

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Real Wages in Manufacturing, 1890-1914

Volume Author/Editor: Albert Rees

Volume Publisher: UMI

Volume ISBN: 0-87014-069-8

Volume URL: <http://www.nber.org/books/rees61-1>

Publication Date: 1961

Chapter Title: APPENDIX A Sources and Characteristics of State Earnings Data

Chapter Author: Albert Rees

Chapter URL: <http://www.nber.org/chapters/c2289>

Chapter pages in book: (p. 131 - 136)

APPENDIX A

Sources and Characteristics of State Earnings Data

THE period covered by our study was the period of greatest activity of state bureaus of labor statistics. This activity was largely inspired by one man, Carroll D. Wright. Wright headed the Massachusetts Bureau of Statistics of Labor, the first in the country, from 1873 to 1885, and was the first United States Commissioner of Labor, 1885-1905. The patterns he set in Massachusetts were followed by many of the other states, as is apparent from the notes that follow. Wright was also instrumental in organizing an association of officials of bureaus of labor statistics which was active during this period and helped to provide uniformity among the statistics of some of the states.

After World War I, as the statistical services of the federal government improved, the statistical activities of the states tended to decline, and many of the series used here were discontinued. Even during our period, though almost every state issued some labor statistics, there were great variations in quality from state to state. In some states, frequent changes in personnel or administrative arrangements produced frequent shifts in the kinds of data collected and gaps in many of the series. For several states, however, especially Massachusetts and New Jersey, we were impressed with the continuity and full coverage of the data, the detail in which they were presented, and the care that seemed to have been exercised in compiling them.

*Massachusetts*¹

Massachusetts is the only state whose data cover our entire period. The information provided is consistent throughout. For all industries combined, and for separate industries, the data include the number of establishments, the average number of wage earners, total wages paid, and average days in operation per establishment. Beginning in 1908, the industry classification used is identical with that of the census. The data for 1909 and 1914 are identical with the census data,

¹ SOURCE: Massachusetts, Bureau of Statistics, *Annual Report on the Statistics of Manufactures*, Nos. 6-29, 1890-1914.

APPENDIX A

but provide, in addition, the number of days in operation. There are no omissions of industries or employees.

Before 1908 the industry classification is not identical with that of the census but is rather similar, and there were a few difficulties in combining series into census industries. The coverage of census employment is not complete, but is generally very high. Through 1905, two sets of data are provided for each year, one covering the same establishments included in the preceding year, one covering the establishments included in the following year. The two differ very little. We have consistently used the second, which generally covers more employees because of the entry of new firms.

*New Jersey*²

The New Jersey data are consistent from 1895 to 1914 and provide the same information as the Massachusetts data. Industry definitions are consistent throughout but do not correspond with census definitions, being generally finer in industries important in the state. After 1899, coverage of employment is generally very high. Before that time the coverage is incomplete, and we have made little use of the data for individual industries.

*Pennsylvania*³

The Pennsylvania data consist of a number of distinct segments, which will be discussed in turn. With exceptions to be mentioned, they provide the same information as the Massachusetts data. In all segments, the industry classification does not correspond to the census and is very much finer in industries important in the state.

For 1892-1905 the data consist of series for two sets of identical establishments, one covering 1892-1901, the second covering 1896-1905. Here, and in Appendix C, we shall refer to these as the "1892 series" and the "1896 series." In general, the 1896 series has greater coverage. The coverage of both series is usually a rather small fraction of census employment. However, the use of identical establishments removes the principal difficulty we find in series with incomplete coverage—erratic movement resulting from changes in coverage from year to year.

² SOURCE: New Jersey, Bureau of Statistics of Labor and Industries, *Annual Reports*, Nos. 18-37, 1895-1914.

³ SOURCES: Pennsylvania, Bureau of Industrial Statistics, *Annual Reports*, Vols. 20-40, 1892-1912; Pennsylvania, Department of Labor and Industry, *Annual Reports*, Vols. 1-2, 1913-1914.

We have no way of knowing whether some of the same establishments are included in both series during the period of overlap. We have, therefore, not averaged the two to avoid double-counting some establishments. Our all-manufacturing series uses the 1896 series for 1896–1901 and links the 1892 series to it for 1892–95. In our industry series we have chosen between the two sets of data for the overlap period on the basis of the number of employees covered and the correspondence of annual earnings with the Census in 1899. Where simple acceptance of both series at the dates of transition (1895–96 or 1901–2) produced a movement similar to that of the continuous series, we have used this method. Elsewhere, we have linked the two series to remove differences in level.

From 1905 to 1912 the coverage of the Pennsylvania data increases greatly and becomes almost complete. The data no longer relate to identical establishments. For some industries sharp changes in coverage or industry definition within the period 1905–12 prevent us from accepting the series.

In 1913, the work of collecting statistics of manufactures was transferred from the Bureau of Industrial Statistics in the Department of Internal Affairs to the Bureau of Statistics and Information in the Department of Labor and Industry. This transfer produced serious discontinuities in the statistics. There were often sharp changes in the number of employees reported, and apparently in industry definition, which were not always reflected in industry titles. The resulting erratic movement of many earnings series led us to reject them. Where Pennsylvania series for 1913–14 were accepted, we had to deal with a special problem. The data for 1913 do not give the number of salaried employees and wage earners separately, although they do give salaries and wage payments separately. We have either assumed that the ratio of wage earners to all employees was the same in 1913 as in 1914, or, in some cases, we have estimated the number of salaried workers in 1913 by making a linear interpolation of average annual earnings of salaried workers from 1912 to 1914 and dividing the 1913 figure into salaries paid.

*Ohio*⁴

The Ohio data are very different in form from those of the other states we have used. The data consist of the number of workers employed, average number of days worked, average daily earnings,

⁴ SOURCE: Ohio, Bureau of Labor Statistics, *Annual Reports*, Nos. 18–37, 1893–1912.

APPENDIX A

and average yearly earnings *by occupations* within industries, separately for Cincinnati, Cleveland, Columbus, Dayton, Toledo, other cities, and villages. No averages whatever are provided for these many observations for occupations. Within industries we have computed employment-weighted averages of daily wages for all occupations and localities combined. These become our interpolating series. For census years only, we have computed similar employment-weighted averages of days worked to be used in estimating average daily earnings from census data. These estimates of daily earnings serve as benchmarks for the interpolating series.

The industry definitions of the Ohio data are consistent throughout, though they do not correspond exactly to census definitions. The coverage of the data is very good after 1900 but often low before that time, so that we have had to reject the early portions of some Ohio series.

*Wisconsin*⁵

The Wisconsin statistics are, in general, similar to those of Massachusetts. However, the series are interrupted by a three-year gap, for 1900, 1901, and 1902, resulting from a fire in the State House that destroyed the records of the Bureau of Labor and Industrial Statistics. The industry classifications are consistent throughout and, for the series we have used, seem to be similar to census definitions. Prior to 1896, there are data for weeks in operation rather than days in operation; we have converted these to days by multiplying them by six. There are also no employment data before 1896, so that Wisconsin is excluded from the 1889 estimates of coverage of census employment in Appendix C, although we use some Wisconsin data back to 1890.

*Connecticut*⁶

The Connecticut data are also similar to the Massachusetts, but cover only five years, 1900–1904. They are on a fiscal-year basis, with years ending June 30. We have converted them to a calendar-year basis by averaging adjacent years. We thus lose one year's observations, and the data appear in some of our series for the four years 1900–1903. Nevertheless, the employment covered by the Connecticut series is included in the employment covered by our series for 1899

⁵ SOURCE: Wisconsin, Bureau of Labor and Industrial Statistics, *Biennial Reports*, Nos. 4–9 and 11–13, 1889–99 and 1903–7.

⁶ SOURCE: Connecticut, Bureau of Labor Statistics, *Annual Reports*, 1900–1904.

APPENDIX A

and 1904 in the coverage tables of Appendix C. The Connecticut industry definitions did not present any special problems.

*Rhode Island*⁷

The Rhode Island data are similar to the Massachusetts, though they cover only the textile industries for 1893–1903. Two sets of data are presented for each year, one comparable with the preceding year and one with the following. In each pair, we have used the data with the largest coverage. The days in operation per year are given separately by establishments; we have computed employment-weighted averages. In the woolen industry some establishments are reported as operating more than 312 days per year, in some cases more than 365 days; the highest figure is 588. All entries of more than 312 days are marked “running overtime.” We have assumed that this refers to two-shift operations and that the number of shifts operated was counted. In computing average days in operation, we have entered all establishments “running overtime” as operating 312 days.

The industry titles are unchanged during the period and correspond to census industries.

*Maine*⁸

The Maine data are for two industries only, cotton and woolen mills. They are for fiscal years ending June 30, for 1895–1903. We have averaged adjacent years and the data appear in our series for 1895–1902. The time in operation is presented in weeks, which we have converted to days by multiplying by six. The data on time in operation are for individual establishments, and we have computed employment-weighted averages.

*South Carolina*⁹

The South Carolina data include average employment, total wages, and average days in operation. We have used only one series, textiles, which is described in Appendix B under cotton.

⁷ SOURCE: Rhode Island, Bureau of Industrial Statistics, *Annual Reports*, Nos. 8–18, 1893–1903.

⁸ SOURCE: Maine, Bureau of Industrial and Labor Statistics, *Annual Reports*, 1895–1903.

⁹ SOURCE: South Carolina, Department of Agriculture, Commerce, and Industries, Labor Division, *Annual Reports*, Nos. 1–6, 1909–14.

APPENDIX A

Other States

A number of other states published some data on employment and earnings during 1890–1914. In one case, Missouri, we made extensive attempts to use the data and discarded them because of rapidly shifting coverage, which produced very erratic wage series. The other data were rejected at the outset. For several western states, the data covered too few workers to have an appreciable effect on any of our series. In two cases, West Virginia and Washington, the data were available only for every second year. New Hampshire data provided the number of days in operation for 1893–98 only, a period that does not include a census benchmark. Virginia data provided days in operation only as establishment-weighted averages. Illinois and Indiana had data only for a few scattered years. The remaining states omitted some essential item from the data, either number of employees or number of days in operation.