Estimating Physical Capital and Land for States and Sectors of the United States, 1850-2000

Turner, Chad and Tamura, Robert and Schoellman, Todd and Mulholland, Sean

Texas AM University Corpus Christi, Clemson University, Arizona State University, Stonehill College

24 June 2011

Estimating Physical Capital and Land for States and Sectors of the United States, 1850-2000

Chad Turner[*], Robert Tamura[†], Todd Schoellman[‡], Sean E. Mulholland[§]

June, 2011

Abstract

This paper introduces new estimates of physical capital and land for the states of the United States covering up to 150 years, from 1850–2000. The estimates of physical capital are decomposed into estimates for agriculture, manufacturing, and a residual sector, while the estimates of land are for agriculture only. The paper describes the data sources and methodology used to generate the estimates. It provides sensitivity analysis when alternative choices are possible and comparison to alternative benchmarks when available.

[*]Department of Finance, Economics, and Decision Sciences, Texas A&M University - Corpus Christi; chad.turner@tamucc.edu

[†]Department of Economics, Clemson University and Federal Reserve Bank of Atlanta; rtamura@clemson.edu

[‡]Department of Economics, Arizona State University; todd.schoellman@gmail.com

[§]Department of Economics, Stonehill College; smulholland@stonehill.edu.

1 Introduction

This paper provides a detailed description of the construction of original estimates of physical capital and land for the states of the United States from 1850 to 2000. The estimates of the physical capital stock are further sub-divided into three sectors: agricultural, manufacturing, and other. The estimates of land are for the agricultural sector only. We use these estimates elsewhere to understand the extent and role of sectoral convergence, as well as the role of input growth in accounting for the long-run growth of states (Schoellman and Tamura 2011, Turner, Tamura, and Mulholland 2011). We also provide the estimates of the physical capital stock and land on the authors' webpages for interested researchers.

The primary historical sources for measured physical capital and land are the U.S. Census of Agriculture and the U.S. Census of Manufactures, with some data tabulated in a more convenient form in the *Statistical Abstracts of the United States* (U.S. Census Bureau various years). We also draw on the United States Department of Agriculture (USDA)'s Farm Balance Sheets, and various special reports to the censuses for additional information when available. These sources allow us to estimate physical capital in the agriculture and manufacturing sectors, as well as land in the agricultural sector. Detailed enumeration of other industries (non-manufacturing, non-agriculture, hereafter the NMNA sector) began much later.[1] Hence, our capital stock series for the NMNA sector relies on stronger assumptions and spans a shorter time period.

We discuss the reliability of our historical data sources. At several points in the analysis we need to make assumptions to derive our estimates. Whenever possible, we provide robustness checks to show the importance of these assumptions. For the most part, reasonable alternative assumptions lead to capital stocks quantitatively similar to the baseline. For some of our estimates we have outside evidence from alternative sources; we compare our estimates to these alternatives whenever possible. Finally, the available data allow us to construct estimates for some states, sectors, and years that we do not use in our papers, but which are available in our data. For example, although we choose not to use Washington, D.C. in our analyses, we can construct a complete data record for the manufacturing sector there. We describe the construction of this additional data as well.

Our work follows in the footsteps of many other researchers. Perhaps the most notable efforts were provided in a series of NBER volumes on the subject of capital and capital

[1] For example, the first Census of Business (now Economic Census) did not occur until 1929; the first Census of Transportation did not occur until 1963; and the Census of Finance, Insurance and Real Estate Industries was not separate from the Census of Construction until 1992. Given the late arrival of detailed state-level census data for these other sectors, we group them together and employ the alternative approach explained below.

formation in the US economy, organized by Simon Kuznets. In particular his overview volume and the volumes on capital in agriculture and manufacturing were useful precedents for our work (Kuznets 1961, Tostlebe 1954, Creamer, Dobrovolsky, Borenstein, and Bernstein 1960). The work described here expands on the NBER studies in four ways. First it adds new data up to the year 2000, whereas the data reported in the NBER volumes stops in 1950. Second, our capital figures are produced on a state basis, whereas theirs are typically on a regional or national basis. Third, our estimates of land in agriculture include a new correction for the relative productivity of land in different states. Finally, we measure a slightly different notion of the capital stock, more in line with modern definitions. Nevertheless the methods, discussions, and caveats in their work have been instrumental in our own, since many of the most challenging issues arise in the earlier portions of the sample and were also dealt with by them.

The rest of the paper proceeds as follows. Section 2 describes the construction of estimated physical capital in the agricultural sector, Section 3 describes the construction of estimated physical capital in the manufacturing sector, and Section 4 describes the construction of estimated physical capital in the NMNA sector. Section 5 describes the construction of productivity-adjusted land input in the agricultural sector. Finally, Section 6 gives the dates at which data become available for the various states, and Section 7 concludes.

2 Physical Capital in the Agricultural Sector

2.1 Data Sources, Coverage, and Reliability

Our primary sources for physical capital in the agricultural sector are the U.S. Censuses of Agriculture, 1850–1997, and the USDA Farm Balance Sheets (hereafter FBS), from 1960 onward. As a general principle our estimates of the physical capital stock use census data for 1850–1959 and FBS data for 1960–2000. The format and interpretation of estimates is largely compatible between these two sources, and indeed responsibility for the 1997 and subsequent Censuses of Agriculture was handed over from the Department of Commerce to the Department of Agriculture. At some points we rely on additional data collected in special reports to the Census of Agriculture. For more detailed data on the different types of farm assets we use the Farm Finance Surveys, which are supplements to the 1969 and 1978 Censuses of Agriculture, and the Agricultural Economics and Land Ownership Surveys, which are supplements to the 1987 and 1997 censuses.

The timing and frequency of censuses changes over time. The 1850 census marks the first serious attempt to measure the nation's physical capital stock, and is the starting point for our work. The census was collected decennially from 1850 to 1920; every five years from 1925 to 1950; every five years from 1954 to 1974; in 1978; and every five years from 1982 onward. FBS data were collected by the USDA annually from 1960 onward. For later years we can compare the two sources of data. For earlier years we rely solely on the census data.

Before proceeding to how we construct our data series, we discuss two historical concerns about the validity of the reported census data. First, our estimates of physical capital in the agricultural sector are all based on estimated values of various components of the physical capital stock; we do not use investment or the perpetual inventory method to calculate agricultural physical capital stocks at any point. However, there is an open question of what exactly the reported census values mean: does value refer to book value, reproduction value, market value, or some other concept? Later censuses eventually clarified that they wanted the market value, but earlier censuses were much less clear. The original census enumeration schedules provide little help. For instance, the 1850 instructions told enumerators to "include the actual cash value.. In this, as in all cases where an amount of money is stated, make your figures represent dollars" (U.S. Census Office 1900). Kuznets (1946) and Gallman (1986) disagree about whether this statement was generally answered using book value (Kuznets) or market value (Gallman). As Gallman notes, interpreting the figures as market values leads to estimates of the capital stock that closely agree with outside estimates derived from the perpetual inventory approach. Lindert (1988a) finds a close correspondence between census values and market values for a few states. We treat reported values as market values. However, we note that the original source material is inconclusive on this matter.

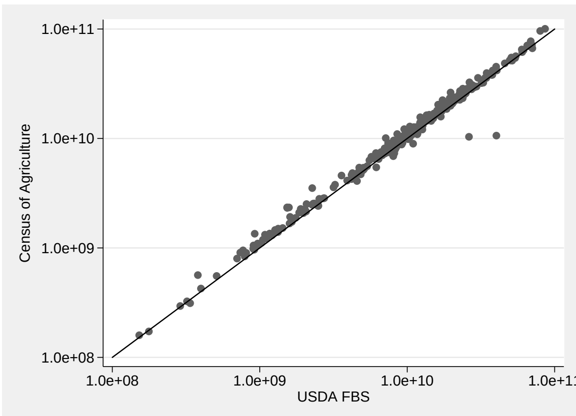
Our second historical concern is that the censuses have always constrained their attention to farms that meet a minimum size requirement. Farms that fail to meet this minimum size are not enumerated. For example, in 1870 the minimum size was three acres or $500 (1870 dollars) in farm sales. This figure changed most often during early censuses, and was relatively fixed from 1974-1997; see Hurley (1962) for an overview of the different definitions in early years. Given that there is a minimum farm size and that it changes by year, the concern is that we are underestimating total farm activity, and that the extent of underestimation varies from year to year. We have two responses to this concern. First, most of our statistics on the agricultural sector are drawn from the censuses, particularly land, capital, and output. Hence, our measurement of the agricultural sector is internally consistent, although it may not include some small farms. Second, we follow Tostlebe in

Table 1: Underlying Data Sources by Year for Physical Capital in Agricultural Sector

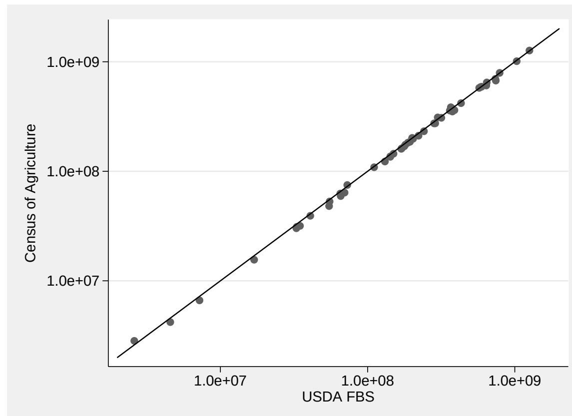
Year	Physical Capital Component			
	Farm Buildings	Machinery	Livestock	Crops
	Panel A: Relevant Measure Reported in Census of Agriculture			
1850–1890	Land/Buildings	Implements/Machinery	Livestock	None
1900–1925	Buildings	Implements/Machinery	Livestock	None
1930	Farm Buildings	Implements/Machinery	Livestock	None
1935	None	None	Livestock	None
1940	Buildings	Implements/Machinery	Livestock	None
1945	Land/Buildings	Implements/Machinery	Livestock	None
1950–1964	Land/Buildings	None	Livestock	None
1969+	Land/Buildings	Machinery/Equipment	None	None
	Panel B: Relevant Measure Reported in USDA Farm Balance Sheets			
1960 +	Farm Real Estate	Machinery/Motor Vehicles	Livestock	Crops

noting that a failure to enumerate small farms has an important impact on the measurement of number of farms, but less of an impact on acreage, sales, or capital. For example, the census responded to a change in farm size definition between 1997 and 2002 by retroactively calculating farm activity in 1997 under both the old and new definitions. Comparison of the figures shows that the change in the definition of a farm resulted in a 15.9% change in the number of farms, but only a 2.4% change in the total acreage of farms. Since we do not use the figures for the number of farms, changes in the definition will be less important for us. However, there are particular concerns about the coverage of Southern states in the 1870 (i.e., post-Civil War) census, as documented by Ransom and Sutch (1975). Hence, we make a point of not using 1870 in our level comparisons or as an endpoint in our growth comparisons, and we urge others to treat the 1870 figures with caution.

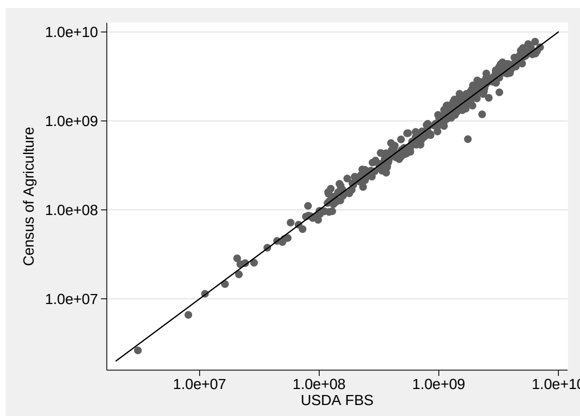
With those caveats in mind we turn to the available raw data. In a perfect world, the physical capital stock in the agricultural sector would include the value of farm buildings, such as barns or silos; of implements, machinery, and equipment, such as wagons and tractors; of livestock; and of crop inventories. We deviate from some other sources in excluding land from the physical capital stock, since our intent is to estimate the land value separately; see Section 5.[2] Table 1 gives the available raw data in the censuses and FBS pertaining to these four components of the physical capital stock. These data represent the primary sources of information for our estimates.



(a) Land, Buildings, and Dwellings



(b) Livestock



(c) Machinery and Equipment

Figure 1: Comparison of Data from Census and Farm Balance Sheets

For some years after 1960 we have measures of livestock, land and buildings, or machinery and equipment from both census and FBS sources. In these cases we can compare our two data sources. Figure 1 plots the census and FBS-measured variables against each other. Land and building values (Figure 1a) agree fairly closely with one another, with a couple of exceptions. The FBS farm real estate series for Nevada and Arizona both reveal

[2]Thus our definition is somewhat different from that in the seminal work of Tostlebe (1954), which includes land but is otherwise identical (p. 4). Otherwise, our results are quite similar, as we show in Section 2.7. Tostlebe also provides a wealth of other information on the changes in the composition or financing of the capital stock, which we do not repeat here.

irregularities (a one-year 25% decline in the value for the former, and a one-year 150% increase in the value for the latter). These irregularities are not present in the census data and seem implausibly large, so we use census values with some interpolation between census years instead. Figure 1b reveals no such irregularities for livestock data, although we are able to make the comparison only for one fairly early year (1964). Finally, Figure 1c shows more disagreement between the two sources for the value of machinery and equipment. As we will discuss below there are good reasons to expect this. For now we confine ourselves to noting that the four clustered outliers in Figure 1c represent four consecutive observations for North Carolina. The USDA reported value for machinery nearly doubles in a single year while the census value remains fairly constant. As before we find the USDA numbers implausible and use interpolated and actual census data instead.

We use these data to estimate the four components of physical capital in the agricultural sector. We proceed from the easiest to the most difficult component to estimate.

2.2 Livestock

Between the census and the FBS we have a complete time series for livestock by state. Further, the two sources overlap in one year, 1964. Comparison of the estimates of the two sources in that year reveals that they agree closely (see Figure 1b above). We combine the two sources to create a continuous series on livestock by state. There may be some incompatibility issues in early years since the date of enumeration has changed over time, and the value of livestock on farms varies predictably over the year. See for example the introduction to the 1940 census for further discussion of this issue.

2.3 Crops

The FBS include crop inventory values from the beginning, but the censuses have never collected data on crop inventories. Tostlebe (1954) generated series for crop inventories in each of ten agricultural regions. His method for constructing the inventory series involves multiplying the quantity of crops held in inventory by the prevailing market price for a number of crops. Some of the underlying raw data for his series was generated by himself and not published, so we cannot follow his footsteps to generate state-level series.

Instead, we use his region-level data to impute the value of crop inventories by state. To perform the imputation, we construct for each region the value of crop inventories relative to other reported capital in Tostlebe (1954).[3] We generate each state's crop inventory by

[3]Since we are working with the reported figures in Tostlebe's book, this means the value of crop inven-

multiplying the state's reported capital by the appropriate regional crop-capital ratio. This process may be reasonable since the grouping of states into agricultural regions is based on the presence of common soil types, weather patterns, and crop production, which are presumably key factors influencing the decisions about how much crop inventory to hold.[4] As a test of the imputation, we plot the time series of the "crop-capital" ratio time series for all states. The values from 1870–1950 come from Tostlebe; the values from 1960 onward come from the FBS, where we compute the same statistic. These series show remarkable continuity, with few large jumps when switching from Tostlebe's independent, regional data to the USDA's state data. We conclude that using Tostlebe's data is reasonable. Further, we note that these series are relatively smooth in the 19th century. This fact motivates us to fit a time trend, project the "crop-capital" ratio back to 1850, and use it to compute crop inventory values for 1850–1860.

This procedure yields state-level estimates of crop inventories for 1850–1950, which we combine with FBS data from 1960 onward. Since this series involves a number of assumptions, we also produce an alternative *no crops* series which excludes crop inventories for all years, including the period 1960–2000. We check the robustness of our results to this alternative below.

2.4 Machinery

For machinery and equipment we combine census data on implements and machinery with FBS data on farm machinery and motor vehicles. We noted above that there is overlap between census and FBS data from 1964 onward, and that there is some disagreement between these two sources. The likely reason is that the USDA estimates and removes the fraction of farm machinery and equipment usage that is for non-farm (personal) purposes. For example, a truck can be used on the farm but also to drive to a restaurant for dinner. Census estimates include the whole value of the truck, while FBS estimates include a fraction of the truck's value, based on its relative usage for farm activity.

We combine the census and FBS data to form a single estimate of machinery, equipment, and implements on the farm. This step could result in a discontinuous series if the personal usage adjustment in the FBS is quantitatively important in the early years. Figure 2 suggests that this is not the case. It compares the census and FBS measures of machinery

tories relative to the value of land, buildings, machinery, and equipment. We construct and use the same value in our data for the crop inventory imputation process.

[4]For example, one region is the Corn Belt, the major corn-growing states: Illinois, Indiana, Iowa, Missouri, and Ohio. See p. 8 of his book for the definition of regions. Alaska and Hawaii are excluded from all regions.

and equipment for the earliest year possible, 1969. The two series agree closely, other than the obvious outlier of North Carolina. In later years there is a growing discrepancy, with census figures systematically higher than FBS figures; this fact suggests that personal usage has become more important over time. We take away that our method of using the FBS from 1960 onward has the desirable feature of introducing the adjustment for personal use of farm capital without introducing any discontinuity with earlier, census-derived estimates.

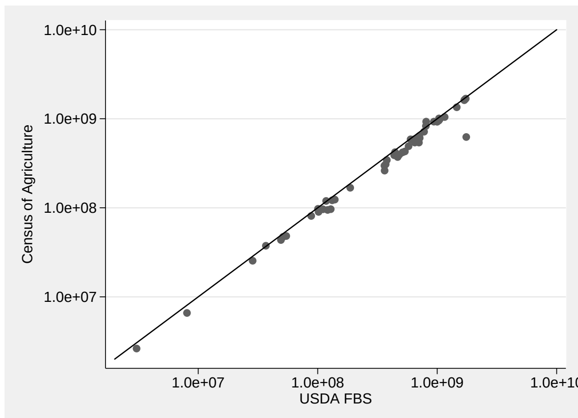


Figure 2: Comparison of Machinery and Equipment from Two Sources, 1969

The censuses collected no information on implements, machinery, or equipment in 1935 and from 1945-1959. We interpolate the series for these years.

2.5 Farm Buildings

Farm building value is the most challenging component of the agricultural sector physical capital stock to estimate. For most years we have one aggregate measure, called land and building value in the censuses and farm real estate in the FBS. We established in Figure 1a that the two series are closely linked, so we combine them to form one continuous series. However, the naming of this series is somewhat misleading. What the census calls "land and building value" actually includes the value of land, farm buildings, and farm dwellings. Below we refer to this aggregated series as "value of land, buildings, and dwellings" for clarity. Our goal is to include only the value of farm buildings; we include the value of farm dwellings in the finance, insurance, and real estate sector, and account for land separately.[5]

[5] Tostlebe includes the value of farm dwellings in his estimates. His argument is that the farm house was the place of some home production as well of residence for the family, and that it is difficult to disentangle the fraction used for each purpose. However, changes in the scale of farming since the time of his publication have reduced the use of farm dwellings for home production and made this inclusion less desirable.

We form our estimate of the value of farm buildings in two steps, starting with the aggregate series that includes farm buildings, farm dwellings, and land values. In the first step, we separate the value of farm buildings and dwellings from the value of land. We perform this step separately because there are more data on the decomposition of land versus farm buildings and farm dwellings, and because the relative value of land in the total fluctuates greatly. In the second step we separate farm building value from farm dwelling value.

The 1900–1940 censuses report the value of farm buildings and dwellings separately from land. For these years we use the direct measurements of farm building and dwelling value. Special reports to the 1969, 1978, 1987, and 1997 censuses include enough information to estimate the fraction of land, building, and dwelling value that is farm building and dwelling value. We interpolate this fraction from 1900–1997, including the longer break between 1940 and 1969. We project the 1997 value forward to 2000, and the 1900 value back to 1850.[6] This method of projecting building value before 1900 is also used in Lindert (1988b), although his interest is to produce an estimate of land value rather than of building and dwelling value.[7] Figure A2 shows the state-year variation in the ratio of land and buildings to land, buildings, and dwellings for all years where we have data.

We then construct building and dwelling value as

$$buildings + dwellings = (land + buildings + dwellings) \times \widehat{\frac{buildings + dwellings}{land + buildings + dwellings}},$$

where the hat indicates that the variable is a projection or an interpolation of the relevant fraction. We use this method to construct the value of buildings and dwellings for years outside the range 1900–1940.

Data for the second step are scarcer. We have direct observations on the value of farm buildings in the 1930 census. We use this observation for our series. Special reports to the 1987 and 1997 censuses include enough information to estimate the fraction of building and dwelling value that is building value as well. Figure A3 shows the data for each year and state for which it is available. There are clear state differences but not time trends, although of course this could be obscured by missing data. We construct the average proportion of farm building and farm dwelling value that is farm building value for each state.

[6] With the exception of Wyoming, where the 1900 value is an obvious outlier; for them, we project the 1910 value back instead.

[7] For further detail see the working paper version (Lindert 1988a).

We then construct building value as

$$buildings = (buildings + dwellings) \times \frac{\widehat{buildings}}{buildings + dwellings}$$

for years other than 1930, where the hat indicates that the variable is the state average over the observations available. When combined with our estimates of livestock, crop inventory, and machinery values, we have a complete estimate of the physical capital stock in the agricultural sector.

2.6 Alaska, Hawaii, and Washington D.C.

Alaska and Hawaii pose special challenges because of irregular data collection. Before 1960 we generally rely on census data covering the components of physical capital, but little information is given for Alaska or Hawaii. Land and building values were reported in 1900, 1910, 1920, and 1930 for both states, and for Alaska but not Hawaii in 1959. The only other reported value for physical capital components is an estimate of livestock in Hawaii in 1959. Hence, we cannot rely on direct observation of physical capital in the censuses. Fortunately, the data from the FBS are quite similar to what are available for other states. We discuss first our approach to the FBS and physical capital for 1960 onward, and then our approach to years where only partial census data are available.

FBS data include both Alaska and Hawaii from 1960 onward, but the data on crop inventories is sporadic. Fortunately crop inventories for these states are small, amounting to less than 1% of the total physical capital stock in all years where data are available. The values of livestock and machinery can both be used as reported. We need to estimate the value of farm buildings from the value of farm real estate reported. We follow the same basic approach as with the 48 contiguous states, but with somewhat less data. We estimate the value of farm buildings and dwellings as

$$buildings + dwellings = (land + buildings + dwellings) \times \frac{\widehat{buildings + dwellings}}{land + buildings + dwellings}.$$

The proportion is known from the 1969 (Hawaii only), 1978, 1987, and 1997 special reports to the census. We project the earliest value back in time, linearly interpolate between the earliest value and 1997, and project the 1997 value forward to 2000. We estimate the value of buildings as

$$buildings = (buildings + dwellings) \times \frac{\widehat{buildings}}{buildings + dwellings}$$

where the proportion is the average of the values observed in 1987 and 1997. Again, everything is similar to the contiguous 48 states, except that we lack the earlier 1900-1940 census data to augment these exercises.

The only consistent data prior to 1960 are the total value of land, farm buildings, and farm dwellings from the census in select years. To extend our measure of physical capital earlier, we estimate each state's agricultural capital stock using the reported data as:

$$capital_{it} = (land_{it} + buildings_{it} + dwellings_{it}) \times \frac{capital_{i,1960}}{land_{i,1960} + buildings_{i,1960} + dwellings_{i,1960}}$$
$$\times \frac{capital_t/(land_t + buildings_t + dwellings_t)}{capital_{1960}/(land_{1960} + buildings_{1960} + dwellings_{1960})}$$

Variables without subscript i denote national totals for the 48 contiguous states. Capital is our constructed measure of physical capital. The raw data for our estimate of capital is the land, building, and dwelling value in Alaska and Hawaii. We multiply by two terms. The first term is the ratio of capital to land, building, and dwelling value in 1960 for each state; we choose 1960 since it is the first year for which these data are available. This term captures the large cross-state variation in capital intensity as opposed to land or dwelling intensity. The second term is the aggregate ratio of capital to land and building value in year t, as compared to the aggregate ratio of capital to land and building value in 1960. This term captures trends or cyclical variation in capital that affect the nation as a whole. This approach gives us approximations for the physical capital stock in 1900, 1910, 1920, and 1930; we interpolate from 1930 to 1960.

Although we do not use data for Washington D.C. in our papers, it is possible to estimate their physical capital stock in the agricultural sector for early years. They were enumerated in the 1850-1950 censuses, but were not separately enumerated in any subsequent censuses or FBS data. In the early censuses the same data were collected in Washington D.C. as in the other states; see Table 1 for a reminder. We can construct an estimate of their physical capital from 1850 to 1945 using this data. We combine census data on livestock and machinery with our derived estimate of farm building value. Again, we estimate farm building value in two steps. We first estimate the proportion of building, dwelling, and land value that is building and dwelling value (drawing on 1900-1940 data, but not 1969, 1978, 1987, or 1997 data). We then estimate the proportion of building and dwelling value that is building value (using only 1930 data). Our data set includes the physical capital stock constructed along these lines.

2.7 Data and Comparisons

Figure A4 presents our estimates of the time series of physical capital in the agricultural sector by state. We include the earliest possible estimates for Alaska and Hawaii although they may not be reliable.

We faced two major choices in the construction of our baseline estimates. First, we included crop inventories although doing so required some imputations; second, we used FBS rather than census data in later years where both were available. We have computed two alternative series: the *no crops* series which excludes crop inventories, and the *census only* series which uses census data throughout. Figure 3 compares their properties at the national level. The three series agree quite closely, reflecting the general close agreement of census and FBS data, and the relatively small role for crop inventories, particularly in later years.

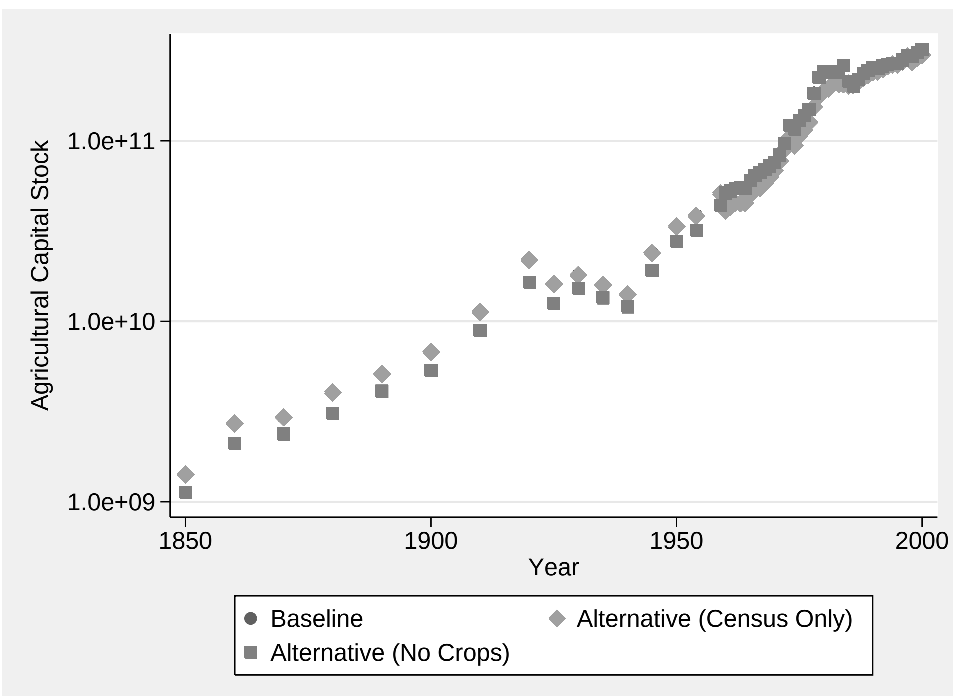


Figure 3: National Agricultural Capital Stock Computed Under Alternative Assumptions, 1850–2000

We can sum our data across states and compare them to two outside sources at the national level. Tostlebe (1954) includes estimates of the physical capital in the farm sector from 1870–1950. His figures differ from ours in that they include the value of land, but we can still assess whether his figures show similar trends. Second, the Bureau of Economic Analysis produces estimates of the aggregate capital stock by SIC industry for 1947-1997 as a part of the National Income and Product Account data.[8] We use their estimate of agricultural sector capital for later years. Figure 4 compares our national results to these benchmarks. Our figures show the same patterns as Tostlebe's, but with the expected level difference. Our figures agree very closely with those of the BEA, except that our figures show a modest decline in agricultural capital in the 1980s that does not appear in the BEA figures.

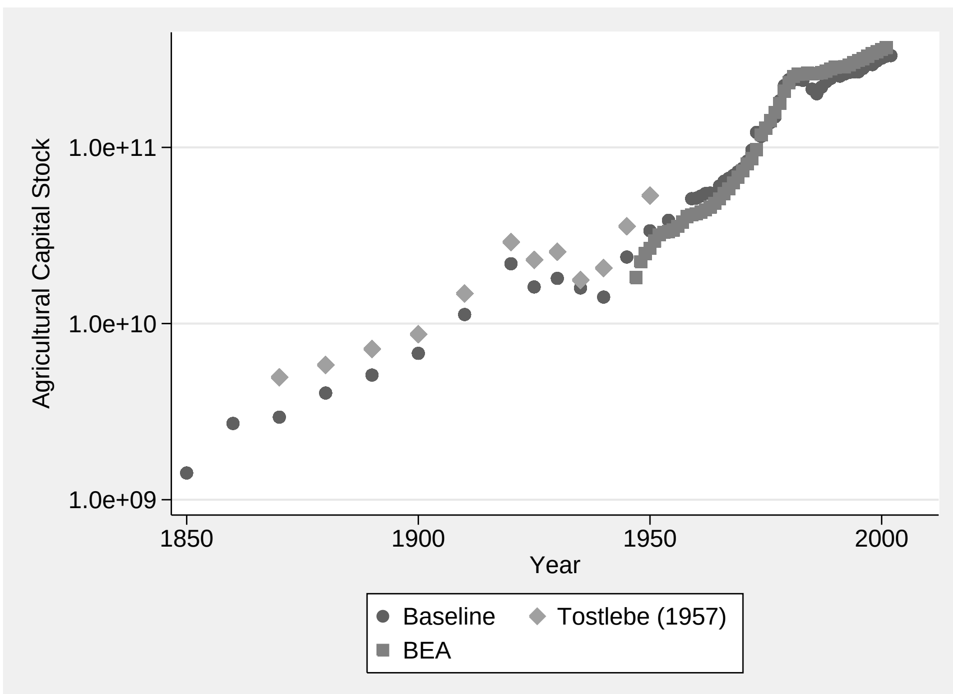


Figure 4: National Agricultural Capital Stock Compared to Benchmarks, 1850–2000

We can also compare our data to Tostlebe's at the regional level by aggregating our states into his ten regions. Figure 5 compares the series at the region level from 1870–1950. As in the aggregate the series agree closely, with Tostlebe's always slightly higher because he includes land value in the physical capital stock.

[8]Currently available online at `http://www.bea.gov/National/FAweb/AllFATables.asp`; Table 3.1ES.

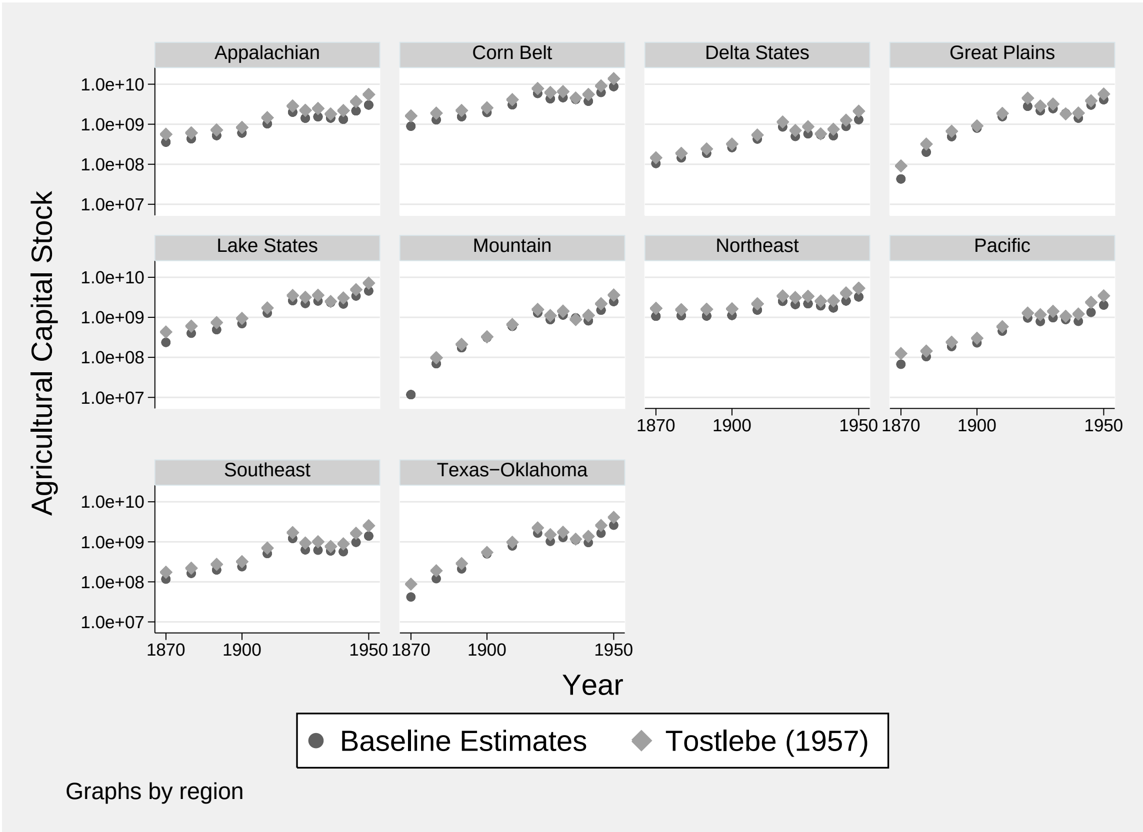


Figure 5: Regional Agricultural Capital Stock Compared to Benchmarks, 1870–1950

3 Physical Capital in Manufacturing

3.1 Data Sources, Coverage, and Reliability

The data used to construct physical capital in manufacturing come from the U.S. Census of Manufactures, 1850-1997, and the Annual Survey of Manufactures (hereafter ASM) which is conducted in intercensal years from 1949 onward. As a general principle our estimates of the physical capital stock use census data whenever available, and use ASM data to fill in between censuses from the 1950s onward. The format and interpretation of statistics between the two sources is essentially identical since the ASM was designed to provide continuous coverage on key statistics between census years.

The timing and frequency of censuses shifts over time. As with agriculture, the 1850 census marks the first serious attempt to measure the nation's physical capital stock, so we start our work in 1850. The census was collected decennially from 1850 to 1900; every five years from 1904 to 1919; every two years from 1921 to 1939; not again until 1947; in 1954, 1958, and 1963; and every five years from 1967 to 1997. The ASM was explicitly started in 1949 to fill in between censuses, so we have annual data from 1949 onward, but only census data before then. Since the ASM is never taken in the same year as a census we cannot compare data from the two sources.

As with the Censuses of Agriculture, there are some historical concerns about the validity of the data worth discussing. The primary concern in this area arises from claims made by Francis A. Walker, Superintendent of the Census in the nineteenth century, that the figures on capital were "flagrantly false." If the chief of the organization that produces the figures places so little confidence in them, then perhaps it is not worth pursuing their use as data? However, Creamer, Dobrovolsky, Borenstein, and Bernstein (1960) produce a number of checks that in general tend to validate rather than repudiate the figures in the historical censuses.[9] We rely on their work as a defense of the general reliability of the statistics at hand.

The second relevant question is again whether the values reported represent book or market value. Creamer, Dobrovolsky, Borenstein, and Bernstein (1960) interpret values as book value, following Kuznets; again, Gallman (1986) disagrees. We treat values as market values. Creamer, Dobrovolsky, Borenstein, and Bernstein (1960) also suggest concerns that mergers in the early 20th century may have changed how assets were reported given the same underlying question; they find that this does not seem quantitatively significant. Finally, the censuses typically do not enumerate small manufacturing firms or establishments, based on sales. Our response is the same as with agriculture: we consistently exclude small establishments in our measures of inputs and outputs; and further, excluding small establishments affects the number of establishments much more than the value of physical capital or output in the manufacturing sector.

A further concern comes from our use of investment data during World War II. Higgs (2004) raises questions about wartime investment. Capital formed during this period often had little civilian use. Further, high rates of utilization may mean that the constant depreciation assumption is not appropriate for this period. However, problems with government investment and capacity utilization also arise in the cross-country development and growth accounting literatures; see Cubas (2010) on the former. Hence, these problems are by no means unique to our work. Further, as Higgs himself notes, he is able to identify the "direction, if not the precise magnitude, of some strongly warranted adjustments." For our purposes the magnitude is of course quite important, even if we decided to make such an

[9]See their Appendix 1. For example, they find that the 1919 census figures agree closely with balance sheet data available from alternative sources.

Table 2: Underlying Data Sources by Year for Physical Capital in the Manufacturing Sector

Years	Measures of Capital or Investment
1850–1880	Physical and Working Capital
1890–1900	Physical Capital
1904–1919	Physical and Working Capital, Horsepower
1921–1937	Horsepower
1939	Horsepower, New Capital Expenditures
1947–1972	New Capital Expenditures
1977–1997	New Capital Expenditures, Gross Book Value

adjustment.

With those caveats in mind we turn to the available data. Most of our data measure buildings and machinery, the two main components of physical capital in manufacturing. Table 2 gives the information available in the censuses and ASM pertaining to physical capital in manufacturing. Roughly speaking, the time series can be split into four eras. From 1850–1919 we have estimates of the stock of physical capital in the census. From 1919–1939 the only data available are measures of installed horsepower. Censuses from 1939 onward, and ASMs from 1949 onward, include data on new capital expenditures, which we can use with perpetual inventory methods to estimate physical capital. Finally from 1977 onward the censuses include measures of gross book value of capital. These data represent the primary sources of information for our estimates and for robustness checks.

We use these data to estimate the physical capital stock in manufacturing in four eras. We proceed chronologically. We construct a baseline series and two additional series for robustness when we face significant choices about the data to use.

3.2 Physical Capital in Manufacturing, 1850–1919

The censuses collected a general measure called capital from 1850 to 1919. However, their notion of capital is broader than the modern one, and corresponds more closely to total assets.[10] The 1890 and 1900 censuses enumerate four types of capital: land; buildings; machines, tools, and implements; and a fourth category called variously live assets or cash and sundries. This final category is a combination of working capital and inventory; it

[10] For more details on the information given and a comparison to other sources such a the Internal Revenue Service *Statistics of Income*, see Creamer, Dobrovolsky, Borenstein, and Bernstein (1960), p.12.

includes “cash on hand, bills receivable, unsettled ledger accounts, raw materials, stock in the process of manufacture, finished products on hand, and other sundries.”[11] Creamer, Dobrovolsky, Borenstein, and Bernstein (1960) include the entire value of this final category in their tabulations of physical capital in manufacturing. One improvement of our work is that we attempt to exclude this category, in line with modern notions of physical or fixed capital. We also exclude land, both because this practice is in line with previous research that uses perpetual inventory methods, and because our data from 1919 onward do not include land in any way. Our measure of physical capital in manufacturing then corresponds to buildings, machines, tools, and implements in the manufacturing sector.

In the 1890 and 1900 censuses the available data correspond closely to our preferred estimate of physical capital in manufacturing. Our measure of the physical capital stock in these years is then the value of buildings, machines, tools, and implements, excluding the value of live assets or cash and sundries. For other years only the total capital stock (including working capital) was collected and reported. Figure A5 shows the relationship between our preferred estimate of capital and reported capital for 1890 and 1900, by state. There seems to be differences across states but no consistent trend over this time period. This leads us to estimate our preferred value of capital in other years using reported capital and the state-average ratio of our preferred estimate of capital to reported capital as

$$capital = reportedcapital \times \frac{\widehat{capital}}{reportedcapital}$$

where we estimate the fraction of reported capital that is our desired estimate of capital using the average of 1890 and 1900 data. Using this formula yields our estimate for 1850–1880 and 1904–1919.

3.3 Physical Capital in Manufacturing, 1919–1939

The census stopped collecting measures of the physical capital stock after 1919, and did not restart until they asked firms about their gross book value in 1977. Starting in 1939, they asked about new capital expenditures, which can be used with the perpetual inventory method to estimate physical capital. But between 1919 and 1939, the only information on physical capital in manufacturing is various measures of horsepower in place. Indeed, this gap in coverage is probably what lead Creamer, Dobrovolsky, Borenstein, and Bernstein (1960) to produce national figures for physical capital stock in manufacturing; at the time

[11] 1900 Census report, volume 8, part 2, p. viii

they compiled their volume, it was not clear that there was any useful state-level data after 1919. Instead, they turned to the national capital stock figures reported by the IRS.

Our baseline approach is to use the information on horsepower and a fitted relationship between horsepower and total physical capital to approximate the physical capital stock by state from 1919 to 1939. Horsepower is an imperfect measure of the capital stock: it really captures (one aspect of) machines but not buildings, tools, or implements. However, horsepower can be useful if there is a reasonably stable relationship between horsepower in place and total physical capital in manufacturing. We proceed by estimating a baseline series under this assumption. We also construct an alternative series that interpolates over the years 1919–1939. We then compare the two series.

The censuses actually present a number of different measures of the installed horsepower capacity available in manufacturing. Further, no single measure of horsepower spans the entire time period. Fortunately, there exists a series of overlapping, highly correlated measures of horsepower that we can use to construct a continuous horsepower variable. Our aggregated variable is constructed from three underlying measures:

1. In 1904, 1909, 1914, and 1919 information was collected on primary horsepower.

2. In 1919, 1925, 1927, and 1929 information was collected on horsepower.

3. In 1929 and 1939 information was collected on horsepower from electric motors.[12]

These series are highly correlated but have different scales. A log-log regression of primary horsepower on horsepower in 1919 yields an R^2 of 1.0000; a log-log regression of horsepower on horsepower from electric motors in 1929 yields an R^2 of 0.9832. We use the predicted values from these log-log regressions to form a single, composite horsepower series that spans 1904–1939. We then multiply the horsepower by a price series for machinery and equipment derived from Kuznets (1961).[13]

Our series for horsepower overlaps with physical capital between 1904 and 1919. Figure 6 compares the two series in logs and finds that they are strongly correlated. Table 3 shows different regressions that aim to fit the relationship between the two series. State fixed effects seem necessary but do a good job of fitting the series, so we use the regression with state fixed effects (column (5)) to predict physical capital out of sample from 1919 to 1939. Note that adding year fixed effects does not improve the fit of the model between 1904 and 1919.

[12] A separate variable, horsepower from prime movers, is also available over the same time period, but is slightly less correlated with the previous measures of horsepower.

[13] Table R33; we divide the nominal by real value of gross producer's durables in each year to recover the price series. A similar approach is taken by Collins and Williamson (2001).

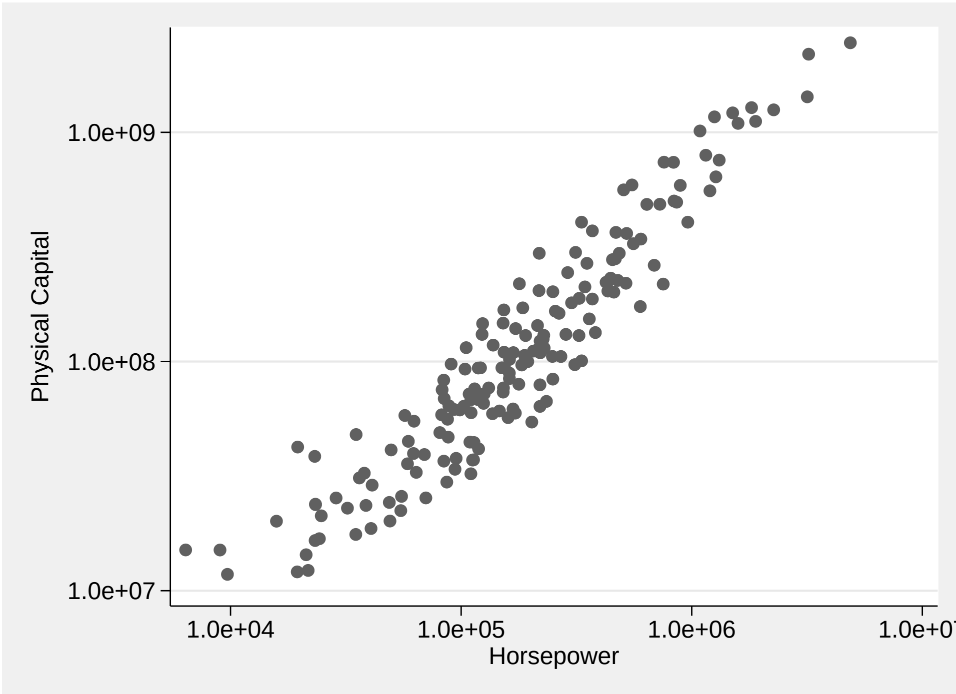


Figure 6: Comparison of Physical Capital to Measured Horsepower in Manufacturing

However, electrification spread more widely in the 1920s and 1930s, changing the structure of manufacturing firms and the equipment that they employed (David 1990, Tostlebe 1954). Electrification could in turn have disrupted the relationship between horsepower and physical capital. We explore this possibility by constructing an alternative series that simply interpolates the available capital stocks from 1919 to 1939.

3.4 Physical Capital in Manufacturing, 1939–1977

Data on new capital expenditures begins in 1939. The data are reported in the 1939 and 1947 censuses, and in the ASM generally annually from 1949 onward.[14] We use investment data and the perpetual inventory approach to construct physical capital estimates from 1939 onward.

Since we are missing data between 1939 and 1947, in 1948, and for a few other states and years, we need to interpolate investment over missing years. Then we have a comprehensive investment series from 1939 onward. We can compare this investment series to an outside

[14]The ASM do not report values for small states in early years, but are nearly comprehensive by 1955. Data were not collected in 1959 or 1960. We also smooth one extreme value, for Ohio 1996, where the reported investment is larger than the combined total of Texas and California; we simply interpolate from 1995 to 1997.

Table 3: Fitted Relationship Between Capital and Horsepower

	(1)	(2)	(3)	(4)
Horsepower, Log	0.921***	0.821***	0.922***	0.836***
	(0.0178)	(0.0205)	(0.0309)	(0.0840)
Constant	7.372***	8.324***	7.300***	8.139***
	(0.211)	(0.271)	(0.395)	(0.992)
Year Dummies	No	No	Yes	Yes
Region Dummies	No	No	Yes	No
Region-Year Int	No	No	Yes	No
State Dummies	No	Yes	No	Yes
Observations	196	196	196	196
R^2	0.932	0.988	0.949	0.989

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

benchmark, NIPA investment data. NIPA data are aggregated to the national level, so we sum our state investment series across states. Figure 7 compares the two series in logs. The NIPA investment series of course fluctuates more between 1939 and 1947, when we are smoothly interpolating. From roughly 1950 onward the two series agree closely. We scale state investment in every year so that the state total matches the NIPA aggregate. This change has the largest impact in the 1940s and is relatively less important afterwards; after 1950, the largest difference is 23% and the average difference is 5%.

Given a series for investment we can form physical capital in manufacturing using the standard perpetual inventory formula:

$$K_{i,t+1} = (1 - \delta)K_{i,t}\frac{P_{t+1}}{P_t} + I_{i,t}.$$

We use a standard depreciation rate of $\delta = 0.06$ annually. P_t is the prevailing price of capital goods at time t; since our estimates are of the nominal value of the capital stock, we adjust each year for the change in the price of the existing stock of capital goods. We combine data from Kuznets (1961) for 1872–1948, and from Cummins and Violante (2002) for 1948–2000. Cummins and Violante (2002) extend the quality-adjusted capital goods

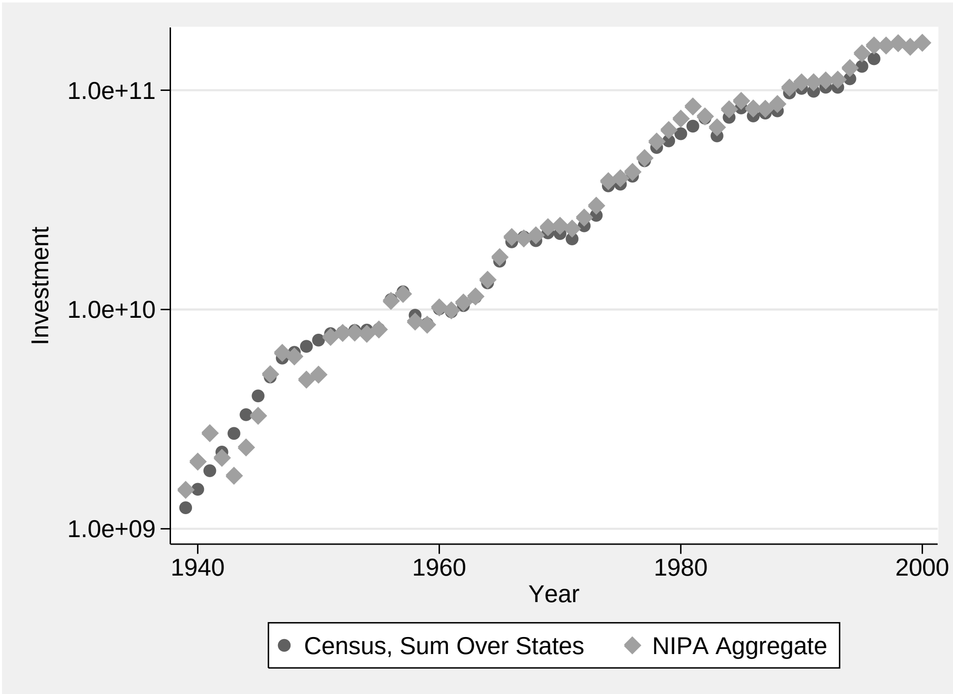


Figure 7: Comparison of Census and NIPA Investment, Aggregate

price series first developed by Gordon (1990). The earlier price series by Kuznets does not include the same quality adjustment; to our knowledge, no such series exists.

Annual investment data begin in 1939. In the *baseline* series we have an estimate of the physical capital stock in 1939 based on the horsepower in that state for 1939. In this case, we use the perpetual inventory method from 1939 onward to estimate physical capital by state and year. We also construct an alternative *steady state* series that does not use the horsepower information available from 1919 to 1939. Instead we follow the development accounting literature and construct the initial physical capital stock as

$$K_{1939} = \frac{I_{1939}}{\delta + g_i}$$

where g_i is the average growth rate of investment in state i (Caselli 2005). This starting condition is justified as the steady-state capital stock in the Solow model. It is standard in the development accounting literature to use the growth rate of investment near the beginning of the sample to estimate g_i. However, in this case the buildup for the war caused high growth in investment after 1939 that does not accurately depict the status of investment over, say, 1930-1939. Instead we set $g_i = 0$ for all states, guided by the fact that national investment figures in 1939 had only just recovered to 1930 levels; most of the intervening years were indeed much lower. Our estimates of physical capital for the 1930s and 1940s will be affected by this choice. However Caselli (2005) provides evidence that the effect of the initial condition diminishes rapidly over time.

Given an estimate of physical capital in 1939 and investment for 1939 onward, we can use the perpetual inventory method to construct physical capital for all years up to 1997. Our *baseline* series and *steady state* series do exactly this. However, in 1977 an alternative data series becomes available. Hence from 1977 onward we construct a second alternative series using this new data.

3.5 Physical Capital in Manufacturing, 1977–2000

In both the *baseline* series and the *steady state* series we rely only on capital expenditure data from 1939 onward. The 1996 ASM was the last source to ask for data on new capital expenditures; beginning with the 1997 census, the data was capital expenditures on new and used capital. The survey questions asked for expenditures on used capital but not sales of used capital, raising concerns about double-counting sold used capital. Since this occurs near the end of our sample, we stop our investment series in 1996. We use this series to construct capital to 1997, and then assume capital in each state continues to grow from 1997 to 2000 at the rate it did between 1992 and 1997.

In the 1977–1997 censuses firms were asked to report the gross book value of their physical capital, which is then reported by state. Gross book value suffers from well-known problems as a measure of the physical capital stock. It does not account for depreciation of the capital stock, which tends to bias values upward. Hence, we choose not to use gross book value in our *baseline* series. Still, it offers a useful check on our results in the last twenty years of the sample, particularly since it is derived from a stock rather than a flow measure. Hence, we construct a second alternative *gross book value* series which uses gross book values as an alternative measure of the physical capital stock from 1977 onward.

For the *gross book value* series we use investment data to construct physical capital until 1972. We use gross book value for census years between 1977 and 1997. We interpolate other years between 1972 and 1997, and we assume that capital in each state continues to grow from 1997 to 2000 at the rate it did between 1992 and 1997.

3.6 Alaska, Hawaii, and Washington D.C.

Until Alaska and Hawaii achieved statehood in 1959, there is almost no record of their capital stock or investment. The capital stock was measured for Alaska in 1890 and 1900, and for Hawaii in 1900. There is no other measure of their capital stock, no measures of installed horsepower, and no measures of investment until 1960. From 1960 onward we have annual investment with a few gaps, none exceeding four years; we interpolate investment through these gaps. From 1977–1997 we have gross book value of capital as in other states.

To construct our *baseline* series for Alaska and Hawaii, we construct the physical capital stock in 1960 using the steady state physical capital equation from the Solow model, with the average growth rate of investment from 1960-1980. We use the perpetual inventory equation to estimate physical capital from 1961 onward. Prior to 1960 we have little data, except for occasional observations on manufacturing value added. From 1960 onward the capital-output ratio is relatively consistent for the two states, at around 1.5 in Alaska and 1 in Hawaii. We project physical capital back using

$$K_{it} = Y_{it} \widehat{\frac{K_i}{Y_i}}$$

where the capital-output ratio for each state is taken as the observed average from 1960 onward. This methodology yields estimates of physical capital back to 1890 for Alaska and 1900 for Hawaii, although we do not use all of this estimated series in our work.

Given the lack of data, we do not compute a series based on horsepower data for Alaska or Hawaii. However, we do compute the *gross book value* series, which omits the physical capital from the perpetual inventory equation after 1972, and uses instead the gross book value of capital from the 1977–1997 censuses.

Manufacturing activity in Washington, D.C. has always been enumerated as in other states, so we provide the estimates for completeness.

3.7 Comparison to Outside Sources

Figure A6 presents our estimates of the time series of physical capital in the manufacturing sector by state. We include the earliest possible estimates for Alaska and Hawaii although they may not be reliable.

We faced two major choices in the construction of our baseline estimates. First, we use perpetual inventory rather than reported gross book value from 1977 onward; and second, we imputed the capital stock from reported horsepower rather than interpolating over the available capital stock data. We have computed two alternative series: the *gross book value* series which uses gross book values of capital from 1977 onward; and the *steady state* series which computes 1939 capital from the Solow steady state assumption, and then interpolates physical capital from 1919 to 1939. Figure 8 compares their properties at the national level. The *gross book value* series is essentially identical to the baseline. However, the *steady state* series shows slightly slower growth in the capital stock during the Great Depression.

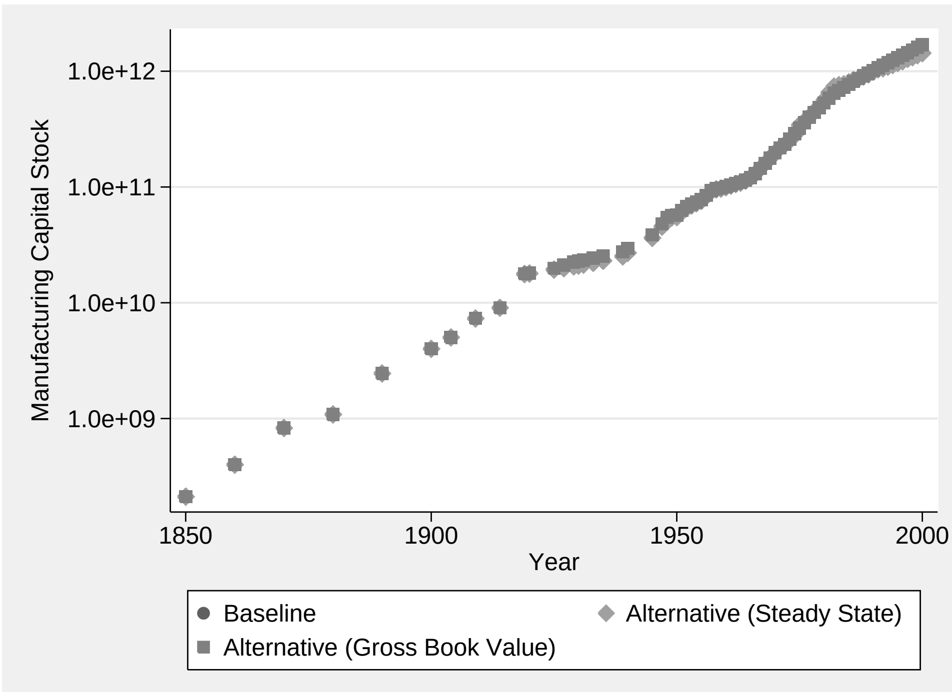


Figure 8: National Manufacturing Capital Stock Computed Under Alternative Assumptions, 1850–2000

We can compare our national data two outside sources. Creamer, Dobrovolsky, Borenstein, and Bernstein (1960) includes estimates of the physical capital in the manufacturing sector from 1870–1950. Their figures differ from ours in that they include the value of working capital, but we can still assess whether their figures show similar trends. This step is particularly valuable between 1919 and 1939. During this time period we rely state-level reports of horsepower to generate our physical capital series, while Creamer, Dobrovolsky, Borenstein, and Bernstein (1960) use national IRS statistics on physical capital. Second, the Bureau of Economic Analysis produces estimates of the aggregate capital stock by SIC industry for 1947–1997 as a part of the National Income and Product Account data.[15]

[15]Currently available online at `http://www.bea.gov/National/FAweb/AllFATables.asp`; Table 3.1ES.

We use their estimate of manufacturing sector capital for later years. Figure 9 compares our national aggregate results to these benchmarks. Our figures are generally similar to Creamer et al's, although lower as expected. The main difference is that their series shows a more pronounced dip during the Great Depression than does ours. Our figures are quite similar to the BEA figures from 1947 onward. On the whole, we take comfort in the general similarity of our series to outside benchmarks.

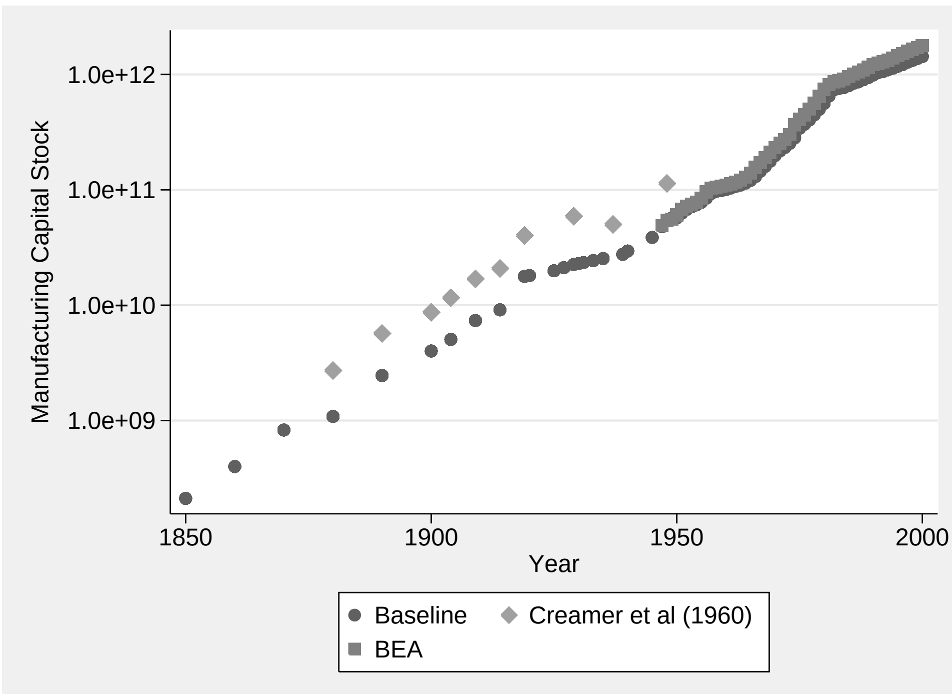


Figure 9: National Manufacturing Capital Stock Compared to Benchmarks, 1850–2000

4 Other Physical Capital

4.1 Data Sources, Coverage, and Reliability

In terms of broad SIC codes, there are seven industries not covered so far: mining; construction; transportation and utilities; wholesale and retail trade; finance, insurance, and real estate; services; and government.[16] Collectively, we refer to these as the non-manufacturing,

[16] In later years the available data distinguish between wholesale trade and retail trade, but before 1958 they did not, so we merge them into a single category. BEA capital data distinguish somewhat finer industries than do the value added, wage, and salary data, such as separating durable and nondurable manufacturing. Since we have to combine the two data sources we ignore these distinctions.

non-agriculture (NMNA) sector. For these seven industries we lack detailed state-level data and historical data. Our approach can be summarized as follows. The ideal way to construct state i capital at time t is to sum across the pertinent industries j:

$$K_{it} = \sum_j K_{ijt}$$

but we lack data on K_{ijt}. Instead, we construct state capital as:

$$K_{it} = \sum_j \frac{K_{ijt}}{Y_{ijt}} Y_{ijt} \approx \sum_j \frac{K_{jt}}{Y_{jt}} Y_{ijt}. \tag{1}$$

Hence, we use aggregate data on the capital-output by industry and the state composition of value added by industry to impute each state's capital stock for the NMNA sector.[17] Doing so requires ignoring any state-level variation in capital-intensity within an SIC industry, an assumption previously made by Garofalo and Yamarik (2002) in their work on regional convergence from 1977 to 1996.

Our data for this section all come from the National Income and Product Accounts. Table 4 summarizes the available data. For Y_{ijt} we use value added by state and industry, available from 1963–1997. Data are not available earlier and 1997 marks the last year of data in the SIC industry classifications (and the beginning of NAICS classifications). From 1929–1962 and 1998–2000 we have wage and salaries by state and industry; we show below that there is a close relationship between wages and salaries and value added that enables us to estimate value added reasonably well for 1929–2000. Before 1929 no data are available that would allow us to estimate value added reliably. NIPA provides capital stock by industry measures from 1947–2000, with a longer sequence for the government spanning 1929–2000. Before 1947 for most industries, and before 1929 for the government, there are no data on physical capital available.

We make two changes to the NIPA industry categories so that the estimates in the NMNA sector will complement our previous estimates in the agricultural sector. First, NIPA includes the stock of farm housing and the associated service flow in the agricultural sector, but we move it to finance, insurance, and real estate (FIRE) to be consistent with our earlier estimates. Second, we move the activity in fishing, forestry, and agricultural services to the service sector, again to be consistent with our earlier agriculture-only estimates. NIPA provides enough detail for these categories to make both switches for all the relevant

[17] Since we are focusing on the 1929–2000 period, we did not use the seminal work of Jorgenson, Gollop, and Fraumeni (1987), which produces sectoral capital measures from 1948 to 1979, inclusive.

Table 4: Underlying Data Sources by Year for Physical Capital in the NMNA Sector

Years	Measure of Capital by Industry	Measure of Value Added by State and Industry
1929–1946	Capital Stock (Gov't Only)	Salary and Wages
1947–1962	Capital Stock	Salary and Wages
1963–1997	Capital Stock	Value Added
1998–2000	Capital Stock	Salary and Wages

years, for physical capital, value added, and wages and salaries.

We use these data to estimate the physical capital stock in the NMNA sector in three eras: 1929–1946, when capital and value added by industry are not observed directly; 1947–1962 and 1998–2000, when capital by industry is observed directly but value added by industry is not; and 1963–1997, when both are observed directly. We proceed in roughly reverse chronological order since more recent estimates are based on more direct observations.

4.2 Physical Capital in the NMNA Sector, 1963–1997

Between 1947 and 1997 we have all the data to implement equation (1) directly. In particular we take Y_{ijt} to be value added in industry j in state i at time t and $Y_{jt} = \sum_i Y_{ijt}$ to be the sum of value added in industry j across all states at time t. We take K_{jt} to be the aggregate capital in industry j from NIPA.

4.3 Physical Capital in the NMNA Sector, 1947–1962 and 1998–2000

From 1947 to 1962, and again from 1998 to 2000, we have the aggregate capital stock by industry, but not value added by state and industry. However, we do have wage and salaries by state and industry from 1929–2000.[18] Our approach is to show that there is a very close relationship between value added and wages and salaries. We then use the fitted relationship to estimate value added by state and industry from 1929–2000. Given capital by industry (from the data) and value added by state and industry (estimated using the

[18] With a few exceptions. We have to interpolate mining values for Delaware in 1933 and 1934, and for Hawaii from 1975–1978; and we have to interpolate mining and services values for Idaho and Minnesota in 1998. We also project mining and services forward to 1999 and 2000 in Delaware using the observed 1993–1998 growth rate; project mining and services forward to 1998–2000 in Washington, D.C. using the 1992–1997 growth rate; and project mining back to 1963–1964 in Washington, D.C. using the 1965–1970 growth rate.

fitted relationship to wages and salaries), we can estimate physical capital in the NMNA sector with equation (1).

For each of the seven industries, we fit the relationship between log value added and log wage and salaries with state fixed effects using the years of overlapping data, 1963 to 1997. Table A1 shows the fitted relationship between the two. The elasticity is around 1 for the industries, and after including state fixed effects the R^2 is high; the lowest R^2 is 0.994 in mining. The state fixed effects (not shown) are important to achieving this good fit: there are economically large and statistically significant differences in the relationship between wages/salaries and value added by state.

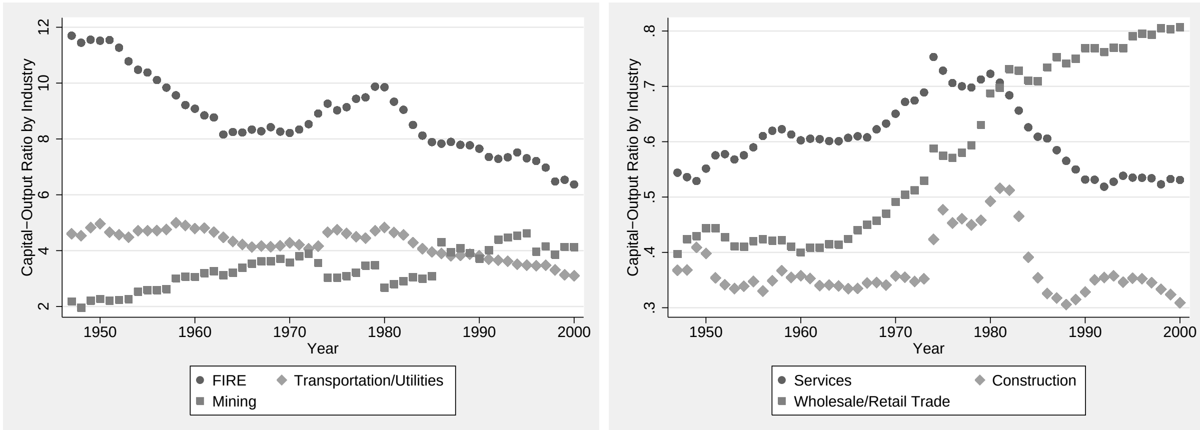
We use the fitted relationships for each industry, along with wages and salaries paid in that industry from 1929–1962 and 1998–2000, to predict value added by state and industry for years where it is not reported. When combined with the direct observations on aggregate capital by industry, we can use equation (1) to estimate physical capital in the NMNA for each state from 1947–2000.

4.4 Physical Capital in the NMNA Sector, 1929–1946

We can form reasonable estimates of the value added by state and industry as far back as 1929 using the methodology explained in the last subsection. However, before 1947 the BEA only has data on the physical capital in the government. Hence, we cannot construct national capital-output ratios for each industry. Instead we project these capital-output ratios using the trends observed from 1947–2000. We use this projected capital-output ratio to construct physical capital per state as

$$K_{it} = \sum_j \frac{K_{ijt}}{Y_{ijt}} Y_{ijt} \approx \sum_j \widehat{\frac{K_{jt}}{Y_{jt}}} Y_{ijt}. \tag{2}$$

Figure 10 plots the time series of aggregate output per worker for the six industries for which we do not have data from 1929–1946. They are clustered into two groups: mining, transportation and utilities, and FIRE all have capital-output ratios significantly greater than unity (Figure 10a), while services, wholesale/retail trade, and construction all have capital-output ratios less than unity (Figure 10b). Three of these series show evidence of systematic trends: FIRE seems to have become less capital-intensive, while mining and wholesale/retail trade have become more capital-intensive. Given that there is some trend in these time series it is important that we be careful in projecting capital-output ratios back to 1929.



(a) High Capital-Output

(b) Low Capital-Output

Figure 10: Time Series of Capital-Output Ratio by Industry

The primary question is whether we should assume the observed trends also held for the 1929–1946 period. For evidence we turn to the historical nineteenth century capital-output ratios computed by Gallman (1986). He reports values for several of the underlying NMNA industries for three different years in the nineteenth century. We report the average values for select industries of relevance in the first two columns in Table 5. By comparing his nineteenth century values to our 1947–1997 values, we can in most cases make a reasonable forecast about whether capital-output trends were part of a longer-term trend or were temporary.

Based on this evidence, we use the average capital-output ratio from 1947–2000 for services, construction, and transportation/utilities, none of which seem to show pronounced trends. Comparison to Gallman's figures shows that our averages agree closely with his reported historical values. For the industries with trends in the capital-output ratio we use different approaches. Comparing our series for FIRE to Gallman's figure for real estate leads us to use the average value rather than allow for a trend, which would imply capital-output ratios over 15 in FIRE in the nineteenth century, against his average value of 7.91 for real estate. For wholesale/retail trade we project the 1947 value back. The series showed no trend at around 0.45 until 1965, when it started to trend upwards; further, a lower value (rather than the average) more closely agrees with Gallman's reported capital-output ratio in commerce in the 19th century. For mining, we take the average from 1947–2000, and project it back. Unfortunately Gallman does not report a separate mining figure, so we cannot compare our results to any historical figures. The last two columns of Table 5

Table 5: Comparison of Projected K/Y Against Outside Evidence

Gallman (1986)		Projected	
Industry	Value	Industry	Value
Transportation & Public Utilities	4.15	Transportation & Utilities	4.19
Real Estate	7.91	Finance, Insurance, & Real Estate	8.63
Commerce & All Other Private Business	0.53	Services	0.58
		Wholesale/Retail Trade	0.39

Values for Gallman (1986) are the average reported for 1840–1900; his figures show some range over this time. Projected value is the value that would prevail in 1900 under the projection chosen for the industry, as discussed in the text.

give our industries and estimated capital-output ratios, lined up against the comparison industries from Gallman.

We use these projected capital-output ratios and the estimated value added by industry to estimate physical capital in the NMNA sector using equation (2). This series covers most states for 1929–1946.

4.5 Alaska, Hawaii, and Washington D.C.

Alaska, Hawaii, and Washington, D. C. all have somewhat incomplete records for value added and/or wages by industry. Alaska and Hawaii have no data from before 1950, but all the necessary data afterwards. Hence, our series for them spans 1950–2000. Washington, D.C. has sufficient data from 1963–2000; for earlier years they lack data on fishing, forestry, and agricultural services, which actually appear at least modestly important from 1963 onward. Hence, we provide estimates for Washington, D.C. from 1963–2000, although they are not used in our empirical work.

4.6 Data and Comparisons

Figure A7 presents our estimates of the time series of physical capital in the non-manufacturing, non-agricultural sector by state. Note that the results begin only in 1950 for Alaska and Hawaii, and even later for Washington, D.C.

Our method for constructing capital stocks captures variation in capital intensity that arises from differences in industry composition across states, but it abstracts from within-industry differences in capital intensity across states. Our method will of necessity understate the total variation in capital intensity across states. A potential concern is whether

Table 6: Variability of Predicted Capital-Output Ratios in the NMNA Sector

Year	Coefficient of Variation	Minimum	Maximum
1929	0.11	3.59	5.91
1947	0.11	3.37	5.97
1997	0.09	2.54	4.31
2000	0.10	2.42	4.13

there is any meaningful variation in capital intensity using our estimates. Figure 10 shows that industries do vary substantially their capital-output ratios, giving us the potential for cross-state variation. Table 6 shows that there is indeed cross-state variation. It shows, for various years, the coefficient of variation, minimum, and maximum capital-output intensity across states. The minimum capital-output ratio is around two-thirds the maximum, with a relatively stable coefficient of variation of around 0.1.

We have no outside sources to which we can compare our results. Note that our national results will match the national BEA results by construction, since we used the BEA data on capital stock by industry to construct capital-output ratios. Further, we have already verified that our capital-output ratios compare well with Gallman's (when possible) in Table 5.

5 Land in the Agricultural Sector

Our last data contribution is to construct estimates of the land input in the agricultural sector for each state, covering up to 1850–2000. While acreage of land is readily available, our primary contribution is to adjust each state's acreage in agriculture for the agricultural quality of the land. We measure quality-adjusted land input in two steps, with two different data sources.

In the first step, we collect estimates of the acreage of land in the agricultural sector for each state and year.[19] We actually collect the acreage of three different types of land, distinguishing between irrigated and nonirrigated cropland as well as pastureland. These data come from the Censuses of Agriculture, 1850–1997, and special reports to the same. In the second step, we estimate the relative productivity of each of the three types of land in each state. We do so using the cash rental price of land, which is the spot market price

[19] We follow the practice of our sources of measuring land in acres. International sources commonly use hectares instead. For purposes of comparability we note that an acre is 0.4047 hectares.

to rent an acre of land of a specified type in that state for a year. In doing so we rely on the standard no-arbitrage condition: if relative prices do not reflect relative productivities, then farmers should prefer to rent land with the highest ratio of productivity to rental price. We used the rental price data from a fixed year.

Given each state i's time t acreage endowment $(\lambda_{it}^{irr}, \lambda_{it}^{nonirr}, \lambda_{it}^{unimp})$ and the assumed time-invariant relative productivity of these types of land $(z_i^{irr}, z_i^{nonirr}, z_i^{unimp})$, we construct the state's quality-adjusted land input $\mathcal{L}_{it}$ as:

$$\mathcal{L}_{it} = \lambda_{it}^{irr} z_i^{irr} + \lambda_{it}^{nonirr} z_i^{irr} + \lambda_{it}^{unimp} z_i^{unimp}.$$

This method of summing different inputs, weighted by the observed market rental price is similar to the methodology used by Bils and Klenow (2000) to aggregate workers with different years of schooling into a single human capital input.

Our use of cash rental prices provides an important advance on previous methods, which typically use the average value or purchase price of an acre. Since land is a stock variable, the purchase price reflects the present discounted value of future expected income flows. Some of the value may then reflect future non-agricultural uses of the land: the high purchase price of an acre of farmland near a city reflects the anticipation of future development, not high agricultural productivity. By using spot market prices, we avoid this problem. Lindert (1988b) documents fluctuations in the ratio of cash rents to land values that suggest this correction may be important.

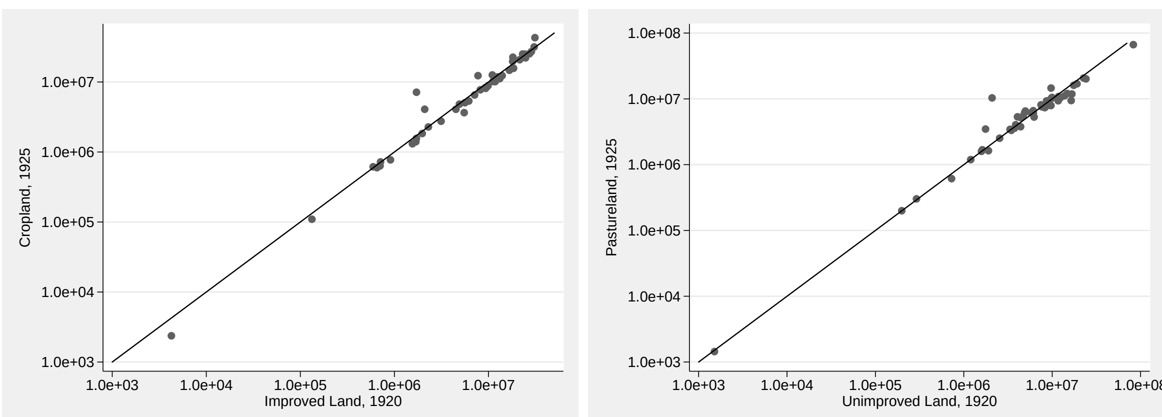
Our estimates account for some but not all of the changes in land inputs over time. We capture changes in the total acreage, as land is added to or removed from the agricultural sector. We also capture changes that arise from discrete improvements that move acreage between our three categories: clearing pastureland for use as cropland, or irrigating cropland; Primack (1977) documents the difficulty and importance of these tasks in early American agriculture. However, we miss any "within-category" changes that improve land but do not change its status between these three discrete groups.

In the next three subsections we describe the data for the acreage of land, the cash rental price of land, and the resulting series.

5.1 Farm Land Acreage

The Censuses of Agriculture have collected data on land used in agriculture for each state since 1850. Since the beginning they have distinguished between land that was ready for crops, and land that was not. Prior to the 1925 census they collected data on acres of improved and unimproved land.[20] From the 1925 census onward they changed the terminology, reporting the total amount of land and cropland. However, they use the term cropland to refer to land that is *available* for crops and not necessarily land that is *used* for crops. Indeed, in later years they ask specifically the acreage of cropland left fallow, the acreage of cropland used as pasture, and so on. For this reason, improved land and cropland correspond closely.

Our first step is to use this data to separate the total acreage into acreage of cropland and unimproved land. We do so by equating improved land to cropland, and unimproved land to all land that is not cropland. We check whether it is reasonable to construct a single series out of these different measures. We ask whether the improved land values in 1920 seem consistent with the cropland values in 1925, and likewise whether the unimproved land values in 1920 seem consistent with the non-cropland values in 1925. Figure 11 gives the scatter plots for these series, with the 45-degree line included. They are largely consistent. The few large increases are for Western states where land was rapidly being settled. Note that there are no examples of states with large decreases in land. This fact suggests that the two series are compatible.



(a) Cropland and Improved Land

(b) Pastureland and Unimproved Land

Figure 11: Comparison of 1920 and 1925 Land Acreage by Type

The next step is to distinguish between irrigated and non-irrigated cropland. As agriculture pushed into the arid Western states, irrigation became more important. Starting in 1890 the censuses include data on irrigation in special reports that covered only the Western states. Beginning in 1940 the censuses include information on irrigation in all states as a

[20]Tostlebe argues that the definitions of improved and unimproved were reasonably consistent over time in early censuses (p. 177).

part of the regular census report. We lack data on irrigated land before 1890 for Western states, and before 1940 for Eastern states.

For most years and most states where we do have data, the value reported is the total irrigated acreage. We proceed in two steps. First, we need to estimate the irrigated acreage for years before 1890 in Western states, and before 1940 in Eastern states. We take for each state the fraction of the total land acreage that is irrigated in the earliest year for which we have data, either 1890 in the West or 1940 in the East. To estimate irrigated land in earlier years we multiply the total endowment of land by the fraction of land that was irrigated in 1890 or 1940. Hence our estimate in Western states prior to 1890 is given by

$$\lambda_{it}^{irr} = \left(\lambda_{it}^{irr} + \lambda_{it}^{nonirr} + \lambda_{it}^{unimp}\right) \frac{\lambda_{i1890}^{irr}}{\lambda_{i1890}^{irr} + \lambda_{i1890}^{nonirr} + \lambda_{i1890}^{unimp}}$$

and in Eastern states prior to 1940 by

$$\lambda_{it}^{irr} = \left(\lambda_{it}^{irr} + \lambda_{it}^{nonirr} + \lambda_{it}^{unimp}\right) \frac{\lambda_{i1940}^{irr}}{\lambda_{i1940}^{irr} + \lambda_{i1940}^{nonirr} + \lambda_{i1940}^{unimp}}.$$

This assumption has important consequences in Western but not in Eastern states, for two reasons. First, much more land in the West is irrigated. Figure A8 plots the raw data on the fraction of land that is irrigated by state. For Eastern states, irrigation is relatively unimportant; none had more than 1% of their land irrigated in 1940. For Western states irrigation is already prevalent in 1890, and so it matters quantitatively whether irrigation was equally prevalent before 1890. Second, as we will show in the next subsection, the cash rental price (and presumably productivity) of irrigated land is similar to nonirrigated land in the East but much higher in the West. For these reasons the treatment of data prior to 1890 is important in the West.

We assume that the fraction was constant because qualitative evidence suggests that irrigation was present well before 1890 (Coman 1911, Hess 1912). Early agriculture and irrigation in the West were often paired to take advantage of accessible local water sources such as rivers or lakes. In some cases it grew naturally out of the diversion of water used for mining purposes, and the profitable opportunity to feed mining camps (Hess 1912). Later, agriculture and irrigation both expanded through the use of large-scale irrigation systems for land with less obvious water sources; 1880–1890 saw a particular spurt of such projects (Coman 1911). Put together, the evidence suggests that irrigation was common and growing rapidly to support agriculture in the West. However, since we lack firm data we also explore an alternative series where we assume that no land was irrigated before our earliest data year, which again is 1890 in the West and 1940 in the East.

At this point we have estimates of cropland, unimproved land, and total irrigated land. The second step in accounting for irrigation is to estimate the fraction of irrigated land that is used for cropland. In most years the Census reports only the acreage of irrigated land, without specifying the use. From roughly 1930–1959 it reports the acreage of irrigated cropland for Western states; the exact dates vary somewhat by state. For the years with overlap between acreage of irrigated land and acreage of irrigated cropland we construct the ratio of the two. As expected, most irrigated land in most states is used for crops. However for a few states a non-trivial fraction of irrigated land is not used for crops; most of this land is used for dairying, and some is used as irrigated pastureland (Selby 1949). The most extreme case is Nevada, where about 43% of irrigated land is not used for crops. This fraction is relatively stable over the time period.

For Western states we estimate the average fraction of irrigated land used for crops from 1930–1959. We take this value and multiply it by the irrigated land for other years to estimate the acreage of irrigated cropland. We assign irrigated land not used for crops to the (homogeneous) unimproved land. For Eastern states we never observe the fraction of irrigated land used for crops. We assume that all irrigated land in these states is used for crops. Given the low level of irrigation and the relatively small productivity difference between different types of land in the East, this assumption makes little difference.

Data for Hawaii and Alaska are sporadic and have some irregularities. Irrigation is important in Hawaii, but data on irrigation are available only in 1900 and then from 1940 onward. Further, the 1900 data suggest substantially more irrigated acreage than the 1940 data. We choose to start our Hawaiian series in 1940 to avoid the uncertainty with the early data point. In Alaska, the reported value of unimproved acreage fluctuates a lot in the early censuses. We smooth the data to avoid large rises and falls in the land input. We also start Alaska in 1940. By starting the data for these states in 1940, we have to interpolate only two missing censuses, in 1945 and 1954. Our series for Washington, D.C. begins in 1850 but ends in 1945; after 1945, Washington D.C.'s land area was no longer enumerated. Acreage in irrigation was never enumerated for D.C.; we assume that no land was irrigated there.

We also interpolate a few missing values in other states, but mostly to cover missing data for a single census. For example, there is no irrigation data for Western states in 1925, and no irrigation data for Louisiana in 1954. These briefer interruptions in data should have little effect on our results. The resulting land series includes λ_{it}^{irr}, λ_{it}^{nonirr} and λ_{it}^{unimp} for states and years. These series match exactly the data on cash rental prices, allowing us to construct a productivity-adjusted series. We now turn to discussing the cash rental price data.

5.2 Cash Rental Prices and Productivity

Our next goal is to weight the acreage of land in different states and of different types by the relative quality or productivity of that land. We use the cash rental price of land as a measure of its productivity, appealing to a standard no-arbitrage condition. The United States Department of Agriculture has collected different measures of cash rental prices intermittently since 1930.[21] The quality and coverage of the data has increased substantially over time, leading us to focus on the most recent data, from 2009 (United States Department of Agriculture National Agricultural Statistics Service 2009). We also provide robustness checks using older cash rental price data, from as far back as 1930.

By the year 2009, the cash rental price data are fairly complete across states and land types, covering 127 of the 153 possible combinations (three types of land in 51 states). Our general strategy is to handle the few missing values by using the cash rental price for the most similar type of land available. We discuss each missing observation in turn. First, six Northeastern states report only a single cash rental price for cropland, without distinguishing by type.[22] We apply this rental price to both irrigated and nonirrigated land in these states. Five of the six states also have no cash rental price for pastureland. We fill these cash rental prices using the average of the observations for Vermont and Massachusetts, two nearby and similar states. We use the rental price of nonirrigated land in West Virginia for irrigated land, and the rental price for New Mexican nonirrigated cropland for nonirrigated cropland in Nevada and Arizona.

Alaska and Washington, D.C. are missing entirely from our data. We have found no cash rental price data for Alaska in any source; the only data we can find for Washington, D.C. is older and unreliable (as we will discuss below). Hence we need to estimate cash rental prices for all three types of land for both states. Washington, D.C. is geographically close to Virginia and Maryland and shares their coastal land; we assume that its cash rental prices are the same as Maryland's. For Alaska, most of the farmed cropland is in two valleys in the southeast coastal portion of the state, which features a short growing season and adequate rainfall. Maine's cropland is probably the closest equivalent, so we use Maine's cash rental prices for cropland for Alaska as well. Most of the unimproved land is in the

[21] They have also occasionally collected information on land rented for share payments, or land rented for a mixture of share payments and cash. These estimates are less useful for our current exercise.

[22] The six are Connecticut, Maine, New Hampshire, Rhode Island, Vermont, and Delaware.

Aleutian islands, where average farm values are extremely low.[23] We associate this land with the lowest pastureland rental price for the continental United States, that of New Mexico; even this assumption may be overly generous.

Using these imputations we have a complete record for z_i^{irr}, z_i^{nonirr}, and z_i^{unimp}. In the next section we combine these cash rental prices with our measures of raw acreage by state and type to produce our estimates of productivity-adjusted land input by state.

5.3 Data and Comparisons

Our baseline estimates are constructed using 2009 cash rental price data:

$$\mathcal{L}_{it} = \lambda_{it}^{irr} z_{i,2009}^{irr} + \lambda_{it}^{nonirr} z_{i,2009}^{irr} + \lambda_{it}^{unimp} z_{i,2009}^{unimp}.$$

Figure A9 gives our baseline estimates of productivity-adjusted land input for the states, plotted in logs. We face one significant source of uncertainty, namely how to estimate irrigation in early years in Western states. Our baseline series assumes that the fraction of cropland that was irrigated in 1850–1880 was the same as the observed fraction in 1890. We create an alternative series that assumes no irrigation before data are available. Figure 12 compares the results of this series to our baseline for the Western states. The difference between the series is largest in New Mexico, but also apparent in Arizona, California, Colorado, Idaho, Nevada, and Wyoming. If irrigation was less widespread before 1890 than our sources indicate, then this would imply lower levels of productivity-adjusted land in the West in the 19th century.

Next we explore whether the productivity adjustment is important for our results and, if so, in which states it matters most. We compare our estimate to one consisting of raw acreage, generated by summing $\lambda_{it}^{irr} + \lambda_{it}^{nonirr} + \lambda_{it}^{unimp}$. Figure 13 plots the relationship between our land estimate and acreage for states in 1997. The primary impact of our results is to reduce the land input in Western states. The size of this correction is significant. To give one example, New Mexico has a similar acreage in agriculture as does Nebraska, but Nebraska provides ten times the quality-adjusted land input in agriculture because New Mexico's land is quite unproductive. To put this difference into perspective, note that Herrendorf and Valentinyi (2008) measure a land share of 18 percent in the U.S. economy. Failing to adjust for land quality would thus lead one to understate Nebraska's land input by 50 percent ($= 10^{0.18}$) and overstate their TFP by the same, as compared to New Mexico.

[23] The 2007 Census estimates the market value of land and buildings at just 67 dollars per acre, lower than every single county in New Mexico.

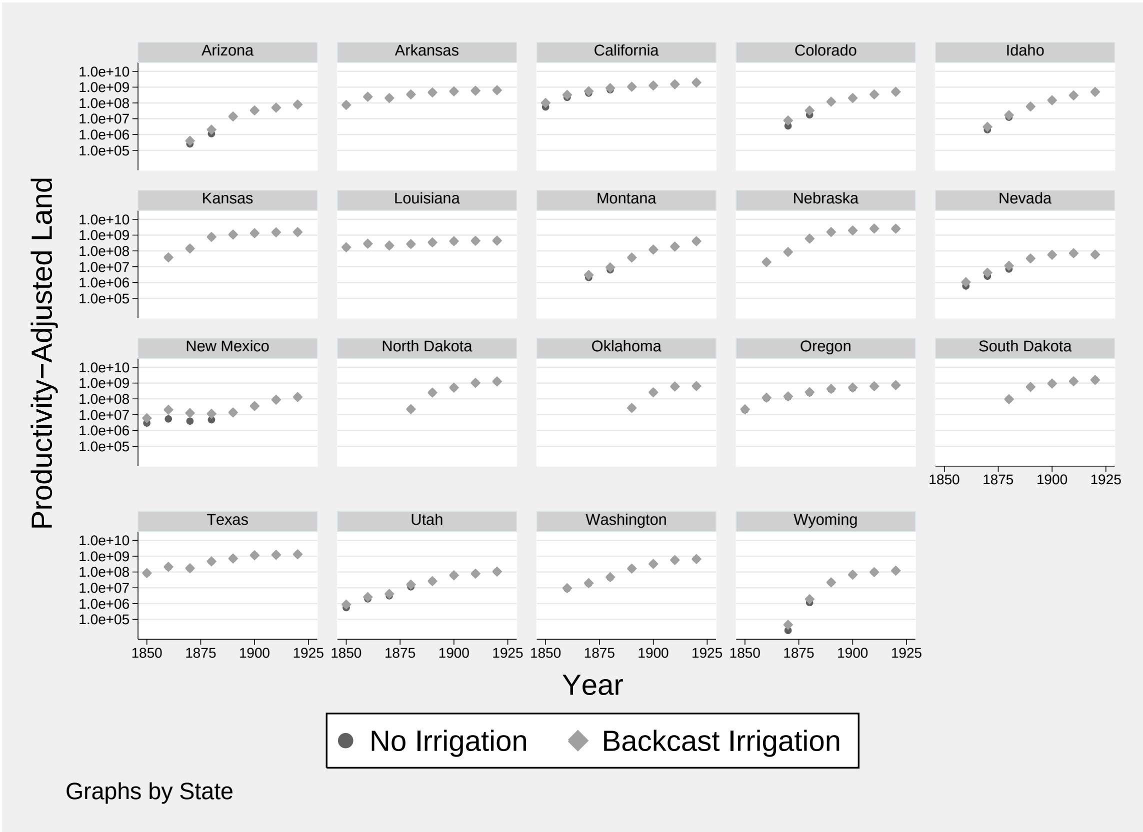


Figure 12: Land Estimates of Western States Under Alternative Assumptions on Early Irrigation

Hence, while making productivity adjustments is challenging in some respects, doing so also has quantitatively important and sensible implications for the measured land input.

Finally, we test whether our results are sensitive to the use of year 2009 cash rental prices. Using earlier years may be important if the relative productivity (and relative cash rental price) of different types of land has changed over time, perhaps as new methods of farming or new crops alter the relative merits of land in different states. We produce alternative productivity-adjusted land data using cash rental prices from three other years. Data in these earlier years cover fewer states and distinguish fewer types of land. The demerits of having fewer states is straightforward. Distinguishing fewer types of land is problematic because the average rental rate will then depend on the type of land rented - irrigated cropland typically rents for much more than pastureland.

In 1930, 1940, 1950, and 1954 the Census of Agriculture collected information on cash rents paid for rented land.[24] These data cover most of the states, but are a single cash

[24] Although the 1959 and 1964 Censuses of Agriculture include similar information, these Censuses include

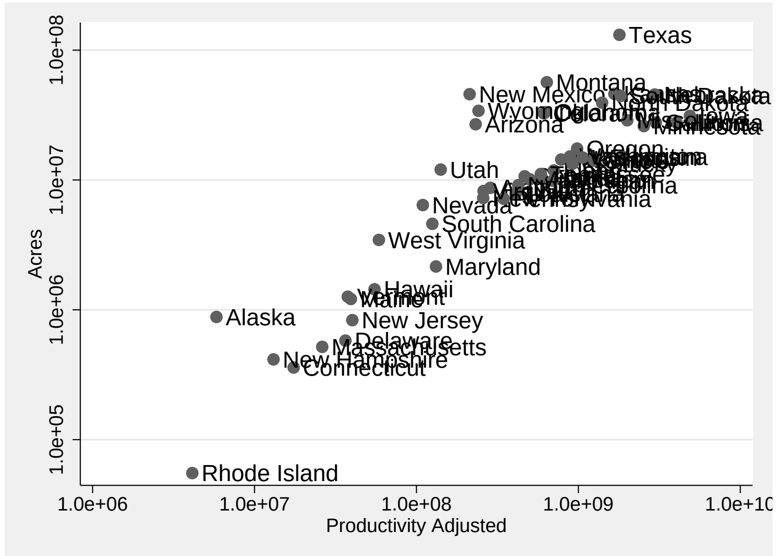


Figure 13: Comparison of Acreage and Productivity-Adjusted Land

rental price for each state. Still, given that this data was recorded nearly eighty years prior to our baseline cash rental price data, we use it. We focus on data for the earliest year, 1930.

The USDA moved measurement of cash rental prices to a separate survey in 1960.[25] The initial surveys covered fewer states than the Census and still offered only a single measure per state. Over time the survey expanded. The 1967 survey was the first to distinguish between cropland and pastureland. We compute a second alternative measure of productivity-adjusted land by summing the acreage of pastureland and total cropland, with each weighted by its reported cash rental price. In 1984 the survey added eleven mostly Western states; for these states the survey distinguished between irrigated and nonirrigated cropland. Because of this change, we compute our final alternative measure for the year 1984, the first to distinguish (in some states) between all three types of land.

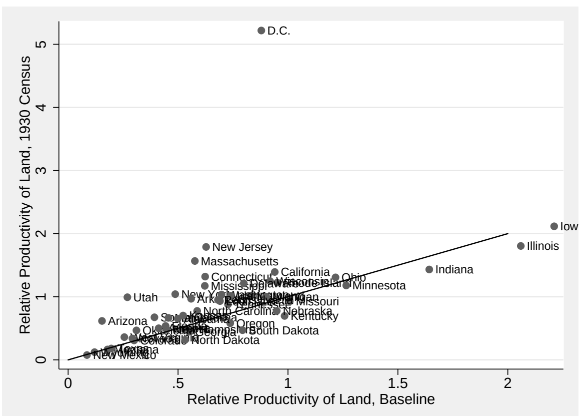
We normalize all series to facilitate comparison between our benchmark and alternative series. In particular, we report the weighted average productivity (rental price) per acre of land for each state. We also normalize each state's average productivity by the acreage-weighted average productivity of the 22 states for which data are available in all years. These normalizations allow us to investigate whether a state's farmland is consistently

the cash rental rate for land where cash was used for only part of the payment, making them less useful for our current purposes (Lindert 1988b).

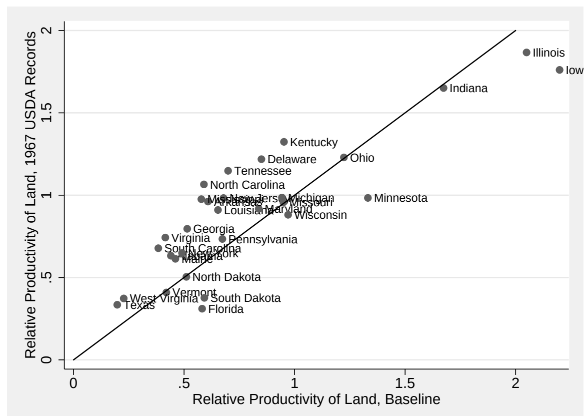
[25] The relevant survey is the Agricultural Land Value Survey from 1960 to 1994 and the June Agricultural Survey from 1994 onward.

rated as above or below average in terms of productivity. We make no attempt to fill in missing values or states for the alternative series.

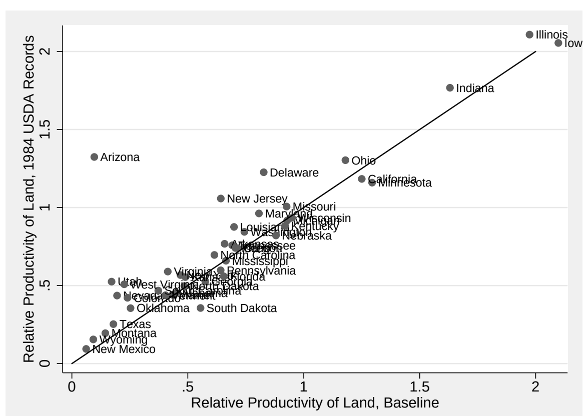
We choose carefully the years in which we compare the benchmark and alternative series. In principle, we can compare their performance in any year. However, most of the alternative series make little or no adjustment for different types of land. Hence, the 1930 cash rental price is influenced by the types of land available in 1930, which may differ from the types of land available today. To minimize this problem we compare the two series in 1930. Likewise, we compare the baseline to the second alternative series in 1967 and to the third alternative series in 1984.



(a) 1930 and 2009 Rental Prices



(b) 1967 and 2009 Rental Prices



(c) 1984 and 2009 Rental Prices

Figure 14: Comparison of Relative Land Productivity Using Different Rental Prices

Figure 14 compares our baseline measure of each state's productivity-adjusted land against an alternative computed using either 1930, 1967, or 1984 cash rental prices. The

45-degree line is included. If relative productivity is well-measured by cash rental prices and stable over time, then we expect the states to cluster along the 45-degree line so that (for example) states that had high rental price and appeared productive using 1930 cash rental prices also had high rental prices and appeared productive using 2009 cash rental prices. Looking across the figures we can see that the estimates agree closely and agree more over time; the R^2 of a simple regression is 0.25, 0.75, and 0.79. 1930 is a bit problematic compared to other years, mostly because the Northeast has much higher implied average productivity (particularly Washington D.C.).[26] We suspect that these deviations represent the value of farm dwellings included on rental land or other non-farm uses. Overall, although results using earlier cash rental price data differ somewhat, they do not suggest any systematic trends in cash rental prices or productivity.

6 Data Availability by State and Series

The data span only a portion of the time period for some states, typically because a state was not enumerated in one of the censuses we use as our primary data sources. Thirty-five states have data covering the entire period: Alabama, Arkansas, California, Connecticut, Delaware, Florida, Georgia, Illinois, Indiana, Iowa, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Utah, Vermont, Virginia, and Wisconsin. The availability of the four data series in the remaining sixteen states is given in Table 7.

One advantage of studying states instead of countries is that the geographic boundaries of states are generally more stable. Most large changes in boundaries happen before 1850 or before the state/territory is enumerated in censuses. We are aware of only one major change in state geography during our series of data availability: the split of West Virginia from Virginia between the 1860 and 1870 censuses as a result of the Civil War.[27]

[26] The other obvious outlier is Arizona in 1984, due entirely to a reported rental value of pastureland 38 percent higher than any other state in 1984, and 20 times larger than the value reported in 2009. We find this value to be implausible.

[27] The Gadsden purchase, ratified by Congress in 1854, added a small amount of land to New Mexico (available from 1850), but was more important for Arizona (not available until 1870). The retrocession of land from Washington, D.C. to Virginia, which had a significant impact on Washington, D.C.'s size, occurred in 1846, just before our data series begin.

Table 7: Date of First Data Availability for States with Incomplete Series

State	Land	Ag. Capital	Man. Capital	NMNA Capital
Alaska	1940	1900	1890	1950
Arizona	1870	1870	1870	
Colorado	1870	1870	1870	
Washington, D.C.*	1850–1945	1850–1945		1963
Hawaii	1940	1900	1900	1950
Idaho	1870	1870	1870	
Kansas	1860	1860	1860	
Montana	1870	1870	1870	
Nebraska	1860	1860	1860	
Nevada	1860	1860	1870	
North Dakota	1880	1860	1890	
Oklahoma	1890	1890	1890	
South Dakota	1880	1880	1890	
Washington	1860	1860	1860	
West Virginia	1870	1870	1870	
Wyoming	1870	1870	1870	

Notes: Table records date of first availability for states and series with partial availability. Empty cells indicate that the series was fully available, i.e., available for 1850–2000 for land, farm capital, and manufacturing capital, or 1929–2000 for NMNA capital.

* Land and farm capital data for Washington, D.C. is available starting in 1850, but availability ceases with the 1945 census.

7 Conclusion

This paper provides details on the construction of original estimates of the stock of physical capital in three sectors and land in the agricultural sector for states in the United States spanning up to 150 years. We focus on three main sets of issues. First, we discuss the underlying historical and contemporary data sources that are useful for such an estimation, and the coverage and reliability of these sources. Second, we discuss the choices that go into creating our estimates and plausible alternative choices, and when possible compare estimates made under different assumptions. Finally, we have compared our results with other results, mostly available at the national level, and discussed the comparability of our results with these outside estimates. We hope these estimates will be of use to other researchers.

References

Bils, M., and P. J. Klenow (2000): "Does Schooling Cause Growth?," *American Economic Review*, 90(5), 1160–1183.

Caselli, F. (2005): "Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, vol. 1A, chap. 9, pp. 679–741. Elsevier.

Collins, W. J., and J. G. Williamson (2001): "Capital-Goods Prices and Investment, 1870-1950," *The Journal of Economic History*, 61(1), 59–94.

Coman, K. (1911): "Some Unsettled Problems of Irrigation," *The American Economic Review*, 1(1), 1–19.

Creamer, D., S. Dobrovolsky, I. Borenstein, and M. Bernstein (1960): *Capital in Manufacturing and Mining: Its Formation and Financing*. UMI.

Cubas, G. (2010): "Accounting for Cross-Country Income Differences with Public Capital," mimeo, University of Iowa.

Cummins, J. G., and G. L. Violante (2002): "Investment-Specific Technical Change in the United States (1947-2000): Measurement and Macroeconomic Consequences," *Review of Economic Dynamics*, 5(2), 243–284.

David, P. A. (1990): "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *American Economic Review: Papers and Proceedings*, 80(2), 355–61.

Gallman, R. E. (1986): "The United States Capital Stock in the Nineteenth Century," in *Long-Term Factors in American Economic Growth*, ed. by S. L. Engerman, and R. E. Gallman, chap. 4, pp. 165–213. University of Chicago Press.

Garofalo, G. A., and S. Yamarik (2002): "Regional Convergence: Evidence From a New State-By-State Capital Stock Series," *The Review of Economics and Statistics*, 84(2), 316–323.

Gordon, R. J. (1990): *The Measurement of Durable Goods Prices*. National Bureaou of Economic Research.

Herrendorf, B., and A. Valentinyi (2008): "Measuring Factor Income Shares at the Sector Level," *Review of Economic Dynamics*, 11(4), 820–835.

Hess, R. H. (1912): "The Beginnings of Irrigation in the United States," *The Journal of Political Economy*, 20(8), 807–833.

Higgs, R. (2004): "Wartime Socialization of Investment: A Reassessment of U.S. Capital Formation in the 1940s," *The Journal of Economic History*, 64(2), 500–520.

Hurley, R. (1962): "Census Concepts: Past, Present, and Future," *Journal of Farm Economics*, 44(2), 616–621.

Jorgenson, D. W., F. M. Gollop, and B. M. Fraumeni (1987): *Productivity and U.S. Economic Growth*. Harvard University Press, Cambridge.

Kuznets, S. (1946): *National Product Since 1869*. National Bureau of Economic Research, New York.

——— (1961): *Capital in the American Economy: Its Formation and Financing*. Princeton University Press.

Lindert, P. H. (1988a): "Long-run Trends in American Farmland Values," Agricultural History Center Working Paper No. 45, University of California-Davis.

——— (1988b): "Long-run Trends in American Farmland Values," *Agricultural History*, 62(3), 45–85.

Primack, M. L. (1977): *Farm Formed Capital in American Agriculture, 1850 to 1910*. Arno Press, New York.

Ransom, R., and R. Sutch (1975): "The Impact of the Civil War and of Emancipation on Southern Agriculture," *Explorations in Economic History*, 12, 1–28.

Schoellman, T., and R. Tamura (2011): "Accounting for the Convergence of the U.S. States," mimeo, Arizona State University.

Selby, H. E. (1949): "The Importance of Irrigation in the Economy of the West," *Journal of Farm Economics*, 31(4), 955–964.

Tostlebe, A. S. (1954): *The Growth of Physical Capital in Agriculture, 1870-1950*. UMI.

Turner, C., R. Tamura, and S. E. Mulholland (2011): "How Important are Human Capital, Physical Capital and Total Factor Productivity for Determining State Economic Growth in the United States, 1850-2000," mimeo, Clemson University.

United States Department of Agriculture National Agricultural Statistics Service (2009): "Land Values and Cash Rents 2009 Summary," .

U.S. Census Bureau (various years): *Statistical Abstract of the United States*. U.S. Government Printing Office, Washington, D.C.

U.S. Census Office (1900): *U.S. Census of Agriculture*. Washington, D.C.

A Larger Tables and Figures

Table A1: Fitted Relationship Between value added and Wages/Salaries by State and Industry

	(1) Mining	(2) Construction	(3) Transportation/Utilities	(4) Wholesale/Retail Trade	(5) FIRE	(6) Services	(7) Government
Wages and Salaries	1.066***	1.065***	1.101***	1.003***	0.961***	0.974***	1.093***
	(0.0023)	(0.0005)	(0.0008)	(0.0002)	(0.0005)	(0.0002)	(0.0002)
Constant	-0.200***	-0.860***	-1.087***	0.675***	2.427***	1.065***	-1.673***
	(0.0480)	(0.0109)	(0.0170)	(0.0050)	(0.0117)	(0.0053)	(0.0058)
State Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3649	3681	3681	3681	3681	3596	3681
R^2	0.994	0.999	0.999	1.000	0.999	1.000	1.000

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dependent variable is value-added in logs; wages and salaries are also measured in logs.

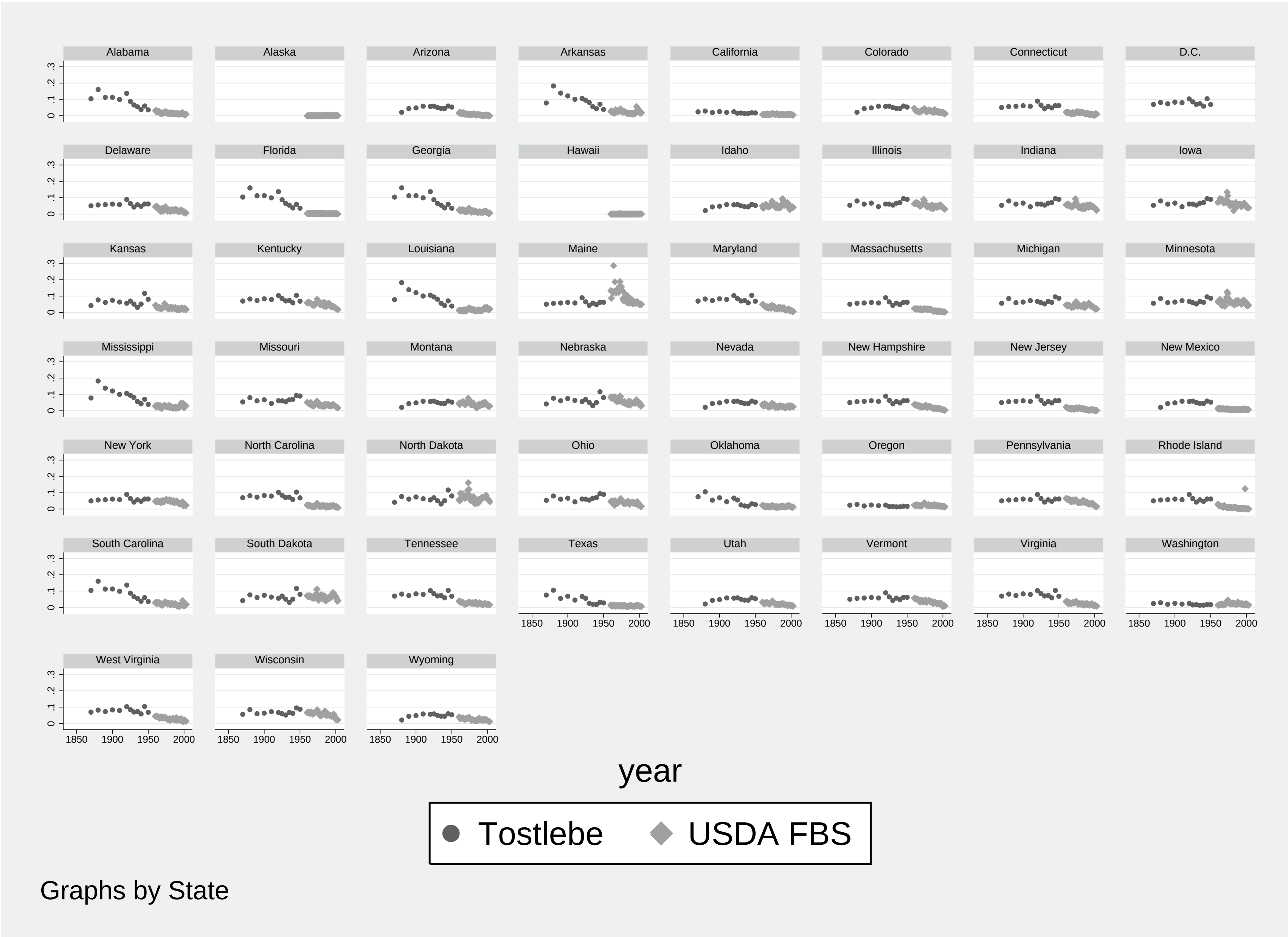


Figure A1: Time Series of Crop Inventories by State

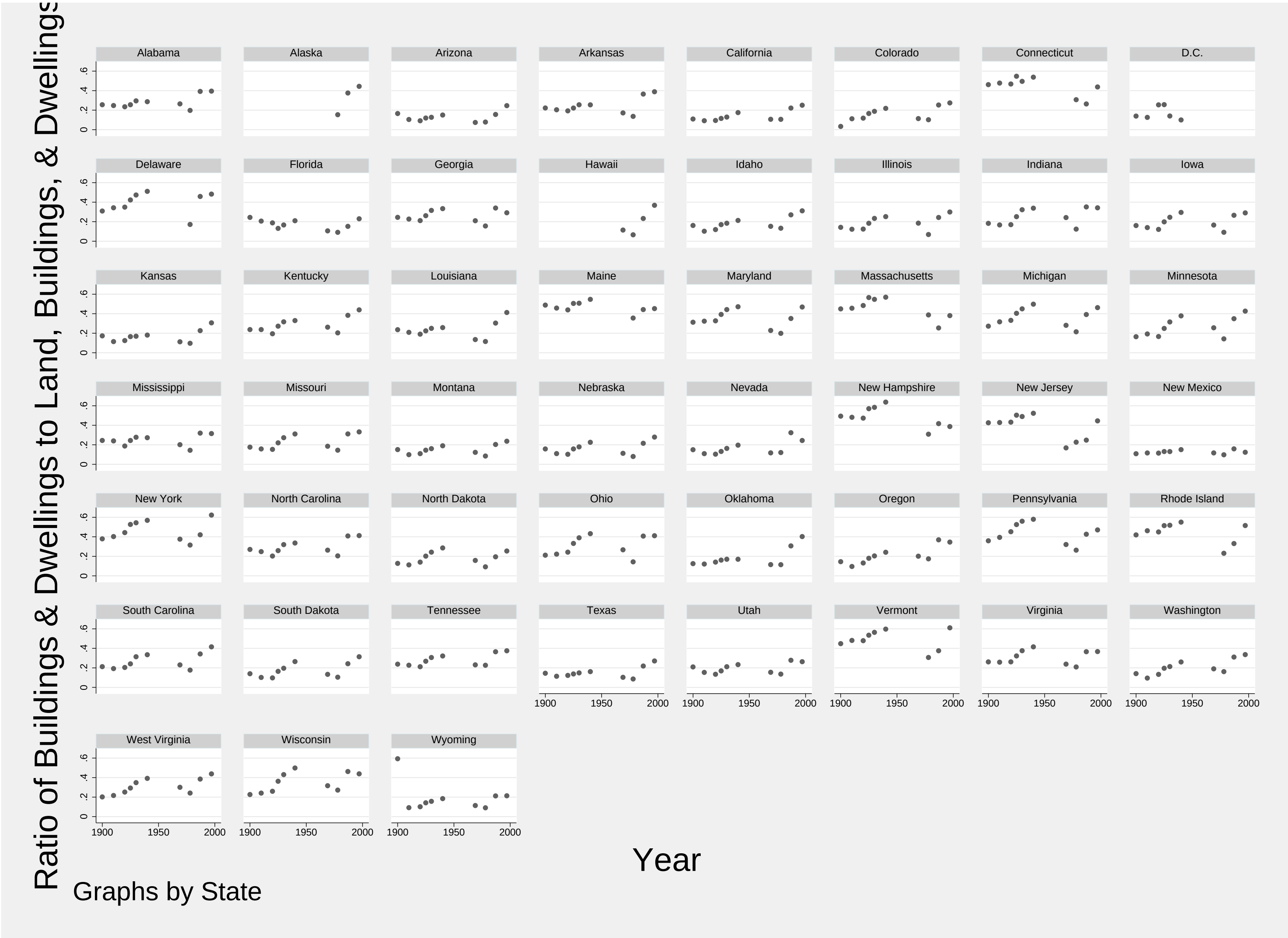


Figure A2: Share of Buildings and Dwellings in Real Estate Value

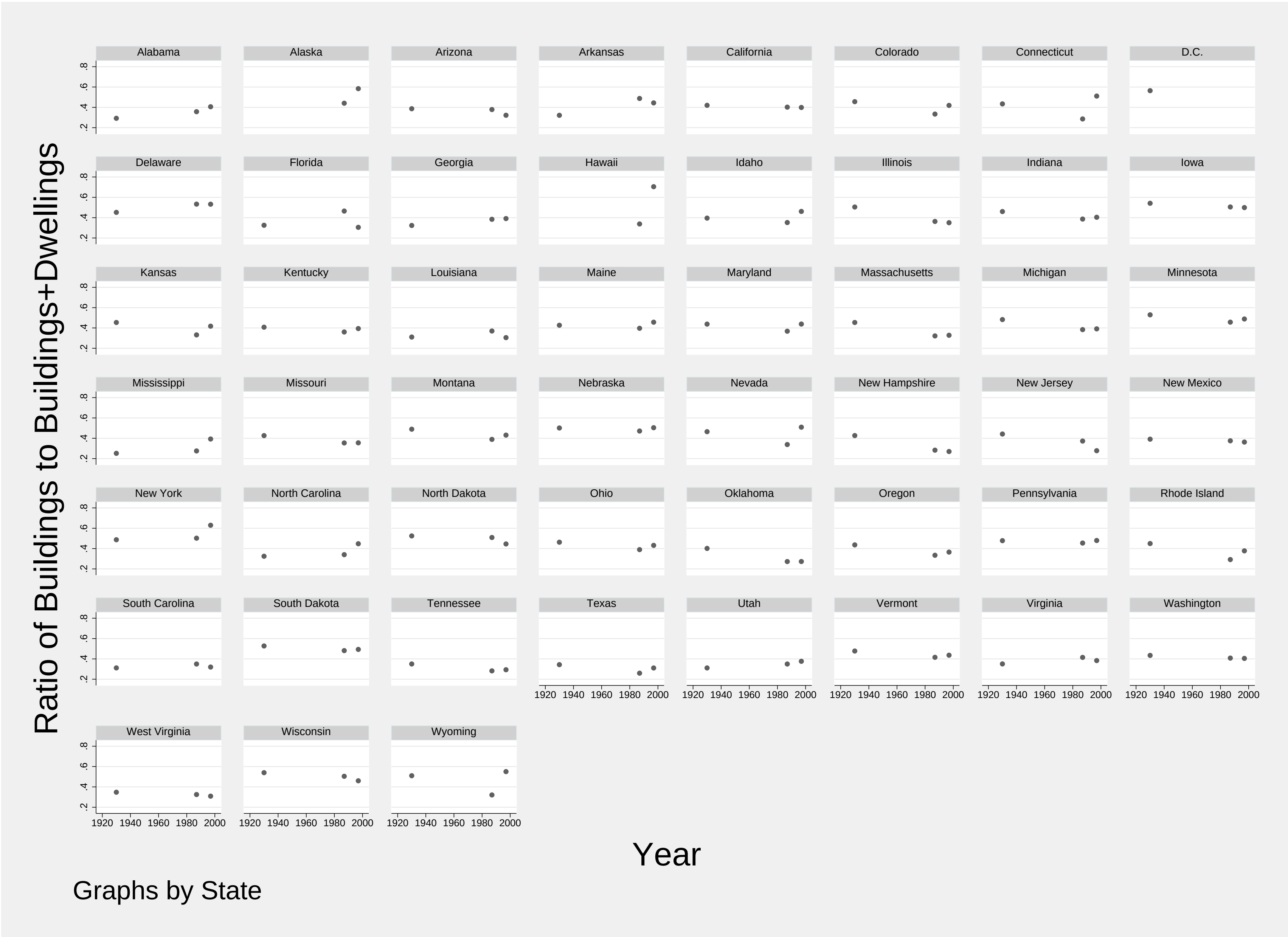


Figure A3: Share of Buildings in Buildings/Dwellings

Figure A4: Time Series of Physical Capital in the Agricultural Sector by State

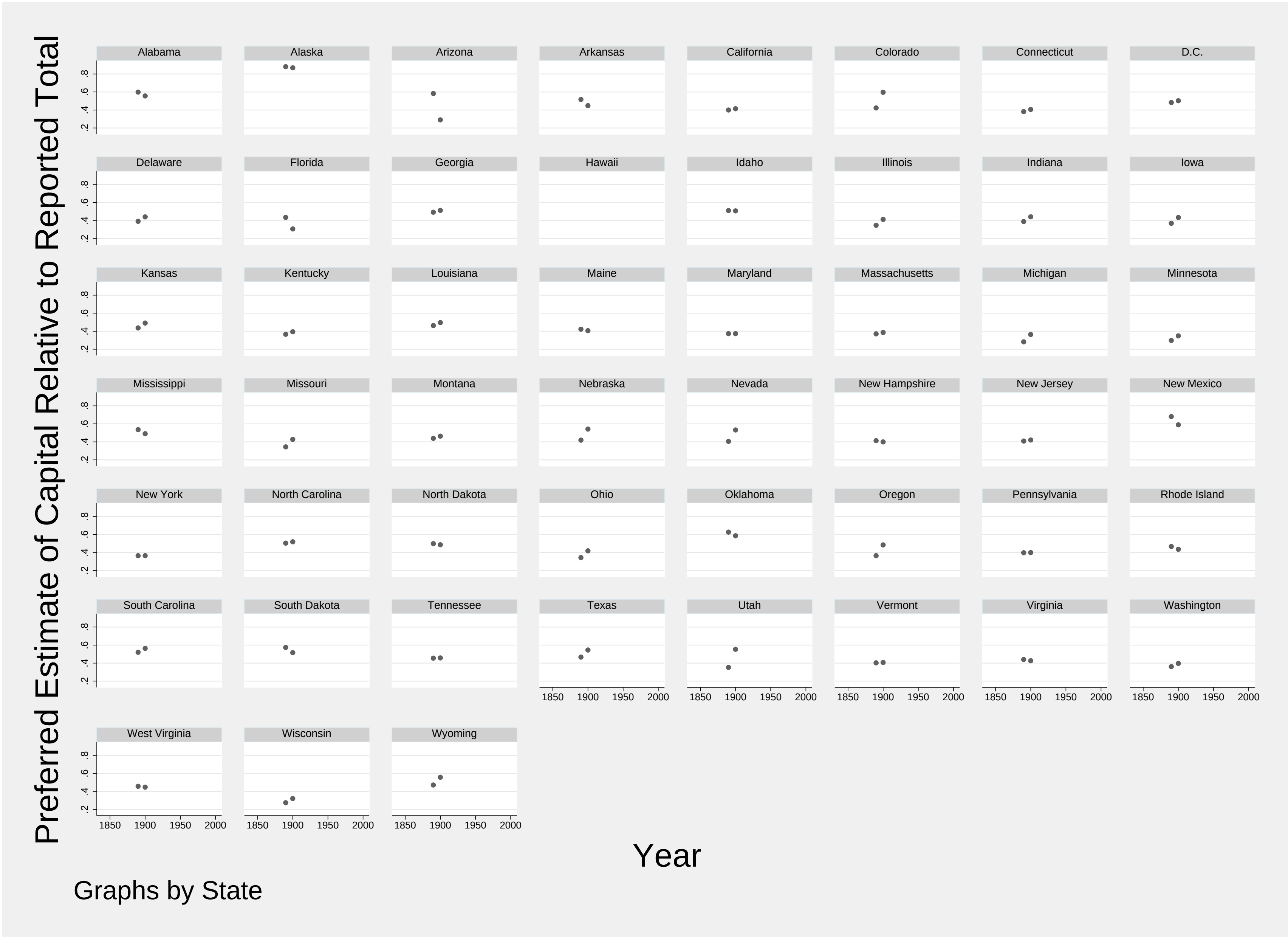


Figure A5: Capital and Reported Capital, 1890-1900

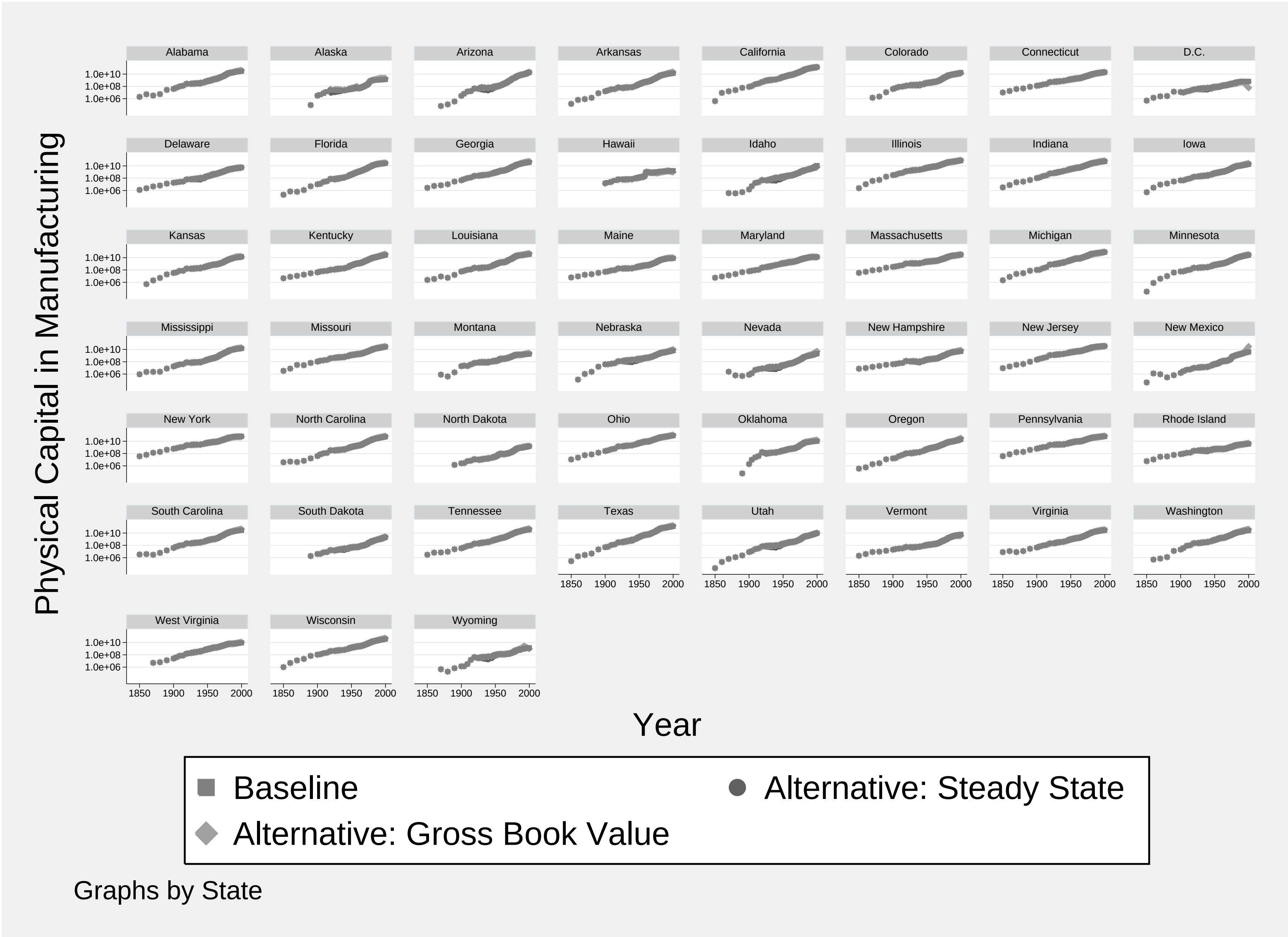


Figure A6: Time Series of Physical Capital in the Manufacturing Sector by State

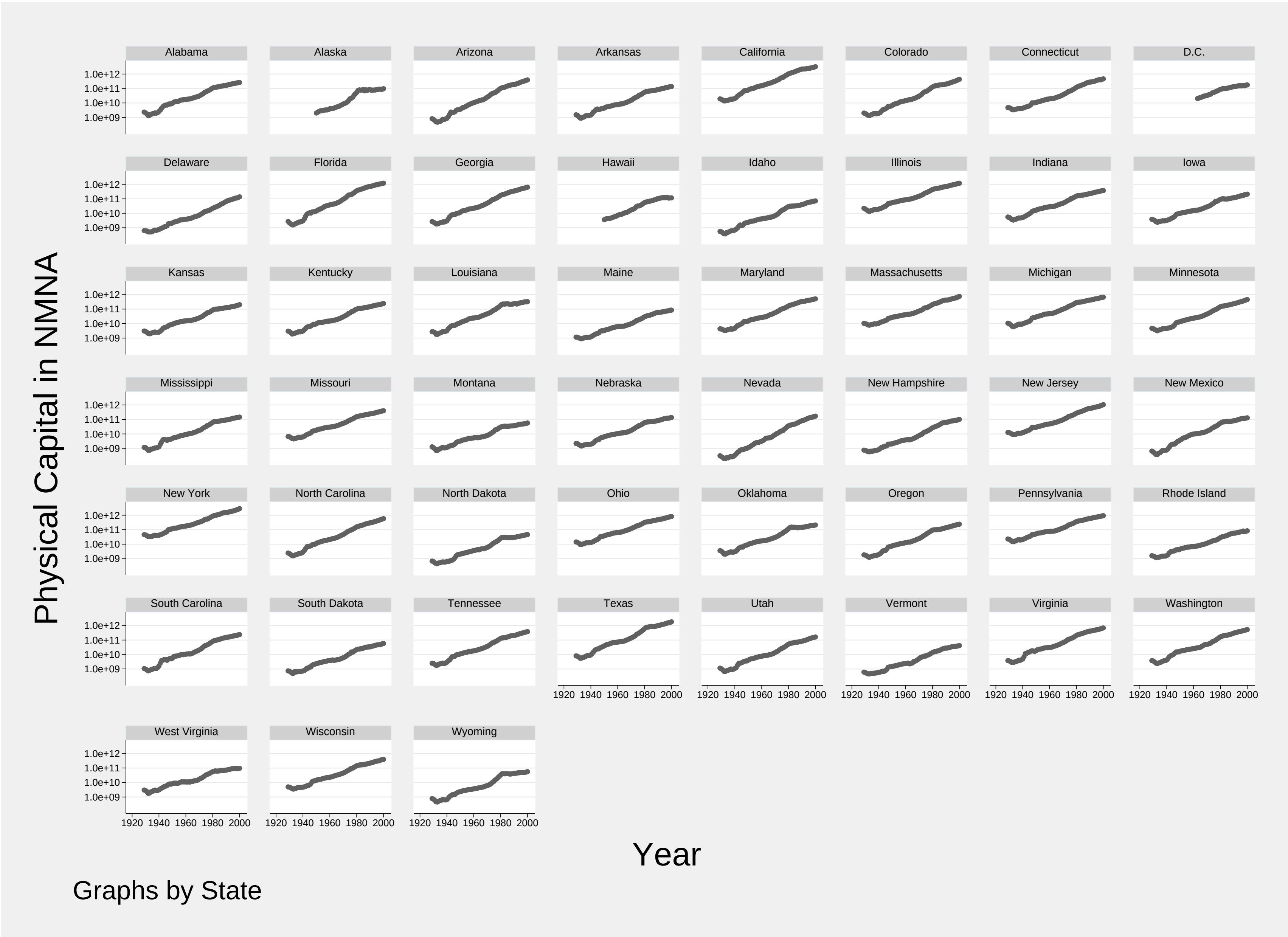


Figure A7: Time Series of Physical Capital in the NMNA Sector by State

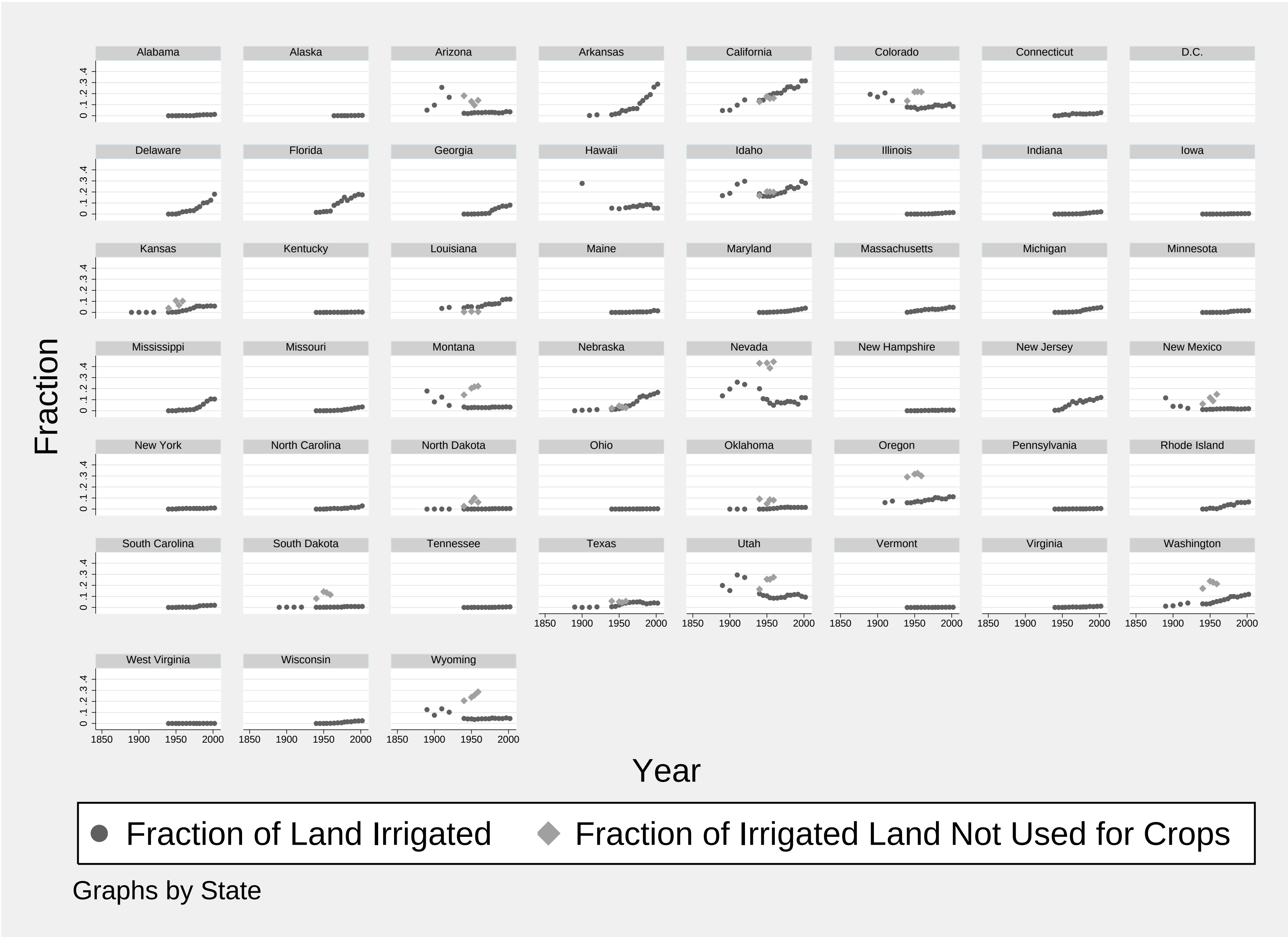


Figure A8: Irrigated Land as a Fraction of the Total

Figure A9: Time Series of Productivity-Adjusted Land in the Agricultural Sector by State