The Sources of Growth in a Technologically Progressive Economy: the United States, 1899-1941

<u>ONLINE APPENDICES</u>

Gerben Bakker

(London School of Economics)

Nicholas Crafts

(University of Warwick)

Pieter Woltjer

(University of Groningen)

Reference to main paper: Gerben Bakker, Nicholas Crafts and Pieter Woltjer, 'The Sources of Growth in a Technologically Progressive Economy: the United States, 1899-1941,' *Economic Journal*, forthcoming.

Version 20 August 2018

Appendix A. Sources for Value Added Weights.

Appendix B. Estimates of TFP Growth for Hard-to-Measure Sectors, 1899-1929.

Appendix C. Measurement and Sources for Output and Input, 1929-41.

Appendix D. Labour Quality.

Appendix E. Dual TFP.

Appendix F. Research and Development Inputs by Industry.

JEL Classification: N11, N12, O47, O51.

Keywords: productivity growth; total factor productivity; great inventions; spillovers; United States — history

Acknowledgements: We are grateful to Herman de Jong for making data available to us and to Reitze Gouma for help with the estimation of capital input. We also like to thank Stephen Broadberry, Alexander Field, Robert Inklaar, Nicholas Oulton, Marcel Timmer, three anonymous referees, the editor Hans-Joachim Voth, and seminar audiences at Carlos III University Madrid, London School of Economics, Oxford University, Queen's University Belfast, the University of Birmingham, UC Davis, Stanford University, and the University of Groningen, as well as at conferences in Montreal and Kyoto. Funding for a pilot project was provided by a seed grant from the Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD) of the London School of Economics, as well as a grant from the Netherlands Organization for Scientific Research (NWO Grant no.: 360–53–102). Morten Jerven provided excellent research assistance at this initial stage of the project. The usual disclaimer applies.

Online Appendices:

Appendix A. Sources for Value Added Weights

Three benchmark years have been chosen to calculate value added weights: 1899, 1929, and 1939. From these benchmarks, mid-period weights have been calculated using linear interpolation. Below we discuss the sources for each set of benchmark estimates in detail.

A1. Value Added Weights for 1899

Estimates of value added for electric utilities, farming, metals, mining, non-metals, oil & gas, and manufacturing industries other than those specified below have been taken from Whitney (1968). Value added in foods is the aggregate of 'processed foods' and 'grain mill products' in Whitney (1968). Within manufacturing, for tobacco, non-electric machinery, electric machinery, transport equipment, furniture, and 'miscellaneous' estimates have been taken from the *Census of Manufactures*. Estimates of value added in wholesale & retail trade and in FIRE have been taken from Carter et al. (2006, series Dh1 and Dh2), and estimates of value added in local transit and in railroads from Gallman and Weiss (1969, p.310).

Other sectors all entailed some computation which was implemented as follows:

Anthracite coal and bituminous coal has been estimated using Whitney's (1968) value added for coal mining and the U.S. Bureau of the Census (1960), series M 13-37, p. 350 to apportion the shares of anthracite and bituminous coal mining.

Manufactured gas and natural gas have been estimated using the gross output data in Gould (1946, Table A17), and then using the ratios of value added to gross output for 1919 from Kuznets (1941, p. 659, 661) to arrive at a value-added estimate for 1899.

Construction has been arrived at by calculating gross output from Abramovitz (1964) and then using the average gross output/value added ratio for 1919-1924 from Kuznets (1941, pp. 641-2) to estimate value added.

Residual transport comprised water transportation, pipelines and transportation services. Water transport value added has been taken from Gallman and Weiss (1969, p. 316). For pipelines a rough bench mark estimate has been made for 1885 based on Chandler (1990, p. 74, 94), who stated that Standard Oil's pipeline network was about 4,000 miles. It is assumed that total installed length was double this, i.e., 8,000 miles. This benchmark is then linked to the time series reported in Carter et al. (2006, series Df1246) for 1921-1939 using geometric interpolation, and an estimate for 1899 is made. Real gross output in 1929 from Kendrick (1961, p. 463) relative to pipeline length is then used to estimate real gross output for 1899. The ratio between the 1929 value added of 'Pipelines except natural gas' reported in the *National Income and Product Accounts*, and 1929 gross output for pipelines in Kendrick (1961) is then used to estimate value added for 1899. Value added for transport services has been estimated using the ratio of this to all other transport sectors in 1929 and applying this ratio to the value added of all other transport sectors in 1899.

Telephone is based on the value added of the Bell system companies for 1899, as reported in U.S. Bureau of the Census (1961, p. 481, series R 14-27), multiplied by the inverse of its estimated share in all telephone value added. The latter has been estimated by taking the shares (weighted by local-exchange and long-distance calls) of the Bell companies and the independent companies in 1900, and back-projecting this ratio to 1899 taking into account the differential growth rates of the number of telephones for the two systems.

Telegraph value added comprises the 'International telegraph industry' and the 'Domestic telegraph industry'. The former has been estimated taking operating revenues from Carter et al. (2006, series Dg18), and using the 1907 value added/revenue ratio to estimate 1899 value added. For the latter, 1899 Western Union revenues were taken from Carter et al. (2006, series Dg 16). To arrive at non – Western Union revenues the growth of this category relative to Western Union growth has been calculated for 1902-1907. This ratio has been used to extrapolate 1902 non-Western Union revenues back to 1899. The ratio between gross income and value added in the telephone industry for 1902, as reported in Department of Commerce and Labor, Bureau of the Census, Bulletin 17, *Telephones and Telegraphs, 1902* (1905, p. 31), has then been used to arrive at an estimate for 1899 value added.

Post Office value added is taken as the sum of wages and capital income. The ratio of 1909 wages as reported in King (1930, p. 364) to total revenue as reported in Carter et al. (2006, series Dg 181-9) has been used to estimate 1899 wages based on 1899 revenue from Carter et al. (2006, series Dg 181-9). It has then been assumed that income of remunerated capital was about 0.1 from 1899 revenue.

Value added for spectator entertainment has been calculated by extrapolating the benchmark estimate for 1900 gross output from Bakker (2012), using the growth rate of output over the population growth rate between 1900 and 1909, and multiplying by the average fraction of value added over gross output for live entertainment between 1929 and 1941. The latter has been estimated from the NIPA by using the share of live entertainment expenditure in all 'Amusements and recreation except motion pictures' expenditure.

Government value added has been estimated using John Wallis' (2006) estimate for 1902 of government expenditure, compensation of government employees and net interest paid by government and government surplus or deficit, as no estimate for 1899 itself was available (John Joseph Wallis, "Total government expenditure, by character and object: 1902–1995" in Carter et al. (2006), series Ea 14, 52, 53 and 59). All these values were then extrapolated back to 1899 using GDP-growth from Johnston and Williamson (2017), to arrive at an estimate of government value added for 1899. Then the value added of the Post Office for 1899 (see Appendices A1 and B3) was subtracted, to arrive at the government value added used in this paper to estimate the size of the private domestic economy. As a further coherence check of this estimate, the average ratio of government gross fixed capital consumption to government total labour compensation for 1929-1941 was taken from the National Income and Product Accounts, and it has been assumed that a similar ratio existed in 1902 and in the estimated 1899, which yielded an estimate of 1899 government value added as 3.0% of GDP, which was just marginally lower than the original estimate above of 3.1% of GDP, and gives further confidence in that (latter) estimate. In any case, the findings of this paper are not very sensitive to the difference between the two estimates.

A2. Value Added Weights for 1929

Value added for most sectors has been obtained directly from the National Income and Product Accounts as reported in Appendix C. Below we discuss the sectors for which value added could not be obtained directly from this source.

FIRE and wholesale & retail trade have been taken from Carter et al. (2006, series Dh2 and Dh25-6). Construction has been calculated from Abramovitz (1964). Manufactured gas has been calculated from Kuznets (1941, pp. 659-676).

The weight of electric utilities starts from the quantity of electricity sold and its price as reported in Bureau of the Census (1960, series S80-1) to arrive at the value of gross output for 1929. This value was then multiplied by the value added / gross output ratio calculated for 'electric light & power and manufactured gas' for 1929 from Kuznets (1941, pp. 659-676).

Natural gas is based on gross output from Gould (1946) multiplied by the ratio of value added to gross output for petroleum and natural gas together for 1929 from Leontief (1953).

Telephone value added has been calculated from Kuznets (1941, pp. 659-670). A second estimate has been made by using the method for calculating telephone value added outlined for 1939 in section 1.3 below. This estimate was slightly less reliable as for the 'Independent Telephone Companies' the geometric interpolation of 'wages and salaries' versus the use of relative factor incomes to estimate 'wages and salaries' yielded two estimates for value added for 'Independent Telephone Companies' that were 33% apart. Using the average of these two estimates yields a total estimate of value added for the entire telephone industry that is only one percent higher than the Kuznets estimate. The latter value has therefore been taken.

For the 'Domestic Telegraph Industry' the value of intermediate inputs has been calculated by subtracting 'wages and salaries' from 'operating expenses' taken from Bureau of the Census (1960, p. 484-5, series R53-67). These 'operating expenses' do not include 'net income' and 'federal income tax', so must be approximately the sum of wages and salaries plus intermediate inputs. Capital income is then taken to be 'operating revenues' minus 'operating expenses'. A similar estimate has been made for the ''International Telegraph Industry' from Bureau of the Census (1960, p. 485-6, series R72-85).

A second estimate for Telegraph was made based on gross output for the domestic and international telegraph industries from the Bureau of the Census (1961, p. 484) multiplied by the ratio of value added to gross output from Kuznets (1941, pp. 659-670). This estimate yields a value that is 2.7% higher. As the former estimate is more precise, that estimate has been taken.

Post Office value added is taken as the sum of wages and capital income. Wages were estimated by taking the ratio of the compensation of postmasters as reported in the *Annual Report of the Postmaster General* (1970, pp. 138-141) to all wages as reported in King (1930, p. 364) for 1925, and establishing that this ratio was fairly stable for 1923-25 (fluctuating between 10.0 and 11.1% of all wages), and then using this ratio to estimate all wages for 1929 based on postmasters' compensation in 1929. It has then been assumed that income of remunerated capital was about 0.1 from 1929 revenue.

For Spectator Entertainment the value added of motion pictures is taken directly from the National Income and Product Accounts. The value added of live entertainment has been estimated by using the share of live entertainment expenditure in all 'Amusement and recreation, except motion pictures' expenditure.

For the government sector, from the National Income and Product Accounts the 'Compensation of employees, general government, federal, state and local' (excluding 'Government enterprises', which are partially in our sector 'Post Office' and partially in our Residual Sector) and the 'Government consumption of fixed capital', excluding the lines for 'Government enterprises' have been taken to calculate government value added.

A3. Value Added Weights for 1939

Value added for most sectors has been obtained directly from the National Income and Product Accounts as reported in Appendix C. Below we discuss the sectors for which the value added could not be obtained directly from this source.

FIRE and wholesale & retail trade have been taken from Carter et al. (2006, series Dh27 and Dh25-6).

Construction value added has been calculated from Abramovitz (1964).

The weight of electric utilities starts from the quantity of electricity sold and its price as reported in Bureau of the Census (1960, series S80-1) to arrive at the value of gross output. This value was then multiplied by the value added / gross output ratio calculated for 'electric light & power and manufactured gas' for 1939 from Kuznets (1941, pp. 659-676).

Manufactured gas is based the estimate for 1938 in Kuznets (1941) extrapolated to 1939.

Natural gas is based on gross output taken from Gould (1946) multiplied by the ratio of value added to gross output for petroleum and natural gas together for 1939 from Leontief (1953).

For the Bell Telephone Companies, the value of intermediate inputs has been calculated by subtracting 'wages and salaries' from 'operating expenses' taken from Bureau of the Census (1960, p. 481, series R14-27). These 'operating expenses' do not include 'interest expenses' and 'federal income tax', so must be approximately the sum of wages and salaries plus intermediate inputs. Capital income is then taken to be 'operating revenues' minus 'operating expenses' plus 'income from Western Electric Co.', which equals the sum of 'interest expenses', 'federal income tax' and 'net income' (Bureau of the Census 1960: 481, series R14-27). A similar estimate has been made for the 'Independent Telephone Companies' from Bureau of the Census (1960, p. 483, series R28-42), with the difference that 'wages and salaries' are not available for 1939 and had to be estimated. We made two estimates: one using the growth rate of 'wages and salaries' between 1934 and 1941 and estimating a 1939 value by geometric interpolation, and one by using the ratio of the factor incomes for the Bell Telephone Companies for 1939. The estimates for value added using these two different 'wages and salaries' estimates differ by only 1.2%, and the average has been taken. Total value added is then the sum of these estimates.

A second estimate has been made using the value-added estimate in Kuznets (1941, pp. 659-676) for 1938 and multiplying it by the growth rate of operating revenue between 1938 and 1939 for Bell and independent telephone companies taken from Bureau of the Census (1960, p. 481-3, series R14-27 and R28-42). This yields a value added that is 4.8% lower than the above estimate. As the second estimate is an extrapolation and based on less information from the actual year (1939), the first estimate has been taken.

For the 'Domestic Telegraph Industry' the value of intermediate inputs has been calculated by subtracting 'wages and salaries' from 'operating expenses' taken from Bureau of the Census (1960, p. 484-5, series R53-67). These 'operating expenses' do not include 'net income' and 'federal income tax', so must be approximately the sum of wages and salaries plus intermediate inputs. Capital income is then taken to be 'operating revenues' minus 'operating expenses'. A similar estimate has been made for the 'International Telegraph Industry' from Bureau of the Census (1960, p. 485-6, series R72-85).

A second estimate was made with the second method used for 'Telephone', using the weighted growth rate of operating revenues of the domestic telegraph industry and the international telegraph industry between 1938 and 1939 from Bureau of the Census (1960, p. 484) to extrapolate Kuznets's value added from 1938 to 1939. Both estimates are very close: the first estimate is only 2.6% higher than the second. Given that this first estimate is based on data from the year itself and not on extrapolations, this first estimate has been taken.

Post Office value added is taken as the sum of wages and capital income. Wages were estimated by taking the ratio of the compensation of postmasters as reported in the *Annual Report of the Postmaster General* (1970, pp. 138-141) to all wages as reported in King (1930, p. 364) for 1925, and establishing that this ratio was fairly stable for 1923-25 (fluctuating between 10.0 and 11.1% of all wages), and then using this ratio to estimate all wages for 1939 based on postmasters' compensation in 1939. It has then been assumed that income of remunerated capital was about 0.1 from 1939 revenue.

Spectator Entertainment value added has been estimated using the same method and sources as for 1929.

The value added for the government sector has been estimated using the same method and sources as for 1929.

Table A1. *Industry Value Added as percentage of the PDE, United States, 1899-1941.*

Industry	Value added (percentage of the PDE)				
	1899-1919	1909-1919	1919-1929	1929-1941	1899-1941
Farming	13.1	11.4	9.8	8.2	10.5
Metals	0.6	0.6	0.6	0.5	0.6
Anthracite Coal	0.3	0.3	0.3	0.2	0.3
Bituminous Coal	0.7	0.7	0.7	0.7	0.7
Oil and Gas	0.4	0.5	0.7	0.9	0.6
Non-metals	0.4	0.3	0.3	0.2	0.3
Foods*	4.1	3.7	3.3	3.5	3.6
Tobacco	0.9	0.9	0.9	1.2	1.0
Textiles	1.9	1.9	1.9	1.9	1.9
Apparel	1.7	1.5	1.4	1.3	1.4
Leather Products	0.9	0.8	0.6	0.6	0.7
Lumber Products	2.3	1.8	1.3	0.8	1.5
Paper	0.5	0.5	0.6	0.7	0.6
Printing Publishing	1.4	1.6	1.8	1.7	1.6
Chemicals	1.0	1.1	1.2	1.4	1.2
Petroleum, Coal Products	0.6	0.9	1.3	1.3	1.0
Rubber Products	0.2	0.3	0.4	0.4	0.3
Stone, clay, glass	0.7	0.8	0.9	0.9	0.8
Primary Metals	2.7	2.6	2.5	2.4	2.5
Fabricated Metals	1.0	1.3	1.6	1.6	1.4
Machinery Non-Electric	2.3	2.2	2.1	2.0	2.2
Electric Machinery	0.4	0.7	0.9	1.1	0.8
Transport Equipment	1.0	1.4	1.9	2.1	1.6
Furniture	0.5	0.6	0.7	0.7	0.6
Miscellaneous	0.9	0.8	0.7	0.7	0.8
Electric Utilities	0.5	1.2	1.8	2.2	1.4
Manufactured Gas	0.1	0.2	0.2	0.2	0.2
Natural Gas	0.1	0.2	0.2	0.3	0.2
Construction*	5.5	4.9	4.4	3.4	4.5
Wholesale & retail trade*	13.9	13.9	14.0	14.6	14.1
Railroads	6.4	6.1	5.8	4.6	5.7
Local Transit	1.1	1.1	1.0	0.9	1.0
Residual Transport	1.1	1.3	1.4	1.8	1.4
Telephone	0.3	0.5	0.7	0.9	0.6
Telegraph	0.1	0.1	0.2	0.1	0.1
Post Office*	0.5	0.5	0.6	0.6	0.5
FIRE*	4.7	8.2	11.7	11.6	9.2
Spectator Entertainment*	0.4	0.5	0.5	0.6	0.5
Manufacturing	24.9	25.4	25.9	26.2	25.6
Great inventions	19.8	22.1	24.4	26.3	23.3
Aggregate measured sectors	75.1	77.9	80.7	78.8	78.2
Government sector	3.5	4.2	4.8	8.2	5.3
Residual sector	24.9	22.1	19.3	21.2	21.8

Notes: benchmark estimates were made for 1899, 1929 and 1939 based on original sources. The period values were then estimated by linear adjacent-year weighting using mid-interval years, for example: 1899-1909 is 25/30 of the 1899 weight and 5/30 of the 1929 weight; 1909-1919 is the average of the 1899 and 1929 weights, and 1919-1929 is 5/30 of the 1899 weight and 25/30 of the 1929 weight.

* = sector measured in this paper but not by Kendrick. FIRE = Finance, insurance & real estate.

Appendix B. Estimates of TFP Growth for Hard-to-Measure Sectors, 1899-1929

B.1. Construction

Kendrick (1961, pp. 489-498) found that capital was a very small production factor in the construction sector, and therefore he only provided labour productivity estimates. From 1970, precise capital income shares of the U.S. construction sector are available from the EU KLEMS dataset (*EUKLEMS database*, November 2009 release, revised June 2010) which report a very small capital income share of 0.1 of value added in 1970. Abramovitz (1964) also suggests that capital was relatively unimportant in this period. Accordingly, for 1899-1929, we have taken labour productivity growth rates from Kendrick (1961, p. 498) to proxy crude TFP growth for 1899-1909, 1909-1919, and 1919-1929. Crude TFP for the periods to 1929 is adjusted by subtracting labour quality growth (see Table 3 in the main text and Appendix D) to arrive at refined TFP growth.

We have checked this estimate using indices of construction output from Abramovitz (1964, pp. 142-4, Table A1, series 6 for 1899-1915 and series 3 for the period after) and indices of labour input (Ibid., p. 125) to calculate output per person-hour growth (and thus TFP growth) for the period 1899-1929. The resulting growth rate for the 30-year period from 1899-1929 is virtually the same as is obtained from Kendrick's growth rates for the three sub periods (1899-1909, 1909-1919 and 1919-1929).

B2. Wholesale and Retail Trade

Kendrick (1961, pp. 499-506) found that capital was also a very small production factor in the wholesale and retail trade sector and therefore he only provided labour productivity growth estimates prior to 1929. He estimated that the capital income share was about 0.13 in both 1937-1948 and 1948-1953, and considerably less than 0.13 in 1929-1937 (Kendrick 1961, p. 505). Kendrick also estimated that about half of all capital stock in 1929 consisted of inventories (1961, p. 504). Accordingly, for 1899-1929, we have taken the labour productivity growth rates from Kendrick (1961, p. 506) to proxy crude TFP growth for 1899-1909, 1909-1919 and 1919-1929. Crude TFP is then adjusted by subtracting labour quality growth (see Table 3 in the main text and Appendix D) to arrive at refined TFP growth.

B3. Post Office

The growth rates of output and person-hours have been calculated from Kendrick (1961, p. 611), who following established convention, included the Post Office, and government enterprises in general, in the PDE. The growth rate of capital is based on the number of first, second, and third-class post offices in existence at the benchmark years, taken from the *Reports of the Postmaster General* (Washington, D.C., various years). Given that between 1909 and 1925 the share of wages in total revenues ranged between 0.62 and 0.80, as reported in King (1930, p. 364), and given that a substantial part of the capital used consisted of the use of government buildings, the income share of total capital (remunerated capital and unremunerated use of government buildings) has been set at 0.4. Using these assumptions, TFP growth rates have been estimated for 1899-1909, 1909-1919 and 1919-1929. Crude TFP is then adjusted by subtracting the labour quality growth rates for the Post Office reported in Table 3 and Appendix B to arrive at refined TFP growth.

B4. Finance, Insurance and Real Estate (FIRE)

Crude TFP growth has been estimated for 1899-1909, 1909-1919, and 1919-1929 as follows. The growth rate of output has been estimated by creating an index consisting of one quarter of the output growth of financial intermediation, one quarter of the growth of life insurance policies, and one half of the growth of rent. These weights follow those reported for the FIRE sector in Central Statistical Office (1956, pp. 364-5).

The output growth of financial intermediation has been taken from Philippon (2015, appendix). The index used consists of the weighted average of Philippon's level index and 8.48 times Philippon's flow index, using Philippon's scaling factor (8.48) to make the two components comparable. Philippon's series have been taken from the file 'Data Series' for his article published on his website http://pages.stern.nyu.edu/~tphilipp/research.htm (accessed on 8 December 2014).

The value of life policies has been taken from Carter et al. (2006), series Cj715, deflated to real values by using the Bureau of Labor Statistics based Consumer Price Index as reported in Carter et al. (2006), series Cc1. Consumer expenditure on rent was taken from Lebergott (1996), Table A2 and deflated to real values using the residential rents index reported in Carter et al. (2006), series Cc4. The growth rate of labour inputs is based on the number of person-hours reported for 'finance, insurance, real estate' by Kendrick (1961, p. 314). The growth rate of the capital stock is based on FIRE tangible assets in 'Banking, 'Insurance' and 'Miscellaneous' reported in Goldsmith (1958), Tables A1, A7 and A15. Factor shares were estimated from Goldsmith (1958) based on 'employee compensation' and 'non-farm proprietors' compensation' for labour and on 'corporate profits, pre-Tax' and 'net interest' for capital, resulting in a 0.5 share for both factors. Using the EU KLEMS data for 1970 for labour and capital compensation in financial institutions and in real estate (*EUKLEMS database*, November 2009 release, revised June 2010), and weighing each sector by a half, also yields factor incomes of 0.5 each. These rates of crude TFP-growth have then been adjusted by subtracting the labour quality growth rates for the FIRE sector reported in Table 3 and Appendix D to arrive at refined TFP growth.

B5. Spectator entertainment

TFP-growth for spectator entertainment has been estimated following the methods and sources set out in Bakker (2012) for 1900 and 1938, extending the estimates to 1899 and 1941, and now including the three intermediate benchmark years, 1909, 1919 and 1929. The output estimates are based on the consumer expenditure series from Carter et al. (2006, series Dh311 and Dh312), the National Income and Product Accounts, the Bureau of Labor Statistics Admission Price Index, and several other price studies for the period before 1929. The labour estimates are based on the census using the same method as outlined for 1900 in Bakker (2012). Labour quality has been estimated using the method outlined in Appendix D. The capital estimates are based on Bakker (2012), extrapolating the 1900 estimate by one year using a composite growth index of capital proxies to arrive at an 1899 estimate, using the method outlined in Bakker (2012) for 1900 to arrive at the 1909 estimate, and for 1938 to arrive at a 1941 estimate. For motion pictures the annual investments in new cinemas have been used to make an estimate for 1929 based on the 1941 estimate. The 1919 estimate has been interpolated from the 1909 and 1929 estimates using the growth in aggregate cinema seating capacity and the capital per seat deflated by Kuznets' (1961) capital goods deflator. Live capital for 1929, and for 1926 (a year before talking pictures arrived) has been estimated assuming capital grew at one fourth the rate of output (given the small share of live capital by this time, the findings are not very sensitive to this assumption). Live capital for 1919 has been estimated by geometrically interpolating the 1909 and 1926 values. For 1929-1941 the resulting stock-based capital growth rate has been modified to a capital services-based capital growth rate by multiplying by the weighted difference between stock- and service-based capital growth estimates of the 'Motion picture and sound recording industries' and 'Performing arts, spectator sports, museums, and related activities' from the Bureau of Economic Analysis' fixed assets table (see Appendix C). A detailed statistical survey from 1909 of all Boston entertainment venues, reported by Jowett (1974), was used to assess the relative importance of live and filmed entertainment at that time.

B6. Foods

We needed to make an adjustment for the Food sector in order to combine Kendrick's (1961) separate series for crude 'Food' TFP-growth and for crude 'Beverages' TFP-growth into one aggregate series for crude Food TFP-growth for 1899-1929 which is comparable with the NIPA data for 1929-1941.

From Fabricant (1940: 608-610), the weights for Beverages value added as share of total value added for Food for 1899, 1909, 1919 and 1929 have been taken to compute average weights for 1899-1909, 1909-1919 and 1919-1929. These weights have then been used to merge Kendrick's separate series of crude TFP-growth. From the resulting rate then has been subtracted the labour quality growth in Food as reported in Table 3 to arrive at refined TFP-growth for Food for 1899-1929 that is comparable with the NIPA data for 1929-1941.

It should be noted that the resulting TFP-growth rate for Food masks two very different trends for 'Food' compared to 'Beverages' that was visible in the Kendrick crude TFP-growth rates, and this may have been the reason why Kendrick found it necessary to list the sectors separately. The disaggregated data show that TFP-growth in Beverages was massively negative (-5.6% per annum) in the 1910s, almost nil (-0.2% per annum) in the 1920s and massively positive in Kendrick's 1930s (1929-1937) with 15.2% per annum, while crude TFP for Food shrank with only 0.4% per annum in the 1910s, grew with 5.3% per annum in the 1920s, and grew with only 1.5% per annum in the latter period (Kendrick 1961: 136). Our Foods sector, combining Kendrick's 'Food' and 'Beverages', was by far the largest manufacturing sector in all sub periods, ranging from 1.2 - 1.6 times the next largest manufacturing sector. The end of Prohibition coincided with a massive 15.2% per annum growth of Beverages crude TFP between 1929 and 1937, and this clearly had a large pull on manufacturing TFP-growth as a whole.

Appendix C. Measurement and Sources for Output and Input, 1929-41

As emphasized by Field (2003), the assessment of productivity trends during the 1930s is highly sensitive to the choice of beginning- and end-point. In order to prevent cyclical effects from influencing the measurement of productivity growth it is best to choose business-cycle peaks as reference years. Kendrick's (1961) choice of comparing the depressed American economy in 1937 to the peak-year of 1929 conflicts with this principle. Field (2003: 1403) argues instead that 1941, with an unemployment rate of 9.9%, compares much more favourably to the fully employed economy of 1929 than the year 1937 (14.3% unemployment). Regrettably, little productivity data is available – at least not beyond the total economy trends – for the early 1940s. This led Field (2006) to restrict his analysis of technological change between 1929 and 1941 to a 4-sector breakdown of TFP growth. As noted in section 1, however, we require a much finer breakdown in order to fully decompose the sectoral contribution to TFP and labour productivity growth. This appendix describes the methods and sources which we use to develop new, industry level estimates extending beyond Kendrick's original 1929-37 figures. These new estimates allow us to measure productivity growth over the period suggested by Field, 1929-41, while matching the full sectoral detail realized by Kendrick.

Output

Instead of estimating value added on the basis of industry output less purchases of materials and services, we obtain nominal gross value added by summing over total compensation, gross operating surplus, and taxes on production less subsidies. The components of value added at the industry level are compiled by the U.S. Bureau of Economic Analysis (BEA, 2009) and listed in the *National Income and Product Accounts* (NIPA). Table C1 provides an overview of the relevant variables, the exact source-tables, the number of industries differentiated, as well as the share of value added covered by each respective variable in the year 1947.

The NIPA tables provide annual data at the industry level, allowing us to estimate net output for a set of 38 (disaggregate) industries, completely covering the domestic economy. As illustrated in table C1, the NIPA provides full industry coverage for the most influential variables which, together, make up over 80% of gross value added. For the remaining variables the BEA supplies data at a higher level of aggregation, distinguishing between either 12 separate industries (e.g. proprietors' income) or listing the total-economy value only (e.g. taxes on production less subsidies). For these variables, we use the detailed industry-level data for the components of value added in 1947 – taken from the BEA's (2011) *Historical Industry Accounts Data* – to distribute the aggregate figures over our complete list of industries.

To obtain real value added we deflate the nominal output figures for agricultural, mining, manufacturing, utilities and wholesale trade on the basis of wholesale prices compiled by the U.S. Bureau of Labor Statistics (1943: 4; 1949: 6; 1958: 26, 34) supplemented with the production prices listed in the *Historical Statistics of the United States* (HSUS 1975: 582-6) and the price index of electrical equipment compiled by the BEA (2010). For the remaining service sectors, we apply the relevant price indices for personal consumption expenditure from the NIPA (BEA 1966: table 8.6; BEA 2009: table 1.5.4) and Kendrick (1961: 543-5, 556, 583-4). We aggregate the price deflators over industries on the basis of an annually chained Fisher index, where nominal gross value added, previously discussed, serves as weights.

Table C1. *Components of Value Added by Industry, United States, 1929-1941.*

Variable	Description	Source [a]	Cov. [b]	Shr. [c] (%)
VA	Value added, by industry			100
COMP	Compensation of employees, by industry	NIPA, table 6.2A	38	54
TXPIXS	Taxes on production and imports less subsidies	NIPA, table 1.7.5 line 18	1	7
GOS	Gross operating surplus, by industry			
NINT	Net interest, by industry	NIPA, table 6.15A	12	1
PROINC	Proprietors' income, by industry (nonfarm)	NIPA, table 6.12A	12	9
FRMINC	Proprietors' income, farm (with IVA and CCadj)	NIPA, table 2.1 line 10	1	6
PBT	Corporate profits before tax, by industry	NIPA, table 6.17A	38	13
CCCA	Corporate capital consumption allowance, by industry	NIPA, table 6.22A	38	3
NCCA	Non-corporate capital consumption allowance, by industry	NIPA, table 6.13A	12	3
BCTP	Business current transfer payments	NIPA, table 7.7 line 1	1	0
IVA	Inventory valuation adjustment, by industry (nonfarm)	NIPA, table 6.14A	12	-3
CCadj	Capital consumption adjustment, by industry (nonfarm)	NIPA, table 7.6	1	0
GCFC	Consumption of fixed capital, government	NIPA, table 7.5 line 21	1	4
RIP	Rental income of persons, FIRE	NIPA, table 2.1 line 11	1	3

[a] *Source*: BEA (2009).
[b] Number of separate industries distinguished in the original source. Note that full coverage corresponds to 38.
[c] Share of total economy value added covered in 1947. Source: BEA (2011).

Labour input

For labour input we rely on estimates of total employment by industry, fully compensated for changes in the average annual hours of work and the growth in the quality of labour. The sources for total employment and the average hours of work are discussed below. The adjustment for labour quality is dealt with in appendix D.

In correspondence with Kendrick (1961: 47-9), we define total employment as the sum of the number of employees, converted to a full-time equivalent basis, and self-employed persons. From 1929 onwards, the NIPA (BEA 2009: table 6.8A) lists this statistic as the total Persons Engaged in Production (PEP) at the detailed industry level.

Estimates for the average annual hours of work between 1929 and 1941 for the majority of industries are based on Kendrick (1961: 310, 360-2, 397-8, 543-7, 556, 583-4, 590-8, 611). For construction, other transportation and trade we rely on the HSUS (1975: 170-3) estimates of changes in the weekly hours of work. In addition, we accounted for differences in the average hours of work in durable and nondurable manufacturing based on data from the HSUS (1975: 169-70), normalized to fit Kendrick's (1961: 465-6) total manufacturing estimates. Our final measure of labour input is then derived by multiplying total employment (PEP) by both the index for the change in the average annual hours of work as well as the index for labour quality.

Capital input

For the period 1929 to 1941 we estimate the capital input on the basis of capital services. As opposed to capital stocks, which measure the total value, or wealth of all capital equipment and structures in place, our measure captures the capital service *flows* derived from these capital assets. The difference between both these methods is that capital services weight the growth of capital assets by their respective rental prices, whereas capital stocks weight assets by their asset price. As noted by Jorgenson et al. (2008: 109), "[c]apital input takes the form of services of the capital stock in the same way that labour input involves the services of the work force", making the resulting capital service indices strictly comparable to the measure of labour input discussed above.

In comparison to the stock measure, the capital service flows will allocate greater weight to assets that have shorter asset lifetimes and/or rapidly falling asset prices, as both of these characteristics will drive up the cost a user would have to pay to hire the asset for a given period. In the 1930s, prime examples of assets that are underweighted by the traditional capital stock measure are: communication equipment, instruments and trucks.

Our capital services differ from the measure of capital adopted by Kendrick (1961) but is consistent with the post-war estimates of capital input by the BLS. This thus allows us to directly compare the 1929-1941 residual in our growth accounting exercise to the official TFP estimates for the decades following the war.

The construction of the indices of capital services proceeds in two phases. First, we estimate the industry-level stock of capital for the private domestic economy between 1929 and 1941 using a Perpetual Inventory Method (PIM) and the investment series taken from the BEA's *Fixed Assets* tables. Second, we estimate the rental price of assets at the industry level based on the imputed industry rate of return, the asset-specific rate of depreciation and capital gains and losses resulting from changing asset prices. Multiplying the stock of an asset by its rental price yields so called 'capital compensation', which in turn can be used as weights to aggregate the capital stocks to the industry and ultimately the total economy level.

For the construction of the capital stocks, we follow the approach set out by the BEA (2003: M-7), where the real investment ($I_{i,k}$) for asset *k* during year *i* is assumed to contribute $N_{i,k,t}$ to the real net stock of capital at the end of year *t*.

$$N_{i,k,t} = I_{i,k}\left(1 - \frac{\delta_k}{2}\right)(1 - \delta_k)^{t-i} \quad \text{, where } t \geq i \tag{C.1}$$

All investments are expected to have been made during the middle of the calendar year and are depreciated at an annual geometric rate of depreciation δ_k. By summing the contributions over all investments up to and including year *t*, the real net stock of capital ($N_{k,t}$) for asset *k* at the end of year *t* can be derived.

$$N_{k,t} = \sum_{i=1}^{t} N_{i,k,t} \tag{C.2}$$

From 1901 onwards, the BEA's (2010) detailed *Fixed Assets* tables provide annual industry-by-asset investment series for private nonresidential capital. To reliably estimate the starting stock of capital in 1900, we supplement this data with the asset-specific constant-cost investment series for the period 1832-1900, listed in the BEA's (1993: 374-81) *Fixed Reproducible Tangible Wealth* report. Unfortunately, the pre-1901 investment series is only available at the total private economy level. We thus distribute the nineteenth century investment data for each of the 37 assets over our entire industries-list on the basis of the average investment shares for the first decade in the twentieth century – for which we have detailed industry-by-asset data. The geometric rates of depreciation for all our assets, with the exception of automobiles, are taken from Fraumeni (1997). The rate of depreciation for autos was derived implicitly from the standard *Fixed Assets* tables (BEA 2010, tables 2.2, 2.8).

On the basis of these investment series, depreciation estimates and equation (C.2) we compile the real net stock of capital between 1929 and 1941 for all assets and industries distinguished by the BEA (with the exception of the government sector). Capital services (*K*) for industry *j* can then be derived by weighting the growth of capital stocks for all *m* assets by its relative share in total industries capital compensations (φ). Dropping the industry subscript *j* for ease of notation, the growth of capital services can be represented as

$$\hat{K} = \sum_{k=1}^{m} \bar{\varphi}_k \hat{N}_k \tag{C.3}$$

Where hats indicate growth rates in natural logs and $\bar{\varphi}_k$ is the average share of capital compensation in year *t* and *t-1* for asset *k*

$$\overline{\varphi}_k = \frac{1}{2}\left(\varphi_{k,t} + \varphi_{k,t-1}\right) \tag{C.4}$$

As noted previously, capital compensation is the product of the rental price ($p_{k,t}^K$) and the real stock ($N_{k,t}$) of this asset. The share (φ_k) is then calculated by dividing the assets capital compensation by the total industry's capital compensation. Note that industry *j's* capital compensation can be obtained from the national accounts as gross operating surplus (GOS) minus the sum of noncorporate income not allocated to labour (see table C1).

$$\varphi_{k,t} = \frac{p_{k,t}^K N_{k,t}}{\sum_{k=1}^m p_{k,t}^K N_{k,t}} \tag{C.5}$$

The calculation of the rental price reflects the fact that in equilibrium, an investor is indifferent between two alternatives: either buying a unit of capital at time *t*-1, collecting a rental fee and then selling the depreciated asset in the next period, or earning a nominal rate of return on a different investment opportunity. The capital services thus depend on the asset-specific depreciation rates (δ_k), the (industry-specific) rate of return (i_t) and the capital gains or losses from changes in the asset-specific investment price ($\hat{p}_k^I$).[1]

$$p_{k,t}^K = p_{k,t-1}^I i_t + p_{k,t}^I \delta_k - \frac{1}{2}\left(\widehat{p}_{k,t-1}^I + \widehat{p}_{k,t}^I\right) p_{k,t-1}^I \tag{C.6}$$

For the calculation of the industry rate of return we follow the ex-post procedure preferred by the BLS to make our capital service estimates comparable to the post-war figures. The rate of return is the sum of total capital compensation and the total capital gains from changes in investment prices, minus total depreciation, divided by the capital stock in prices of year *t-1*.

$$i_t = \sum_{k=1}^m \left(p_{k,t}^K N_{k,t} - p_{k,t}^I \delta_k N_{k,t} + \frac{1}{2}\left[\widehat{p}_{k,t-1}^I + \widehat{p}_{k,t}^I\right] p_{k,t-1}^I N_{k,t}\right) \Big/ \sum_{k=1}^m p_{k,t-1}^K N_{k,t} \tag{C.7}$$

For the estimation of the rental prices we again rely on Fraumeni's (1997) depreciation rates, the BEA's (2010) price index of investment and the industry-level capital compensation from the NIPA tables (BEA 2009).

Table C2 shows the difference between Kendrick's original capital input measures, the average annual growth of the capital stock measured using the BEA's investment series and the growth in capital services. Kendrick's estimates are very similar to the growth figures for the capital stocks but differ substantially from the estimates based on capital services. As previously noted, capital service flows will allocate greater weight to assets that have shorter asset lifetimes (e.g. instruments, machinery and trucks), the stock of which expanded more rapidly than for long-lived assets (i.e. structures and long-lived equipment such as ships and train rolling stock) during the 1930s. This explains why the growth figures for capital services exceed the capital stock-based measures for both the PDE and PNE as well as most of the underlying industries during the years 1929-1941.

[1] In equations (C.6) and (C.7) we rely on the two-period average change in the asset-specific investment prices to smooth out incidental price shocks. The subscript *t* refers to growth between year t-1 and t.

Table C2. *Average Annual Rates of Growth of Capital Input, United States, 1929-1941, in percent per annum.*

Economic Aggregate	Annual growth (% per annum)
Private domestic economy (PDE)	
Kendrick	-0.08
Capital stocks	-0.09
Capital services	0.37
Private domestic nonfarm economy (PNE)	
Kendrick	-0.13
Capital stocks	-0.05
Capital services	0.48

Source: Kendrick (1961), pp. 333-335; 338-340.

Table C3 shows the impact that the different measures of capital input have on total factor productivity. With the exception of the three manufacturing industries Petroleum, Coal Products, Electric and Non-Electric Machinery, as well as Residual transport, Telephone and the Post Office, industries' capital-service based TFP-growth was lower or equal than the stock-based estimates (first three columns). The difference in TFP-growth ranged from minus 0.5 percentage points for Metal mining, to plus 0.1 percentage points for the six industries mentioned above. Of the aggregate growth rates, only the residual sector (minus 0.4 percentage points), and the PDE-growth rate (minus 0.1 percentage points) were affected. The mean industry TFP-growth decreased by 0.03 percentage point. The coefficient of variation and the range only increased marginally.

The capital services-based intensive growth contribution (IGC) showed a similar pattern (Table C4, first three columns), with the same six industries having positive differences and the minimum value being -0.015% per annum for Wholesale & retail trade. The aggregate measured sectors' IGC decreased with 0.029%, the residual sector with minus 0.083%, and the PDE with 0.113%. The mean and coefficient of variation decreased marginally, and the range decreased by 0.013%. Overall, the use of capital-services based TFP-growth and IGC showed a small but not insignificant difference with capital-stock based TFP growth.

For the whole period 1899-1941, using stock-based estimates for 1899-1929 and service-based estimates for 1929-1941, the differences in TFP-growth were very small, and differed in only two sectors from 0.0, plus 0.1 in both cases. Likewise, for the IGC there were only significant differences (in the third decimal) for five industries.[2]

Variable retirement

Gordon (2016: 659-663) argues that the official investment and depreciation rates from the BEA severely underestimate the growth in capital input for the period between 1925 and 1945. In particular, he questions whether the depreciation rates, which are fixed over time, are representative for the Depression era. As investment collapsed, Gordon would expect equipment and structures to be scrapped and depreciated at a slower rate; i.e. he proposes that the expected life-time of all assets should increase substantially during the 1930s. This lower rate of depreciation would lead to a greater increase in the capital stock than the estimates by the BEA would suggest.

As a crude proxy for these varying rates of depreciation, Gordon suggests comparing the ratio of investment to the official capital stock for each year with the average for 1925-1972. A low ratio, as was the case for 1933, would indicate an increase in the asset lifetime, whereas a relatively high ratio would indicate a reduction in the time producers hold on to their ageing capital assets.

[2] The tabulated results are available from the authors.

As a robustness check to our capital input estimates, we apply the same procedure. We estimated the ratio of investment to stock separately for equipment capital and structures. The data was taken from the BEA's (2010) *Fixed Assets*, basic tables 2.1, 2.2, 2.7 and 2.8, lines 3 and 37. The official depreciation rates discussed in the previous section were multiplied by the ratio of investment to capital in the respective year, relative to the average of the period 1925-1972. On the basis of these depreciation rates we re-estimated the capital stocks and capital services. Tables C3 and C4 show the resulting TFP and IGC based on these revised capital inputs.

The effect of the adjustment on industry TFP-growth for 1929-1941 was negligible or significantly negative for all industries. It varied substantially, from 0.0 percentage-points for Coal Mining, Residual Transport and Spectator Entertainment, to minus 0.7 percentage-points for Oil and Gas Mining and minus 0.5 percentage-points for Metal Mining, Tobacco, and Petroleum & Coal products. The overall effect on the TFP-growth rate of the PDE was substantial, minus 0.2 percentage-points, which constitutes a ten percent downward adjustment. The impact on the intensive growth contribution (IGC) was small but significant for many industries, as only 6 industries had no difference at three decimals, and the change in the IGC was substantial for Farming, Wholesale & Retail trade and FIRE—all large sectors. For both TFP and IGC, the mean decreased and the coefficient of variation increased by about ten percent.

For the whole period 1899-1941, using standard depreciation-based estimates for 1899-1929 and variable retirement-based estimates for 1929-1941, the differences in TFP-growth were very small, minus 0.1 or 0.0 in most sectors and minus 0.2 in only one sector, Oil and Gas Mining. Likewise, for the IGC there were only marginally significant differences (in the third decimal) for 17 industries and no significant differences in 19 industries.[3]

Electrical equipment

To estimate the share of electrical equipment assets in the total stock of equipment – used to measure the impact of the installed electrical horsepower in manufacturing in section 6 – we also compile an annual series of the nominal net stock of capital for electrical equipment and total equipment. We apply the methods described above but focus on the period 1921 to 1929 instead.[4] The nominal net stock of electrical equipment is the aggregate of the nominal value of the BEA (2010) assets EI60 (electrical transmissions, distribution and industrial apparatus) and EO70 (electrical equipment not elsewhere classified). Total equipment includes electrical equipment in addition to transportation equipment, instruments and non-electrical equipment. We estimate the share of electrical equipment assets by dividing the nominal stock of electrical equipment by the total stock of equipment for the years 1921 and 1929 and then taking the average over both these years.

[3] The tabulated results are available from the authors.

[4] Note that the shorter asset lifetimes for machinery and equipment allows us to safely estimate the real and nominal stocks for these assets beginning in 1921. The greater rate of depreciation reduces the sensitivity of these assets to the assumptions made regarding the pre-1900 rate of investment by individual industries.

Table C3. *Comparative growth in Total Factor Productivity (TFP) by industry, using capital stock vs. capital services and standard depreciation vs. variable retirement, United States, 1929-1941.*

Industry	TFP-growth (percent per annum)					
	Effect capital services			Effect variable retirement		
	Stock	Serv.	Diff.	SD	VR	Diff.
Farming	2.6	2.5	-0.1	2.5	2.2	-0.3
Metals	1.1	0.6	-0.5	0.6	0.1	-0.5
Anthracite Coal	0.4	0.3	0.0	0.3	0.3	0.0
Bituminous Coal	1.9	1.8	-0.1	1.8	1.8	0.0
Oil and Gas	2.1	2.1	0.0	2.1	1.3	-0.7
Non-metals	3.9	3.6	-0.3	3.6	3.3	-0.3
Foods*	3.7	3.7	0.0	3.7	3.5	-0.2
Tobacco	5.8	5.8	-0.1	5.8	5.3	-0.5
Textiles	3.4	3.3	-0.1	3.3	3.1	-0.2
Apparel	-0.4	-0.4	0.0	-0.4	-0.5	-0.1
Leather Products	-0.1	-0.1	0.0	-0.1	-0.2	-0.1
Lumber Products	-1.7	-1.7	0.0	-1.7	-1.9	-0.1
Paper	1.2	1.1	-0.1	1.1	1.0	-0.1
Printing Publishing	0.3	0.3	0.0	0.3	0.2	-0.1
Chemicals	2.2	2.1	0.0	2.1	1.9	-0.2
Petroleum, Coal Products	-1.2	-1.1	0.1	-1.1	-1.6	-0.5
Rubber Products	1.6	1.5	-0.1	1.5	1.4	-0.1
Stone, clay, glass	1.9	1.7	-0.1	1.7	1.5	-0.2
Primary Metals	2.3	2.3	0.0	2.3	2.0	-0.3
Fabricated Metals	1.3	1.3	0.0	1.3	1.2	-0.1
Machinery Non-Electric	2.1	2.2	0.1	2.2	2.0	-0.2
Electric Machinery	4.6	4.7	0.1	4.7	4.6	-0.1
Transport Equipment	3.6	3.6	0.0	3.6	3.4	-0.2
Furniture	1.4	1.4	0.0	1.4	1.4	-0.1
Miscellaneous	1.7	1.6	-0.1	1.6	1.4	-0.2
Electric Utilities	5.2	5.2	0.0	5.2	4.7	-0.4
Manufactured Gas	2.0	2.0	0.0	2.0	1.7	-0.3
Natural Gas	3.8	3.8	0.0	3.8	3.4	-0.4
Construction*	0.4	0.3	0.0	0.3	0.3	-0.1
Wholesale & retail trade*	3.5	3.4	-0.1	3.4	3.3	-0.1
Railroads	2.5	2.6	0.0	2.6	2.4	-0.1
Local Transit	0.5	0.4	-0.1	0.4	0.2	-0.1
Residual Transport	5.5	5.6	0.1	5.6	5.6	0.0
Telephone	1.3	1.4	0.1	1.4	1.1	-0.2
Telegraph	0.8	0.9	0.0	0.9	0.7	-0.1
Post Office*	0.6	0.8	0.1	0.8	0.4	-0.3
FIRE*	-1.3	-1.4	0.0	-1.4	-1.8	-0.4
Spectator Entertainment*	4.4	4.4	0.0	4.4	4.4	0.0
Great Inventions*	3.2	3.2	-0.1	3.2	3.0	-0.2
Aggregate measured sectors	1.8	1.7	0.0	1.7	1.5	-0.2
Residual sector	2.8	2.4	-0.4	2.4	2.3	0.0
PDE	2.0	1.9	-0.1	1.9	1.7	-0.2
Memorandum:						
Kendrick's aggregate measured sectors	2.5	2.5		2.5	2.5	
Kendrick's residual sector	2.0	2.0		2.0	2.0	
Kendrick PDE	2.3	2.3		2.3	2.3	
Minimum	-1.7	-1.7	0.0	-1.7	-1.9	-0.1
Maximum	5.8	5.8	-0.1	5.8	5.6	-0.2
Range	7.5	7.5	0.0	7.5	7.5	0.0

Notes: **Stock** = TFP-growth rate is calculated from input data that include stock-based estimates of capital. **Serv**. = TFP-growth rate is calculated from input data that include service-based estimates of capital. **Diff**. = the difference between the rates in the two preceding columns. **SD** = TFP-growth rate is calculated from input data based on the standard depreciation method using capital services set out in this appendix. **VR** = TFP-growth rate is calculated from input data based on using both capital services (as set out in this appendix) and the variable retirement of capital method introduced by Gordon (2016). Kendrick's aggregate measured sector comprises Farming, Mining, Manufacturing, Transportation, Communication & Public Utilities). ***** = sector measured in this paper but not by Kendrick. For 1929-1941 and 1899-1941 the TFP-growth of Kendrick's measured and residual sectors was estimated using our data for Kendrick's measured sectors for 1929-41 and the relationship between measured, residual and aggregate TFP growth in Kendrick's numbers for his 4 sub-periods during 1899 through 1937. TFP here is refined TFP, i.e. after correcting crude TFP for the growth of labour quality (see text).
Source: Kendrick 1961, pp. 136-7; own calculations, see the text and Appendices B, C and D.

Table C4. *Comparative intensive growth contribution by industry, using capital stock vs. capital services and standard depreciation vs. variable retirement, United States, 1929-1941.*

Industry	Intensive Growth Contribution (VA share x TFP growth)					
	Effect capital services			Effect variable retirement		
	Stock	Serv.	Diff.	SD	VR	Diff.
Farming	0.212	0.206	-0.006	0.206	0.179	-0.027
Metals	0.006	0.003	-0.003	0.003	0.001	-0.003
Anthracite Coal	0.001	0.001	0.000	0.001	0.001	0.000
Bituminous Coal	0.013	0.013	0.000	0.013	0.012	0.000
Oil and Gas	0.018	0.018	0.000	0.018	0.012	-0.006
Non-metals	0.008	0.008	-0.001	0.008	0.007	-0.001
Foods*	0.132	0.131	-0.001	0.131	0.123	-0.008
Tobacco	0.067	0.067	-0.001	0.067	0.061	-0.005
Textiles	0.063	0.062	-0.001	0.062	0.058	-0.004
Apparel	-0.005	-0.006	0.000	-0.006	-0.007	-0.001
Leather Products	0.000	0.000	0.000	0.000	-0.001	-0.001
Lumber Products	-0.014	-0.014	0.000	-0.014	-0.015	-0.001
Paper	0.008	0.008	0.000	0.008	0.007	-0.001
Printing Publishing	0.005	0.005	0.000	0.005	0.003	-0.002
Chemicals	0.031	0.030	-0.001	0.030	0.027	-0.004
Petroleum, Coal Products	-0.015	-0.014	0.001	-0.014	-0.020	-0.007
Rubber Products	0.007	0.007	0.000	0.007	0.006	0.000
Stone, clay, glass	0.017	0.016	-0.001	0.016	0.014	-0.002
Primary Metals	0.054	0.054	0.000	0.054	0.048	-0.006
Fabricated Metals	0.021	0.021	0.000	0.021	0.019	-0.002
Machinery Non-Electric	0.042	0.044	0.002	0.044	0.041	-0.003
Electric Machinery	0.049	0.050	0.001	0.050	0.049	-0.001
Transport Equipment	0.078	0.078	-0.001	0.078	0.073	-0.005
Furniture	0.010	0.010	0.000	0.010	0.009	0.000
Miscellaneous	0.011	0.010	-0.001	0.010	0.009	-0.001
Electric Utilities	0.112	0.112	0.000	0.112	0.102	-0.009
Manufactured Gas	0.004	0.004	0.000	0.004	0.003	-0.001
Natural Gas	0.013	0.013	0.000	0.013	0.012	-0.001
Construction*	0.013	0.012	-0.001	0.012	0.009	-0.003
Wholesale & retail trade*	0.509	0.494	-0.015	0.494	0.483	-0.011
Railroads	0.118	0.119	0.000	0.119	0.112	-0.007
Local Transit	0.004	0.003	-0.001	0.003	0.002	-0.001
Residual Transport	0.097	0.099	0.002	0.099	0.099	-0.001
Telephone	0.012	0.012	0.001	0.012	0.010	-0.002
Telegraph	0.001	0.001	0.000	0.001	0.001	0.000
Post Office*	0.004	0.005	0.001	0.005	0.003	-0.002
FIRE*	-0.156	-0.158	-0.003	-0.158	-0.207	-0.049
Spectator Entertainment*	0.025	0.025	0.000	0.025	0.025	0.000
Great Inventions*	0.845	0.830	-0.015	0.830	0.783	-0.047
Aggregate measured sectors	1.387	1.358	-0.029	1.358	1.180	-0.178
Residual sector	0.586	0.503	-0.083	0.503	0.498	-0.005
PDE	1.974	1.861	-0.113	1.861	1.679	-0.182
Mean	0.041	0.041	-0.001	0.041	0.036	-0.005
Coefficient of variation	2.290	2.283	-0.007	2.283	2.584	0.301
Minimum	-0.156	-0.158	-0.003	-0.158	-0.207	-0.049
Maximum	0.509	0.494	-0.015	0.494	0.483	-0.011
Range	0.665	0.653	-0.013	0.653	0.690	0.038

Notes: see table C3. *Source*: see table C3.

Appendix D. Labour Quality

D1. Discussion of the Kendrick Labour Quality Estimates

Kendrick (1961: 31-34) assessed the effect of skill changes on the composition of the labour force between 1869 and 1957. Instead of measuring changes in education attainment, gender and experience directly, however, he measured the changes in the occupational structure. He adjusted labour input by weighting the person-hours of work in separate occupations and industries by their average hourly earnings for a given base year. Kendrick's measure of labour quality thus captures two effects: (1) the relative shifts of workers between occupations, and (2) the relocation of employment between industries. The first effect, the shift of workers from low-paying positions (e.g. laborers) to better-paying jobs (e.g. operatives or clerical staff), reflects a change in the potential output per worker. The higher earnings (measured in terms of base-period compensation) imply a rise in the marginal productivity of that worker and thus a rise in the quality of the labour force – in line with the Jorgenson approach discussed below. Likewise, the shift of workers to better-paying industries also show up as an increase in labour quality.

Kendrick assumes under (1) that labour quality will only change over time if a worker transfers from one occupation to another or if an individual joins (or leaves) the labour force in an occupation that is better (worse) paid than the national average. Kendrick (1961: 33) surmises that "the inherent average physical and mental capacity of the person employed in each occupation is constant over time." The rapid increase in educational attainment during the late nineteenth and early twentieth century casts serious doubt on this assumption, however. The average years of schooling for cohorts born between 1880 and 1950 nearly doubled, increasing from approximately 8 to 14 years (Goldin and Katz, 2008: 20; 113; 170). Part of this increase in skill translated into a shift of employees between occupations and industries, but part also translated into a rise of the labour quality *within* occupations. For instance, the likelihood for a blue-collar worker born around 1885 to have attended high school was substantially greater than it was for its counterpart born only 10 years prior, around 1875. The high-school education gave the blue-collar worker basic knowledge of chemistry, electricity and algebra, allowed him to read manuals and blueprints and made it much easier for him to effectively converse with managers and other professionals, raising his marginal productivity in the process. In addition to undervaluing the impact of the rapid increases in education attainment during the late nineteenth and early twentieth century, Kendrick's method ignores other demographic changes as well which thus biases his labour quality figures downwards compared to our own (see table D1). Changes in the educational attainment, average age, or experience of the workforce and shifts in the gender composition are generally considered to be determining factors in the quality of labour, as we will illustrate below.

Table D1. *Average Annual Rates of Growth of Labour Quality, United States, 1899-1941, percent per annum.*

	Kendrick	This study
Private domestic economy (PDE)	0.32	0.79
1899-1929	0.36	0.87
1929-1941	0.20	0.59
Private domestic nonfarm economy (PNE)	0.15	0.36
1899-1929	0.16	0.40
1929-1941	0.14	0.27

Source: Kendrick (1961), pp. 333-335; 338-340; this paper.

D2. Discussion of the Labour Quality Estimates in this Paper

In order to fully assess the impact of the substantial investments in schooling as well as the structural changes in the gender and age composition of the American workforce during the early twentieth century, we turn to an approach developed by Dale Jorgenson and Zvi Griliches (1967). The key innovation in their work was to adjust the traditional measure of labour input – i.e. total hours of work – for improvements in quality. The main principle behind the labour quality adjustment is the distinction among several different types of labour inputs characterized by one or more quantifiable factors that affect the productivity potential of the worker (e.g. educational attainment, age, gender). By then assigning weights to these categories – usually in the form of average wages and earnings – one can measure the change in the productivity 'potential' of the workforce. The rationale for this procedure is that differences in average earnings between the labour categories can be thought of as reflecting differences in their marginal productivity. When this new measure of labour input is used in a growth accounting framework, output growth as a result of better educated and trained workers is ascribed to input growth, rather than productivity or technology growth (Jorgenson et al., 2008). Previous studies have shown that this quality adjusted measure can account for a substantial part of the residual or Total Factor Productivity (TFP) growth within traditional growth accounting studies (Denison, 1962; Griliches, 1963; Denison and Poullier, 1967; Gordon, 2010). Therefore, the labour quality adjustment allows for a purer measure of both labour input as well as technical change within a growth accounting framework.

Methodology

To construct an index of labour input for each individual sector, we assume that labour input ($L_{j,t}$) for industry *j* at time *t* can be expressed as a translog function of its individual components (Jorgenson et al., 1999: 92-3). We form indices of labour input from data on hourly employment by industry (H), cross-classified by gender, age and education.[5] Dropping the industry subscript *j* for ease of notation, the growth of labour input (log growth denoted by a hat) for industry *j* can thus be represented as

$$\hat{L} = \sum_{l=1}^{q} \bar{\mu}_l \hat{H}_l \qquad \text{(D.1)}$$

where H_l is total hours of work at the industry level for a given set of *q* characteristics of the labour force *l* (gender, age and education) and $\bar{\mu}_l$ is the two-period average of this employment group's share in the total labor income at the industry level.

$$\bar{\mu}_l = \frac{1}{2}\left(\mu_{l,t} + \mu_{l,t-1}\right) \qquad \text{(D.2)}$$

The share of labour income ($\mu_{l,t}$) at time *t* is derived as the product of the average hourly wage ($p_{l,t}^{L}$) and hours of work ($H_{l,t}$) for each combination of labor characteristic *l,* divided by the total wage sum

$$\mu_l = \frac{p_{l,t}^{L} H_{l,t}}{\sum_{l=1}^{q} p_{l,t}^{L} H_{l,t}} \qquad \text{(D.3)}$$

Alternatively, the index of labour input can also be expressed as the product of total hours (H) and an index of labour quality (Q) or, in growth terms, as

$$\hat{L} = \hat{H} + \hat{Q} \qquad \text{(D.4)}$$

[5] Note that age, in our estimate for labour input, serves as a proxy for (work) experience. We thus assume that an individual has held a job his entire life since leaving high-school or college; depending on his educational attainment.

Rearranging terms in equation (D.4) and substituting the index for labour input by (D.1) we arrive at a direct measure of sectoral labour quality growth

$$\hat{Q} = \sum_{l=1}^{q} \bar{\mu}_l \hat{H}_l - \hat{H} \tag{D.5}$$

The change in labour quality thus reflects the difference between the growth rates of the compensation-weighted index of labour input and sectoral employment.

The drawback of this approach is that it requires highly disaggregate data on hourly employment and compensation, generally not available in the published census reports or secondary sources for the early twentieth century. Fortunately, the Integrated Public Use Microdata Series (IPUMS) has made samples from the decennial population censuses publicly available, providing detailed records for nearly 10 million individuals between 1900 and 1950 (Ruggles et al, 2010). We utilize the microdata from this source to construct our measure of labour quality.

Rather unfortunately, the 1900-1930 American population censuses did not inquire into the educational attainment of the general population, the compensation of workers and employees, nor the average hours of work. To overcome these data issues, we follow a three-tiered approach to the data preparation for the labour quality estimation. First, we estimate educational attainment at the micro level for the pre-1940 census samples on the basis of the 1940 returns. Second, we construct an employment matrix for the entire period that groups workers according to their (predicted) educational attainment, gender, age and by industry. Lastly, we derive the compensation matrix on the basis of average hourly wages for each labour category taken from the 1940 census of population.[6] These employment and compensation matrices can then be used to calculate labour quality on the basis of equation (D.5).

Educational attainment

For the first stage, we estimate the educational attainment *y* for an individual *i* on the basis of his or her occupation, gender, age and place of residence (x_i). This approach takes both the long-run changes in the average years of schooling as well as the effects of changes in the occupational structure and the gender/age composition of the workforce into account. We define four education categories (see table D2) and we predict the likelihood that an individual *i* belongs to each of these specific educational categories (e.g. $Pr\{y_i = 1\}$). This probability should be bounded by 0 and 1, continuous and nonlinear; conditions which are all met by an (ordered) logit model

$$Pr\{y_i \leq k | x_i\} = \frac{e^{x_i' \beta}}{1 + e^{x_i' \beta}} \quad \text{, where } k = 1, 2, 3 \tag{D.6}$$

The right-hand side of equation (D.6) is a cumulative distribution function with mean 0 and standard deviation 1. The coefficients are estimated using maximum likelihood, which is the optimal parametric estimator in this context (Long and Freese 2006).[7]

---

[6] The 1940 census was the first census of its kind to ask about schooling, labour compensation and working hours to all citizens surveyed. In the wake of the depression the 1940 population census dedicated a substantial part of its inquiry into the issue of employment and productivity. Note that for the estimation of labour quality we assume the average hours of work per employment category to remain unchanged relative to 1940.

[7] Note that we estimate the cumulative probability for the first three educational categories, since all individuals that are not part of either the first, second or third category will be part of the fourth category. The fourth category can thus be implicitly derived and should be excluded from the model.

Table D2. *Categorical Variables Logit and Labor Quality Models.*

Logit model	Labor quality model
Education:	Education:
See labor quality model	(1) 1-4 years grade school
	(2) 5-8 years grade school
Gender:	(3) 1-4 years high school
See labor quality model	(4) 1 or more years college
Occupation:	Gender:
(1) professional, technical	(1) male
(2) farmers (owners and managers)	(2) female
(3) managers, officials, and proprietors	
(4) clerical staff	Age:
(5) sales workers	(1) 16-17 years
(6) craftsmen	(2) 18-24 years
(7) operatives	(3) 25-34 years
(8) service workers (household)	(4) 35-44 years
(9) service workers (other)	(5) 45 years and over
(10) laborers	
(11) unemployed/retired	Industry:
	See main text
Region:	
(1) South	
(2) Midwest	
(3) West	
(4) Northeast	

Data

For the estimation of the logit model we rely exclusively on the 1940 1-percent sample included in the IPUMS dataset. This sample is limited to include only those citizens aged 16 years and above, leaving approximately 975,000 observations for the logistic regression. The dataset includes a measure of the highest year of schooling or degree completed. As illustrated in table D2, we reclassify this variable to encompass four distinct educational attainment classes. The reason we reclassify the education variable is twofold. First, by treating it as a categorical variable as opposed to a continuous variable (e.g. years of education), we avoid the assumption that the distances between classes are equal; i.e. that an additional year of grade school is identical to one additional year in college. Second, we limit the number of classes to 4 to ensure that each class is covered by a sufficient number of observations. This is important not just for the estimation of educational attainment, but also for the construction of the compensation matrix.[8] For the independent variables, we follow the literature on US labour quality and mark four variables as important predictors of educational attainment, namely: occupation, birth cohort, gender and region.

Individuals are classified into one of eleven main occupational groups which differ markedly in terms of their average educational attainment. For instance, the probability of a professional (e.g. engineers, economists) having attended high school was substantially greater than was the case for the average labourer. The importance of gender and year of birth is illustrated by Goldin and Katz (2008: 18-22; 170). They observe a rapid increase in the average years of schooling throughout the late nineteenth and early twentieth century. Each successive cohort spent a substantially greater number of years in school compared to the previous generation. In addition, Goldin and Katz (2008: 19) show that women generally attended school for longer than men did throughout most of the early twentieth century. The gender variable was taken directly from the IPUMS dataset while the year of birth was rounded to the

[8] Limiting the number of classes for the education variable allows us, for instance, to test the 'parallel regression assumption'; meaning that for each education class (grade-school, high-school, college, etc.) the coefficients for the independent variables (beta) are identical. As it turns out the assumption is violated. Hence, we effectively estimate separate regressions for all education classes, obtaining different betas for each.

nearest decade. The log of the relative distance in decades to 1930 was then taken as the birth cohort measure. Lastly, the literature points to widespread differences in state support for education and shows that the rise in both high school graduation rates as well as college enrolment rates for states in the North and West of the country were considerably more impressive than for the rest of the nation (Goldin and Katz, 2008: 271-7). We incorporate a variable in the model that differentiates between the four main regions of the country (see table D2).

For the second stage of the labour quality estimation, the construction of the employment matrix, we rely on the IPUMS 1-percent census samples for the decades between 1900 and 1950. To estimate educational attainment we include the occupation, birth cohort, gender and region variables discussed earlier, supplemented by data on the age of the individual, number of weeks worked in the previous year, the average hours of work per week, and industry in which the subject is engaged. The employment sample is limited to include only those citizens between the ages of 16 and 84, who are part of the labour force. For the employment-matrix our dataset includes roughly 3,135,000 individual observations.

In the third stage of the data preparation we again rely on the 1940 sample to estimate relative hourly compensation per labour category. Here we limit the sample to include only those citizens between the ages of 16 and 84 having worked at least 48 weeks in the previous year and earning an income greater than 0 (Goldin and Katz, 2008). These individuals are allocated to the cells of the matrix cross-classified by gender, age, education and industry as summarized in table D2. Compensation is reported in the census as the respondent's total pre-tax wage and salary income for the previous year, expressed in current dollars. To obtain total personal income, which also includes non-wage income, we multiplied the 1940 compensation figures by the industry specific ratio between wage and salary income and total personal income taken from the 1950 census returns. Nonwage income generally represented only a small part of total personal income, with the notable exception of the agricultural sector. The samples for the logistic regression, the compensation matrix and the employment matrix are all weighted by the IPUMS `person weight' variable.

Robustness

Ideally, we would like to allow the weights for our labour quality index to vary over time, reflecting potential changes in relative compensation between the labour categories. Unfortunately, the censuses prior to 1940 did not inquire into either wages or earnings, impeding the accurate measurement of labour compensation for these earlier decades. Reassuringly, Goldin and Katz (2008: 53-63) demonstrate that the wage structure observed in 1940 was fairly typical for the pre-war period. Although they do observe a gradual compression of the wage distribution for production workers between 1890 and 1940, Goldin and Katz conclude that the gap in the skilled/unskilled wage level for 1920 was virtually identical in comparison to 1940.

On the basis of Goldin and Katz's (2010) data for the state of Iowa we can perform a more conclusive sensitivity check of our labour quality figures. The *Iowa State Census Project* provides detailed compensation data, cross-classified by most of the categories that make-up labour input for the year 1915. Below we will compare the results from the third stage of the data preparation – the estimation of the compensation matrix – for the 1915 Iowa data and the original 1940 census data. We will then use these new estimates to provide an alternative estimate of labour quality for the first half of the twentieth century and decompose these estimates to trace the sources of divergence. In addition, we perform the same sensitivity check based on comprehensive income data for 1950, taken from the IPUMS dataset. Overall, on the basis of this evidence presented here, we feel confident using solely the 1940 compensation figures as weights for the construction of our labour quality index.

Table D3. Labour Income Estimates for Iowa and the United States, 1915, 1940 and 1950, Dependent Variable: Log of Labour Income

	US 1940 (1)	US 1940 (2)	Iowa 1940 (3)	Iowa 1915 (4)	US 1950 (5)
Intercept	6.82***	6.90***	6.64***	6.48***	7.54***
	(0.005)	(0.003)	(0.029)	(0.012)	(0.009)
Female dummy	-0.70***	-0.53***	-0.52***	-0.65***	-0.52***
	(0.008)	(0.007)	(0.071)	(0.033)	(0.014)
Age 16-17 dummy	-0.81***	-1.23	-1.21***	-0.96***	-1.30***
	(0.019)	(0.021)	(0.155)	(0.043)	(0.036)
Age 18-24 dummy	-0.35***	-0.44***	-0.53***	-0.54***	-0.29***
	(0.004)	(0.005)	(0.039)	(0.017)	(0.008)
Age 35-44 dummy	0.24***	0.29***	0.28***	0.10***	0.15***
	(0.003)	(0.004)	(0.034)	(0.017)	(0.006)
Age 45+ dummy	0.28***	0.35***	0.35***	0.13***	0.16***
	(0.003)	(0.004)	(0.032)	(0.016)	(0.006)
1-4 yrs. grade school dummy	-0.34***	-0.46***	-0.28***	-0.30***	-0.22***
	(0.006)	(0.006)	(0.100)	(0.025)	(0.009)
1-4 yrs. high school dummy	0.20***	0.27***	0.33***	0.30***	0.20***
	(0.003)	(0.003)	(0.028)	(0.017)	(0.005)
1+ yrs. college dummy	0.47***	0.55***	0.54***	0.52***	0.38***
	(0.004)	(0.004)	(0.038)	(0.021)	(0.007)
Industry dummies	YES	NO	NO	NO	YES
Interaction terms	YES	YES	YES	YES	YES
Observations	207,436	207,436	3,456	14,403	88,071
Adjusted R-squared	0.50	0.37	0.31	0.26	0.35

Notes: Robust standard errors in brackets; * significant at 10%, ** significant at 5%, *** significant at 1%.
Reference category: male worker, aged 25 to 34, 5 to 8 years of grade school.

Table D3 above provides a summary of the relative compensation in 1915, 1940 and 1950. The log of total compensation is regressed against a set of dummies for gender, age and education. We controlled for the full range of industries in our sample and included a set of interactions terms between gender and our main explanatory variables. We included samples from 1940 and 1950 for the whole of the US as well as the state of Iowa, which, as before, are derived from the IPUMS dataset by Ruggles et al. (2010). The 1915 data is taken from the Iowa State Census Project by Goldin and Katz (2010).

A drawback of the 1915 Iowa data is that Goldin and Katz do not report the industry in which the worker was active. Consequently, as relative compensation by industry is unavailable in the 1915 data, we are unable to fully capture the effects of the reallocation of labour between these industries. This reallocation effect turns out to have had a significant impact on overall labour-quality growth between 1900 and 1950, as we will show below. In addition, the sole reliance on income data from Iowa introduces a bias in the compensation estimates, as wages for the different categories were not uniform across all states. Iowa may not be the most representative state for this sensitivity check, but unfortunately it is the only source of micro-data on labour income we have prior to 1940.

To tackle these issues, we include the results from 5 separate estimations in table D3. Column (1) reports the full model that we have relied on so far, based on the 1940 data for the US as a whole, including controls for industries. We drop the industry dummies in (2) and restrict the sample to Iowa in (3). Column (4) shows the results based on the 1915 Iowa data from Golden and Katz. The coefficients from (3) can be directly compared to the estimates from (4). The other columns can be used to gauge the bias introduced by the exclusion of industry specific compensation figures and the sole reliance on data from Iowa. The final column (5) shows the regression based on 1950 census data for the US as a whole. This regression includes a full set of industry dummies and interaction terms and can be directly compared against the results in column (1).

Looking first at columns (3) and (4), table D3 shows that the gap between male and female wages in Iowa was significantly bigger in 1915 than in 1940; the lower and upper 95% confidence limits for the female dummy in model (4) are -0.72 and -0.58 respectively. The estimate for the female dummy in (3) clearly falls outside these bounds. The returns to experience (proxied by age) were smaller in 1915 than 1940. The returns to education were roughly equivalent in the 1915 sample compared to the 1940 Iowa sample; the upper bounds for the high school and college coefficients in (4) are 0.33 and 0.56 respectively. Were we to base our compensation estimates on 1915 (instead of 1940), we would expect the effects on our labour quality index to be mixed. Ignoring the interaction terms, the reduced weight given to female labour in the 1915 would dampen the growth in labour quality, as we observe a sizable increase in the share of women in the labour force over the twentieth century. Similarly, the ageing of the workforce between 1900 and 1950 would show a less pronounced positive effect on labour quality growth. However, the increase in the educational attainment of the workforce during the early twentieth century should have a comparable impact when 1915 weights are used. Comparing column (5) to column (1) we would expect the effect of both gender and age on labour quality growth to be slightly higher based on the 1950 compensation weights, while the effect of education is expected to be lower. The latter can be inferred from the fact that the relative spread between the coefficient for the highest and lowest educational classes is lower based on 1950 data than for the original 1940 data.

As we will show below, of the three labour characteristics (age, gender and education) education is the driving force behind the growth in labour quality over the first half of the twentieth century. The change in educational attainment – particularly the rapid rise in the number of workers that attended high school or even college – is also the most important factor missing from Kendrick's (1961) measure of labour quality. The fact that the 1940 compensation weights allocated to the four educational classes appears to be representative for earlier years is thus reassuring. Based on identical sources, Goldin and Katz (1999: 22, 45) even show that the returns to a year of high school and college education was greater for young men and at least equal for all men in 1915 compared to 1940 when one adjusts the 1915 Iowa data to cover the national economy as a whole. This would mean the contribution of education to labour quality growth would come out even higher if we would include 1915 compensation weights into our analysis. Goldin and Katz also show that in 1950, the returns to education had indeed fallen substantially compared to the pre-war era. This, Goldin and Margo (1992: 32) say, "was primarily the result of a particular confluence of short-run events affecting the demand for labour and of institutional changes brought about by the war and the command economy that accompanied it." The post-war figures are thus less likely to approximate the relative compensation weights for the early twentieth century.

As previously noted, the 1915 Iowa data summarized in table D3 cannot be used to determine whether the 1940 relative wages by industry are relevant for earlier years, since the earlier population census does not reveal which industries the employees were engaged in. For data on pre-1940 labour compensation by industry we turn to the *National Income and Product Accounts* by the BEA (2009), which provides aggregate data from 1929 onwards. Comparing the industry-specific wages in 1929 to those derived from the 1940 census reveals that, over the course of the 1930s, relative wages by industry did not change much. The three worst-paying industries in 1940 were agriculture, personal and public services and the lumber industry. In 1929, agriculture and personal services also recorded the lowest average compensation per worker, while the lumber industry ranked as the seventh worst paying employer. The highest average annual compensation was recorded in the petroleum and coal products industry for both years. Wage data prior to 1929 is not readily available for the entire US economy, but the 1909 *Census of Manufactures* does report wages, salaries and persons employed for the major 2-digit manufacturing industries. Comparing 1909 to 1940 we observe that the textile and lumber mills consistently paid the lowest wages, while the printing and publishing, petroleum and transportation equipment industries always ranked near the top of the list of best-paying industries. This appears to suggest that the industry-specific wages observed in 1940 are a decent proxy for earlier years. The 1940-based compensation data is thus likely to adequately capture the effects of the reallocation of workers between industries on labour quality.

Decomposition

Although the coefficients from the income regression provide a rough overview of the changes of relative compensation of workers between 1915, 1940 and 1950, the effect on our labour quality estimates can only be properly observed by incorporating the new compensation matrices into our full model. We will re-estimate labour quality change between 1900 and 1950 for the private domestic economy based on the compensation weights derived on the basis of estimations (2) through (5) in table D3 and compare them to our baseline estimate from column (1). To fully assess the impact of the different compensation weights – both for the development of labour input as well as aggregate production – we should decompose the labour quality index into its underlying constituents. Jorgenson et al. (1999, p. 239) suggest a breakdown of the index on the basis of its distinctive characteristics. They propose the construction of partial indices of labour input in which only a subset of the characteristics is incorporated. To construct such a partial index, we sum the hours worked and the corresponding value shares over some of the characteristics and construct a translog index over the remaining characteristics.

Previously, we used a single subscript *l* to represent the categories of labour input cross-classified by all characteristics except for industry. Below we use a separate subscript for each of the individual characteristics: two sexes, represented by the subscript *s*; five age-groups, represented by *a*; four educational classes, represented by *e*; and thirty-eight industries, still represented by *j*. An example of the partial labour input index for gender is given below.

$$\hat{L}_s = \sum_{s=1}^{2} \mu_s \sum_{a=1}^{5} \sum_{e=1}^{4} \sum_{j=1}^{38} \hat{H}_{s,a,e,j} \qquad \text{(D.7)}$$

Equation (D.7) is based on equation (D.1), the basic labour input equation introduced in this appendix. However, now the compensation shares $\bar{v}_s$ solely distinguish between the two gender categories and is multiplied by the log change in male and female workers respectively. The resulting partial labour input index only reflects changes in the relative share of men and women in the workforce and ignores the effects of the other characteristics. As before, labour-quality growth can still be derived as the difference between the growth rates of the compensation-weighted, partial index of labour input and hours worked.

Partial indices for all four characteristics can be computed, which are referred to as first-order indices. In addition to these first-order indices, second-order indices of labour input can also be defined. These depend on any two characteristics of labour input, by adding hours of work and the corresponding value shares over other characteristics and again constructing a translog index (Jorgenson et al. 1999, p. 270). Similarly, we can define third- and fourth-order indices. In our full model, there are six second-order indices, four third-order indices and one fourth-order index. The fourth-order index reflects compositional shifts among all characteristics, as in equation (D.1).

The first row in table D4 reports the results from the decomposition of labour-quality growth for the American labour force based on the full US sample for 1940, including controls for industries. The first column in this table displays the annual average log growth over the entire period, while the subsequent columns report the partial, first order indices for education (*e*), age (*a*), gender (*s*) and industry (*j*) respectively. The final column reports the sum of the residual; i.e. second-, third- and fourth-order effects. The other rows in table D4 report the decomposition of labour quality growth based on the four remaining compensation estimates introduced in table D3.

Table D4. *Contribution to Labour Quality Growth for the Private Domestic Economy, United States, 1900-1950, in percent per annum.*

		Total	Educ. (*e*)	Age (*a*)	Gender (*s*)	Industry (*j*)	Resid.
US 1940	(1)	0.83	0.41	0.15	-0.12	0.48	-0.09
US 1940	(2)	0.52	0.38	0.19	-0.14	…	0.10
Iowa 1940	(3)	0.47	0.33	0.20	-0.16	…	0.10
Iowa 1915	(4)	0.35	0.28	0.15	-0.15	…	0.08
US 1950	(5)	0.69	0.34	0.13	-0.12	0.37	-0.04

Note: May not sum to total due to rounding. Educ. = Education. Resid. = Residual.
Sources: see text.

The first row in table D4 shows that the growth of labour quality, at the total economy level, appeared to be driven primarily by the change in educational attainment and shifts in the industrial structure. The contribution of education was positive for all decades and showed a rising trend over time, reflecting the findings by Goldin and Katz (2008). The relocation of labour from low-skill/low-productive sectors (e.g. agriculture) to high-skill sectors (e.g. trade and FIRE), reflected an improvement in the utilization of the workforce, greatly raising the potential output per worker. To a lesser extent, the gradual rise in the experience level of the American workforce, as illustrated by the increase in the average age, also positively contributed to labour-quality growth. In contrast, the rising share of women in the labour force tended to depress the growth of labour quality. Particularly the period between 1940 and 1950 – as a result of the war effort – observed a marked increase in the number of female workers.

The results from estimation (2) – where the variations in income between industries are no longer taken into account – shows a marked drop in the annual average growth of labour quality. The reallocation of workers between industries contributed a little over 0.30% per annum to labour quality growth. Note that the contributions of the remaining first order indices changes slightly as well, as the variations in income among individuals is now attributed to these categories instead of to the differences in compensation between industries (see table D3). If we narrow the 1940 sample in (3) to include compensation figures from Iowa only, we again observe a modest downward adjustment of 0.05%. Compared to (2), the difference in annual labour quality growth appears to come from a lower contribution of education as a result of the reduced returns to education we observed for (3) in table D3.

The penultimate row in table D4 reports the results based on the 1915 Iowa sample. If we compare the estimates from (4) directly to (3), we see that using the earlier weights would lower labour quality growth by about 0.12% per annum. Half of this difference comes from a reduced contribution of education and half from a lower contribution of work experience. Taking the bias for the Iowa sample and the mismeasurement of the reallocation of labour into account – observed in estimations (3) and (2) respectively – we would expect the average labour quality growth for the private domestic economy to be approximately 0.70% per annum based on the 1915 compensation weights. The labour quality estimates for the individual industries based on the 1915 income regression appear to be very similar to our baseline findings as well. The correlation between the labour quality estimates based on (1) and (4) for the disaggregate industries measured for each decade individually is a strong 0.97.

The findings on the basis of the 1950 compensation weights in estimation (5) paint a strikingly similar picture. Overall labour quality growth is reduced by 0.14% per annum compared to our original estimates in the first row of table D4. Again, the difference stems primarily from a reduced contribution of education and a lower reallocation effect (*j*). Based on the 1950 compensation data, annual labour quality growth is still approximately 0.70%. Once again, the correlation between the labour quality estimates for the individual industries based on (1) and (5) is very high: 0.98.

Overall, the modest difference between the labour quality results at the total economy level based on the 1915, 1950 and the original 1940 weights of approximately 0.12-0.14% per annum shows that our results are quite robust. This conclusion is reinforced by the striking similarity between the disaggregate results based on the two sets of weights. The benefits of the detailed 1940 estimate, that not only covers the income differences for the full US sample but can also take the reallocation effects of the shift in employment between industries into account, outweighs the need to incorporate changes in the relative incomes over time into the analysis. We prefer the 1940 weights over the 1950 weights as the latter falls outside the period we study in this paper. The post-war figures are also less likely to capture the relative compensation between the labour categories for the early twentieth century, particularly in the case of the educational classes.

Table D5. *Growth in Labour Quality by Industry, United States, 1899-1941.*

Industry	Growth in labour quality (percent per annum)				
	1899-1909	1909-1919	1919-1929	1929-1941	1899-1941
Farming	0.00	0.73	0.31	0.48	0.38
Metals	0.16	0.54	0.54	0.44	0.42
Anthracite Coal	0.19	0.49	0.56	0.51	0.44
Bituminous Coal	0.19	0.49	0.56	0.51	0.44
Oil and Gas	-0.05	-0.22	0.73	0.65	0.30
Non-metals	0.13	0.55	0.14	0.54	0.35
Foods*	0.06	-0.06	0.26	0.58	0.23
Tobacco	-0.38	0.10	0.20	0.65	0.17
Textiles	0.52	0.67	0.66	0.71	0.64
Apparel	0.24	0.63	0.40	0.06	0.32
Leather Products	-0.71	0.23	0.26	0.31	0.04
Lumber Products	-0.06	0.47	0.46	0.45	0.33
Paper	0.93	0.63	0.67	0.60	0.70
Printing Publishing	0.07	0.42	0.39	0.47	0.34
Chemicals	-0.09	0.51	0.43	0.62	0.38
Petroleum, Coal Products	0.19	0.61	0.50	0.84	0.55
Rubber Products	0.48	0.75	0.56	0.76	0.65
Stone, clay, glass	-0.03	0.37	0.57	0.44	0.34
Primary Metals	0.05	0.57	0.61	0.47	0.43
Fabricated Metals	-0.02	0.45	0.56	0.42	0.35
Machinery Non-Electric	0.02	0.40	0.75	0.54	0.43
Electric Machinery	0.84	0.48	0.51	0.51	0.58
Transport Equipment	-0.20	0.15	0.60	0.60	0.30
Furniture	-0.36	0.45	0.23	0.32	0.17
Miscellaneous	0.00	0.44	0.71	0.43	0.40
Electric Utilities	0.13	0.39	0.40	0.98	0.50
Manufactured Gas	-0.19	0.14	0.45	0.71	0.30
Natural Gas	-0.19	0.14	0.45	0.71	0.30
Construction*	-0.14	0.49	0.15	0.13	0.16
Wholesale & retail trade*	-0.04	0.40	0.19	0.09	0.16
Railroads	-0.08	0.53	0.75	0.76	0.50
Local Transit	0.14	0.60	0.60	0.57	0.48
Residual Transport	0.14	0.12	0.55	0.49	0.33
Telephone	0.18	0.00	0.80	1.14	0.56
Telegraph	-0.06	0.17	-0.19	0.84	0.22
Post Office*	0.30	0.35	0.46	0.44	0.39
FIRE*	-0.36	-0.25	0.41	0.54	0.11
Spectator Entertainment*	0.21	0.77	0.30	0.24	0.37
Manufacturing	0.21	0.71	0.56	0.43	0.47
Great inventions*	0.25	0.62	0.29	0.24	0.35
Aggregate measured sectors	0.99	1.09	0.81	0.75	0.90
PDE	0.85	1.12	0.65	0.59	0.79
Memorandum:					
Kendrick PDE	0.50	0.41	0.17	0.20	0.32
Minimum	-0.71	-0.22	-0.19	0.06	0.04
Maximum	0.93	0.99	0.80	1.14	0.70
Range	1.64	1.20	0.98	1.08	0.67

Notes: * = sector measured in this paper but not by Kendrick.
Source: own calculation, see text and Appendix D. Average annual growth rates calculated using continuous compounding.

Appendix E. Dual TFP

So far, we have relied on a conventional or 'primal' definition of TFP, where the rate of TFP growth is defined as the difference between the growth of real output and the weighted average growth of real physical and human capital. For 1929-1941, the growth of real factor input is given by the change in the flow of respectively capital- and labour services, both weighted by their respective factor shares in the total value of output.[9] The intuition behind the calculation of primal TFP is that real output growth not covered by the growth in real factor input represents a shift in the production function resulting from productivity advances, efficiency change or, potentially, measurement error (Solow 1957).

Griliches and Jorgenson (1967) demonstrate that TFP can also be computed from indices of prices for output and input, instead of quantities. The reasoning behind this 'dual' approach to productivity change is very similar to that of primal TFP: any price reduction for output not resulting from the (weighted) change in nominal wages or the rental price of capital represents a shift in the production function. Griliches and Jorgenson show that under the same assumptions – e.g. constant returns to scale and perfect competition – the primal and dual approaches will yield identical results. Even if these assumptions were to be violated, however, Antràs and Voth (2003: 57) show that primal and dual estimates of TFP would be biased to the same degree, as both measures are theoretically equivalent. The frequently observed differences between estimates for the primal and dual approach are more likely the result of inconsistencies in how prices and quantities are measured in the national accounts or the use of a different production function or disparate factor shares (see Hsieh 2002; Aiyar and Dalgaard 2005).

For the period 1929-1941 we have all the building blocks required to estimate dual TFP. Denoting TFP by *A*, the dual growth accounting identity is given by:

$$\widehat{A} = \alpha\widehat{p}_K + (1 - \alpha)\widehat{p}_L - \widehat{p}_Y \qquad \text{(E.1)}$$

where hats indicate growth rates in natural logs, *p* the price for capital (*K*), labour (*L*) and output (*Y*) and α the factor share of capital.

The prices in equation (E.1) are derived implicitly from the inputs for the primal TFP estimates, as described in the main text and appendices C and D. The growth of output prices ($\hat{p}_Y$) is estimated as the residual of nominal- and real value-added growth at both the sector and the industry levels.

We adopt the definition by the BLS (2014) for nominal labour income, i.e. the sum of compensation to employees (COMP) and a portion of noncorporate income (INC).[10] Note that both compensation and noncorporate income are directly available from the NIPA for 1929-1941; see table C1. From the growth in nominal labour income, we deduct the growth in total persons engaged in production (PEP), the average hours of work for employees and labour quality to obtain the growth of the price of labour ($\hat{p}_L$). Persons engaged is listed in the NIPA, average hours of work are taken from the HSUS (1975) and Kendrick (1961), and the construction of labour quality is described in detail in appendix D.

As discussed in appendix C, the rental price of capital assets depends on the rate of depreciation, the (industry) rate of return and any capital gains or losses from price changes in the price of capital assets; see equation (C.6). To arrive at the average growth rate of the price of capital by industry ($\hat{p}_K$), we weighted the rental price for capital assets by that asset's share in total capital compensation; see equation (C.5).

[9] Prior to 1929, for the capital input we rely on the growth rate of the stock of capital instead.

[10] The portion of noncorporate income allocated to labour income is the sum of proprietors' income (PROINC) and farm income (FRMINC) times the ratio of compensation (COMP) and nominal value added minus total noncorporate income.

The factor share of capital (α) is calculated as the average price of capital assets times the total stock, which is identical to gross operating surplus (GOS) minus the sum of noncorporate income not allocated to labour. Note that the factor-share of capital used in the dual approach deviates from that used previously for primal TFP. The factor shares taken from Kendrick (1961), used throughout this paper to retain consistency with Kendrick's original estimates as well as between the different periods, allocate greater weight to labour.[11] Still, as the growth of the price of labour and capital is reasonably similar for the period 1929-1941 (0.7 and 1.2% p.a. respectively), the factor shares only modestly impact the estimate of dual TFP.

Table E1 compares the results for the dual TFP estimates to their primal counterparts for the years 1929-1941, as well as for 1948-1960 and 1960-1973. The post-war growth rates are taken directly from the Bureau of Labor Statistics' (2014) Multifactor Productivity Measures.

Table E1. *Average Annual Rates of Growth of TFP, United States, 1929-1973, percent per annum*.

	Primal	Dual
Private domestic economy (PDE)		
1929-1941	1.86	1.82
1948-1960	1.98	1.87
1960-1973	2.21	2.25
Private domestic nonfarm economy (PNE)		
1929-1941	1.93	1.89
1948-1960	1.68	1.59
1960-1973	1.99	2.02

Sources: primal TFP see tables 5 and 7; dual TFP, 19291941: see appendices C and D; 1948-1973: and Bureau of Labor Statistics, "Historical Multifactor Productivity Measures", http://www.bls.gov/mfp/home.htm (October 2014).

Table E1 shows that for pre- and post-war periods, primal and dual TFP estimates were very similar. For the private domestic economy, dual TFP was still a very strong 1.82% growth p.a. during 1929-1941, but not as strong as during the 1960s which showed productivity increases at 2.25% p.a. Overall, the dual TFP estimates bolster our finding that the 1930s was not the most technologically progressive decade of the century, as claimed by Field (2011, 2013).

[11] Kendrick (1961: 285) allocates a 23 percent share to capital for the PDE, whereas the method described for the dual approach estimates this share at 37 percent.

Appendix F. Research and Development Inputs by Industry

The purpose of this appendix is to show how our new sectoral TFP growth data can be linked to sectoral data on R&D inputs to explore whether or not there is a clear link between R&D inputs and TFP growth, as some authors suggest (Bloom et al. 2017). The data on R&D inputs is derived from the National Research Council survey data on R&D in manufacturing as reported by Mowery and Rosenberg (1989), based on Mowery (1981). The data consists of laboratory foundations per decade, the number of research scientists and the number of research scientists per 1,000 of all wage earners. For the R&D laboratory foundations data, our sector Miscellaneous is matched with the category 'Instruments' and with the miscellaneous laboratories from the survey data. The latter have been computed by subtracting the sectoral total from the grand total. The cumulative foundations of new laboratories from 'prior to 1899' through 1918 is what is reported Table F1; data on laboratories subsequently closed is not available. For the number of research scientists, Mowery and Rosenberg (1989) report that the miscellaneous category is not available, and in this instance only Instruments have been used to match with our Miscellaneous sector.

At the disaggregated level, comparing R&D indicators for manufacturing industries with TFP growth shows a weak relationship, with limited correlation in some instances but none in others. A simple visual inspection of several R&D indicators (Table F1), shows that the ratio of TFP growth acceleration over R&D inputs differed sharply by industry, making it unlikely that R&D came anywhere close to dominating TFP growth. Simple coefficients of rank order correlation for R&D indicators with various accelerations in TFP growth run from 0.15 to 0.61, suggesting some influence of R&D but not a lot (Table F2).

Table F1. *R&D inputs and TFP Growth Acceleration in U.S. manufacturing industries, 1919-1929.*

	R&D labs founded pre-1919 (%)	Share of scientists in 1921 (%)	Scientist intensity in 1921 (average)	Growth scientists 1921-27 (% p.a.)	Growth intensity 1921-27 (% p.a.)	TFP growth acceleration (%-pt p.a.)
Foods	9	4	0.3	19	17	5.1
Tobacco	0	0	0.0			1.0
Textiles	3	1	0.0	28	26	1.8
Apparel	0	0	0.0			2.2
Leather Products	1	1	0.2	6	3	2.2
Lumber Products	0	1	0.1	9	22	3.1
Paper	4	3	0.9	13	10	3.3
Printing Publishing	0	0	0.0			0.2
Chemicals	28	40	9.3	8	4	7.0
Petroleum, Coal Products	3	6	3.3	18	16	8.3
Rubber Products	3	7	3.6	9	4	2.8
Stone, clay, glass	6	3	0.7	24	19	3.8
Primary Metals	9	11	1.4	10	3	4.1
Fabricated Metals	6	4	0.5	20	14	2.2
Machinery NonElectric	10	5	0.4	20	16	1.5
Electric Machinery	9	7	2.0	22	16	3.1
Transport Equipment	3	3	0.4	19	16	3.7
Furniture	0	0	0.0			4.6
Miscellaneous	5	5	0.7	13	8	4.0
Total	100	100	1.0	14	7	3.7
Coefficient of variation	1.19	1.64	1.73	0.41	0.54	0.57

Notes: The TFP-growth acceleration has been calculated by subtracting the growth rate for 1899-1919 from the growth rate of 1919-1941. Share of scientists has been calculated from the absolute numbers and total reported for 1921 by Mowery and Rosenberg (1989), p. 64. R&D labs founded and scientist data exclude independent research laboratories, whose scientists numbered 15.2% of all research scientists in manufacturing in 1921. Scientist intensity = the number of research scientists per 1,000 of all wage earners.

Sources: Mowery and Rosenberg (1989); Bakker, Crafts and Woltjer (2017).

Table F2. *Correlation between R&D inputs and TFP growth indicators in U.S. manufacturing, 1919-1929.*

	TFP growth acceleration	IGC	IGC/VA (TFP growth)
Share R&D Labs foundations before 1919	0.49	0.31	0.29
Share of scientists in 1921	0.55	0.24	0.44
Scientist intensity in 1921	0.61	0.15	0.58
Growth of			
Scientists, 1921-1927	-0.16	0.15	-0.11
Scientist intensity, 1921-1927	-0.17	0.04	-0.33

Notes: The TFP-growth acceleration has been calculated by subtracting the growth rate for 1899-1919 from the growth rate of 1919-1941. R&D labs founded and scientist data exclude independent research laboratories, whose scientists numbered 15.2% of all research scientists in manufacturing in 1921. Scientist intensity = the number of research scientists per 1,000 of all wage earners.
IGC = intensive growth contribution, share of total; VA = share in total value added.
Sources: Mowery and Rosenberg (1989); Bakker, Crafts and Woltjer (2017).

Lorenz curves plotting the two key indicators and the IGC show the same picture (Figure F1, panels A and B). If there would be a clear one-to-one relationship between R&D indicators and the IGC, one would expect the Lorenz curves to be close to the diagonal. In practice, they are very far away from that line, showing that the effect of R&D was extremely variable: if we look at the share of intensive growth contribution (IGC) over the share in research scientists employed in 1921 per industry (the cumulative of which is the Lorenz curve's tangent), the range covers two orders of magnitude, or more than fortyfold, from 0.2 in Chemicals to 7.7 in Textiles. The IGC over R&D labs founded before 1919 ranges over three orders of magnitude across industries, or more than a hundredfold, from 0.3 in Chemicals to 29.3 in Apparel. If TFP-growth was largely driven by R&D outlays, we would not expect to see these large differences between R&D and IGC across industries. Likewise, the five industries with no R&D Labs whatsoever before 1919, accounted for 18% of the IGC in the 1920s, and obversely, industries accounting for 50% of R&D labs founded before 1919 accounted for just 16% of the IGC in the 1920s.

The big differences are reflected in the 1920s TFP growth acceleration over research intensity, as the highest ratio (Textiles) was ninety times the lowest (Chemicals). The three sectors growing fastest in research intensity between 1921-1927 (Textiles; Stone, clay, glass; Electric Machinery) accounted for 3.7% of value added and 10.7% of the IGC in the 1920s, and 3.8% vs. 22.9% in the 1930s. An OLS regression on the log growth acceleration of TFP for the 1920s on the log of research intensity in 1921 suggests that increasing research intensity by one percent near the mean of one scientist per 2,000 wage earners, would increase TFP growth acceleration by a little under 0.2 percentage points (Figure 1 in the main text). Yet the same exercises for the 1930s yield negative coefficients and using other R&D indicators yields even less robust relationships.

A special case is chemicals, which was by far the most R&D intensive industry. One reason for this might be the discovery of the periodic table in the 1860s, which sharply decreased the costs of the R&D needed to reach a given quality level and provided a clear 'roadmap' for future research efforts. Chemicals accounted for almost 30% of labs founded before 1919, 40% of all research scientists employed in 1921, and 20% of additional scientists hired until 1927, but for only 7% of TFP growth in the 1920s and 5% in the 1930s. While Chemicals had the second-highest TFP growth acceleration in the 1920s, it had the third-lowest in the 1930s. This underperformance surprises all the more because other evidence also suggests chemical R&D became extremely well-developed, especially by absorbing knowledge from abroad: it obtained various patents

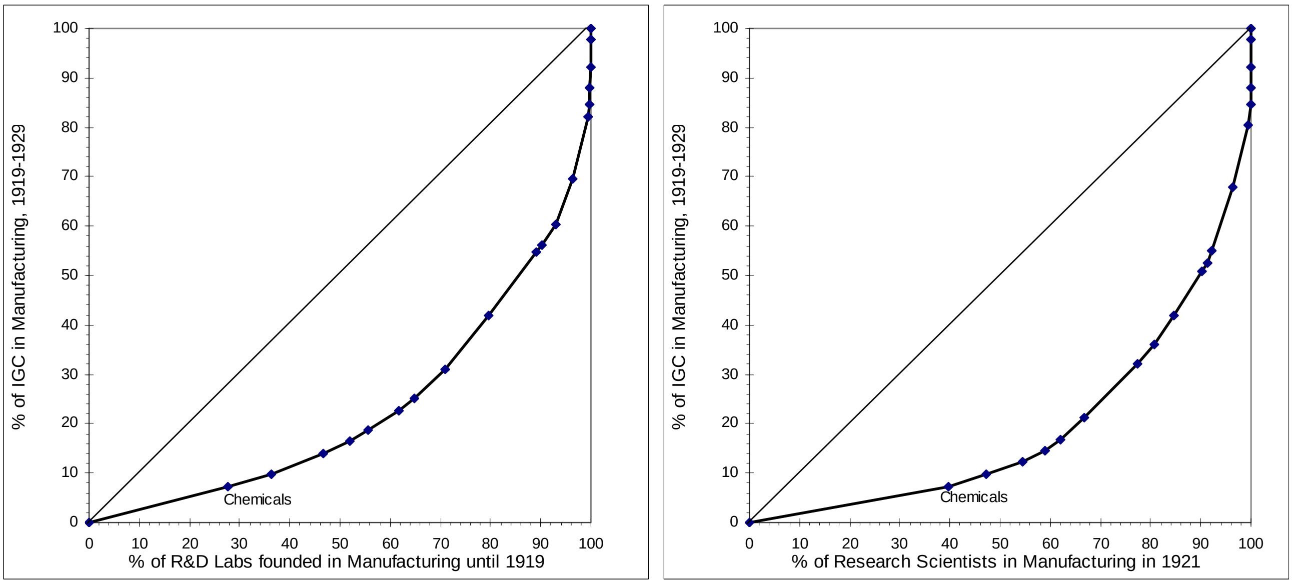

Figure F1. Panels A and B. *Lorenz curves for the share of R&D Labs founded before 1919, the share of research scientists in 1921, versus intensive growth contribution, for US manufacturing sectors, 1919-1929.*

Source: National Research Council surveys on R&D in Manufacturing, reported by Mowery and Rosenberg (1989); Table 2 of this paper.

through wartime expropriations (Moser and Voena, 2012), and patent acquisition continued throughout the 1920s. Standard Oil of New Jersey even licensed the entire oil patent portfolio of IG Farben, the German chemicals cartel, for an amount equal to about 30% of total U.S. industrial R&D outlays in 1930 (Bakker, 2013: 1801-2; Enos, 1962). During the 1930s, many top German scientists fled to the United States, generating quantifiable knowledge spillovers (Moser et al., 2014). Clearly, if R&D was driving TFP-growth one would expect Chemicals to have had a much higher IGC.

The survey data used as the source here is of course not perfect and one should take into account the following qualifications. Independent research labs are not taken into account, and in 1921 15.2% of research scientists in industry worked in independent R&D labs, declining to 8.7% by 1940 (Mowery and Rosenberg 1989). This means that some of the growth in some industries may not reflect an increase in R&D but also partially a substitution from external to internal R&D. Noise might also be caused by whether all those research scientists really only worked in R&D. Edgerton (1996) writes that in industry most scientists are employed outside of R&D. According to the source, what is counted are scientists employed in R&D, but we do not know to what extent they were also used in day to day management of the manufacturing process especially in industries such as Chemicals and Petroleum. This fraction may differ by industry. Some noise might also be caused by the circumstance that only a very minor part of R&D is generally is Research such as inventions and innovations; the overwhelming majority of research scientists in industry were employed in Development (Edgerton 1996).

Appendix References

Abramovitz, Moses. 1964. *Evidences of Long Swings in Aggregate Construction since the Civil War.* New York: Columbia University Press.

Aiyar, Shekhar and Dalgaard, Carl-Johan. 2005. "Total Factor Productivity Revisited: A Dual Approach to Development Accounting." *IMF Staff Papers* 52(1): 82-102.

Antràs, Pol and Voth, Hans-Joachim. 2003. "Factor prices and productivity growth during the British industrial revolution." *Explorations in Economic History* 40: 52-77.

Bakker, Gerben. 2012. "How Motion Pictures Industrialized Entertainment." *Journal of Economic History*, 72(4): 1036-1063.

Bakker, Gerben. 2013. "Money for Nothing: How Firms Have Financed R&D-Projects since the Industrial Revolution." *Research Policy* 42(10): 1793-1814.

Bloom, Nicholas, Charles I. Jones, C. I., John Van Reenen, J. and Michael Webb. 2017. "Are Ideas Getting Harder to Find?" CEPR Discussion Paper No. 12294.

Bureau of Labor Statistics. 2014. "Historical Multifactor Productivity Measures", http://www.bls.gov/mfp/home.htm (October 2014).

Carter, Susan B., Scott S. Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright (Eds.). 2006. *Historical Statistics of the United States: Earliest Times to the Present*. Cambridge: Cambridge University Press.

Central Statistical Office. 1956. *National Income Statistics: Sources and Methods.* London: HMSO.

Chandler, Alfred D., Jr. 1990. *Scale and Scope. The Dynamics of Industrial Capitalism.* Cambridge, Mass.: Belknap Press.

Denison, E. 1962. *The Sources of Economic Growth in the United States and the Alternatives before Us*. New York: Committee for Economic Development.

Denison, E., and J. Poullier. *Why Growth Rates Differ: Postwar Experience in Nine Western Countries*. Washington D.C.: Brookings Institution Press, 1967.

Edgerton, David. 1996. *Science, Technology and the British Industrial 'Decline', 1870-1970*. Cambridge: Cambridge University Press.

Enos, John L. 1962. *Petroleum Progress and Profits: A History of Process Innovation*. Cambridge, Mass.: MIT Press.

Fabricant, Solomon. 1940. *The Output of Manufacturing Industries, 1899-1937*. New York: National Bureau of Economic Research, Inc.

Field, Alexander J. 2003. "The Most Technologically Progressive Decade of the Century," *American Economic Review* 93(4): 1399-1413.

Field, Alexander J. 2006. "Technological Change and U.S. Productivity Growth in the Interwar Years," *Journal of Economic History* 66(1): 203-236.

Fraumeni, Barbara M. 1997. "The Measurement of Depreciation in the U.S. National Income and Product Accounts," *Survey of Current Business*, 77(7): 7-23.

Gallman, Robert E. and Thomas J. Weiss. 1969. "The Service Industries in the Nineteenth Century." In *Production and Productivity in the Service Industries,* edited by Victor R. Fuchs, 287-381. New York: Columbia University Press.

Goldin, C., and L. Katz. 1999. "The Returns to Skill in the United States across the Twentieth Century." *National Bureau of Economic Research Working Paper* 7126: 1–49.

Goldin, C., and L. Katz. 2008. *The Race between Education and Technology*. Cambridge: Belknap Press of Harvard University Press.

Goldin, C., and L. Katz. 2010. *The 1915 Iowa State Census Project*. ICPSR28501-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-12-14. http://doi.org/10.3886/ICPSR28501.v1

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

Goldsmith, Raymond W. 1956, *A Study of Saving in the United States*, Volume III, Princeton: Princeton University Press

Goldsmith, Raymond W. 1958. *Studies in Capital Formation and Financing III: Financial Intermediaries in the American Economy since 1900.* Princeton: Princeton University Press.

Gordon, R. 2010. "Revisiting U.S. Productivity Growth over the Past Century with a View of the Future." *National Bureau of Economic Research Working Paper* 15834: 1–30.

Gordon, Robert J. 2016. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton, NJ.: Princeton University Press.

Gould, Jacob M. 1946. *Output and Productivity in the Electric and Gas Utilities, 1899-1942*. New York: National Bureau of Economic Research.

Griliches, Z. 1963. "The Sources of Measured Productivity Growth: United States Agriculture, 1940–60." *Journal of Political Economy* 71: 331–346.

Griliches, Zvi and Jorgenson, Dale. 1967. "The explanation of productivity change." *Review of Economic Studies* 34: 249–283.

Hsieh, Chang-Tai. 2002. "Productivity Growth and Factor Prices in East Asia." *The American Economic Review* 89: 133-38.

Johnston, Louis and Samuel H. Williamson. 2017. "What Was the U.S. GDP Then?" Measuring Worth, http://www.measuringworth.org/usgdp/, accessed 10 April 2017.

Jorgenson, D., F. Gollop, and B. Fraumeni. 1999. *Productivity and US Economic Growth*, 2nd ed. New York: toExcel.

Jorgenson, D., and Z. Griliches. 1967. "The Explanation of Productivity Change." *Review of Economic Studies* 34: 249–283.

Jorgenson, D., M. Ho, and K. Stiroh. 2008. "A Retrospective Look at the U.S. Productivity Growth Resurgence." *Journal of Economic Perspectives* 22: 3–24.

Jowett, Garth S. 1974. "The First Motion Picture Audiences,' *Journal of Popular Film*, 3 (Winter):39-45.

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

King, Willford I. 1930. *The National Income and its Purchasing Power*. New York: National Bureau of Economic Research.

Kuznets, Simon S. 1941. *National Income and its Composition, 1919-1938.* New York: National Bureau of Economic Research.

Kuznets, Simon S. 1961. *Capital in the American Economy: Its Formation and Financing.* Princeton: Princeton University Press .

Lebergott, Stanley. 1996. *Consumer Expenditures: New Measures and Old Motives.* Princeton: Princeton University Press.

Leontief, Wassily W. 1953. *Studies in the Structure of the American Economy: Theoretical and Empirical Explorations in Input-Output Analysis.* New York: Oxford University Press.

Long, J. Scott and Freese, Jeremy. 2006. *Regression Models for Categorical Dependent Variables Using Stata. Second Edition.* College Station, TX: StataCorp LP.

Moser, Petra, and Alessandra Voena. 2012. "Compulsory Licensing: Evidence from Trading with the Enemy Act." *American Economic Review*, 102(1): 396-427.

Moser, Petra, Alessandra Voena and Fabian Waldinger. 2014. "German Jewish Emigres and US Invention." *American Economic Review*, 104(10): 322-355.

Mowery, David C. 1981. "The Emergence and Growth of Industrial Research in American Manufacturing, 1899-1945." PhD Diss. Stanford University.

Mowery, David C. and Rosenberg, N. 1989. *Technology and the Pursuit of Economic Growth*. Cambridge: Cambridge University Press.

National Bureau of Economic Research, "US Business Cycle Expansions and Contractions," http://www.nber.org/cycles.html (accessed 28 November 2015).

Philippon, Thomas. 2015. "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation." *American Economic Review* 105: 1408-38.

Ruggles, S., T. Alexander, Genadek K., R. Goeken, and M. Schroeder. 2010. *Integrated Public Use Microdata Series*. 5.0. Minneapolis: University of Minnesota.

U.S. Bureau of the Census. 1913. *Thirteenth Census of the United States*, Vol. 8. Washington, D.C.: U.S. Government Printing Office.

____. 1960. *Historical Statistics of the United States: Colonial Times to 1957.* Washington D.C.: US Department of Commerce, Bureau of the Census.

U.S. Bureau of Economic Analysis. 1966. *The National Income and Product Accounts of the United States, 1929-1965: Statistical Tables*, Washington, D.C.: U.S. Government Printing Office.

____. 1993. *Fixed Reproducible Tangible Wealth in the United States, 1925-89*, Washington, D.C.: U.S. Government Printing Office.

____. 2002. *Fixed Assets*, http://www.bea.gov/National/FAweb/Index2002.htm (accessed April 4, 2014).

____. 2003. *Fixed Assets and Consumer Durable Goods in the United States, 1925–99*, Washington, D.C.: U.S. Government Printing Office.

____. 2009. *National Income and Product Accounts* (accessed January 11, 2010).

____. 2010. *Fixed Assets* (accessed December 8, 2011).

____. 2011. "Value Added by Industry, Gross Output by Industry, Intermediate Inputs by Industry, the Components of Value Added by Industry, and Employment by Industry, 1947-1997," *Historical Industry Accounts Data,* http://bea.gov/industry/io_histannual.htm (accessed April 4, 2014).

U.S. Bureau of Labor Statistics. 1943. *Wholesale Prices, July-December and Year 1942*, Bulletin No. 736, Washington, D.C.: U.S. Government Printing Office.

____. 1949. *Wholesale Prices, 1947*, Bulletin No. 947, Washington, D.C.: U.S. Government Printing Office.

____. 1958. *Wholesale Prices, and Price Indexes 1957*, Bulletin No. 1235, Washington, D.C.: U.S. Government Printing Office.

U.S. Bureau of the Census. 1975. *Historical Statistics of the United States, Colonial Times to 1970*, Washington, D.C.: U.S. Government Printing Office.

Whitney, W. G. 1968. *The Structure of the American Economy in the Late Nineteenth Century.* Cambridge, Mass.: Harvard University, Ph.D. dissertation.