese Population*)，天津人民出版社.

Chiao, Chi-ming (1945)《中國農村社會經濟學》(*The Social Economics of Rural China*)，上海：商務出版社(Commercial Press).

China Continuation Committee (CCC) (of the National Missionary Conference, Shanghai), cited in Hou (2001) "…"

Department of Statistics (DOS), Ministry of Internal Affairs (MIA) (中華民國内政部統計處) (1935)《全國各縣市土地人口調查》(*National Land and Population Surveys in Counties and Cities*).

Department of Statistics (DOS), Ministry of Internal Affairs (MIA) (中華民國内政部統計處) (1936)《全國選挙區戶口統計》("Household Statistics of National Election Districts"), *Internal Affairs Statistics Quarterly*, No. 1.

Feuerwerker, Albert (1983) "Economic Trends, 1912–1949," in Denis Twitchett and John K. Fairbank (eds.), *The Cambridge History of China*, Vol. 12, *Republican China 1912–1949*, Part I: pp. 28–127.

Hou, Yangfang (候楊方) (2001)《中國人口史，第六卷：1910–1953》(*Chinese Population History*, Vol. 6: pp. 1910–1953), Fudan University Press.

Kuznets, S. (1966) *Modern Economic Growth: Rate, Structure, and Spread*. New Haven.

Kuznets, Simon. (1957) "Quantitative Aspects of the Economic Growth of Nations," *Economic Development and Cultural Change*, Vol. 5, No. 4, Supplement pp. 1–111.

Lieu, Da-Kuin (劉大均) (1937)《中國工業調查報告》(*Report on a Survey of Chinese Industry*), Nanking.

Liu, Ta-Chung and Kung-Chia Yeh (1965) *The Economy of the Chinese Mainland: National Income and Economic Development, 1933–1959*, Princeton University Press, Princeton.

Maddison, Angus (1998a) *Chinese Economic Performance in the Long Run*, OECD Development Centre, Paris.

Maddison, Angus (2007) *Chinese Economic Performance in the Long Run, 960–2030*, OECD, Paris.

Maddison, Angus and Harry X. Wu (2008) "Measuring China's Economic Performance," with Angus Maddison, *World Economics*, Vol. 9, No. 2, April-June 13–44.

Makino, Fumio and Huanzhen Lou (2014) "The Prewar Population and Labor force," in Ryoshin Minami, and Fumio Makino (eds.), *Asian Historical Statistics: China*: pp. 51–62, 東洋經濟新報社.

Minami, Ryoshin and Fumio Makino, *Asian Historical Statistics: China*, Vol. 3 of Konosuke Odaka, Osamu Saito and Kyoji Fukao (eds.) *Asian Historical Statistics*, 東洋經濟新報社.

Ministry of Agriculture and Commerce (MAC) (1923)《第八次農商統計表，1919》, Published by 中華民國農商部總務廳統計科.

Ministry of Industry and Commerce (MIC) (中華民國實業部) (1933)《中國經濟年鑒》(*China Economic Yearbook*).

Ministry of Internal Affairs (内務部 MIA), National Government (1936)《内務部年鑒》(*Internal Affairs Yearbook*). Commercial Press.

Notestein, Frank W. and Chi-ming Chiao (1937) Population, in John Lossing Buck (ed.), *Land Utilization in China*: pp. 358–399, University of Chicago Press.

Odaka, Konosuke, Osamu Saito and Kyoji Fukao (eds.) (1988) *Asian Historical Statistics*. 東洋經濟新報社.

Ou, Pao-san *et al.* (巫寶三主編：編者：汪馥蓀，章季閔，马黎之，南鐘萬，贝友林) (1947) 1933 年中國國民所得 (China's National Income, 1933), 2 vols., Shanghai: Commercial Press.

Peng, Zeyi (彭澤益・編) (1957)《中國近代手工業史資料，1840–1949》，第二卷 (Historical Materials of China's Contemporary Handicrafts, 1840–1949)，北京：中華書局.

Peng, Zeyi (彭澤益・編) (1962)《中國近代手工業史資料，1840–1949》，第一卷 (Historical Materials of China's Contemporary Handicrafts, 1840–1949)，北京：中華書局.

Rawski, Thomas G. (1989) *Economic Growth in Prewar China*, University of California Press, Berkeley.

Shi, Jianyun (史建云) (2005)《第一次世界大战后十年间中国手工业的转型》(The Transformation of Chinese Handicraft Industry: Ten Years after the World War I)，中国社会科学院近代史研究所民国史研究室，四川师范大学历史文化学院编：《1920 年代的中国》(*China in the 1920s*)．北京：社会科学文献出版社.

Skinner, G. W. (1964–1965) "Marketing and Social Structure in Rural China," Journal of Asian Studies, Part I: 24(1) November 1964, pp. 3–34, Part II: 24(2) February 1965, pp. 195–228 and Part III: 24(3) May 1965, pp. 363–399.

Smith, Adam (1976) *An Inquiry into the Nature and Causes of the Wealth of Nations*, E. Cannan (ed.), 5th edition. Chicago.

Wang Qingbin *et. al.* (王清彬等・編) (1928)《中國勞動年鑒》(*China Labor Almanac*) 北平社會調查部編輯出版.

Wang, Jingyu (汪敬虞・編) (1957)《中國近代工業史資料，第二輯：1895–1914》(下冊) (Historical Materials of China's Contemporary Industry, 1895–1914)，北京：科學出版社.

Wang, Shida (王士達) (1932 and 1933)《民政部户口調查及各家估計》(Household Survey by Ministry of Civil Affairs and Alternative Estimates)《社會科學雜志》(*Journal of Social Sciences*), Vol. 3, No. 3 and Vol. 4, No. 1.

Wang, Shida (王士達) (1935) "最近中國人口的新估計 (A New Estimate of Recent Chinese Population)," *Journal of Social Sciences* (《社會科學雜志》), Vol. 6, No. 2, pp. 191–266.

Wu, Harry X (2002) "Industrial Output and Labour Productivity in China 1949–1994: A Reassessment," in Maddison, A., W. F. Shepherd and D. S. Prasada Rao [eds.], *The Asian Economies in the Twentieth Century*, Edward Elgar Publishing Ltd., London pp. 82–101.

Wu, Harry X (2013) "Measuring Industry Level Em-

ployment, Output and Labor Productivity in The Chinese Economy, 1987–2008," *The Economic Review*, Institute of Economic Research, Hitotsubashi University, Vol. 64, No. 1, pp. 42–61.

Wu, Harry X (2014a) "China's Growth and Productivity Performance Debate Revisited – Accounting for China's Sources of Growth in 1949–2012," *The Conference Board Economics Working Papers*, EPWP1401.

Wu, Harry X (2014b) "The Growth of "Non-Material Services" in China – Maddison's "Zero-Labor-productivity-growth" Hypothesis Revisited, *The Economic Review*, Institute of Economic Research, Hitotsubashi University, Vol. 65, No. 3, pp. 265–283.

Wu, Harry X. and Ximing Yue (2012) Accounting for Labor Input in Chinese Industry, 1949–2009, *RIETI Discussion Paper*, 12-E-065.

Wu, Harry X., George G. Zhang and Ximing Yue (2015) "Constructing Annual Employment and Compensation Matrices and Measuring Labor Input in China," with Ximing Yue and George G. Zhang, *RIETI Discussion Papers* 15-E-005.

Yan, Zhongping *et. al.* (嚴中平等・編) (1955)《中國近代經濟史統計資料選輯》(*Selected Statistical Materials of China's Contemporary Economic History*), 北京：科學出版社.

Zhang, Jihong(章季閎) (1947) "中國就業人數的估計 (Estimation of Chinese Employment)", *Journal of Social Sciences* (《社會科學雜志》), Vol. 9, No. 2, pp. 71–91.

Zhang, Youyi(章有義・編) (1957)中國近代農業史資料, 第二輯: 1912–1927 (Historical Materials of China's Contemporary Agriculture, 1912–1927), 北京：生活, 讀書, 新知三聯書店.

Zhao, Wenlin and Shujun Xie (趙文林, 謝淑君) (1988)《中國人口史》(The Population History of China), 北京：人民出版社.