RIETI Discussion Paper Series 15-E-005

Constructing Annual Employment and Compensation Matrices and Measuring Labor Input in China

Harry X. WU
Hitotsubashi University

Ximing YUE
Renmin University of China

George G. ZHANG
University of Wisconsin, Madison

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 15-E-005

January 2015

Constructing Annual Employment and Compensation Matrices and Measuring Labor Input in China*

Harry X. WU

Institute of Economic Research, Hitotsubashi University, Tokyo

Ximing YUE

School of Public Finance, Renmin University of China, Beijing

George G. ZHANG

Department of Economics, University of Wisconsin, Madison

ABSTRACT

This paper documents the procedures in constructing China's employment and compensation metrics and measuring labor input in the Chinese economy. We begin with discussions of major conceptual, coverage, and classification problems in the official labor statistics and then propose strategies to solve the problems using a labor indexing methodology. This methodology allows the human capital attributes of any employed person weighted by their specific costs and the sum of the costs is controlled by the national income accounts. It guides our basic data work to adjust breaks, fill gaps, and restore consistencies in annual statistics and in benchmark marginal matrices. We use the iterative proportional filling (IPF) approach to obtain full-dimensioned employment and compensation matrices for benchmarks and, based on which, we then construct the matrices in time series. We show that China's total (cost-weighted) labor input grew by 4.1% per annum in 1980-2010, of which changes in quality and in hours worked made almost equal contributions. We also find that nearly half of the 2% quality improvement was attributable to changes in the industrial structure, followed by changes in education attainment and age structure.

Keywords: Labor employment and compensation by age, gender and education; Labor input indexing; Main and interactive effects of age, gender and education; Chinese economy

JEL classification: C82, E24, O47

* We are indebted to Kyoji Fukao and Mun S. Ho for constructive comments and suggestions. We thank participants at World and Asia KLEMS Conferences for helpful discussions. What reported in this work are interim results of China Industry Productivity (CIP) Project under the RIETI East Asian Industrial Productivity Project (http://www.rieti.go.jp/en/projects/program/pg-05/008.html). We acknowledge financial supports from RIETI (Research Institute of Economy, Trade and Industry, Japan) and Institute of Economic Research of Hitotsubashi University. We also acknowledge Zhan Li and Shuai Yuan for timely research support.

1. INTRODUCTION

Due to data problems a proper measure of China's labor input that takes into account the contribution of skill levels of different workers across industries remains as a big gap in the literature. This study extends earlier studies on Chinese industry by Wu and Yue (2003, 2010 and 2012) to the entire economy for the post-reform period 1980-2010. It aims to tackle major data problems and provides researchers not only with a consistent quantitative time series in concept, coverage and classification, but also a matching cost-weighted qualitative time series across 37 industries.

Hicks once said: "The measurement of capital is one of the nastiest jobs that economists have set to statisticians" (1981). Our long pains-taking experiences have shown that measuring labor input is by no means an easier task for a huge and complex economy like China that has been undergoing perhaps the most rapid income growth and structural changes in human history. Most studies have attempted to fix inconsistencies in the natural numbers of employment, which is important but it alone does not conform to the theory of homogenous production function. This is because natural numbers are heterogeneous, thereby the production function cannot be homothetically separable (Jorgenson 1990).

Thanks to the theoretical debates and empirical endeavours in earlier studies such as Denison (1962, 1974), Jorgenson and Griliches (1967), Kendrick (1961, 1973) and later contributions by Chinloy (1980), Gollop and Jorgenson (1980, 1983) and Jorgenson, Gollop and Fraumeni (1987),[1] the core issue in measuring labour input is no longer a dispute. That is how to hold the quality of hours worked constant when there are actually changes in the composition of the age, gender, education, occupation and industry of the workforce. In other words, how to convert heterogeneous hours worked into homogenous volume of labour input by weighting the cost of different types of labor in which the total costs are controlled by the national accounts. If this ignored, in the case of an improving labour quality, the growth of labour input will be underestimated and thereby the growth of total factor productivity (TFP) will be exaggerated.

While the Chinese economy has experienced significant changes in working hours and the composition of workforce caused by income growth, structural changes of the economy, as well as demographic transition, the role of labour input in the economy has not been properly measured. This has affected our understanding of the real sources of growth in China. Most studies have simply used the numbers employed implicitly as a proxy for labour input, irrespective of improperness and inconsistencies in various aspects (e.g. Borensztein and Ostry 1996; Chen *et al*. 1988; Chow 1993; Hu and Khan 1997; Bosworth and Collins 2008; Perkins and Rawski 2008).

Given serious inconsistencies in concept, coverage and classification in the official labor statistics, in this study we emphasize the importance of taking a *holistic* or *coherent* view about the relationship between economy-wide labor accounts and national accounts. We argue that with the control totals for production and income given in the national accounts and the control totals for employment from population and economic censuses, the key to the construction of a consistent labor input series is a careful examination of the quantitative relations of various labor

[1] See a comprehensive review of these studies by Jorgenson (1990, pp. 32-41).

statistics with these control totals. In so doing, with some proper assumptions, different data sources can be integrated in a "conceptually coherent system".

With this principle in our data work and following Chinloy (1980) and Jorgenson, Gollop and Fraumeni (1987), we have constructed China's employment and compensation matrices in time series and based on which we are able to express changes in the quality of China's workforce as the sum of main effects by demographic, educational and sectoral attributes and their interactions. We show that over the entire period in question the total labor input in China grew by 4.1 percent per annum, of which changes in quality and hours worked made almost equal contributions (about 2 percent each). This means that holding others constant, ignoring the effect of labor quality improvement (measured as positive change in the composition of hours across classified by sector, gender, age and education) will exaggerate TFP growth by at least one percent per annum (assuming that labor and capital take equal share in the national income). We also show that about half of the quality improvement is attributable to the change of sectoral (industrial) structure, whereas the rest is attributed to the changes in education attainment and age structure.

To document our procedures for obtaining these results this paper is organized as follows. After reviewing major data problems in the next section, we first introduce the labour input indexing methodology in Section 3 and then deal with the structural break in the official employment statistics and establish the control totals for major sectors in Sections 4. This is followed by Section 5 focusing on the construction of marginal employment and compensation matrices in time series and Section 6 explaining how to use census and survey data to complete the full-dimensioned the matrices for benchmarks. In Section 7 we estimate hours worked using census and survey data and construct hours-based employment matrix. Finally in Section 8, we interpolate the full-dimensioned benchmark matrices to obtain time series and based on which calculate the main and interactive effects of demographic, educational and industrial attributes on the changes of labor input. Section 9 concludes the study.

2. DATA PROBLEMS

The Chinese labour statistics cannot reflect significant changes in labor employment and compensation as the consequences of major policy changes and economic restructuring while maintaining historical consistency. The great challenges to researchers are how to handle structural breaks and fill gaps in time series and how to reconcile cross-sectional data from various official sources collected by different approaches and serving different policy objectives. Besides, as a long legacy of the material product system (MPS) under central planning the statistical system emphasizes more on quantitative indicators rather than value measures, which is a major obstacle to a proper measure of labor input that relies on labor compensation-weighted changes in quantity. Below we briefly review some of the main problems from both quantity and compensation perspectives.

Quantity-wise, we are interested in how to get the basic numbers correct and consistent in time and space for the total economy and its inherent industries and sectors. To an industry-level study the first and foremost problem in the Chinese labor statistics is perhaps the inconsistent industrial classifications over time. The official statistical system underwent significant changes in its standard industrial classification (CSIC) from 1972, 1985 and 1994 to 2002 and further to

2011. The changes reflect shifts from a system that mainly facilitates administrative and planning controls over vertically connected industries to a new system that is more in line with the technical nature of industries. However, there has been no official adjustment to the historical statistics following the changes of CSIC.

Second, the official employment statistics has long suffered from some significant structural breaks. As far as our research period is concerned, the great 1990 break cannot be ignored. From 1989 to 1990, the total number of employment is reported as a jump from 553.3 to 647.5 million, suggesting an astonishing increase by 94.2 million or 17 percent in one year (Maddison and Wu 2008)! This new total is available with three-sector breakdowns (primary, secondary and tertiary) linking to the same breakdowns prior to 1990, but not with official estimates at sub-sector level. When the sub-sectoral labor statistics discontinued in 2002, the gap maintained at a size of nearly 100 million (NBS, 2009, Table 4-5). Two decades have passed since the gap first emerged, yet there has been neither explanation nor adjustment by the statistical authority. Researchers have either ignored the break or made some rough adjustment.

Third, the official employment statistics is never measured in its natural unit, i.e. hours worked. In the absence of systematic surveys and estimates on hours, almost all studies accept the official indicator of numbers employed (e.g. Borensztein and Ostry 1996; Chen et al. 1988; Chow 1993; Hu and Khan 1997; Bosworth and Collins 2008; Perkins and Rawski 2008), which assumes that the average hours worked per employed person remained unchanged. This is certainly untrue. Studies have shown that institutional working hours were never the same across industries even under central planning (Zhu, 1999). In the reform era, changes in actual hours worked have been affected by two opposite factors: on the one hand, there have been several important reductions in institutionally regulated weekly working hours, and on the other hand, there has been a trend of increasing working hours in manufacturing especially since China's WTO entry.

Last but not least, as a long tradition in the central planning era, industrial enterprises, especially large state-owned corporations provide various internal services in education, health care, child care and social work and political organizations, as well as commercial outlets. This situation has to some extent changed since the later 1990s along with state-owned enterprise reform. However, how to adjust historical employment statistics remains as a big challenge. There is no easy way to estimate their numbers, costs and outputs, and thereby allocating them to proper service industries.

Turning to data problems related to the measurement of labor quality, to properly account for the change in labor quality one needs to take the ratio of user cost-weighted index of hours worked cross-classified by demographic, educational and industrial attributes to un-weighted quantity index (Denison 1961; Chinloy 1980; Gollop and Jorgenson 1980 and 1983). This requires matching employment and compensation matrices at least for some benchmark years. Directly using numbers employed as the practice in many existing studies implicitly accepts a strong assumption that different types of workers embodied with different human capital stocks are paid by the same marginal product.

To more properly measure labor input, one has to use available census and survey data. However, censuses, being population or industrial censuses, are not always compatible and usually do not provide researchers with cross-classified data for all demographic, educational and industrial dimensions. Economic and industrial censuses seldom provide matching income data. Besides, it is also a challenge to match censuses with annual labour statistics based on the state statistical reporting system. The quality of some censuses is even questioned e.g. for upward bias in education attainment (Young 2003).

To bypass the difficulties, many studies tend to rely solely on education statistics. Following Barro and Lee (1997 and 2000) and Maddison (1998), growth accounting studies such as Wang and Yao (2002) and Wu (2014) apply the perpetual inventory method (PIM) to measure human capital stock. Wu (2014) emphasizes the importance of weighting different flows of graduation by returns on education. However, this approach excludes other knowledge accumulated through employment experiences together with demographic and occupational/industrial factors. It may be a useful proxy for the aggregate economy but surely insufficient for measuring labor services at industry level.

Compared to human capital data, there is even less information on labour compensation at the required level of details. The only available time series data are annual wage bills and average wage per employee by industry. Some researchers have to rely on limited compensation surveys (Li *et al.* 1993). Young (2003, pp. 1245-46) used a larger sample from NBS household surveys for 1986-92 and from CHIP household surveys for 1988 and 1995. However, he did not attempt to examine the human capital effect at industry level. Besides, it is also difficult to estimate income paid in kind which may vary greatly across industries and ownership types (e.g., employees of the state sector enjoyed heavily subsidized housing and various welfare payments in kind).

3. LABOR INPUT INDEXING

Labour input indexing must be discussed coherently with a production function aggregating services provided by different types of labour and capital. The essential idea of constructing labour input index roots in the heterogeneity of labour in the sense that different types of labour have different marginal products in a given period. For example, an increase in the share of hours worked by skilled labour or by labour with better human capital will increase labour input even if the total hours worked remain unchanged (Denison 1962; Jorgenson and Griliches 1967). Directly using the numbers employed as a proxy for labour input in a production function implicitly assumes that labour is homogenous and the same hours worked by different types of labour will provide identical volume of services, which will certainly affect the reliability of the estimated residual.

Suppose that we have the following production function at time *t*, separable between labour and capital inputs, with a Hick's neutral shift parameter *A*:

(1) $$Y_t = A_t f(L_t, K_{1t}, ..., K_{jt})$$

where Y_t represents output, L_t labour input, and $K_{1t}, ..., K_{jt}$ the services of different capital inputs. Now let us define the labour aggregate as a function of hours worked by different types of labour:

(2) $$L_t = \phi(H_{1t},..., H_{nt})$$

where H_{it}, $i = 1, \ldots, n$, represents hours worked by type *i* labour.[2] Following Chinloy (1980) and Jorgenson, Gollop and Fraumeni (1987), if assume efficient labour market and linear homogeneity of ϕ, then we have:

(3) $$\frac{\partial \ln L_t}{\partial t} = \sum_{i=1}^{n} s_{it} \frac{\partial \ln H_{it}}{\partial t}$$

where s_{it} is the share of the *i*th type of labour in total labour compensation, which is equal to its logarithmic marginal output under the efficiency assumption:

(4) $$s_{it} = \frac{w_{it} H_{it}}{\sum_{i=1}^{n} w_{it} H_{it}} = \frac{\partial \ln \phi}{\partial \ln H_{it}}$$

In Equation (4), the hourly wage of the *i*th type of labour is w_{it} and its compensation is $w_{it} H_{it}$. The growth rate of labour input is a convex combination of growth rates of total hours for each type of labour, with compensation shares as weights. Equation (4) also indicates that the necessary condition for producer equilibrium is given by equality between the share of the *i*th type of labour in the labour aggregate and the elasticity of the aggregate with respect to the *i*th type of labour.

Let total hours worked by all types of labour be $H_t = \sum_{i=1}^{n} H_{it}$. Then, the growth rate of H_t is the sum of the weighted growth rates of hours worked by each type of labour:

(5) $$\frac{\partial \ln H_t}{\partial t} = \sum_{i=1}^{n} b_{it} \frac{\partial \ln H_{it}}{\partial t}$$

with $b_{it} = H_{it} / \sum_{i=1}^{n} H_{it}$ the weight of the *i*th labour type. Therefore, average labour quality per hour can be defined as labour input divided by hours worked:

(6) $$Q_t = L_t / H_t$$

and its growth rate is:

(7) $$\frac{\partial \ln Q_t}{\partial t} = \sum_{i=1}^{n} (s_{it} - b_{it}) \frac{\partial \ln H_{it}}{\partial t}$$

[2] It is standard to use hour worked here. In the case of using numbers employed in the equation, when hours are not available or for the comparison between hours and numbers, *N* (for numbers) can be used to substitute for *H*.

which is the sum of growth rates of hours worked by each type of labour, weighted by the *difference* between the shares in labour compensation and hours worked.

Now following Christensen, Jorgenson and Lau (1973), we specify the labour aggregate in the translog form:

(8) $$\ln L_t = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln H_{it} + 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln H_{it} \ln H_{jt} ,$$

where α_0, α_i, i = 1, …, n, and β_{ij}, i, j = 1, …, n, are parameters and where $\beta_{ij} = \beta_{ji}$ to satisfy the required symmetry conditions. Under the assumption of linear homogeneity, we have $\sum_{i=1}^{n} \alpha_i = 1$ and $\sum_{j=1}^{n} \beta_{ij} = 0, \quad i = 1, \ldots, n$.

With the efficiency assumption, the relative share of the ith type of labour equals its logarithmic marginal product:

(9) $$s_{it} = \alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln H_{jt}$$

where s_{it} is defined the same as in equation (4).

Equations (8) and (9) as well as the symmetry conditions $\beta_{ij} = \beta_{ji}$ imply that the growth rate of the translog index of labour input l_t is:

(10) $$l_t \equiv \Delta \ln L_t = \sum_{i=1}^{n} v_{it} \, \Delta \ln H_{it}$$

where $v_{it} = (s_{it} + s_{it-1})/2$ and Δ demotes the first difference operator.

From equation (6), the growth rate of quality q_t is defined as:

(11) $$q_t \equiv \Delta \ln Q_t = \Delta \ln L_t - \Delta \ln H_t = l_t - h_t$$

where h_t represents the growth rate of total hours worked by all types of labour. Clearly, the growth rate of quality will be positive if hours worked by relatively high wage labour increase more rapidly than total hours worked.

Next, the contribution of each attribute of labour to quality change can be decomposed into two types of effects: the main effect of the attribute and the interactive effects of the attribute with each of the rest attributes. The main effect of the ith attribute is defined as the difference between the growth rates of labour input due to the ith attribute and total hours worked, regardless the time subscript:

(12) $$q_i = l_i - h$$

where h is exactly the same as h_t defined in equation (11) and l_i is growth rate of labour input due to the ith attribute or the factor i. In the case of $q_i > 0$, as noted in Jorgenson and Griliches (1967), labour input measured as total hours worked is biased downward and hence TFP is biased upward.

Suppose that there are two attributes of labour, i and k, as a proper subset from n factors, a first-order interactive effect is derived from the partial index growth rate l_{ik} for the two factors and the single factor indices l_i and l_k:

(13) $$q_{ik} = (l_{ik} - h) - (l_i - h) - (l_k - h) = l_{ik} - h - q_i - q_k$$

that is, the joint effect of i and k or $(l_{ik} - h)$ less the main effect of each. If there are only two factors, i and k, the growth rate of labour quality is defined as the summation of the main effects of two factors and their first-order interactive effect:

(14) $$q = l_{ik} - h = q_i + q_k + q_{ik}\ .$$

As for labour input with j factors, $l_{1,\ldots,j}$, interactive effects up to (j – 1)th order are obtainable following the same principle (Chinloy, 1980, p.111).

4. Quantitative "Control Totals" for major Sectors

This section deals with the problem of "control totals". Following Wu (2014) we adjust the 1990 break in the employment series by investigating the nature of the break and the fundamental forces that might affect the demand and supply of labor at the time of the break. However, different from Wu who focused on the aggregate economy, we here focus on sector breakdowns. Furthermore, we also adopt Wu's new estimates (2014) for the missing military personnel prior to 1990, a problem which Maddison (1998) made the first attempt to tackle.

Examining the "structural break": 1990 or 1982?

A quick look at the 1990 "structural break" of a 17-percent jump against the background of labor supply and macroeconomic situation gives an impression that the break is rather artificial. On one hand, the change of working-age population around that time was stable, i.e. without any significant deviation from the trend. On the other hand, it was impossible for the demand for labor to have a faster-than-normal increase in the middle of a serious growth slowdown—even by official estimates the growth of GDP dropped sharply from 10.5 percent in 1988 to 3.3 percent in 1989 and stayed at around a similar rate (3.2) in 1990, which was the slowest growth since the reform.

As discussed in Yue (2005), the gap is caused by inappropriately linking the results of the 1990 Population Census results to the annual estimates based on a regular employment registration and reporting system that was established in the early planning time. The population census discovered a large number employed who had been missed by the regular reporting system, yet NBS did not integrate the new results with its annual estimates at the industry level. Nonetheless, without any good reason to ignore the census results, between 1990 and 2002 the NBS continued its census-based estimation for total employment supported by annual population sample surveys and published the results parallel to annual industry-level estimates in a way that disguised the huge underlying inconsistency between the reported totals and the (implicit) sum of industry-level estimates (Maddison and Wu 2008).

It is reasonable to assume that this 94.2 million of additional workforce recorded in the 1990 Census did not come out all of sudden in one year but had already existed in the economy for some time though it was not captured by the regular statistical reporting system. This means that the outside-system employment might have begun much earlier and at a time when changes in employment policy allowed informal employment to emerge but it was not yet covered by the registration system. A proper investigation should be conducted along two lines: checking earlier or pre-1990 population censuses or sample surveys to see if a similar discrepancy had already existed and examining changes in employment policy that might have resulted in outside-system employment some time before 1990.

Before the 1990 Population Census, China only conducted three population censuses in 1953, 1964 and 1982, and only the 1982 Population Census provided employment data. Wu (2014) reports that the 1982 census actually shows that China's total number of employment was 521.5 million, or 68.6 millions more than the annual estimate of 452.9 million for that year. This is to a large extent in line with the 1987 one-percent population sample survey which gave an estimate of 584.6 million or 56.7 million more than the annual estimate of 527.8 million. It is now clear that the structural break should have emerged at least in 1982 rather than in 1990.

What behind the 1982 "structural break"?

The next question is when this additional employment began to emerge. According to Wu (2014), there is evidence to suggest that the government began to relax its employment regulations in the early 1970s to allow rural enterprises (then labeled as "commune and brigade factories") and "neighborhood factories" in cities to absorb surplus labor. However, since the new jobs were created in an informal way and many of new workers were temporal and seasonal in nature, they were not fully covered by the state labor planning and registration system. Following Wu (2014) we also assume that the discrepancy began in the early 1970s, which was the time when China first politically settled down from the chaotic situation at the early stage of the Cultural Revolution and reemerged internationally marked by its normalized relationship with the US and Japan.

An adjustment to the 1982 "structural break"

In Wu's adjustment (2014), he first revised the official employment estimates between 1982 and 1990 by taking into account the results from the 1982 and 1990 Population Censuses and the 1987 Population Sample Survey, as well as their relationships with the official annual employment estimates. We take his reconstructed 1982-90 time series of employment as "control totals" for this time period. Not surprisingly, his adjustment has pushed back the "structural break" from 1990 to 1982 and results in a 19.3-percent jump in 1982.[3] Wu proposes three scenarios to adjust the 1982 break. His first scenario does not consider the policy change effect in the early 1970s and assumes that all the employment data prior to 1982 are underestimated to the same degree as suggested by the 1982 Census. This raises the level of employment from 1981 back to 1949, though maintaining the original official growth rates. His second scenario assumes that the gap identified by the 1982 census originated from 1971 when the government began relaxing planning controls over employment. In his adjustment, the level of employment in 1981 is first raised to iron out the break. Then, by incorporating a new trend between 1970 and 1982 with annual deviations from the original trend over the same period, he obtained a new employment series for the period 1971-81. This added increasingly more numbers employed to each year, for example, 4.8 million for 1971 and 69.3 million for 1981.

In the present study, we follow Wu's third scenario because it has industry implications. It takes the same approach as the second scenario to adjusting the 1982 break, but proposes a different approach to allocating the additional employment. It assumes that the additional employment is only engaged in labor-intensive, non-farming activities. In services, since these additional laborers are likely least educated and largely part-time and multi-job migrant workers, they are assumed to mainly engage in the "material services" (a concept adopted in the Chinese statistical system under MPS) including transportation, wholesales and retails, and hotel and catering services, and not in "non-material services" such as financial, governmental, healthcare and education.

Military personnel as part of "non-material service" employment

To complete our estimation for the control totals, we need to consider bringing the missing military personnel back for the period prior to 1990 that are excluded in official estimates.[4] Maddison made the first ever attempt for the adjustment. But with very limited information, he simply assumed the size of the military personnel was a constant 3 million for the period 1952-1996 (Maddison 1998, pp. 168-9). In his later work, he revised his estimate by noticing that the official employment statistics began to include military personnel in services from 1993 onwards (Maddison 2007, p. 170).

In our final estimate of the "control totals", we adopt Wu's recent estimates with new evidence (2014a and 2014b). He shows that first, the official practice of excluding military personnel actually ended in 1990 not 1993 and second, the size of military personnel prior to

[3] The adjustment is made for four sectors, namely, agriculture, industry, construction and services. However, the number of agricultural employment in the 1982 Census (384.2 million) looks too high, almost the same as that of the 1990 Census (389.1). Its share in the census total employment is 74 percent, which is higher than that from the regular statistical report system, 68 percent. This is unreasonable in that the census is expected to pick up more non-agricultural employment that is not covered by the reporting system. Taking 1990 and changes between 1980 and 1990 as references, Wu (2014a) reduces the number of agricultural employment by 10 percent and reallocate the difference to other sectors by their weights. The results look more plausible with agricultural employment accounting for 66.3 percent, industry 18 percent, construction 2.2 percent, and services 13.5 percent.

[4] In addition to defense service, military personnel also engaged in construction, transportation, farming and government services in the early period of the People's Republic. Assuming that they only engaged in "non-material (and non-market) services" may exaggerate the input and output of these services, but it will not affect the aggregate analysis.

1990 was not a constant 3 million over time as Maddison assumed. His evidence shows that China's armed forces were numbered at about 5.5 million in 1949. After four rounds of demobilization between 1950 and 1956, the number was substantially reduced to 2.4 million by the end of 1958. It, however, increased again between the mid 1960s and the mid 1970s in response to the border tensions and conflicts with India and the Soviet Union, respectively. By the end of 1975, the Chinese armed forces picked at 6.8 million. There were, however, two rounds of demobilization in the post-Mao period initiated by Deng aiming at maintaining smaller but more modernized armed forces at around 3 million from the end of the 1980s (see Wu 2014a, Table C, Appendix C).

Finally, this exercise yields the "control totals" for broad sectors used in the construction of our employment matrices, namely, primary, secondary that is further divided into industry and construction, and tertiary that is further divided into two categories, "materials services" (Services 1) and "non-material services" (Services 2) for a convenient concordance with the official statistics. The "control totals" are however subject to further but minor adjustments between agriculture and services in Section 5. The final results (with the further adjustments) are reported in Table A2 (Appendix).

5. MARGINAL EMPLOYMENT AND COMPENSATION MATRICES

In this section, we focus on the construction of two types of marginal matrices in time series, i.e. employment and compensation matrices. They are "marginal" because they are not yet fully cross-classified by all designated attributes in Table 1. These marginal matrices are constructed as industry-specific "control totals". Importantly, the employment matrix discussed in this section is a quantitative base for the compensation matrix discussed in the next section.

Coverage and cross-classification

Our work on measuring labor input covers the entire Chinese economy at industry level for the period 1980-2010. It is well known that inconsistent, incomplete or overlapped coverage of the Chinese official employment statistics, reported by different authorities using different criteria, has caused great confusions to researchers. Following the earlier work in Wu and Yue (2010 and 2012), and taking into account international compatibility especially with EU-KLEMS, the CIP Project establishes a classification system of 37 industries economy-wide (Appendix Table A1). Based on the reconstructed Chinese national accounts especially input-output tables, Wu and Ito (2014) have constructed annual supply-use tables and input-output tables for 37 CIP industries. Our work on employment aims to exactly match the CIP output accounts. This also conceptually ensures that the industry-specific labor income in the input-output accounts can be used as the "control total" in the compensation matrix (see Section 6).

The number of factors that are considered affecting labour quality determines the dimensions of the matrices to be constructed. As listed in Table 1, in this study we aim to construct matching labour employment and compensation matrices cross-classified by four primary attributes, reflecting the heterogeneity of different types of labor, that is, gender/sex (g=2), age (a=7 groups), education attainment (e=5 levels) and industry/sector (s=37) that captures the change of labor quality through occupational and technological experiences. Therefore, each full-dimensioned matrix contains four cross-classified attributes (dimensions) and in total 2590 cells for each time point of the full period.

In the present section, we first explain how our full labor accounts in numbers employed are constructed to match the 37 CIP industries in the national accounts and input-output accounts. As introduced earlier, we have already established the "control totals" for major sectors (i.e. agriculture, industry, construction and two categories of services, Table A2). We now focus on the data construction for individual industries within each of these sectors, mainly the industrial and service sectors.

TABLE 1

CLASSIFICATION FOR HUMAN CAPITAL ATTRIBUTES

Gender: (*g*)	*Education Attainment*: (*e*)
1. Male	1. Illiteracy or semi-illiteracy
2. Female	2. Primary school
Age Group: (*a*)	3. Junior high school
1. 15-19	4. Senior high school
2. 20-24	5. Tertiary education
3. 25-29	*Industry/Sector*: (*s*)*
4. 30-39	1. Agriculture
5. 40-49	2. Coal mining
6. 50-54	…
7. >54	37. Personal & other services

Source: See text. *See the full list in Table A1 in Appendix.

Industry-specific employment series for the industrial sector

Following the discussion of the problems in the official employment statistics in Wu and Yue (2012), the enterprises that are monitored by the state statistical reporting system can be considered representing the "formal component" of each industry. Wu and Yue have identified three layers within the boundary of the "formal component". The first layer consists of all state-owned enterprises (SOEs) which can be treated as the "hard core". The second layer is called "township layer" that includes all enterprises at or above the rural township level prior to 1998 and at or above the "designated size" (see below) afterwards. This is followed by the "outer layer" of the "formal economy" that is made up by small but registered including village-level enterprises and other industrial activities.

Outside the "formal component" it is the rest of industrial workforce engaged in "informal" industrial activities, mainly labor-intensive industries. Wu and Yue (2012) define the workers engaged in this informal part of the economy as "outside-the-system employment", which is considered as the forth layer of the entire industrial employment. They treat the first three layers fundamentally important because they lay out the basic structure of China's industrial economy. The more accuracy the estimates for three layers, the closer to the reality of the final structure. Here, we accept the same principle although we do not consider the ownership or administrative dimension.

Under NBS there are two specialized departments that regularly publish data on industrial employment: Department of Industrial and Transportation Statistics (DITS, later changed to Department of Industrial Statistics or DIS) and Department of Population and Employment

Statistics (DPES). However, researchers often find that for the same indicator the two official sources may report different data. This requires a good understanding of the definition used by the different authorities. Of all the available indicators, the most compatible indicator from the two sources is the state "staff and workers". From 1952 to 1997, DITS and DPES reported almost identical data for this indicator although for the period 1980-84 and 1994-97 there are slight differences between the two sources.

However, the biggest discrepancy between DITS and DPES is found with the *total* "staff and workers". The DITS series began to diverge from the DPES series from 1978 and surpassed the latter by 11.2% in 1978 and 35.5% in 2000. We find significant incompatibility in the definition of "staff and workers" between the two sources. In fact, the DITS definition refers to the employment at or above the rural township level prior to 1998 and afterwards engaged in all state enterprises and non-state enterprises with at least 5 million yuan of annual sales (DITS, 2000, p.16, footnote). By contrast, the DPES definition refers to the employment located in cities including so-called "industrial areas" that are administratively treated as urban areas. Obviously, the DITS definition has a wider coverage than that of the DPES because it includes employment outside urban areas. For the indicator of "persons engaged", which has a wider coverage including both "staff and workers" and those defined as "non-staff and workers", DPES provides a long time series back to 1949 that largely maintains conceptual consistency. By contrast, DITS has never published similar indicator.

Following Wu and Yue (2012), we concentrate on SOEs (the first layer) for which longer and more consistent data are available and both DITS and DPES sources can be used and cross checked. The border line of the "township layer" (the second layer) is basically controlled by the DITS statistics on "staff and workers". Besides, we also use the industry-level labor statistics on foreign-invested enterprises to help the estimation for this layer because more industry-level information may make our estimation closer to the reality.

Our work on the "outer layer" or the third layer of the formal sector relies on the DPES series of "persons engaged" to control the border line. But the series lacks industry details. To obtain industry-level information we rely on data from two industrial censuses for 1985 (including some historical information for 1980 as well) and for 1995, data from two economic censuses for 2004 and 2008, and data on rural enterprises by Ministry of Agriculture. After reclassifying the census data to the CIP 37 industry system, we estimate annual series based on interpolated census structures and the "control totals" from the DPES series of "persons engaged".

Finally, we deal with those employed "outside the system", i.e. the fourth layer of China's full labor accounts. We define the forth layer by the gap between the population-census-based "total numbers employed" (Table A1) and the DPES concept of "total persons engaged". This gap is used as the "control totals" for this layer. By assumption, the distribution of the "control totals" is based on the structure of the labor intensive industries in the "outer layer". Table 2 presents the structure of each layer of the industrial workforce by industry for four selected years.

TABLE 2

EMPLOYMENT STRUCTURE OF CHINESE INDUSTRY BY OWNERSHIP TYPE OR ADMINISTRATIVE LEVEL

(Industry total employment = 1)

CIP Code		1980				1990				2000				2010			
		1st	2nd	3rd	4th	1st	2nd	3rd	4th	1st	2nd	3rd	4th	1st	2nd	3rd	4th
	Total Industry	***0.464***	***0.269***	***0.200***	***0.066***	***0.382***	***0.300***	***0.187***	***0.131***	***0.250***	***0.314***	***0.199***	***0.237***	***0.112***	***0.471***	***0.149***	***0.268***
2	Coal mining	0.783	0.117	0.100	0.000	0.675	0.194	0.131	0.000	0.634	0.172	0.195	0.000	0.551	0.290	0.159	0.000
3	Oil, gas extraction	0.760	0.240	0.000	0.000	0.847	0.153	0.000	0.000	0.984	0.008	0.008	0.000	0.925	0.039	0.036	0.000
4	Metal mining	0.620	0.231	0.150	0.000	0.648	0.184	0.169	0.000	0.386	0.325	0.289	0.000	0.199	0.577	0.224	0.000
5	Non-metallic mining	0.260	0.359	0.381	0.000	0.229	0.283	0.488	0.000	0.211	0.197	0.592	0.000	0.102	0.508	0.389	0.000
6	Food products	0.453	0.122	0.248	0.177	0.352	0.158	0.165	0.325	0.145	0.189	0.159	0.508	0.037	0.400	0.085	0.479
7	Tobacco products	0.980	0.013	0.007	0.000	0.804	0.192	0.004	0.000	0.873	0.110	0.016	0.000	0.948	0.044	0.008	0.000
8	Textiles	0.555	0.300	0.145	0.000	0.481	0.401	0.119	0.000	0.339	0.481	0.180	0.000	0.033	0.752	0.214	0.000
9	Apparel	0.053	0.545	0.235	0.167	0.042	0.425	0.179	0.354	0.023	0.361	0.146	0.469	0.006	0.356	0.096	0.541
10	Leather goods	0.188	0.446	0.213	0.152	0.120	0.396	0.163	0.321	0.019	0.446	0.127	0.408	0.002	0.420	0.087	0.492
11	Saw mill, furniture	0.087	0.224	0.402	0.286	0.106	0.391	0.169	0.334	0.033	0.162	0.192	0.614	0.007	0.239	0.114	0.641
12	Paper & publishing	0.392	0.275	0.195	0.139	0.250	0.271	0.161	0.318	0.120	0.219	0.157	0.503	0.020	0.196	0.118	0.666
13	Petroleum	0.677	0.269	0.053	0.000	0.839	0.042	0.119	0.000	0.590	0.274	0.136	0.000	0.439	0.494	0.067	0.000
14	Chemicals	0.748	0.184	0.068	0.000	0.654	0.244	0.102	0.000	0.494	0.356	0.150	0.000	0.166	0.653	0.181	0.000
15	Rubber & plastics	0.183	0.438	0.221	0.158	0.138	0.346	0.174	0.342	0.059	0.266	0.161	0.514	0.015	0.301	0.103	0.580
16	Building materials	0.346	0.267	0.387	0.000	0.216	0.355	0.429	0.000	0.183	0.338	0.479	0.000	0.061	0.584	0.355	0.000
17	Basic metals	0.831	0.122	0.046	0.000	0.719	0.206	0.076	0.000	0.578	0.279	0.144	0.000	0.365	0.584	0.050	0.000
18	Metal products	0.172	0.598	0.230	0.000	0.135	0.584	0.281	0.000	0.095	0.513	0.392	0.000	0.038	0.737	0.225	0.000
19	Indus. machinery	0.550	0.358	0.092	0.000	0.622	0.237	0.141	0.000	0.408	0.353	0.239	0.000	0.113	0.612	0.275	0.000
20	Electrical equip.	0.513	0.391	0.097	0.000	0.322	0.559	0.118	0.000	0.166	0.652	0.182	0.000	0.057	0.781	0.162	0.000
21	Electronic equip.	0.763	0.183	0.054	0.000	0.517	0.389	0.094	0.000	0.217	0.671	0.112	0.000	0.074	0.852	0.073	0.000
22	Instruments	0.789	0.172	0.039	0.000	0.497	0.441	0.062	0.000	0.484	0.302	0.215	0.000	0.100	0.765	0.135	0.000
23	Transport equip.	0.768	0.200	0.033	0.000	0.451	0.493	0.056	0.000	0.524	0.358	0.118	0.000	0.273	0.543	0.184	0.000
24	Miscellaneous	0.137	0.145	0.420	0.299	0.053	0.185	0.256	0.505	0.032	0.243	0.172	0.552	0.009	0.367	0.094	0.530
25	Power, gas & water	0.789	0.188	0.023	0.000	0.934	0.044	0.022	0.000	0.854	0.103	0.043	0.000	0.721	0.131	0.148	0.000

Sources: Authors' calculation based on Wu and Yue (2012).

Notes: 1st layer: employment in state-owned enterprises; 2nd layer: non-state employment within the township layer; 3rd layer: employment engaged in the "outer layer"; the 4th layer: employment outside the system. Sum of the four layers (total industrial employment) = 1.

Industry-specific employment series for services

Available data on services are much more problematic than the industrial employment data for three reasons. First, traditionally the official statistical coverage for services is insufficient because of the Marxian dogma that considers "non-material services" being non-productive. Second, under central planning many services were provided within industrial enterprises especially large-sized state-owned enterprises. Not all of these "internal services" have been separated from industrial employment despite reforms. Third, the mushrooming of small services, especially those of family-based and self-employed, has made maintaining reliable records difficult.

The official statistics for service employment are provided by DPES. In the DPES 16-sector statistics that began in 1978 and terminated in 2003 (NBS, 2003: 128-131), there are 11 sectors of services. The data suffer from the 1990 structural break as examined at the aggregate level. However, from 1994 the official labor statistics includes a subset of 105 sectors for "urban units", of which 57 sectors belong to services. After this 16-sector series was discontinued in 2003, the total number of sectors under "urban units" increased to 120. According to the official data for 2010, the service employment in "urban units" accounts for 26.2 percent of the total service employment (NBS, 2011: 113-114). It may be qualified for the "formal component" of the total services. However, since there are no sector breakdowns for the service employment in urban "non-units", it is difficult to work out urban totals for each service industry. Besides, there are also no sector breakdowns for rural service employment. Therefore, it is difficult to establish a full service employment accounts at industry level using the DPES data.

TABLE 3
EMPLOYMENT STRUCTURES OF CHINESE SERVICES BASED ON POPULATION CENSUS DATA

CIP Code	EU-KLEMS		1982	1987	1990	1995	2000	2005	2010
		Services 1	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*
27	G	Wholesale & retail trades	0.493	0.552	0.570	0.570	0.540	0.551	0.574
28	H	Hotels & restaurants	0.081	0.080	0.075	0.118	0.158	0.161	0.168
29	I	Transport, storage & post	0.394	0.336	0.328	0.189	0.283	0.244	0.212
30	71t74	Information, comp. services	0.032	0.032	0.028	0.123	0.019	0.044	0.045
		Services 2	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*
31	J	Financial intermediation	0.035	0.034	0.036	0.031	0.031	0.059	0.072
32	K	Real estate activities	0.011	0.011	0.007	0.009	0.010	0.094	0.059
33	71t74	Business services	0.041	0.044	0.061	0.036	0.027	0.030	0.122
34	L	Public administration	0.220	0.263	0.177	0.118	0.106	0.203	0.227
35	M	Education	0.383	0.382	0.367	0.360	0.318	0.220	0.204
36	N	Health and social security	0.137	0.138	0.136	0.147	0.144	0.104	0.103
37	O&P	Other services	0.173	0.128	0.216	0.298	0.363	0.290	0.212

Source: Authors' calculation.

To bypass the difficulties, we opt for using population census data on employment. The key steps are explained as follows. First, we adjust available census data to align them with the CIP classification for services (Table A1). Second, based on our "material services" and "non-material services" grouping, for which we have constructed "control totals" (Table A2), we work out intra-group structure for the census years and then make structure interpolations between the census years. Note that the service structures of the 1982 Population Census are applied to the period 1980-81. The benchmark structures are presented in Table 3. Third, based on the annual structures, sector-specific employment data are derived.

Marginal employment matrix in time series

However, before finalizing the industry-specific marginal employment matrix in time series, we need to have further some adjustments to the "control totals". The adjustments are related to agriculture. The "control totals" presented in Table A2 are census-based, broad-sector (primary, secondary and tertiary) estimates. Up to that point, we had accepted the primary sector as agriculture without any adjustment. After working on the industrial and service sectors, we need to make further adjustments.

The first adjustment is between agriculture and industry. China's industrial classification before CSIC/2011 defines "logging" as part of the mining group that is included in the broad industrial sector (Sector II.2, Table A2). However, following the international standard, "logging" should be relocated from industry to forestry. The shares presented in Table 3 have already excluded "logging". We should finally relocate it from industry to agriculture.

The second adjustment is based on a crosscheck of the agricultural employment statistics published by two sources, DPES and the Ministry of Agriculture (MoA). We have mentioned that the DPES 3-broad-sector data are inconsistent with its 16-sector-data. It is difficult to set up concordance between the 3-sector data with the 16-sector data without making strong assumptions. It is reasonable to argue that the MoA data on rural employment have narrowed down the scope and put agriculture in a more relevant setting, that is, rural rather than national economy. We assume that the MoA source is more reliable than the DPES source. Since the agricultural employment in the MoA statistics is smaller than that from DPES, we calculate the gap and relocate it to the following labor-intensive services: "whole sales and retails", "hotels and restaurants", and "personal services". The adjusted "control totals" are also reported in Table A2.

Marginal compensation matrix in time series

We use the industry-specific labor compensation data from the CIP-reconstructed national accounts to control for the costs of labor input in each industry (Wu and Ito 2014). This two-dimensioned, i.e. compensation-by-industry marginal matrix in time series conceptually matches the numbers employed by industry. It will be first used in the construction of the full-dimensioned compensation matrices for benchmarks in the next section and then used as "control totals" in the construction of the full-dimensioned compensation matrices in time series in Section 8.

6. Full-dimensioned Employment and Compensation Matrices for Benchmarks

To construct full-dimensioned employment and compensation matrices, we need data for other dimensions that are not yet integrated in the above constructed marginal matrices in time series. We first look for such data in available cross-classifications (various forms of marginal matrices) and then construct a full-dimensioned matrix for each designated benchmark by using an iterative proportional filling (IPF) method. The required data are collected from all available population censuses and household surveys. In what follows, we begin with an introduction to the IPF method and then explain the features and sources of newly added marginal matrices for employment and compensation, respectively, which are used in our IFP exercise.

Methodology for constructing full-dimensioned matrices

In this study, following Bishop, Fienberg and Holland (1975),[5] we develop an IPF approach to estimate full-dimensioned benchmark matrices based on all available or constructed marginal matrices including the relevant data for the benchmarks of the constructed marginal matrices in time series. The IPF approach is designed to integrate the marginal matrices that contain incomplete data by generating the maximum likelihood estimate of each element of a matrix. It is also used to fill missing cells.

The IPF procedure is expressed as follows. Suppose that we have a set of marginal matrices, J, for a specific benchmark. For the simplicity in our notation, we let J be the number of elements in set J. A matrix $j \in J$ is constructed from the intersection of some of the 4 human capital dimensions (Table 1) and the related number of employment in a specific cell i of j is denoted as N_{ij}. We use Σ as an operation of summation of the employment numbers of all cells in a given marginal matrix. In the estimation, we search for $\widehat{N}_i$ in a full-dimensioned matrix such that $\widehat{N}_i$ is the MLE of the expected value N_i

Let $N_{i1}^{(0)} = 1 \; \forall i,$ and for $1 \leq t \leq T$ we run the following loop:

$$N_{i1}^{(Jt-J+1)} = \frac{N_{i1}^{(Jt-J)} N_{i1}}{\sum N_{i1}^{(Jt-J)}}$$

$$\ldots$$

$$N_{ij}^{(Jt-J+j)} = \frac{N_{ij-1}^{(Jt-J+j-1)} N_{ij}}{\sum N_{ij-1}^{(Jt-J+j-1)}}$$

$$\ldots$$

$$N_{iJ}^{(Jt)} = \frac{N_{iJ-1}^{(Jt-1)} N_{iJ}}{\sum N_{iJ-1}^{(Jt-1)}}.$$

A full-dimensioned matrix is constructed when we get the expected value of N_i : $\lim_{T\to\infty} N_{iJ}^{(JT)}$.

It is recommended that $N_{iJ}^{(JT)}$ generally converges when setting up $T = 150$, which is what we did in the calculation. Actually, in most of the cases in our exercise, $N_{iJ}^{(JT)}$ converges when $T = 3$.

---

[5] Interested readers may refer to for example the discussions of a five-dimension model with missing variables (Bishop, Fienberg and Holland, 1975, pp. 45-46), of the rules for detecting existence of direct estimates (pp. 76-77), and of the basic elements of the IPF (p. 83).

As shown in the above loop, the value of jth estimate of N_i in the tth loop, $N_{ij}^{(Jt-J+j)}$, is dependent on the value of total sum of $(j-1)$th matrix in J. Thus the resulting value of $\widehat{N_i}$ might depend on the order of the marginal matrices in the above loop if they have different total values. A typical case of this problem is that the marginal matrices are from various data sources. However, this problem is not severe, since the resulting value will be further adjusted to make it consistent with the control totals in Table A2.

The resulting value of N_i might depend on the order of partial matrices in the above loop if they have different the control totals. A typical case of this problem is that the partial matrices are from various data sources. However, this problem is not severe and the resulting value will be further adjusted to make it consistent with the control totals in Table A2.

Full-dimensioned employment matrix for benchmarks

As given in Table 1, our full-dimensioned employment (quantity) matrix for any time point (benchmark) is defined as 37 sectors (industries) cross-classified by two genders, seven age groups and 5 education attainment levels. The constructed marginal employment matrix in time series is only two-dimensioned as industry by time. In order to construct a full-dimensioned matrix, we use IPF to integrate the industry dimension of the constructed time series with the marginal matrices available from all population censuses and surveys as summarized in Table 4 including data sources and number of sectors for 7 benchmarks.

TABLE 4
DATA SOURCE AND DESCRIPTION OF MARGINAL EMPLOYMENT MATRICES

Benchmark	Data Source	Number of Sectors	Marginal Matrices
1982	3rd population census	38	*s×e; s×g×a*
1987	1% population sample survey	57	*g×a×e; s×e*
1990	4th population census	75/13	*s×g×a; s×g×e*
1995	1% population sample survey	16	*g×a×e; s*
1980-2009	Wu and Yue (2012)	24	*S*
2000	5th population census	92	*g×a×e; s×g*
2005	1% population sample survey*	96	*s×g×a×e*
2010	6th population census	95	*s×g×a; s×g×e*

Sources: As specified in the table.
Note: *In addition to the published national average data, we have also obtained 20 percent of the raw data.

In constructing the full-dimensioned matrix, in general sectors are first reclassified in every available marginal matrix in Table 4 to make them consistent with the CIP 37-sector classification. We then apply the IPF method to the marginal matrices. This results in a full-dimensioned employment number matrix for each benchmark first. It is then adjusted to the control totals of the benchmark. This procedure is used in the case of 1982, 1987, 2000 and 2010 population censuses.

We need more work to do for the 1995 and 2005 one-percent population sample surveys. In the case of 2005, we can get full-dimensioned matrix directly from the survey, so we do not need to apply IPF for this benchmark. In the case of 1990 sample survey, we get two marginal matrices with different sector classifications: one $s \times g \times a$ matrix with 13 sectors and another $s \times g \times e$ matrix with 75 sectors. We cannot directly apply IPF in this case. Hence, before conducting IPF, we first convert the $s \times g \times a$ matrix into a $g \times a$ matrix by summating the employment numbers in the same $g \times a$ cross-defined dimension. In the case of 1995, there are only 16 sectors available. To satisfy the CIP 37 industry-classification, we have to take the structure from the Wu-Yue estimates for the industrial sectors and make an interpolation between the available 1990 and 2000 benchmarks to obtain a mid-point (1995) structure for the service sectors. Besides, we also decompose services in the 1995 population survey to obtain a sector-specific marginal matrix that is in line with the CIP classification. After these adjustments, we then conduct the same IPF procedure as we do for other benchmarks.

Full-dimensioned compensation matrix for benchmarks

In Table 5 we summarize the available matrices for labor compensation for each benchmark from various sources. Although these matrices appear to be "full-dimensioned", except 1982, some further adjustment is needed to make the number of sectors in line with the CIP classification of 37 industries. Since population censuses do not report labor compensation, we mainly rely on data from two surveys: Chinese Household Income Project (CHIP 1988, 1995 and 2002) and Rural, Urban and Migrant in China Project (RUMiC 2009).[6] However, since the CHIP data do not provide breakdowns for industries within the industrial sector, we adopt the structures of industry-specific compensation estimated in Wu and Yue (2012). Our "control totals" are industry-level labor compensation from the reconstructed annual Chinese input-output tables in Wu and Ito (2014).

The CHIP survey covers all provinces for sampled households in both rural and urban areas whereas the RUMiC project focuses on 9 provinces for samples in three categories, i.e. rural, urban and migrant workers.[7] To us the RUMiC/2009 survey is the only income survey that is available for the estimation of the 2010 benchmark. Although the CHIP project is designed to cover the whole country and the RUMiC project carefully selects 9 provinces for a better representativeness, they cannot be simply viewed as random sampled surveys. For example, the available data set of CHIP/1995 includes 18,925 rural households out of the rural subsample of 34,739 households and 13,873 urban households out of the urban subsample of 21,698 households, with a rural-urban ratio of 1.4. However, according to our reconstructed national employment accounts, the ratio is 2.6 (490 and 190 million workers in rural and urban areas, respectively). Thus, the available CHIP/1995 sample is biased towards urban workers. As for RUMiC/2009, it includes 8,000 rural and 5,000 urban households in the sampled nine provinces, with a rural-urban ratio of 1.6. Besides, since missing values are very common in the RUMiC rural and migrant subsamples, if ignoring them, the results would be biased towards urban workers. To deal with the sampling bias, we adjust the samples of all the available surveys to make them in line with the structure of workers in three categories, i.e. rural, urban unit, which is

[6] CHIP and RUMiC data are now available from the website of China Institute for Income Distribution (CIID).

[7] They are Shanghai, Jiangsu, Zhejiang, and Guangdong from eastern China; Anhui, Henan, and Hubei from central China; Chongqing and Sichuan from western China for the urban and migrant subsample and Shanghai was substituted by Hebei in rural subsample.

a category in the official employment statistics in more details, and urban non-unit, which can be assumed to mainly include migrant workers.

TABLE 5

DATA SOURCE AND DESCRIPTION OF MARGINAL COMPENSATION MATRICES

Benchmark	Data Source	Number of Industries	Available Matrices[1]
1982	3rd population census	16	*s*
	CHIP/1988	Aggregate	$g\times a\times e$
	Wu and Yue (2012)	24	$s\times g\times a\times e$
1987	CHIP/1988	16	$s\times g\times a\times e$
	Wu and Yue (2012)	24	$s\times g\times a\times e$
1995	CHIP/1995 (Urban)	13	$s\times g\times a\times e$
	CHIP/1995 (Rural)	18	$s\times g\times a\times e$
	Wu and Yue (2012)	24	$s\times g\times a\times e$
2000	CHIP/2002 (Urban)	16	$s\times g\times a\times e$
	CHIP/2002 (Rural)	18	$s\times g\times a\times e$
	CHIP/2002 (Migrant)	25	$s\times g\times a\times e$
	Wu and Yue (2012)	24	$s\times g\times a\times e$
2005	1% population sample survey[2]	96	$s\times g\times a\times e$
2010	RUMiC/2009	20	$s\times g\times a\times e$
	1% population sample survey 2005	24	$s\times g\times a\times e$

Sources: As specified in the table.

Note: 1) The available number of industries/sectors (s) is not CIP-classification standard except those from Wu and Yue (2012) which is for the industrial part of the economy only. 2) The available data are 20 percent of the raw data.

Our next task is to fill up the empty cells in the available matrices constructed based on the CHIP and RUMiC data. For example, to construct the full-dimensioned compensation matrix, we need 910 cells ($13 \times 2 \times 7 \times 5$) for the non-industrial sectors (agriculture, construction and 11 services). Taking 1995 as an example, we only have 576 cells filled by data from CHIP/1995. The remaining 334 empty cells are filled by regression-based predictions. For this purpose, we first reclassify the non-industrial sectors in line with the CIP standard.[8] And then using the CHIP and RUMiC data, we run the Mincer equation-type OLS regressions for labor compensations across sectors on gender, age and education dummies for each benchmark year. Next, with the estimated coefficients we predict the missing values in the matrices. Both the regression model and the results are reported in Appendix. The results are finally integrated with the 24 industrial sectors from Wu and Yue (2012), updated.

In the case of the 1982 benchmark, we can only establish a sector-specific marginal matrix from the population census data (Table 5). Considering that the planned employment and compensation systems remained largely intact in the 1980s, we assume that the industrial

[8] The industrial classification of the 2005 survey is much more detailed than that of the CIP, but not in the case of other benchmarks. In our procedure, for all the benchmarks prior to 2005 except for 1982, we reclassify the non-industrial sectors by excluding the (Wu-Yue) 24 industrial sectors from the full-dimensioned compensation matrix. For the last benchmark 2010, we adopt the compensation structure of the 2005 survey.

compensation structure of CHIP/1988 can also be used for this benchmark. This is conducted by the same IPF procedure as we did for the quantitative matrices.

7. Converting Numbers Employed to Hours Worked for Benchmarks

Unlike the numbers employed and labor compensation, we have no “control totals” for hours worked. The final estimates are certainly affected by our choice of average weekly hours worked. The information on hours worked in the official annul statistics is very limited. The occasional available data focus on either state owned enterprises or large firms in selected cities. They are in most cases only available for weekly averages and often without industry breakdowns. Wu and Yue (2010 and 2012) made perhaps the first attempt to estimate industry-specific hours worked within the industrial sector. They rely on studies on “institutional working hours” according to the “day-shift and working-hour” regulations under central planning (Zhu 1999) and assumptions for post-reform changes taking into account widely reported extra working hours among migrant workers.

In this study, we incorporate the Wu-Yue estimates with the available census and labor survey data. Table 6 lists all available marginal matrices for the benchmarks since 1990. There are however no data for our two earlier benchmarks 1982 and 1987. This is understandable because the labor system in the 1980s was still controlled by the planning system and the statistical authority did not bother recording actual hours that might be different from the regulated “institutional working hours”.

Table 6
Data Source and Description of Marginal Matrices for Hours Worked

Benchmark	Data Source	Number of Sectors	Marginal Matrices
1990	CHIP/1988	16	$s \times g \times a \times e$
	Wu and Yue (2012)	24	$s \times g \times a \times e$
1995	CHIP/1995	Aggregate	$a \times e$
	1% population sample survey	16	$s \times g$
	Wu and Yue (2012)	24	$s \times g \times a \times e$
2000	CHIP/2002	Aggregate	$g \times a \times e$
	5th population census	16	$s \times g$
	Wu and Yue (2012)	24	$s \times g \times a \times e$
2005	RUMiC/2007	Aggregate	$g \times a \times e$
	1% population sample survey*	96	$s \times g$
2010	RUMiC/2009	Aggregate	$g \times a \times e$
	6th population census	95	$s \times g$; $g \times a$

Sources: As specified in the table.
Note: *Data available for 20 percent of the raw data.

The method of constructing hours-worked matrix for benchmarks is the same as what used in constructing the employment number and compensation matrices. Before constructing the marginal matrices from the CHIP and RUMiC data, we first need to adjust the rural, urban and migrant parts of the data so that they conform to the structure of our reconstructed national employment accounts. Next, we assess the data quality based on weekly hours worked.

As shown in Table 7, in general the CHIP and RUMiC data show higher weekly hours worked than those of from censuses or population sample surveys, which is supportive to the Wu-Yue conjecture about common extra hours worked among migrant workers. However, there are many data points, especially their migrant subsamples, that show extremely high weekly hours worked that are hardly to be representative, e.g. more than 70 hours a week and no weekends and holidays which is equivalent to annually 3,650 hours.[9] We adjust the marginal matrices obtained from the CHIP and RUMiC data by dropping all the samples with annual hours worked over 3,400 and fill the gaps by regression predictions. The 3,400 hour-threshold is equivalent to about 10 hours per day with two days off per month (for 11 months excluding the month of the Chinese New Year) and ten days annual leave (holidays including CNY).

TABLE 7

WEEKLY HOURS WORKED: CENSUSES VIS-À-VIS CHIP/RUMIC

Weekly Hours Worked	1995	2000	2005	2010
Wu and Yue (2012)[1]	43.0	43.4	46.0	-
Census	40.7	46.8	46.1	45.2
CHIP/RUMiC	45.3[2]	54.3[2,3]	53.1[2]	51.1[2]

Sources: Author's estimates based on data from censuses and CHIP and RUMiC surveys.

Note: 1) Only for the industrial sector; 2) We use CHIP/1995 for 1995, CHIP/2002 for 2000, RUMiC/2007 for 2005 and RUMiC/2009 for 2010; 3) Due to poor quality of the rural data of CHIP/2002, we use the rural part of CHIP 1995 as a substitute.

Specifically, we can only obtain full-dimensioned hours worked matrix directly from CHIP/1988. For other benchmarks, we need to apply the IPF procedure to first construct their marginal matrices and then derive their full matrices. One problem that should be mentioned is that for every of these benchmarks there are generally two data sources, population sample surveys and the CHIP/RUMiC surveys, which give us different values on total numbers of hours worked (Table 7). Since we have no control totals for hours worked matrix to adjust the results by IPF, the choice of which marginal matrix to be the last matrix in the loop of the IPF procedure will affect the results. Convinced by widely reported extra hours worked by migrant workers, now more than 242 millions in China engaged mainly in labor intensive manufacturing and services (*People's Daily*, February 14, 2011), we argue that the CHIP/RUMiC surveys are likely more reliable than other sources and thus use them to control the IPF exercise.

Up to this step, the constructed full-dimensioned hours-worked matrices are not yet in line with the CIP classification for each benchmark. For those benchmarks without detailed industry

[9] This threshold is not implausible. According to a study by Jian and Huang (2007) based on a sample survey of 1000 migrant workers in six locations in north, central and south China, the average of hours worked per day is 10 and the average of days worked is 6.5 per week, and 54.1% workers surveyed did not have national holidays. If we assume that on average they could have one-week holiday in the Chinese New Year, their total hours worked per year will be 3,315. In 2006, the General Union of Labor, Hebei Province, north China, also conducted a survey on 2,200 workers of 56 factories. Its report shows that 12.3% workers worked for 10 to 12 hours per day and 1.8% for over 12 hours per day. Besides, migrant workers in construction worked for 11 hours each day normally and 14 hours in the summer time. The survey also shows that migrant workers in non-state enterprises virtually had no holiday at all (General Union of Labor, Hebei, 2006).

breakdowns such as 2005 and 2010, we first conduct further IPF procedures to repeat what was done in Wu and Yue (2012) using the number of hours worked from the CHIP and RUMiC surveys as "control totals" and the Wu-Yue estimates for 24 industries of the industrial sector as the "control structures", then drop the 24 industries and reclassify the non-industrial sectors in line with the CIP standard and finally integrate them to complete the final matrices.

8. RESULTS AND DISCUSSIONS

Full-dimensioned matrices in time series

Based on the above constructed full-dimensioned employment (numbers and hours) and compensation matrices for benchmarks and marginal matrices in time series, in this section we are now in a position to construct the full-dimensioned employment and compensation matrices in time series. Since our construction of the time-series marginal matrices pursued consistencies in concept, coverage and classification, they should be deemed the best candidate for capturing annual movements of the Chinese workforce and changes in their compensation as given by the annual input-output accounts constructed by Wu and Ito (2014). Thus, the most important issue in the construction of the full-dimensioned matrices in time series is how to best use the marginal matrices in time series and integrate them with the full-dimensioned matrices for the benchmarks.

In the procedure, we mainly take two steps explained as follows. In the first step, we connect the neighboring benchmarks by linear interpolations to obtain all the three full-dimensioned matrices in time series, i.e. numbers, hours and compensation. For the years before the first benchmark of 1982 we assume they are the same as the first benchmark in the case of numbers and compensation, but in the case of hours worked, we use our results for 1990 for reasons discussed in Section 7. However, the time series results do not have the same value as given in the marginal matrices in time series. Our second step is to rescale the non-benchmark estimates to make them consistent with the time series marginal matrices. This not only ensures consistency in values but also in terms of annual changes. The results are full-dimensioned employment, for both numbers employed and hours worked, and compensation matrices in time series for the entire period 1980-2010.

Changes in labor input, quantity, quality and prices

Tables 8a and 8b summarize our estimates of labor inputs and related quantitative and qualitative indicators in annual growth rate and in value or in level index, respectively. To examine the impacts of major policy regime shifts on the labor input in the Chinese economy, it is important to interpret our results over different sub-periods. In these tables we divide the entire period 1980-2010 into five sub-periods, namely 1980-1984, 1985-1991, 1992-2001, 2002-2007 and 2008-2010.

The first sub-period covers early reforms in agriculture which with encouraging positive effects prepared for urban and industrial reforms. The next sub-period covers the nation-wide industrial reform began in 1984-85 with the double-track price reform. This period experienced various institutional shocks as well as the 1989 political turmoil. The third sub-period began with Deng's call for bolder reforms and the official adoption of the "socialist market economy", which kicked off the reforms to the state sector. This was followed by China's post-WTO period during which there were, however, mixed changes. On the one hand, WTO-induced wider opening up to foreign trade and direct investment moved China further towards the market system, but on the other hand, consolidated and enlarged state corporations resurged and meanwhile growth-motivated local governments became more involved in local business. The last sub-period began with the Global Financial Crisis (GFC) in which the role of the state was further enhanced by an unprecedented fiscal injection to rescue the economy from the crisis.

TABLE 8A

CHANGE OF LABOUR INPUT AND BASIC LABOR INDICATORS IN THE CHINESE ECONOMY, 1980-2010

(Percent per annum)

	Labor Input (*d*)	Price of labor input (*p*)	Labor comp. (*c*)	Quality (*q*)	Numbers (*n*)	Hours (*h*)	Hours per week (*hw*)	Hourly comp. (*ch*)
1980-1984	4.44	10.09	14.99	0.97	3.35	3.44	0.09	11.16
1984-1991	3.31	12.92	16.66	0.76	2.48	2.54	0.06	13.78
1991-2001	2.57	15.56	18.53	0.79	1.15	1.77	0.61	16.47
2001-2007	6.57	5.13	12.04	4.93	0.96	1.56	0.59	10.31
2007-2010	5.55	10.53	16.66	4.68	0.59	0.83	0.24	15.70
1980-2010	4.08	11.56	16.11	2.00	1.66	2.03	0.37	13.80

Sources: Authors' estimates.

Notes: The estimation is hours based. Refer to labor indexing for the methodology used. Here: $c=p+d+p*d/100$, $d=q+h+q*h/100$, $c=h+ch+h*ch/100$ and $h=n+hw+n*hw/100$.

TABLE 8B

LEVEL OF LABOUR INPUT AND BASIC LABOR INDICATORS OF THE CHINESE ECONOMY, SELECTED YEARS

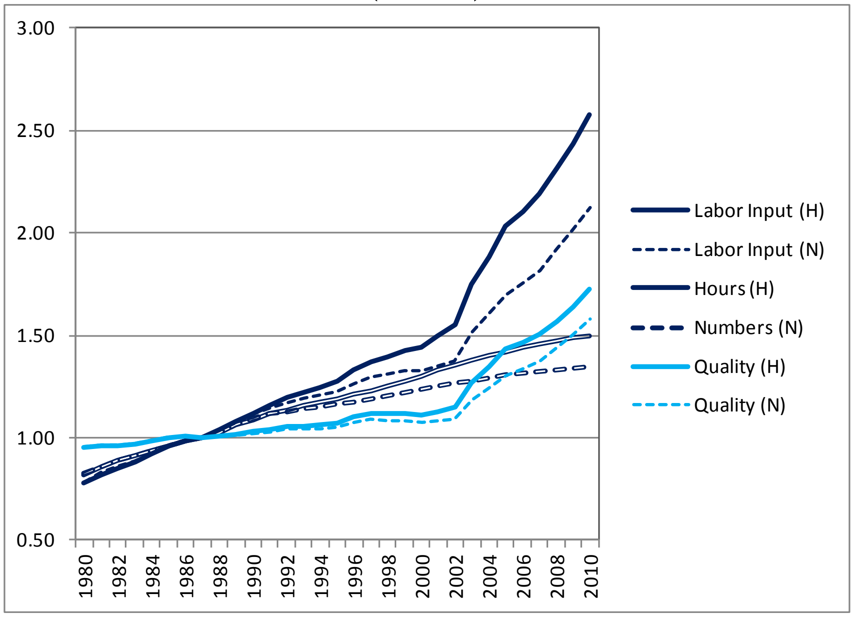
	Labor input (1987=1)	Price (1987=1)	Labor comp. (bl. yuan)[1]	Quality (1987=1)	Numbers (x1000)	Hours (ml. hours)	Hours per week[2]	Hourly comp. (yuan)[3]
1980	0.8	0.4	200	0.9	479,120	1,170,661	48.9	0.2
1984	0.9	0.6	350	1.0	546,710	1,340,410	49.0	0.3
1991	1.2	1.5	1,030	1.0	648,980	1,597,456	49.2	0.6
2001	1.5	6.2	5,638	1.1	727,850	1,903,332	52.3	3.0
2007	2.2	8.4	11,150	1.5	770,850	2,088,768	54.2	5.3
2010	2.6	11.4	17,704	1.7	784,560	2,141,154	54.6	8.3

Sources: Authors' estimates.

Notes: 1) In current RMB yuan. 2) Estimated assuming 50 working weeks per year. 3) In current RMB yuan.

As shown in Table 8a, on average China's labor input grew by about 4.1 percent per annum for the full period, to which changes in hours worked and changes in quality almost made the same contribution by about 2 percent. The results suggest that ignoring the contribution of the quality change in China's workforce would underestimate the labor input in the economy, hence exaggerating total factor productivity, *ceteris paribus*. The total compensation to labor grew at a fast rate of over 16 percent a year, of which over 70 percent is attributed to the rise of prices for labor input. Meanwhile, the average labor compensation rose while hours worked per person engaged also increased. As Table 8b shows, the number of hours worked per person increased from 2,443 in 1980 to 2,729 in 2010 (derived from total numbers and total hours), opposite to the general trend of declining hours per person in the course of economic development. Besides, the hourly compensation rose by 13.8 percent per year or from 0.2 yuan in 1980 to 8.3 yuan in 2010, and hours worked per week increased by about 0.4 percent per year or from 48.9 hours in 1980 to 54.6 hours in 2010. In order to trace the impacts we also demonstrate the annual changes of these indicators in Figure 1.

FIGURE 1
INDICES OF QUANTITY AND QUALITY OF EMPLOYMENT AND LABOUR INPUT IN THE CHINESE ECONOMY
(1987 = 1)



Sources: Authors' estimates.

The agricultural reform period in 1980-84 which also triggered the unprecedented growth in rural enterprises saw the most rapid growth in numbers employed by about 3.4 percent per year. This was followed by the early industrial reform in 1984-1991 which experienced the growth of numbers employed by 2.5 percent per year. In both periods, the growth rates of numbers and hours were similar, suggesting that the quantity change was almost entirely attributed to the increase in the number of workers employed rather than more hours worked by each worker.

This situation began to change since the early 1990s following deepening reforms, especially the restructuring of the state-owned enterprises, following Deng's reform push. As Table 8a shows, it appears that the employment system was substantially impacted by reforms introduced in 1991-2001. While the price of labor input underwent the most rapid rise in the entire period by 15.6 percent per year, the labor input had the slowest growth of 2.6 percent per year. The hourly compensation also experienced the most rapid growth by about 16.5 percent per year, so did the hours worked per week by 0.6 percent per annum. However, the increase of quality still appeared to be very slow by less than 0.8 percent per annum.

Nevertheless, as shown in Table 8a China's post-WTO period (2001-07) completely turned the situation around. While China's labor input grew by about 6.6 percent per year compared to 2.6 percent per years in 1991-2001, which was the fastest ever, the growth of prices significantly slowed down to only about 5 percent per annum compared to nearly 16 percent per year in 1991-2201, which was the record low rate over the entire reform period. At the same time, the growth of hourly pay and the growth of total compensation also experienced the slowest growth ever. Even so, 75 percent of the most rapid labor input growth was due to the growth of labor quality or the increase of skill, knowledge and experience. Other things being equal, such changes could significantly raise China's competitiveness in labor intensive industries, which was indeed what happened. This is perhaps the main reason for China to quickly emerge as the world factory during this period.

The most distinct feature of the post-GFC period in 2008-10 is that it was the first time for China to experience the slowest growth in the number employed (0.6 percent per annum) and in the hours worked (0.8 percent per annum), though the latter still grew slightly faster than the former. Meanwhile, the growth of hours per week significantly slowed down to about 0.2 from about 0.6 percent per annum in the post-WTO period. This may signal that China had nearly exhausted its demographic bonus. Our estimated labor input growth for this short period is still fast or the second fastest over the whole period, but nearly 85 percent of the growth was attributable to the growth of labor quality. Thus, it is not surprisingly to see that hourly compensation accelerated dramatically to 15.7 percent per annum from about 10.3 in the post-WTO period. Such changes will significantly affect China's competitiveness in the years to come. Indeed, as Table 8b shows, the hourly pay jumped from 3 yuan at the eve of China's WTO entry in 2001 to 5.3 yuan in 2007 and further to 8.3 yuan in 2010.

Impact of human capital factors

We have also estimated the main effect of each human capital factor, i.e. industry/sector, gender, age and education, upon the quality change of the Chinese workforce in Table 8a. The results are reported in Table 9 and the dynamic change of each factor is given in Figure 2.

TABLE 9

CHANGES OF LABOR QUALITY BY FACTOR IN THE CHINESE ECONOMY, 1980-2010

(Percent per annum)

	Quality	Sector (*s*)	Age (*a*)	Education (*e*)	Gender (*g*)
1980-1984	0.97	1.10	0.26	0.36	-0.02
1984-1991	0.76	0.56	0.54	-0.02	-0.02
1991-2001	0.79	0.08	0.93	0.08	0.02
2001-2007	4.93	3.00	0.36	1.24	0.00
2007-2010	4.68	0.94	0.31	3.14	-0.16
1980-2010	2.00	0.99	0.58	0.63	-0.02

Sources: Authors' estimates.

In general, structural changes of the economy, which is captured by "sector", made the largest or half contribution to the overall quality change of the Chinese workforce.[10] This was followed by changes in education attainment and age structure. We observe the important positive impacts of structural changes on the quality of the workforce in the earlier agricultural reform period (1.1 percent per annum), especially in the post-WTO period in 2001-07 (3 percent). This suggests that workers were reallocated to industries with higher labor productivity

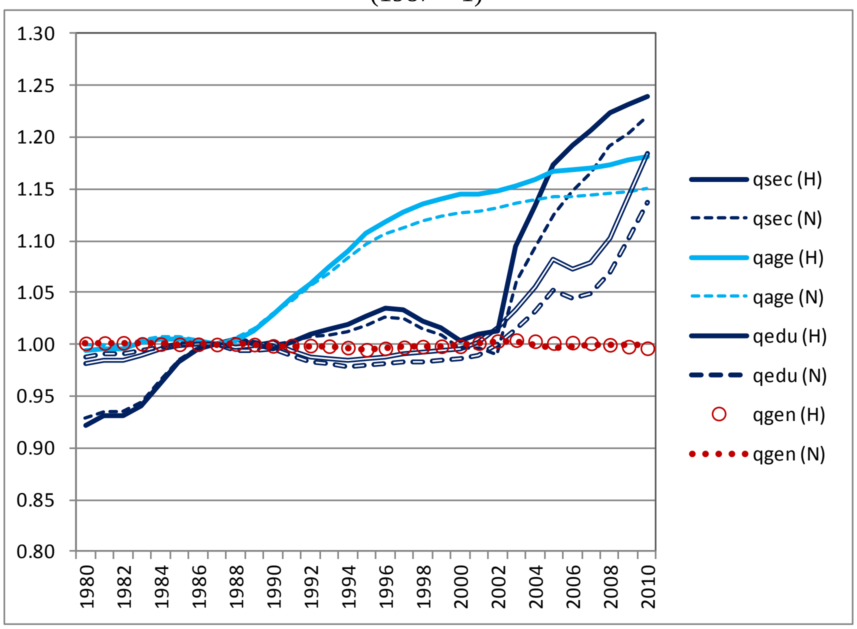
[10] For the results for individual sectors see Tables A4a and A4b.

or faster labor productivity growth. The effect of seniority (age) made the most important contribution in the deepening industrial and SOE reform period in 1991-2001 (0.9 percent per annum), perhaps reflecting age-embodied experiences responded better in the early period of transition. The role of education in quality change appeared to rise first following China's WTO entry (1.2 percent per annum), but then most importantly, in the wake of global financial crisis (3.1 percent). In the case of the post-WTO period, this effect should be mainly attributed to demand for better educated workers in a more world market-driven restructuring, whereas in the case of the post-GFC period, it could be caused by the technological upgrading of manufacturing industries and perhaps also by the growth of business services as well as government sectors. Figure 2 shows that the impacts of WTO and GFC appear to be evident in the effect of structural change and in the effect of education attainment change, respectively. On the other hand, the quality effect of changes in age structure, especially in gender, was insignificant.

FIGURE 2

QUALITY INDEX OF EMPLOYMENT BY FACTOR IN THE CHINESE ECONOMY

(1987 = 1)



Sources: Authors' estimates.

Impact of changes in the gender ratio

The results reported in Tables 10a and 10b allow us to have a closer look at the impact of changes in the gender ratio on the Chinese workforce by holding other attributes constant. Table 10a shows gender differences in the growth rates of labor input and related indicators. It appears that changes in hours worked per week and hourly compensation from the gender perspective have important implications. The annual growth of the hourly compensation of female workers (13.4 percent per annum) was slower than their male counterparts (14 percent) whereas the annual growth of the weekly hours worked of the former (0.5) was faster than the latter (0.3 percent). These are reflected by our level estimates reported in Table 10b. By 2010, on the average of labor compensation (wage plus welfare), male workers were paid for 9.8 yuan per hours and female workers were only paid 6.3 yuan per hour or less than two thirds of the former.

Meanwhile, female workers worked longer hours than their male counterparts or 55.2 compared to 54.1 per week in the case of male workers. Consequently, female workers accounted for 43 percent of the total numbers employed but only 33 percent of the total labor compensation. On average, this seems to suggest that female workers were over supplied and not as competitive as their male counterparts in the labor market.

TABLE 10A

GROWTH OF LABOUR INPUT AND BASIC LABOR INDICATORS BY GENDER, 1980-2010

(Percent per annum)

Gender	Labor Input (*d*)	Price of labor input (*p*)	Labor comp. (*c*)	Quality (*q*)	Numbers (*n*)	Hours (*h*)	Hours per week (*hw*)	Hourly comp. (*ch*)
Males	3.99	11.78	16.24	2.00	1.64	1.95	0.30	14.02
Females	4.26	11.13	15.87	2.07	1.68	2.15	0.46	13.43
Total	***4.08***	***11.56***	***16.11***	***2.00***	***1.66***	***2.03***	***0.37***	***13.80***

Sources: Authors' estimates.

Note: See Table 8a.

TABLE 10B

THE 2010 LEVEL OF LABOUR INPUT AND BASIC LABOR INDICATORS BY GENDER

Gender	Labor input (1987=1)	Price (1987=1)	Compensation (bl. yuan)	Quality (1987=1)	Numbers (x1000)	Hours (ml. hours)	Hours per week	Hourly comp. (yuan)
Males	2.5	12.7	11,837	1.7	445,859	1,206,393	54.1	9.8
Females	2.6	9.3	5,867	1.7	338,701	934,761	55.2	6.3
Total	***2.6***	***11.4***	***17,704***	***1.7***	***784,560***	***2,141,154***	***54.6***	***8.3***

Sources: Authors' estimates.

Note: See Table 8b.

Impact of changes in age structure

The results reported in Tables 11a and 11b allow us to have a closer look at the impact of changes in the age structure on the Chinese workforce by holding other attributes constant. Perhaps, the most distinct finding in Table 11a is that the aged 25-29 group experienced the slowest growth (about 0.2 per annum) among all age groups (except the aged 16-19 group who declined), but at the same time it also showed the most rapid growth in hourly compensation (15.4 percent) and labor quality (1.9 percent). Besides, over 70 percent of the labor input growth by this group was attributed to quality change rather than the growth of numbers or hours (based on Table 11a). In 2010, as Table 8b shows, this group was paid for 10.6 yuan per hour or the highest and worked for 52.1 hours per week or the shortest among all groups.

By contrast, the older groups experienced just the opposite, particularly the aged 40-49 and followed by the aged 54+ groups. The growth of their labor input was largely attributed to the growth of quantity rather than the growth of quality, 24 and 13 percent, respectively (based on Table 11a). According to Table 11b, in 2010 the aged 40-49 group was were paid nearly as high as the aged 25-29 group. Besides, both the aged 40-49 and 54+ groups worked for 55.9 hours per week, the longest hours worked among all groups. The striking differences between the aged 25-29 group and the older groups in terms of the contribution by quality and the similarity in terms of compensation are provoking and worthwhile for further investigations, but it is beyond the scope of this data construction paper.

TABLE 11A

GROWTH OF LABOUR INPUT AND BASIC LABOR INDICATORS BY AGE GROUP, 1980-2010

(Percent per annum)

Age Group	Labor Input (*d*)	Price of labor input (*p*)	Labor comp. (*c*)	Quality (*q*)	Numbers (*n*)	Hours (*h*)	Hours per week (*hw*)	Hourly comp. (*ch*)
16-19	-2.51	13.24	10.39	1.34	-3.93	-3.80	0.14	14.75
20-24	2.77	11.64	14.73	1.38	0.98	1.36	0.38	13.19
25-29	2.54	13.27	16.16	1.85	0.48	0.68	0.19	15.37
30-39	4.26	10.34	15.03	1.68	2.11	2.54	0.42	12.19
40-49	5.40	11.52	17.55	1.29	3.41	4.06	0.63	12.96
50-54	4.41	10.67	15.55	1.38	2.78	2.99	0.20	12.20
54+	4.80	12.63	18.04	0.61	3.46	4.17	0.69	13.32
Total	***4.08***	***11.56***	***16.11***	***2.00***	***1.66***	***2.03***	***0.37***	***13.80***

Sources: Authors' estimates.
Note: See Table 8a.

TABLE 11B

THE 2010 LEVEL OF LABOUR INPUT AND BASIC LABOR INDICATORS BY AGE GROUP

Age Group	Labor input (1987=1)	Price (1987=1)	Compensation (bl. yuan)	Quality (1987=1)	Numbers (x1000)	Hours (ml. hours)	Hours per week	Hourly comp. (yuan)
16-19	0.4	17.0	260	1.4	24,324	68,308	56.2	3.8
20-24	1.1	13.7	1,281	1.4	88,910	241,251	54.3	5.3
25-29	1.9	18.1	2,597	1.4	94,259	245,626	52.1	10.6
30-39	2.5	8.6	3,781	1.6	207,199	569,675	55.0	6.6
40-49	4.5	11.0	6,036	1.5	214,682	599,736	55.9	10.1
50-54	3.0	7.9	1,213	1.5	60,235	151,349	50.3	8.0
54+	3.3	12.5	2,535	1.2	94,951	265,209	55.9	9.6
Total	***2.6***	***11.4***	***17,704***	***1.7***	***784,560***	***2,141,154***	***54.6***	***8.3***

Sources: Authors' estimates.
Note: See Table 8b.

Impact of changes in education attainment

We report estimates for the education dimension in Tables 12a and 12b. They allow us to have a closer look at the impact of changes in the structure of the Chinese workforce by education attainment holding other attributes constant. The most striking finding in Table 12a is that the labor input by the group with college and above-college level (college level hereafter) education grew most rapidly over the entire period, about 10.9 percent per annum compared to the average growth of 4.1 percent per annum. It is also striking to see that the growth of the quantity (11.3 percent) was even faster than that of labor input by this group, suggesting that there was no contribution at all by the quality change of this group. We consider that this was largely due to the oversupply of the workers with the college-level education that obstructed the growth of their compensation, though their hourly pay was already the highest in 2010 (19.1 yuan compared to the average of 8.3 yuan, Table 12b). This problem is rooted in the structural problems of the economy in general which are responsible for the misallocation of resources and the mismatch between the education industry (supply) and the goods and market service industries (demand) in particular.

Another important finding in Table 12b is that by 2010 the workers with junior-high education attainment, who were very likely to start working at age 16, still accounted for 46 percent of the total numbers employed and 51 percent of the total hours worked. On average, they worked for 60.3 hours per week, the highest of all education groups, but were only paid by 6 yuan per hour, which were the lowest except for the illiterate group.

TABLE 12A

GROWTH OF LABOUR INPUT AND BASIC LABOR INDICATORS BY EDUCATION ATTAINMENT, 1980-2010

(Percent per annum)

Education Level	Labor Input (*d*)	Price of labor input (*p*)	Labor comp. (*c*)	Quality (*q*)	Numbers (*n*)	Hours (*h*)	Hours per week (*hw*)	Hourly comp. (*ch*)
Illiteracy	-4.54	12.12	7.03	1.16	-5.78	-5.64	0.15	13.42
Primary	1.10	12.30	13.53	1.44	-0.19	-0.34	-0.15	13.92
Junior High	4.94	10.91	16.38	0.70	3.33	4.21	0.86	11.68
Senior High	4.16	10.74	15.34	0.89	2.49	3.24	0.72	11.72
College & +	10.85	14.32	26.72	-0.36	10.74	11.25	0.46	13.90
Total	***4.08***	***11.56***	***16.11***	***2.00***	***1.66***	***2.03***	***0.37***	***13.80***

Sources: Authors' estimates.
Note: See Table 8a.

TABLE 12B

THE 2010 LEVEL OF LABOUR INPUT AND BASIC LABOR INDICATORS BY EDUCATION ATTAINMENT

Education Level	Labor input (1987=1)	Price (1987=1)	Compensation (bl. yuan)	Quality (1987=1)	Numbers (x1000)	Hours (ml. hours)	Hours per week	Hourly comp. (yuan)
Illiteracy	0.2	6.9	247	1.2	19,689	54,014	54.9	4.6
Primary	1.0	15.2	2,526	1.4	150,823	364,623	48.4	6.9
Junior High	3.2	11.5	6,476	1.3	359,653	1,084,913	60.3	6.0
Senior High	2.7	11.6	2,837	1.3	127,936	343,175	53.6	8.3
College & +	12.4	23.0	5,618	0.9	126,459	294,430	46.6	19.1
Total	***2.6***	***11.4***	***17,704***	***1.7***	***784,560***	***2,141,154***	***54.6***	***8.3***

Sources: Authors' estimates.
Note: See Table 8b.

9. CONCLUDING REMARKS

China's economic reform in the past over three decades not only brought this world most populous country out of a low income trap but also trigged perhaps an unprecedented transformation of labor in human history in both quantitative and qualitative terms. In this measurement-oriented, data-driven study, using labor indexing methodology, we construct China's employment and compensation metrics and measure labor input in the Chinese economy. Our pains-taking data work is based on our understanding of the major conceptual, coverage and classification problems in the official labor statistics. We first adjust breaks, fill gaps and restore consistencies in official annual statistics and benchmark marginal matrices, and then use the IPF approach to obtain full-dimensioned employment, for both numbers engaged and hours worked, and compensation matrices that are controlled by the national income accounts for the benchmarks. We finally obtain the full-dimensioned matrices in time series by interpolating the benchmark results controlled by the available marginal matrices in time series.

Despite insufficient data, the results appear to be plausible. We find that China's total labor input grew by 4.1 percent per annum in the period 1980-2010, of which changes in quality and in hours worked made almost equal contributions (about 2 percent). Besides, nearly half of the 2-percent quality improvement was attributable to the changes in industrial structure, followed by the changes in education attainment and age structure. The observed strongest sectoral effect on quality change mainly reflects substantial corrections to the distortions in resource allocation under central planning thanks to market-oriented reforms and China's WTO entry following which workers moved to sectors or industries with higher labor productivity or faster labor productivity growth. Furthermore, changes in the age structure made important contribution in the earlier reform period, reflecting age-embodied experiences responded better in the early reform period, whereas changes in the structure of education attainment had more impact following China's WTO entry and even in the wake of the GFC, suggesting higher education responded better in a more international trade-exposed economy and perhaps competition-induced technological upgrading.

We also find that in general hours worked grew more rapidly than numbers engaged across board over the entire period, reflecting the persistent dominance of labor-intensive manufacturing in the economy. This observation is further supported by the negative gender effect upon the quality change. Although female workers accounted for 44 percent of Chinese workforce, they worked longer hours than their male counterparts but only paid two thirds of the latter.

APPENDIX

Industrial classification

TABLE A1

CIP INDUSTRIAL CLASSIFICATION AND CODE

CIP	EU-KLEMS	CSIC/2011	National Accounts	Sector
1	AtB	01t05 (A)	I	Agriculture, forestry, animal husbandry & fishery
2	10	06 (B)	II.1	Coal mining
3	11	07 (B)	II.1	Oil & gas excavation
4	13	08t09 (B)	II.1	Metal mining
5	14	10t11 (B)	II.1	Non-metallic minerals mining
6	15	12t14 (C)	II.1	Food and kindred products
7	16	15 (C)	II.1	Tobacco products
8	17	16 (C)	II.1	Textile mill products
9	18	17 (C)	II.1	Apparel and other textile products
10	19	18 (C)	II.1	Leather and leather products
11	20	19t20 (C)	II.1	Saw mill products, furniture, fixtures
12	21t22	21t22 (C)	II.1	Paper products, printing & publishing
13	23	24 (C)	II.1	Petroleum and coal products
14	24	25t27 (C)	II.1	Chemicals and allied products
15	25	28t29 (C)	II.1	Rubber and plastics products
16	26	30 (C)	II.1	Stone, clay, and glass products
17	27t28	31t32 (C)	II.1	Primary & fabricated metal industries
18	27t28	33 (C)	II.1	Metal products (excluding rolling products)
19	29	34t35 (C)	II.1	Industrial machinery and equipment
20	31	37 (C)	II.1	Electric equipment
21	32	38 (C)	II.1	Electronic and telecommunication equipment
22	30t33	39 (C)	II.1	Instruments and office equipment
23	34t35	36 (C)	II.1	Motor vehicles & other transportation equipment
24	36t37	23,40t41 (C)	II.1	Miscellaneous manufacturing industries
25	E	42t44 (D)	II.1	Power, steam, gas and tap water supply
26	F	45t48 (E)	II.2	Construction
27	G	61t62 (H)	III.2	Wholesale and retail trades
28	H	63t64 (I)	III.3	Hotels and restaurants
29	I	49t57 (F)	III.1	Transport, storage & post services
30	71t74	58t60 (G)	III.6	Information & computer services
31	J	65t68 (J)	III.4	Financial Intermediations
32	K	69 (K)	III.5	Real estate services
33	71t74	70t75 (L,M)	III.6	Leasing, technical, science & business services
34	L	76t78 (N) 90t95 (S,T)	III.6	Government, public administration, and political and social organizations, etc.
35	M	81 (P)	III.6	Education
36	N	82t84 (Q)	III.6	Healthcare and social security services
37	O&P	79t80 (O) 85t89 (R)	III.6	Cultural, sports, entertainment services; residential and other services

Source: Wu and Ito (2014).

Broad sector control totals

TABLE A2

ADJUSTED NUMBERS EMPLOYED BY MAJOR SECTORS OF THE CHINESE ECONOMY

(Thousand 1990 yuan, midyear)

	Total	I	I*	II.1	II.1*	II.2	III.1-3	III.1-3*	III.4-6	III.4-6*
1980	47,912	28,878	28,551	10,793	10,688	1,025	3,607	3,607	3,609	4,041
1981	50,090	29,450	29,116	11,747	11,641	1,093	3,989	3,989	3,812	4,251
1982	51,915	30,318	29,975	12,245	12,137	1,183	4,244	4,244	3,925	4,375
1983	53,320	31,005	30,654	12,458	12,353	1,331	4,459	4,459	4,066	4,522
1984	54,672	31,010	30,659	12,668	12,556	1,628	4,828	4,828	4,537	5,000
1985	56,024	30,999	30,648	12,767	12,653	2,060	5,264	5,264	4,934	5,399
1986	57,209	31,193	30,410	12,984	12,865	2,399	5,533	5,533	5,100	6,002
1987	58,242	31,459	30,669	13,173	13,048	2,614	5,652	5,652	5,344	6,259
1988	59,815	31,956	31,163	13,517	13,395	2,779	5,936	5,936	5,627	6,542
1989	61,908	32,737	31,949	14,191	14,071	2,813	6,315	6,315	5,852	6,760
1990	63,707	33,855	32,889	14,350	14,234	2,773	6,961	6,961	5,769	6,851
1991	64,899	34,717	33,761	14,185	14,069	2,785	7,009	7,009	6,203	7,276
1992	65,600	34,921	34,112	14,163	14,047	2,882	6,665	6,665	6,970	7,895
1993	66,480	34,580	33,648	14,079	13,961	3,171	6,793	6,793	7,857	8,907
1994	67,132	33,934	32,974	13,989	13,876	3,436	7,128	7,128	8,644	9,718
1995	67,760	33,237	32,513	13,893	13,766	3,565	7,691	7,691	9,373	10,226
1996	68,508	32,686	32,298	13,889	13,771	3,741	8,185	8,185	10,007	10,512
1997	69,385	32,648	32,348	13,924	13,804	3,932	8,463	8,463	10,418	10,838
1998	70,229	33,100	32,530	13,680	13,548	4,191	8,415	8,415	10,842	11,544
1999	71,016	33,830	32,719	13,194	13,102	4,429	8,398	8,780	11,164	11,986
2000	71,740	34,539	32,734	12,684	12,605	4,555	8,393	9,287	11,569	12,559
2001	72,555	35,150	32,400	12,330	12,259	4,680	8,295	9,884	12,100	13,333
2002	73,383	35,759	31,875	11,927	11,869	4,725	8,426	10,864	12,545	14,050
2003	74,086	35,984	31,240	11,893	11,843	4,510	8,675	11,759	13,024	14,734
2004	74,816	35,404	30,297	12,644	12,595	4,184	8,854	12,189	13,730	15,551
2005	75,513	34,188	29,158	13,569	13,569	4,221	8,862	12,035	14,672	16,530
2006	76,113	32,760	28,033	14,451	14,451	4,548	8,851	11,732	15,502	17,348
2007	76,695	31,428	26,985	15,418	15,418	4,908	8,895	11,523	16,045	17,861
2008	77,235	30,407	26,198	16,132	16,132	5,189	9,008	11,429	16,500	18,288
2009	77,738	29,474	25,474	16,311	16,311	5,583	9,219	11,426	17,150	18,943
2010	78,138	28,451	24,670	16,515	16,515	6,101	9,347	11,343	17,724	19,509

Source: See text.

Results of Mincer Regression

In constructing the full-dimensioned compensation matrix, we have missing values in the resulting matrix. In order to make the matrix to be a full one, we run Mincer regression for based on the resulting matrix and use the predictions to fill the missing values. To run Mincer regression, we estimate the following model:

$$lncomp_i = \alpha + \sum_{j=2}^{37} \beta_j 1\{s_i = j\} + \gamma 1\{g_i = 2\} + \sum_{l=2}^{7} \delta_l 1\{a_i = l\} + \sum_{k=2}^{5} \eta_k 1\{e_i = k\} + \epsilon_i$$

where $lncomp$ is the logarithmic value of compensation, s is the code of sector, g is the code of gender, a is the code of age, e is the code of education and $1\{\cdot\}$ is the indicator function. See Table 1 for definition of j, l and k.

We estimate the above model for 1987 based on CHIP1988, 1995 based on CHIP1995, 2000 based on CHIP2002, 2005 based on 1% population sample survey and 2010 based on RUMiC2009. We report the results in Table 9

TABLE A3

MINCER REGRESSION FOR 1987, 1995, 2000, 2005 AND 2010

(Dependent variable: Labor compensation)

Variables	1987	1995	2000	2005	2010
Female	-0.094***	-0.149***	-0.147***	-0.254***	-0.297***
	(0.008)	(0.005)	(0.000)	(0.000)	(0.007)
20-24	0.472***	0.401***	0.218***	0.156***	0.438***
	(0.025)	(0.021)	(0.003)	(0.000)	(0.045)
25-29	0.437***	0.562***	0.276***	0.249***	0.889***
	(0.025)	(0.021)	(0.003)	(0.000)	(0.044)
30-39	0.723***	0.806***	0.370***	0.309***	1.058***
	(0.023)	(0.020)	(0.003)	(0.000)	(0.044)
40-49	0.943***	0.991***	0.471***	0.318***	1.018***
	(0.023)	(0.020)	(0.003)	(0.000)	(0.044)
50-54	0.953***	0.987***	0.541***	0.304***	0.935***
	(0.026)	(0.021)	(0.003)	(0.000)	(0.045)
54+	0.995***	0.798***	0.598***	0.159***	0.880***
	0.028)	(0.022)	(0.003)	(0.000)	(0.045)
Primary school	-0.034	0.223**	-0.270***	0.194***	0.224***
	(0.033)	(0.035)	(0.000)	(0.000)	(0.045)
Junior High School	0.063**	0.481***	-0.351***	0.335***	0.217***
	(0.031)	(0.033)	(0.001)	(0.000)	(0.042)
Senior High School	0.131***	0.633***	-0.027***	0.531***	0.423***
	(0.031)	(0.033)	(0.002)	(0.000)	(0.042)
Tertiary	0.225***	0.766***	0.311***	1.025***	0.824***
	(0.032)	(0.033)	(0.004)	(0.000)	(0.041)
Constant	4.141***	7.342***	7.798***	7.773***	6.288***
	(0.049)	(0.042)	(0.005)	(0.000)	(0.057)
Industry sectors included	No	No	Yes	Yes	No
N	10867	9126	430052	1388164	5724
R^2	0.2813	0.6119	0.9101	0.9828	0.7074
Adjusted R^2	0.2798	0.6109	0.9101	0.9828	0.7062

Note: Standard errors are in parentheses. ***Indicates the coefficient is significant at 1%, ** at 5% and * at 10%.

Results by industry

TABLE A4A

ANNUAL CHANGE OF LABOUR INPUT AND BASIC LABOR INDICATORS BY INDUSTRY, 1980-2010

(Percent per annum)

	CIP Industry	Labor Input (*d*)	Price of labor input (*p*)	Labor comp. (*c*)	Quality (*q*)	Numbers (*n*)	Hours (*h*)	Hours per week (*hw*)	Hourly comp. (*ch*)
1	Agriculture	-0.33	14.06	13.68	0.17	-0.44	-0.50	-0.06	14.26
2	Coal mining	2.50	12.48	15.30	0.45	0.97	2.04	1.06	12.99
3	Oil & gas extraction	5.75	13.68	20.22	1.71	3.34	3.98	0.61	15.62
4	Metal mining	3.88	14.85	19.31	0.61	1.98	3.24	1.24	15.56
5	Non-metal mining	-0.48	15.34	14.79	0.55	-2.24	-1.02	1.24	15.98
6	Food products	1.87	17.34	19.54	0.46	0.66	1.40	0.74	17.88
7	Tobacco products	3.80	19.44	23.98	1.80	1.61	1.97	0.35	21.59
8	Textile products	2.73	9.48	12.47	0.26	1.51	2.46	0.94	9.76
9	Apparels	4.11	13.44	18.09	0.08	2.67	4.02	1.31	13.53
10	Leather products	5.00	11.44	17.02	0.04	3.56	4.96	1.36	11.49
11	Saw mill & furniture	1.32	18.00	19.56	0.43	-0.39	0.89	1.28	18.51
12	Paper, printing & publishing	3.38	12.08	15.87	1.01	1.28	2.35	1.05	13.21
13	Petroleum & products	6.02	13.82	20.67	1.41	3.91	4.55	0.61	15.43
14	Chemicals	4.45	11.97	16.95	0.91	2.77	3.51	0.72	12.99
15	Rubber & plastics	4.35	11.70	16.55	0.41	2.79	3.92	1.10	12.15
16	Building materials	1.91	14.37	16.55	0.50	0.24	1.40	1.15	14.93
17	Basic metals	3.68	12.22	16.35	0.67	2.16	3.00	0.82	12.96
18	Metal products	2.52	13.24	16.09	0.36	0.98	2.15	1.16	13.65
19	Machinery	2.19	12.25	14.72	0.43	0.97	1.75	0.77	12.74
20	Electric equipment	6.20	10.96	17.85	0.50	4.76	5.68	0.88	11.52
21	Electronic equipment	8.11	13.23	22.41	0.42	6.78	7.66	0.83	13.70
22	Instruments & office equip.	4.56	10.44	15.48	0.62	3.21	3.91	0.68	11.13
23	Transport equipment	4.62	17.33	22.75	0.53	3.41	4.06	0.63	17.96
24	Other manufactures	-0.33	13.80	13.43	0.33	-2.07	-0.65	1.45	14.17
25	Utilities	6.11	11.87	18.71	1.44	4.16	4.60	0.43	13.49
26	Construction	7.03	8.74	16.38	0.38	6.13	6.63	0.47	9.15
27	Wholesale & retail	5.64	10.67	16.91	0.70	4.26	4.91	0.62	11.44
28	Hotels & catering	7.47	7.71	15.76	0.26	6.52	7.20	0.64	7.99
29	Transport services	2.81	12.57	15.74	0.67	1.57	2.13	0.55	13.33
30	Tele. & computer services	8.06	14.95	24.21	2.52	4.98	5.40	0.41	17.84
31	Financial services	11.78	8.67	21.47	2.31	8.64	9.25	0.56	11.18
32	Real estate services	14.20	1.83	16.28	1.76	11.52	12.22	0.63	3.62
33	Business services	7.75	10.07	18.59	1.35	5.43	6.31	0.83	11.55
34	Government	9.50	7.39	17.59	2.14	6.77	7.21	0.42	9.68
35	Education	6.25	10.86	17.80	2.35	3.16	3.82	0.64	13.46
36	Healthcare & social security	7.35	10.74	18.89	2.24	4.35	5.00	0.62	13.23
37	Other services	7.60	8.50	16.74	0.77	5.97	6.77	0.76	9.33
	Total economy	***4.08***	***11.56***	***16.11***	***2.00***	***1.66***	***2.03***	***0.37***	***13.80***

Sources: Authors' estimates.

Notes: The estimation is hours based. Refer to labor indexing for the methodology used. Here: $c=p+d+p*d/100$, $d=q+h+q*h/100$, $c=h+ch+h*ch/100$ and $h=n+hw+n*hw/100$.

TABLE A4B

THE 2010 LEVEL OF LABOUR INPUT AND BASIC LABOR INDICATORS BY INDUSTRY

	CIP Industry	Labor input (1987=1)	Price (1987=1)	Compensation (bl. yuan)	Quality (1987=1)	Numbers (x1000)	Hours (ml. hours)	Hours per week	Hourly comp. (yuan)
1	Agriculture	0.8	16.2	3,844	1.0	249,890	639,355	51.2	6.0
2	Coal mining	1.6	30.4	461	1.2	7,690	21,544	56.0	21.4
3	Oil & gas extraction	3.4	69.2	156	1.7	1,100	2,852	51.8	54.7
4	Metal mining	2.3	45.8	150	1.3	1,550	4,331	55.9	34.7
5	Non-metal mining	0.6	73.7	81	1.2	2,750	7,814	56.8	10.4
6	Food products	1.4	46.6	441	1.2	12,170	34,811	57.2	12.7
7	Tobacco products	1.5	113.9	64	1.7	210	553	52.6	115.4
8	Textile products	1.4	14.7	241	1.0	14,490	43,508	60.1	5.5
9	Apparels	2.6	19.7	169	1.0	12,100	36,860	60.9	4.6
10	Leather products	2.9	19.8	81	1.0	6,190	19,058	61.6	4.3
11	Saw mill & furniture	3.0	18.7	142	1.1	7,630	22,151	58.1	6.4
12	Paper, printing & publishing	1.9	17.9	131	1.2	5,850	16,835	57.6	7.8
13	Petroleum & products	3.7	82.7	203	1.6	980	2,592	52.9	78.5
14	Chemicals	2.6	20.4	389	1.4	8,370	23,068	55.1	16.9
15	Rubber & plastics	2.5	18.5	150	1.1	9,440	27,621	58.5	5.4
16	Building materials	1.0	28.6	331	1.2	8,360	24,347	58.2	13.6
17	Basic metals	2.3	31.7	398	1.3	5,640	15,747	55.8	25.3
18	Metal products	1.7	20.7	171	1.2	7,890	22,884	58.0	7.5
19	Machinery	1.6	25.3	475	1.2	11,850	33,436	56.4	14.2
20	Electric equipment	4.3	13.2	231	1.2	11,010	31,658	57.5	7.3
21	Electronic equipment	7.4	22.1	374	1.2	10,230	29,469	57.6	12.7
22	Instruments & office equip.	3.3	17.1	57	1.3	2,090	5,894	56.4	9.6
23	Transport equipment	3.3	37.3	565	1.3	6,890	19,206	55.7	29.4
24	Other manufactures	0.8	12.2	51	1.1	6,700	19,693	58.8	2.6
25	Utilities	4.4	30.7	357	1.6	3,970	10,285	51.8	34.7
26	Construction	3.2	9.3	1,392	1.2	61,010	176,936	58.0	7.9
27	Wholesale & retail	3.5	3.8	850	1.4	68,180	199,606	58.6	4.3
28	Hotels & catering	6.4	3.3	236	1.1	21,200	63,189	59.6	3.7
29	Transport services	1.6	21.3	539	1.3	19,800	56,716	57.3	9.5
30	Tele. & computer services	6.2	28.9	210	2.4	4,250	11,357	53.4	18.5
31	Financial services	18.0	18.4	618	2.2	12,740	32,819	51.5	18.8
32	Real estate services	38.8	3.4	263	1.9	10,540	28,559	54.2	9.2
33	Business services	13.9	3.7	750	1.6	15,800	41,702	52.8	18.0
34	Government	8.0	9.7	1,435	2.5	40,270	100,584	50.0	14.3
35	Education	4.7	12.2	945	2.2	36,190	92,358	51.0	10.2
36	Healthcare & social security	6.5	9.9	450	2.2	18,280	48,879	53.5	9.2
37	Other services	6.7	5.4	302	1.4	61,260	172,878	56.4	1.7
	Total economy	***2.6***	***11.4***	***17,704***	***1.7***	***784,560***	***2,141,154***	***54.6***	***8.3***

Sources: Authors' estimates.

Note: See Table 8b.

REFERENCES

Bai, Hejin (2002), *A Guideline to the Major Economic Events of the People Republic of China* [*Zhonghua renmin gongheguo jingji dashi jiyao*], China Planning Press, Beijing

Barro, R. J., & Lee, J. W. (1997). International measures of schooling years and schooling quality. *American Economic Review*, Papers and Proceedings, 86(2), 218–223.

Barro, R. J., & Lee, J. W. (2000, April and August). International data on educational attainment: updates and implications. CID Working Paper No. 42. Center for International Development at Harvard University, Cambridge, MA.

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul M. Holland [1975], *Discrete Multivariate Analysis*, Cambridge, MA: MIT Press.

Borensztein, Eduardo and Jonathan D. Ostry, 1996 "Accounting for China's Growth Performance", *American Economic Review* 86 (2), 224-28.

Bosworth, Barry and Susan M. Collins (2008), "Accounting for Growth: Comparing China and India", *Journal of Economic Perspectives*, Volume 22, Number 1: 45–66

CASS [Chinese Academy of Social Sciences], 1988 and 1995, *Household Survey*, 1988 and 1995. Beijing: Institute of Economic Research, Chinese Academy of Social Sciences

Chen, Kuan, Gary H. Jefferson, Thomas G. Rawski, Hongchang Wang, and Yuxin Zheng, 1988, "Productivity Change in Chinese Industry: 1953-1985", *Journal of Comparative Economics* 12 [4], 570-91

Chinloy, Peter T., 1980, "Sources of Quality Change in Labor Input", *America Economic Review* 70 [1], 108-19

CHIP (China Household Income Project), 1988, 1995, 2002 and 2007, *China Household Income Survey*. Beijing: Institute of Economic Research, Chinese Academy of Social Sciences

Chow, Gregory C. (1993). "Capital Formation and Economic Growth in China", *Quarterly Journal of Economics* 108 (3), 809-42.

Christensen Laurits R., Dale W. Jorgenson and Lawrence Lau, 1973, Transcendental Logarithmic Production Frontiers. *Review of Economics and Statistics* 55, no. 1 (February): 28-45.

CSIC (China Standard of Industrial Classification), *China Standard of Industrial Classification*, 1975, 1985, 1994 and 2002. Beijing: China Statistical Press.

Denison, Edward F., 1961, "Measurement of Labor Input: Some Questions of Definition and the Adequacy of Data". In *Output, Input, and Productivity Measurement*, 347-72. NBER Studies in Income and Wealth. Princeton, N.J.: Princeton University Press

Denison, Edward F., 1962, *Sources of Economic Growth in the U.S. and the Alternatives before Us*, New York, Committee for Economic Development

Denison, Edward F., 1974, Accounting for United States Economic Growth, 1929-1969, Washington DC, Brookings

DITS [Department of Industrial and Transportation Statistics, National Bureau of Statistics, China], various years, *Gongye Jiaotong Tongji Nianbao* [Annual Statistical Report on Industry and Transportation] [Internal Publication]. Beijing: National Bureau of Statistics

DITS, 1978, 1979, 1980, 1981, 1982 and 1983, *Annual Bulletin on Commune Factories*. Beijing: NBS Archives

DITS, 1984 and 1985, *Annual Bulletin on Township Enterprise*. Beijing: NBS Archives

DITS, 1999, *Xinbian Gongye Tongji Gongzuo Zhinan* [New Handbook of Industrial Statistical Work], Beijing: China Statistical Publishing House

DITS, 2000, *Zhongguo Gongye Jiaotong Nengyuan Wushinian Tongji Ziliao Huibian: 1949-1999* [Comprehensive Statistics of Fifty Years in Chinese Industry, Transportation and Energy: 1949-1999], Beijing: China Statistical Publishing House

DPES, 1955, 1957 and 1963 *Labour and Wage Annual Bulletin* (Internal), National Bureau of Statistics of China

DPSSTS, 1998, [Department of Population, Social and Science and Technology Statistics], *Xinbian Laodong Tongji Shiyong Shouce* [A New Handbook of Practical Labour Statistics], Beijing: China Statistical Publishing House

DPSSTS, 2001 and 2006, *China Labor Statistical Yearbook*. Beijing: China Statistical Publishing House

DSS [Department of Social Statistics, National Bureau of Statistics, China], 1987, *Laodong Gongzi Tongji Ziliao 1949-1985* [Labour Employment and Wage Statistics 1949-85]

DSS, various years, *Laodong Gongzi Tongji Nianbao* [Annual Statistical Report on Labour and Wage] [Internal]. Beijing: National Bureau of Statistics

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

General Union of Labor, Hebei Province, 2006, "Survey Report on the Basic Conditions of Migrant Workers in Hebei", mimeo in Chinese, Government of Hebei.

Gollop, Frank M. and Dale W. Jorgenson, 1980, "U.S. Productivity Growth by Industry, 1947-1973", in John W. Kendrick and Bertrice Vaccara [eds.], *New Developments in Productivity Measurement*, Chicago, University of Chicago Press, 17-136

Gollop, Frank M. and Dale W. Jorgenson, 1983, "Sectoral Measures of Labor Cost for the United States, 1948-1978", in Jack E. Triplett [ed.], *The Measurement of Labor Cost*, Chicago, University of Chicago Press, 185-235, 503-20

Hicks, John. 1981. *Wealth and Welfare: Collected Essays in Economic Theory*. Cambridge, Mass.: Harvard University Press.

Holz, A.C. and Yi-min Lin, 2001, "Pitfalls of China's Industrial Statistics: Inconsistencies and Specification Problems", *The China Review*, 1 [1], 29-71

Hu, Zuliu F. and Mohsin S. Khan, 1997, "Why is China growing so fast?", *IMF Staff Papers*, 44 (1), 103-131.

Jian, Xinhua and Huang, Kun, 2007, "On the Recent Living Conditions of Chinese Migrant Workers" (in Chinese), *Population Studies*, 31 (6), 37-44

Jorgenson, Dale W. 1990, Productivity and Economic Growth, in Ernst R. Berndt and Jack E. Triplett (eds.), *Fifty Years of Economic Measurement*, NBER, Studies in Income and Wealth Volume 54. Chicago and London: University of Chicago Press

Jorgenson, Dale W. and Zvi Griliches, 1967, "The Explanation of Productivity Change," *Review of Economic Statistics*, Vol. 34 [3], 249-283

Jorgenson, Dale W., Frank Gollop and Barbara Fraumeni, 1987, *Productivity and U.S. Economic Growth*, Harvard University Press, Cambridge, MA

Kendrick, John W., 1961, *Productivity Trends in the United States*. Princeton, N.J.: Princeton University Press

Kendrick, John W., 1973, Post-war Productivity Trends in the United States, *1948-1969*, New York: NBER

Kong, S. T., 2010, "Rural-Urban Migration in China: Survey Design and Implementation," in Meng, Xin and Manning, Chris [eds.] with Shi, Li and Effendi, Tadjuddin, *The Great Migration: Rural-Urban Migration in China and Indonesia*, Edward Elgar Publ. Ltd.

Li, Jingwen, Dale W. Jorgenson, Zheng Youjing and Masahiro Kuroda, 1993, *Productivity and Economic Growth in China, USA and Japan*, China Social Science Publishing House, Beijing

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

Maddison, Angus. 1998. *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 (2), April-June

MOLSS [Ministry of Labour and Social Security, China] 1999 and 2000, *Zhongguo Laodongli Shichang Gongzi Zhidao Jiawei 2000* [Guidance of Wage Rates for Labour Market in China]. Beijing: Zhongguo Laodong Shehui Baozhang Chubanshe [China Labour and Social Security Press].

NBS (National Bureau of Statistics, China) (various years), *China Statistical Yearbook*, China Statistical Press, Beijing

NBS, 1959, *The Great Ten Years*, People’s Press, Beijing.

NBS, 1990, *A Revision the 1955 Regulations on the Structure Total Wage Bill*, Decree No. 1, 1990. Beijing: National Bureau of Statistics

NBS, 1992, *On the Structure of Total Labour Compensation Including Both Wage and Non-wage Incomes*, Circular No. 257. Beijing: National Bureau of Statistics

NBS and MOL [National Bureau of Statistics and Ministry of Labour], 1994, *Laodong tongji Zhibiao Jieshi ji Wenjian Xiuanbian* [Explanation of Labour Statistical Indicators and Selected Documents]. Beijing: China Statistical Publishing House.

Perkins, Dwight H. and Thomas G. Rawski, (2008), “Forecasting China's Economic Growth to 2025”, Chapter 20, in Loren Brandt, Thomas G. Rawski (eds.) *China's Great Economic Transformation*, Cambridge University Press

SC/DRC (State Council/Development Research Center), 1962, Adjustment for Industrial Employment to Support Agricultural Production and Reduce Demand for Foodstuff, in SC/DRC, 2000, V.4 (II).

STA (State Taxation Administration), 2006, On Further Adjustment of Local Floating Band for Wage Rate. Beijing: State Taxation Administration

State Council, 1985a, 1985b, 1985c, *Circular on Regulation of Wage Bills*. Beijing: Sate Council

Wang, Haibo and Dong Zhikai, 1995, *Xin zhongguo gongye jingjishi* [History of Industrial Economy of the New China]. Beijing: Jingji Guanli Chubanshe

Wang, Yan and Yudong Yao, 2002, “Sources of China’s Economic Growth 1952-1999: Incorporating Human Capital Accumulation,” *China Economic Review*

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

Wu, Harry X. 2010. “Accounting for China’s Growth in 1952-2008 - China’s Growth Performance Debate Revisited”, Prepared for the 2010 AEA Meetings, Growth Session (O3), Atlanta, GA

Wu, Harry X. 2013. “Accounting for Productivity Growth in Chinese Industry – Towards the KLEMS Approach”, presented at the 2nd Asia KLEMS Conference, Bank of Korea, Seoul, August 22-23, 2013

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 (3) 2014

Wu, Harry X. and Keiko Ito, 2014. “Reconstructing China’s Supply-Use and Input-Output Tables in Time Series”, *RIETI Discussion Paper*, 2014 forthcoming

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

Wu, Harry X. and Ximing Yue, 2010. “Accounting for Labor Input in Chinese Industry”, presented at the 31st IARIW General Conference, St. Gallen, Switzerland, August 22-28, 2010

Wu, Harry X. and Ximing Yue, 2003. “Sources of Quality Change in the Labour Input of Chinese Industry, 1955-2000”, presented at the Western Economic Association International Pacific Rim Conference, Taipei, January 9-12, 2003

Wu, Harry X., Ximing Yue and George G. Zhang, 2013. “Accounting for Labor Input in the Chinese Economy: Effects of Changes in Hours Worked and the Quality of the Workforce”, IER, Hitotsubashi University, mimeo

Young, Alwyn (2003), “Gold into Base Metals: Productivity Growth in the People’s Republic of China during the Reform Period,” *Journal of Political Economy*, 111, 1221–61

Yue, Ximing (2005), “Problems of Current Employment Statistics in China”, *Economic Research Journal* (Jingji Yanjiu)

Zhu, Cishou, 1999, *Zhongguo gongye laodong shi* [Industrial Labour History of China]. Shanghai University of Finance and Economics Press