DAE
7/11/06
2:10 pm

The Returns to Education in the Early Twentieth Century: New Historical Evidence

Joseph Kaboski Trevon D. Logan[1]

June 2006

Preliminary Draft – Please Do Not Quote or Cite Without Permission

Abstract

There is a large literature that looks at skill biased technical change, both in developed and developing countries. The emerging consensus is that returns to education were U-shaped over the twentieth century in the United States. We construct a model which highlights the role that resource endowments play in the returns to education and their interaction with skill biased technical change. Given the regional labor markets and different sectoral structures in different areas of the country there is reason to expect considerable geographic variation in returns to education in the early 20th century. We use a new data source, a report from the U.S. Commissioner of Education in 1909, to estimate the returns to education of high school teachers in the early twentieth century. Overall, we find significant regional variation in the returns to education, with large (within-occupation) returns for the Midwest (7%), but much lower returns in the South (3%) and West (0.5%). We reject the hypothesis that the returns in the Midwest are equal to the returns in the South. We provide evidence that our results are generalizable to returns to education in the United States and that returns to education for teachers tracked quite closely with the overall returns to education from 1940 onward. These results suggest that we should expect variation in returns to education with skill biased technological change if there are significant differences in resource endowments before the technological change.

JEL Classifications: I2, J2, J3, N3

[1]Kaboski: Department of Economics, The Ohio State University, 410 Arps Hall, 1945 N. High Street, Columbus, OH 43210; e-mail: kaboski.1@osu.edu. Logan: Department of Economics, The Ohio State University, 410 Arps Hall, 1945 N. High Street, Columbus, OH 43210 and NBER; e-mail: logan.155@osu.edu. We gratefully acknowledge the financial support of the OSU Economics fund for Junior Faculty Research. We thank Joshua Cepluch and Sharmistha Dey for excellent research assistance. The usual disclaimer applies.

1 Introduction

Research on the returns to education in the early twentieth century has given rise to a new literature linking education, technology, and economic growth. Goldin and Katz (1999, 2000) used the Iowa state Census of 1915 to sketch out the returns to education in the early twentieth century, and a new view of the U-shaped returns to education over the twentieth century has emerged. There is now a large literature that looks at skill biased technical change (SBTC) and the general complementarity between technology and skill, whose origins are discussed by Goldin and Katz (1998). The consensus now is that high rates of return to education at the beginning of the century signaled the end of the period of early industrialization where physical capital, raw materials, and unskilled labor substituted easily for skilled labor. With the use of new large-scale processing technologies, and the increasing electrification of the industrial workplace, the returns to education increased dramatically in the early years of the last century. Returns declined again with the advent of the high school movement, perhaps intensified by the wage controls used in the second World War (Goldin and Margo 1992), but rose again in the second half of the century.

It is not clear, however, if U-shaped returns to education hold for the United States as a whole. There are several reasons to expect more variation in the returns to education in the first half of the twentieth century than the second. First, the new large-scale processing technologies that led to the rise in the returns to education did not diffuse evenly across the United States. The industrial states of the Northeast and Midwest had larger concentrations of such industries than the South and West. The South and West employed older-technology industries for the most part (including traditional agriculture), and it is therefore unlikely that the returns to education in those regions would be large in the early twentieth century. Similarly, the increasing technological sophistication of agriculture, which led states like Iowa to invest heavily in education, was largely a phenomena of the Midwest and Northeast. The South was still able to exploit its large supply of unskilled labor, and the West had relative abundance of raw materials. In general, the correlates and precursors of high rates of return to education were not evenly distributed across the U.S. at the start of the last century.

Not only would changes in demand for skill have varied across regions, but the relative supply of skilled workers was also variable across regions. Investment in education and infrastructure more generally varied considerably, and the high school movement diffused unevenly as well. In the South and Southwest in 1910, high school graduation rates were only four percent, while they were triple that in the Midwest and Pacific Coast, and still higher in New England (Goldin 1999). Finally, the transportation and information technologies at the beginning of the twentieth century were not uniform, and as such the first half of the twentieth century saw a significant integration of regional and local labor markets into a national market, spurred not only by wage controls, but also by the national minimum wage, which spurred capital investment in Southern agriculture (Wright 1987).

When one takes this regional heterogeneity into account there are very good reasons to believe that the Northeast and Midwest had U-shaped returns to education but that other regions of the country had steadily increasing returns to education over the twentieth century. Any discussion of the technology-skill complementarity should take into account such regional differences as they highlight the importance of the technological and capital endowment at the beginning of skill biased technical change. This is especially important for policies in developing countries, who have different capital endowments and may have different short-term responses to skill-biased technological change. We formally develop these ideas of capital-complementing skill-biased technical change in a two-sector model and show that the initial level of capital is an important piece of the returns to education relationship. The model shows that capital-rich markets would experience the largest increase in returns to education and complementing this, skill-rich markets will have the largest increases in the returns to capital.

In describing the origins of skill biased technical change, the rates of return to education in the early twentieth century reported by Goldin and Katz could very well be relative outliers for this time, and without additional data from the early twentieth century we would not know if that is the case. What was lacking was a data source that would allow us to estimate the returns to education by region, to see if significant differences existed. In this paper we use a new data source, a report from the U.S. Commissioner of Education in 1909, to estimate the returns to education of high school teachers in the early twentieth century. Our data list not only the education and earnings of the teachers individually for a number of different states, but also includes actual years of experience in the teaching profession, allowing us to estimate the returns to schooling while controlling for experience directly.[2] These returns are for a single occupation– the absolute levels

[2]Note that the requirements to teach at the secondary level varied greatly in the past, and the professionalization

may understate returns generally, since one of the important gains of schooling comes from enabling workers to choose higher paid occupations. Nonetheless, secondary teachers returns are of interest since they likely reflect the rising demand for high school education relative to the current stock of high school educated workers (the pool of potential teachers). Overall, we find significant regional variation in the returns to education, with large returns congruent to Goldin and Katz's estimates for the Midwest (7%), with substantially lower returns in the South (3%) and West (0.5%).

In considering the generalizability of our main finding, we uncover several facts which strengthen our conclusion. We find that teachers' returns to education are indicative of overall returns to education. The geographical patterns we find hold for male teachers and for less experienced teachers, for whom outside options may be more relevant and may therefore be more closely connected to the wider labor market. Finally, we use IPUMS returns to show that the returns to education for teachers track quite closely with the overall returns to education from 1940 onward, and that the returns to education for the states used in our data track well with the national returns.

The paper proceeds as follows. The next section reviews the facts about regional economic heterogeneity at the beginning of the twentieth century and presents our two sector model which highlights variation in returns to education at the start of skill biased technical change. The third section presents the empirical results, which are based on the 1909 Commissioner of Education Report. They show that there was significant regional variation in the returns to education in the early twentieth century. The fourth section addresses the robustness and extensions of the results of the third section. The final section concludes.

2 Skill, Technology, and the Returns to Education

The existing explanation for the trend in returns to education has not acknowledged, for the most part, the substantial variation in the preconditions for the rise in the returns to education. There are two ways in which the existing theory should be modified to fit the regional histories of the United States. First, the differences in the resource and capital endowments in different regions of the country must be accounted for. Second, the preconditions for the increasing returns to education, as required by the theory, must be reconciled with the historical record. Below, we

of the teaching profession, in terms of certification and degree requirements, did not begin until after the high school movement.

sketch out these two issues, augmenting the theory to yield predictions of the returns to education for different regions of the United States at the beginning of the twentieth century.

2.1 The Historical Record

The differences in the capital and resource endowments in different regions of the U.S. in the early twentieth century are well known. Capital development in South, from the end of the Civil War to World War I, was rather inefficient (Davis 1965, Sylla 1969, Wright 1987, Ransom and Sutch 2001), and financial institutions in the South were not structured in the same way as those in the Northeast and Midwest, with Southern banks much smaller than the national average and with higher interest rates in the South. This is important to the extent that capital markets in the U.S. were segmented in the early twentieth century. The South did not have as many capital intensive industries as the Northeast and Midwest at the beginning of the twentieth century. Similarly, the South, with its sharecropping system and Jim Crow legislation, had a large supply of unskilled labor of both races (Ransom and Sutch 2001, Collins 1998). Furthermore, black unskilled labor was locked in the South by the large flows of immigrants from Europe and racial discrimination in non-farm employment in the U.S. in general (Collins 1998). The West, with its relatively sparse population, had an abundant resource endowment that was only beginning to be exploited in the early twentieth century (Nelson and Wright 1992). In general, this implies that the trade offs where physical capital, raw materials and unskilled labor substituted for skilled labor would have been more prevalent in the South and West since they had an abundance of the former.

The historical record also tells us that the processes that led to the increasing returns to education were less prevalent in the South and West. Given the low levels of capital intensity, and the need for air conditioning in southern manufacturing, there were relatively few of the new large-scale processing technologies highlighted by Goldin and Katz (1998) in the South and West at the end of the nineteenth century. As Wright (1987) has shown, the South was simply not in a position to industrialize (beyond the harvesting of raw materials) in any large scale before the first World War. A possible exception would be textiles, an older industry that has permeated the South before the early twentieth century, was bolstered by cheap Southern labor, and would later exceed its northern competition (Calrson 1981, Wright 1981). Similarly, the South's agriculture, with its dependence on labor-intensive work, was not as sophisticated as the agriculture of the

Northeast and Midwest, nor the cattle ranching seen in the West. Indeed, Goldin and Sokoloff (1984) have argued that industrialization first appeared in the Northeast and Midwest because of the crops grown in those areas, which led to agricultural technology that made women relatively less productive than men in agriculture. In terms of the educational structures necessary to see large returns to education develop, Goldin and Katz (1998) have shown that large investments in education took place most successfully in homogeneous populations, and recent research has shown that households have strong preferences for public goods expenditures to be made over homogeneous populations (Boustan 2005). As such, racial diversity in the South would have caused lower investments in schooling, and it is certainly true that low investments in education left large portions of the southern workforce relatively unskilled at the turn of the last century (Margo 1994).[3]

There is also research that details the extent to which the labor market in the United States was fragmented in the late nineteenth and early twentieth centuries. Rosenbloom (1990, 1996) has shown that the labor market in the early years of the twentieth century was fragmented, and North-South differentials in wages suggest that a national labor market did not exist before the first World War. While it is not true that every locality had its own independent labor market, it is true that the South and North had different labor markets that were not fully integrated to any large degree until after the first World War. Wright (1987) contends that the Southern labor market was not integrated until the New Deal forced the South to invest in capital for the agricultural sector, and that the South was finally brought into the rest of the national labor market by the end of the second World War.

There is also evidence that the usual interpretation of early industrial work is more in line with the view of skill biased technical change than previously thought. Carter and Savoca (1988) have shown that the duration of the average industrial job in the nineteenth century was longer than previously thought. With long tenure at a particular occupation, the increased investment in firm-specific human capital would be expected to be higher, and this would perhaps lead to greater investment in technology by firms. This would naturally favor high returns to education in localities that had a significant industrial presence before the beginning of the twentieth century, further supporting the hypothesis of heterogeneous returns to education in the early twentieth

[3] Bleakley (2006) notes that the eradication of hookworm in the South, begun in 1910, raised the return to education in the region.

century.

This literature suggest that there are both supply and demand factors related to the returns to education in the early twentieth century. Supply factors would include the general education level of the population in a particular region, as it would be evidence that a particular region would have the skill in the workforce to readily adopt to new technologies. Demand factors would relate to the capital stock, both in manufacturing and agriculture, and the share of the labor force employed in manufacturing. If a region had high supply and low demand for skill, we would expect relatively low returns to education. If a region had low supply and high demand for skill we would expect for the returns to education to be large. A priori, we cannot form firm hypotheses about the returns to education in regions that had high supply and high demand or low supply and low demand for skill. While some regions fall easily into high/low or low/high supply/demand categories, direct estimates of the returns to education are necessary for regions with indeterminate predictions for the returns to education based on supply and demand proxies. We also note that, generally, supply of skill is slow to adjust to technology-based demand for skill. As such, levels of skill at a point in time will be exogenous, and returns in the short run would reflect primarily demand factors while long run returns would reflect the endogenous nature of the supply and demand for skill.

Table 1 summarizes these supply and demand factors, presenting evidence on the literacy, manufacturing labor force size, agricultural machinery values, and livestock values in 1900 and 1910.[4] Illinois, Ohio, and Wisconsin have both high literacy rates (each above 95%) and large and growing shares of the labor force employed in manufacturing. The manufacturing share of the labor force in those three states grows by an average of 88% from 1900 to 1910. These summary measures appear to be consistent with a large supply of and very large demand for skilled labor in those states. Iowa has particularly high literacy rates and very high demand for skill (the manufacturing share of the labor force grows by 118% between 1900 and 1910). Texas appears to be somewhat of an anomaly. While they have a relatively low literacy rate, the percentage employed in manufacturing nearly doubles between 1900 and 1910, increasing by 91%. Similarly, the value of machinery per the agricultural workforce increases substantially between 1900 and 1910. This seems to suggest relatively low supply of skill but high demand (nearly as high as in

[4] The literacy rate is also correlated with the graduation rates and secondary school enrollment rates reported by Goldin (1999). This suggest that literacy is a good proxy for other supply of skill factors.

the Midwest) in Texas. Georgia lagged well behind the other states in its literacy rate, suggesting a low supply of skill in the state. Similarly, the growth of Georgia's labor force in manufacturing was relatively modest (only 64%), suggesting relatively low demand for skill in Georgia as well. California has literacy rates in line with the Midwestern states, but the manufacturing sector grows the least in California between 1900 and 1910 (only 51%). This appears to be consistent with high supply of skill in California but low demand for skill in California.

These regional differences do not fit well into a monolithic model of skill-biased technical change and U-shaped returns to education over the twentieth century. The Midwestern states appear to be an example of high supply and high demand for skill, Georgia an example of low supply and low demand for skill, California has high supply but low demand, and Texas has low supply and high demand. Overall, Table 1 shows striking heterogeneity in the factors and proxies related to the return to education. A movement towards new capital and skill-intensive processing technologies would not have the same effect on the relatively unindustrialized South as it would on the Midwest and Northeast. Given these regional differences in the technological endowments at the beginning of the twentieth century, we would expect skill biased technical change to produce regional differences in returns to education in the early twentieth century. How would the predictions of the technology-skill story look like given this historical evidence of significant regional differences? Below, we present a model that captures features of the technology-skill story in a two-sector model, to highlight the importance of factor endowments, especially the size of the capital-intensive sector experiencing technical change.

2.2 A Model of Skill Biased Technical Change with Heterogeneous Endowments

Assume initially that there are two sectors of production, a land-dependent sector, t (for traditional agriculture) and a capital-dependent sector o (for old capital-dependent sector).[5] The two sectors

[5] Here t could represent any sector that is natural resource intensive, but not capital intensive. T would then represent all natural resources.

use skilled and unskilled labor together with either capital K or land T to produce output:

$$\begin{aligned} Y_i &= K_o^{\alpha} H_o^{\beta} L_o^{1-\alpha-\beta} \\ Y_t &= T^{\alpha} H_t^{\beta} L_t^{1-\alpha-\beta} \\ H_t + H_o &= H \\ L_t + L_o &= L \\ K_o &= K \end{aligned}$$

We model each region as a small open economy that takes the relative price of output in each sector as given. We normalize this relative price of output to one. High and low skill labor are mobile across sectors, and so in equilibrium they each get paid their marginal product and these wages are equalized across sectors. Solving the equilibrium labor allocation and wages is straightforward: in equilibrium, the fraction of high-skilled workers employed in the capital-dependent sector is increasing in the capital/land ratio, and equal to the fraction of low-skilled workers employed:

$$\begin{aligned} \frac{K}{K+T} &= \frac{\tilde{H}_o}{H} = \frac{\tilde{L}_o}{L} \\ \frac{\tilde{w}_H}{\tilde{w}_L} &= \frac{\beta}{1-\alpha-\beta} \left(\frac{H}{L}\right)^{-1} \end{aligned}$$

The expressions show that the fraction of workers employed in industry is increasing in the capital/land ratio, and the relative wage of high skill workers is decreasing in their relative abundance. (The tildas signify the initial equilibrium.) Assumption (A1) assures that high skill workers are scarce enough to earn a premium over low skill workers:

$$\frac{H}{L} < \frac{\beta}{1-\alpha-\beta} \tag{A1}$$

Now consider the introduction of a new capital-dependent sector (n):

$$Y_i = A K_n^{\alpha} H_n^{\gamma} L_n^{1-\alpha-\gamma}$$

Since capital is now mobile across the two capital-dependent sectors, the capital constraint becomes:

$$K_o + K_n = K$$

The new capital intensive sector differs from the old sector in that it is more skilled labor-intensive. Mathematically, this assumption is expressed:

$$\gamma > \beta \tag{A2}$$

This assumption captures the skill-biased nature of the new technology.

We show that if the new capital-intenisve technology is a large enough improvement over the old technology, the new equilibrium has the following characteristics.

Proposition 1 *Given (A1)-(A2), if the productivity of the new technology is sufficiently large, the new capital-intensive sector displaces the old capital-intensive sector, and the new capital-intensive technology sector employs a higher fraction of high skilled workers than low skilled workers. That is,*

$$\begin{aligned} \exists A^* \; s.t. \; for \; A &> A^* \\ K_n &= K \\ \frac{H_n}{H} &> \frac{L_n}{L} \end{aligned}$$

Proposition 2 *Given (A1)-(A2), given the same level of productivity, the number of high skilled employed in the new capital-intensive technology exceeds the number of high skilled previously employed in the old capital-intensive technology. The relative wage of high skilled workers also exceeds the previous relative wage. That is,*

$$\begin{aligned} for \; A &> A^* \\ H_n &> \tilde{H}_o \\ \frac{w_H}{w_L} &> \frac{\tilde{w}_H}{\tilde{w}_L} \end{aligned}$$

Furthermore, if the productivity is even larger, the number of low skill workers in the new capital-intensive technology exceeds the number employed in the old capital-intensive technology. In partic-

ular,

$$\begin{aligned} \exists \hat{A} &> A^* \ s.t. \ for \ A > \hat{A} \\ L_n &> \tilde{L}_o \end{aligned}$$

Proposition 3 *Given (A1)-(A2), the higher the capital/land ratio, the higher the fraction of skilled and unskilled workers employed in the new capital-intensive technology and the higher the relative wage of high-skilled workers. That is,*

$$\begin{aligned} for \ A &> A^* \\ \frac{dH_n}{d(K/T)} &> 0 \\ \frac{dL_n}{d(K/T)} &> 0 \\ \frac{d(w_H/w_L)}{d(K/T)} &> 0 \end{aligned}$$

Proposition 4 *Given (A1)-(A2), the introduction of the new technology raises the return to capital relative to land. Furthermore, the higher the ratio of skilled/unskilled labor, the larger is this increase in the relative rental rate of capital.*

$$\begin{aligned} for \ A &> A^* \\ \frac{R_K}{R_T} &> \frac{\tilde{R}_K}{\tilde{R}_T} \\ \frac{d(R_K/R_T)}{d(H/L)} &> 0 \end{aligned}$$

Proof. See appendix. ■

Together Propositions 1 and 2 show that the model replicates the story of Goldin and Katz (1998). That is, the new capital-dependent sector expands, increasing the relative demand for skilled workers and also their relative wage. If the new technology is a dramatic enough advance it furthers industrialization.[6] This is the standard skill-biased technical change.

Proposition 3 has strong implications that predict higher returns and more labor employed in

[6] Given the static nature of the problem, Proposition 1 implies extreme displacement. In the real world, the changeover of capital from the old to new technology is clearly a slower process.

the new technology in areas with high relative endowments of capital. Thus, there will be variation in the returns to education that go hand-in-hand with the nature and extent of industrialization before the technological change. The result is entirely intuitive– the region that is technologically backward sees little increase in the returns to education because the technological change is skill intensive, but the backwards region has little of either the old or new capital-intensive technologies. In order for batch processing and electrification to induce high returns to education, there had to be industries that could implement and successfully take advantage of the new technologies. Proposition 3 highlights the role that the technological endowment has with the return to education. If a region did not have the infrastructure or extensive industry before the diffusion of skill intensive technologies, it would not lead to large returns to education in that region. Proposition 3 therefore provides us with the central test of the theory in the next section. Also note that areas that had a large endowment of capital intensive agriculture should also have high returns to education as well.

Finally, Proposition 4 shows how the new technologies increased the incentives to invest in physical capital, especially in areas with high levels of human capital. The model therefore offers an explanation for increased levels of industrialization experienced in the first half of the century, but faster industrialization in the Northeast, Midwest and West (where schooling levels were high) than in the South, where they were lower. We test these implications in the following section.

3 The Returns to Education in the Early 20th Century

3.1 Data

We estimate the returns to education with a new and unique data source, a 1909 report from the U.S. Commissioner of Education which allows us to estimate the returns to education of secondary teachers in the early twentieth century. This data is, to our knowledge, the earliest which can be used to estimate the returns to education in the United States, and the only source that can capture geographic variation in the returns to education before the 1940 Census. The data come from a report prepared for then U.S. Commissioner of Education Elmer Ellsworth Brown on the labor force of teachers. The report, entitled *"The Teaching Staff of Secondary Schools in the United*

States" by Edward L. Thorndike was the first report in a five-year, five-report plan to collect data on secondary education (the other four focused on the student body, curriculum, finances, and special education, respectively). It presented tabulated data on the (i) income, (ii) experience, (iii) education, and (iv) gender of U.S. secondary school teachers in 1908. The data were collected via survey for approximately five thousand teachers, chosen to be a representative sample of the nation's secondary teaching workforce at the time.

The data were collected using a two-part survey sent by the Office of Education to administrators for a sample of secondary schools. The first survey collected the salaries, years of secondary and post-secondary education, and actual years of experience of all teachers in the schools surveyed. The fact that years of experience are directly reported is a major strength of the data, since imputed "potential years of experience" (i.e., the traditional $age - years\ of\ schooling - 6$) can diverge strongly from actual experience. This is particularly true for women, who are not as closely tied to the labor force and who constituted a significant share of secondary school teachers at the time. The second survey was a follow up survey sent with the intent of measuring any biases or measurement error in years of education (e.g. adding in primary schooling) and experience (e.g. reporting years of service at the particular school surveyed). Thorndike spent a great deal of effort discussing potential sources of measurement error and trying to quantify or minimize them. The second survey showed that the larger initial survey did not suffer from any aggregate biases. The data we use comes from the first survey.

In general, the data appear to be of extremely high quality. There appear to be no more than the standard measurement errors that arise in all survey data. For example, Thorndike mentions that income may vary somewhat due to varying lengths of the school year, such that low salaries in the South are partially explained by shorter school years. The data would nonetheless reflect the actual income received. Thorndike also mentioned that private schools who underpay may feel pressure to overestimate their incomes. For years of education, the distinction between secondary and post-secondary education was not always clear, but this will not affect our results since we look only at the sum of these two. For experience, Thorndike mentioned a tendency to report roughly and to include the current year of service.

Unfortunately, we do not have the original survey returns, only the processed data from the report. We focus on two sets of tables (Tables 7-10 in the original Thorndike Report) for our purposes. The first is individual public school teacher data tabulated separately for California, Georgia, and Texas, and tabulated together for Illinois, Ohio and Wisconsin.[7] These states cover the West Coast (California), Southwest (Texas), Southeast (Georgia), and Midwest (Illinois, Ohio and Wisconsin). These tables allowed creation of a dataset of teachers including their state, gender, income, experience, and education levels. The second set of tables gives separate details on Illinois, Ohio and Wisconsin teachers, but only provides the median income level for difference experience-education-gender cells. This allows creation of a dataset of median incomes by gender, experience, education and state for all six states separately.[8] We combine this data with IPUMS data on the industrial composition of the workforce in 1910, 1920 and 1930, and later information on teachers and workers earnings in Census records from 1940 on.

3.2 Summary Analysis

Table 2 presents summary statistics for the individual data by state and gender. Several interesting observations can be gleaned from these. First, secondary teachers averaged 12.6 to 13.8 years of education, having completed not much beyond high school education themselves. Men and women had similar levels of education with women having slightly more education on average in the Midwest states, and men having a slight advantage in the other states. Education levels are also fairly similar across states, with the exception of California, whose secondary teachers had an additional year of education on average.

Second, average salary levels vary greatly, both between men and women and across states. For example, men in California earn about three times what women in Georgia earn. Salaries in California are substantially higher (roughly $300/year or 35% higher) than in Georgia and the Midwest states, while those in Texas are significantly lower (about $100/year or 12%). As expected, women earn lower salaries in all states with the largest difference in Georgia and the smallest differences in the Midwestern states. Beyond gender discrimination, a possible reason for this male wage premium is that a significant number of male secondary teachers performed a dual role of teacher and administrator. Thorndike notes this fact in his report, although we cannot

[7] Thorndike explained that the data were calculated together because the data were similar.

[8] All data from the tables was entered twice, in separate files, to assure accurtate data entry. The use of tabulated data does introduce additional sources of measurement error in the data as both income and experience are grouped into small ranges. Neither of these should substantially change our estimates of the returns to education, and indeed replicating the corresponding groupings in the U.S. census data does not alter the results substantially.

distinguish in the data which teachers were also administrators.

Third, teachers average between 8.2 and 9.6 years of experience, with male teachers having on average 2.0 to 3.6 more years of experience. The additional years of experience may also be related to the previously-mentioned dual teacher-adminstrator role that males often play. Finally, it should be noted that secondary teaching is a mixed-gender occupation. For the sample overall about half (fifty five percent) of the teachers are men. In the Southern states of Georgia and Texas, this is closer to 2/3 of teachers, while in California women constitute 2/3 of secondary school teachers in the data from that state.

3.3 Regional Variation in the Returns to Education

We estimate the returns to education using a standard Mincerian regression

$$\log(w) = \alpha + \beta_1 s + \beta_2 x + \beta_3 x^2 + \beta_4 g + \varepsilon$$

where w is the wage of a person with years of schooling s, years of experience x, and gender g. Table 3 presents the regression results for each of the states. The estimates show considerable geographic variation in the Mincerian return to schooling.[9] The three Midwestern states and Texas had high returns, 7.0 and 7.1 percent, respectively. Recall that the Midwestern states had high levels of industry, and so likely rapidly rising demand for skill, while in Texas, rapidly growing manufacturing created high demand, especially given the low education levels in Texas. In contrast, the returns are much lower in Georgia and especially California. The return in Georgia is just 3.3 percent which is significantly different from the returns in Texas and the Midwest states, despite the smaller sample size in Georgia and the consequently larger standard error. This suggest that the low supply of educated labor and low demand for skill in Georgia combined to yield low returns to education. The return in California is a miniscule 0.5 percent and not statistically significant. Recall that California teachers averaged 1.2 more years of education than teachers in the other states, and that they also had relatively high literacy rates, reflective of high education levels.

[9] We formally tested for differences in the regional returns to education in a pooled regression (unreported), in which we rejected the hypothesis of equal returns between the Midwest and South. As the purpose of this study is to document the extent of the variation and to estimate the returns by region, we present the separate regressions throughout.

As described in the previous section, the individual data for Illinois, Ohio and Wisconsin is pooled together. We do, however, have data on median incomes (by sex, education, and experience) separately for each of these three states. Regressions can therefore be run separately using median-level data for these three as a way to disaggregate the returns to education estimates. A key question in interpreting this data is the extent to which estimates from median-level regressions are comparable to individual-level regression results. To answer this question, we construct comparable median-level data for the states that have individual data and compare regression results. If the estimates for the returns to education are similar in the median and individual regressions, then we would surmise that median regressions for the individual Midwestern states will give us reliable estimates of the returns to education for each Midwestern state. Table 4 shows that median regressions do in fact express much of the same information about returns to schooling that individual-regressions do and the qualitative interpretations remain the same. Focusing on the schooling coefficients, Mincerian returns are low in California and Georgia, and relatively high elsewhere.[10] Still, comparing across states, the median estimates show the higher wages in California.[11] We conclude that the qualitative patterns in the median-level estimates are strongly indicative of patterns in the individual-level estimates. As such, we believe that estimates for the return to education for each Midwestern sate will not be biased by the median representation of the data.

The median level estimates are presented separately for Illinois, Ohio and Wisconsin in Table 5. The main lesson from Table 5 are that wage returns to schooling found in the Midwest using the individual-level data do not appear to be at the same high level across all three states. Mincerian returns are high in Illinois and Ohio, but much lower in Wisconsin. These lower Mincerian returns are accompanied by higher wage levels in Wisconsin, as evidence by the significantly larger intercept. Thus, even amongst similar states in the same geographic region, there appear to be important differences in the returns to schooling.

Overall, there is substantial variation in the returns to education for these secondary teachers in 1909. The returns in the Midwest are significantly greater than the returns in the South, suggesting

[10] Focusing on the constant term results, we see that constant terms are somewhat higher in the median regressions. The difference in levels is not surprising since the individual- and median-level regressions weight individuals differently; the median regressions give each experience-education cell equal weight, while the individual data use the weights in the sample population.

[11] The patterns by sex across states also match up well. The one exception is that returns are typically lower for women in the median-level estimates and slightly higher for men.

that high supply and high demand for skill regions will have high rates of return while low supply and low demand regions will have low rates of return to education. Consistent with our earlier predictions, high supply and low demand regions such as California have low returns to education while low supply and high demand regions such as Texas have high returns. If these returns to education are indicative of general returns to education for these regions the theory of U-shaped returns to education over the twentieth century would have to be augmented to reflect this regional heterogeneity. In the next section we consider the generalizability and robustness of the results presented in this section.

4 Extensions and Implications

4.1 Generalizability

Applying the evidence for secondary teachers to our story of relative endowments in the overall economy raises the question of whether these estimated returns to education are informative about the returns to education in the labor force overall. Specifically, does variation in teachers' returns to education track with the variation in returns to education of the labor force overall? To answer this question we must look at comparisons of teachers with workers more generally in the U.S. We use U.S. census data from IPUMS to underscore the relationship between teachers' returns and overall returns to education.

To gain insight into the question of geographic variation, we compare (by state) the returns to education experience for teachers and the general labor force in later Census data. Since teachers are not oversampled, only the 5% samples (available in 1980, 1990, and 2000) have enough teachers to reasonably estimate returns to teachers by state with any reasonable precision. Using these IPUMS returns, we estimate for each state a return to education for both teachers and workers overall. We then regress the overall returns on teacher returns to determine the relationship. The regression takes the form

$$\rho_{G,M,t} = \alpha + \beta_1 \rho_{\tau,M,t} + \sum \delta_t D_t + \varepsilon$$

where ρ_G are the overall returns to education, ρ_τ the returns for teachers, M indicates state,

D is a time dummy for each year, and t is the Census year. In this regression, each state-year combination is an observation, but we allow for fixed time dummies so that it is the geographic variation driving the results. Table 6 presents the results of a regression of the estimates of overall returns on teachers' return estimates. The positive coefficient of 0.09 on teachers' return is small but statistically significant, indicating that, while the relationship is not extremely tight, the data nevertheless give a valuable indication that overall returns to education have varied significantly by state to the extent that teachers' returns tracked with overall returns.

4.2 Robustness

The robustness of the estimates of the returns to education reported in the previous section must be established. It is important to establish, first, that the returns reported here indeed reflect the return to education across region and not another measure that varies by region such as teacher salary. If the returns simply reflected salary differences by region for teachers, we would not predict the very low returns in California given the substantially higher salaries in that state. Similarly, the high returns in Texas would not be consistent with the low salaries in that state and the similar years of schooling in Texas and in the Midwestern states. We believe that the extent to which the returns to education vary across regions in ways that are not predicted by the summary statistics in Table 2 establishes that the returns reported here are indeed estimates of the return to education and not simply teacher salary variation across states.

We have argued that the geographic variation uncovered in teacher's returns to schooling is indicative of variation in returns in the overall labor force. We use two checks to test the robustness of this assumption. First, we estimate the returns for men only. We divide the sample by gender for two reasons. First, women teachers may have had fewer outside options in the broader labor market. Competition for their services may not have been strong enough for teacher's returns to reflect returns overall. Men would have been more integrated into the market, however, and would have had fewer restrictions placed on their supply of labor. Secondly, Carter and Savoca (1991) have suggested that different levels of education and wages by gender were due to the fact that women were expected to be less attached to the labor market than males, making it unwise to invest heavily in education and lowering the wages that they received in the labor market. Although this point is related to the first, it also suggest that the education of women in teaching occupations would be different from those of men, which was shown earlier. To the extent that variation in schooling identifies the returns to education in a Mincerian regression, separating the sample by gender would tell us if the total returns were biased.

Second, we estimate the returns for teachers with few years of experience. We focus on teachers with little experience because these teachers would presumably have invested less in teacher-specific human capital, and so would hold relatively more general human capital for potential use in the broader labor market. The general idea is that the returns to education for teachers with more experience in teaching may not reflect labor market conditions as much as they would reflect occupation or firm-specific investments or skills. We also posit that teachers with less experience also have more and varied outside options. We take "little experience" to be five years or less in the teaching profession. Considering that teachers in the sample averaged more than eight years of experience, this cut-off certainly captures the less experienced teachers while at the same time being a large enough sample to yield robust estimates of the returns to education for the group of teachers with the least attachment to the profession.

Table 7 shows that the pattern in overall returns shows up in men's returns as well (despite the fact that women were an important fraction of teachers) and in returns for the young. The returns in California and Georgia are low, while those in the Midwest and Texas were high. These results further support the contention that our estimates of the return to education do not simply reflect regional salary differentials. The returns for men in Georgia would be higher than those in Table 7 if their high salaries and the same average schooling, as reported in Table 2, were used to predict the return to education. Overall, the results of Table 7 give us further confidence that the geographic variation in teachers' returns to schooling reflect geographic variation in schooling returns of the workforce overall.

4.3 Secular Implications

A final question is the question of secular variation: what can these data tell us about the overall returns to schooling in 1909 relative to 1940? We use IPUMS census data to confirm the relationship between teacher's returns and overall returns over time, and then compare our 1909 results with later results. One caveat is that the census occupational code for teachers includes all teachers (except for professors/instructors and music, dance or art teachers), and not just secondary teachers. To the extent that education selects people into higher paying secondary school teaching, the IPUMS data will overstate the return to schooling *within* secondary education and thus exceed our estimates.

Figure 1 answers the question about secular variation, showing that the relationship between teachers' returns and overall returns is strong over time. The four different series represent teachers' returns for the states we examine, teachers' returns for all states, overall returns for the states we examine, and overall returns for the nation as a whole. Again the number of teachers in the sample states is relatively few (especially in the 1950 census), so we present robust regression results.[12] All four series move substantially together with a mid-century decline followed by rising returns. Indeed, the results for all workers across the nation and all workers in the sample states are nearly identical. While the estimates of teachers' returns in the sample states have perhaps the weakest relationship with overall returns across the nation, the relationship is still quite strong. The correlation between the two series is 0.81 and a regression of overall returns on teachers' returns in the sample states explains 66 percent of the variation in overall returns. We therefore again conclude that comparing teachers' returns over time can give us a strong indication of patterns in overall returns over time.

Table 8 does precisely this, comparing the 1909 return to several benchmarks from the 1940 census. The 1909 return is based on a weighted regression of the individual data in 1909. Since the sample sizes varied greatly over region and were not entirely representative, the weights were chosen to make the sample representative of the sample of teachers in 1940. This required weighting the California sample by a factor of 1.92, the Georgia sample by a factor of 1.94, and the Texas sample by a factor of 1.74, and weighting the Midwest sample by a factor of 0.59. The resulting estimate for the return to education was 8.3 percent in 1909. At first glance, the returns seem quite high for a within-occupation return to schooling. Comparing with 1940, however, the return is not overly high. Indeed it is slightly less than the Mincerian return of 9.1 percent estimated for teachers in these same states in 1940 though not significantly different. The returns for all workers in these states were somewhat higher at 9.6 percent per year in 1940, while those for the nation overall were

[12] Robust regressions incorporate a recursive algorithm for reweighting observations that downweights outliers that have too strong an influence on regression results. Robust regressions produced substantially lower estimates than OLS in 1950 (0.075 vs. 0.096), but otherwise similar results. Robust regression also has little effect on the 1909 sample estimates.

slightly lower at 8.9 percent.

Using 1940 as a benchmark, we would surmise that since the return to schooling for teachers in the sample states was representative of the returns to workers for the nation overall, that the same is true for 1909. In this case, returns in 1909 would be relatively high (since 1940 preceded the Great Compression and was a year of relatively high returns to schooling), but lower than the most comparable evidence, the returns in Iowa in 1914. Recall the caveat that the returns from 1940 are for all teachers, not just secondary teachers. The 1940 teachers' regressions include both primary and secondary teachers, while the 1909 estimates are based on only secondary teachers. Secondary teachers tend to be more educated and substantially better paid. To the extent that schooling enables teachers to sort into higher paying secondary education jobs, the 1940 estimates would be biased upward as an estimate of the return to education of secondary teachers.

5 Conclusion

We have shown, using historical evidence on the returns to education for secondary teachers in the U.S., that the returns to education showed marked geographic variation. In our model of skill biased technical change, we showed how returns to education vary with technological change. The shape of these changes over time, we have argued, will be related to the technological endowment before the skill-biased technical change. Our model predicted that regions with greater degrees of capital intensity would experience higher returns to education than those with less capital intensive endowments. With skill biased technical change, we should expect U-shaped returns in regions with large capital endowments, but steadily increasing returns in regions with relatively small capital endowments. Our data on the returns to skill for secondary teachers in the very earliest part of the twentieth century is consistent with that predictions. Teachers in the Midwest had greater returns to education than those in the South. Furthermore, we found that this result is robust– the returns to teachers tracks with the returns to skill more generally, and our result was robust to considering only men and younger teachers. In sum, we find strong evidence that returns to education were large in 1909 in the Midwest, consistent with Goldin and Katz, but that they may have varied considerably across states. As such, the study of U-shaped returns to education should be modified to reflect the fact that returns for some regions would rise continuously throughout the twentieth century.

The variation in returns to education has important implications for the study of the returns to skill more generally, and for education and immigration policies in many developing nations in particular. Rather than states or regions of one nation, our model easily generalizes to different nations, where nations with large capital endowments will see large returns to education, while those with relatively small capital endowments will see small returns initially. As Uwaifo (2005) has shown, there is considerable debate over the size and shape of the returns to education in Sub-Saharan Africa, and her estimates of the return to education in Nigeria in the 1990s are similar to the returns we found in Georgia in 1909. Our results suggest that while skill biased technological change will eventually lead to universal large returns to skill in the long run as markets integrate and capital intensity diffuses, in the short run nations with relatively small capital endowments may see negligible returns to education. This has important implications for immigration and emigration policies in nations with low resource endowments– to encourage their educated workforce to remain when the returns to education are low in their home country, but large in other parts of the world.

References

[1] Bleakley, Hoyt (2006). "Disease and Development: Evidence from Hookworm Eradication in the American South." Forthcoming, *Quarterly Journal of Economics.*

[2] Boustan, Leah P. (2005). "Inside the Black Box of 'White Flight': The Role of Suburban Political Autonomy and Public Goods." Mimeo, Harvard University.

[3] Carlson, Leonard (1981). "Labor Supply, the Acquisition of Skills and the Location of Southern Textile Mills, 1880-1900," *Journal of Economic History* 41 (1): 65-71.

[4] Carter, Susan and Elizabeth Savoca (1988). "Labor Mobility and Lengthy Jobs in 19th Century America." *Journal of Economic History* 50 (1): 1-16.

[5] Collins, William J. (1998). "When the Tide Turned: Immigration and the Delay of the Great Black Migration." *Journal of Economic History* 57 (3): 607-632.

[6] Davis, Lance (1965). "The Investment Market, 1870-1914: The Evolution of a National Market." *Journal of Economic History* 25 (3): 355-393.

[7] Goldin, Claudia D. (1999). "A Brief History of Education in the United States." NBER Working Paper No. h119.

[8] Goldin, Claudia D. and Lawrence F. Katz (1998). "The Origins of the Technology-Skill Complementarity." *The Quarterly Journal of Economics* 113 (3): 693-732.

[9] Goldin, Claudia D. and Lawrence F. Katz (1999) "The Returns to Skill in the United States across the Twentieth Century." NBER Working Paper No. 7126.

[10] Goldin, Claudia D. and Lawrence F. Katz (2000). "Education and Income in the Early Twentieth Century: Evidence from the Prairies." *Journal of Economic History* 60 (3): 782-818.

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

[12] Goldin, Claudia D. and Kenneth Sokoloff (1984). "The Relative Productivity Hypothesis of Industrialization: The American Case 1820 to 1850." *The Quarterly Journal of Economics* 99 (3): 461-488.

[13] Margo, Robert A. (1994). *Race and Schooling in the South, 1880-1950: An Economic History.* Chicago: University of Chicago Press.

[14] Nelson, Richard R. and Gavin Wright (1992) "The Rise and Fall of American Technological Leadership: The Postwar Era in Historical Perspective," *Journal of Economic Literature*, 30 (4):1931-1964.

[15] Ransom, Roger L. and Richard Sutch (2001). *One Kind of Freedom: The Economic Consequences of Emancipation.* New York: Cambridge.

[16] Rosenbloom, Joshua L. (1990). "One Market or Many? Labor Market Integration in the Late Nineteenth-Century United States." *Journal of Economic History* 50 (1): 85-108.

[17] Rosenbloom, Joshua L. (1996). "Was there a National Labor Market at the End of the Nineteenth Century? New Evidence on Earnings in Manufacturing." *Journal of Economic History* 56 (3): 626-656.

[18] Ruggles, Steven and Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004.

[19] Sylla, Richard (1969). "Federal Policy, Banking Market Structure, and Capital Mobilization in the United States, 1863-1913." *Journal of Economic History* 39 (4): 657-86.

[20] Thorndike, Edward L. (1909) "The Teaching Staff of Secondary Schools in the United States" United States Bureau of Education Bulletin No. 4. Washington, D.C.: Government Printing Office.

[21] United States Bureau of the Census (1917) Statistical Abstract of the United States, 1916, 39th edition. Washington, D.C.: Government Printing Office.

[22] United States Bureau of the Census (1925) Statistical Abstract of the United States, 1924, 47th edition. Washington, D.C.: Government Printing Office.

[23] Uwaifo, Ruth (2005). "Africa's Education Enigma? The Nigerian Story." Mimeo, University of California, Berkeley.

[24] Wright, Gavin (1981). "Cheap Labor and Southern Textiles, 1880-1930." *Quarterly Journal of Economics* 96 (4): 605-629.

[25] Wright, Gavin (1987). "The Economic Revolution in the American South." *Journal of Economic Perspectives* 1 (1): 161-178.

[26] Wright, Gavin (1990). "The Origins of American Industrial Success." *American Economic Review* 80 (4): 651-668.

6 Appendix

6.1 Proof of Proposition 1:

A) Proving $K_n = K$

Since the land-intensive technology and the old capital-intensive technologies have the same factor shares, it can be readily shown that they will always employ the same ratio of inputs. It is also trivial to show that K_n, H_n and L_n are increasing in A. Thus, A^* can be derived as the level that equates the marginal return to capital in the new capital-intensive technology when all capital is employed in that sector (i.e., $K_n = K$) to the marginal return to land in the land-intensive sector (which equals the potential marginal return to capital in the old capital-intensive sector):

$$\begin{aligned} R_{K,n}(A^*) &= R_{K,o}(A^*) \\ A^*\alpha K^{\alpha-1} (f_H H)^{\gamma} (f_L L)^{1-\alpha-\gamma} &= \alpha T^{\alpha-1} ((1-f_H) H)^{\beta} ((1-f_L) L)^{1-\alpha-\beta} \end{aligned}$$

Solving for

$$A^* = \frac{(1-\alpha-\beta)^{1-\alpha-\beta} \beta^{\beta}}{(1-\alpha-\gamma)^{1-\alpha-\gamma} \gamma^{\gamma}} \left(\frac{H}{L}\right)^{\beta-\gamma} \left[\frac{(1-\alpha)(T+K)}{(\beta T+\gamma K)} - 1\right]^{\beta-\gamma}$$

For all levels higher than A^*, $R_{K,n}(A^*) = \tilde{R}_{K,o}(A^*)$ and so $K_n = K$.

B. Proving $\frac{H_n}{H} > \frac{L_n}{L}$

We prove by contradiction. Defining $f_H \equiv \frac{H_n}{H}$ and $f_L \equiv \frac{L_n}{L}$ we assume $\frac{H_n}{H} \leq \frac{L_n}{L}$, which is $f_H < f_L$.Optimality again requires

$$\begin{aligned} w_H &= \gamma A \left(\frac{K}{f_L L}\right)^{\alpha} \left(\frac{f_H H}{f_L L}\right)^{\gamma-1} \\ &= \beta \left[\frac{T}{(1-f_L) L}\right]^{\alpha} \left[\frac{(1-f_H) H}{(1-f_L) L}\right]^{\beta-1} \qquad (1) \\ w_L &= (1-\alpha-\gamma) A \left(\frac{K}{f_L L}\right)^{\alpha} \left(\frac{f_H H}{f_L L}\right)^{\gamma} \\ &= (1-\alpha-\beta) \left[\frac{T}{(1-f_L) L}\right]^{\alpha} \left[\frac{(1-f_H) H}{(1-f_L) L}\right]^{\beta} \qquad (2) \end{aligned}$$

Dividing the two equations by each other yields:

$$\begin{aligned} \frac{\gamma}{(1-\alpha-\gamma)} \frac{(1-f_H)}{(1-f_L)} &= \left(\frac{f_H}{f_L}\right) \frac{\beta}{(1-\alpha-\beta)} \\ \frac{\gamma}{(1-\alpha-\gamma)} &\leq \frac{\beta}{(1-\alpha-\beta)} \\ \gamma &\leq \beta \end{aligned}$$

But $\gamma > \beta$, by assumption.

6.2 Proof of Proposition 2

A) Proof of $H_n > H_o$

Again, one can trivially show that H_n is increasing in A, so it suffices to show that at $A = A^*$, $H_n > H_o$. We prove equivalently that $f_H > \tilde{f}_H$. Consider the first order conditions above. Dividing the top by the bottom yields and expression for which we define an implicit function:

$$g_1(f_H, f_L) = \left[\frac{\frac{\gamma}{(1-\alpha-\gamma)}\frac{(1-f_H)}{f_H} -}{\frac{(1-f_L)}{f_L}\frac{\beta}{(1-\alpha-\beta)}} \right] = 0$$

It is trivial to show that $\partial g_1/\partial f_H < 0$ and $\partial g_1/\partial f_L < 0$. Since $f_L < f_H$, it suffices to show that $g_1(\tilde{f}_H,\tilde{f}_L) > 0$. Substituting in $\tilde{f}_H = \tilde{f}_L = K/(T+K)$ yields

$$g_1(\tilde{f}_H, \tilde{f}_L) = \frac{T}{K}\left(\frac{\beta}{(1-\alpha-\beta)} - \frac{\gamma}{(1-\alpha-\gamma)}\right) < 0$$

since $\gamma > \beta$.

B) Proof of $\frac{w_H}{w_L} > \frac{\tilde{w}_H}{\tilde{w}_L}$

We prove by contradiction assume:

$$\begin{aligned}
\frac{w_H}{w_L} &< \frac{\tilde{w}_H}{\tilde{w}_L} \\
\frac{\beta}{(1-\alpha-\beta)}\left(\frac{(1-f_H)\,H}{(1-f_L)\,L}\right)^{-1} &< \frac{\beta}{(1-\alpha-\beta)}\left(\frac{\tilde{f}_H H}{\tilde{f}_L L}\right)^{-1} \\
\left(\frac{1-f_H}{1-f_L}\right)^{-1} &< 1 \\
f_H &< f_L
\end{aligned}$$

which contradicts Proposition ~~1~~.

C) Proof of $L_n > \tilde{L}_o$ for $\forall A$ for $\hat{A} > A^*$

It is trivial to show that both H_n and L_n are increasing in A. We show that $H_n + L_n < \tilde{H}_o + \tilde{L}_o$ for $A = A^*$ and then derive $\hat{A}$.

Assume $A = A^*$ and $L_n > \tilde{L}_o$. By construction at A^*, the marginal product of capital and land are equated, as are the marginal product of low skilled workers:

$$\begin{aligned}
\alpha A^* K^{\alpha-1}\left(f_H H\right)^{\gamma}\left(f_L L\right)^{1-\alpha-\gamma} &= \\
\alpha T^{\alpha-1}\left(\left(1-f_H\right)H\right)^{\beta}\left(\left(1-f_L\right)L\right)^{1-\alpha-\beta}
\end{aligned}$$

$$\begin{aligned}
\left(1-\alpha-\gamma\right)A^* K^{\alpha-1}\left(f_H H\right)^{\gamma}\left(f_L L\right)^{-\alpha-\gamma} &= \\
\left(1-\alpha-\gamma\right)T^{\alpha-1}\left(\left(1-f_H\right)H\right)^{\beta}\left(\left(1-f_L\right)L\right)^{-\alpha-\beta}
\end{aligned}$$

Dividing these two expressions by each other yields:

$$\begin{aligned}
\frac{1}{(1-\alpha-\gamma)}\left[\frac{f_L L}{K}\right] &= \frac{1}{(1-\alpha-\beta)}\left[\frac{(1-f_L)\,L}{T}\right] \\
\frac{T}{(1-\alpha-\gamma)}\left(1-f_L\right) &= \frac{K}{(1-\alpha-\beta)}f_L \\
\frac{(1-\alpha-\gamma)\,K}{(1-\alpha-\gamma)\,K + (1-\alpha-\beta)\,T} &= f_L
\end{aligned}$$

The expressions can be solved for f_L and $f_{H:}$. Now we start with the assumption:

$$\begin{aligned}\frac{(1-\alpha-\gamma)K}{(1-\alpha-\gamma)K+(1-\alpha-\beta)T} &> \frac{K}{T+K} \\ \gamma &< \beta\end{aligned}$$

which contradicts (A2)

We now derive $\hat{A}$ by assuming:

$$\hat{f}_L = \tilde{f}_L = \frac{K}{T+K} \tag{3}$$

and solving for the implied $\hat{A}$. The first order conditions for high and low skilled labor again yield the following expression:

$$\begin{aligned}\left(\frac{1-\alpha-\gamma}{\gamma}\right)\frac{\left(1-\hat{f}_L\right)}{\hat{f}_L} &= \left(\frac{1-\alpha-\beta}{\beta}\right)\frac{\left(1-\hat{f}_H\right)}{\hat{f}_H} \\ \left(\frac{1-\alpha-\gamma}{\gamma}\right)\frac{T}{K}\hat{f}_H &= \left(\frac{1-\alpha-\beta}{\beta}\right)(1-f_H) \\ \hat{f}_H &= \frac{\gamma(1-\alpha-\beta)K}{[\gamma(1-\alpha-\beta)K+\beta(1-\alpha-\gamma)T]}\end{aligned}$$

Substituting $\hat{f}_L$ and $\hat{f}_H$ into the first order condition on high-skilled labor, we solve for $\hat{A}$:

$$\hat{A} = \frac{\beta}{\gamma}\left[\frac{\beta(1-\alpha-\gamma)(T+K)}{[\gamma(1-\alpha-\beta)K+\beta(1-\alpha-\gamma)T]}\frac{H}{L}\right]^{\beta-\gamma}$$

6.3 Proof of Proposition 3

The relative wage equals the ratio of the marginal products in agriculture, which can be simplified to:

$$\begin{aligned}w_H/w_L &= \frac{\beta}{1-\alpha-\beta}\left(\frac{L_a}{L_a}/\frac{L}{H}\right) \\ \frac{d\log(w_H/w_L)}{d\log(K/T)} &= \frac{d\log(1-f_L)}{d\log(K/T)} - \frac{d\log(1-f_H)}{d\log(K/T)}\end{aligned}$$

so we proceed by showing that $\frac{d\log(1-f_H)}{d\log(K/T)} < \frac{d\log(1-f_L)}{d\log(K/T)} < 0$, which implies $\frac{dH_n}{d(K/T)} > 0$ and $\frac{dL_n}{d(K/T)} > 0$, and, by the above equation, $\frac{d(w_H/w_L)}{d(K/T)} > 0$. To simplify presentation, we change notation to work directly with the fractions of labor in agriculture, ${}^a f_L \equiv 1-f_L$ and ${}^a f_H \equiv 1-f_H$, and use the implicit function defined by the log of the first-order conditions for comparative statics

$$
\begin{aligned}
\log\left(\frac{\gamma}{\beta}\right)+\log A+(\alpha+\beta-1)\log^{a} f_L+ & \\
(1-\alpha-\gamma)\log\left(1-^{a} f_L\right)+(1-\beta)\log^{a} f_H+ & \\
(\gamma-1)\log\left(1-^{a} f_H\right)+\alpha\log\left(\frac{K}{T}\right)+(\gamma-\beta)\log\left(\frac{H}{L}\right) &= 0 \\
\log\left(\frac{1-\alpha-\gamma}{1-\alpha-\beta}\right)+\log A+(\alpha+\beta)\log^{a} f_L+ & \\
(-\alpha-\gamma)\log\left(1-^{a} f_L\right)-\beta\log^{a} f_H+ & \\
\gamma\log\left(1-^{a} f_H\right)+\alpha\log\left(\frac{K}{T}\right)+(\gamma-\beta)\log\left(\frac{H}{L}\right) &= 0
\end{aligned}
$$

Now solving the first order conditions for the change $d\log^{a} f_H$ and $d\log^{a} f_L$ as $\log(K/T)$ yields the following system of equations:

$$
\begin{bmatrix}
\left[\begin{array}{c}(1-\beta)+\\(1-\gamma)\left(\frac{^{a}f_H}{1-^{a}f_H}\right)\end{array}\right] & \left[\begin{array}{c}(\alpha+\beta-1)-\\(1-\alpha-\gamma)\left(\frac{^{a}f_L}{1-^{a}f_L}\right)\end{array}\right] \\
\left[\begin{array}{c}-\beta-\\\gamma\left(\frac{^{a}f_H}{1-^{a}f_H}\right)\end{array}\right] & \left[\begin{array}{c}(\alpha+\beta)+\\(\alpha+\gamma)\left(\frac{^{a}f_L}{1-^{a}f_L}\right)\end{array}\right]
\end{bmatrix}
\begin{matrix}\frac{d\log^{a} f_H}{d\log(K/T)}\\ \frac{d\log^{a} f_L}{d\log(K/T)}\end{matrix}
=
\begin{matrix}-\alpha\\-\alpha\end{matrix}
$$

Defining the 2 by 2 matrix as M. Given $^{a}f_L >^{a} f_H$ (which follows immediately from Proposition 1), we show after algebraic simplification that that the determinant of M is positive:

$$
\begin{aligned}
|M| = & \ \alpha+\alpha\left(\frac{^{a}f_H}{1-^{a} f_H}+\frac{^{a}f_L}{1-^{a} f_L}\right)+ \\
& \alpha\left(\frac{^{a}f_H}{1-^{a} f_H}\right)\left(\frac{^{a}f_L}{1-^{a} f_L}\right)+ \\
& \alpha(\gamma-\beta)\left(\frac{^{a}f_L}{1-^{a} f_L}-\frac{^{a}f_H}{1-^{a} f_H}\right)>0
\end{aligned}
$$

Applying Cramer's rule, we show that the resulting solutions are therefore negative:

$$
\begin{aligned}
\frac{d\log^{a} f_H}{d\log(K/T)} &= \frac{-\alpha\left[1+\left(\frac{^{a}f_L}{1-^{a}f_L}\right)\right]}{|M|}<0 \\
\frac{d\log^{a} f_L}{d\log(K/T)} &= \frac{-\alpha\left[1+\left(\frac{^{a}f_H}{1-^{a}f_H}\right)\right]}{|M|}<0
\end{aligned}
$$

and the difference between the first exceeds the second:

$$\frac{d\log^a f_L}{d\log(K/T)} - \frac{d\log^a f_H}{d\log(K/T)} = \frac{\alpha\left(\frac{^a f_L}{1-^a f_L} - \frac{^a f_H}{1-^a f_H}\right)}{\begin{bmatrix} \alpha + \alpha\left(\frac{^a f_L}{1-^a f_L} + \frac{^a f_H}{1-^a f_H}\right) \\ +\alpha\left(\frac{^a f_L}{1-^a f_L}\right)\left(\frac{^a f_H}{1-^a f_H}\right) + \\ \alpha((\gamma-\beta)\left(\frac{^a f_L}{1-^a f_L} - \frac{^a f_H}{1-^a f_H}\right) \end{bmatrix}} > 0$$

6.4 Proof of Proposition 4

The fact that $\frac{R_K}{R_T} > \frac{\tilde{R}_K}{\tilde{R}_T}$ follows directly from Proposition 1. We know that $\frac{\tilde{R}_K}{\tilde{R}_T} = 1$, and from the proof in Proposition 1, we show that $R_K > R_T$. We show now that the relative return to capital and labor is increasing in H/L:

$$\begin{aligned} \frac{R_K}{R_T} &= \frac{\alpha A K^{\alpha-1}(f_H H)^{\gamma}(f_L L)^{1-\alpha-\gamma}}{\alpha T^{\alpha-1}((1-f_H)H)^{\beta}((1-f_L)L)^{1-\alpha-\beta}} \\ &= A\left(\frac{K}{T}\right)^{\alpha-1}\left(\frac{f_H}{f_L}\right)^{\gamma}\left(\frac{1-f_L}{1-f_H}\right)^{\beta}\left(\frac{H}{L}\right)^{\gamma-\beta} \end{aligned}$$

$$\frac{1-\alpha-\gamma}{\gamma}\left(\frac{f_H}{f_L}\right) = \frac{1-\alpha-\beta}{\beta}\left(\frac{1-f_H}{1-f_L}\right)$$

$$\begin{aligned} \frac{R_K}{R_T} &= A\frac{\beta(1-\alpha-\gamma)}{\gamma(1-\alpha-\beta)}\left(\frac{K}{T}\right)^{\alpha-1}\left(\frac{f_H H}{f_L L}\right)^{\gamma-\beta} \\ \frac{d(R_K/R_T)}{d(K/L)} &= C\left(\frac{K}{T}\right)^{\alpha-1}\left(\frac{H_n}{L_n}\right)^{\gamma-\beta-1}\frac{d(H_n/L_n)}{d(K/L)} > 0 \end{aligned}$$

where

$$C = A(\gamma-\beta)\frac{\beta(1-\alpha-\gamma)}{\gamma(1-\alpha-\beta)}$$

Table 1
Summary of Suppy and Demand Factors for the Returns to Education, 1900 and 1910

1900	Supply Indicator	Demand Indicators				
	Literacy Rate	Percent of Labor Force in Manufacturing	Value of Livestock Per Capita	Value of Machinery per Capita	Value of Livestock per Ag Workforce	Value of Machinery Per Ag Workforce
State						
California	94.47	25.07	45.3	14.4	306.0	96.9
Georgia	66.94	13.72	15.9	4.4	53.2	14.8
Illinois	95.9	33.63	40.2	9.3	279.1	64.8
Iowa	97.59	13.21	124.9	26.0	482.5	100.3
Ohio	95.84	41.36	30.3	8.7	190.1	54.9
Texas	83.95	8.41	78.9	9.9	251.6	31.5
Wisconsin	95.12	33.02	46.6	14.1	237.8	72.2

1910	Literacy Rate	Percent of Labor Force in Manufacturing	Value of Livestock Per Capita	Value of Machinery per Capita	Value of Livestock per Ag Workforce	Value of Machinery Per Ag Workforce
State						
California	95.48	37.81	53.6	15.4	367.8	105.3
Georgia	76.74	22.51	31.0	8.0	99.2	25.7
Illinois	95.73	62.96	54.8	14.0	449.4	114.6
Iowa	97.98	28.87	176.6	42.9	751.9	182.7
Ohio	96.46	74.85	41.4	10.7	300.9	78.1
Texas	87.77	16.08	81.8	14.6	269.6	48.1
Wisconsin	96.34	64.89	67.9	22.7	357.3	119.4

Sources:
Value of Machinery and Value of Livestock: 1916 Statistical Abstract of the United States
Percent of the Labor force in Manufacturing, Population Size, and Ag. Workforce Size: 1924 Statistical Abstract of the United States
Literacy Rate Caluculated for those above the age of 10 based on IPUMS 1900 and 1910 5% samples.

Table 2: Summary Statistics by State

	California	Georgia	Ohio, Illinois, & Wisconsin	Texas
Mean Annual Salary (Overall)	1142	828	848	733
	(316)	(377)	(379)	(278)
Mean Annual Salary (Men)	1375	1001	918	823
	(344)	(331)	(403)	(290)
Mean Annual Salary (Women)	1020	474	757	575
	(219)	(145)	(323)	(159)
Mean Years Schooling (Overall)	13.8	12.6	12.6	12.6
	(1.4)	(1.7)	(1.9)	(1.9)
Mean Years Schooling (Men)	13.9	12.9	12.4	12.8
	(1.6)	(1.6)	(2.1)	(2.0)
Mean Years Schooling (Women)	13.7	11.9	12.9	12.2
	(1.4)	(1.6)	(1.6)	(1.6)
Mean Years Experience (Overall)	8.3	8.2	9.1	9.6
	(7.0)	(5.8)	(7.2)	(7.1)
Mean Years Experience (Men)	10.6	9.2	10.0	10.3
	(7.3)	(6.1)	(7.1)	(7.3)
Mean Years Experience (Women)	7.1	6.2	8.0	8.3
	(6.5)	(4.6)	(7.1)	(6.6)
Fraction Male	0.34	0.67	0.57	0.64
Number of Observations	658	137	3141	381

Source: Authors' Calculations from Thorndike Report
Standard errors are listed in parentheses.

Table 3: Mincerian Regressions Across States Using Individual Data

	California	Georgia	Ohio, Illinois, & Wisconsin	Texas
Schooling	0.005	0.033	0.070	0.071
	(0.006)	(0.016)	(0.003)	(0.009)
Experience	0.034	0.012	0.048	0.034
	(0.004)	(0.013)	(0.004)	(0.007)
Exper. Sq.	-0.0007	0.0004	-0.0008	-0.0009
	(0.0001)	(0.0006)	(0.0001)	(0.0003)
Male Dummy	0.22	0.64	0.16	0.27
	(0.02)	(0.06)	(0.01)	(0.03)
Intercept	6.67	5.63	5.35	5.26
	(0.08)	(0.20)	(0.04)	(0.10)
R^2	0.45	0.61	0.39	0.42
N	658	137	3141	381

Source: Authors' Calculations from Thorndike Report
Dependent variable is the log of the wage in all regressions.
Robust Standard errors are listed in parentheses.

Table 4: Comparing Mincerian Regression Estimates From Individual- and Median-Level Regressions

	Schooling Coefficient		Male Dummy		Intercept	
	Individual Data	Median Data	Individual Data	Median Data	Individual Data	Median Data
California	0.005	0.013	0.22	0.22	6.67	6.59
Georgia	0.033	0.026	0.64	0.64	5.63	5.70
Ohio, Illinois, & Wisconsin	0.070	0.058	0.17	0.17	5.35	5.52
Texas	0.071	0.071	0.27	0.28	5.26	5.24

Source: Author's Calculations from Thorndike Report.
Dependent variable is the log of the wage in all regressions.

Table 5: Median Mincerian Regression Results for Illinois, Ohio, and Wisconsin, 1909

	Illinois	Ohio	Wisconsin
Schooling	0.073	0.080	0.034
	(0.009)	(0.009)	(0.010)
Experience	0.026	0.030	0.042
	(0.009)	(0.009)	(0.010)
Exper. Sq.	0.0001	-0.0003	-0.0006
	(0.0003)	(0.0003)	(0.0003)
Male Dummy	0.192	0.079	0.274
	(0.048)	(0.047)	(0.051)
Intercept	5.819	5.766	6.029
	(0.078)	(0.104)	(0.085)
N	122	133	99
R^2	0.69	0.59	0.55

Source: Authors' calculations from Thorndike Report
Dependent variable is the log of the wage in all regressions.
Robust standard errors listed in parentheses.

Table 6: Regression of Overall Returns by State on Teachers Returns by State, 1980-2000

Variable	Coefficient
Year=1980	**0.053**
	(0.003)
Year=1990	**0.072**
	(0.005)
Year=2000	**0.080**
	(0.005)
Teachers' Return	**0.089**
	(0.046)
R^2	0.65
N	152

Source: Authors's calculations from IPUMS.
Dependent variable is the overall return to education for each state.
Robust standard errors listed in parentheses.

Table 7: Mincerian Regressions for Men, Women, and Less Experienced Teachers, 1909

	Individual Data				Median Data		
	California	Georgia	OH, WI, IL	Texas	Illinois	Ohio	Wisconsin
Men Only							
Schooling	0.004	0.020	**0.075**	0.085	**0.084**	**0.088**	**0.044**
	(0.010)	(0.021)	(0.004)	(0.009)	(0.013)	(0.011)	(0.015)
Experience	**0.041**	0.006	**0.048**	**0.035**	0.020	0.033	**0.040**
	(0.007)	(0.017)	(0.004)	(0.009)	(0.013)	(0.124)	(0.014)
Exper. Sq.	**-0.0010**	0.0006	**-0.0009**	**-0.0009**	0.0003	-0.0005	-0.0003
	(0.0002)	(0.0007)	(0.0001)	(0.0003)	(0.0004)	(0.0004)	(0.0005)
Intercept	**6.89**	**6.47**	**5.48**	**5.36**	**5.979**	**5.816352**	**6.221**
	(0.13)	(0.26)	(0.05)	(0.13)	(0.114)	(0.104)	(0.124)
Number of Obs.	226	92	1776	243	60	67	47
R^2	0.22	0.15	0.33	0.34	0.69	0.611	0.58
Women Only							
Schooling	0.007	**0.067**	**0.063**	**0.035**	**0.064**	**0.072**	0.023
	(0.007)	(0.026)	(0.005)	(0.013)	(0.012)	(0.013)	(0.014)
Experience	**0.028**	0.0144	**0.046**	**0.029**	**0.033**	**0.027**	**0.0468555**
	(0.004)	(0.030)	(0.004)	(0.011)	(0.014)	(0.013)	(0.014)
Exper. Sq.	**-0.0004**	0.0007	**-0.0005**	-0.0007	-0.0001	-0.00008	-0.0009
	(-0.0002)	(0.0019)	(0.0001)	(0.0005)	(0.0005)	(0.0004)	(0.0005)
Intercept	**6.64**	**5.19**	**5.42**	**5.73**	**5.834**	**5.798**	**6.111**
	(0.10)	(0.32)	(0.07)	(0.18)	(0.104)	(0.109)	(0.109)
Number of Obs.	432	45	1365	138	62	66	52
R^2	0.30	0.29	0.41	0.14	0.62	0.57	0.43
Less Experienced Teachers Only							
Schooling	0.020	0.012	0.052	0.071	0.047	0.059	0.032
	(0.010)	(0.030)	(0.04)	(0.010)	(0.011)	(0.010)	(0.010)
Male Dummy	0.231	0.769	0.201	0.280	0.224	0.152	0.254
	(0.027)	(0.091)	(0.014)	(0.037)	(0.063)	(0.053)	(0.052)
Intercept	6.538	5.884	5.690	5.355	5.799	5.602	6.013
	(0.134)	(0.362)	(0.053)	(0.130)	(0.137)	(0.119)	(0.123)
Number of Obs.	298	54	1278	155	51	54	50
R^2	0.20	0.59	0.20	0.43	0.42	0.48	0.43

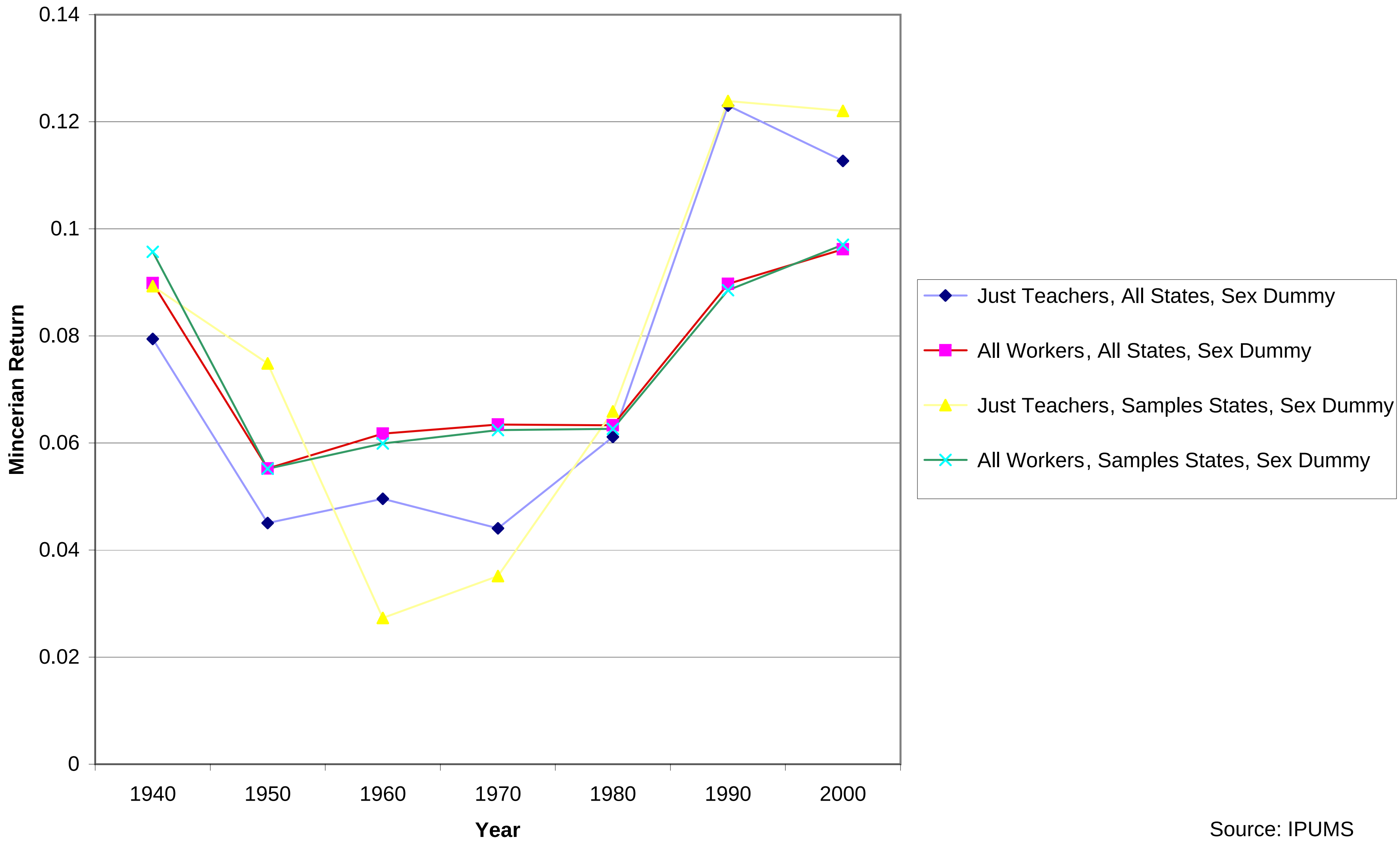
Source: Authors' calculations from Thorndike Report.
Dependent variable is log of the wage in each regression.
Robust standard errors listed in parentheses.

Table 8: Comparing Mincerian Regression Estimates 1909 and 1940

Variable	Weighted Teachers, Sample States, Sex Dummy (1909)	Just Teachers, Sample States, 1940 Normal Regression	Just Teachers, Sample States, 1940 Robust Regression	All Workers, Sample States, Sex Dummy (1940)	All Workers, All States, Sex Dummy (1940)
Schooling	0.083	0.091	0.089	0.096	0.090
	(0.003)	(0.010)	(0.008)	(0.001)	(0.000)
Sex	0.140	-0.162	-0.125	-0.452	-0.459
	(0.010)	(.056)	(0.044)	(0.007)	(0.004)
Exp	0.034	0.035	0.039	0.047	0.516
	(0.002)	(0.007)	(0.005)	(0.001)	(0.005)
Exp2	-0.001	-0.001	-0.001	-0.001	-0.058
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)

Source: Author's calculation from Thorndike Report (for 1909 estimates) and IPUMS (all others)
Dependent variable is the log of thw wage in each regression.

Figure 1: Comparison of Mincerian Returns over Time



Source: IPUMS