REAL WAGES AND SKILL PREMIA IN CHINA, 1858-1936*

Se Yan
Guanghua School of Management
Peking University
seyan@gsm.pku.edu.cn

What happens to wage inequality when a country begins to trade and industrialize? I construct new wage series for China from 1858 to 1936. I collect the nominal wages from the records of the China Maritime Customs, and estimate real wages for unskilled and skilled workers using new group-specific cost of living indices. I find that unskilled real wages were stagnant, but skilled wages rose rapidly before 1920 and fell thereafter. My findings suggest that technological advances increased skill demand, driving up the skill premium. Educational progress eventually increased the supply of skilled workers, thereby reducing the skilled wage.

Key Words: Wage, Inequality, Skill Premium, Education, International Trade, China

JEL Classification Codes: I20, O15, E24, J31, N35, N75,

* The author would like to thank Kenneth Sokoloff, Naomi Lamoreaux, Dora Costa, Jean-Laurent Rosenthal, Matthias Doepke, Leah Platt Boustan, Loren Brandt, Hongbin Cai, Ramon Myers, Hans van de Ven, and the audiences at many seminar and conference presentations for helpful comments and suggestions. Financial support from UCLA Center for Economic History, the National Science Foundation Grant SES-0922531 and China National Social Science Foundation Grant 09CJL009 is gratefully acknowledged. The errors are mine.

Economists have devoted much effort to documenting and understanding the patterns of inequality in leading industrial countries. A large body of academic work has shown both theoretically and empirically that technological changes in the first and second industrial revolutions and more recent periods had different skill biases and affected inequality patterns differently.[1] Scholars have also observed that the expansion of mass education contributed to a decline in income inequality in the US in the twentieth century.[2] It is also well studied how the expansion of trade between developing and developed countries affects wage inequality in the US and Europe.[3] External shocks, such as wars, regime changes, minimum wage regulations, etc. are also found to greatly affect wage and income inequalities in the US and Europe.[4]

An equally interesting question is how wage inequality changes in developing countries and what factors drive these changes. For example, it is heavily debated how openness and globalization has altered wages in developing countries. Classical trade theory, such as Heckscher-Ohlin-Vanek model and Stolper-Samuelson theorem, predicts that inequality should be falling when a country opens up and become more specialized in producing and exporting labor-intensive commodities. On the other hand, these countries usually begin to industrialize by importing many new skill-biased technologies from developed countries when they open up. This leads to a relatively higher demand for skilled workers complementary to the imported new technologies, and raises wage inequality. While some studies find empirical evidence supporting the prediction of classical trade theories,[5] many other empirical studies suggest that globalization has likely increased inequality in developing countries in the last three decades.[6]

---

[1] See, for example, Katz and Murphy (1992), Goldin and Katz (1996, 1998), Acemoglu (1998, 2002a, 2002b, 2003), Caselli (1999), Krusell *et al*. (2000), O'Rourke, Rahman and Taylor (2007).
[2] See, for example, Goldin and Margo (1992), Katz and Murphy (1992), Card and Lemieux (2001), Autor, Katz and Kearney (2008), Goldin and Katz (2009).
[3] See, for example, Leamer (1988), Revenga (1992), Lawrence *et al.* (1993).
[4] See, for example, Piketty and Saez (2003), Piketty, Postel-Vinay and Rosenthal (2004), Atkinson and Piketty (2006).
[5] For example, using city-level data, Wei and Wu (2001) find evidence that inequality has fallen within China, and that the decline in rural-urban inequality have been most pronounced in areas that increased their openness.
[6] For articles on the effects of trade on skill premia in specific countries see, for example, Robbins (1996), Beyer, Rojas, and Vergara (1999), Gasparini (2003), Hanson (2004), and Robertson (2000, 2004).

Determining the driving factors of wage inequality in today's world is complicated since all these possible factors are intertwined together and difficult to be disentangled. For example, in thinking about the causes of rising wage premium in the US since the 1970s, empirical researchers may need to account for growing trade with developing countries, declining union power, falling minimum wages, increased rates of immigration of unskilled workers, and greater skill-biased technological change. Similarly, to study factors changing wage inequality in developing countries, like China, scholars also face an array of challenging empirical issues that make identification difficult, including assessing the impacts of technological change, foreign direct investment, and state intervention on factor prices.

In this paper, I use the lens of history to better understand how rapid openness and industrialization affects real wages and skill premium for a developing country. In particular, I ask how the real wages and the skill premium in China responded to an earlier era of openness and industrialization. Prior to the 1840s China was a closed, largely agrarian economy. Under international pressure, China gradually opened its economy to foreign trade and, later, to foreign direct investment. By the early twentieth century, the coastal regions were industrializing and commercializing, and Chinese-made manufactures had begun to replace imports and even to be exported. Historians and economists have made substantial progress in sketching the broad outlines of this transformation.[7] What has been less understood, however, is how these accelerating economic changes affected people's real income and standard of living.

I assemble new data on wage records and prices and present the first detailed real wage series both in general and by skill group from 1858 to 1936. I construct nominal wage series from the records of employees in the China Maritime Customs service (hereafter "CMC") for nearly fifty Chinese cities. I also construct group-specific cost of living indices from price data and household budget information contained in CMC trade records. The resulting nominal wage series and cost of living indices make it possible to estimate long-run trends in real wages and skill premia for three categories of Chinese workers: unskilled, skilled, and highly skilled.

[7] See, for example, Chang (1962), Liu and Yeh (1965), Chang (1969), Rawski (1989), and Ma (2009).

These newly constructed real wage series reveal that skill premia rose rapidly during the first three decades of industrialization. After the 1920s, the wage gap between skilled and unskilled labor began to decline, while the gap between highly skilled and unskilled labor continued to rise. These changes in the skill premia were driven by movements in the wages of skilled and highly skilled labor. China's enormous reservoir of unskilled labor kept unskilled real wages stagnant until a rapid growth spurt in the 1930s. By contrast, skilled and highly skilled labor became increasingly scarce, driving up their real wages and the skill premia up to 1920. After 1920, however, progress in primary and secondary education increased the supply of skilled labor, lowering the skilled wages. In contrast to the declining skilled wages, the wages for highly skilled workers continued to grow after 1920, because this group of workers required high education and progress in high education was much slower. These findings on China about 100 years ago cast light to the question why inequality in today's developing countries has aggravated, and indicate possible solutions to rising inequality in the developing countries.

The rest of this paper is organized as follows. Section II offers a brief overview of the economic changes that occurred in China during this period. Section III introduces the primary data source, the CMC archives, and studies the representativeness of the dataset. In Section IV I use hedonic regression techniques to generate the nominal wage series and subject my findings to robustness checks. Section V builds cost of living indices for the different income groups. In Section VI I calculate real wages and skill premia using the nominal wage series and cost of living indices, and then look at what factors drove the changes in the skill premia. Section VIII concludes.

II. Background

The late nineteenth and early twentieth centuries were a critical period of China. This period witnessed economic openness, rapid industrialization and commercialization, and drastic political and social transformations. Prior to the 1840s China was largely a closed, agrarian economy. Under pressure from foreign nations led by Great Britain, China gradually opened its economy to international trade and later on foreign direct investment. The Treaty of Nanking in 1842 permitted foreigners to trade with the

Chinese in five open ports and stipulated a general 5 percent ad valorem tariff on all goods leaving and entering China.[8] By the early twentieth century, the number of cities open to trade had climbed to 48 cities. The transition from a closed to an open economy, with a virtually free trade policy, sparked rapid, sustained growth in international trade (see Figure 1). From 1871-84 to 1920-29 exports as percentage of GDP increased from 2.5 percent to 7.3 percent, the per capita value of foreign trade grew from 0.58 to 3.01 US dollars, and China's share of world trade increased from 1.3 percent to 2.4 percent (Dernberger, 1975).

China's defeat in the Sino-Japanese war in 1895 ushered in a second wave of economic change. The Treaty of Shimonoseki forced China to allow Japanese businesses to invest directly in China and produce goods and services that could be sold to other nations as well as marketed within China. This privilege was quickly extended to other foreign powers by dint of the most-favored-nation clauses included in their trading agreements with China. Foreign capital poured in to finance enterprises in the railway, telecommunications, shipping, and manufacturing sectors (Hou, 1965; Dernberger, 1975). It was during this second wave of economic change that the industrialization of China really began.

Although the scarcity of data makes it virtually impossible to construct annual time series of GDP or other major economic indicators,[9] scholars have compiled various estimates of the speed and magnitude of industrial expansion and economic growth. For example, Chang (1969) estimates that industrial output grew 9.8 percent annually from 1912 to 1936. Liu and Yeh (1965) estimate an annual growth of per capita GDP of 0.33 percent in the same period, while Rawski (1989) contends that per capita GDP growth in advanced regions of China had already attained a similar rate to that of Japan in the Republican era.

[8] Except for a few special commodities such as opium and tea, China's tariff rates for both imports and exports were fixed at five percent ad valorem until 1929. However, because commodity prices rose rapidly after 1858, the effective tariff rates were often below 3 percent and were never above 4 percent. See Cheng (1956).

[9] The historical record for Chinese economiy mainly comes from the records kept by China's rulers, officials, and local elites. These records, such as tax records and land contracts, however, reveal more about how and why the Chinese people were organizing their lives to establish a special kind of civilization but very little quantitative information about what China's economic performance was all about.

While this view of China's accelerating economic change is shared by many historians and economists, it has been poorly documented how this transformation affected people's real income and standard of living. This situation calls for the collection of new information, which would make it possible to construct new wage series and cost of living indices, and contribute to a better measurement of the speed and magnitude of economic changes in China in the late nineteenth and early twentieth centuries.

III. Data

A. The China Maritime Customs and Its Archives and Publications

My wage and price data are collected from CMC archives and publications. These archives are among the richest resources for the study of trade, economics, politics, society, international relations, and many other aspects of Chinese history from the 1850s to 1949. The CMC was answerable to the Chinese central government, but its top administration consisted largely of British and later Japanese and American officials and a major proportion of its middle-rank technical and management staff were foreigners.

The CMC was probably the only bureaucratic organization in China that continued to operate uninterruptedly throughout the period from 1858 to 1949. The CMC's primary tasks, of course, were collecting revenues from, and recording and publishing data on, China's foreign trade. But it took on more and more functions over time, including collecting revenues for domestic trade, administering the postal system, developing inland and coastal waterways, and representing China at international fairs. The CMC's geographical reach grew from just fourteen stations during the 1860s to nearly fifty during the 1920s, covering not only the coastal regions but also inland cities including those near the border with Burma and along the Amur River on the northern tip.

The most studied of CMC's publications are its trade statistics.[10] These trade statistics record the quantities and the values of all commodities passing through each treaty port and therefore provide a precious source for the study of trade patterns and commodity prices in these cities. The CMC also published annual and decennial analytical reports on economic activities in those port-cities, as well as in the provinces where they were located.[11] These reports provide extremely rich information on many aspects of the Chinese economy, such as trade and shipping, the natural environment, agriculture, industry, currency and finance, mines and minerals, communications, education, administration, justice and police, health and sanitation, and population, and are of enormous values for scholars of these interests.

The CMC's archives have been much less studied than its trade publications. They consist of about 55,000 volumes stored in the Second Historic Archive of China in Nanjing. Around half these volumes pertain to labor. They include surveys of local wages and standards of living, CMC wage scales, and most importantly, the Service Lists – that is, the individual personnel records of CMC employees. The CMC started to keep the Service Lists in 1875 and continued until the end of the service in 1949, maintaining the series in almost the same format and with almost the same contents for this entire time. In each year the Service Lists recorded each employee's name, home town, year of joining the service, year of being promoted, year transfer to the current customhouse, rank, and monthly salary.

B. New Wage Data

The Service Lists are the primary sources for my wage data. However, the CMC did not start compiling the Service Lists until 1875 and stopped reporting wages in the 1930s, so I must supplement the Service Lists with other information from the CMC's personnel management archives. For the more recent period I use the Service Correspondences and the Memoranda of Services. These records provide information on temporary employees not covered by the Service Lists. Unfortunately, these records are

[10] See, for example, Cheng (1956), Hsiao (1974), Lyons (2003).
[11] The most well-known and cited reports are the so-called "*Decennial Reports*", which are a series of publications covering a wide range of natural and social conditions in all the treaty ports in China from 1882 to 1931. The annual publications, "*Returns of Trade and Trade Reports*", also contain reports on trade and economy in each treaty port.

not available for the earlier period because the archives in which they were stored were destroyed during the Boxer Rebellion in 1900. The only sources available for the period before 1875 are correspondence and dispatches between the CMC's head office and local branches. Although data on wages were scattered through these documents, I have been able to garner hundreds of data points from them and to extend my wage series back as far as 1858.[12] My data set stops at 1936 because the Sino-Japanese War broke out in 1937, and the economy was disrupted by war from then until the Communist victory in 1949.

The complete data set contains 44,600 observations.[13] Each observation records, for one particular year from 1858 to 1936, the wage received by a Chinese employee of the CMC, as well as biographical information about the employee, such as name, job title, home town, current port, the year beginning the service, year promoted to current position and year transferred to the current port. I excluded foreign workers from the data set because their wages were significantly higher than those in local Chinese markets.

Table 1 reports the geographic coverage of the data set.[14] The CMC treaty "ports" were not just coastal or river cities: 12,700 data points came from inland cities, constituting 28.48 percent of the whole sample. Some of these inland ports were along the upper and middle Yangtze River, some in northernmost Manchuria, some in the northwestern Shanxi and Shaanxi provinces, some in the southwest Yunnan province and some near the borders of Burma and Vietnam. There is even one Tibetan treaty port, Yatung. I categorized these cities into nine regions (see Appendix A1).

The CMC's personnel hierarchy was very complex. There were more than 100 specific positions in the CMC system. My wage database focuses on the 35 most common occupations, which I categorize into three groups: unskilled, skilled, and highly skilled labor according to the education requirements of

[12] There are several hundred volumes of dispatches and correspondences in early periods, many of which contain precious information on the CMC's internal management and its observations of local society. I have only been able to examine a small sample of these documents.

[13] The same employee often shows up in the dataset for a few times in different years. This feature of the data set allows tracking the career paths of the CMC employees in future researches.

[14] Due to the nature of the CMC system, its personnel distribution was highly skewed toward skilled labor, and in particular, office clerks. In addition, as noted earlier, the CMC's senior officials corresponded to government officials in local governments, and their wages were extraordinarily high, as shown in Table 1. Therefore the current study excludes the sample of senior officials.

these occupations.[15] Workers in the unskilled occupations were mostly illiterate. No formal schooling was required for these jobs. Biographical data in the personnel records suggest that most employees in skilled occupations had high school educations, and that most in highly skilled jobs had college degrees. Appendix 2 lists the occupations in all three worker groups, and Table 1 reports the group distribution of my wage database.

C. Data Coverage and Representativeness

The purpose of this paper is to estimate the long-run trend of wages for different income groups in China, utilizing the wage data developed from the CMC internal personnel archives. Since all these data points are the wages earned by the CMC employees, this methodology has several intrinsic problems.

First of all, the CMC employees all lived in cities. So to the extent that labor markets are not in equilibrium, their incomes do not tell us much about trends in wages in rural areas.[16] Furthermore, only treaty ports, not all cities, were included in the data set. Treaty ports were usually large and middle cities, and were more commercialized and integrated into national and even world markets. It is possible that wages in the treaty ports might be different from those in other rural areas. However, as industrialization drew unskilled labor from villages to cities, the labor market became more integrated. Therefore unskilled wages in treaty ports were able to indicate wages in smaller cities and even in rural areas.

Second, the wages paid by the CMC might have deviated from those in the local labor markets. Certainly, foreign employees received wages much higher than those paid to their Chinese colleagues. Their wages were not able to reflect any information of the wages in local markets. In addition, the CMC's senior officials were usually paid extremely high. Their wages were not able to reflect those paid

[15] Because the criterion is education requirement, the categorization of unskilled, skilled, and highly skilled workers in this paper is different from the usual categorization in the literature of European or American economic history. For example, carpenters here are considered unskilled workers because they did not need to get formal education to be carpenters. In this sense, this categorization is more similar to Lindert and Williamson (1983).

[16] Rural income and economic development have been extensively studied by quite a few scholars, such as Dwight Perkins, Ramon Myers, Philip Huang, Loren Brandt etc. However, studies on urban income and wages are very scarce. The CMC's annual and decennial reports also include studies on rural incomes and standard of living.

by private-sector firms. Instead, they were analogous to local government officials in terms of income.[17] Therefore observations of these employees are excluded from my sample. The question here is, whether the wages of other Chinese employees, such as coolies, engineers, mechanics, and clerks, were able to reflect market wages.

Because the CMC's wage system did not adjust instantaneously to reflect labor market fluctuations and rising living cost, it is possible that the CMC temporarily paid less than the market wage. Certainly, the CMC would not have been able to hire if in the long run, it had paid wages that, all things being equal, were below those comparable in the cities in which they operated. Existing archives in the early twentieth century show that CMC officials were fully aware of market wages and made concerted efforts to match them. The Inspector General's Circular No. 3002 in 1920 clearly shows such efforts.[18] Because the "pay of the Chinese outdoor staff … has been found insufficient in these days of increased cost of living to attract good men or to retain desirable employees," the CMC's head office instructed its branches to survey the market wages and living costs very carefully: "as local conditions vary so much at different ports, what would be adequate at one being inadequate or excessive at another ... I have to request you therefore to forward: a return showing the local average wages and average expenses of men performing work similar to that performed by the corresponding Customs employee".

Another important document, Circular No. 3839 in 1928, expresses more explicitly the CMC's efforts to match its wage to the market wage and accommodate the rising cost of living: "In order that I may be in a position to understand clearly whether the pay at each individual port of the lower grade ranks of employees is on a scale which takes into account the actual cost of living as well as the current market value of the services rendered by such employees, I have to instruct you to prepare and forward to me … a statement … showing the average market prices prevailing at your port during the previous six

[17] I did a rough comparison of salaries received by senior CMC officials and incomes of senior gentry-officials estimated by Chang (1962), and found them to be on the same level.

[18] Circulars were issued by the Service's head, the Inspector General. They were confidential documents to which only senior Customs officials such as Customs Commissioners had access. Like Imperial Edicts, they were laws until explicitly superseded by new Circulars. They were the key texts of the Customs Service. Circulars tell us how the Custom Service organized itself, discharged its routines, and responded to events. For details, see Bickers and van de Ven (2007).

months for the various necessaries of life, as well as the average wage paid by Chinese and /or foreign firms at your port to the specified classes of employees."

On the other hand, the question remains that the CMC, as a special bureaucratic organization, might have paid significantly higher than the market wage in order to attract best people. This question can be answered by carefully comparing the wages from the CMC and from the local private-sector firms. Fortunately, the CMC conducted six nationwide wage surveys in the three years from 1929 to 1931. These surveys compared wages of employees from the CMC's local customhouses, local Chinese firms, and local foreign firms. Table 2 reports the mean wages of observations from the 1929-31 surveys by dates, locations and occupations of these observations. This table revealed mixed information: the wages from three types of employers do not seem to differ too much from each other by dates and locations, but the wage discrepancies by occupations were conspicuous. Therefore a better study is needed to show whether the CMC had successfully paid their employees fair market-level wages. To do so, I run the following regressions for the survey data:

$$\log(wage_i) = \alpha + \sum_{m=1}^{3}\sum_{j=1}^{3} \beta_{m,j} \cdot source_{m,i} \cdot skill_{j,i} + \sum_{k=1}^{9} \gamma_k \cdot region_{k,i} + \varepsilon_i \ . \qquad (1)$$

The variable "source" indicates different types of wages. Three sources of wages were collected in the 1929-31 surveys: wages of Chinese firms, wages of local foreign firms, and the CMC wages. The variable "region" indicates the regions these wage observations came from. The 1929-31 surveys contain 25 occupations. These occupations are categorized into three groups: unskilled, skilled and highly skilled, following the categorization rule described in the previous subsection. This specification includes an interaction term of types and skills. It allows me to compare the differentials of the CMC wages and those of private-sector firms by three skill groups.

Table 3 reports the regression coefficients, which indicate the log wage differences compared with the base group: unskilled CMC wage. It is clear that, for the unskilled and highly skilled labor, wages from the CMC and foreign firms were statistically indistinguishable, while for skilled labor foreign firms paid about 6.6 percent higher. This was not a big gap, and it is possibly because the occupations of skilled

labor from these employers are highly heterogeneous. Therefore, the regression strongly suggests that the wages from the CMC and local foreign firms do not differ greatly from each other.

Table 3 also indicates that the CMC and local foreign firms paid significantly higher than local Chinese firms. The differences between the CMC wages and the wages of local Chinese firms were 22.2, 11.3 and 7.6 percent for three skill groups respectively. The wage differences between local Chinese and foreign firms were also significant. However, this result is somewhat illusory. Chinese firms generally provided food, accommodations, and other non-pecuniary benefits to their employees, while the CMC and foreign firms seldom did so. Such a difference in compensation method has been thoroughly discussed in the reports from local customhouses.[19] According to these survey reports, the board and accommodation benefits usually value 12 to 20 percent of the monetary income.[20] With such large in-kind benefits, the total incomes received by employees in Chinese firms were no lower than those worked in foreign firms and the CMC.

Given that the CMC was a large bureaucratic organization that was not able to adjust wage scales instantaneously, the CMC wages might have been temporarily different from the market level. However, the CMC tried hard to set its wage scales to match the market equilibrium level, and the CMC's survey data also suggest that the wages paid by the CMC, local Chinese firms and foreign firms did not statistically distinguish from each other. This study is in favor of the view that local labor markets were highly integrated. Different types of employers were competing for skilled and unskilled laborers, and equilibrium wages were attained from the labor market competition.

IV. Nominal Wages

[19] For example, a large number of surveys from local customhouses indicate that wages from Chinese firms did not include "free food and housing, value $7" (China Maritime Custom Archive No. 679(1)-16231 and 16232, *"Ports Returns of Living Expenses, etc, of Native Outdoor Staff (Excluding Chinese Tidewaiters) , Submitted in Reply to Circular No.3002"*); also see, for example, most Chinese firms provided "lodging and boarding", or "mess allowance valuing $3 to $4", or sometimes even "food, clothing, quarter, and festival bonus" (China Maritime Custom Archives No. 679(1) 16233, 16234, 16235, 16236, 16237, 16238, *"Cost of Living Returns for Half Year Ending 31th March, 1929 to September 1931"*).

[20] See the same sources of archives listed in the previous footnote.

A. Hedonic Wage Regressions

A major goal of this paper is to construct annual time series of nominal and real wages for each of the three groups of workers and calculate the trend in the wage premia between these groups.[21] In order to generate these estimates, I use hedonic regressions to adjust for changes in the attributes of the sample over time (for example, the number and location of the customhouses), as well as to control for characteristics of the individual employees.[22] The dependent variable is the log wage earned by a particular CMC employee, and the independent variables are characteristics of the employee under analysis: the year and the location of the observation, the worker's occupation, his length of service (the variable "tenure") and tenure squared.

As noted in the previous section, my data set covers 35 jobs, which I group into three categories: unskilled, skilled and highly skilled, based on their education requirements. I run the following hedonic regression separately for each of the skill groups, yielding three separate nominal wage series:

$$\begin{aligned} \log(wage_i) = \alpha + \sum_{k=1}^{K} \delta_k \cdot occupation_{k,i} + \sum_{j=1}^{9} \phi_j \cdot region_{j,i} \\ + \gamma_1 \cdot tenure_i + \gamma_2 \cdot tenure_i^2 + \sum_{t=1}^{21} \beta_t \cdot period_{t,i} + \varepsilon_i \end{aligned}, \qquad (2)$$

where $\log(wage_i)$ is the log of the monthly wage of employee *i*; α is the constant term; $occupation_{k,i}$ is a dummy for the employee's occupation;[23] $period_{t,i}$ is the date of the observation, recorded as one of

[21] This nominal wage data set can also be used to study trends in regional wage differentials. A preliminary study, in which I divide China into inland and coastal regions, shows that regional differentials were small in early period. The gap expanded in the 1860s, after the economy opened to trade, and grew particularly fast after the 1890s, when the ban on foreign direct investment was lifted. The regional gap narrowed steadily during the WWI and then leveled off in the 1920s and 1930s. These results show that economic openness and industrialization impacted the coastal regions first, raising wages there. Wages in the inland regions did not begin to catch up until the 1910s.

[22] See Margo (2000) for a similar study.

[23] The omitted groups of three regressions are: lighthouse keeper for the regression of unskilled wages; postal clerk for the regression of skilled wages; harbor master clerk for highly skilled wages.

the 21 sequential periods from 1858 to 1936; [24] $region_{j,i}$ indicates the location of the observation.[25] $tenure_i$ indicates the number of years this employee had worked for the CMC; and ε_i is the error term.

The coefficients of $period_t$, β_t, for the three skill groups are plotted in Figure 2.[26] The figure shows that the nominal wages of all three skill groups began to grow rapidly in the mid 1890s and that those of highly skilled workers rose fastest.

B. Robustness Tests of the Hedonic Regression

In order to test how reliable my estimation of nominal wage indices are, I conduct four types of robustness tests. These robust tests do not change the qualitative features of my results obtained in previous subsections.

My categorization of the 35 occupations into three groups was based on the educational requirements of these jobs. In order to see whether my results are sensitive to this classification scheme, I estimated nominal wages for categories based on wage levels (low, middle and high income workers) and work properties (unskilled physical workers, skilled physical workers and white collar workers). The trends in nominal wages and the skill premia for these alternative groups are very similar to the current results.

The second robustness test is to pool the three hedonic regressions into one single regression, and produce the wage series for the three skill groups simultaneously. The biggest problem of this method is to control for occupation variation within each skill group. My wage data set is an unbalanced panel, and

[24] In the regression I combine all the observations before 1875 into one period because the early data are very scarce. This is the first period, and is the omitted category of the period dummies. Beginning with 1876, I use three-year intervals. This method cannot produce an annual wage series, but it is reasonable because the CMC wages intrinsically reflect long-run equilibrium wages and annual variations are not able to be reflected precisely. An alternative estimation method is to estimate annual variations by assigning each year as a dummy. The advantage of this method is that it constructs annual wage series. But when the sample size of each year is not big enough, the coefficients become volatile. So an acceptable way to accommodate this problem is to smooth the annual wage series by moving average. These two methods do not show significant differences.

[25] Appendix A1 reports how nine regions are defined. The omitted group of the region dummies here is Manchuria.

[26] Due to length constraint, the full regression results are not reported. The regressions fit the data well. The R^2 of three regressions are 0.54, 0.50, and 0.52 respectively.

the composition of the occupations changes greatly over time. I solve this problem by running a weighted hedonic regression, where the weights are generated to represent the changing occupational composition. I have tried various ways of generating weights. The trends are very similar to those in Figure 2 but more volatile, indicating that these estimation methods do not control for occupational composition as effectively as the original method. But the qualitative results do not change.

A more direct way to assess whether my method of estimation has done a good job of controlling for changes in the composition of occupations is to estimate wage trends for several representative occupations and see if they are similar. Estimates for wages of lighthouse keepers, postal clerks and customs clerks do not differ significantly from the trends in Figure 2.

Another possible problem in my wage estimations is that the CMC's customhouses were expanding geographically over time, so my wage data cover more inland regions in later periods. If trends in wages in the coastal regions were significantly different from those in inland, my estimations could be biased. To test this possibility, I run the regressions only for observations from Shanghai. Even though I do not have enough data points to generate wage indices in Shanghai for each period, the patterns are not significantly different from the current results.

C. Estimations of Actual Nominal Wages and Nominal Wage Gaps

Figure 2 reports nominal wage indices for the three groups. It is useful to estimate a specific monetary wage of each group in each period, so that wage gaps between the skill groups can be assessed. The actual monetary wage series are also of great interest for many purposes -- for example, to sharpen international comparisons.

The method for estimating nominal wage levels from wage indices is to obtain a benchmark monetary wage for each group from other sources, and to calculate wages in other periods based on wage indices reported in Figure 2. However, it was difficult to obtain a widely accepted set of wages for three

skill groups in China for the Qing period.[27] The benchmarks I adopt are the CMC's wages developed from the CMC's surveys from 1929 to 1931, which indicate the nominal wages in period 20 in my wage estimation. I categorize these survey wage records into three skill groups using the same rule, and then calculate the medium wage of each skill group. [28]

Once I obtain these benchmark wages, the calculation of nominal wage series is straightforward. Let $wage_t$ be the nominal wage in period t, then

$$wage_t = wage_{20} \times I(t), \tag{3}$$

where $I(t) = \exp(\hat{\beta}_t - \hat{\beta}_{20})$ and $\hat{\beta}_t$ is the coefficient of the dummy for period t. The nominal wage series for unskilled, skilled and highly skilled labor are reported in Table 4. I use these wages to compute the skill premia between skilled and unskilled workers and between highly skilled and unskilled workers. Two skill premia are plotted in Figure 3 after being indexed. [29]

How do my estimations of nominal wage series compare with the scattered data on wages available from other sources? There are a few existing studies about wages in China at that time, but these studies cover different occupations and different regions, and are very unsystematic. The comparisons between my nominal wage series and existing wage information show that the differentials are within reasonable ranges.

D. Discussion of Nominal Wage Series and Nominal Wage Gaps

[27] There are wage surveys from various sources in the early twentieth century. However, they usually cover only low-income group, and the geographic and occupation coverage are limited. So I do not use these wage survey data as benchmarks.

[28] These surveys have a relatively complete coverage of geography, occupation and tenure. So using medium wage is a reasonable choice. Certainly this simple average is unable to accommodate the different weights of population distributions among regions and occupations.

[29] I have also used the "fixed-worker" method to test the robustness of the estimations of actual monetary wages. In this method, a fixed value series of occupations, regions and tenure is chosen for each skill group. In this robustness test, the tenure value is the average of each skill group; the occupation values are the equal weights of each skill group; the region values are obtained from the regional distribution of the CMC employees taken from the CMC's surveys between 1929 and 1931. The wages from this method and the benchmark method do not significantly differ from each other.

Figure 2 shows that the nominal wages for all three groups of workers stagnated until the end of the nineteenth century and then began to grow rapidly. Figure 3 reports that the nominal wage gap between skilled and unskilled workers was flat from 1880 to 1936, while the gap between highly skilled and unskilled workers declined slowly until a sharp increase after the mid 1920s.

Changes in the nominal wage series largely resulted from movements in the general price level. The widely cited Nankai Index (reported in Figure 4) shows that the general price level declined by nearly 30 percent in the 1870s, leveled off until the late 1890s, and then began to grow rapidly.[30] By the 1920s, the price level was 2.5 times that of 1867. Deflating the nominal wage series yields the preliminary real wage series reported in Figure 5. The figure shows that after a 20 to 30 percent increase from 1858 to 1880, the real wages of all three groups declined by more than 30 percent between 1880 and 1900. After that, real wages began to grow slowly until a rapid increase after 1930. Of course, adjusting for inflation in this way has no effect on the skill premia reported in Figure 3.

V. Cost of Living Indices

It is unlikely that price movements affected all groups of workers in the same way. The poor and the rich usually consume different bundles of goods. The nominal wages of the three skill groups in my data set differ greatly from each other. Table 4 shows that the skilled wages were about four times of the unskilled wages, and the highly skilled wages were about ten times of the unskilled wages. Unskilled workers devoted 55 percent of their expenditure on food, while highly skilled workers spent only 26 percent. If the rice price rose faster than prices of other commodities, the unskilled would have been hurt more than the skilled and the highly skilled. Figure 4 shows that this happened. The rice price index rose much faster than the general price index, which includes many manufactured goods. Therefore, real inequality should have increased more than nominal inequality. To capture the differences in real wage

[30] The Nankai Index is probably the only long-run price index covering the period from the 1860s to 1949 (Hsiao, 1974). It is constructed with data obtained from the CMC's trade statistics. But this index suffers from two problems: it is a simple average rather than a weighted average; and rather than a consumer price index it is a general index including a large number of industrial intermediate goods.

movements, I estimate different expenditure shares for the three skill groups in this section. Applying these expenditure shares to the price series generates three different cost of living indices for three skill groups respectively. This method gives a better measurement of real wages and wage inequality.[31] The three group-specific cost of living indices come from the price data in the CMC trade statistics and the CMC's family budget surveys. These are the first nation-level cost of living indices that span a long period from 1864 to 1936.[32]

The consumption behavior changes over time, and ideally we want to use different consumption quantities or weights for different periods. In this way the substitution effect could be taken into account. However, I do not have enough data to generate weights for the different skill groups for each period, and must use fixed-weight estimation.[33] Fixed weights can possibly lead to substitution bias as people's consumption behavior might have changed during this long period of nearly eighty years. But this is not a serious problem in this paper. The nominal wages the three skill groups in my study differ from each other so greatly (as shown in Table 4), and the substitution bias is not possible to be so large as to make up for differences in the consumption patterns caused by huge wage differentials.

A. Weights

My commodity weights are developed from the family budget surveys conducted by the CMC's customhouses across the nation.[34] As mentioned in Section III, the CMC collected information about cost of living expenses and family budgets in all the treaty ports in the 1920s and the 1930s for the purpose of setting appropriate salary standards. Nearly fifty port-cities have been surveyed, so these archives are able

[31] A similar work can be found in Hoffman *et al*. (2002).
[32] The Nankai Index, as noted earlier, is a general index that includes many manufactured goods, and is only a simple average of all the price series. There are some other indices that focus on consumer goods. Prices and expenditure weights were collected from the market directly. However, it only covers big cities such as Beijing, Shanghai, and Tianjin in the 1920s and 1930s (the index for Beijing begins from 1900).
[33] In the estimation of historical price indices, usually the scanty information of family budgets does not allow the application of spliced indices with shifting budget shares. In this case the fix-weighted method is often used. See Sokoloff and Villaflor (1992) for example.
[34] In the early twentieth century sociologists and local governments began to collect information about household budgets and expenses, and these surveys have been widely used to study the standard of living for Chinese people in that era. Most of these surveys focus on family budgets of low-income residents, and were conducted in big cities. Therefore I am not able to develop national average consumption weights of three skill groups from these surveys.

to disclose more detailed information about geographical variation in consumers' behavior. The CMC budget surveys covered a wide range of occupations, which I categorize into the same three groups shown in Table A2.

Two batches of surveys from the CMC archives, the one for 1920 and the other for 1929/1931, reported expenses on major consumer goods. I group these commodities into five categories, deriving the expenditure weights reported in Table 5. The weight of food in total expenditures declined slightly over these ten years, and that of sundry items increased, suggesting that the living standard may have improved during this decade. Table 5 also suggests that households in the low-income group devoted a larger share of their resources to food, while the high-income group spent relatively more on sundry items. Housing expenses were relatively small for all income levels, ranging from 8.71 to 13.89 percent of the household budget, and were not very income elastic.[35] Table 6 gives a detailed breakdown of the weights on the commodities in the three class-specific baskets.

B. Basket of Representative Consumption Goods and Their Prices

28 consumer goods are selected into the consumption basket. These commodities are highly representative of the total household consumption because they are of high weights in the household budgets. Their prices are calculated from the trade statistics recorded in the CMC's trade publications. The CMC's trade publications provide a regular, continuing source of trade data for commodities leaving and entering Chinese customhouses. The CMC's records are believed to be "the only reliable and systematic material" (Cheng, 1956) to study China's trade patterns and commodity prices in that era. No other data source has such a long time span and geographical coverage as the CMC's publications. The CMC collected and published quantities and values of all the commodities passing through each customhouse from 1858 to 1949. Because the data from different ports over such a long time are usually in different measurement and currency units, I convert them into same units, and then aggregate the port-

[35] The small shares of housing expenses in modern China were not unique. The housing expenses in the US in the nineteenth century were estimated between 13 to 18 percent, which was also not high. The housing expenses in Europe in early periods were also small. See David and Solar (1977) and Sokoloff and Villaflor (1992) for example.

level quantities and values into national totals. I divide the national aggregated values by the quantities, and obtain the invoice prices of each selected commodity by each trade type. Finally, following the adjustment schemes proposed by Cheng (1956), I convert the invoice prices into market wholesale prices. The details are reported in Appendix A3.

Incidentally, data for certain years of each commodity might be missing from trade publications. In this case, I use two other sources to supplement the data. One supplementary data source is *China's Foreign Trade Statistics* (Hsiao 1974). The other source is the appendices of the CMC's Decennial Reports. Data recorded in these two sources are also cleaned from the CMC's trade publications, and their reliability is widely recognized.[36] If there are still data missing, I have to interpolate. Since such cases are scarce, interpolation does little harm to the precision of the price indices.

These commodities are grouped into five categories: food, clothing, fuel and lighting, housing and sundry. The first three categories are easy to measure because all of them are tradable goods. Food is divided into cereals, meat, vegetables, fruits and condiments. In my data set, rice, millet and wheat flour represent cereals; meat (a combination of pork, beef and mutton), fish and egg represent all meat consumption; vegetables are represented by turnip, bean and "other vegetables"; fruits are represented by pears and persimmons; and condiments are represented by bean oil, sugar and salt. The clothing consumption is represented by native cloth (nankeen cloth), mechanized cotton cloth (shirting), and silk products. The fuel is represented by coal, and the lighting item is largely kerosene.

The other two consumption categories are composed largely of nontradable goods, and accurate prices are virtually impossible to obtain. There is no systematic evidence on real estate markets in China at that time, so I use two alternative methods to estimate housing prices. The first method is to put together sporadic records of housing expenses from the CMC's Decennial Reports, which was one of the very few sources providing such information. This calculation is very crude. The brief descriptions that the CMC provided of local housing prices for scattered cities at different points of time were difficult to

[36] But because they only cover a much smaller collection of commodities in a limited amount of time, I cannot directly construct price series from these two sources, and have to go back to the original trade records.

aggregate into a national average. The second method is to construct a housing price series from the geometric average of labor costs and construction material costs.[37] I use my unskilled nominal wage series to proxy labor costs and the prices of timber planks and cement to proxy construction materials.[38]

Sundry consumption items, such as education, communication, sanitation, health, entertainment, religion and taxes are seldom priced in the data sets. The pricing method I use is the following. First, since a large part of sundry consumption consists of services, I use the nominal unskilled and skilled wages developed from the previous section to proxy service prices; second, I compute the prices of a few commodities pertaining to sundry items: tobacco, liquor, green tea, blankets and carpets, chinaware, paper, and clocks. A simple average of these prices is used to proxy cost of sundry expenses.

C. Cost of Living Indices

By applying these weights to the price series, I construct the indices of different groups as well as the general cost of living index. Table 7 reports the cost of living indices of the three skill groups and Figure 6 plots them.

Two features are conspicuous in Figure 6. First, the general trend of the cost of living indices of all three skill groups is the same. The cost of living declined by about 20 percent from 1858 to 1890, possibly as the result of silver appreciation in terms of gold. Inflation became rampant in the last decade of the nineteenth century. The first two decades of the twentieth century witnessed a gradual inflation, followed by rapid growth of cost of living expenses after the WWI. This wave of inflation stopped in 1930, most likely because of the Great Depression.

The other feature evident in Figure 6 is that the cost of living of the unskilled group rose significantly faster than that of the skilled and the highly skilled groups. This shows that low-income people suffered more from the inflation than high-income people. This was likely driven by a faster rise in

[37] This method was used by David and Solar (1977). Appendix 4 explains how the rent estimation is generated with this method.

[38] As studied in the next subsection, the shares of housing expenses were only 8.71 to 13.59 percent of total expenses. Therefore the cost of living indices were not much affected even if the estimations on housing expenses were not precise.

food prices than the prices of manufacture goods. As low-income people spent a much larger share of their expenses on food, higher food prices hurt the poor much more. The significant differences in these three cost of living indices suggest the necessity of constructing group-specific indices. The real wages would be biased and wage inequality would be underestimated if simply using the same index for all three groups

VI. Real Wages, Skill Premia and Explanations

Deflating the nominal wage series of the three skill groups by the respective cost of living indices yields the real wage series in Table 7 and Figure 7.[39] Real wages for unskilled labor rose by about 20 percent in the deflation period between 1858 and 1890, fell back in the next decade, and then leveled off until experiencing significant growth in the 1930s. Skilled real wages rose much faster from 1858 to 1890, declined after 1890, and then resumed growth during the first two decades of the twentieth century. Although skilled real wages began to fall again around 1920, the drop was followed by another rise in the 1930s. Highly skilled real wages rose from 1858 to 1890 and fell back from 1890 to 1900, just like wages for the other two groups. But the highly skilled real wage increased continuously from around 1900 to 1930, followed by faster growth during the 1930s.

I divide skilled and highly skilled wages by unskilled wage and obtain two skill premia. The indexed skill premia are plotted in Figure 8. The two skill premia rose by 15 percent from 1858 to 1880, leveled off after that, and began to grow rapidly again from 1900 to 1920. After 1920, the wage gap between the highly skilled and the unskilled continued to rise, but the gap between the skilled and the unskilled began to decline.

The real wage series for the three groups of Chinese workers suggest that changes in the skill premia were driven by the wages of skilled and highly skilled workers. Demand for all three types of workers increased rapidly as a result of two waves of economic change: trade liberalization beginning in

[39] The cost of living indices are annual. I group them into three-year averages to match the three-year nominal wage series.

1842 and rapid industrialization beginning in 1895. The large stock of unskilled labor was more than sufficient to meet any increase in demand, but, given China's traditional Confucian system of education, it was more difficult to find workers equipped with the knowledge and skills demanded by the modern sectors of the economy, including engineering and mechanics, reading and writing in native and foreign languages, bookkeeping and accounting, and business management.

Indeed, China has a long tradition of placing an inordinately high value on education. Literacy was widespread compared to most other pre-industrial economies. Rawski (1979) has estimated that from 30 to 45 percent of men and from 2 to 10 percent of women could read and write.[40] But the curriculum, which was designed to prepare students to take the examinations required for entrance into the bureaucracy, emphasized the Confucian classics, history, philosophy and literature, and excluded practical fields of knowledge such as medicine, agriculture, and engineering, which were regarded more as crafts. Therefore, the supply of skilled and highly skilled workers was limited, and the skill premia accordingly rose.

Beginning around 1920, however, the two skill premia began to diverge. While the wage gap between highly skilled and unskilled workers continued to rise, that between skilled and unskilled workers declined. It is unlikely that the fall in the premium for skilled labor resulted from demand-side factors. Industrial output continued to grow at a rapid pace, estimated at about 9.8 percent from 1912 to 1936 (Chang, 1962), so the demand for skilled as well as highly skilled workers should have continued to grow accordingly. Instead, it makes sense to look for factors that differentially affected the supply of skilled and unskilled workers. A good candidate is changes in the educational system that occurred after 1904, changes that fundamentally transformed the traditional Confucian system of education.

The rapid process of openness and industrialization in the late nineteenth century generated a large demand for talents equipped with certain degree of modern education, particularly in science,

[40] According to Rawski (1979), the literate group "included the fully literate members of the elite and, on the opposite pole, those knowing only a few hundred characters." She adds that, with such literacy, late imperial China was an "advanced society" by the standards of preindustrial Europe, "with many modern characteristics that helped ease its transition to modernity in the 20th century."

technology, and management. The old educational test system was not able to adapt itself to the drastically changing social and economic conditions, failed to train students to meet the demand from the modernization and self-strengthening movement, and called for a revolutionary reform. Although the attempts to build modern schools started in some coastal cities as early as around 1860, the abolishment of the Imperial Civil Service Examination in 1904 by the Qing court marked a nationwide remodel of the education system. The education revolution was continued and deepened after China became a republic in 1912, and the new system was expanded and gradually dominated.

The new education system was composed of primary, secondary and tertiary schools. Secondary schools, including general high schools and vocational high schools, were the major suppliers of skilled workers. According to the *Imperial Prospectus on Modern High Schools*, which was issued by the Qing court in 1903, the mission of secondary education was "to prepare students to work in modern industries or governments after graduation". This mission was further emphasized in the Republican period. The subjects taught in the general high schools were redesigned to meet the new demand, which included, for example, reading and writing of Chinese and foreign languages, physics, chemistry, law, economics, and geography. Parallel to the general high schools, the vocational high schools were organized purely for the purpose of training students with practical skills essential for the new industries. The mission of these vocational schools, stipulated in the *Regulations of Education System* published in 1911, was "to teach practical knowledge and skills as well as general subjects to those who are already employed in industries and who intend to". In order to do so, the vocational schools were further classified to agricultural schools, manufacture industry schools, commerce schools, shipping schools and so on, and the subjects taught in these schools were tightly connected to the demand from the related industrial sectors.

Although in many laws and regulations the governments committed themselves to financially support modern schools, they often failed to do so because of the unstable political regimes and frequent conflicts and wars. Instead, in order to have sufficient labor force equipped with up-to-date skills needed in their businesses, the rising entrepreneur and businessman class was one of the major patrons in financing the education revolution and modernization. Their rapidly growing economic power also

enabled them to do so. According to the *First Education Yearbook of China*, among all 19,830 new schools in 1906, 74 percent were private, while in 1904 only 16 percent were private. The total amount of education donations from private parties mounted to the astronomical figure of 11,414,253 silver dollars. The leading two provinces in private donation were Jiangsu and Zhejiang, where modern industry and commerce were most prosperous. Modern industries began to grow even faster in the early twentieth century. The fast growing industrial sector called for a larger modern education system. While the governments failed to provide sufficient funds to do so, these entrepreneurs invested heavily in modern education.[41]

As a result of the policy change and collective efforts, modern high schools began to grow rapidly in the early twentieth century. Table 8 shows that by 1936 the number of enrolled high school students had grown to 627,246, a level 6 times that of 1912. I estimate the stock of high school graduates by dividing the total of enrolled high-school students each year by six (the number of years of study) and cumulating the results over time. For this calculation I assume that there was no high school graduates before 1912, the first year for which data is available. This procedure yields an estimate of the stock of high school graduates of around 200,000 in 1919 and as much as 1 million in 1936. This still underestimates the stock of high school graduates, because they do not include the graduates from vocational education. Many of the graduates from modern schools specialized in the fields directly related to modern industry and business: 37 percent majored in law, 12 percent in engineering, 6.1 percent in science, and 5.7 percent in business.

This number was small relative to the total population. It was also only about 8 percent of total employment in the modern sector, which is estimated at about 2,885,000 by 1919 (Liu and Tang, 1998).

[41] A good example of this is Zhang Jian, who was one of the most famous entrepreneurs in modern China. He founded his first factory, Dasheng Cotton Textile, in 1899, and in 1903, not long after his first factory, he founded a vocational school with specialties in pedagogy, sericulture, agriculture and engineering. He spent one tenth of his industrial profits in education. In 1923, Zhang Jian and his family set up 25 schools in his home town, and the total expenses amounted to 3.5 million silver dollar, constituting one seventh of his total business assets..

However, the modern sector employed many more unskilled than skilled workers,[42] and the demand for skilled workers should have been only a small fraction of total employment. Furthermore, the estimated stock of high school graduates has not included those who attended vocational schools and who attended but did not high school. Taking all these factors into account, it is highly plausible that the rapid developments in high school substantially increased the supply of skilled workers and drove down the skill premium.

In sharp contrast to the rapid development of secondary education, the progress in tertiary education was much slower. Table 8 shows that the number of students enrolled in college was 40,114 in 1912, and it was still only 41,922 as late as 1936. Such educational statistics as are available suggest that the number of colleges and universities actually decreased from 115 in 1912 to 89 in 1918, and the annual government expenditure on tertiary education decreased from 4 million silver dollars in 1912 to 3.78 million in 1918. The number of college graduates in 1918 was only 900, but that number in 1913 was already 976 (Xiong, 1990). Compared to the total employment of nearly 3 million in the modern sector in 1919, the supply of college students was tiny. This stagnant supply, together with the rapidly growing demand, led to a persistently rising wage premium between highly skilled and unskilled workers.

VII. Conclusion

There are virtually no estimates of GDP or other measures of the size of China's economy for the late nineteenth and early twentieth centuries. Nonetheless, there is abundant evidence that the coastal regions, if not the entire country, began to industrialize rapidly in the late nineteenth century. This paper presents the first long-run real wage series for different skill groups of workers from 1858 to 1936. Industrial development had no positive effect on the wages of unskilled workers: they stagnated or even

[42] A survey about a cotton textile mill in Shanghai shows that only 20% of male workers and 3% of female workers could write their own names (*First Labor Yearbook of China*). The education level of workers in Shanghai was among the highest over the nation. So it was evident how low workers' education level was in the entire country. Literacy in coalmining industries was even lower. A survey about a coal mine in Shandong shows that 78.4% of the total employees never received any education; Employees in chemical industries had a higher education level: the literacy ratio was around 60%. There was little information about the proportion of skilled workers in the total employment, but it must be much lower than the literacy ratio. A detailed study can be found in Liu and Tang (1998).

declined until the late 1920s. Rather, the economic changes of the period disproportionately benefited highly skilled workers, whose wages grew rapidly compared to those of other groups. Skilled workers occupied an intermediate position. Their wages rose relative to unskilled (though not highly skilled) workers during the early industrial period, but during the 1920s they lost some of the ground they had previously gained.

These changes in the skill premia were driven by movements in the wages of skilled and highly skilled labor. China's enormous reservoir of unskilled labor kept unskilled wages flat throughout the period. By contrast, supplies of skilled and highly skilled labor initially were extremely scarce, driving up the skill premia in the first few decades of industrialization. After around 1920, however, progress in education increased the supply of skilled labor. Skilled wages started to decline, but wages for highly skilled labor stayed relatively high because the supply of this type of labor remained insufficient. This pattern is consistent with the story of a "race between technology and education".

References

Acemoglu, Daron, "Why Do New Technologies Complement Skills? Directed technical Change and Wage Inequality," *The Quarterly Journal of Economics*, 113 (1998), 1055-1089.

——, "Directed Technical Changes," *The Review of Economic Studies*, 69 (2002), 781-809.

——, "Technical Change, Inequality and the Labor Market," *Journal of Economic Literature*, 40 (2002), 7-72.

——, "Patterns of Skill Premia," *Review of Economic Studies*, 70 (2004), 199-230.

Atkinson, Anthony, and Thomas Piketty, *Top Incomes over the Twentieth Century*, (Oxford, UK: Oxford University Press, 2006).

Autor, David H., Lawrence F. Katz and Melissa S. Kearney, "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 90 (2008), 300-323.

Beyer, Harald, Patricio Rojas and Rodrigo Vergara, "Trade Liberalization and Wage Inequality," *Journal of Development Economics*, 59 (1999), 103-123.

Bickers, Robert, and Hans van de Ven, *China and the West: The Maritime Customs Service Archive from the Second Historical Archives of China, Nanjing* (London, UK: Thomson Gale, 2007).

Brady, Dorothy S. "Consumption and the Style of Life." In American Economic Growth, An Economist's History of the United States, by Lance E. Davis et al., 61-89. New York: Harper & Row, 1972.

Card, David, and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 116 (2001), 705-746.

Caselli, Francesco, "Technological Revolutions," *American Economic Review*, 89 (1999), 78-102.

Chang, Chung-li, *The Income of the Chinese Gentry* (Seattle, WA: University of Washington Press, 1962).
Chang, John, *Industrial Development in Pre-Communist China* (Chicago, IL: Aldine Publishing Co., 1969).

Cheng, Yu-Kwei, *Foreign Trade and Industrial Development of China* (Washington, D.C.: The University Press of Washington, 1956).

David, Paul A. and Peter Solar, "A Bicentenary Contribution to the History of the Cost of Living in America," in *Research in Economic History Vol. 2*, Paul Uselding, ed. (Greenwich, CT: JAI Press Inc., 1977).

Dernberger, Robert F., "The Role of the Foreigner in China's Economic Development, 1840-1949," in *China's Modern Economy in Historical Perspective*, Dwight H. Perkins ed. (Stanford, CA: Stanford University Press, 1975).

Gasparini, Leonardo, "Different Lives: Inequality in Latin America and the Caribbean," Unpublished Paper, Universidad Nacional de la Plata, 2003.

Goldin, Claudia, and Lawrence F. Katz, "Technology, Skill, and the Wage Structure: Insights from the Past," *American Economic Review*, 86 (1996), 252-257.

——, "The Origins of Technology-Skill Complementarity," *The Quarterly Journal of Economics*, 113 (1998), 693-732.

——, "Why the United States Led in Education: Lessons from Secondary School Expansion, 1910 to 1940", in *Human Capital and Institutions: A Long-Run View*, David Eltis, Frank D. Lewis, and Kenneth L. Sokoloff, eds. (Cambridge, UK: Cambridge University Press, 2009).

Goldin, Claudia, and Robert A. Margo. "The Great Compression: The Wage Structure in the United States at Mid-Century," *The Quarterly Journal of Economics*, 107 (1992), 1-34.

Hanson, Gordon,"What Has Happened to Wages in Mexico since NAFTA?" in *Integrating the Americas, FTAA and Beyond*, Antoni Estevadeordal, Dani Rodrik, Alan M. Taylor and Andrés Velasco, eds. (Cambridge, MA: Harvard University Press, 2004).

Hoffman, Philip T., David Jacks, Patricia A. Levin, and Peter H. Lindert. "Real Inequality in Europe since 1500." *Journal of Economic History*, 62 (2002): 381-413.

Hou, Chi-ming, *Foreign Investment and Economic Development in China, 1840-1937* (Cambridge, MA: Harvard University Press, 1965).
Hsiao, Liang-lin, *China's Foreign Trade Statistics, 1864-1949* (Cambridge, MA: Harvard University Press, 1974).

Katz, Lawrence F., and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107 (1992), 35-78.

Krusell, Per, Lee E. Ohanian, Jose-Victor Rios-Rull, and Giovanni L. Violante, "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68 (2000), 1029-1053.

Lawrence, Robert Z. and Matthew J Slaughter, Robert E. Hall, Steven J. Davis and Robert H. Topel, "International Trade and American Wages in the 1980s: Giant Sucking Sound or Small Hiccup?" *Brookings Papers on Economic Activity. Microeconomics*, (1993), 161-226.

Leamer, Edward E., "In Search of Stoloper-Samuelson Linkages," in *Imports, Exports, and the American Worker*, Susan M. Collins, ed. (Washington, D.C.: Brookings Institution Press, 1998).

Lindert, Peter H., and Jeffrey G. Williamson, "English Workers' Living Standards during the Industrial Revolution: A New Look," *The Economic History Review*, 36 (1983), 1-25.

Liu, Mingkui, and Yuliang Tang, "History of Worker Movements in China" (Guangzhou, China: People's Press of Guangdong, 1998).

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

Lyons, Thomas P., *China Maritime Customs and China's Trade Statistics, 1859-1948* (Trumansburg, NY: Willow Creek Press, 2003).

Ma, Debin, "Modern Economic Growth in the Lower Yangzi in 1911-1937: a Quantitative, Historical and Institutional Analysis," *The Journal of Economic History*, 68 (2008), 355-392.

Margo, Robert A., "Wages and Labor Markets in the United States, 1820-1860" (Chicago, IL: The University of Chicago Press, 2000).

Minister of Education, China, *First Education Yearbook of China*,1934.

O'Rourke, Kevin H., Ahmed S. Rahman, and Alan M. Taylor, "Trade, Knowledge, and the Industrial Revolution," NBER working paper No. 13057, 2007.

Piketty, Thomas, and Emmanuel Saez, "Income Inequality in the United States, 1913-1998," *The Quarterly Journal of Economics*, 118 (2003), 1-39.

Piketty, Thomas, Gilles Postel-Vinay, and Jean-Laurent Rosenthal, "Wealth Concentration in a Developing Economy: Paris and France, 1807-1994," *American Economic Review*, 96 (2006), 236–256.

Rawski, Evelyn S., *Education and Popular Literacy in Ch'ing China* (Ann Arbor, MI: University of Michigan Press, 1979).

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

Revenga, Anna L., "Exporting Jobs? The Impact of Import Competition on Employment and Wags in U.S. Manufacturing," *The Quarterly Journal of Economics*, 107 (1992), 255-284.

Robbins, Donald J., "Evidence on Trade and Wages in the Developing World," OECD Development Centre Working Paper No. 119, 1996.

Robertson, Raymond, "Trade Liberalization and Wage Inequality: Lessons from the Mexican Experience, " *World Economy*, 23 (2000), 827-849.

——, "Relative Prices and Wage inequality: Evidence from Mexico, " *Journal of International Economics*, 64 (2004), 387-409.

Sokoloff, Kenneth L. and Georgia C. Villaflor, "The Market for Manufacturing Workers during Early Industrialization, the American Northeast, 1820 to 1860," in *Strategic Factors in Nineteenth Century American Economic History, A Volume to Honor Robert W. Fogel*, Claudia Goldin and Hugh Rockoff, eds. (Chicago, IL: The University of Chicago Press, 1992).

The Imperial Maritime Customs Services of China, "*Returns of Trade and Trade Reports, 1859 to 1936*" (Shanghai, China: the Statistical Office of the Imperial Maritime Customs Services of China).

——, "*Decennial reports on the trade, navigation, industries, etc* " (Shanghai, China: the Statistical Office of the Imperial Maritime Customs Services of China).

Wei, Shang-Jin, and Yi Wu, "Globalization and Inequality: Evidence from within China." NBER Working Paper No. 8611, 2001.

Xiong, Ming'an, *Education History of Republic of China* (Chongqing, China: Chongqing Publishing House, 1990).

Table 1. Data Distribution

	N
Job Categories	
Unskilled	7,634
Skilled	18,727
Highly Skilled	17,257
Periods	
1858-1875	736
1876-1889	4,230
1890-1899	4,504

1900-1909	12,834
1910-1919	10,427
1920-1929	11,094
1930-1936	775
Regions	
Manchuria	1,770
North China	6,427
Lower Yangtze	13,044
Upper and Middle Yangtze	7,045
Southeast Coast	12,613
Guangxi and Hainan	2,081
Southwest	1,238
Taiwan	219
Northwest	163
Total	44,600

Table 2. Means of Monthly Wages Recorded in CMC's 1929-1931 Surveys

	CMC	Chinese firms	Foreign firms
Dates			
29-Mar	11.76	12.22	13.24
29-Sep	12.16	12.34	13.49
30-Mar	13.08	12.76	14.19
30-Sep	13.2	12.6	14.4
31-Mar	14.08	13.5	15.52
31-Sep	14.11	13.56	15.23
Regions			
Manchuria	15.91	16.36	19.98
North China	11.08	10.13	11.30
Lower Yangtze	12.70	11.83	12.47

Upper and Middle Yangtze	11.59	10.62	12.44
Southeast Coast	15.78	17.08	17.74
Guangxi Hainan Taiwan	12.12	11.96	16.01
Southwest	7.82	5.19	7.80
Occupations			
boatman	10.20	8.33	9.13
boatswain	26.78	19.26	
bricklayer	20.06	25.03	
cabin hand	11.93	15.19	12.18
carpenter	15.74	15.95	18.45
coolie	10.43	9.05	10.43
coxswain	18.75	33.39	22.77
deck hand	12.34	12.12	12.37
engineer	24.98	23.22	23.35
fireman	15.13	13.31	14.41
gardener	10.12	11.11	10.59
gatekeeper	10.27	8.19	9.85
hulkkeeper	11.16	11.24	17.09
junk boat owner	20.15	20.19	20.82
mason	14.46	15.26	17.92
mechanic	27.58	25.87	27.04
messenger	9.58	8.35	10.23
motorman	22.27		18.61
office boy	11.12	8.86	12.25
sailor	15.26		10.91
seaman	15.20	11.00	14.00
watchman	10.30	8.71	10.96

Source and Note: China Maritime Custom Archives No. 679(1) 16233-16238; Wages are in silver dollars

Table 3. The Logarithms of Wage Differences, Surveys 1929-31

	unskilled	skilled	highly skilled
CMC	0	0.274***	0.787***
		[0.016]	[0.021]
Chinese firms	-0.222***	0.161***	0.711***
	[0.014]	[0.024]	[0.031]
foreign firms	0.0161	0.3397**	0.7752
	[0.013]	[0.024]	[0.029]

Notes: *: Significant at 10%; **: 5%; ***: 1% Standard errors in brackets.
Occupation dummies are not reported due to length constraint.
Omitted groups are logarithms of monthly wages from CMC, in Manchuria and of the unskilled.

Table 4. Nominal Wage Series

Unit: silver dollar

year	unskilled	skilled	highly skilled
1858-1875	11.23	41.94	120.19
1876-1878	11.20	44.95	135.69
1879-1881	11.18	46.65	132.71
1882-1884	10.98	45.60	129.35
1885-1887	10.79	45.62	126.01
1888-1890	10.78	46.78	122.89
1891-1893	10.98	46.06	119.94
1894-1896	10.88	45.76	118.53
1897-1899	11.50	46.97	122.01
1900-1902	12.36	48.26	135.70
1903-1905	12.81	49.31	143.13
1906-1908	12.87	51.51	144.22
1909-1911	14.40	56.19	156.25
1912-1914	14.99	60.81	169.55
1915-1917	15.51	64.52	169.41
1918-1920	16.55	68.30	176.15
1921-1923	17.89	73.81	190.12
1924-1926	18.75	74.68	219.25
1927-1929	20.29	81.31	250.68
1930-1932	23.09	97.20	310.96
1933-1936	26.31	109.98	402.28

Note: Estimations of cross terms of years and job categories are plotted in Figure 2, but not reported in this table due to length constraint. Estimations of controlling variables “proportion” are not reported here either.

Table 5.

Weights from the CMC's Family Budget Surveys

sources	1920		1929-1931			
Skill groups	Unskilled	Skilled	Unskilled	Skilled	Highly skilled	Average
average wage (in silver dollar)	9.85	30.51	7.58	27.89	46.67	12.52
food (%)	56.67	42.67	55.89	37.49	26.24	50.48
clothing (%)	12.96	15.34	8.08	8.80	10.18	8.39
rent (%)	13.59	13.89	8.71	12.70	13.06	9.67
fuel and lighting (%)	10.45	9.86	9.44	7.48	6.15	8.85
sundry (%)	6.33	18.24	18.01	33.86	44.37	22.75

Table 6. Weights of Representative Commodities

	Low Income	Middle Income	High Income	Average
average wage (in silver dollar)	7.58	27.89	46.67	12.52
Rice	16.21	5.00	2.63	11.06
Millet	5.40	1.67	0.88	3.69
Wheat flour	8.88	7.20	3.04	8.10
Bean oil	6.42	5.11	2.85	5.56
Meat	9.66	11.58	7.13	11.21
Egg	0.21	1.47	2.14	1.35
Fish	0.41	1.47	4.09	2.45
Turnip	0.68	0.24	0.11	0.50
Other vegetables	0.68	0.24	0.11	0.50
Bean	5.35	1.78	0.86	3.91
Pear	0.11	0.05	0.89	0.35
Persimmon	0.11	0.05	0.89	0.35
Sugar	0.41	0.33	0.18	0.38
Salt	1.22	0.98	0.55	1.14
Native cloth	3.29	2.67	1.83	2.98
Shirting	3.29	2.67	1.83	2.98
Silk	1.51	3.46	6.53	2.43
Rent	8.71	12.70	13.06	9.67
Coal	3.48	3.76	4.68	4.22
Kerosene	5.97	3.72	1.36	4.43
Chinaware	0.86	1.61	2.11	1.08
Paperware	0.86	1.61	2.11	1.08
Clock	0.86	1.61	2.11	1.08
Liquor	0.86	1.61	2.11	1.08
Blanket and carpet	0.86	1.61	2.11	1.08
Green tea	0.86	1.61	2.11	1.08
Tobacco	0.86	1.61	2.11	1.08
Unskilled wage	6.00	11.29	14.79	7.58
Skilled wage	6.00	11.29	14.79	7.58
Total	100.00	100.00	100.00	100.00

Table 7. Group-specific CPIs and Real Wage Indices

year	CPI			Real wage indices		
	low	middle	high	unskilled	skilled	highly skilled
1858-1875	100	100	100	100	100	100
1876-1878	91.12	89.86	89.49	109.41	119.27	126.16
1879-1881	85.66	85.12	84.77	116.18	130.67	130.26
1882-1884	80.89	81.89	81.36	120.79	132.78	132.28
1885-1887	83.32	81.68	79.48	115.33	133.18	131.91
1888-1890	81.63	80.56	79.30	117.60	138.45	128.95
1891-1893	81.31	79.61	78.97	120.25	137.96	126.37
1894-1896	93.01	90.67	87.30	104.11	120.33	112.96
1897-1899	108.70	102.11	96.24	94.21	109.69	105.48
1900-1902	111.78	106.01	101.30	98.41	108.54	111.46
1903-1905	127.51	111.30	102.66	89.43	105.63	116.00
1906-1908	125.91	110.01	101.45	90.98	111.64	118.27
1909-1911	139.69	121.05	111.76	91.74	110.67	116.32
1912-1914	137.69	115.13	104.15	96.91	125.94	135.45
1915-1917	136.12	117.10	107.05	101.45	131.38	131.67
1918-1920	166.14	138.71	125.00	88.66	117.39	117.25
1921-1923	167.70	138.95	126.74	94.96	126.66	124.81
1924-1926	186.89	152.52	136.56	89.33	116.74	133.58
1927-1929	205.52	178.07	164.28	87.87	108.88	126.96
1930-1932	230.76	203.38	189.05	89.07	113.96	136.86
1933-1936	206.86	192.59	183.78	113.20	136.16	182.12

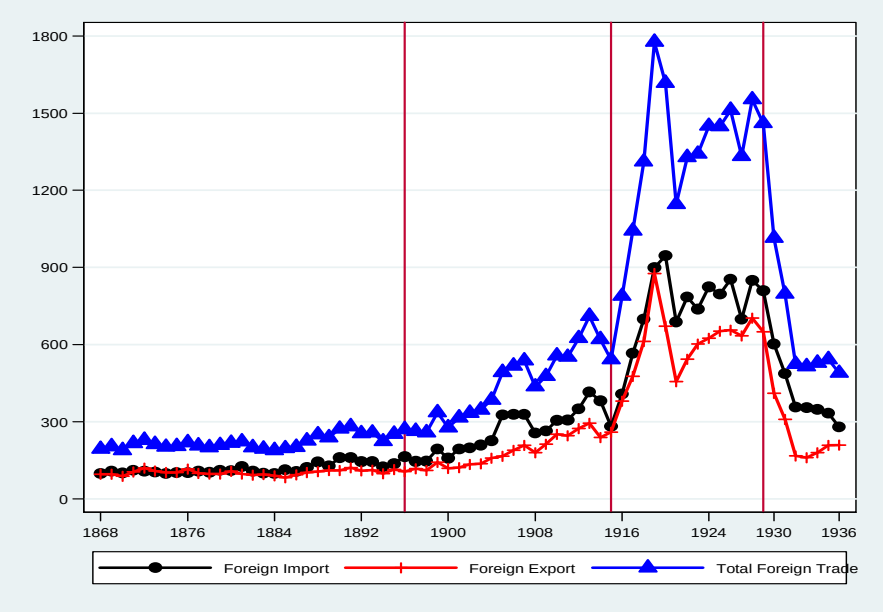
Table 8. Education Statistics in Modern China

Year	Number of Modern Schools	Number of Students	Number of High School Students	Number of College Students
1912	52,500	2,933,387	97,798	40,114
1913	87,272	3,643,206	119,026	38,373
1914		4,075,338		32,076
1915	122,280	4,294,251	164,981	25,242
1916	129,528	3,974,454	146,561	17,241
1921		4,987,647		
1922	178,847	6,615,772	118,656	
1923		9,000,733		
1924		11,067,635		
1925			129,978	36,321
1928			234,811	35,198
1929	214,572	9,252,222	341,022	29,123
1930	253,917	11,501,152	514,609	37,566
1931	262,992	12,301,611	536,848	44,167
1932	266,578	12,812,983	547,207	42,710
1933	262,328	12,985,735	559,320	42,936
1934	263,915	13,771,380	541,479	41,768
1935	294,724	15,694,589	543,262	41,128
1936	323,452	19,034,124	627,246	41,922

Source: Xiong (1990)

Figure 1. China Foreign Trade Statistics, 1868-1936

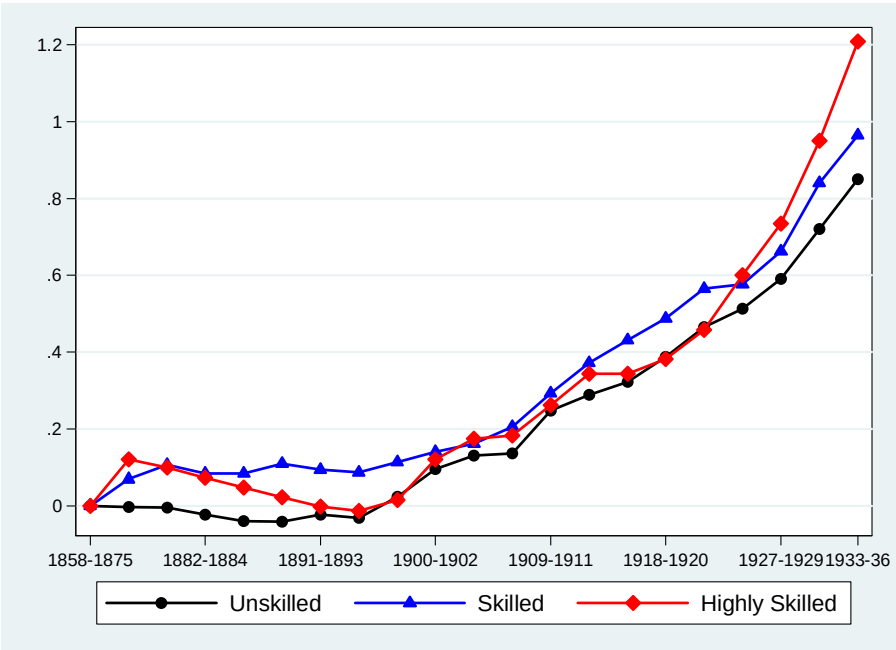
Unit: thousand U.S. Dollars



Sources:

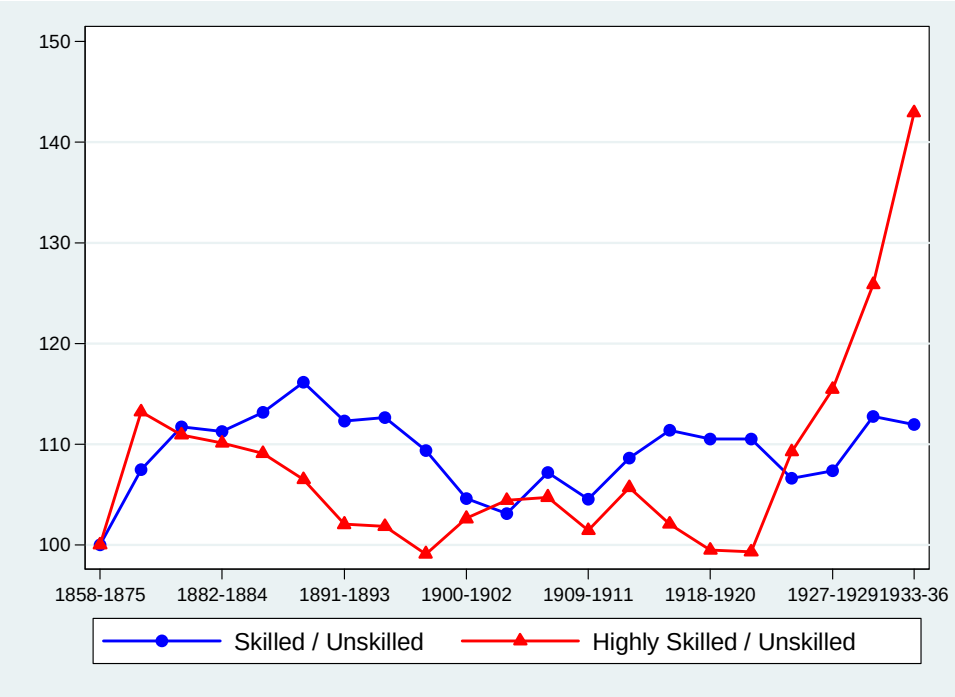
Raw data are from Hsiao (1974) and are originally in silver taels. They are converted to US dollars using exchange rates recorded in Hsiao (1974). 1 Tael roughly equals 1.558 US dollars.

Figure 2. Log Nominal Wages, 1858-1936



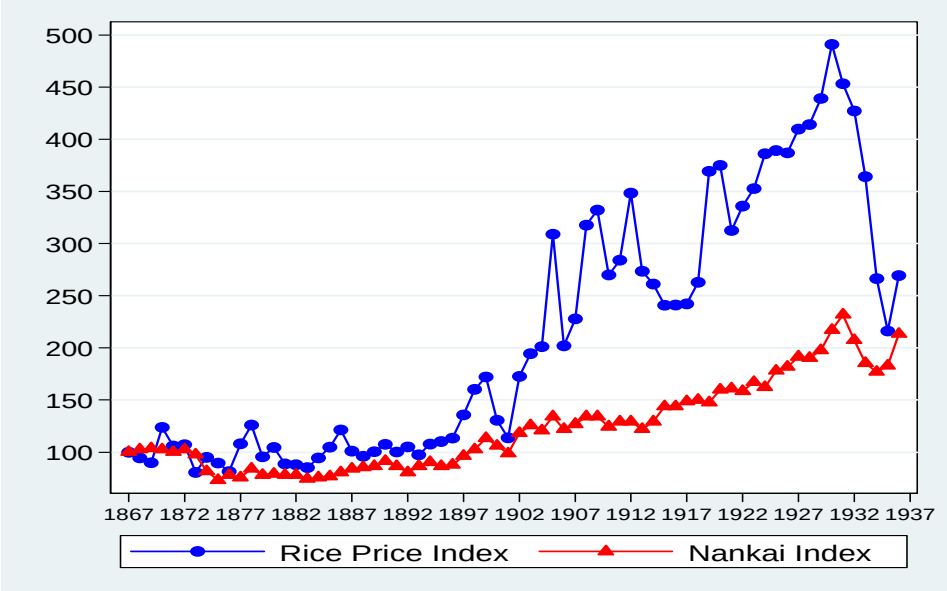
Source: Authors' calculation as described in the text.

Figure 3. Skill Premia on Nominal Wages, 1858-1936



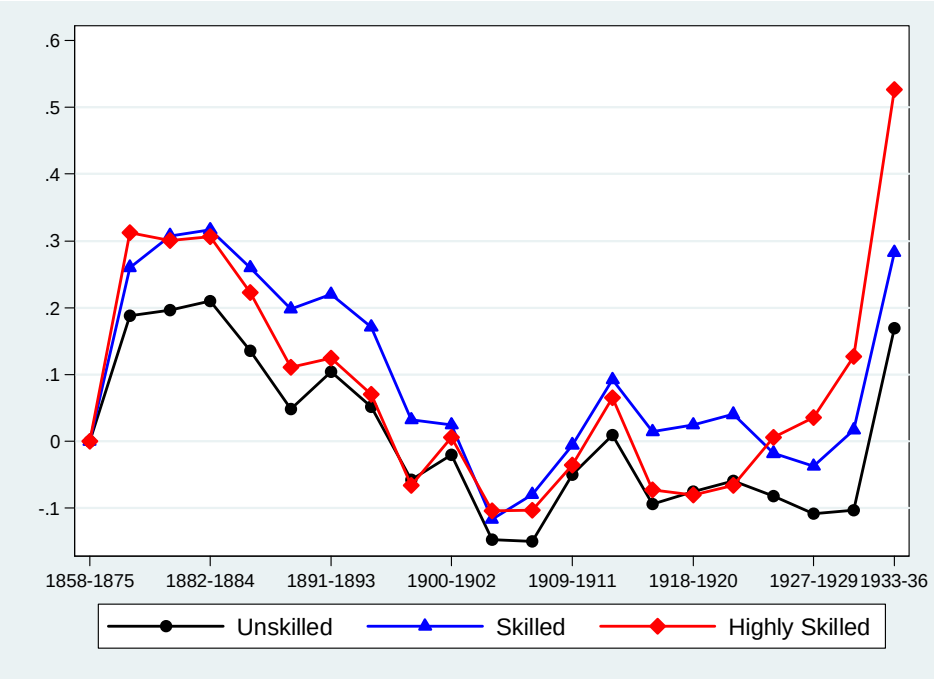
Source: Authors' calculation as described in the text.

Figure 4. General Price Index and Rice Price Index



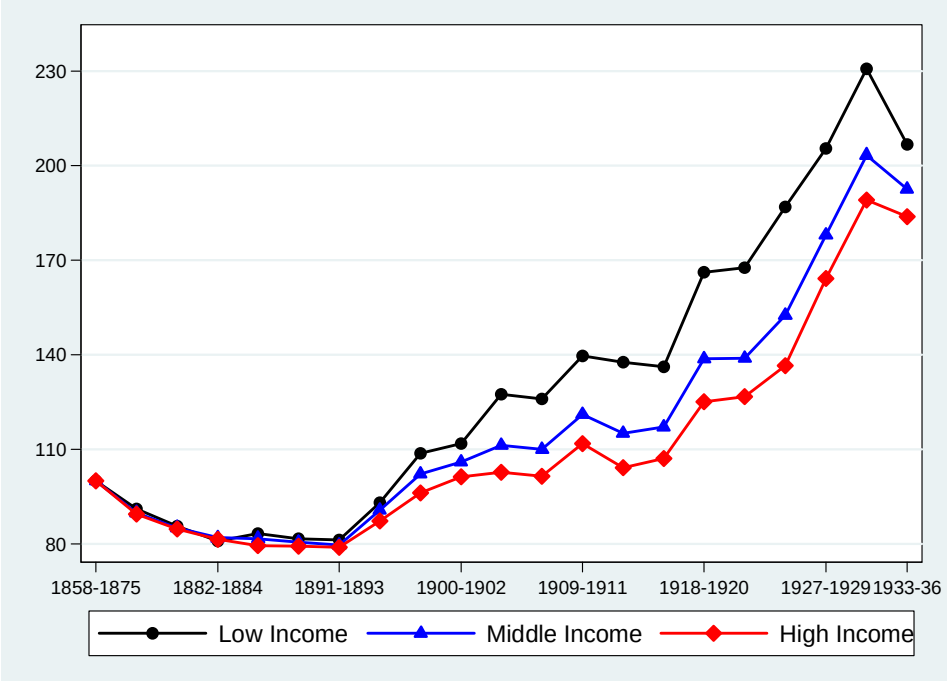
Sources: General Price Index is calculated with the data from Kong (1988);
Rice Price Index is calculated with the data from Hsiao (1974)

Figure 5. Real Wages for Three Skill Groups Using General Price Index as Deflator



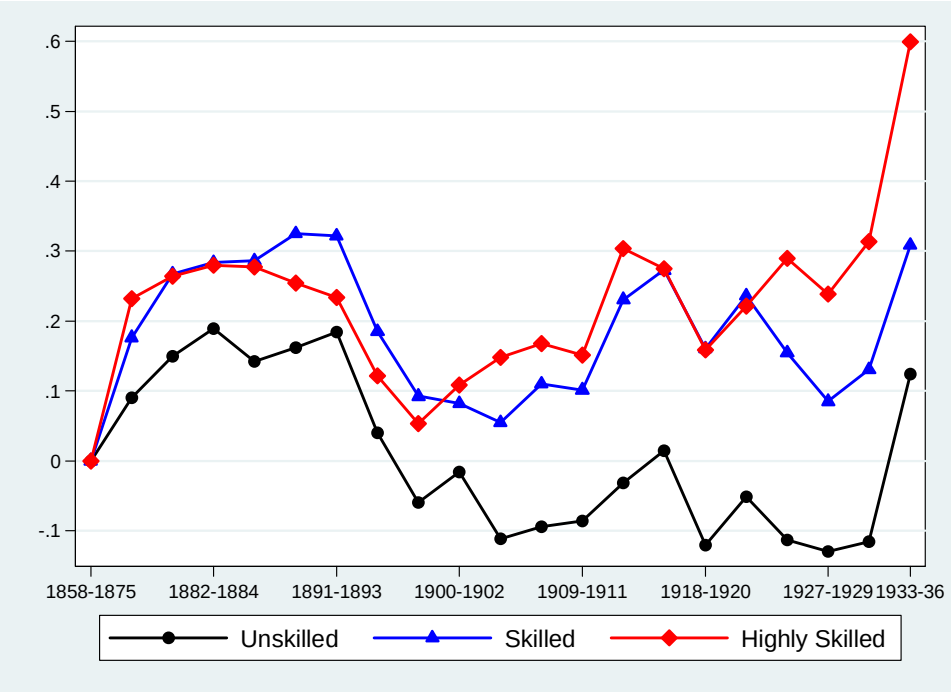
Source: Authors' calculation as described in the text.

Figure 6. Cost of Living Indices for Three Income Groups



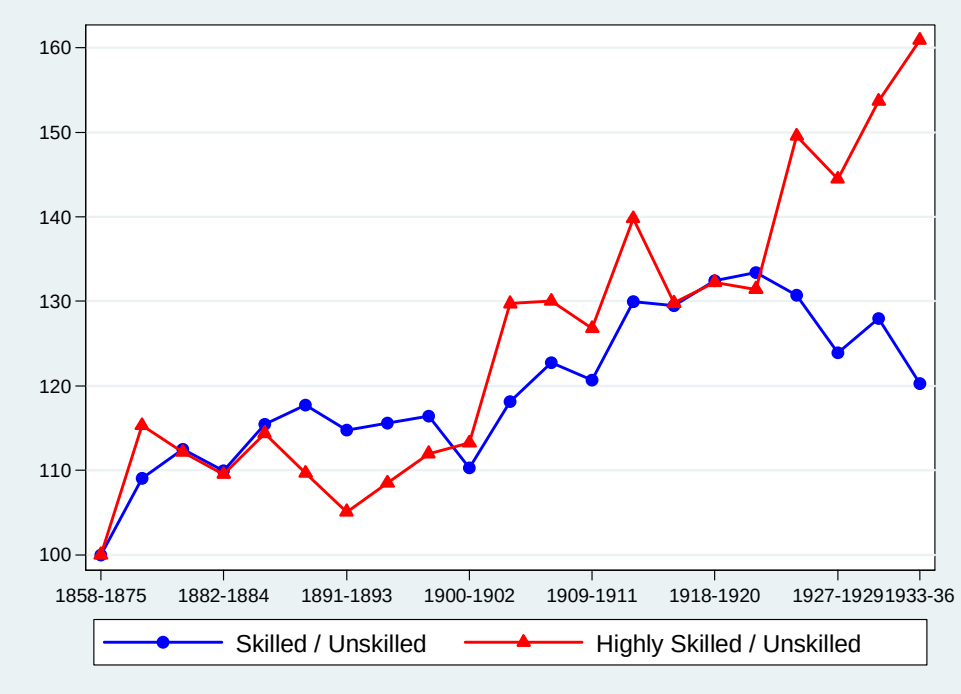
Source: Authors' calculation as described in the text.

Figure 7. Indices of Log Real Wages, 1858-1936



Source: Authors' calculation as described in the text.

Figure 8. Skill Premia of Real Wages, 1858-1936



Source: Authors' calculation as described in the text.

Appendix 1. Sampling Cities and Their Regional Groupings

Table A1. Regional Distribution of Cities

Region	Number of Cities	List of Cities
Manchuria	12	Aihui; Andong; Dalian (Dairen); Harbin; Hunchun; Jilin; Longjingcun; Manzhouli; Shenyang (Moukden); Sanxing; Suifenhe; Dadongkou
North China	8	Beijing (Peking); Yantai (Chefoo); Qinhuangdao; Qingdao (Kiaochow); Longkou; Newzhuang; Tianjin (Tientsin); Weihaiwei
Lower Yangtze	7	Zhengjiang; Hangzhou; Nanjing; Ningbo; Shanghai; Suzhou; Wenzhou
Upper and Middle Yangtze	8	Changsha; Chongqing; Yichang; Jiujiang; Shashi; Wanxian; Wuhu; Yuezhou
Southeast Coast	9	Xiamen (Amoy); Guangzhou (Canton); Fuzhou; Jiangmen (Kongmoon); Kowoon; Lappa; Sanshui; Sanduao; Shantou (Swatow)
Guangxi and Hainan	6	Haikou; Qiongzhou; Longzhou; Nanning; Beihai (Pakhoi); Wuzhou
Southwest Frontier	4	Mengzi, Simao; Tengyue; Yadong
Taiwan	3	Tainan; Takow; Tamsui
Northwest	3	Gansu; Shaanxi; Shanxi

Appendix 2. Categorization of Occupations

Table A.2. Summary Statistics of Data by Occupation

Unit: silver dollar

Unskilled	**N**	**median pay**	**Skilled**	**N**	**median pay**
lighthouse keeper	3115	9.91	postal clerk	5836	28.33
watchman	324	11.22	teacher	99	28.33
coolie	665	11.42	office secretary	5781	30.59
gatekeeper	93	11.42	pilot	46	39.66
cook	110	12.24	engineer	247	40.21
office assistant	86	12.65	shroff	125	42.49
office boy	277	16.32	writer	3431	46.74
soldier	263	12.24	copyist	4108	46.74
boatman	420	12.75	**subtotal**	**19673**	**35.41**
boat crew	70	12.75			
sampanman	96	14.00	**Highly Skilled**	**N**	**median pay**
seaman	327	14.28	harbor master clerk	199	49.57
fireman	144	18.36	tidewaiter	1932	63.74
weigher	296	18.41	revenue auditor	635	80.73
carpenter	267	19.74	probationary clerk	248	70.82
boatswain	132	21.81	clerk	13715	84.98
junk boat owner	164	25.21	assistant	280	226.62
watcher	778	27.53	**subtotal**	**17009**	**77.90**
subtotal	**7622**	**12.46**	**Total**	**44304**	**42.49**

Appendix 3. Steps of Constructing Commodity Price Series

The sources of price data for tradable goods are a series of publications called "*Returns of Trade and Trade Statistics*" compiled by the CMC's statistical department. For each year, quantities and values of each commodity in every treaty port were recorded in the publications. Totally there are more than 160 volumes of trade records, and the size of the database is incredibly huge.

Step 1 of constructing the price series is to computerize these 160-volume hardcopies of the trade publications. This is an extremely time-consuming and costly project. Headed by Robert Bickers and Hans van de Ven, historians from University of Cambridge and University of Bristol and I together have spent more than three years in computerizing these publications. The database will eventually publish on the Internet for everybody's use.

Step 2 is to calculate the national average prices from trade publications. First, I add up port-level data for each selected commodity into national aggregates. I divide the national aggregated values by the quantities, and obtain the invoice prices of each selected commodity by each trade type. Trade types include foreign export, foreign import, domestic export, domestic import, and re-export. I focus on foreign exports and imports because of data availability.

This step is very difficult to do, because measurement and currency units varied tremendously over locations and over time. The details of converting measurement and currency units are reported in Appendix 4.

Step 3 is to convert invoice prices into wholesale prices. Before 1903 the CMC recorded the market values of commodities, so no adjustment is needed. Since 1904, the CMC began to report FOB and CIF values for exports and imports. The following equations are adopted to recover the market values from FOB and CIF values:

CIF Value = (Market Value – Import Tariff) * (1 – Miscellaneous Fees)

FOB Value = Market Value * (1 + Miscellaneous Fees) + Export Tariff

Both import and export tariffs were 5%. The miscellaneous fees of imports and exports are estimated to be 7% and 8% respectively. Therefore the adjustment equations are:

Market price of Imported Goods = CIF Value / 0.88

Market price of Exported Goods = FOB Value/ 1.13

The last step is to adjust the systematic bias in the CMC's trade statistics. Most scholars agree that the CMC has consistently undervaluated the export values, but they have different estimations on the degree of this bias. On the basis of existing studies, the following estimation scheme is adopted: export values were undervaluated by 5% before 1914; by 8% in 1914-1919; by 10% in 1920-1929; by 12% in 1930-1936. The export values and prices are adjusted accordingly.

Appendix 4. Detailed Processes of Constructing Commodity Price Series

Food consumption:

1. Rice: rice price series are available from various sources. Hsiao (1974) recorded quantities and values of imported rice. The CMC also recorded average export and import rice prices in their decennial reports from 1862 to 1921. In this project, the long-run export rice prices from 1859 to 1936 are drawn from all CMC's raw trade statistics.

2. Millet: millet and other grains such as sorghum are substitutes for rice, and were mainly consumed by the poor. The millet export data were also recorded in the trade statistics, their prices showed the similar trend with rice price.

3. Wheat flour: all prices are extracted from foreign import trades. The CMC collected prices of imported wheat flour and exported wheat in 1862/1921. This project calculates imported wheat flour prices from the raw trade data.

4. Meat: the CMC recorded meat imports and exports, which includes a collection of pork, beef, mutton and etc. Here I use meat export prices.

5. Fish: the CMC's trade records report dried and salted fish of all kinds. The fish price series here is generated from the export records of dried and salted fish.

6. Egg: the egg price series is generated from native export of preserved eggs, recorded in the CMC trade publications.

7. Beans: green, yellow and red beans were recorded under the category of bean in the CMC's trade statistics. Beans are large items in the Chinese diets. They are either directly consumed, or made into bean curds and other products. The data source is the export of green beans from the CMC trade records.

8. Turnip: turnip, dried and salted, is one of the most widely consumed condiments in China. Its native export was recorded and the price was calculated here.

9. Vegetables: other vegetables together were recorded in the raw trade statistics. The native export data were recorded here, and the price was calculated.

10. Bean oil: bean oil is the most widely used cooking oil in China, supplemented by groundnut oil and sesame oil. The CMC's Decennial Reports recorded the prices of bean oil and groundnut oil, and they show quite similar trend. Therefore using any of these oils does not affect the long-run trend. This project uses the bean oil price from the Decennial Reports, supplemented by prices from 1922 to 1936 calculated from raw trade data.

11. Sugar: sugar is one of the most important condiments in the Chinese diet. Sugar foreign import price here is calculated from trade statistics recorded in Hsiao (1974). Data from 1864 to 1866 are calculated from the original trade statistics.

12. Salt: salt used to be an expensive condiment in China. Its native export was recorded in the CMC's records, and was used to calculate salt price.

13. Pear: pear is a common fruit in China. Its native export was recorded in the CMC's records, and was used to calculate its price.

14. Persimmon: persimmon is also a common fruit in China. Its native export was recorded in the CMC's records, and was used to calculate its price.

Clothing consumption:

15. Shirting: shirting is used to represent all the mechanized cotton cloth. There were different types of imported cloth, such as grey shirting, white shirting, sheeting, drills, lastings, etc. Using shirting import data to calculate its price and represent other cotton cloth does not lead to a significant bias.

16. Nankeen cloth: many Chinese people, mostly with low income, consumed a large amount of native cloth. Nankeen cloth is a typical native cloth, and its price is calculated from the export record.

17. Silk: white silk was a luxury and was mostly consumed by the rich. China is a major silk-exporting country, and white silk export price is calculated from trade statistics.

Fuel and lighting consumption:

18. Coal: coal is the most important fuel and heating material in China. Its export data were recorded in the trade statistics and the price is calculated accordingly.

19. Kerosene: kerosene was the most important lighting material, and was all imported. Its price is calculated from the raw trade data.

Housing consumption

It is common that housing costs were not available in economic history of many countries. To remedy this, I adopt the estimation method proposed by David and Solar. This method approximates rent by an index of the reproduction costs or prices of new structures. This index is calculated as a geometric average of indices for building materials prices and common labor wages. This method corresponds to the dual of a constant return to scale Cobb-Douglas production function characterized by output elasticities of 0.5 for its labor and material inputs. My common labor wage series was developed from the CMC's unskilled employees' wage dataset. I extract prices of two building materials from the CMC's trade publications: timber plank and cement, each is weighted by 0.25.

An alternative way of estimating housing costs is to collect scattered or anecdotal records from historical materials. Among them, the CMC's Decennial Reports provide relatively richer observations. However, all these observations are port-level. There is no practical way of aggregating national averaged rents from these fragmentary records. The impression from these observations is that the housing costs started to grow rapidly in the early twentieth century due to fast urbanization. Improved transportation condition lowered the cost of migration for rural labor, and industrialization provides more job opportunities for them. Booming population in cities led to a sharp increase in housing costs. This coincides with the picture generated by the first estimation method.

The sources of price series for two building materials follow below:

20. Cement: foreign imported cement records are used to calculate prices.

21. Timber plank: native export records of timber plank were recorded in the CMC's trade publications, and the price is calculated accordingly.

Sundry consumption

22. Unskilled wages: unskilled wages are generated in Section 4. Unskilled labor represents many types of service costs in sundry category.

23. Skilled wages: skilled wages are also generated in Section 4. Skilled labor represents expenses on teachers, doctors, etc.

24. Green tea: green tea is widely consumed in China. Its export price is calculated from trade statistics recorded in Hsiao (1974). Data from 1864 to 1866 are calculated from the original trade statistics.

25. Tobacco: foreign imported prepared tobacco records are used to calculate prices.

26. Liquor: foreign imported liquor records are used to calculate prices.

27. Blanket and carpet: their foreign import records are used to calculate prices

28. Paper: paper represents consumptions on reading and writing. Its import records are used to calculate prices.

29. Chinaware: chinaware represents housing furnishing expenses, and its native export records are used to calculate prices.
30. Clock: clock import records are used to calculate prices.