Appendix: Data Description

Chapter 4

These two appendices explain in detail the sources used and methods employed to construct the data. All data that was generated or modified by me will be provided here in the final form it was used. Data that was used in its original form and is publicly available will not be provided here, but a citation to the source is supplied.

Chapter 4, Data Appendix 4.A

Some Preliminaries

It is important to acknowledge, at the outset, that the quality of data is inversely related to their age. Data prior to 1850 are of very limited value, and statistics for 1850 -1870 have some merit but must be examined with great care. Although the quality of data improves in the late 19th century, some skepticism about the estimates for that period is justified. But such doubts should not prevent a serious discussion of late 19th century economic developments, since that period is of great importance for the study of economic growth and of wealth and income inequality. Hence, we ought to double our efforts to solve the problem at hand with the data that can be reconstructed for that period. There is no doubt the results contain significant errors in measurement, but there is no evidence that after 1889, the starting date of my study, such errors have any systematic bias. With the same confidence I suggest that the *order of magnitude* of these results is correct and just for that reason we must pay attention to them.

I started in 1889, by using Kendrick's (1961) data as a base, which is an integration of the vast work done before him by pioneers such as Kuznets, Goldsmith, King and Douglas. However, Kendrick focused on measuring productivity, hence his database is not directly applicable for income and wealth distribution. Given these limitations, my goal was to use his data files as a foundation for reconstructing data for the economic sector where market power is exercised. In the later period 1925-2017 I stud the sector defined by BEA as the "domestic corporate sector." For the earlier period 1889-1929 the closest one can approximate the BEA definition is by constructing the data for the "private, domestic, non-farm, non-residential economy," which is what I did. Kendrick's sector which comes closest to my need is his *Domestic Private Economy,* although for some variables data are available for the Private, domestic, *non-farm economy* or the private, domestic, *non-farm, non-residential* economy. Kendrick's data files contained virtually no information on capital goods prices. They offer no detailed files on values of capital goods employed in the farm sector, which had to be subtracted from the capital stock employed in the "private domestic economy." They also provided no information on the service value of residential structures in the domestic economy, needed for estimating gross value added of the private, domestic, non-farm, non-residential economy. In this appendix I detail how I reconstructed these.

General Guidelines

Here I provide a few comments that aim to convey the general ideas used in construction of the series.

(4A.1) Kendrick's data and productivity analysis use only real variables and therefore all his data files were built, from the start, in 1929 prices. No implicit price deflator is available for the value added of the domestic, private, non-farm, non-residential sector. The need to use real terms data requires capital rental functions to be built in real terms. This is different from the later study for 1925-2017, which is conducted in nominal terms and accurate implicit prices of all goods in the model are available from the BEA for that period. Since Kendrick does not provide capital goods' prices, it was necessary to go back to his original source, which was the three volumes of Goldsmith (1955) Tables W-1 and W-3 of Volume III that report on various categories of capital goods in current and 1929 prices. Unfortunately, Goldsmith's study covers only the period of 1896-1929, requiring extrapolations back to 1889, made possible with data in Kuznets et al. (1946) and Kuznets (1961) where economic variables are reported as averages of overlapping decades. Extrapolations of such variables are more complex and the precise sources are provided in Data Appendix A.

(4A.2) Kendrick's Gross Value Added for the domestic, private, non-farm economy *includes* imputed service value of residential structures, which is a significant component of value added. Unfortunately, Kendrick (1961) does not provides the data for this service value variable, which is needed to exclude the residential component from value added. To carry out this imputation I used the information available in Kuznets et al. (1946) with which I could estimate this component of value added. Both the Kenrick Value Added and my estimates of the service value of residential structures are provided in Data Appendix A.

(3) Although Kendrick's Tables A-XVI and A-XV record detailed capital stock information for the private, domestic non-residential sector, he does not separate the *farm sector* from the capital data. It was thus necessary for me to reconstruct the detailed capital structure of the farm sector for 1889-1929, which was made possible by information provided by Kendrick (see footnote on page 325). The final, detailed data files of the structure of the capital stock for the domestic, private, non-farm, non-residential sector are then reported in Data Appendix A.

(4A.4) Total number of full-time equivalent employed workers, and the average yearly earnings of private, non-farm employed workers. The files in Data Appendix A are put together from several sources. Annual earnings of private, non-farm employed for 1900-1929 are from Lebergott (1964), and for 1889-1899, I used two files from Historical (1985). Full-time equivalent private non-farm employed for 1900-1929 are from Historical (1985) and for 1889-1899, I used Kendrick (1961).

(4A.5) Inventory data are provided by Kendrick (1961). Inventory Valuation Adjustment for 1896-1929 is from Goldsmith (1955) and extrapolated for 1889-1899, by regressing the available data on the rate of change of GNP.

As noted, all 1929-2017 data used for computing relative factor shares are stated in current values. Since all of Kendrick's (1961) data files were formulated in 1929 prices, it was convenient to carry out all computations for 1889-1929 using real values at 1929 prices. This requires other variables, such as rental rates and interest rates, to be converted to real values. This procedure was supported by the fact that Goldsmith (1955) was the main source for capital goods' prices, and although these prices were the main source used by Kuznets for constructing the national accounts, those prices are viewed as less accurate than desired. To avoid this error in variables, I used only the average price of capital goods in computing rents and used wherever I could Goldsmith's (1955) values in 1929 prices.

Data Sources for the Private, Domestic, Non-Residential, Non-Farm Sector 1889-1929

Gross Value Added:

- 1889-1929- Kendrick (1961) Table XXIII Column 2 for the Private Non-Farm sector.
- To subtract the rental service value of residential structures I first construct the price index for residential rent: for 1889-1913 interpolate the means of overlapping decades, Table III-10 Columns 1 and 4 in Kuznets (1946, Part III) and for 1913-1929 use Series Cc4 in Table Cc3-5 in Millennial Historical (2006). To construct values in 1929 prices: for 1889-1913 use Column 4 of Table III-10 in Kuznets (1946, Part III) and for 1913-1929 use Column 1 and the constructed price index which is provided

Compensation to Employees: multiply annul earnings by number of full-time equivalent employed

- Annual earnings of private, non-farm employed: 1900-1929 from Lebergott (1964) Table A-17 and for 1889-1899 I used Historical (1985) Series D-780, based on Douglas (1930).

- Full-time equivalent private non-farm employed for 1900-1929 are from Historical (1985) Series D-7 for all non-farm employees and Series D-139 for Government employed (update of Lebergott (1964)). For 1889-1899 I used Kendrick (1961) Table A-XXIII (column 3). Since this variable is computed by using economy wide averages, it includes compensation to officers of firms.

Taxes on Production

- For total private economy in Kendrick (1961) Table A-IIa Column 2. Then allocated to the private, non-farm, non-residential economy in proportion to value added.

Real Non-Residential Capital Stock: Structures, Equipment and Inventories

- Initial values: Kendrick (1961) Table A-XVI
- Then use Kendrick (1961), footnote a on page 325 and Table B-III to reconstruct the real three capital stocks for the farm sector
- Subtract farm capital stock values from the values in Table A-XVI

Capital goods' prices: Structures, Equipment and Inventories

- 1896-1925 Goldsmith (1955). For 1929 prices, Table W-3 Columns 6, 7 12 and 17 for Current prices, Table W-1 Columns 5,6,11 and 16.
- For 1889-1896 my interpolation using Goldsmith's estimates for 1880,1890 reported in Millennial Historical Table Ce209-232 page 3-325

Depreciation of Non-Residential Fixed Assets

- Kendrick's (1961) depreciation rates for the aggregate economy are about 2.5%, reported in Table A-1 column 2, and these are dominated by depreciation rates of structures. Since the composition of non-residential capital changed over time, I used the BEA depreciation rates for 1925-1929 which were 2.56% for structures and 11.79% for equipment.

Inventory Valuation Adjustment

- 1896 - 1929 - Goldsmith (1955) Table P-19, Column 7 and then use Goldsmith's implicit price deflator for inventories (using Goldsmith's (1955) capital good prices implied by Tables W-1 and W-2), deflate the IVA to 1929 prices
- 1889-1895 - Extrapolated by regressing the Goldsmith IVA data on the rate of change of the GNP, using the Balke and Gordon (1989) GNP data

Interest Rate:

- Unadjusted Index of Yields of American Railroad Bonds, Historical Series X- 476

Rate of Change of the CPI, Annual Rate

- Millennial Historical Series Cc1, page 3-158

Chapter 4: Data Appendix 4.B

Data Sources for the Domestic, Non-Residential, Corporate Sector 1925-2017

Gross Value Added:

- 1929-2017- BEA Table 1.14
- 1925-1929 - interpolated using Balke and Gordon (1989) GNP data

Compensation to Employees:

- 1929-2017- BEA Table 1.14
- 1925-1929 - interpolated using "Compensation to Employees 1889-1929" in Appendix B

Taxes on production and imports less subsidies

- 1929-2017 - BEA Table 1.14
- 1925-1929 - used tax categories to compute 1922 and 1927 values in History (1985), Series Y590-604 and interpolated in between

Officers' Compensation

- 1990-2013 Table 13 of the IRS OSI
- 1929-1990 In Table B1 of Piketty-Saez Database, updated 2015 courtesy of Emanuel Saez.
- 2013-2017 interpolated using total wage income

Current Cost Net Stock of Non-Residential Fixed Assets

- Nominal Value: Structures, Equipment and Intellectual Property Assets - BEA Table 4.1
- Real Index: Structures, Equipment and Intellectual Property Assets - BEA Table 4.2

Current Cost Depreciation of Non-Residential Fixed Assets

- Nominal Value: Structures, Equipment and Intellectual Property Assets - BEA Table 4.4
- Real Index: Structures, Equipment and Intellectual Property Assets - BEA Table 4.5

Non-Financial, Non-Farm Business Inventories (current value)

- 1945-2017 - Federal Reserve Bank of St. Louis (FRED), Series BOGZ1FL145020011A provide inventory data for the non-farm business sector. To deduce figures for the corporate sector the data are adjusted by the ratio of fixed assets employed by the corporate sector divided by the fix assets employed by the non-farm business sector
- 1925-1945 - Goldsmith (1955), Volume III, Table W-1, Column 16

Non-Financial, Corporate Business Inventories (current values)

- Compute percent corporate non-financial fixed assets in private non-financial non-farm fixed assets - BEA Table 6.1, and multiply by the non-financial, non-farm business inventories.

Inventory Valuation Adjustment (current values)

- 1929-2017 - Federal Reserve of St. Louis (FRED), Series B058RC1A027NBEA
- 1925-1929 - Extrapolated by regressing on the first difference of inventory values

Interest Rate:

- Moody's Seasoned Aaa Corporate Bond Yield Annual Average - Federal Reserve Bank of St. Louis (FRED), Series, AAA

- War Adjusted Moody's Seasoned Baa Corporate Bond Yield Annual Average - Federal Reserve Bank of St. Louis (FRED), Series, AAA and computations explained in the text.

Rate of Change of the CPI, Annual Rate

- BLS Series CUUR0000AA0
- 1930-1932 deflation adjusted CPI - BLS Series CUUR0000AA0 in which the change in 1930, 1931 and 1932 are set to zero.

Chapter 5

Three main data sources were used to construct the data files employed in assessing GE's market power: GE's financial reports; data obtained from the Museum of Innovation and Science in Schenectady, New York; and *Historical Statistics of the United States Colonial Time to 1970*, ("Historical," for short). The GE annual financial statements for 1892-1939, which are part of the annual reports, are my major data source, but I was unable to obtain any background source data used by accountants to prepare these reports. Consequently, I do not have data on the composition of cost or sales, and no direct information was provided on total wages.

5A.1 *Separating affiliates' data*. The problem of GE's affiliates arose upon GE's formation in 1892 as a merger of the Edison General Electric Company of Schenectady, New York, and the Thomson-Houston Electric Company of Lynn, Massachusetts. Both firms owned diverse assets, some of which arose from their active participation in the formation of local electric utilities across the country. In acquiring these assets, GE planned to dispose of some of them, and with that in mind, it took on some debt that was scheduled to be repaid by selling these assets. The recession of 1893-1894 resulted in a decline of their values, which caused GE to sustain serious losses. The company was forced to engage in financial management and trading of its debt and diverse assets, making it difficult to distinguish between normal production and sales decisions, on the one hand, and emergency financial decisions, on the other. For this reason, *I chose 1894 as the starting date for my analysis*, not 1892, when GE was formed.

GE ended up keeping most of the affiliates it received in 1892, and over time, it added many more. Its policy was not to interfere in the decisions or activities of the local firms, but the cash flow from these affiliates helped maintain GE's financial stability. The company's financial statements provide detailed information on the value of securities that GE owned as well as interest and dividends that GE received from its ownership of the securities of its affiliated companies. The value of these assets constituted about 25% of the value of GE's total assets before the Great Depression and approached 50% during the 1930s. Since these assets and incomes reflect the production activities of other companies and since my aim is to study GE's monopoly power in its own operations, these assets and the related income had to be separated from the main reports for GE. This was not entirely easy.

The reports provide sufficiently detailed information that enabled me to separate the assets, but the flows were more difficult. Interest income is often an aggregate of income received from affiliates, and income received from cash holdings of the firm and the amount of cash held was motivated by both GE's sales and its investments. As I suggest in the text, I consider GE a combination manufacturer of electric products and financial holding company. All estimates and results reported in this chapter exclusively relate to GE the manufacturing company, *excluding all affiliated companies and all transactions associated with financial*

management. It is, however, useful to stress that such a division entails some risk since GE's association with affiliates strengthened its monopoly power because the affiliates had intimate business relationships with it and many were its direct customers, buying equipment from GE or, in some cases, selling GE's products.

5A.2 *Revenue and capital*. I used the following definitions for revenue and capital employed:

- *Total sales*. Provided directly in most reports, except for those from the early years, a period when an arcane accounting system was used and for which sales have to be constructed from the data provided.
- *Fixed assets*. Provided directly as an aggregate of buildings and equipment, which cannot be separated. Additional real estate used by the corporation is provided separately in some years and needs to be added. The amounts reported are the undepreciated value of the assets.
- *Inventories*. Provided in all reports.

5A.3 *Total cost, income taxes and depreciation*. GE's reports provide information on the financial side of GE's operations, which includes debts, interest payments, asset holdings, and cost of operation. Detailed income taxes paid and depreciation charged are only separated out from total costs in some later years, while in most years, they are included in total costs. Fortunately, the data missing from the financial tables can be reconstructed, in each year, from the detailed information provided in the body of the investments section of the reports, enabling me to accurately reconstruct an accounting of taxes and depreciation charges. I was thus able to construct from the financial reports:

- *Cost of goods sold, including depreciation but before interest cost and income taxes* but without excluding labor cost
- *Depreciation charges*

Once I obtained data on the total labor cost, the above two items enabled me to compute

- *Value added = sales - (cost of goods sold including depreciation - depreciation) + labor cost*
- *Depreciation rate* = (total depreciation)/(value of fixed assets).

5A.4 *Rentals*. As a result of the aggregation of structures with equipment, I constructed risk premia and price indexes for the combined fixed assets based on data that I used for the entire private sector in Chapter 4. This means that the risk premia and the rate of relative price change of GE's aggregated fixed assets are the same as those used for the private sector in Chapter 4. The interest rates that I used for 1894-1929 are the unadjusted index of yields of American Railroad Bonds. For 1930-1939, I used the Moody's Aaa nominal interest rate, which is appropriate for the risk level of GE.

5A.5 *Constructing labor input and labor cost*. Precise data on labor employed and the total labor cost is, unfortunately, unavailable in the annual reports before 1927. Information about the number of GE's employees is scattered in the text of some annual reports, albeit not in a systematic manner. This missing information led me to the Museum of Innovation and Science in Schenectady, New York, which houses most of GE's records. One source that was made available is a 1922 unpublished internal report by Mr. M. F. Westover, entitled "*General Electric*

Company: Historical Summary." It offers a brief summary of the company's directors, officers and key events, but it also reports the approximate number of GE's total employees for 1896-1921. Finally, due to Swope's growing interest in labor relations, the GE annual reports for 1927-1939 contain extensive discussions of the labor force employed by GE, policy toward workers, and various benefits that GE provided workers. These reports provide the precise number of workers employed by GE and their average compensation. I combined these sources as follows:

- Employment and wages 1927-1938: I used the employment and compensation data from GE's annual reports under Swope.
- Employment for 1894-1926: I identified information on GE's labor employed from the Westover report, which suffered from some inconsistency across years, and therefore I matched with the scattered figures in the annual reports, such that I needed to impute labor employment for 1922 and 1924 only.
- Compensation for 1894-1926: For 1894-1919, I used the data on the mean annual wage of manufacturing workers in the US in the Historical Statistics (1971, Series D735-D740). GE's actual mean wages were reported in the annual reports for 1919, for 1923 and for 1927-1939. Based on this information, I was able to establish that in 1919, the mean GE wage was virtually equal to the mean annual wage in US manufacturing. During the 1920s, all US wages rose rapidly, but GE's wages rose faster. In 1923, the mean wage at GE was 8.9% higher than the mean wage in US manufacturing, and the gap rose to 19.3% in 1927, enabling interpolation. Apart from the geometric interpolation of wages for 1924, 1925 and 1926, my wage and employment data for 1920-1939 are reasonably reliable.

Data Sources

Price of Residential Electricity-All Consumption (Figure 1): *Historical Statistics of the United States Colonial Time to 1970* Volume 2, Series S 108-119 page 827. Data for 1889-1890 from Lebergott (1984): The Americans An Economic Record.

Cost of Living Index: David and Solar (1977), Table 1 page 16.

Proportion of Dwellings and Farms with Electric Service (Table 1): *Historical Statistic of the United States Colonial Time to 1970* Volume 2, Series S 108-119 page 827 All dwellings: 108, Farm:109. Data for 1900 from Lebergott (1976), Table 16 page 279.

Proportion of Manufacturing Mechanical Power Capacity Electrified (%): DuBoff (1977) Table 15 Col. 6 page 60.

Electric Utility Real Output Index; Electric Utility Capital Input Index (Figure 1): Kendrick Table H-VI, Pages 590-591.

British Price of Electricity: pounds per million lumen hours: Fouquet and Pearson (2012). Data at the LSE's website or at https://ourworldindata.org/light

Chapter 6

A6 Data adjustments and errors

The Compustat data includes a broader set of firms than those relevant to the study of

market power in the U.S. I limit the universe of firms to those headquartered in the U.S., with stock listed as trading in dollars, and without missing value for total assets. Many of the computations involving monopoly wealth also exclude firms with no market value. I attempt to compute market value, if it is missing, using share price and the number of shares outstanding, but this procedure still leaves market value missing for firms representing about 15% of total assets, mostly in the utilities sector. Market value includes the outstanding value of preferred stock.

A6.1. Date of data. Because I use calendar years as my unit of analysis, while firm annual reports take place on a fiscal year basis, which is a moving target based on the choices of the accountants, it's possible for a firm to have zero or two reports in the same calendar year. In the case of two reports, I take the later one, and in the case of zero reports, I use the report from the first half of the following year.

A6.2. Missing data. Missing values are a persistent problem in the Compustat data. To compute monopoly wealth, I need values for total assets, intangible assets, total liabilities, and market value. I exclude entirely firms with missing values for total assets and market value, but 10% of firms are missing intangible assets and 1% are missing liabilities. In these cases, I first attempt to use the values from a different year for the same firm (the nearest available), and assume that the ratio of liabilities or intangible assets to total assets is constant. This works for most firms; for the remaining, I use year and two-digit industry fixed effects to impute the missing values. There is a risk firms with fewer intangible assets are less likely to report their intangible assets, which could be a small source of bias.

A6.3. Adjustment of asset values for historical cost. An important issue arises from the accounting convention of reporting asset values at depreciated historical cost, which understates their value and would therefore lead to overstatement of monopoly wealth. The problem is less severe for equipment because of its depreciation, but for real estate the problem can be severe since a rise in the market value of real estate leads the value of firms' real estate holdings to be understated by a factor that can exceed 2 in some years. I get that value from the Z.1 Integrated Accounts, which includes the balance sheet of non-financial corporate businesses as part of the national accounts.

The Z.1 Financial Accounts, Table B.103, reports the assets of non-financial corporate businesses in five categories: financial, real estate, equipment, intellectual property, and inventories. For equipment, intellectual property products, and inventories, they report aggregate asset values at historical cost and at current cost. For real estate, they offer the historical cost and the current market value, and for financial assets, no adjustment is made.

I endeavor to match up the Compustat asset categories to these five categories, and the match is reasonably successful. For example, in 2019, I have 51% financial assets, 20% real estate, 15% equipment, 5% inventories, and 9% intellectual property products, compared to their values of 55%, 18%, 14%, 6% and 7%. I chose a slightly different adjustment factors that ensure that the total upward adjustment for the non-financial firms assets is equal to the total adjustment in the Z.1 data, while minimizing the squared distance to the adjustment factors from the Federal Reserve. I am then able to apply these adjustment factors to the assets of the financial firms,

because they report the same asset categories. None of real estate, equipment, and intellectual property make up more 1% of the assets of financial firms, but 4% of their assets are categorized as inventories, largely from real estate transactions in progress. In 2019, this results in an adjustment upward in the asset values of the non-financial firms of 23%, and of 1% for the financial firms.

Note that since most change in prices was due to changed value of real estate holdings, this adjustment ensures that monopoly wealth does not contain any land value. This is so since market price of equity reflects the value of land owned while *current value* of assets contains those same values, hence by (6.1) the two cancel each other in computing monopoly wealth.

A6.4. Data errors. There are two data errors which I note. In aggregating capital employed and wealth created by firms, I add variables that I would have liked to control. First, when aggregating private debt, inter-firm holdings would be cancelled by indebtedness of firms that issued the debt. However, since most debt of firms is owned either by households or by financial institutions, and these are excluded from my aggregation, the effect of ignoring such inter-firm holdings is small. Since monopoly wealth is computed by subtracting aggregate debts from aggregate assets, my inability to account for inter-firm holdings of debt *has no effect on monopoly wealth*. The only effect is a small upward bias in the size of total wealth generated by non-financial firms in my Compustat data. Second, I am unable to estimate the inter-firm holdings of equity within the non-financial Compustat universe, and this component causes an upward bias in the estimated size of capital employed. If firms own securities in firms outside of this universe, their values are correctly recorded. Again, this data error *has no effect on computed monopoly wealth* since any asset added to the balance sheet alters the market value of a firm and cancelled in the estimated monopoly wealth. No doubt some measurement errors remain. Since the same procedure is used in all years 1950-2019, and since I mostly focus on ratios like (monopoly wealth)/(market value), the behavior of proportions over time is a reasonably accurate measure.

A6.5. Definition. Compustat firms included in the study satisfy the following criteria:
(i) firms with headquarters in the US
(ii) non-financial firms: exclude firms with Standard Industrial Classification Code from 6000 to 6499
(iii) firms with positive assets
(iv) firms for which market value can be constructed
(v) firms with a positive number of common shares outstanding (firms in the process of dissolving show up in Compustat with 0 shares outstanding)