Essays on the Gender Gap and the Effects of Secondary School Expansion: Evidence from the Early Twentieth Century's High School Movement

By

Michael Quinn Moody

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in

Economics
August, 2016
Nashville, Tennessee

Approved:
William J. Collins, Ph.D.
Kathryn H. Anderson, Ph.D.
Robert Dale Ballou, Ph.D.
Claudia R. Rei, Ph.D.

ACKNOWLEDGEMENTS

Financial support from the Noel Dissertation Fellowship and the National Science Foundation is gratefully acknowledged.

The employees of the Tennessee State Library and Archives provided excellent assistance throughout the process.

I thank my committee members, Kathryn Anderson, Dale Ballou, and Claudia Rei, for their advice and guidance throughout the process.

I am most indebted to my adviser, William Collins. His patience and guidance throughout the process has been outstanding. This project would not have been possible without his advice and I am eternally grateful for his assistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Rapid economic and social change characterized the United States in the late 19th and early 20th centuries. One salient aspect of this change concerned investment in formal educational attainment, which had significant implications for human capital accumulation in the labor force (Goldin and Katz 2007). Publicly funded high schools diffused throughout the U.S., and by the middle of the twentieth century four years of high school became a normal part of life. At the same time, many women entered the formal labor market, and relatively well-educated women often found employment as schoolteachers. In addition, the South began to industrialize and urbanize, and improvements in transportation made geographic mobility less costly than ever before.

I construct and study a new dataset of school records from the Tennessee State Library and Archives to investigate the economic history of high schools and, specifically, to understand important aspects of labor markets in this period. There are two main questions I seek to answer:

1. How large was the gender wage gap among high school teachers in Tennessee during this period, and what factors contributed to this gap?
2. What impact, if any, did changes in high school access have on young men's choices between agricultural and non-agricultural activities?

For the first question, I use the highly detailed teacher-level information from the archives to calculate the average wage gap in each county. After documenting that large wage gaps remain in the data even after controlling for each teacher's education, classes taught, experience, and administrative responsibilities, I try to find the cause of the remaining gender gap. I find two potential explanations to be compelling: differential employment rates for higher

paying positions such as agricultural science teacher or school principal and monopsony hiring power of the school boards.

For the second question, I take advantage of a 1917 law that greatly expanded access to high school in Tennessee. Using a linked set of census records for children who resided in rural areas in 1910, I perform a difference-in-differences analysis of their occupational outcomes in 1930 as reported in the census. Departing from most literature on the subject, I find that rural children were more likely to remain in farming after gaining access to high schools. This may well be because rural Tennessee high schools focused on teaching modern agriculture techniques, which may have improved a student's returns to farming.

High schools in Tennessee can be particularly enlightening to study, as Tennessee was arguably a microcosm of the South. The geography varies widely over the state, with the Mississippi River bounding the western part of the state and the Appalachian Mountains dominating the eastern part. The southwestern part of the state is bordered by Mississippi and Arkansas, whereas Virginia and North Carolina border the northeastern part. Agriculture was the dominant economic activity in Tennessee, like the South as a whole, though specialization varied depending on soil and geography. Farmers in West Tennessee grew a significant amount of cotton, similar to farmers in the Deep South. However, farms in the central and eastern part of the state tended to focus on crops such as tobacco and wheat (Corlew 1989, p. 368), which is more like the agriculture that can be found in North Carolina and Virginia. The state also included major cities, such as Memphis and Nashville, which were involved in trade and industry important to the region, including textiles, lumber, transportation, and coal mining.

I propose that schools were one of the major promoters of social change in this era and studying how Americans invested in their children's education can give insight into and help

reconcile two separate branches of the economic history literature. The first major branch studies the high school movement, which itself is part of the long-run process of human capital accumulation, economic growth, and inequality. Documented most thoroughly by Goldin and Katz (2009), the high school movement represented "an extraordinary increase in the education of the nation's youth" (p. 195). In 1910, the economic return to an investment in a high school education was extremely high, but only 9 percent of all American 18-year-olds graduated from secondary school. By 1940, the graduation rate for 18-year-olds was around 45 percent (Goldin and Katz 2009, Figure 6.1). As this was occurring, the public high school replaced private academies as the main instrument of high school instruction. With this change came a shift in the focus of high schools' curricula. Whereas the 19th century high schools mostly trained children for college, the high school diploma of the early 20th century was seen predominantly as a terminal degree.

The high school movement was slow to arrive in the South. In 1910, about 15 percent of children in the East North Central region of the country graduated from high school. The rate for southerners was about 5 percent (Goldin and Katz 2009, Figure 6.2). By 1940, the graduation rate was 62 percent in East North Central compared to 40 percent in the South Atlantic and 30 percent in the East South Central (Goldin and Katz 2009, Figure 6.2). Some of the regional disparity is due to low rates of schooling among southern African Americans. However, the issue affected both races. For example, in the South Atlantic region in 1940, whites had a graduation rate of about 48 percent. So, although high schools (and high school graduation) expanded rapidly in the South through the early twentieth century, the region still trailed the U.S. average, and this had significant implications for the distribution of human capital within the

southern-born workforce. The first task of the dissertation is to document the spread of high schools throughout Tennessee in unprecedented detail in terms of location, timing, and character.

The second branch of the economic history literature I engage concerns the increasing numbers of women in the teaching force as part of the larger history of women's labor force participation. Perlmann and Margo (2001) track the female dominance of the elementary school teaching force and find that it started in New England and spread to the Midwest and West. The South, again, lagged behind the rest of the country in educational change. Perlmann and Margo give many reasons for the transition but among the most important are: (1) the Civil War, which forced school boards to hire women in place of the absent men and may have changed opinions about women's ability to do the job; (2) graded schools, which allowed women to handle younger children and alleviated contemporaneous concerns that they may be unable to discipline older boys; (3) a relatively cheaper woman's wage which allowed 19th century reformers to argue that the hiring of women allowed for more schools to be opened at the same cost; (4) general changes in attitudes towards women's ability to perform as teachers. The economic history of women's employment in high schools is less well documented and understood. My data suggests that the transition from male to female labor among high school teachers was not as extreme as in elementary schools. Women entered high school teaching in large numbers but worked alongside men instead of completely displacing them.

The phenomena discussed in this dissertation occur against the backdrop of increased integration of the South with the rest of the country. Wright (1986) paints a multifaceted story whereby northern industrialists moved south to take advantage of cheap labor and that labor then benefitted from increases in the minimum wage during the New Deal. Coleman and Caselli (2001) suggest that much of the South's wage equalization with the North was a result of

increased education and human capital. Connolly (2004) also argues for the important role of education in equalizing wages, claiming that the transition from informal methods of human capital transmission in agriculture to more formal methods of human capital transmission in schools was a driver of economic convergence between the North and South.

This dissertation expands our knowledge of the southern high school movement by using newly discovered data to explore the high school's role as both an employer for skilled women in an era when women entered the workforce in large numbers for the first time as well as the impact on job choice of children who attended high school. The findings confirm that significant wage gaps existed between men and women at this time even when they performed very similar work. Furthermore, I find evidence that the Great Depression served to decrease women's movement into more important administrative roles. In this case, women became less likely to work as principals over time. At the same time, rural southern high schools emphasized agriculture. Perhaps surprisingly, my work suggests that access to high school increased the likelihood that a child would become a farmer later in life. Thus, while southern high schools were based on their northern counterparts and intended to make the South more like the North, headwinds against change continued to show up in the historical experience.

CHAPTER 2. HISTORICAL BACKGROUND ON HIGH SCHOOLS IN TENNESSEE

This chapter reviews the history of high school education in the State of Tennessee, with emphasis on the first part of the twentieth century. I focus on legislative and social changes in the late nineteenth and early twentieth century which directly related to two aspects of education in Tennessee: the demand for and supply of teachers, particularly female teachers, and access to high schools by the rural population.

Goldin and Katz (2009) argue that the American high school movement in the early twentieth century was a critical component in the history of human capital acquisition. It opened a wide gap in average educational attainment between U.S. workers and those elsewhere, with long-lasting implications for international differences in productivity. It also tended to narrow measures of skill-based wage inequality within the U.S. This was a time when the rising supply of skills outstripped the demand for skills in the "race between education and technology" (Tinbergen 1974, Goldin and Katz 2009). For most students in this period, the high school diploma was a terminal degree rather than preparation for university-level work.

Tennessee, and the South more broadly, lagged behind the North in the high school movement (Goldin 1998, Goldin and Katz 1998 and 2009, Ramcharan 2002). In the late 19th and early 20th centuries, a collection of northern philanthropists began endowments such as the General Education Board, the Southern Education Board, the Peabody Fund, and the Rosenwald Fund (Woodward 1951, Aaronson and Mazumder 2011). Due in part to these endowments, southern educational achievement improved significantly over the course of the early 20th century.

Ideally, I would like to study the impact of high schools throughout the South. However, the data and the legal history I utilized for this project are specific to Tennessee. Therefore, it is

important that Tennessee is an appropriate representative for the South as a whole. I turn to this in the next section before discussing the legislative history of high schools in Tennessee.

2.1. Tennessee as a Southern Microcosm

My dissertation, like studies by Margo (1982), Donohue, Heckman, and Todd (2002), Reber (2010), and Baker (2015), relies heavily on detailed information retrieved from archival sources for a particular state. I review the sources and methods of data collection in the next chapter. Here, I simply point out that Tennessee is reasonably representative of the South in terms of demographic and economic composition and that much can be learned about the ramifications of the southern high school movement from Tennessee's historical experience.

During the period under study, agriculture was the dominant economic activity in Tennessee. The IPUMS 1% sample (Ruggles et al. 2010) shows that 60 percent of the male labor force in 1920 was employed in agriculture. The same held for the South as a whole (59 percent of the male labor force). Statistics from Haines, Fishback, and Rhode (2014) suggest that the types of agriculture practiced were similar in Tennessee and the other southern states. Major crops grown in Tennessee include: corn (37.4 percent of improved farmland), hay (19.8 percent), cotton (7.2 percent), wheat (6.1 percent), oats (1.5 percent), and tobacco (1.2 percent). In the rest of the South (the Confederate States and West Virginia), the major crops were: cotton (22.5 percent), corn (22.3 percent), hay (10.5 percent), wheat (7.6 percent), oats (3.4 percent), barley (1.8 percent), and tobacco (1.0 percent). While cotton was more important in the South as a whole than in Tennessee proper, the mix of crops grown across the South in general and Tennessee in particular was very similar.

Tennessee's major cities, such as Memphis and Nashville, boasted industries similar to other southern states including textiles, lumber, transportation, and coal mining.[1] Approximately 16 percent of Tennessee's population in 1900 resided in urban areas, compared to 18 percent of the South's population (17 percent excluding Washington DC). These cities housed some of the nascent industry emerging in the South at the time. This was particularly true of East Tennessee, with more urban population and small industrial plants. For example, Wright (1986, p. 165) mentions the importance of iron production in the rapid growth of Chattanooga.

In the period under study, people in Tennessee lived and learned similarly to other southerners but differently from the rest of the country. For instance, Figure 1 shows the share of adults over age 25 living on farms from the late-1800s to the mid-1900s, covering the study's timeframe. The data are drawn from the IPUMS samples of manuscript data from the decennial Census of Population (Ruggles et al. 2010). The trends for Tennessee and the South as a whole track one another closely, and they are clearly higher than the levels for other regions. There is a sharp decline in farm residence over the period in the graph, but the difference between Tennessee and the whole South is always between 2.1 and 5.5 percentage points after 1900.

Figure 2 and Figure 3 illustrate school attendance rates of whites and blacks, respectively, between ages 6 and 18. School attendance in the census is defined as any attendance during a specified time period ending with the census enumeration.[2] Both charts show a jump in school attendance rates between 1900 and 1910, which is particularly strong for whites. Margo (1990)

[1] Data calculated from Ruggles et al. (2010).

[2] This definition of this variable varies census-to-census. The first issue is that the census day (the day that the count was actually made) changed from year to year. From 1850 to 1900, census day was June 1. In 1910, census day was April 15. In 1920 census day was January 1. From 1930 to 2000, it was April 1. Next, the question changed both the definition of school and the window over which a person needed to attend. From 1850 to 1900, the question asked if someone had attended school in the past year. For 1910 to 1930, the window was September 1 to census day (seven and a half months, four months, and seven months, respectively).

finds the same jump in the data. It is driven by an especially strong increase in attendance for those between 6 and 10. It is difficult to explain this sizeable jump in school attendance. The census questions changed over this period but in a way that should have led to lower schooling rates in 1910 compared to 1900. In 1900 the question asked if an individual had attended any school over the last twelve months. In 1910, the question changed to ask if an individual had attended "any school, college, or educational institution" over the last seven and a half months. The Tennessee Superintendent of Public Education's *Annual Report*, which will be discussed in more detail later, shows average attendance rates for white children of school age of 47.5 percent in 1900, 50.2 in 1910, and a jump to 57.8 percent in 1920. Nevertheless, assuming that the source of the abnormality is consistent across states, the data show Tennessee and the South moving together very closely over the 70 years charted. The South and Tennessee start at a much lower rate than in the rest of the country but move upward as the regional gap narrows dramatically over time.

Prior to 1940, the only census question that gauged educational attainment was a question about the ability to read and write. Figure 4 compares literacy rates for adults across regions, and the upshot is similar to the discussion above. Tennessee and the South move closely together and narrow the regional gap over time. The kink in the data between 1930 and 1940 is due to a change in the definition of literacy. Prior to 1940, the census specifically asked if a person was literate. After 1940, literacy has to be imputed. I use education to do so. Collins and Margo (2006) explore the relationship between schooling and literacy using data from manufacturing employees in the early 20th century and Census Bureau data collected just after World War II. Collins and Margo find that somewhere between two and three years of schooling is sufficient to ensure literacy. I consider anyone with three or more years of education to be

literate in the 1940 and 1950 data. To the extent that people could learn to be literate without three years of formal schooling, this method will tend to underestimate literacy rates. The significant deficit in southern literacy speaks to the extent formal schooling in the South ran behind the rest of the country.

Figure 5 displays the average number of children per head of household between the ages of 30 and 50. A decrease in children per head-of-house is consistent with the demographic transition associated with industrialization. Wannamaker (2012) shows such a transition occurred in South Carolina in the early 20th century. Figure 5 suggests this transition occurred much later in Tennessee and the South than elsewhere in the U.S. However, Tennessee closely followed southern trends in this aspect, as well. Figure 6 looks at the same information for whites only to ensure that gaps between different areas of the country are not entirely due to differences in racial composition.[3] The rates are roughly the same as those in Figure 5 suggesting that the demographic transition in the South was, in fact, a shift in behavior at the population level.

2.2. The Role of Private Academies in Tennessee Education

This section discusses the development of the schooling system in early Tennessee and the role that private academies played as a bridge from elementary school to college. The goal is

[3] The graphs suggest that African American heads of house had fewer children than white heads of house. This is because of the definition of head of household. In the IPUMS 1% samples (Ruggles et al. 2010) the graphs are derived from, African American households have more multi-generational households and to have children slightly earlier in life. Thus, the figures capture households where a parent supports one of their children and their grandchildren in the house. The head of house would show as having 1 child whereas the real figure of interest is the number of children the head of household's child has. I tested this calculating the average number of children between 0 and 18 with a common parent over the same period and African-Americans, as expected, have more children per parent.

to establish that the academies' importance waned in the wake of competition from public institutions and, therefore, their influence on my study is negligible.

For much of Tennessee's history, the state maintained two separate schooling systems: the white system and the "colored" system. This section and most of what follows focuses on the white high schools primarily because they have more complete and reliable data. The superintendent did not consistently give information about the black schools in the *Annual Report*, and most of the original data have been lost in the intervening years.

In 1794, two years before Tennessee became a state, the territorial legislature created Blount College.[4] Over the next 200 years, the institution would change names many times and is known today as the University of Tennessee. Despite this early commitment to higher education, it would be almost 100 years before the state established a program to build and fund public high schools. During this time, the job of training young men and women for college was left to "academies." Academies were quasi-public institutions that were privately managed but chartered by the legislature and were, by law, "considered as schools preparatory to the introduction of students into the colleges of this state" (State law quoted in White 1929).

Information on the academies is scarce as the private records have disappeared or are unavailable, and Tennessee records were unreliable prior to the 20th century.[5] Nonetheless, historians describe an extensive network of schools. White claims that over 500 academies were chartered between 1794 and 1889 (1929, p. 10). Holt claims there were 1,450 private schools and academies in the state in 1873 (1938, p. 26). There are no attendance records during the earliest days of the state. In 1875, two years after significant investment in public education

[4] Much of this discussion relies on White (1929) and Holt (1938) who specifically study the history of Tennessee schools.
[5] The problem of getting accurate information was so severe that a later law would require accurate reports to be filed in order to collect state aid (White (1929)).

began, authorities estimated there to be 43,138 students in private academies. This is roughly one fifth of the total number of students ages 6 to 21 educated in Tennessee. As public education – and specifically public high schools -- became increasingly important in Tennessee, the importance of the academies was reduced. The 1912 *Annual Report* lists 106 private schools, of which 44 were colleges. The 62 that were not colleges employed 413 teachers and taught 11,294 students in grades 1 to 12.[6] In the same year, public schools in the state employed 11,086 teachers and taught 539,911 students, suggesting that private schools were relatively unimportant suppliers of elementary and high school education in the period under study.

2.3. Tennessee High School Legislation

This section discusses the legislative acts on which the Tennessee Public High School system of the early 20th century was built. I pay particular attention to two specific points: first, that the system developed in such a way that male and female teachers had both a common floor to their skill sets and performed teaching jobs with similar descriptions; second, the laws dictating high school funding created differential distribution of high school access across the state which I can take advantage of for difference-in-differences analysis.

Governor William G. Brownlow signed the first comprehensive school law in the state's history during Reconstruction. First elected in January 1865, Brownlow's election was largely due to the large numbers of men fighting for the Confederacy who were unable to vote because by then the Union controlled the state and prohibited men in rebellion against the Union from voting. To protect his office, Brownlow enacted two pieces of legislation. The first denied the

[6] University School in Memphis is listed as having 140 teachers. This is likely a typo as the school only has 135 students. I count the school as having 14 teachers.

vote to all ex-Confederates and the second granted the vote to freed male slaves.[7] These laws not only guaranteed his election in 1867 – with 77 percent of the vote – but a state legislature favorable to the Radical Republican coalition.

In 1867, the legislature passed a bill that re-instituted the office of the State Superintendent of Public Schools (which had been abolished in 1844), county superintendents for each county, and a tax system to pay for the schools. The tax rate was established at $0.02 per $100 of taxable property and a tax of $0.25 on each male citizen between the ages of 21 and 50 (White 1929).

In 1869, Brownlow was appointed to the U.S. Senate, and the Radical Republicans split into two factions. Brownlow's replacement, DeWitt C. Senter, won re-election both by campaigning on a platform of restoring the vote to ex-Confederates and by replacing county registrars with men loyal to him who would register former Confederates for the election. Senter's victory ended Reconstruction in Tennessee, and the new state legislature methodically repealed the entirety of Brownlow's legislation, including the schooling law (Corlew 1989).

Almost immediately following repeal, the nascent State Teachers Association campaigned for a new law. Their work resulted in passage of such a bill in 1873. This law, according to White, "was the parent act that *established and the one that has maintained* a permanent system of public schools in Tennessee" [Emphasis his] (1929, p. 122). The bill put in place much of the bureaucracy that oversaw Tennessee high schools during the period of study, including the State Superintendent of Public Education and County Superintendents of

[7] The strife between the two factions was intense. Brownlow, for his part, believed that Andrew Johnson would start another revolution after which Johnson would be executed and the South made "as God found the earth when he commenced the work of creation, 'without form and void.'" Confederates responded to Brownlow by forming the Ku Klux Klan under ex-Confederate General Nathan Bedford Forrest (Corlew 1989, p. 332).

Public Education. The latter were selected by the county school boards ensuring some measure of local control over the schools. The 1873 bill also established a fund for the schools and taxation to maintain the fund (Acts of Tennessee, 1873, Chapter 25).

Following this bill, the legislature passed a series of bills that reformed the Tennessee school system. The period from 1873 to 1903 had fewer bills because extreme financial hardship following the Civil War caused many further reforms to be put off for the next 30 years. 1903 to 1913, however, "the Campaign Era," saw significant legislative action. I will give a brief review only of the bills directly affecting the high schools in early 20th century Tennessee. A full account of the Tennessee school laws of this period can be found in White (1929).

The public high school in Tennessee started with an 1885 law, which permitted municipal corporations to raise money for graded high schools, known colloquially as "city schools." The state, however, provided no additional funding for that purpose. The lack of state funding was, in part, a reflection of the state's previously referenced financial difficulties following the Civil War. The bill also specifically forbade mixed race high schools, a rule that would apply until the *Brown v. Board of Education* decision by the Supreme Court in 1954 (*Acts of Tennessee* Extra Session 1885, Chapter 19).

In 1899 the legislature expanded on the earlier bill and permitted public high schools to be established in the unincorporated areas of Tennessee. Known as the County High School Law, it established a maximum tax of 15 cents per $100 of taxable property to establish and maintain high schools. More importantly, however, the bill gave control of the county high schools to the county board of education. This was different from the elementary schools, where control was vested in local school boards, and allows me to study hiring and funding decisions at the county level (*Acts of Tennessee* 1899, Chapter 279).

Although this law allowed "county high schools" to be built, progress was slow. The first school built under the law was in Lake County five years later. Despite this, the law's passage had significant impact on the development of the schools as the decision to make the county the administrative seat of the high schools meant that hiring and wage decisions were made by the county superintendent and, later, the county school boards (Hoffschwelle 2014).

While the law was established to enable schools to be built in unincorporated areas, the county high schools – a name that distinguished them from the city high schools built under the earlier law – were also built in incorporated areas and served students from these areas. The distinguishing feature was who controlled the school: the county board of education or the city board of education. The two types of schools are, unfortunately, indistinguishable in the school-level records. However, according to the administrative records, there were fewer than 20 city high schools by 1923 compared to 589 county high schools.

In 1909 the legislature passed the General Education Bill, which increased funding and established state oversight of the state public school system (*Acts of Tennessee* 1909, Chapter 264). This bill was the first that systematically funded county high schools out of state taxes and a period of significant high school expansion followed. [8] The State Superintendent's Annual Reports claim that there were 111 county high schools in 1910, growing to 499 by 1921 and 621 in 1930.

In addition, the bill established the position of the State Superintendent of High School Instruction. The legislation gave the superintendent two jobs relevant to my work. First, the superintendent was charged with certifying high school teachers. He met this goal by instituting

[8] The county high schools are separate from "city high schools," built under the auspices of an 1885 law. The city schools were separate entities and not bound by the same curriculum rules as the county schools. Specifically, city schools did not need to teach agricultural science and home economics. Additional discussion can be found in the appendix.

a test that potential teachers had to pass to be certified.[9] An early example of this test can be found in the Biennial Report for 1909/1910 (p. 293). The questions are rigorous and require knowledge similar to that taught in high schools today. Second, the superintendent was responsible for determining the exact course of study in the county high schools. The exception to this was that, by law, all county high schools had to offer classes in agricultural science to boys and classes in home economics to girls. To fill in the rest of the course of study, the superintendent called a meeting of "high school principals and college presidents … to consider a course of study for high schools" (Biennial Report 1909/1910, p. 765). Following the inspector's meeting, specific courses of study for high schools were printed in the Biennial Report, including a list of subjects to be taught, subjects required for graduation, and how requirements would be altered for the two- and three-year high schools. Students were required to complete the course of study to receive a diploma.[10] At the same time, the inspector instituted a rigorous test to qualify high school teachers, which ensured a baseline level of ability for all teachers. To generate a supply of teachers for the new schools, the state built four normal schools to train teachers.[11] The schools were spread across the state to maximize access to prospective teachers. In total, the county school boards hired teachers to do jobs with similar requirements and qualifications who were often trained at the same institutions.

The ultimate effect of the 1909 law and meeting was to homogenize the county high school curriculum and the teaching profession.[12] School boards across the state hired teachers to

[9] Unfortunately, the annual reports have no information on what constitutes a passing grade.
[10] Students had some ability to choose electives. This allowed a high school to serve diverse functions: produce new students for college, farmers comfortable with modern scientific techniques, and business workers with knowledge of bookkeeping and typing.
[11] There were three schools for white teachers and a fourth "Agricultural and Industrial" school for black teachers.
[12] The city high schools, built under the auspices of the 1885 law, were separate entities and not bound by these rules.

do a job with essentially identical requirements and qualifications. A male English teacher hired by Shelby County (the home of Memphis) should have had the same job description as a female English teacher hired in a rural district.

In 1917, the legislature required every county in the state to adopt a property tax of at least $0.05 per $100 of taxable property to fund expenditures on high schools (*Acts of Tennessee* 1917, Chapter 96). The law appears to have had an impact on the educational opportunities of children in rural areas. According to S. W. Sherrill, the Superintendent of Public Education:

> Up to 1917 about seventy counties had levied such a tax [i.e., that provided for by the 1899 bill] and established one centralized four-year high school. These schools failed to reach the boys and girls in the rural districts…With the funds derived from [the 1917] tax the high school boards have established one- and two-year high schools within reasonable reach of all boys and girls in the county (*Biannual Report* 1917/1918, p. 14).[13]

This is a slight overstatement. Some counties with no urban population according to the 1920 census did have a four-year high school in 1916 (see Table 1). However, as I show below, the law clearly expedited the construction of high schools in the state.

The first compulsory schooling law in Tennessee was passed in 1913, requiring children of ages 8 to 14 to attend school (White 1929). This was amended in 1919, further requiring school attendance by children 7 to 16. This law remained unchanged until 1992 when it was increased again to the present 6 to 18 (Morgan 2004).

The impact of the compulsory schooling laws in Tennessee is difficult to see in the contemporaneous data. Since it was specific to age rather than grade completion, it may have had a minimal impact on high school attendance per se, let alone graduation, in the period under

[13] Interestingly, this is the only reference to this law in the annual reports. White does not mention it at all. It was a smaller section of the law defining primary school as grades 1-8 and high schools as grades 9-12.

study. The IPUMS microdata sample suggests that 79.6 percent of white children 14 to 16 were attending school in 1910 as opposed to 51.1 percent of 17 and 18-year olds. By 1920, these numbers were 78.4 percent and 38.4 percent, respectively. From this perspective, the law does not seem to have increased attendance of 14-16 year olds. However, given that attendance decreased for children the law did not affect (17-18 year olds), it is possible that the law may have prevented a decline in census enumerated attendance rates for the 14 to 16 year olds (Morgan 2004).

The 1940 census shows some evidence that the 1919 law impacted schooling decisions. Figure 7 shows total years of schooling for white birth cohorts who turned 14 before and after both laws. The 1913 law seems to have little effect on years of schooling completed. Children who turned 14 in 1919 saw a spike in years of schooling completed of about 0.4 years. This then fell in for cohorts who turned 14 in 1920 and 1921 before jumping again 1922.

2.4. First, Second, and Third Class High Schools

Three different types of high schools emerged in Tennessee in the early twentieth century. The first was the "first-class" (or "four-year") high school, often called the Central High School. This class of school had at least three teachers – usually more – with specialized training in the subjects they were hired to teach. The classrooms were designed around the subject or subjects taught in them. Libraries, auditoriums, science labs, and athletic facilities were likely to be found in a four-year high school. The "third-class" (or "two-year") high school is similar in concept to modern community colleges but of course at a lower level. They allowed students to get some of the benefits of the four-year high school. After completing the two-year course, students had the option to transfer into a four-year high school or enter the labor force. Two-year high schools required at least one teacher, who was often the principal of an attached

grade school (Summitt 1935). The "second-class" high school was rare in practice and consisted of two teachers and three years of instruction.

Contemporaneous observers believed the first-class high school to be clearly superior to the other two classes. State Superintendent J.W. Brister said,

> Some counties have attempted too many (high schools) and the result is a number of small, indifferent schools doing a poor grade of high school work.… [O]ne first class high school … is worth a dozen inferior schools which only pretend to give high school instruction (Biannual Report 1911/1912, p. 52).

The State Superintendent's preference for first-class high schools ultimately led to a law in 1921 that required all counties to build at least one first-class high school. However, both first-and third-class high schools experienced strong growth into the mid-1930s, when third-class high schools began to die out.

In Tennessee, tension between farmers and the central school authorities was strong in some communities.[14] Perhaps in reaction to this, the Superintendents made it clear that rural high schools were designed to make children better farmers. J.W. Brister, the Tennessee State Superintendent of Public Education, said:

> I hold strongly to the opinion that a county high school has no right to exist for any length of time unless it does something in the way of teaching agriculture and domestic science. Every county high school should have a demonstration farm, with adequate equipment to give scientific instruction, should have laboratories and apparatus for genuine work in home economics. But it will be a long time, perhaps, before adequate equipment will be furnished

[14] Hoffschwelle (1998) discusses the struggles to establish new schools. Speaking of the related consolidation movement – which involved consolidating small, one room school houses into more modern, graded school buildings and also involved primacy of state education authorities over local communities – she quotes Lewis County Superintendent John White: "When all other counties get their schools consolidated and consolidation has shown to be a modern necessity, when all the arguments against it have been 'shown up' as falsehoods, then… [w]e will consolidate with a boom – but until that time comes, let us alone. You may have trouble insuring your houses if you don't (p. 37)."

> and high school people cannot afford to wait (Biannual Report 1911/1912, p. 52-53).

The focus on farm living underscores the extent to which county high schools were intended to differ from city schools.[15] County high schools were supposed to teach children who would (in theory) stay in the community and continue farming. In concept, the idea is similar to the effect identified by Parman (2012), who discovered that high schools played an important part in the improvement of agricultural productivity in Iowa at the turn of the century. Using the 1915 Iowa state census, Parman argues that not only did schooling improve farmers' private returns by 5.2 percent per year of high school but also had spillover effects on their neighbors, increasing the neighbors returns by 2.8 to 3.8 percent per year of high school attended.

Table 2 presents economic and school data for white high schools, workers, and residents in the early 1920s.[16] I divide the state into county groups based on rural population: 100 percent rural residents ("rural counties"); 80 to 100 percent rural residents ("mostly rural counties"); fewer than 80 percent rural residents but no major urban center, defined as a city with a population of 25,000 or more, ("less rural counties"); and counties containing a major urban center ("counties with urban centers").[17] Not surprisingly, the percent of farmers decreased and the share of white-collar workers increased as the urban share of population rises across columns.[18]

[15] County schools in large, urban counties like Shelby or Davidson were expected to have agricultural science classes and they did employ teachers for such classes. The demand from the students for farming classes in major urban counties is unclear.
[16] 1923 data is used for the schools. Demographic data comes from the 1920 Census (Ruggles et al. 2010).
[17] In 1920 Tennessee, these were Shelby (Memphis), Davidson (Nashville), Knox (Knoxville), and Hamilton (Chattanooga).
[18] White collar workers are defined as those involved in professional, clerical, or managerial work as defined in Ruggles et al. (2010).

The connection between schools and farming can be seen in the pattern of the different classes of schools in Tennessee. While first-class high schools were found in all four groups of counties, they tended to be larger and have more students the less rural a county became. The four highly urban counties had a relatively large number of first-class schools per county (7.75), and these schools were large compared to those found elsewhere (more than 700 students per high school on average). As one might expect, increased population density led to increased classrooms and students per square mile across the four groups of counties. Interestingly, whereas classrooms and students per square mile increase by factors of about 3 and 5, respectively, across the county groups, schools-per-square-mile increase by a factor of less than 2. Thus, in this period, before school busses and automobiles were ubiquitous means of transport for high school students, the more rural areas tended to build smaller but more numerous (per capita) high schools to serve less dense populations (e.g., "rural counties" had 89 high schools for 34 percent of the state's population, whereas "urban centers" had 31 high schools for 28 percent of the population).

The rural and mostly rural counties used the small, cheap third-class high schools more extensively than the mostly urban counties and urban centers. Middle Tennessee, composed of many of these smaller, sparsely populated counties, built more third-class high schools than the other two divisions combined. East Tennessee, on the other hand, already had more first class high schools than third class high schools by 1923. This, again, may be a result of more cities in the east. I will return to the geographic distribution schools in subsequent chapters.

2.5. Summary

Over half a century, Tennessee's Legislature built a high school system with statutory requirements on taxation, school quality, and oversight. The legislature gave oversight to the

State Superintendent of Public Education. The Superintendents of the time used their power to standardize curriculum, graduation requirements, and the certification of teachers.

Because of this legislation, teaching jobs were similar across the state and across genders, an historical feature that takes on more significance later in the dissertation when comparing wages across space (Chapter 4). In addition, the so-called "equalization funds" ensured that school funding was based more on total students in a district than on local taxable property, and the establishment of three separate normal schools across the state led a more even distribution of young teachers.

The laws also spurred growth in the number of high schools and a corresponding increase in attendance. In particular, the 1917 law which required taxes for high school construction preceded a jump in the total number of high schools (shown in detail in the next chapter), suggesting a causal relationship. Because the law was plausibly exogenous from the standpoint of localities, it may provide a reasonable basis to study the impact of changes in high school access on the outcomes of Tennessee students (Chapter 5).

CHAPTER 3. ARCHIVAL DATA COLLECTION AND DESCRIPTIVE STATISTICS

3.1. Original Sources

In the United States, education has historically been planned and funded at the local and state levels. Therefore, state boards of education, rather than the federal government, often collected and archived the richest sources of historical information on schooling. I collected and digitized two new data sources originally archived by the Tennessee State Superintendent of Public Education for this project. The first dataset consists of individual high school reports, which principals submitted to the State Superintendent of Public Education annually. The second consists of administrative records compiled in the *Annual Report of the State Superintendent of Public Education*.[19] Both the annual reports and the school-level reports are available in the Tennessee State Library and Archives. The annual reports are in their original published format whereas the individual reports are on microfilm.

The individual reports are five-page forms with detailed information on the school. A copy of the first three pages of one such form can be seen in Figure 8. Each form contains information on the monetary cost of building the school, the classes taught, class schedules, students-by-grade, and detailed information on the teaching corps. The information on the teachers includes their salary, education, type of certificate held, and years teaching. The latter is broken up into total years teaching, years teaching secondary school, and years teaching at the current school. I created digital copies of the original reports and had a third party enter the data into spreadsheet files.

The individual reports prior to 1920 have been lost, as have the reports for the 1920-21 and 1921-22 school years. For the 1922-23 school year to 1945, the records cover about 80% of

[19]Biannual Report from 1907-1908 to 1919-1920.

the public schools open at the time and those schools contained over 90% of all students.[20] For example, the annual report says that there were 589 white high schools in 1923. There are extant individual records for 499 of these schools. School counts for the two sets of reports are found in Table 3. My estimated distribution school class does not equal the *Annual Reports'* in many cases. It seems very likely that many schools were classified by the State Superintendent differently than the rules technically required. In addition, there is no information in the *Annual Reports* for Shelby County and Overton County, which had records for 18 and 2 schools, respectively, in the school level reports. It is apparent, though, that there are significant missing individual reports in counties with large numbers of second and third class schools. This will bias the dataset towards the bigger, more modern high schools. However, I control for school class in the analysis contained in the later chapters to minimize the impact of this shortcoming.

The Superintendent of Public Education used the individual-level reports and other data to create the *Annual Reports*, where data are aggregated at the county level. In addition, the *Annual Reports* contain county-level data collected from other sources, such as school age population, local tax base, and elementary school data. *Annual Reports* are available as far back as 1875, though before the twentieth century they tend to be more qualitative than quantitative.[21] I entered the data from the annual reports into spreadsheet files. Every page had summations by county and state-wide that enabled me to check that the data were entered correctly.

3.2. From the Annual Reports: Charting the Diffusion of High Schools

I chart the growth of public high schools using the State Superintendent's *Annual Reports*. This presents two challenges: (1) the records are somewhat inconsistent prior to 1919;

[20] One exception to this is the 1928-29 school year, which has slightly less than half the schools available of the surrounding years.
[21] No report was filed for 1899.

(2) information on schools authorized under the 1885 law is incomplete in the administrative records.

The first challenge is a lesser problem. There is no record of the total number of high schools in the 1905, 1907, 1909, 1917, or 1918 *Annual Reports*. Fortunately, the 1906 report gives the number of high schools for 1905 and 1906. Other pieces of data are missing for specific years. While this is not ideal, I can still provide a reasonably accurate characterization of the rise of high schools in Tennessee with the data that have survived.

The second challenge is harder to deal with. Unfortunately, information for so-called "city schools" is missing for large periods from 1900 to 1930, and the data that do exist suggest that such schools were a non-trivial part of the history. For example, city high schools accounted for 38.5 percent of all high school enrollments in 1917 and 20.3 percent of all enrollments in 1930. Where the records allow the two series to be combined, I do so. In most cases, however, I omit the city schools.

There is a notable break in the city school data between 1921 and 1922. The total count of city schools went from 74 in 1921 to 13 in 1922, which seems implausible in reality. At the same time, city high school enrollments dropped by roughly 5,000. It is possible that the students shifted to the "county school" category, where enrollments increased by 9,000. However, the total number of county schools only increased by 23, suggesting that city schools were not merely reclassified. While the narrative histories, the annual reports, and the extant legal records do not provide an explanation for this change, there is a clue in the Annual Reports. After the change, there were only city high schools in seven counties compared to 33 before the change. These counties were Davidson, Hamilton, Knox, Madison, Obion, Shelby, and Sullivan,

all home to major cities. I conclude that smaller cities were folded into the county high school system.

Total county high schools are charted in Figure 9.[22] The trend break surrounding the passage of the 1917 tax is obvious. As the jump precedes the troublesome issue with city schools mentioned above, it cannot be explained by new high school construction to accommodate former city school students.[23] Despite the excitement expressed by contemporaneous sources, the impact of the 1909 funding law appears muted on this graph—county high schools continued an upward trend from a low level, doubling in number between 1906 and 1910.

The plotted points for 1923, 1924, and 1925 are curious. The sharp jump in 1923, followed by a reduction for 1924 and 1925, followed by another jump in 1926 is difficult to explain. It seems likely that 1924 and 1925 are both missing data, but I cannot confirm this. The annual reports do not include a list of schools. All I can confirm is that some counties have fewer schools in 1924 than they did in 1923. Since the microdata are also incomplete, adding up total number of schools in the individual reports cannot solve the problem. This era also occurs right after the reclassification of city schools.

Total city high schools, as listed in the annual reports, are charted in Figure 10. There are no data available prior to 1913. Following the mysterious break in 1921, there are no data from 1922 to 1927, when city schools have roughly doubled. There are fewer total city high schools, but the enrollment numbers suggest that the schools were educating more students per school

[22] The numbers prior to 1909 are likely less accurate than numbers after 1909. Persistent problems getting principals to turn in school reports led to a specific provision of the funding law that required accurate reports to be filed with the State Superintendent in order for state funds to be disbursed.
[23] The reclassification of city high schools discussed above is not for the jump seen in the data. The numbers were stable between 1915 and 1919, during which the jump occurred.

than their county counterparts. Again, because the city high school data are inconsistent, I primarily study county schools. Indeed, Chapter 7 focuses solely on children living in rural areas in part to minimize the impact of city school classifications.

As emphasized in the Introduction, the increase in high schools in Tennessee took place during a general increase in schooling across the U.S., known as the high school movement (Goldin 2008, Goldin and Katz 2009). During this period, large numbers of Americans entered high school and obtained a diploma before entering the workforce.

Figure 11 compares Tennessee schooling patterns to the rest of the country from 1880 to 1950, drawing on the IPUMS cross-sections and focusing on school attendance among 14 to 18 year olds. In 1880, the South had the lowest rate of schooling among 14 to 18 year-olds, just behind the Northeast.[24] Because the 1890 census manuscripts were destroyed, the next year in which we have micro-level data is 1900. By then, Tennessee schooling rates had jumped and were roughly equivalent to the Midwest and the West. The Northeast, however, had experienced a slight decrease in attendance rates and lagged behind the other regions. Over the next 50 years, attendance rates in the Northeast rose to equal those in the Midwest and West while the rates in Tennessee and the South lagged behind. By 1930, rates in the Northeast equaled the South. By 1950, southern enrollment rates lagged roughly 8 percentage points behind the rest of the country. The mystery of southern enrollment rates stagnating during the time of rapid high school expansion might be explained by the phenomena of ungraded schools, where students of

[24] The Northeast figure is surprisingly low. Goldin (1998) looks at enrollment figures starting in 1910 and finds enrollment rates in the Middle Atlantic and East North Central lag behind New England and the Pacific before showing strong growth in the '20s and '30s to catch the Pacific region. Two explanations present themselves. First, large immigrant populations may have skipped high school reducing enrollment rates. Indeed, given Ramcharan's (2002) theory that immigration rates drove up the returns to skilled labor, the growth in attendance may well have been an active response by native parents. The second possibility is that in the pre-compulsory schooling days of the 1880s high returns to work in the industrial North convinced students with a low discount rate to forgo high school.

all ages would be taught by one teacher without explicitly defined grades (Margo 1990).[25] If students between the ages of 14 and 18 shifted from ungraded schools to graded high schools, then overall attendance rates (as gauged by the census) would stay the same while the quality of the education and actual high school enrollments would improve.

To further investigate the relationship between high school openings and attendance rates, I considered enrollment figures from the high schools' administrative records. Although high school enrollment is straightforward to observe in the annual reports, establishing the denominator for an enrollment rate requires some estimation. The annual reports do report the total number of school-age children, defined as those ages 6 to 21 until 1924 and ages 6 to 18 afterwards. The census of school age children is reported annually until 1923 and biannually thereafter. To get the number of *high-school-age* children, I find the number of Tennessee children between 15 and 18 years in the 1920 IPUMS microsample as a percent of all children 6 to 21 (6 to 18 for years after 1925) and multiply this ratio times total number of school-age children in a given year.[26] As the acceptable age range of high school children shrank over time, this method may overestimate the percentage of children in high school in early years.

Figure 12 shows the percent of high school age children enrolled in high school. The result paints a very different picture than the census data. The solid line represents the county

[25] Concurrent with the high school movement in Tennessee was a movement to replace one-room schoolhouses with larger "consolidated" schools with more formal grades and requirements for progression. Many southern states were consolidating rapidly. Mississippi spent almost $16 million to build 969 consolidated schools. By 1934, South Carolina had only 350 one room schoolhouses left in the state (Tindall 1967). Consolidated schools in Tennessee in 1923 were in the early days of this overhaul. Of the state's 6,331 schools only 526 were consolidated. This, however, reflected 66 newly built consolidated schools for the 1922-23 school year (Annual Report 1923).

[26] 15 to 18 is chosen because there would only be 4 annual cohorts of children in high school at any one time. Since the Census took place on January 1 in 1920, it's possible that some 14 year olds were in fact high school age. This is balanced by having some 18 year olds who would have been beyond the "standard" ages for high school (e.g., those turning 19 in Spring 1920).

schools. From a low base, there was a steady rise in high school attendance rates. The rate doubled from 5 to 10 percent between 1910 and 1916, and it doubled again between 1916 and 1922, largely due to a jump in the early 1920s. By 1930, attendance had climbed to about 42 percent. The dashed line attempts to track numbers statewide by including city high school information. This information is only available after 1917. There is a slight gap between the two lines in the earlier years, suggesting total attendance in 1910 was higher than the 4.6 percent for the county schools. Even given this underestimation, the attendance rate is lower than the 47.6 percent calculated with the census information.

The main takeaway from comparing the census-based and the administrative-based graphs is that the increase in high schools did not, on net, induce many children to attend school when they otherwise would not have attended at all. Rather, the rise of high schools appears to have shifted teenage students from ungraded schools to graded high schools. At the margin, it is possible that new, higher quality schools for teenagers might have induced some students to continue with their education beyond what they otherwise would have done (or to attend more days of school conditional on enrollment) because the increase in human capital per time-in-school would have increased while the opportunity cost would have remained the same. But the existing aggregate records are not sufficient to uncover such an effect.

3.3. From the Annual Reports: The Growth of the Teaching Force

This section studies trends in the number and gender of high school teachers over time to establish that (a) high school teacher represented a significant improvement in financial status to women compared to elementary school teacher and (b) that roughly equal numbers of men and women worked as high school teachers, easing comparisons between the two groups. Figure 13 shows the growth of county high school teachers between 1910 and 1930, broken into men and

women. One striking feature is that the number of teachers did not increase as much as the number of schools between 1916 and 1919. This is because the increase in schools in that period was largely an increase in the one-teacher, third-class high schools.

Figure 14 uses the available data from city high schools to measure teachers in the cities. In the early years, city teachers made up a large number of teachers in the state as a whole. By 1927, however, the number of high school teachers in cities had largely stagnated while county teachers had grown significantly. By 1930, city-school teachers made up only 19.4 percent of all high school teachers in the state.

Comparing Figure 13 to Figure 14, we can see that women were more likely than men to be teachers in city high schools but, in the county high schools, the numbers were extremely similar throughout the 1910s and 1920s. The results in the city could represent the availability of higher paying or higher "status" jobs in cities that led to fewer men electing to be teachers.

High school teachers were far better paid than their elementary school counterparts. In 1910, county elementary school teachers were paid an average of $84.39 per year compared to $169.85 for county high school teachers.[27] By 1920, salaries had increased to $347.53 for elementary teachers and $861.20 for high school teachers, increases of 312% and 407%, respectively. Over the decade, teacher pay rose across the board, but high school teachers benefitted more in both relative and absolute terms. For comparison, average manufacturing earnings in Tennessee in 1920 were $854.89. Average manufacturing earnings are unavailable for 1910, but in 1900 the average manufacturing earnings were $793.75 (Haines 2010). So by

[27] The average manufacturing earnings are calculated as the total manufacturing wages divided by the total manufacturing workers (Haines (2010)). All numbers in this section are given in 1920 dollars. In 1910 dollars, elementary teachers earned $39.25 and high school teachers earned $79.00. Conversion used the discount factors on the Minneapolis Fed's website, https://www.minneapolisfed.org/community/teaching-aids/cpi-calculator-information/consumer-price-index-1800

1920 high school teacher wages were competitive with manufacturing work, while elementary school lagged behind. It seems likely that high school teachers were rewarded for greater training as well as, perhaps, the need to attract men to the profession. High school teachers were roughly evenly split between men and women whereas elementary schools were almost entirely women (Perlmann and Margo 2001).

Figure 15 shows the increase in both high school teacher salaries and overall high school spending over time. The chart shows that teacher salaries moved, in percentage terms, roughly the same as total spending over this period. The teachers were the beneficiaries of increased spending on high schools rather than receiving a greater share of school funding.

These stylized facts paint a different picture for high schools than the picture Perlmann and Margo (2001) paint for grade schools. The findings can be reconciled, though. Perlmann and Margo suggest that men were still desired to handle discipline of older boys which should lead to more male teachers in high schools. Another reason men were not completely pushed out of the market may be that the higher wages relative to elementary school increased incentives for men to become high school teachers. At the same time, higher wages for men meant that the situation facing administrators was different than the situation in elementary schools, where Perlmann and Margo (2001) theorize that one of the reasons women came to dominate the lower grades was because reformers could hire more teachers for the same money.

For potential women employees, high school employment offered a position with relatively high pay compared to elementary school teaching. As opposed to other lines of work which may have paid similarly, women would have the opportunity to have some degree of autonomy within their classrooms as teachers. The increases in female teachers over this time

period suggests that high school teacher was an attractive job to well-educated young women entering the workforce.

3.4. School Level Data: Teachers' Names and the Assignment of Gender Codes

One remarkable aspect of the school-level reports is the detailed information provided about the faculty at each school. I created the baseline "teachers dataset" with information from the 1923 school reports. Chapter 4 and Chapter 5 of the dissertation document and interpret the gender wage gap that existed among high school teachers in this period. By historical standards, it is an exceptionally rich dataset, and it covers one of the few occupations where men and women often worked side by side in the same establishment doing the same kind of job.[28]

This paper uses a new dataset I collected and digitized with records for 1,598 public high school teachers in Tennessee. The data were originally recorded in handwritten, high-school-level reports, which school principals submitted annually to the State Superintendent of Public Education. The records are currently stored at the Tennessee State Library and Archives.

As opposed to the aggregate employment information that many studies of this period are based on, I extracted information on individual teachers from the reports including: teacher's name, salary, years of education, university or normal school attended, classes taught, and teaching experience. In addition, each report contains information on the monetary cost of building the school, the classes offered, class schedules, and student enrollments by grade, all of which I also included in the dataset. I supplemented the school-level data with county- and state-level records from the Biennial Reports of the State Superintendent of Public Schools.

[28] For comparison, the 1940 census was the first to inquire about educational attainment and earnings. The teachers dataset appears nearly 20 years earlier, includes not only educational attainment and earnings but also each teacher's place of training, certification, experience in the field, and current teaching assignments. Of course, the original forms also indicate the school in which the teacher worked and report similar information on all other teachers at the same school.

I focus on the information contained in the extant reports for the 1922-1923 school year. Few school-level reports are available before 1923, but 1923 has relatively good coverage and comes long before any effects from the Great Depression. Micro-level records exist for 504 separate public high schools, of which 497 are white schools and 7 are black schools.[29] The Biennial Report indicates a total of 589 white high schools in Tennessee in 1923 but the report is missing counts for Shelby and Overton counties in 1923. Based on data for 1922 and 1924 covering Shelby and Overton, I estimate there were approximately 609 white high schools in Tennessee in 1923. On that basis, I have school-level data for approximately 82 percent white high schools in 1923. Similar estimations for number of teachers suggest there were 1,802 white teachers employed in 1923 and that I have data for roughly 87 percent of white high school teachers. Black schools are harder to identify. According to Biennial Reports, there were 13 black high schools in 1921 and 47 in 1927, the next available count. Assuming that black high schools increased monotonically, I have micro-level data for between 15 and 53 percent of all black high schools. I have made an effort to find schools that are 'missing' from the 1923 data in later years. The levels of pay by gender are similar in the 'missing schools' when compared to others, and so I proceed with a focus on the 1923 dataset and believe they are fairly representative of all high schools in the state.[30]

[29] School race is inferred from teacher race which, in turn, is inferred from college attended. More information is included below.

[30] Using the school reports from 1926, I found schools that did not appear in the 1923 dataset and that were founded prior to 1923. These schools were considered missing and were compared to the remainder of the schools in the 1926 dataset. For first class schools, men in the missing schools made $1,535 (with a standard deviation of $503) and women made $1,049 ($286). For all other first-class schools, men made $1,692 ($623) and women made $1,059 ($247). In third class missing schools, men earned $1,167 ($348) and women earned $752 ($359) while for all other third class schools men earned $1,229 ($400) and women earned $920 ($520). The numbers are close, and in all cases but women in third class schools high standard errors make it impossible to reject the hypothesis that the salaries are equal at the 95% confidence interval.

The original forms do not explicitly indicate gender, and so I used a simple computer program to assign genders based on the teacher's name and related information.[31] The first pass used a series of regular expressions to assign "female" to teachers with "Miss" or "Mrs." associated with their name. [32] Similarly, I assigned "male" to anyone with a "Mr." or "Jr." in their name. Next, I used Excel macros to standardize the names' format, putting periods and spaces after initials and rewriting names entered "Last, First" as "First Last" for subsequent coding.

The teacher's first name is especially useful in determining his or her gender. In the simplest scenario, which had just a first and a last name, I focused on the first name. I did the same if there was a first, middle, and last name. Where the first name was an initial, but there was a full middle name, I used the middle name. And when the name was two initials followed by a surname, I used the two initials as first name.

Once the first name was identified, it had to be classified as a male name or a female name. For most classifications, I relied on information from the IPUMS 1920 1 percent sample. When more than 95% of the people with a given name in the census records were male, I assigned teachers with that name male status. The same rule was used to determine female names. I assign a missing code to those below 95%.

Cases in which the first name was listed simply as two initials, however, poses a challenge. There is no guidance in the reports for how these names should be interpreted. Correctly placing the 515 teachers who have two initials for a first name is important because in the dataset without these teachers, only 204 teachers were identified as men and 672 as women.

[31] The majority of the work was done with Excel macros, which are available upon request.
[32] A regular expression is used by computer programming languages such as VBA and Python to find specific text patterns in a larger block of text and possibly manipulate that text.

As mentioned earlier (see Figure 13), the aggregate administrative records indicate that high school teachers were roughly evenly distributed between men and women.[33] This provides a way of comparing the gender mix implied by different coding schemes against the "true" gender mix. When "initials" are assigned male status, the dataset has 48.3% women, and this is clearly close to the correct ratio. When those with initials are simply omitted from the dataset, it would appear that 76.7 percent of high school teachers were women, which is far too high. From this perspective, classifying "initials" as male seems like the best course.

An important component of assigning gender was accurately identifying the teacher's name. Many of the names transcribed from the original forms could not be found in the 1920 Census information and were utterly unfamiliar. Because the forms were handwritten, copied from microfilm, and transcribed by non-native English speakers, there is considerable room for transcription and keypunch errors. Therefore, I went back to the original forms to check for accuracy of transcription in cases where the names were unfamiliar. Often, this resolved the issue. If I still could not make out the correct name, the observation was assigned a missing code.

Table 4 has summary statistics on the methods of assigning gender. Most men were identified using the two initials rule. Women were roughly evenly split between having a "Miss" or "Mrs." in their name and the "95 percent rule." Only 6% of the total names were assigned by hand due to transcription errors.

3.5. School Level Data: Teachers' Subjects and Education

[33] It is unclear where the annual reports get their information on gender. It is possible that high school inspectors counted teachers by gender during inspections.

I coded information from the school-level reports on each teacher's classes taught and college attended. I created a group of indicator variables (e.g., "English," "Calculus," "Biology") to capture classes taught. In addition, variables such as "Math" and "Science" capture all subsets of the two broader fields. Finally, to account for whether a teacher was the sole teacher at a school, there is an indicator variable for "All."

Strings listing the teachers' college attended were more challenging to code consistently. Many of the colleges were small or no longer exist, making them difficult to track down. Further, several colleges went by multiple names. For example, Tennessee Polytechnic Institute (now known as Tennessee Tech) was also entered as both Poly and T.P.I. Moreover, many times commas were omitted when entering the data and many teachers attended more than one college. The results of coding the colleges are summarized in Table 5, which lists every college attended by at least ten teachers. It shows that teachers tended to be educated in Tennessee, particularly the University of Tennessee, Peabody, and the three state normal schools discussed previously.[34]

Although the principals' reports do not explicitly say whether the high school served white or black students, this can be inferred from the college attended by the school's teachers, under the assumption that black teachers (who attended black colleges) did not teach at white high schools. As in the entire South in this period, segregation was the law in Tennessee. Five colleges appear in the dataset that could be definitively identified as black colleges in the 1920s: The Agricultural and Industrial Normal School, Knoxville College, Lane University, Roger Williams College, and the Tuskegee Institute.[35] This helps identify seven black high schools:

[34]"Normal schools" were post-secondary institutions specifically dedicated to training teachers.
[35] A sixth school that was correlated with several black teachers was Walden University, likely Walden Seminary in Little Rock, Arkansas. However, it also showed up on a teacher at Humboldt High School, which appears to be a white high school, and did not identify any other high schools as black high schools.

Bruce High (Dyer County), College Hill High School (Bradley County), Colored High (Bedford County), Soddy High (Hamilton County), Trenton Junior High (Gibson County), Union High (Sumner County), and Wilson County Junior High School (Wilson County).

3.6. School Level Data: Summary Statistics on Teachers

The raw data reveal a gender wage gap that is consistent with Goldin's (1990) estimates, which may be surprising given that teaching jobs are more similar than men's and women's jobs in other industries at this time. Summary statistics on the teachers' dataset, reported in Table 6, indicate that teachers earned an average salary of $1,156.70 and a standard deviation of $481.02; 95 percent of teacher's salaries ranged between $540 and $2,000.00. The data also show a statistically significant difference between male and female salaries. Men earned an average salary of $1,363.18, with a standard deviation of $538.65, while women earned $937.91, with a standard deviation of $286.00. Thus, the average female teacher earned about 69 percent of the average male teacher, without adjustment for location or observable characteristics. The statistics in the table exclude 206 observations with missing data. There were 129 cases where gender could not be determined, 29 cases without salary information, 45 cases without educational background, and 3 cases where school principal information was missing.[36]

Figure 16 graphs kernel density estimates of male and female teacher salaries. The graph confirms the implication of Table 6 that women's earnings were, indeed, concentrated in a much tighter range and were centered far to the left of the men's distribution. Whereas 13.9 percent of the male distribution is located above $2,500, only 0.5 percent of women earned in that range.

[36] One of the teachers has a salary of $18,000 listed on the form which must be a large error. The salary is roughly equivalent to Babe Ruth's 1921 salary of $20,000 (source: baseball-refence.com) and nearly five times higher than the dataset's second highest salary ($3,750). I omit this observation from the calculations.

Summary statistics also reveal that male high school teachers on average had 10.3 years of experience teaching at all levels (6.9 at secondary school and 2.8 at the current school), whereas women had 6.8 years of experience (4.4 at secondary level and 2.4 at the current school). Importantly, men were more likely to be responsible for administrative functions: 54.0 percent of men were school principals compared to only 5.8 percent of women. The extremely high number of men counted as principals is a reflection of the relatively small size of schools in this period compared to modern high schools. Men and women were similar in their likelihood of having attended both normal school and college. Men were slightly more likely to have attended only college, whereas women were more likely to have attended only normal school. Finally, men tended to be concentrated in teaching math and science whereas women were more likely to teach languages and home economics. "Domestic sciences" includes classes in both agricultural sciences and home economics. Unsurprisingly, men tended to teach the former and women the latter. These differences across gender in experience, administrative responsibilities, and other observables might account for some of the gender wage gap, which I explore in detail in Chapters 4 and 5.

3.7. School Level Data: Principals 1926 – 1935

Chapter 6 considers the evolution of the gender makeup of principals over time. This dataset includes information for 1926, 1928, 1932, and 1935. The key variables were determined as described above for 1923. However, race, college attended, and classes taught were not calculated for this, which I call the "principals dataset." The extant reports represent 86.0%, 93.1%, 93.4%, and 98.9% of the high schools listed in the annual reports.

Summary statistics for the principals dataset are contained in Table 7. Consistent with the teachers data given above, there are significantly more male principals than female principals

and male principals tend to make more than female principals. However, the median age of a principal is quite similar across genders.

3.8. Access to High Schools

The "school access dataset" uses a sample of males from the 1910 U.S. Census linked to the 1930 U.S. Census (hereafter the "linked sample"), created by Collins and Wanamaker (2014).[37] The linked sample originated with the 1910 1% IPUMS sample, and Collins and Wanamaker found the same men again in the 1930 census manuscripts through Ancestry.com. The men are assigned IPUMS-style occupation and industry codes in 1930. I kept only those people living in Tennessee in 1910 who turned 14 between 1908 and 1924. This resulted in a dataset of 1,146 people. Similar to Card and Krueger (1990), I make the simplifying assumption that children would have attended school in the county they were observed in during the 1910 Census.

I added county-level financial information from the *1916 Annual Report*, including: the tax rate for high schools, taxable property, and number of children, by race, in the county. Finally, I add county-level demographic information from the published volumes of the 1920 Census of Population (Haines 2010).

I geocoded all of the high schools.[38] The majority of the longitude and latitude points were taken from the United States Board on Geographic Names (USBGN) database (www.nhgis.org, University of Minnesota 2011). The main benefit of this database is that it contains the actual location of many historical schools. If the information could not be found in

[37] Collins and Wannamaker generously shared the dataset. The linked dataset is composed only of men. Since women changed their names at marriage, it is difficult to track married women between censuses.
[38] By geocoded, I mean finding the longitude and latitude of every school and entering it into mapping software. In this case, ArcGIS 10.3.

this database, which was rare, I used Google Maps. I used the following priority list to assign location, number of schools in parentheses:

1) Historical school location (106).
2) Present school location in USBGN database (270).
3) Present school location in Google (15).
4) Location of the city or town the school was closest to in USBGN database (105).
5) Location of the city of town the school was closest to in Google (11).
6) The river the school was named after (1).

Figure 17 shows the results of this work superimposed on a map of Tennessee. The counties are further color-coded by the percent of the population that lives in rural areas. Figure 18 shows the density of the high-school age (14-20) population in each county. These graphs help explain why entirely rural counties saw such a steep drop-off in number of schools and teachers per student. In the very lightly populated counties in the southwest and middle of the state, authorities elected to build just one first-class high school, usually built in a central location to ease travel concerns. This is most apparent in the counties towards the south central part of the state. However, among the more densely populated rural counties, school distribution looks similar to the distribution in the 80% to 100% counties.

3.9. Potential Sources of Errors

State administrators compiled the early Annual Reports without the aid of calculators. Summations across counties and for the full state sometimes do not add up to the same number as in the original documents. When this happened, I assumed the Superintendent and his staff made an addition error compiling the reports. A related concern is that similar errors were made aggregating individual data into the county-level data points. I can say with a great deal of

certainty that the larger counties (e.g., Shelby, Davidson, Knox) have errors in aggregation. Unfortunately, I cannot say whether these errors are great or small, but I believe them to be rare based on their performance summing across rows and columns.

There are potential errors in the individual school reports as well. Any mistakes the principal made entering the data, either intentional or unintentional, would lead to errors in my data set. Another potential issue is that the forms were often handwritten. In Figure 19, for example, the first college listed is actually "Georgia Robertson Christian College." For non-native speakers trying to transcribe letters, this would be difficult to decipher. Indeed, it was originally transcribed as "Georgia Roletra Christian College." Every attempt has been made to correct these errors, including the work described above to clean the college data. Unfortunately, some errors may not have been corrected. For example, if "John" was accidentally entered as "Joan," I would not have noticed.

The linked dataset from Collins and Wannamaker (2014) is subject to the issues of all linked datasets. In particular, there is a danger of false matches between the 1910 census and the 1930 census. Matches were based on locating exactly one person in the 1930 census with the same name and birthdate as a person in the 1910 1% census sample. If a man in the 1910 census had died and a second man with the same name and birthdate had lived and appears in the 1930 census records, the dataset considers this to be the same person.

CHAPTER 4. THE GENDER GAP IN WAGES

One of the extraordinary features of the school-level data described in the previous chapter is the rich information on Tennessee's teachers' earnings, professional background, course assignments, and school characteristics. In this chapter, I study the gender wage gap and try to measure observables' impact on the observed gap. The empirical findings are specific to a particular historical setting, yet they may have much to tell us about the history of American women's labor force participation and earnings, especially at the upper end of the educational distribution.

Major challenges to studying the history of gender disparities in labor market outcomes include the scarcity of data on earnings, the inability to match employer-reported wage rates to specific workers and their characteristics, and unknowable biases from differences in working conditions and tasks across employers, occupations, and industries. In this context, high school teachers are a particularly interesting group to study. High school teacher was one of the few professions in the early-to-mid 20th century in which women and men worked together doing similar jobs for the same employer. The earnings and background data are of high quality and pertain to well-defined jobs in specific places spread over a wide geographic area. Therefore, it is possible to get a more precise measure of a gender wage gap than is typically possible with historical data.

This chapter establishes the existence and size of a sizable gender wage gap among high school teachers in Tennessee. Further, I examine the impact on the gender wage gap of observables such as quantity and quality of education, experience, subjects taught, and whether or not the teacher was also the principal of the high school. The following chapter extends the analysis extends the analysis by asking (a) whether school districts in Tennessee had significant

market power over women (but not men) and (b) whether the market for female high school teachers in Tennessee behaves the way theory predicts monopsony should. In this case, the latter implies that women's wages should be lower in regions where one employer has more market power.

4.1. Previous Literature on the Gender Wage Gap

Blau and Khan (2016) give a succinct overview of theories of the observed gender wage gap. Among the factors they illustrate as being responsible for the gender wage gap are changes in labor force participation, selection into the labor force, occupational sorting, and career disruption due to motherhood.

Goldin has done significant work on women's wages and labor force participation relative to men. Goldin (1990) traces the gender gap in labor market outcomes from 1815 to 1987. She finds labor force participation rates by women in the early 20th century starting at almost 20 percent in 1900, increasing to 30 percent by 1930, and to over 50 percent by 1980.

Economists have suggested several supply-side forces that may have contributed to women's increased labor market participation over the twentieth century. Changes in social norms (Goldin 1990, Boustan and Collins 2013); technical innovation freeing women from time-consuming household chores (Greenwood, Seshadri, and Yorukoglu 2005); a decrease in child care costs (Connelly 1992, Attanasio, Low, Sanchez-Marcos 2008); and the local diffusion of information about women's employment (Fogli and Veldkamp 2011) all may have affected women's labor supply throughout the century. Later in the twentieth century, the invention and diffusion of the birth control pill may have been particularly important for women's entry into professional fields that in practice require delaying childbirth (Goldin and Katz 2000, Bailey 2006).

Other economists have suggested that changes in the demand side of the market pulled women into the labor force. One argument states that rising relative wages for women caused the increase in labor force participation from 1950 to 2000 (Jones, Manuelli, and McGrattan 2015), Attanasio, Low, Sanchez-Marcos 2008). Perlmann and Margo (2001) look to one industry, schoolteachers, and suggest that school reformers shifted to hiring female teachers, in part, because of their low wages in the late 19th and early 20th century.

Selection into the labor force is likely to be a particularly troublesome issue in any analysis of female teachers in the early 20th century. Blau and Khan (2016) point out "the closer the wage sample is to 100 percent of the underlying population, the smaller the selection bias." Selection is an issue because the observed wage represents an agreed upon wage. If many women receive wage offers so low they elect not to take the offer and remain out of the workforce, the observed wages will understate the gender wage gap. Table 8, which presents labor force participation statistics for white women aged 25 to 65 in the United States as a whole and in Tennessee specifically throughout the 20th century, shows that female labor force participation was relatively low throughout this period.[39] Single women in Tennessee started almost 30 percentage points lower in labor force participation than in the rest of the country in 1900 but had reduced the gap to about 16 percentage points by 1930. Married women in Tennessee followed a similar pattern to married women nationally. Women as a group had modest upward ticks in labor force participation in the beginning of the 20th century and had a participation rate of about 25 percent prior to World War II. The 1950s saw significant growth in female employment, and by 2000 almost 70 percent of women 25 to 65 were in the labor

[39] Labor force participation is taken from Ruggles et al. (2010). The statistics prior to 1940 are slightly different from those after. Prior to 1940, labor force participation is recorded as one working in a "gainful occupation." Post 1940, labor force participation uses the modern definition, whether or not a person was working or seeking employment.

force. Married women started from a much lower base, moved at a similar rate to all women until the 1950s, when their labor force participation grew rapidly and roughly equaled the labor force participation rate for all women by 1970. Further complicating selection issues is that women who graduated college in 1910 largely had to make a choice between careers and marriage (Goldin 1992).

The same kind of comparison can be done for college-educated women after 1940, when the census first inquires about the level of educational attainment. The bottom six rows of Table 8 track labor force participation rates for women who have had one or more years of college education. I chose this cutoff to approximate the amount of training a typical graduate from normal school would have had. (A Normal School is a post-secondary institution designed to train elementary and high school teachers.) The numbers suggest that even women with high levels of academic achievement were slow to break into the workforce. In 1940, roughly 40 percent of college-educated women were in the workforce. In the same year, only 20 percent of married, college-educated women were in the workforce. Similar to all married women, married, college-educated women made strong gains over the next 30 years.

The labor market impacts of marriage and child rearing were also particularly acute in the early 20th century because, in some cases, so-called "marriage bars" either prohibited the hiring of married women or required the firing of a woman after marriage (Goldin 1988, 1990). Implemented by individual firms and school boards, marriage bars emerged in the late 19th century and were particularly important in the teaching and clerical fields. At the time, the costs of terminating an employee were rising as salary schedules necessitated salary increases based on tenure. Marriage bars may have been a way to control rising wages. They largely

disappeared in the 1950s and their removal corresponds with a large increase in married women's labor force participation.

Unfortunately, because such policies were determined at the local level, state reports and laws do not provide systematic evidence regarding the presence, absence, or strictness of marriage bars. In the teachers dataset for Tennessee, there are unambigously married women present (e.g., Mrs. Smith), though marital status is not stated explicitly in the reports. But this does not imply the absence of a marriage bar, only that the county in question did not require the teacher to be fired upon marriage. Moreover, the absence of married female teachers would not allow me to infer that there is a marriage bar, since the social norm against married women's work was so strong.

One upshot is that female teachers were, on average, likely to be younger and less experienced than male teachers, something that I can observe and adjust for directly. Another is that women who persist in teaching might differ on unobserables from women who drop out and from men who persist.[40] This is a more difficult empirical challenge, but one that additional data collection might help address.

Explorations of the gender wage gap have found a significant gap for the period under study. In 1923, using data for 21 separate industries in the National Industrial Conference Board, Goldin estimates women's weekly earnings at 60.7 percent of men's and hourly earnings at 67.2 percent.[41] Goldin further breaks down gender gaps by occupation and industry groups using the

[40] It is possible that more determined and individualistic women would enter and persist in the workforce in the presence of a marriage bar. If these traits are associated with unobservables that lead to higher wages, this could bias the result. If marriage bars are associated with counties that have greater levels of market power, it would bias the results downwards. If marriage bars are associated with counties that have lower levels of market power, it would bias the results upward.

[41] The hourly ratio is higher because women worked fewer hours per week than men. Information is taken from Table 3.1.

Historical Statistics from the U.S. Bureau of the Census. For professional workers in 1930 – which includes teachers – Goldin finds that the earnings of women were 38.5 percent of the earnings of men. This likely represents occupational sorting, with women confined to teaching roles while men took higher paying professional jobs. Unskilled female workers made 57.5 to 60.7 percent of men's earnings in 1930 (Goldin 1990, p. 60–64).[42]

Perlmann and Margo use annual reports from four cities – Grand Rapids, Michigan; Houston, Texas; Portland, Oregon; and Paterson, New Jersey – between 1875 and 1923 to study gender differences in pay for elementary school teachers. The gap was quite large, with women earning between 41.7 percent and 57.2 percent of the earnings of men. Perlmann and Margo speculate that this gap may, in fact, have been a major reason women came to dominate the elementary teaching profession. One interesting outcome in their analysis is that tenure has a negative impact on women's salaries relative to men. They suggest monopsony power by the school districts may explain this outcome.

My research has some features in common with Perlmann and Margo's (2001) exploration of gender wage gaps among teachers. While Perlmann and Margo have data for four large cities, my data come from a variety of demographic environments, including farming communities and smaller urban environments. Further, Perlmann and Margo only have limited information on teacher training and that is strictly for the Houston sample. I have schooling data for all teachers in the data set as well as detailed information on experience, years in school, classes taught, and school characteristics. Using my more complete information, I can analyze the gender gap while controlling for many factors that might impact salaries and I can analyze the gap over a wide range of geographic regions.

[42] Unskilled workers are those workers in sales, manufacturing, service, and agriculture.

4.2. How Large Was the Gender Gap for High School Teachers?

Summary statistics in Chapter 3 show a significant gap between men's salaries and women's salaries but do not shed any light on what caused the gap and how much of the gap observables can explain. This section uses simple regressions and the decomposition method from DiNardo, Fortin, and Lemieux (1996) method to answer this question. I find that gaps of 19 to 21 log points remain even after controlling for all observables. Of the observables, principal status has the largest impact on the gender wage gap.

Before considering the gender gap for high school teachers using the principals' reports, it is instructive to consider the gender wage gap in Tennessee wages for teachers and other occupations using census data. While I would have preferred to do this with data from 1920, the census does not include wage information until 1940. I used the 1940 data to estimate a simple Mincerian wage equation on a sample of all people with 36 or more weeks worked the prior year (to capture the 9 month employment window of most teachers) who were wage earners. In keeping with Mincer (1974), the left hand side of the equation is the log of annual income and the right hand side has years of schooling, years of potential experience (age – years of schooling – 5) and years of potential experience squared. The right had side also includes a dummy equal to one for women, which will give an estimate for the gender wage gap. The coefficients and standard errors for the estimate of female dummy are given in Table 9. This table suggests that for skilled and white-collar jobs, Tennessee tended to have higher gender gaps than the rest of the country, both among whites (the first two columns) and overall (the second two columns). [43] On the other hand, Tennessee had lower gender gaps than the rest of the country for unskilled

[43] I classified anyone performing professional, clerical, or managerial work as a white-collar worker. A white-collar worker or craftsman is considered "skilled." All others are classified as unskilled workers. Teachers are those classified as Teachers (N.E.C.) in the OCC1950 variable.

work. The overall gender gap among teachers in Tennessee was similar to that in the rest of the country.

Next, I estimate the size of the gender wage gap using the data described in Chapter 3. Table 10 contains the results of fitting the following wage equation:

$$\ln(salary_i) = \alpha + \beta_1 female_i + \gamma X_i + \mu Z_i + \varepsilon_i$$

where *X* is a set of person-specific controls, which may include years of experience teaching and its square and years teaching at the current school (tenure) and its square, whether or not the teacher also served as a principal, dummies for whether the teacher attended high school, college, or both college and normal school, with attending only normal school as the omitted category, fixed effects for the specific college or normal school attended, and classes taught. Z is a vector of control variables for class of the school and county-specific controls or fixed effects. β_1 is the coefficient of interest and measures the difference between male and female salaries. If men make more than women conditional on X and Z, β_1 is negative.

In interpreting these results, some caution is required due to the missing third-class schools discussed in Chapter 3. The remaining schools suggest that most of these had a male principal who also did the majority of the teaching. If these schools paid significantly less than other schools that hired more women, it could bias the results by overestimating the wage gap. However, I include controls for class of school and, to the extent that wages in the missing schools behaved the same as wages in the third-class schools I have data for, the impact should be negligible.

The regression analysis includes different sets of controls for background information on the teacher, the teacher's specific job duties, and the county and school the teacher worked in. Column (1) of Table 10 includes no covariates and simply reports the gender wage gap for the set of teachers with no missing variables. Columns (2) and (3) control for experience and

educational background of the teacher. There are two different measures of educational background: type of institution attended (high school, only normal school, only college, or both college and normal school) and specific college or normal school attended.[44] The type-of-institution dummies captures the average return for attending the given type of school relative to normal schools. I choose dummies instead of years attended because normal schools, in particular, taught to completed coursework regardless of how long it took to complete. There is no reason to believe additional years of coursework added any extra ability to students. The specific institution dummies are included to measure unobserved skill, assuming that individual schools admit students from a similar place on the skill distribution.[45] Column (4) adds control variables for tasks undertaken on the job, including principal status and subjects taught. Column (5) includes all of the above and local fixed effects at the county level.

The regression analysis suggests that job characteristics were important factors in explaining the wage gap. Column (1) shows a large gap of 34 log points in a simple regression with only the female dummy on the right hand side. Little change is seen in Columns (2) and (3) when controls for experience and education are added. Both columns use different controls for education, with Column (2) using controls for type of institution attended and Column (3) using the specific college or normal school attended. However, the gap drops to 21 log points in Column (4) when controls for classes taught as well as whether or not the teacher acted as the principal are added. From the previous section we know there were significantly more male than female principals. That principals earned 13 log points more than non-principals suggests that it

[44] Many teachers have both college and normal school study in their backgrounds. The most likely reason for this phenomena is that the potential teachers were accumulating both teaching expertise (the normal school) and subject expertise (college).

[45] A prospective teacher may have elected to attain both a normal school and high school degree for two reasons (1) to apply for jobs both inside and outside the teaching profession and (2) to increase proficiency in their chosen subjects to make them even more valuable to prospective employers.

was an important factor in the observed gap. Another factor is differences in pay for agricultural science teachers. Agricultural science teachers made up 8.6 percent of the teaching force, were 91.0 percent male, and on average earned a salary 30 log points higher than other teachers. My preferred specification, in Column (5), adds controls for county and school characteristics and the gender wage gap shows little decline compared to Column (4). Nevertheless, a large gap remains unaccounted for even with the extensive observables that are available in the principals' reports.

4.3. Wage Distributions

Having identified a difference in mean salaries in the preceding section, I now consider wage distributions and, specifically, what impact observables have on the wage distribution. Accounting for observables is important when comparing estimated wage distributions for the same reason it is important in regressions. It is no surprise that women, who were less likely to be principals or to teach high paid subjects like agricultural science, had a lower wage distribution than men. A more interesting question is what would the salary distribution look like if men and women were doing similar jobs with similar backgrounds. DiNardo, Fortin, and Lemieux (1996) proposed a method for comparing distributions that reweights the female wage distribution as if their observables were the same as men's. Explicitly, they consider the likelihood of an observation being a man based on observables.[46] Then, female teachers are reweighted with those women who, based on observables, are more similar to men receiving more weight. Finally, the kernel density estimator is recalculated using these new weights.

[46] DiNardo, Fortin, and Lemieux actually consider the likelihood of an individual being in a union but the theory applies equally well to male/female as union/nonunion.

Mathematically, this is equivalent to creating a distribution where a hypothetical group with men's distribution of observables is compensated for those observables as if they were women.

In my implementation of DiNardo, Fortin, and Lemieux, I keep their original assumptions and use a probit model to estimate the likelihood of being male and a Gaussian kernel estimator. The observed variables in the probit regression are: principal status, experience and experience squared, tenure and tenure squared, fixed effects for type of school attended, class of school, and race.[47] The results can be seen in Figure 20. The first panel shows the actual distribution of male salaries as well as the counterfactual distribution where women are re-weighted to appear more like men. The second panel compares the reweighted female distribution to the actual female distribution. In the counterfactual, mean earnings for women jump to $1167.93, compared to the male mean of $1363.18 and the female mean of $943.91. The counterfactual women's median is $1,000 compared to the actual women's median of $900 and the men's median of $1277.50. The third panel shows the difference between the female distribution and the counterfactual female distribution. This chart indicates the areas of the wage distribution where most of the change is occurring. A negative number represents an area of the distribution where women are better represented after the reweighting. The large negative numbers around $1,000 and $1,700 annual salary confirms that women's average wages would be higher if they had training and promotion to principal rates similar to men.

As can be seen in panels 1 and 2, the counterfactual women's wage distribution is now bimodal. There remains a large cluster of women in the counterfactual making around $1,000. The upper mode is the result of two dozen women – including female principals – who have

[47] College or normal school attended is omitted because, in order to make accurate estimations, every category included needs to have both men and women. Because many colleges were segregated by gender, I feel the lost data is more significant than increased precision by including the school of attendance.

been statistically "upweighted" so that the female distribution of observables is more similar to men. In particular, these women are more likely to be high school principals and teach classes such as math and agricultural science that are usually taught by men.. It is possible that second mode is too high and should be a smoother distribution at the upper end since it is based on relatively few observations. However, more traditional methods of decomposition, such as Blinder-Oaxaca, would also base women's returns to being a principal or to teaching agricultural science on relatively few observations.

The exercise suggests that if women had men's observables, the mean gap would have been narrowed substantially, driven mainly by an increase in women acting as principals, and as such being paid at the upper end of the wage distribution. There is also a smaller increase in average salary driven by an increase in salary for women at the lower end of the distribution, which can be seen in Panel 3 as the negative number around $1000 in salary.[48]

4.4. Summary

This chapter finds a significant gender wage gap among high school teachers in Tennessee in the early 1920s. This gap shrinks, but is not eliminated, after analysis with both simple regression and DiNardo-Fortin-Lemieux analysis which control for observables related to education, experience, school class, classes taught, and county fixed effects. Given that the hiring was performed at the county level and that the jobs were, by state regulation, similar from school to school, this may be a surprising outcome.

[48] I can also calculate male salaries if men had the same observables as women. Although not reported, this regression also suggests that some of the gender wage gap was attributable to observables but that women also had lower returns to observables than did men.

Of the observables, I argue that gender differentials among high school principals and classes taught are particularly important. I will return to the issue of principals in Chapter 6 but now turn to addressing the residual portion of the gender wage gap.

CHAPTER 5. MONOPSONY POWER AND WAGES

The analysis in Chapter 4 indicates that the gender wage gap among Tennessee high school teachers in the early twentieth century was large. After controlling for observables much of the gender wage gap remains unaccounted for. For this chapter, I focus on what might be learned from variation in the gender wage gap across local labor markets. I conclude that monopsony power may be an underappreciated cause of the gender wage gap in early twentieth century America.

5.1. Monopsony Power in Labor Markets

Robinson (1933) demonstrated that in a labor market with only one employer, wages will necessarily be less than the marginal product of labor. This is because when a monopsonist raises the wage to draw more workers into the labor pool, he also has to raise the wage he pays to every other worker. Robinson went further and argued that monopsony power, combined with different elasticities of labor supply by two groups could lead to wage differentials even in a situation where there is no underlying discrimination. Robinson's argument extends to markets with a dominant employer or relatively few employers, as in the classic Stackelberg or Cournot duopoly models. In these markets, wages would be higher than in the pure monopsony market but still below marginal product of labor. [49] In all cases, however, monopsony power over labor derives from limited employers in a market.

Empirical tests of monopsony's impact on wages emerged in the mid-twentieth century but have had limited success identifying an effect. Bunting (1962) constructed one of the first

[49] It should be noted that the market for teachers also diverges from Robinson's scenario because school boards only have market power over the labor of the two groups, women. However, if the market for male labor is perfectly competitive, that implies that the labor supply curve perfectly elastic and Robinson's conclusions still hold.

empirical tests of the theory that monopsony could lead to lower wages. He calculated "concentration ratios" which measured the amount of labor employed by a single employer, the top four employers, and the top ten employers. Theory states that companies with more monopsony power would pay lower wages, leading to a negative relationship between the two. Instead, when Bunting regressed wages on concentration ratios for nine cities he found a weak positive relationship between concentration ratios and wages.

Other economists have chosen to look at specific markets to find evidence of monopsony. (For a full discussion, see Boal and Ransom 2002.) I survey the literature on three different industries, the "company" coal town, nursing, and sports leagues and find that only one, sports leagues, shows evidence of monopsony power.

One classic industry in this sense is the coal industry, where mining companies often built a "company town" around the mine. The company often owned the housing and the stores in the town. Nevertheless, Fishback (1992) finds that, since companies often had to buy labor on a competitive market at a nearby town, the company housing served to reduce transaction costs and not to extract monopoly rents. Boal (1995) also looks at coal mining operations in West Virginia at the turn of the century, estimates inverse labor supply elasticities, and finds little evidence of monopsony power. He points to high labor turnover rates as further evidence that monopsony may not hold.

More recently, economists have attempted to find evidence of monopsony power by examining the wages of nurses (see Adamache and Sloan 1982 and Hirsch and Schumacher 1995 and 2005). Nurses are a logical subgroup to study for evidence of monopsony because they require highly specialized training that is not easily transferrable to another occupation and the

number of accessible hospitals, the primary employers of nurses, varies across the country. Despite these features the above studies do not find a clear impact of monopsony on wages.

In direct contrast to coal towns and nursing, economists have found evidence of monopsony power in sports leagues, particularly baseball. Rottenberg (1956) was the first to discuss the monopsony power inherent in the so-called "reserve clause." Prior to 1976, baseball players signed a uniform contract for a specified salary. The reserve clause allowed teams to renew the players contract at the value the teams felt was fair. Scully (1974) estimates that the costs of the reserve clause and concludes that an average player only earned between 11 and 20 percent of his fair value. An important factor in maintaining this monopsony was the collusion of the major league teams and the lack of another league to bid up the offered wages.

These three industries suggest labor mobility is an important factor in determining a monopsony. Coal towns and hospitals are unable to maintain low wages, in part, because workers are free to move to another community with higher wages. On the other hand, baseball managed to effectively limit mobility by ensuring that there was nowhere players could move to find another league.

The monopsony theory has enjoyed a renaissance recently based on the pioneering work of Mortensen (1990) and Mortensen and Burdett (1998). These works employ search theory to slightly redefine monopsony. Instead of identifying market power through the presence of limited employers in a given area, Mortensen and Burdett claim all employers derive some degree of market power from search costs faced by workers associated with finding a new job. Search costs imply that all employers face an upward sloping labor supply curve, which is an important trait of monopsonies. The search theory of monopsony was further developed in Manning (2003). This work has led to a growing literature on the impact of monopsony power

on the gender wage gap, such as Bowlus (1997), Ransom and Sims (2010), Ransom and Oaxaca (2010), and Ransom and Lambson (2011). Generally, these models use employee turnover to calculate labor supply elasticities by gender and, then, use these supply elasticities to calculate the wage gap due to monopsony.

Bowlus (1997) uses a version of the Burdett and Mortensen model to estimate job search behavior of 4,300 men and women from the National Longitudinal Survey of Youth. She discovers that search behavior by women – notably women being more likely to leave the workforce temporarily – accounts for between 20 and 30% of the observed gender wage gap.

Ransom and Lambson (2011) also consider the impact of monopsony power on school teachers. They use information on Missouri school teachers in 1990 and posit that married women are less capable of moving to districts that pay higher wages. They find that single women earn about 5% more than single men, though the result is statistically insignificant, but married women earn almost 15% less than married men. They show that this disparity among married men and women accounts for the gender wage gap among teachers. This is consistent with their theory on married women's mobility. This particular paper is similar to mine in that, lacking the panel that search models of monopsony generally use, they find a proxy for search costs and use that proxy to make conclusions on monopsony power.

Ransom and Oaxaca (2010) consider supermarket employees in a regional grocery store chain in the southwest. Using panel data they discover that men have a flatter supply curve than women and that the differences in labor supply elasticity could lead to women earning roughly 20% less than men. However, the actual gender wage gap is only 9-11% which they attribute to various laws and company employment policies. They argue their results are consistent with

Robinson's claim that differences in labor supply elasticities can lead to a gender wage gap even in the absence of overt discrimination.

5.2. Monopsony Power in Tennessee

I now turn to arguing that the local school boards in Tennessee had significant market power over the women's labor market. There is no question that teaching was a significant employer of young women in the time period under study. In Tennessee, a plurality of skilled white women worked in this field, reaching a high of 42 percent in 1910, as seen in Table 11. ("Skilled" is defined as women working in professional, clerical, managerial, or craftsman jobs as specified in Ruggles et al. 2010) The other major occupation filled by skilled women was secretary. Even as more occupations opened to women over time, over half of skilled women were performing one of two occupations each year from 1910 to 1930: teacher or secretary. In rural districts, where there were fewer businesses, female employment was even more biased towards teachers. In 1920, 51 percent of skilled women who worked in rural districts were teachers compared to 12 percent who worked as a secretary. In 1930, it was 59 percent and 11 percent, respectively.

Table 12 compares labor force trends between men and women in 1920. The South in general and Tennessee in particular tend to have fewer people working in skilled work and those in skilled work are more likely to be teachers. For men, teaching was a relatively small part of portfolio of job opportunities (in Tennessee, 0.53% of men in the labor force were teachers). However, teaching was a major employer of women. 11.72% of all women in the labor force in Tennessee were teachers and, considering only skilled workers, teaching made up 30.96% of all employment. This implies that local school boards may have had significant market power over

female applicants but, since they were a marginal employer, would have had to take male wages as given.

Labor mobility is as important a factor in monopsony as market power by the employer. One factor that may have influenced women's labor force participation and outcomes relates to living arrangements for single women. In 1920, approximately 87 percent of all female teachers in Tennessee (primary and secondary) were "never married" with an average age of 29.7 (compared to 38.2 percent and 36.0 for male teachers). If social convention dictated that women ought not live alone, then single women would be compelled to find work within a short distance of their relatives.[50] This, in turn, would influence the range of markets over which they could search for employment and the costs of search more generally. Table 13 considers the marital and residential status of all teachers. The figures suggest that, while living with a relative was common among both single men and women, it was overwhelmingly so for single women, with 93% of single women in rural areas living with relatives. The differential is exacerbated as 73.2% of male high school teachers were married whereas 81.9% of female high school teachers were single. Thus, while most female teachers were single women who were likely to live with a close relative, most male teachers were married and capable of moving to find the best wage.

[50] To test whether this was the case in Tennessee, I searched for 20 random female teachers from the Tennessee teachers dataset in the 1920 census manuscripts via FamilySearch.org. Six were matched, two were married, three lived with their parents, and one lived with her uncle. All six resided with male relatives. Of particular interest is that one of the women living with her parents is 33, further solidifying the idea that it was socially unacceptable for women to live alone even when they were employed in full time teaching jobs.

5.3. A Model of Monopsony Power and Women

I create a theoretical model to explain how monopsony power by local school boards could impact women's earnings and motivate the regressions that follow. The model presented in this section is based on work by Manning (2003) and Burdett and Mortensen (1998). Consider a market where workers are divided into two groups, male and female, and can be either employed or unemployed. There are a number of firms serving the market. Assume, that no one firm has a large share of the market, that all workers have the same productivity (p), and that firms have constant returns to scale.

Assume a firm has $N_{g,t-1}$ employees of gender g at time t - 1. Between time t - 1 and time t, the firm will separate from a certain share of its workers, s(w), and recruit a number of new workers r(w).[51] Both the separation rate and recruitment are a function of w since a higher wage will lead to less separation (i.e., $s'(w) < 0$) and more recruitment ($r'(w) > 0$). Thus, the number of employees at time t is given by:

$$N_{gt} = \left[1 - s\left(w_{gt}\right)\right]N_{g,t-1} + r(w_{gt})$$

In a steady state, recruitment must equal separation so $N_{gt} = N_{g,t-1}$. Using this fact and rearranging yields:

$$N\left(w_g\right) = \frac{r(w_g)}{s(w_g)}$$

Taking the first derivative yields:

$$\frac{\partial N}{\partial w_g} = \frac{s(w_g)r'(w_g) - s'(w_g)r(w_g)}{s(w_g)^2}$$

Since r' > 0 and s' < 0, this derivative must be positive and there is a positive long run relationship between the wage offered and total employment. Thus, firms have some degree of

[51] Note that s(w) is a rate and r(w) is a number. This is in keeping with earlier work.

market power since increasing wages also increases employment. Each firm in the market makes a permanent wage offer of less than p for both genders and earns surplus of p – w. Wages will be set to maximize profits in the steady state. In equilibrium, wages will be set such that all firms make the same profits. Firms that pay more will have more employees but earn less surplus per employee.

These wages form a distribution F_g where $F_g(w)$ gives the fraction of firms who pay gender g wages lower than w. If all employees have the same reservation wage and no firm will set a wage below the reservation wage, b, (since it would receive no recruits), the distribution must lie between b and p. For now, I assume that $F_m = F_f = F$.[52]

Assume that workers have a measure of 1 and receive new offers at a rate λ_g and are separated from employment at rate δ. (Note that both men and women have the same separation rate.) A higher λ represents more offers. When an offer is received, a worker accepts the offer if they are unemployed. If the worker is already employed, they will accept the new offer if it is higher than their current wage. Note that λ captures search costs in this model. Because the model is used to study firm behavior, employees are relatively passive and receive job opportunities randomly. However, in a labor market job offers are the result of a matching process between employers and employees. If search costs make the matching procedure more difficult for the potential employee, job offers will come in less often and λ will drop.

In this scenario, there will be a non-employment rate, u. People will transition from employment to unemployment at a rate of:

$$\delta(1-u)$$

[52] Note that there must be a distribution of wages. If all firms paid the same wage, a firm could raise their wage infinitesimally and have little impact on wages but recruit employees away from all the other firms, raising their profits significantly.

At the same time, people will transition from unemployment to employment at the rate of:

$$\lambda_g u$$

Again, in the steady state transition out of employment will equal transition into employment yielding a steady state non-employment rate of:

$$u_g = \frac{\delta}{\delta + \lambda_g}$$

Now consider the distribution of wages to the employed. Since people tend to concentrate in jobs that pay higher wages, this will be different from the market distribution of wages F(w) and is designated $G_g(w;F)$, which is interpreted as the share of workers of gender g earning less than wage w given wage distribution F. The share of people earning less than wage w is given by:

$$(1-u)G(w;F)$$

Outflows from this group will be given by:

$$[\delta + \lambda_g(1 - F(w))](1 - u_g)G(w;F)$$

This is the separation rate plus the number of workers who get higher paying jobs times the number of people in the group to start with. Inflow is given by this equation:

$$\lambda_g F(w) u_g$$

Since these are equal in the steady state, setting the two values equal and using the definition of the non-employment rate above yields:

$$G(w;F) = \frac{\delta F(w)}{\delta + \lambda[1 - F(w)]}$$

There are three different factors affecting women's labor market participation and the gender gap that need to be considered separately. First, cultural norms may have reduced women's ability to seek work away from a close relative's home. Second, school boards may

have had monopsony power because they employed a large share of female teachers. Finally, women may have been systematically excluded from higher paying occupations.

If women were restricted to jobs close to a relative's home or there was less competition outside of the teaching profession for female labor market participants, it would enter the model as a reduction in λ, a reduction in (acceptable) incoming offers. Considering G(w; F), a reduction in λ shifts the distribution of G(w; F) to the left. In this situation, $G(w_m; F)$ would stochastically dominate $G(w_f; F)$, which in turn implies that there would be a gender gap despite no overt discrimination in the market.[53]

There is a secondary impact to consider, as well. If the likelihood of finding another job is lower for women, this implicitly reduces the cost of offering women a lower wage because the cost of recruitment and training would drop. The cost drops because it becomes less likely that a woman finds a better job and leaves and, consequently, the likelihood of needing to find and train a replacement drops. Because of this, the original assumption that F(w) is the same for both sexes is unfair. Women will have a lower distribution of offered wages which will also increase an observed gender gap. The distributions of F(w) could be further biased in favor of men if women were excluded from higher paying, long-term jobs as Goldin (1990) describes. One implication of a wage offer curve skewed upward for men relative to women is that men who remained in teaching would tend to be higher paid because low-paid men would have been more likely to find a higher paying and leave the profession.

This section has argued that monopsony power over the women's labor market on the part of the school boards could have shifted the wage distributions between men and women. This shift, even in the absence of overt discrimination by high school administrators, would have

[53] By "no overt discrimination," I refer to Becker's (1971) concept that people obtain a disutility from employing a certain group and, as such, will only hire that group if the group accepts lower wages.

created a gender wage gap. If there was discrimination by the administrators (the presence of marriage bars, for example), it would have served to widen the gap.

5.4. Outside Options for Women and School Board Market Power

Before we can turn to analyzing the impact of market power on the gender wage gap, I first need an effective way to measure the degree of market power by county. In keeping with prior work, I would like to use a measure of the proportion of educated women who work for a given employer. In this case, the relevant employer is the school board. Unfortunately, I cannot construct such a variable because information on education does not appear in the census until 1940. Instead, I construct an index that measures potential outside options for well-educated women.

The Market Power Index is calculated at the county level because county school boards made hiring and firing decisions.[54] The process the county boards used to hire teachers is described in the Nashville Tennessean, the major newspaper in Nashville. The teachers in Bedford were "elected" by the county board in 1922, not the individual principals. The principal and assistant principal of the major high school in the county, Shelbyville High, were both extended at the same meeting ("Teachers May Lose Licenses," Nashville Tennessean, Jun 18, 1922; ProQuest Historical Newspapers: The Nashville Tennessean p. 3). Another story about the Montgomery County School board in 1921 tells how the board reelected two teachers a month late and cut other salaries by 10% because of budget issues ("Two Teachers in High School are Reelected." Nashville Tennessean; Jun 26, 1921; ProQuest Historical Newspapers:

[54] There is no evidence that principals played a role in recommending teachers to school boards; however I also cannot firmly rule it out. Nonetheless, school boards had absolute authority over the school funds. Even if specific teachers were recommended to the school boards by principals, the school board was a factor in any salary package given to the teachers.

The Nashville Tennessean p. 5). An earlier article for the City School Board in Huntington implies that all positions, including the principal, were reelected every year. ("City Board Elects High School Teachers." Nashville Tennessean and the Nashville American; Jun 24, 1918; ProQuest Historical Newspapers: The Nashville Tennessean, p. 6).

Tennessee's counties also exhibited variance in the gender wage gap, as can be seen in Figure 21.[55] In general, gender wage gaps tend to be larger in the more rural, sparsely populated counties in central and south-central parts of the state. The gender gap is smaller in the east, where there were more cities and fewer farms, as well as the counties near Memphis, Cookeville, and Chattanooga.[56] It seems likely that more urban counties had more options for female laborers and, as such, the school boards had less market power in wages.

To measure the degree of market power individual counties had in the labor market for skilled women, I constructed a proxy measure of the local employment options for well-educated women outside of teaching. The measure is driven by differences across industries in the intensity of employment of well-educated women interacted with variation across counties in industry composition. Specifically, the index is calculated using the education and industry-of-employment information for women contained in the IPUMS 1940 Census of Population 1-

[55] Specifically, I estimate county-level gender wage gaps using the following regression equation:

$$\ln Y_{ic} = \sum_{c} \beta_{1c} I_c + \sum_{c} \beta_{2c} I_c F_i + \gamma X_i + \varepsilon_{icx}$$

where Y_{ic} is income of teacher i in county c. I_c is a dummy for the county, X_i is a set of individual and school level controls. The individual and school-level controls include experience and experience squared, tenure and tenure squared, whether individual i was the principal of the school, dummies for the subject taught, dummies for the specific college or normal school attended, and dummies for the class of the school (e.g., first class, second class, third class). ε_{icx} is a stochastic error term. All standard errors are clustered at the county level. In this regression, β_2, the coefficient on the interaction of county dummies and the female dummy, captures the gender wage gap for each county. Note that there is no intercept in the regression equation. This means that there is no omitted county.

[56] Davidson County, home of Nashville, does not follow this pattern. However, it does have a lower wage than many of the immediately adjacent counties.

percent sample (Ruggles et al. 2010). I started with everyone in the country between the ages of 18 and 65 who was in the labor force. Then, I identified well-educated women (defined as one or more years of college instruction) in the workforce. Next, I divided workers into industry groups (IND1950 in the IPUMS dataset) and calculated the percent of workers in each industry who were well-educated women at the national level. At the county level in Tennessee, I used the 1920 Census microdata sample to calculate the percent of the workforce employed in each industry, excluding education. I multiplied the 1920 industry percentages (county level) by the 1940 percent of well-educated women in the given industry (national level) and summed the product over industries within counties.

I refer to this sum as the county's "Women's Option Index," which is inversely related to the market power of the school boards. I then transform the Women's Option Index (I_c) into the Market Power Index using the following:

$$\frac{I_H - I_c}{I_H - I_L}$$

where I_H is the high value of the Women's Option Index and I_L is the low value. This "Market Power Index" varies between zero, which reflects the lowest market power a school board has in the state, and one, which represents the highest market power. A benefit to this specification of the Market Power Index is that the corresponding regression coefficient on the index represents the maximum gender wage gap added by market power. The Market Power Index has a mean of .596 and a standard deviation of .246.

The Market Power Index is somewhat different than traditional measures of monopsony, but it is consistent with prior estimators and avoids some potential sources of endogeneity. For instance, Bunting (1962) used employment concentration ratios, and several authors in the literature have followed. However, local employment concentration ratios are potentially

endogenous to the gender wage gap. Since there were relatively few women with high-educational attainment in this period, a county school board that has been forced through competition to offer educated women a higher wage may attract a large number of the pool of educated women, making it seem that options outside of teaching are low. By using a measure of potential outside employment, I control for the possibility that actual women's employment patterns reflect the rejection of opportunities outside of teaching.

In this setup, the local Market Power Index captures employment options for educated women outside of teaching. Furthermore, female employment rates at the national level are unlikely to be endogenous with number of teachers hired at the county level in Tennessee. It is easiest to think of the result as "potential" female employment.

However, there is still potential endogeneity in the model if industrial structure is not quasi-randomly distributed with respect to women's wage structure. Fortunately, there is reason to believe that businesses did not locate specifically considering women's wages and labor force participation. Outside of teaching, educated women made up relatively little of the workforce. [57] Educated women made up 48% of employees in education but only 13% of employees in hospitals, the industry with the next most educated women, and 11% in the IPUMS defined industry of "welfare and religious services." It seems unlikely an entrepreneur would select a location for their business based on wage rates for 10% or less of their workforce.

5.5. Regression Results

I use the Tennessee teacher data to estimate the following equation:

$$w_{ic} = \alpha + \beta_1 F_i + \beta_2 \varphi_c + \beta_3 \varphi_c F_i + \gamma X_i + \varepsilon_{ic}$$

[57] Educated women refers to women with 1 or more years of college education.

where w_{ic} is the natural log of annual income of person i in county c, F_i is an indicator for female, φ_c is the Market Power Index, X_i is a set of individual-level controls, and ε_{ic} is a stochastic error term. β_3 is the coefficient of interest because it measures the change in the gender wage gap associated with variation in monopsony power, at least as proxied by the market power index. Controls are the same as in the regressions in Table 10.

If monopsony power affected salaries as hypothesized above, we would expect to find $\beta_2 = 0$ and $\beta_3 < 0$. That is, men should receive no extra returns for working in an environment with more job opportunities for women.[58] Women, however, should see their salaries decrease as the school board's market power over skilled female labor increases.

Table 14 reports the results of the regressions with the Market Power Index. Column (1) is a baseline regression which only regresses log income on a female dummy, the market power dummy, and the cross product. The coefficient on the cross product is a relatively large -0.282. This implies that a one standard deviation decrease in market power is associated with a 6.9 log point decrease in the gender wage gap. In localities where skilled women had more opportunities for employment, based on local industrial structure, the gender gap among teachers was smaller. Table 10 showed that the total wage gap is 34 log points, meaning the market power coefficient identified here is relatively large compared to the total gap. The coefficient on market power alone (β_2) is statistically insignificant, suggesting that male wages were not closely associated with the measure of market power.

Column (2) adds controls for teacher background, and Column (3) further adds controls for principal status and subjects taught. In these regressions, the coefficient of interest drops. Columns (2) and (3) suggest a one standard deviation decrease in market power is associated

[58] In fact, since men are hired from a competitive labor market in this theory, they should not receive any increased salary for changes in outside employment options.

with a 3.6 log point and 3.7 log point decrease in the wage gap, respectively. The relatively small change in the coefficient of interest from Column (2) to Column (3), where principal status is included as a control, suggests principal status is not correlated with market structure. The full impact of principal status is caught in an 11.6 log point drop the base female wage gap, which is what we would expect.

Specifications in Columns (4) and (5) add more controls, including some county-level characteristics. Column (4) includes controls for the county's black population share, crop value per farm, and the manufacturing wage per worker (Haines 2010).[59] In this regression, the one-standard deviation decrease in market power is associated with a 4.8 log point decrease in the wage gap. While the coefficient on Market Power (not interacted with female) switches from negative to a positive 0.082, the estimate is not statistically significant.

Column (5), which is my preferred specification, replaces the county-level attributes with county fixed effects. In this specification, a one standard deviation decrease in market power is associated with a 5.3 log point decrease in the wage gap. Because the Market Power Index is calculated at the county level, β_2 cannot be identified in this specification. However, the coefficient on the cross product between Market Power Index and female, β_3, is still identified.

These results suggest that there is a strong correlation between labor market opportunities for skilled women outside teaching due to industry structure and the gender gap in wages. Any explanation for the correlation must explain both the negative correlation with women's salaries and the lack of correlation with men's salaries. The evidence is consistent with monopsony

[59] Black population share is used following Margo's (1990) suggestion that southern counties often distributed money based total children but the money was often spent biased towards white children. This would imply white salaries should be higher in counties with more black children.

models in which difficulty obtaining wage offers drives down workers' wages. In addition, a monopsony model is also consistent with the wage distributions found in Figure 20.

5.6. Wage Distribution Conditioned on Market Power

One of the implications of these results is that counties with significant market power would populate the low wage cluster for women while counties with less market power would pay wages closer to men's. I test this by splitting the dataset into two equal groups based on the Market Power Index. Table 15 reports the quartiles for all counties, low market power counties, and high market power counties. Table 15 is further grouped by men's salaries, women's salaries, and a DiNardo-Fortin-Lemiuex counterfactual of women's salaries. Since the counterfactual women's distribution is calculated, a bootstrapped standard error is included.[60]

If monopsony power impacts the gender wage gap, we would expect the counterfactual women's salaries to be lower in high market power counties. If counterfactual women's wages are more similar to male wages, this implies that returns to observables are similar between men and women. Correspondingly, counterfactual women's wages more similar to actual women's wages implies that women's returns to observables are much lower than men's returns to observables. The impact of monopsony power should be to drive down returns to observables for women compared to men.

Table 15 shows exactly this relationship. The entire dataset shows that women's wages increase when their observables are similar to men's but a large gap remains. However, in low market power counties the top of the counterfactual women's distribution increases and is closer to the top of the men's distribution than the top of the actual women's distribution. In these counties, women were rewarded relatively well for their observables but an observed gap appeared due to a lesser accumulation of observables. In high market power counties, two things

[60] The bootstrap is a Monte Carlo simulation where a new dataset is drawn, with replacement, from the original dataset. I perform the full DiNardo-Fortin-Lemieux reweighting on the newly drawn dataset and obtain median values for the counterfactual distribution. I ran the simulation 20,000 times and calculated a standard error from the distribution of medians created this way.

are of particular note. First, the actual women's distribution is lower than the actual women's distribution in low market power counties, which we would expect if market power is driving gender gap differences. Further, the counterfactual women's wages changes relatively little suggesting that observables played a lesser role in the observed gender wage gap.

Some caution is required interpreting Table 15 because of the relatively high standard errors associated with the upper end of the counterfactual women's wage distribution. The high standard errors are likely a consequence of having relatively few women in the upper end of the wage distribution.

5.7. Robustness Test

One potential issue with the analysis earlier in this chapter is that the coefficient of interest, Female × Market Power, may be capturing an omitted variable. If there is another effect at the county level that is correlated with market power and responsible for the observed gender wage gap, it would be correlated both with the coefficient of interest and women's salaries.

I test whether there is an omitted variable by re-running the regression in Table 14, Column (4) but include cross products for female and percent black, average value of farm products, and average manufacturing wages. If the coefficient of interest is, in fact, capturing the impact of an omitted variable and the omitted variable is also correlated with one of the newly added variables, I would expect the addition of these other county-level cross products to reduce the significance of the coefficient of interest.[61]

The results are reported in Table 16 and suggest that there is no impact on the significance of the coefficient of interest. In fact, while the coefficient of interest is still

[61] Significantly, the percent of the county that is rural is correlated with both the market power index (.82 coefficient of correlation) and the manufacturing wage (-.42 coefficient of correlation).

significant and the point estimate slightly larger in absolute value, the other cross products are all statistically insignificant. Removing the Market Power Index and its cross product from the equation does not impact the significance of the other cross products.

5.8. Discussion and Conclusion

This preceding two chapters expand on our current knowledge of the gender wage gap by bringing individual-level wage data on white-collar female employees to an era where most wage data are aggregate and come from the industrial sector. My results suggest that regional effects had an important impact on the observed gender wage gap. The results are consistent with the theory that monopsony power is at least partially responsible for the historical gender wage gap.

To the extent that monopsony power was a driving force behind wages, my results suggest changing social norms which allowed women both to pursue jobs previously held only by men and to live apart from their family were driving forces behind wage equalization over the course of the twentieth century. The increase in demand for women's labor would have had two effects that brought about this result. First, more competition for female labor would have reduced the leverage of employers such as school boards and the companies hiring clerical workers. Second, wages would have necessarily been bid up to bring more women into the labor the market and meet increased demand. Third, labor force mobility would have equalized women's wages across the state and, eventually, the country. Proving this is beyond the capability of this dataset and is left open for future work.

CHAPTER 6. HIGH SCHOOL PRINCIPALS, 1923 – 1935

6.1. Introduction

Chapters 4 and 5 suggested that one factor in the observed gender wage gap was the differential rate of promotion to principal between men and women. This chapter uses an expanded dataset of teachers to evaluate both whether women were promoted to the position of principal more often as time passed and whether female principals' wages were more equal to male principals' wages. An affirmative answer to both questions would imply that women were making inroads into the labor market. I study the gender makeup of principals in 1923, 1926, 1928, 1932, and 1935.[62] I find that, far from making inroads into a male dominated profession, women became less likely to act as principals over time. This is particularly true of first class high schools and during the Great Depression.

The period under study covers both the "Roaring 20s" as well as the worst parts of the Great Depression. The respective macroeconomic regimes of the time would have had an impact on employment, in particular women's employment and labor force participation. Goldin (1990 Table 2.2) estimates that women's labor force participation generally increased throughout this period. Married white women between 25 and 34 increased labor force participation from 7.7 percent in 1920 to 11.5 percent in 1930 to 16.7 percent in 1940.[63] Older married women experienced a similar trajectory although at a lower level. Single white women between 25 and 34 also saw an increase in labor force participation, rising from 67.7 percent in 1920 to 75.4 percent in 1930 to 79.4 percent by 1940. Single white women between 15 and 24 were an

[62] The data for 1926, 1928, 1932, and 1935 was collected and digitized through a dissertation enhancement grant by the National Science Foundation.
[63] Black married women had higher participation rates throughout this period. For a discussion of this phenomenon, see Goldin (1990) and Boustan and Collins (2013).

exception and saw their labor force participation decrease from 46.6 percent to 42.6 percent to 40.8 percent.

The growth in women's labor force participation is somewhat surprising given that the Great Depression led to greater social acceptance of discrimination against women (Goldin 1990 p. 112). At the time, gender discrimination was legal and the pressures of the Depression led employers to give jobs with the promise of promotion to men since they were the "head of household." In practice, this took the form of stricter marriage bars. Overt discrimination could partially explain the job market participation of young women in Goldin's data, although increased high school attendance by this group seems a more likely explanation.

The increase in married women's employment observed in the data could be explained by Finegan and Margo (1994). They argue that the added worker effect (the likelihood that a wife would enter the workforce should her husband lose his job) caused more women to be in the labor force during the Depression than would have been the case in the counterfactual, though this effect was offset by work relief programs. The picture that emerges of the Depression is an era where jobs with high pay or chance of promotion went to men, the breadwinners, while women took "dead end" jobs such as stenographer, typist, and telephone operators. Applied to the teaching profession, that suggests that men should be more likely than women to become principals during the Depression.

In this chapter, I examine trends in both the share of principals who were female and relative salaries between male and female principals and find differential trends between the variables. While the ratio of male to female principals increased significantly between 1928 and 1935, wages between the genders tended to converge. The cause of the change in the gender ratio is somewhat unclear, though the data suggests that changes in the location and class

makeup of the school system (i.e., first, second, or third class schools) over that period played an important role.

6.2. Within Gender Comparison of Principals Over Time

This section considers changes in principal characteristics within genders. Unfortunately, no information has survived on exactly what decision metrics were used to select high school principals. However, we can infer some of the qualifications quantitatively. Principals tended to have significant experience as high school teachers. 75 % of all principals had the most experience at their particular school and an average of 9.3 years of experience compared to 5.3 years of experience for all other teachers. Principals also tended to have more years of tenure at the current school, with 4.2 years compared to 3.3 years for all other teachers.

Principals were more likely to have both college and normal school in their backgrounds (21.9% of all principals) than all other teachers (11.6%). Other teachers were more likely to have only a college education (79.1% compared to 63.0%). Most principals taught full course loads in addition to administrative work.

Table 17 gives information on the salaries (given in 1928 and real dollars)[64], experience in all high schools, tenure at the current school, education type, and class of school managed by the principals by year and by gender. Real figures for both genders show a spike in salaries in 1932 before falling off again in 1935. The nominal figures suggest that some of this is a result of deflation during the depression, as nominal figures tend to peak in 1928 before falling in the Depression. It seems likely that the severity and length of the Depression potentially explains the real wage drop in 1935. The tenure and experience for both groups tends to increase over

[64] The relevant inflation figures were taken from the Bureau of Labor Statistics, http://data.bls.gov/cgi-bin/cpicalc.pl.

time, possibly reflecting the impact of the a more professional teaching force created under the auspices of the 1909 General Education Bill. It is also possible that the increase in tenure and experience are impacted by increasingly poor alternatives in rural job markets.

Looking at the class of school the principal ran, we see that men were increasingly likely to be principals of first class schools, with first class schools employing almost 4 out of every 5 male principals by 1935. On the other hand, women tended to work in a similar ratio in all 3 classes of school throughout this period, even if the raw number of female principals decreased over time.

6.3. Gender Ratios among Principals

Figure 22 shows raw numbers of men and women serving as a principal between 1923 and 1935 as well as the share of all Tennessee high school principals who are women. The share falls in 1926 and then remains relatively static until 1935, when a moderate increase in number of male principals and a relatively large decrease in female principals leads to a large increase in the ratio of male to female principals. This is broadly consistent with the theory that supervisory jobs were more likely to go to men during the Depression. It is also noteworthy that there are relatively low numbers of female principals throughout this period.[65]

I further separated the principals into groups based on the "class". Figure 23 shows principals of first class schools. The first class schools experienced a similar, if slightly stronger, decrease in the share of women working as a principal. In 1935, only 3% of all first class

[65] It is possible that consolidation among high schools, where smaller third class high schools run by women were replaced by one first class high school run by a man, explains some of this trend. However, consolidation of the third class schools started in the late 1920s and so should have impacted the gender ratio earlier.

schools were run by women. The total share of students who studied under a female principal also decreased over time, from 4.7% in 1923 to 2.1% in 1935.

It is possible the explanation for increasing numbers of men working as first class high school principals can be found in the makeup of the first class schools. Over the course of the 1920s, first class schools were becoming larger institutions. The mean number of teachers in a first class school increased from 4.9 in 1923 to 6.57 in 1935 and the median increased from 4 to 5. The largest school went from 21 teachers in 1923 to 35 in 1935. It is possible that men were viewed as the best option to teach larger schools for one of two reasons: (1) a larger teacher base implied more male teachers, and the school boards may have felt male teachers should report to male principals and (2) larger schools required more organizational work and administrators may have thought men were better suited to such work.

Third class schools are shown in Figure 24 and a unique pattern emerges. In third class schools – schools with one teacher who taught two years of high school – there was an increase in the number of female principals in the 1920s coupled with a decrease in the male/female principal ratio. However, in 1932 and 1935, after the Depression had hit, there was a decrease in both raw numbers of women and share of women working as principals.

While Figure 24 offers some support to the hypothesis that women were being pushed out of the principal ranks during the Depression, I fit the following equation to test whether women were less likely to be principals in later years after controlling for school level observables:

$$Principal's\ Gender_{sc} = \alpha + \beta_1 Teachers_s + \beta_2 year + \beta_3 county_c$$

where gender is a dummy equal to one if individual i living in county c at school s was a woman, teachers is the number of teachers employed at school s, year is a set of dummy variables for 1926, 1928, 1932, and 1935 (1923 is the omitted variable), and county is a set of dummy

variables for the counties. Since each school has only one principal, this is in effect a regression where schools are the observation. The number of teachers is included to control for the possibility that men were more likely to be selected as the principal of larger, more complex schools. The gender variable on the left hand side should capture the same information as the ratio in the figures above. In this case, a higher number on the βs implies a lower male/female ratio. The variables of interest are the β_2 dummy coefficients which capture the relative likelihood in each year that a woman would be a principal relative to 1923.

Table 18 displays the results. A similar picture to the graphic analysis emerges from the regression analysis. Columns 1 and 2 consider all schools. Column 1 suggests that women were increasingly less likely to be principals through this period although the effect lessens in Column 2 when I add controls for number of teachers in the school and "class" of the school. Number of teachers has no statistical impact on the likelihood of employing a female principal.

Columns 3 and 4 consider only first class schools and a similar pattern emerges in Column 4 as in Column 2. While the point estimates suggest that women were less likely to be principals in the 1930s and larger schools were less likely to employ women as principals, none of the variables is statistically significant. Column 5 considers only third class schools. While the point estimates suggest that there were slightly more female principals in 1926, 1928, and 1932 compared to 1923 and slightly less in 1935, only 1928 is statistically significant. The lack of significance in the regressions suggest that other factors may have been impacting the gender of principals.

One explanation for the decrease in female principals is that the school makeup was shifting away from second and third class schools, which had relatively more female principals, to first class schools, which had relatively fewer. I calculated the share of female principals there

would have been if the classes kept the same share of female principals as in 1923 and altered the mix of schools to match what they were in later years. The counterfactual resulted in about 1 percentage point more female principals in 1935 than were observed in the 1935 data. The changing mix of schools appears to be a factor but not the only cause of the decrease in female principals.

Overall, I find a decrease in female principals over my period of study. The decrease appears to have accelerated during the Great Depression. Further exploration is required to determine how important a factor the Great Depression was. It also appears that changes in the average size and complexity of the high schools could have led to more male principals.

6.4. The Gender Wage Gap among Principals

I next turn to analyzing the gender wage gap between male and female principals over this time period. I use a regression model similar to that in Chapters 4 and 5:

$$\ln(Salarly_{iysc}) = \alpha + \beta_1 F_i + \beta_2 I_y + \beta_3 F_i I_y + \gamma_1 X_i + \gamma_2 X_s + \gamma_3 X_c$$

where F is an indicator for being female, I_y are a set of dummies for the year (1926, 1928, 1932, or 1935, 1923 is the omitted category), and the X_s are controls at the individual level (whether the principal attended high school, normal school, college, or normal school and college), the school level (first, second, or third class school), and county fixed effects. β_3 is the coefficient of interest because it will give the increase or decrease women earned relative to men in a given year. β_2 gives the change in male salaries relative to 1923.

Table 19 contains the results of the above regression. Column (3) contains my preferred specification which contains controls for class of school and county fixed effects. The regression shows that the gender wage gap among principals jumped in 1926 relative to 1923 and slowly decreased until 1935, when the gender gap was lower than 1923 for the first time, although there

was still a gender wage gap of about 5 log points based on point estimates. It is unclear if this represents a permanent catch-up by female principals or a temporary decrease in male salaries due to the Depression.

Experience at the current school seems to have little impact on salaries but total experience teaching secondary school does have a slight positive correlation with salary. The impact of high school, college, and both college and normal school attendance are given relative to normal school attendance. College attendance appears to have been an important factor in determining salaries. It is perhaps surprising that principals with only a high school degree earned slightly more than principals with a normal school degree. One possibility for this is that the principals with only a high school degree had, on average, two more years of experience and tenure than principals with a normal school degree and received a bonus for their knowledge that is not fully captured by experience and tenure.

6.5. Conclusion

The statistics given in this chapter suggest that men dominated, and increased their dominant position in, the position of high school principal through the period under study. Most high school principals had college degrees and the share with college degrees increased in the period under study. To the extent that more women entered the labor market in this period, their job choices may have been constrained by societal factors in the Depression.

I believe future research is required to better understand the interplay between the Depression and the hiring process for principals. While it is clear that men were even more likely to be a principal during the Great Depression, it is unclear what factors were most responsible for the change. Among the possible explanations are: an increase in the supply of

qualified men as the labor market collapsed in the 1930s, and a decision by local school boards to hire the “head of household” during the Depression.

CHAPTER 7. SCHOOL ACCESS

7.1. Introduction

A persistent problem in measuring returns to education entails finding an instrument for educational attainment that is plausibly exogenous to the system and affects large numbers of children. Card (2001) suggests that instrumental variable approaches inherently have problems estimating the impact of schooling since an instrument that changes the rules of the existing schooling system will also change the composition of the schooling body, biasing comparisons.

In this context, school access models can be particularly enlightening. School access looks at scenarios where a community goes from having no viable schooling options to having at least one. The process that provides access needs to be exogenous for the model to provide valid estimates of the impact of schooling on outcomes. Duflo (2001) uses a major school construction project in Indonesia to identify returns to education in Indonesia. In the United States, such models are generally confined to measuring the impact of college education (e.g. Currie and Moretti 2003) since almost every community had some sort of school house in place from its inception. However, many rural communities in the early 20th century had no access to public high schools until legislative change forced counties to establish such schools.

Scholars who have studied southern economic growth, particularly Coleman and Caselli (2001), have posited that increases in education were an important factor in the transition of southern labor out of agricultural work. Wright (1986) argues that another important factor was internal migration in which skilled labor flowed from North to South and unskilled labor did the opposite.

However, it is also possible that rural southern high schools, by increasing scientific knowledge, increased returns to farming and, subsequently, increased incentives to work in

agriculture. Previous literature has found a positive correlation between education and returns to farming. Most germane to this study is work by Parman (2012) who used the 1915 Iowa state census to show that schooling improved farmers' private returns by 5.2 percent per year of high school and had spillover effects on their neighbors, increasing the neighbors' returns by 2.8 to 3.8 percent per year of high school attended. Other authors consider the impact of schooling on agriculture in developing countries. Asadullah and Rahman (2006) conducted a study in Bangladesh similar to Parman's, although they considered total years schooling and not exclusively high school education, and found that education improved agricultural output of households but had no discernable spillover effects. Lin (1991) found that educated farmers in China are more likely to adopt new technologies.

I propose studying the impact of education on the transition out of farming by specifically studying the outcomes of rural children who were exposed to high school in Tennessee in the early 20th century. I use a 1917 law, which differentially impacted access to high school by rural students, to identify the impact of high schools on future career choice. I consider the impact access to high school had on rural children's decision to remain farmers or enter white collar work. For the purposes of this chapter, farmers are only those who own or operate their own farms. Farm laborers are not considered "farmers" since the benefits to attending high school – decisions about which crops to plant and which methods to use on the farm – are more likely to be made by the operators and, as such, any returns to schooling would be to the operators and not the laborers.[66] I consider white collar work to include professional, clerical, or managerial occupations. My results suggest that some white children used high schools to become more proficient farmers.

[66] The results are similar when farm laborers are included.

While this chapter focuses specifically on Tennessee, Tennessee was not alone in training its citizens in agricultural science. Mississippi chose to build separate agricultural high schools for students who wanted to learn farming. These were boarding schools with an attached farm and no county had more than one. While they clearly focused on agriculture, they also required training in "standard" high school subjects. In 1917, there were 47 white agricultural high schools compared to 155 non-white agricultural high schools in 1920 (the closest year for which statistics are available). Florida's Biennial Report of the Superintendent of Public Instriuction for 1918-1920 describes increased demand in the state to teach agriculture at the high school level and the beginnings of agricultural education in the state.

7.2. Previous Literature

The impact of southern high schools on their students is an important component of both the literature on the high school movement and the literature on regional wage convergence. Goldin and Katz (2009) document that the high school movement came to the South later than it did the rest of the country. They further argue that the national growth in high school attendance paralleled increased returns to high school education in the country. The returns came from high demand within the national economy for skilled workers. Although not specifically mentioned, one implication of this argument is that rural children who attended high school in the early 20th century would earn higher returns in the skilled labor market. Using a general equilibrium framework, Coleman and Caselli (2001) explicitly posit that a decrease in education costs led to a transition from agricultural into nonagricultural work.

The connection between education and regional wage convergence is an important theme in the economic history of the American South. Collins (2006) also finds strong regional convergence is salaries between the North and the South, but only after 1940. Collins argues that

some of the South's issues stemmed from deep agrarian traditions. Connolly (2004) suggests that this was exacerbated by racial discrimination, which led to lower allocation of schools to both blacks and whites. In the period under study, schooling for white and black children was conducted in completely segregated and highly unequal systems (Margo 1990, Collins and Margo 2006, Baker 2015). The Rosenwald Fund offered matching funds to local school districts specifically to support black schools, and this benefited some black children (Aaronson and Mazumder 2011).

This paper extends the current literature by using an external shock to high school access to assess the impact of high school on the occupational choice of rural children. My work uses a similar methodology to Duflo (2001). The Indonesian school construction program Duflo studies is similar in spirit to Tennessee's experience of forcing counties to raise funds specifically for high school construction and maintenance. She takes advantage of this program to measure the impact of access to primary school on wages. Duflo employs a two stage procedure where the first stage finds the impact of school access on years of education and the second stage calculates the return to an extra year of schooling.

I ask a slightly different question due in part to data limitations. I do not have years of schooling or wages in my dataset so I cannot exactly follow Duflo's approach. Instead of finding the financial impact of an additional year of schooling, I consider the impact of high school attendance on occupational choice. To this end, I use a one-stage difference-in-differences model which makes occupational choice a function of high school access. Similar to Duflo, however, I use a law that differentially affected children's access to high schools to identify the impact of school access on my variable of interest.

7.3. Empirical Strategy

I employ a difference-in-differences approach to measure the impact of school access on rural children. This takes advantage of the 1917 law described in Chapter 2 (see Figure 9). My identification strategy exploits differential tax rates between counties prior to 1917. Of Tennessee's 96 counties in 1916, 29 had no high school tax in 1916. The 1917 law was an exogenous shock to school access for rural children in the counties with no tax. These children form the treatment group in this paper. The control group consists of rural children in counties that already had high school taxation in place in 1917. The children are further separated into groups who turned 14 just before and just after 1917. I consider the group who turned 14 between 1908 and 1914 to be the "just before" group since a child who turned 14 in 1914 would be 17 when the law was passed and unlikely to receive much benefit from the addition of new high schools. Because it is unclear how quickly the law impacted school availability in the immediate aftermath, I choose the 7-year cohort from 1918 to 1924 as the "after" cohort. Duflo has a group in the middle that is partly affected by the building program. It is unclear that a new elementary school for a 10-year old child would have the same impact as a graded high school would on a 16-year old. Therefore, I choose to simply omit the children who turned 14 between 1915 and 1917, who would have been in high school when the law was passed.

One issue raised by my chosen birth cohorts is that not all of the counties that had a high school tax in place in 1917 had a tax in place for all of the children in my sample. To control for this, I went back to find when the first tax had been established for each county and dropped children who turned 14 before the tax when into effect.[67] This eliminated 72 observations. There were also three counties that had no tax in 1916 but did have taxes in earlier periods. These

[67] There is no tax information for 1909 or 1911. This means that if a county instituted a tax in 1911, my dataset would consider the tax as having started in 1912.

counties were dropped from the sample. This eliminated 56 observations. Finally, Overton County showed a 15 cents on the one hundred dollars of property tax for 1910 but had no other taxes at any point before the 1917 law nor did it open any high schools prior to the law. Therefore, I include it with the "no tax" group.

I consider the impact of high school access on the occupational choices of children living in rural areas. I test whether increased access to high schools, measured as a dummy for whether a high school tax was in place prior to 1917, is associated with differential changes in the variables of interest. There are two variables of interest: likelihood of being a farmer in 1930 and likelihood of working in a white collar occupation in 1930. I choose the variables because, in the absence of income information, the jobs people were doing provide some insight into the returns to high school and shed some light on how the high school movement affected the South.

Difference-in-differences analysis rests on the assumption that trends in the two types of counties would have been similar had it not been for the outside intervention. My analysis is somewhat different than traditional difference-in-differences because traditional analysis assumes that the two groups are trending similarly before the intervention and one of the groups diverges afterward. The situation in Tennessee was the reverse, where the counties' school environments are different before the legal intervention and more similar afterwards. I test the validity of this assumption in two steps. First, I compare the observables in both sets of counties and find that the counties looked very similar to one another. Then, I test that trends in county employment are similar when the legal environment is the same in both sets of counties and find that the trends did tend to converge after the 1917 law.

Table 20 contains descriptive statistics of the two groups of counties.[68] There is, notably, a difference in the share of farmers in both sets of counties, with the counties without a high school tax tending to be more agricultural. However, the counties appear very similar across all of the other variables. Farms tend to be a little bigger in counties without a tax in 1916. This is driven by a slight increase in mid-level farms and not by the larger plantation-style farms (500+ acres). The difference in mid-level farms may be explained by racial demographics. The counties without a tax also had fewer blacks on average and blacks would have operated smaller farms.

The third column contains the p-value associated with a two-sided t-test of the hypothesis that both sets of counties are equal. The p-values confirm there is not a statistically significant difference between any of the variables except the share of people working as farmers. Because of this, I directly control for father's occupation in my regressions. While the counties without a tax prior to 1916 had more farms, individual farms appear to have been quite similar in the two sets of counties. Since the agriculture sector is similar between the two sets of counties, it seems likely that the childhood experience of rural children would have been similar between counties. It does, however, suggest that controlling for whether or not a child lived on a farm may be important to get unbiased results.

I now turn to the observables on the children in the sample. I focus on their father's occupations, their parents' literacy, and elementary school attendance between the two groups. These figures are contained in Table 21. Parent's literacy is a rough proxy for a parent's schooling since some base level of education is required to read and write. As mentioned earlier in another context, literacy is not ideal for my purposes as it does not differentiate between

[68] Table 20 excludes the four counties with major urban centers as the presence of a major urban center may have undue influence on the results.

higher levels of schooling (e.g., primary school graduate or high school graduate). However, as years of schooling was not included in the census until 1940, literacy is the best measure of parental education available. Columns 1 and 2 in Table 21 suggest that both parents had similar levels of education and the p-value confirms that the difference in parental education between the two sets of counties are not statistically significant.

Columns 3 and 4 of Table 21 show that there is a statistically significant difference in the share of fathers working as farmers and the share of fathers working in white collar employment. This is not surprising as the counties without a tax had more farmers in general but it could bias the estimates since father's occupation is likely to impact son's occupation. I add dummies for father's occupation (farmer and white collar worker) to control for this issue.

Finally, column 5 looks at school attendance by primary school aged children in the two groups in 1910. Because the sample is further restricted to children between 6 and 12 in the 1910 census, it further reduces the sample to 102 observations in the counties with a tax and 50 observations in counties without a tax. While there is a 9 percentage point difference in the means, this may be because of the small sample size. The p-value shows that the difference is not statistically significant at the 80% confidence level.

The analysis also requires that the two sets of counties not exhibit significantly different economic trends when they are under the same legal regime. Since my difference-in-differences relies on an identification strategy that starts with counties being different prior to the law and similar afterwards, I hope to see similar trends in employment rates between the two sets of counties after 1917. I'm more concerned about underlying economic factors that might cause differences in occupational distribution. However, I'm only looking at occupational distribution in cross-section, I don't know anything about where the farmers or white collar workers were

born and educated. Therefore, any changes in occupational distribution after the law was passed is a weak indicator of the impact of high school access.

Figure 25 examines employment trends among Tennessee men between 1900 and 1940. Counties with a high school tax experienced a slight increase in the share of farmers between 1900 and 1910 and then a decrease between 1910 and 1920 while counties without a high school tax did the opposite. However, after 1920 both sets of counties saw declines and were nearly identical in 1930 and 1950. Skilled jobs show a similar pattern though not as strong. The rates trend upwards in both sets of counties throughout the period under study, even though they grow at different rates.

I employ the difference-in-differences method described in Angrist and Pischke (2008) using two regressions. First I estimate this equation:

$$Y_{irc} = \alpha + \beta_1 post_r affected_c + \beta_2 post_r + \beta_3 affected_c + \Gamma X_i + \varepsilon_{irc}$$

where Y_{irc} is the variable of interest for person i of cohort r in county c. $Post_r$ takes the value of 1 for cohorts who turned fourteen after 1917 and $affected_c$ is equal to 1 for counties with no tax prior to 1917. β_1 is my coefficient of interest and gives the impact of the increased access to schooling on the outcomes. X_i is a set of individual-specific controls, including father's occupation, father's occscore and parents' literacy.[69] ε_{irc} is a stochastic error term. All standard errors are clustered at the county level.

I run a second set of regressions which include more detailed county and birth cohort controls. This specification offers more specificity in estimating the impact on individual birth

[69] Occscore, taken from Ruggles et al. (2010) uses occupation to estimate a rough income level for the individual. The levels are assigned based on the median annual income for the given occupation in 1950's census.

cohorts. However, the estimates tend to be made with larger standard errors because there are relatively few observations in each cohort.

$$Y_{irc} = \alpha + \beta_1 post_r affected_c + \sum_r \beta_r I_r + \sum_c \beta_c I_c + \Gamma X_i + \varepsilon_{irc}$$

where I_r and I_c are indicator variables for cohort and county, respectively, and the rest of the variables are as described above.

In both formulations, β_1 represents an "intent to treat." Essentially, my results will measure the impact of high school education multiplied by the increased likelihood of an affected student attending high school. This is the best that I can do with the current dataset since the 1930 census does not include information on years of schooling implying that I cannot directly see the "first-stage" linking "access" to "attendance." However, the next section provides evidence that this connection did exist. While the 1930 census does include information on literacy rates, this information is not well suited to distinguishing between elementary education and secondary education.

β_1 may also capture general equilibrium effects. As discussed above, agriculture in Tennessee employed a large proportion of the population. It is conceivable that changes in productivity within one group of the population could impact the labor market decisions of another group. Prior to the equalization experienced under the 1917 law, only rural children living in counties with high schools would learn scientific farming methods. These new farming techniques would increase returns to farming for rural children in counties with high schools relative to rural children in counties without a high school. If the returns to farming are greater than higher returns to non-agricultural jobs as a result of attending high school, children in counties with high schools would opt to become farmers. As a result, rural children in counties without high schools would fall behind in labor productivity. In a general equilibrium

framework, this would increase the number of children who choose to become farmers in counties with a high school as those children have a relative advantage in farming. This would lead children in counties without a high school to opt for other pursuits where they have a relative advantge, such as unskilled work or jobs with significant on-the-job training. After high schools were opened in all counties following the 1917 law, potential young farmers would be on a more equal footing, leading to fewer students becoming farmers in counties that had a high school prior to 1917 and more children becoming farmers in counties without a high school prior to 1917.

7.4. Impact of the 1917 Law on School Attendance

While I cannot directly the test the impact of the 1917 law on years of schooling completed, I can test whether it had a differential impact on enrollment rates in the affected counties. Figure 26 shows white enrollment rates for counties with a tax prior to the new law and counties without a tax prior to the new law. There is a visible increase in enrollment rates for counties without a tax after the law is put in place while the enrollment rates in counties with a tax prior to 1917 are fairly static. This resulted in a gap of 13.1 percentage points in enrollment rates in 1917 dropping to 4.8 percentage points by 1920.

While there appears to be a slight jump in enrollment rates for counties without a tax in 1917, this is an artifact of the data. Several large counties instituted a high school tax in 1915. As these schools slowly enrolled more students, the total enrollment rate for all counties dropped. However, the enrollment rates for existing counties remained similar throughout. The 1919 drop in enrollment among counties with a pre-existing tax is harder to explain. I believe it is likely improperly recorded data in the original report.

One striking feature of this analysis is that enrollment rates for all schools are relatively low throughout the period, with the maximum enrollment rate about 15 percent. This suggests that spillover effects, such as those identified by Parman (2012), may be an important factor in any identified intent to treat effects of high school access in Tennessee.

The movement of children from ungraded schools or private schools to public high schools is not captured in Figure 26. If children are moving from ungraded schools (i.e., a 15-year old spending his 9th year in school in a one-room, rural school house), it is fair to assume that the quality of instruction increases in public high school and, therefore, may have an impact on outcomes. However, to the extent that the increase in enrollment at public high schools was coming from students who otherwise would have attended private high schools, this could overestimate the impact of the 1917 law. However, the importance authorities at the time put on spreading public schools to rural districts suggests that these institutions were not widespread.

7.5. Results

I start by comparing the proportion of children in rural areas, separated by pre- and post-law and affected and unaffected counties, who became farmers and white collar workers. The comparison, found in Table 22, suggests that in raw proportions the share of men who elected to become farmers in unaffected counties dropped for children who turned 14 after 1917 whereas the share was unchanged in affected counties. Further, a p-value of 0.818 prior to the law being passed indicates no statistical difference between children in the two types of counties prior to 1917, however children who turned 14 after the law was passed tended to become farmers more in affected counties and a p-value of .056 indicates a difference statistically significant at the 90% confidence level. This is very different than trends in Tennessee from the state census – which also include migrants – that were discussed above. The results for white collar workers

are even more striking. There is no statistically significant difference between children who turned 14 in the two sets of counties before or after 1917, which is in keeping with the results above. However, fewer children who turned 14 after 1917 entered white collar work than prior to 1917. This is the opposite of what statewide trends say should happen. I will continue with my analysis but the inconsistent results reinforce the need to expand the dataset in the future to understand if this is an artifact of a small dataset or a real phenomenon that needs to be understood.

Now I turn to the results of the linear probability models. Table 23 has the results of the "simple" difference-in-difference regressions which has a dummy for post 1917 and a dummy for affected counties. Columns 1-3 look at the impact of school access on the share of students who became farmers as adults. Counties without a tax tended to have fewer children become farmers than counties with a tax but this, again, is not statistically significant. After the tax, the share of children in affected counties who chose to become farmers increased by between 10.1 and 11.8 percentage points. In my preferred specification in column 3, this increase is significant at the 80 percent confidence level.[70]

The "simple" results for white collar work are in Columns 4-6 of Table 23. Here the impact of high school access is a decrease in white collar workers of between 7.9 and 9.4 percentage points. As was mentioned above, this seems to be coming about because white collar workers dropped faster in counties with a high school tax in 1917 than in counties without a high school tax. This is inconsistent with the 1930 census sample. It is possible that white collar jobs were being given to urban children in later years.[71]

[70] The results consider farm laborers to be farmers. The results if I consider only farm owner/operators increase by about 6 percentage points.
[71] Unreported results suggest that high school access had no impact on either interstate or intrastate migration.

Table 24 has the results of the complete difference-in-difference regressions with full controls for county and cohort of birth. Column 1-3 looks at the impact of school access on the share of students who became farmers as adults. The birth cohorts have the expected sign as earlier cohorts were more likely to be farmers than later cohorts. This is consistent with a slow movement out of farming in Tennessee. The coefficient on school access is positive, implying that the introduction of a high school encouraged more rural children to become farmers as adults. The outcome is fairly large, implying a 15.8 percentage point increase in the number of farmers after the law was passed.

Adding in controls for father's occscore and parents' literacy increases the impact of school openings and also increases their statistical significance. However, the magnitudes are still comparable. I believe a larger dataset could yield more significant results. The numbers are higher than for the "simple" calculation but, with the large standard errors, are not statistically different.

As in the "simple" regression, the increase in the number of farmers appears to have led to fewer children entering white collar work, as can be seen in Columns 4-6 of Table 24. The law is associated with a 7.6 percentage point decrease in the share of skilled workers for affected children after the law went into effect, although the high standard error makes it difficult to draw too much from the point estimate. It is, however, similar in magnitude to the "simple" regression. Similar to farmers, adding in controls increased the magnitude of the effect somewhat. The cohort results suggest that younger men were less likely to be involved in white collar work than older men.

The numbers given above are high, especially given the enrollment rates. However, the small dataset also yields large standard errors. I believe that the small sample yielded high

numbers for both farmers and skilled workers. Since occupations are mutually exclusive, drawing a high number of farmers implies that you will draw fewer skilled employees. However, while the magnitudes may be high, the standard error also implies that the signs are in the right direction. I leave it to future research to expand the data set and find more exact point estimates.

7.6. Discussion and Conclusion

This chapter used a linked data set to trace the impact of increased access to schooling on the job choices of Tennessee youth. The raw numbers suggest that increased access to high school increased farming and decreased movement into skilled labor. Since one of the high school's stated goals was to teach children to be better farmers, this is a reasonable outcome if the courses in agricultural science increased the production function of some young farmers more than it increased the value of alternative employment.

Interpreting the mechanisms for this result is not straightforward. One interpretation is that, far from urbanizing Tennessee, the county high schools taught farming methods which led more Tennesseans to be farmers than would have been the case if the high schools were never built. If this is the case, high schools actively discouraged the movement out of agriculture and into skilled work.

General equilibrium may impact these numbers. Figure 25 shows that the number of farmers in both sets of counties shrinks after 1920. It is conceivable that the share of people working as farmers in counties with a tax in 1917 would have been higher had the other counties not opened up high schools as a result of the 1917 law. On the other hand, the large increase in the share of farmers without a high school between 1910 and 1920 suggests that any general equilibrium story is complex.

I cannot distinguish between the two theories with the present data set. Indeed, a different dataset would significantly improve the results of this paper. I plan to create a linked 1910-1940 census dataset using the full 1910 census as a starting point. This should have two benefits for my analysis: (1) it will increase the number of observations in my data; (2) it will allow me to separate the effect of new high schools into increases in high school attendence and the impact of high school attendance.

Nevertheless, this chapter suggests that southern high schools in rural areas did increase the returns to farming. The effect seems to have been relatively large. Future work will determine whether these changes retarded the transition of the South out of agriculture by encouraging children to remain farmers or encouraged the process by increasing productivity and reducing the number of people required to remain in agriculture.

REFERENCES

Public Records

State of Florida, *Biennial Report of the Superintendent of Public Instruction*, 1918-1920

State of Mississippi, *Biennial Report and Recommendations of the State Superintendent of Public Education*, 1916-1917, 1924-1925.

State of Tennessee, *Public Acts of Tennessee*, 1885, 1891, 1899, 1909.

State of Tennessee, *Report of the State Superintendent of Public Instruction*, 1885-1935

United States. Bureau of the Census. *Historical Statistics of the United States, Colonial Times to 1970*. No. 93. US Department of Commerce, Bureau of the Census, 1975.

Newspapers

The Nashville Tennessean, 1816 to 1922, Proquest Historical Newspapers

Journal Articles and Books

Aaronson, Daniel, and Bhashkar Mazumder. "The Impact of Rosenwald Schools on Black Achievement." *Journal of Political Economy 119, no. 5* (2011): 821-888.

Altonji, Joseph G., and Rebecca M. Blank. "Race and Gender in the Labor Market." *Handbook of Labor Economics 3* (1999): 3143-3259.

Angrist, Joshua D., and Alan B. Krueger. "Does Compulsory School Attendance Affect Schooling and Earnings?." *Quarterly Journal of Economics 106, no. 4* (1991): 979-1014.

Angrist, Joshua D., and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2008.

Arrow, Kenneth. "The Theory of Discrimination." In Orley Ashenfelter and Albert Rees, editors, *Discrimination in Labor Markets*. Princeton University Press, 1973: 3-33.

Askildsen, Jan Erik, Badi H. Baltagi, and Tor Helge Holmås. "Wage Policy in the Health Care Sector: A Panel Data Analysis of Nurses' Labour Supply." *Health Economics 12, no. 9* (2003): 705-719.

Atack, Jeremy. "On the Use of Geographic Information Systems in Economic History: The American Transportation Revolution Revisited." *Journal of Economic History 73, no. 2* (2013): 313-338.

Attanasio, Orazio, Hamish Low, and Virginia Sanchez-Marcos. "Explaining Changes in Female Labor Supply in a Life-Cycle Model." *American Economic Review 98, no. 4* (2008): 1517-1552.

Bailey, Martha J. "More Power to the Pill: the Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply." *Quarterly Journal of Economics 121, no. 1* (2006): 289-320.

Baker, Richard B. "From the Field to the Classroom: The Boll Weevil's Impact on Education in Rural Georgia." *Journal of Economic History 75, no. 4* (2015): 1128-1160.

Becker, Gary S. *The Economics of Discrimination*. Second Edition. University of Chicago Press, 1971.

Bergmann, Barbara R. "Occupational Segregation, Wages and Profits when Employers Discriminate by Race or Sex." *Eastern Economic Journal 1, no. 2* (1974): 103-110.

Blau, Francine D., and Lawrence M. Kahn. "Gender Differences in Pay," *Journal of Economic Perspectives 14, no. 4* (Fall 2000): 75-99.

Blau, Francine D., and Lawrence M. Kahn. "Rising Wage Inequality and the US Gender Gap." *American Economic Review 84, no. 2* (1994): 23-28.

Bleakley, Hoyt. "Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics 122, no. 1* (2007): 73.

Blau, Francine D., and Lawrence M. Kahn. "The Gender Wage Gap: Extent, Trends, and Explanations." No. w21913. National Bureau of Economic Research, 2016.

Blinder, Alan S. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources 8, no. 4* (1973): 436-455.

Boal, William M. "Testing for Employer Monopsony in Turn-of-the-Century Coal Mining." *The RAND Journal of Economics 26 no. 3* (1995): 519-536.

Boal, William and Michael Ransom. "Monopsony in American Labor Markets". EH.Net Encyclopedia, edited by Robert Whaples. January 23, 2002. URL http://eh.net/encyclopedia/monopsony-in-american-labor-markets/

Boustan, Leah Platt, and William J. Collins. The Origins and Persistence of Black-White Differences in Women's Labor Force Participation. No. w19040. National Bureau of Economic Research, 2013.

Bowlus, Audra J. "A Search Interpretation of Male-Female Wage Differentials." *Journal of Labor Economics 15, no. 4* (1997): 625-657.

Bunting, Robert L. *Employer Concentration in Local Labor Markets*. University of North Carolina Press, 1962.

Burdett, Kenneth, and Dale T. Mortensen. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review 39 no. 2* (1998): 257-273.

Card, David. "Earnings, Schooling, and Ability Revisited." In Solomon Polachek, editor, *Research in Labor Economics*, vol. 14, Greenwich Connecticut: JAI Press, 1995.

Card, David. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica 69, no. 5* (2001): 1127-1160.

Card, David. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In L.N. Christofides, E.K. Grant, and R. Swidinsky, editors, *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp* Toronto: University of Toronto Press, 1995.

Caselli, Francesco, and Wilbur John Coleman II. "The US Structural Transformation and Regional Convergence: A Reinterpretation." *Journal of Political Economy 109.3* (2001): 584-616.

Collins, William J. "Education, Migration, and Regional Wage convergence in U.S. History" In Hatton, Timothy J., Kevin H. O'Rourke, and Alan M. Taylor, editors, *The New Comparative Economic History: Essays in Honor of Jeffrey G. Williamson*. MIT Press, 2007.

Collins, William J., and Robert A. Margo. "Historical Perspectives on Racial Differences in Schooling in the United States." *Handbook of the Economics of Education 1* (2006): 107-154.

Collins, William J., and Marianne H. Wanamaker. "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data." *American Economic Journal: Applied Economics, 6, no. 1* (2014): 220-252.

Connelly, Rachel. "The Effect of Child Care Costs on Married Women's Labor Force Participation." *The Review of Economics and Statistics 4, no. 1* (1992): 83-90.

Connolly, Michelle. "Human Capital and Growth in the Postbellum South: A Separate but Unequal Story." *Journal of Economic History 64, no. 2* (2004): 363-399.

Corlew, Robert E. *Tennessee: A Short History*. Second Edition. The University of Tennessee Press, 1989.

Daymont, Thomas N., and Paul J. Andrisani. "Job Preferences, College Major, and the Gender Gap in Earnings." *Journal of Human Resources (1984)*: 408-428.

DeLeire, Thomas, and Helen Levy. "Worker Sorting and the Risk of Death on the Job." *Journal of Labor Economics 22, no. 4* (2004): 925-953.

DiNardo, John. "Propensity Score Reweighting and Changes in Wage Distributions." University of Michigan, mimeograph (2002).

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica 64, no. 5* (1996): 1001-1064.

Dohmen, Thomas, and Armin Falk. "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender." *American Economic Review 101, no. 2* (2011): 556-590.

Donohue III, John J., James J. Heckman, and Petra E. Todd. "The Schooling of Southern Blacks: The Roles of Legal Activism and Private Philanthropy, 1910-1960." *Quarterly Journal of Economics 117, no. 1* (2002): 225-268.

Duflo, Esther. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review 91, no. 4* (2001): 795-813.

Fishback, Price V. "The Economics of Company Housing: Historical Perspectives from the Coal Fields." *Journal of Law, Economics, and Organization 8, no. 2* (1992): 346-365.

Fogli, Alessandra, and Laura Veldkamp. "Nature or Nurture? Learning and the Geography of Female Labor Force Participation." *Econometrica 79, no. 4* (2011): 1103-1138.

Gerhart, Barry. "Gender Differences in Current and Starting Salaries: The Role of Performance, College Major, and Job Title." *Industrial & Labor Relations Review 43, no. 4* (1990): 418-433.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics 118, no. 3* (2003): 1049-1074.

Goldin, Claudia. "America's Graduation From High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century." *Journal of Economic History 58, no. 02* (1998): 345-374.

Goldin, Claudia. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review 104, no. 4* (2014 (1)): 1091-1119.

Goldin, Claudia. "Marriage Bars: Discrimination Against Married Women Workers, 1920's to 1950's." NBER Working Paper w2747 (1988).

Goldin, Claudia. "The Meaning of College in the Lives of American Women: The Past One-Hundred Years." No. w4099. National Bureau of Economic Research, 1992.

Goldin, Claudia. "A Pollution Theory of Discrimination: Male and Female Differences in Occupations and Earnings." In L.P. Boustan, C. Frydman, and R.A. Margo, editors, *Human Capital and History*. Chicago, IL: University of Chicago Press, pp. 313-348. 2014 (2).

Goldin, Claudia. *Understanding the Gender Gap: An Economic History of American Women*. Oxford University Press. 1990.

Goldin, Claudia, and Lawrence F. Katz. "The Origins of State-Level Differences in the Public Provision of Higher Education: 1890-1940." *American Economic Review 88, no. 2* (1998): 303-308.

Goldin, Claudia, and Lawrence F. Katz. The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. No. w7527. National Bureau of Economic Research, 2000.

Goldin, Claudia, and Lawrence F. Katz. *The Race Between Education and Technology*. Harvard University Press, 2009.

Goldin, Claudia, and Cecilia Rouse. "Orchestrating Impartiality: The Impact of" Blind" Auditions on Female Musicians." *American Economic Review 90, no. 4* (2000): 715-741.

Goldin, Claudia, and Kenneth Sokoloff. "Women, Children, and Industrialization in the Early Republic: Evidence From the Manufacturing Censuses." *Journal of Economic History 42, no. 4* (1982): 741-774.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. "Engines of Liberation." *The Review of Economic Studies 72, no. 1* (2005): 109-133.

Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. http://doi.org/10.3886/ICPSR02896.v3

Haines, Michael, Price Fishback, and Paul Rhode. United States Agriculture Data, 1840 - 2010. ICPSR35206-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-12-22. http://doi.org/10.3886/ICPSR35206.v1

Hanushek, Eric A. "Efficient Estimators For Regressing Regression Coefficients." *The American Statistician 28, no. 2* (1974): 66-67.

Hirsch, Barry T., and Edward J. Schumacher. "Classic or New Monopsony? Searching for Evidence in Nursing Labor Markets." *Journal of Health Economics 24, no. 5* (2005): 969-989.

Hirsch, Barry T., and Edward J. Schumacher. "Monopsony Power and Relative Wages in the Labor Market for Nurses." *Journal of Health Economics 14, no. 4* (1995): 443-476.

Hoffschwelle, Mary S. *Rebuilding the Rural Southern Community: Reformers, Schools, and Homes in Tennessee, 1900-1930*. Univ. of Tennessee Press, 1998.

Hoffschwelle, Mary S. "Public Education in Tennessee." *Trials and Triumphs: Tennesseans' Search for Citizenship, Community, and Opportunity.* Middle Tennessee State University, 2014. Web.

Holt, Andrew David. *The Struggle for a State System of Public Education in Tennessee, 1903-1936*. New York: Teachers College, Columbia University Bureau of Publications, 1938.

Hirsch, Barry T., and Edward J. Schumacher. "Monopsony Power and Relative Wages in the Labor Market for Nurses." *Journal of Health Economics 14, no. 4* (1995): 443-476.

Jann, Ben. "The Blinder-Oaxaca Decomposition for Linear Regression Models." The Stata Journal 8, no. 4 (2008): 453-479.

Jones, Larry E., Rodolfo E. Manuelli, and Ellen R. McGrattan. "Why Are Married Women Working So Much?" *Journal of Demographic Economics 81, no. 1* (2015): 75-114.

Juhn, Chinhui, and Kevin M. Murphy. "Wage Inequality and Family Labor Supply." *Journal of Labor Economics 15, no. 1* (1997): 72-97.

Krueger, Alan, and David Card. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy 100, no. 1* (1990): 1-40

Manning, Alan. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press, 2003.

Margo, Robert A. 1988. "Schooling and the Great Migration." NBER Working Paper 2697.

Margo, Robert A. *Race and Schooling in the South, 1880-1950: An Economic History*. University of Chicago Press, 1990.

Mincer, Jacob A. "Age and Experience Profiles of Earnings." In *Schooling, Experience, and Earnings*, pp. 64-82. NBER, 1974.

Miller, Amalia R. "The Effects of Motherhood Timing on Career Path." *Journal of Population Economics 24, no. 3* (2011): 1071-1100.

Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, MN: University of Minnesota 2011. http://www.nhgis.org

Morgan, John G. "The Education Improvement Act: A Progress Report." Nashville, TN: Tennessee Comptroller of the Treasury, Office of Research and Educational Accountability (2004).

Mukherjee, Debasri. "Monopsony'in the Market for Nurses? A Semiparametric Note." *Regional and Sectoral Economic Studies 11, no. 2* (2011).

Neumark, David, Roy J. Bank, and Kyle D. Van Nort. "Sex Discrimination in Restaurant Hiring: An Audit Study." *Quarterly Journal of Economics 101, no. 3* (1996): 915-941.

Oaxaca, Ronald. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review 14, no. 3* (1973): 693-709.

Parman, John. "Good Schools Make Good Neighbors: Human Capital Spillovers in Early 20th Century Agriculture." *Explorations in Economic History 49, no. 3* (2012): 316-334.

Perlmann, Joel, and Robert A. Margo. *Women's Work? American Schoolteachers, 1650-1920*. University of Chicago Press, 2001.

Ramcharan, Mr Rodney. Migration and Human Capital Formation: Theory and Evidence from the US High School Movement. No. 2-123. International Monetary Fund, 2002.

Ransom, Michael R., and Val E. Lambson. "Monopsony, Mobility, and Sex Differences in Pay: Missouri School Teachers." *American Economic Review 101, no. 3* (2011): 454-459.

Ransom, Michael R., and Ronald L. Oaxaca. "New Market Power Models and Sex Differences in Pay." *Journal of Labor Economics 28, no. 2* (2010): 267-289.

Ransom, Michael R., and David P. Sims. "Estimating the Firm's Labor Supply Curve in a "New Monopsony" Framework: Schoolteachers in Missouri." *Journal of Labor Economics 28, no. 2* (2010): 331-355.

Reber, Sarah J. "School Desegregation and Educational Attainment for Blacks." *Journal of Human Resources 45, no. 4* (2010): 893-914.

Robinson, Joan, *The Economics of Imperfect Competition*, MacMillan and Company, 1933

Rottenberg, Simon. "The Baseball Players' Labor Market." *Journal of Political Economy 64, no. 3* (1956): 242-258.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

Scully, Gerald W. "Pay and Performance in Major League Baseball." *American Economic Review 64, no. 6* (1974): 915-930.

Stephens, Melvin, and Dou-Yan Yang. "Compulsory Education and the Benefits of Schooling." *American Economic Review 104, no. 6* (2014): 1777-1792.

Summitt, James Euel, "Comparison of the Two-Year Approved High Schools in Tennessee." Unpublished Thesis. 1935. Peabody Library.

Tinbergen, Jan. "Substitution of Graduate by Other Labour." *Kyklos 27, no. 2* (1974): 217-226.

Tindall, George Brown. *The Emergence of the New South, 1913-1945*. Vol. 10. LSU Press, 1967.

Waldfogel, Jane. "Understanding the "Family Gap" in Pay for Women with Children." *Journal of Economic Perspectives 12, no. 1* (1998): 137-156.

Wanamaker, Marianne H. "Industrialization and Fertility in the Nineteenth Century: Evidence from South Carolina." *Journal of Economic History 72, no. 1* (2012): 168-196.

Webber, Douglas A. "Firm-Level Monopsony and the Gender Pay Gap." Unpublished Paper (2013).

Webber, Douglas A, "Firm Market Power and the Earnings Distribution," Unpublished Paper (2012).

White, Robert H. *Development of the Tennessee State Educational Organization, 1796-1929*. Kingsport, TN: Southern Publishers, 1929.

Woodward, C. Vann. *Origins of the New South, 1877--1913: A History of the South*. Vol. 9. LSU Press, 1951.

Wright, Gavin. *Old South, New South. Revolutions in the Southern Economy Since the Civil War*, Basic Books, Inc., New York (1986).

Yun, Myeong-Su. "A Simple Solution to the Identification Problem in Detailed Wage Decompositions." *Economic Inquiry 43, no. 4* (2005): 766-772.

.

TABLES

TABLE 1: FIRST CLASS SCHOOLS AND PERCENT OF THE COUNTY WHICH IS RURAL, TENNESSEE 1917

County	Percent Rural	First Class Schools
Anderson	100.0	1
Bedford	86.6	0
Benton	100.0	0
Bledsoe	100.0	0
Blount	75.4	0
Bradley	65.0	0
Campbell	89.2	3
Cannon	100.0	0
Carroll	100.0	0
Carter	87.2	0
Cheatham	100.0	1
Chester	100.0	1
Claiborne	100.0	1
Clay	100.0	0
Cocke	86.8	0
Coffee	79.9	1
Crockett	100.0	0
Cumberland	100.0	1
Davidson	29.5	0
Decatur	100.0	0
Dekalb	100.0	0
Dickson	100.0	0
Dyer	78.5	1
Fayette	100.0	1
Fentress	100.0	0
Franklin	100.0	0
Gibson	84.6	0
Giles	91.0	1
Grainger	100.0	0
Greene	88.5	3
Grundy	72.6	0
Hamblen	61.0	0
Hamilton	40.0	6
Hancock	100.0	0
Hardeman	100.0	1
Hardin	100.0	0
Hawkins	100.0	0
Haywood	87.9	1
Henderson	100.0	1
Henry	82.6	1
Hickman	100.0	0
Houston	100.0	0
Humphreys	100.0	0
Jackson	100.0	0
Jefferson	100.0	0
Johnson	100.0	0
Knox	31.1	6
Lake	100.0	1
Lauderdale	100.0	1
Lawrence	100.0	1
Lewis	100.0	0
Lincoln	85.9	1
Loudon	74.1	1
McMinn	79.7	1
McNairy	100.0	0
Macon	100.0	0
Madison	57.0	1
Marion	100.0	1
Marshall	84.4	1
Maury	84.4	2
Meigs	100.0	0
Monroe	100.0	3
Montgomery	74.9	1
Moore	100.0	1
Morgan	100.0	0
Obion	84.5	0
Overton	100.0	0
Perry	100.0	0
Pickett	100.0	0
Polk	100.0	3
Putnam	100.0	1
Rhea	100.0	1
Roane	64.8	1
Robertson	84.9	0
Rutherford	83.8	1
Scott	100.0	0
Sequatchie	100.0	0
Sevier	100.0	0
Shelby	27.3	4
Smith	100.0	2
Stewart	100.0	0
Sullivan	62.1	0
Sumner	90.0	2
Tipton	88.7	1
Trousdale	100.0	0
Unicoi	70.7	0
Union	100.0	0
Van Buren	100.0	1
Warren	83.7	0
Washington	63.5	2
Wayne	100.0	0
Weakley	90.9	0
White	100.0	1
Williamson	86.7	1
Wilson	84.4	0

Note: Number of first class high schools comes from 1916 State Superintendent's Annual Report. Percent Rural is taken from Haines (2010).

TABLE 2, SCHOOLS AND DEMOGRAPHICS IN TENNESSEE, 1920, WHITES

	Total	Rural Counties	80% to 100% Rural Population	Less than 80% Rural Population	Conuties with Urban Centers	East	Middle	West
Percent of Population	100%	34%	23%	15%	28%	37%	37%	26%
Farmers	26.42%	37.65%	34.96%	24.56%	5.66%	22.35%	29.19%	28.50%
White Collar Workers	9.74%	4.76%	6.79%	9.45%	18.83%	10.13%	8.41%	11.00%
High Schools per Sq. Mile	0.014	0.010	0.023	0.017	0.018	0.013	0.014	0.016
HS Classrooms per Sq. Mile	0.067	0.049	0.080	0.096	0.149	0.078	0.057	0.069
Students per Sq. Mile	3.679	2.780	3.362	4.755	14.162	4.689	3.139	3.146
Students per High School	233.3	275.8	147.2	251.6	692.1	341.6	199.9	199.2
Students per HS Classroom	51.1	54.8	42.1	48.1	84.2	58.0	48.7	46.8
First Class High Schools	214	89	63	31	31	77	72	65
Second Class High Schools	110	32	56	16	6	21	50	39
Third Class High Schools	283	87	141	47	8	67	141	75
N	95	56	21	14	4	33	41	21

Sources: 1920 IPUMS 1% sample, Haines (2010), Annual Report of the State Superintendent of Public Schools, 1923, Prinicipal Reports found in the Tennessee State Library and Archives

Notes: Rural residents are defined in Haines (2010) with the rur920 variable. Urban Centers are those counties from Haines with at least one city with 25,000+ in population. Schools data is from 1923 as that is when the individual reports are available. Demographic data is from the 1920 census. White collar workers are those involved in professional, clerical, or managerial jobs. Demographic data is restricted to all people (of both genders) of working age, 21-65.

TABLE 3: SCHOOL COUNTS IN THE PRINCIPAL REPORTS AND THE SUPERINTENDENT'S ANNUAL REPORT, 1923

	School Level Reports				Annual Reports		
County	First Class	Second Class	Third Class	County	First Class	Second Class	Third Class
Anderson	4	0	4	Anderson	4	0	4
Bedford	3	1	7	Bedford	3	0	8
Benton	1	1	3	Benton	1	1	3
Bledsoe	1	0	1	Bledsoe	1	1	0
Blount	2	2	3	Blount	1	2	4
Bradley	2	2	4	Bradley	0	0	6
Campbell	3	3	3	Campbell	3	1	6
Cannon	1	0	0	Cannon	1	0	6
Carroll	5	2	1	Carroll	5	0	5
Carter	1	0	15	Carter	1	0	15
Cheatham	2	0	1	Cheatham	2	0	1
Chester	1	0	0	Chester	1	0	0
Claiborne	1	1	0	Claiborne	1	1	0
Clay	1	0	3	Clay	1	0	3
Cocke	1	0	0	Cocke	1	0	1
Coffee	2	0	3	Coffee	1	0	8
Crockett	4	1	3	Crockett	4	1	3
Cumberland	1	0	0	Cumberland	1	0	0
Davidson	6	0	4	Davidson	6	0	4
Decatur	2	1	0	DeKalb	2	1	2
Dekalb	2	1	1	Decatur	2	0	2
Dickson	1	1	1	Dickson	1	1	1
Dyer	3	5	3	Dyer	3	6	3
Fayette	2	4	4	Fayette	1	3	0
Fentress	0	0	3	Fentress	0	0	4
Franklin	3	0	0	Franklin	3	0	0
Gibson	4	5	7	Gibson	5	9	14
Giles	4	0	5	Giles	4	2	3
Grainger	1	0	3	Grainger	1	0	3
Greene	8	1	2	Greene	6	1	2
Grundy	1	0	1	Grundy	1	0	4
Hamblen	2	0	2	Hamblen	3	0	2
Hamilton	8	0	0	Hamilton	8	0	0
Hancock	2	0	0	Hancock	1	0	0
Hardeman	3	1	1	Hardeman	3	1	1
Hardin	2	1	4	Hardin	1	1	0
Hawkins	4	0	3	Hawkins	4	2	1
Haywood	1	0	0	Haywood	1	0	0
Henderson	1	0	3	Henderson	1	0	3
Henry	4	0	0	Henry	5	0	2
Hickman	1	1	5	Hickman	1	0	7
Houston	1	0	0	Houston	1	0	0
Humphreys	0	1	3	Humpreys	3	1	4

TABLE 3: SCHOOL COUNTS, PRINCIPAL REPORTS AND SUPERINTENDENT'S ANNUAL REPORT, 1923 (CONT.)

County	First Class	Second Class	Third Class	County	First Class	Second Class	Third Class
Jackson	1	1	2	Jackson	1	2	1
Jefferson	1	2	2	Jefferson	1	1	2
Johnson	1	0	1	Johnson	1	0	1
Knox	7	1	1	Knox	7	3	0
Lake	2	0	0	Lake	2	1	0
Lauderdale	2	2	1	Lauderdale	2	2	1
Lawrence	1	0	0	Lawrence	1	0	0
Lewis	1	0	0	Lewis	1	0	0
Lincoln	1	4	12	Lincoln	1	4	17
Loudon	2	1	2	Loudon	2	1	2
Macon	0	1	1	Macon	0	0	2
Madison	1	2	9	Madison	4	3	6
Marion	2	0	0	Marion	1	0	0
Marshall	1	0	0	Marshall	1	0	0
Maury	2	1	1	Maury	2	0	0
McMinn	2	1	3	McMinn	2	1	4
McNairy	3	3	2	NcNairy	3	2	4
Meigs	1	0	0	Meigs	1	0	1
Monroe	3	0	1	Monroe	3	0	1
Montgomery	1	2	4	Montgomery	1	0	5
Moore	1	0	0	Moore	1	0	0
Morgan	1	2	1	Morgan	1	2	2
Obion	5	0	6	Obion	5	2	8
Overton	1	1	0	Overton	0	0	0
Perry	0	1	2	Perry	0	1	2
Pickett	0	0	1	Pickett	0	0	1
Polk	3	0	0	Polk	3	0	0
Putnam	3	0	1	Putnam	4	0	1
Rhea	3	1	0	Rhea	2	0	2
Roane	3	0	2	Roane	4	0	1
Robertson	8	1	0	Robertson	6	2	8
Rutherford	4	1	7	Rutherford	4	4	8
Scott	3	1	1	Scott	3	2	5
Sequatchie	1	0	0	Sequatchie	1	0	0
Sevier	1	0	0	Sevier	1	0	0
Shelby	9	5	4	Shelby	0	0	0
Smith	3	0	5	Smith	2	2	2
Stewart	2	0	0	Stewart	2	0	2
Sullivan	3	1	0	Sullivan	3	1	0
Sumner	2	0	2	Sumner	2	10	12
Tipton	3	1	2	Tipton	2	3	1
Trousdale	1	0	0	Trousdale	1	0	0
Unicoi	1	0	0	Unicoi	1	0	0
Union				Union			
Van Buren	1	0	0	Van Buren	1	0	1
Warren	2	2	1	Warren	2	2	1
Washington	5	2	2	Washington	5	2	2
Wayne	2	0	0	Wayne	1	1	0
Weakley	5	0	2	Weakley	4	1	15
White	1	0	2	White	1	2	3
Williamson	1	2	7	Williamson	1	9	10
Wilson	6	1	4	Wilson	4	6	10

Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64..

Notes: Annual Report numbers are taken directly from State Superintendent's Annual Report. School level reports are first assigned a class by the Superintendent's classification system (see text) and the totaled by county.

TABLE 4: METHODS OF ASSIGNING GENDER

	Male	Female	Total
Total Names in the Dataset			1610
Names assigned given "Mr/Mrs/Miss"	16	362	378
Names assigned given "95 percent" rule	188	309	497
Names assigned given "Two intials"	515		515
Names assigned by checking transcription	31	59	90
Total Names Assigned a Gender	750	730	1480

TABLE 5: COLLEGES ATTENDED BY TENNESSEE HIGH SCHOOL TEACHERS, BY NUMBER OF TEACHERS

University of Tennessee	333	Hall Moody	19
Peabody College	283	University of Virginia	18
East Tennessee State Normal	154	Milligan College	16
Middle Tennessee State Normal	153	Tusculum College	16
West Tennessee State Normal	110	Athens School (Tennessee)	15
Other Normal	83	University of Georgia	13
Union University	83	Lincoln Memorial University	13
Maryville College	66	Southern Normal School and Business College	13
Vanderbilt University	58	University of Kentucky	12
Carson Newman College	52	Martha Washington College	12
Columbia University	41	Valparaiso	12
University of Chicago	40	Martin Methodist College	11
University of Chattanooga	33	Winchester Normal	10
Tennessee Polytechnic Institute	23		
Cumberland University	21		

Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

Notes: Taken from Individual School Reports. Every college with at least 10 alumni is shown.

TABLE 6: SUMMARY STATISTICS, TENNESSEE HIGH SCHOOL TEACHERS, 1923

	All	Men	Women	p-value on gender difference
Ln(Salary)	6.95	7.12	6.78	0.00
	(0.50)	(0.50)	(0.44)	
Salary	1156.71	1363.18	943.91	0.00
	(481.02)	(538.65)	(286.00)	
Experience at Elementary School	2.92	3.44	2.38	0.00
	(4.31)	(4.80)	(3.66)	
Experience at Another High School	3.00	4.03	1.93	0.00
	(4.78)	(5.58)	(3.49)	
Experience at Current School	2.64	2.83	2.45	0.02
	(2.96)	(3.23)	(2.64)	
College and Normal School	0.24	0.25	0.24	0.96
College	0.50	0.51	0.48	0.17
Normal School	0.19	0.16	0.22	0.01
High School	0.07	0.08	0.06	0.26
Principal	0.30	0.54	0.06	0.00
Courses Taught				
Math	0.34	0.49	0.18	0.00
Science	0.22	0.30	0.14	0.00
Social Science	0.07	0.08	0.05	0.02
Languages (including English)	0.49	0.40	0.57	0.00
Domestic Sciences	0.22	0.18	0.26	0.00
Education	0.05	0.08	0.02	0.00
History	0.30	0.33	0.27	0.02
N	1391	706	685	

Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

Notes: Principal is equal to one if the name of the principal on the form matched the teacher. Training (e.g., college, normal school, and high school) is equal one if years in the given school is greater than zero. High school represents teachers with a high school degree as a terminal degree. "College" and "Normal School" categories indicated teachers who attended only that type of school. "College and Normal School" indicate a teacher who attended both types of school. Experience is total years teaching. Tenure is years teaching at the current school. Courses taught dummies equal to 1 if the teacher taught the given course. As many teachers taught multiple subjects, courses taught will add up to more than 1. Domestic sciences are agricultural sciences and home economics.

TABLE 7: SUMMARY STATISTICS FOR SCHOOL PRINCIPALS, 1923 TO 1935

	N	5% Salary Cutoff	Mean Salary	95% Salary	Median Age
1923					
Male Principals	402	$720	$1,461	$2,400	35
Female Principals	43	$504	$1,130	$2,000	33
1926					
Male Principals	406	$882	$1,616	$2,750	35
Female Principals	38	$147	$1,093	$1,958	34
1928					
Male Principals	427	$975	$1,675	$2,700	35
Female Principals	38	$149	$1,165	$2,258	34
1932					
Male Principals	438	$876	$1,556	$2,653	34
Female Principals	36	$434	$1,169	$2,190	35
1935					
Male Principals	465	$630	$1,273	$2,200	35
Female Principals	24	$541	$999	$2,475	32

Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64

Notes: Figures are given in nominal dollars.

TABLE 8: WOMEN'S LABOR FORCE PARTICIPATION, 1900 TO 2000, U.S. AND TENNESSEE

		1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
All Women	U. S.	14.2%	17.5%	18.0%	19.9%	24.3%	30.1%	38.6%	46.5%	57.0%	68.5%	70.3%
	Tennessee	9.5%	13.3%	12.2%	15.2%	20.6%	25.8%	36.4%	46.0%	57.3%	67.2%	66.6%
Married Women	U. S.	3.2%	6.4%	6.4%	9.1%	14.0%	21.4%	32.0%	40.9%	52.0%	64.8%	67.1%
	Tennessee	1.5%	5.2%	4.8%	7.3%	13.1%	19.7%	31.8%	42.6%	53.8%	64.6%	64.0%
Single Women	U. S.	52.4%	60.1%	65.1%	68.0%	72.0%	74.3%	75.7%	74.9%	79.3%	83.3%	81.6%
	Tennessee	22.6%	38.4%	38.1%	52.3%	58.2%	57.1%	64.6%	61.4%	77.2%	79.9%	79.1%
College Educated Women	U. S.					39.9%	50.5%	47.5%	53.5%	67.3%	77.6%	78.0%
	Tennessee					40.6%	47.6%	48.8%	52.9%	67.5%	78.1%	76.7%
College Educated Married Women	U. S.					20.0%	31.9%	37.3%	45.3%	60.5%	73.0%	74.1%
	Tennessee					24.3%	34.0%	41.6%	46.7%	61.7%	74.4%	73.0%
College Educated Single Women	U. S.					85.5%	87.8%	88.6%	87.5%	89.7%	91.4%	89.3%
	Tennessee					75.1%	91.7%	86.3%	80.5%	89.4%	93.7%	89.3%

Source: Ruggles et al. (2010)

Notes: LFP is defined as the labforce variable being equal to 2. Married is from the marst variable, and includes women with husbands present and husbands absent. College educated women includes women with one or more years of college.

TABLE 9: GENDER GAP BY OCCUPATION, 1940

	White		Total	
	National	Tennessee	National	Tennessee
White Collar	-0.424	-0.474	-0.423	-0.465
	(0.004)	(0.038)	(0.004)	(0.038)
Skilled	-0.366	-0.419	-0.366	-0.415
	(0.004)	(0.034)	(0.004)	(0.034)
Unskilled	-0.447	-0.242	-0.516	-0.436
	(0.005)	(0.044)	(0.004)	(0.033)
Teachers	-0.252	-0.317	-0.254	-0.296
	(0.015)	(0.083)	(0.015)	(0.081)

Source: IPUMS 1940 Census 1% Sample

Notes: The results above come from 16 separate regressions. The sample consists of wage workers with 36 or more weeks worked the prior year, restricted to the given occupational group. The table displays the coefficient on a female dummy in a Mincerian wage regression with gender (female = 1), years of schooling, potential experience (age – years of schooling – 5), and potential experience squared on the right hand side. The coefficient on gender is reported above. White collar includes people working in professional, managerial, or clerical occupations. Skilled represents people working in white collar or craftsman jobs. Unskilled is every other occupation. Teachers is the group classified as Teachers (not elsewhere classified) in OCC1950.

TABLE 10: BASELINE EARNINGS REGRESSIONS FOR TENNESSEE HIGH SCHOOL TEACHERS, 1923

	(1)	(2)	(3)	(4)	(5)
Female	-0.34	-0.30	-0.32	-0.21	-0.19
	(0.03)	(0.03)	(0.04)	(0.04)	-0.04
Principal				0.13	0.29
				(0.04)	-0.03
High School Experience		0.03	0.03	0.03	0.03
		(0.01)	(0.01)	(0.01)	-0.01
HS Experience ^ 2		-0.98	-0.89	-0.78	-0.76
		(0.31)	(0.35)	(0.31)	-0.26
Current School		0.01	0.01	0.01	0.00
		(0.01)	(0.01)	(0.01)	-0.01
Current School ^ 2		-0.38	-0.48	-0.07	0.32
		(0.70)	(0.72)	(0.65)	-0.62
College and Normal School		0.23		0.19	0.13
		(0.03)		(0.04)	-0.04
College		0.22		0.19	0.14
		(0.03)		(0.06)	-0.06
High School		-0.03		-0.02	0.04
		(0.04)		(0.06)	-0.06
Constant	7.12	6.79	6.88	6.70	6.42
	(0.03)	0.04	0.04	0.07	0.08
Control for College/School Attended	N	N	Y	Y	Y
Control for Subjects Taught	N	N	N	Y	Y
Controls for Class of School	N	N	N	N	Y
Controls for County of Residence	N	N	N	N	Y
N	1391	1391	1391	1391	1391
Adjusted R-squared	0.114	0.197	0.200	0.256	0.347

Notes: Sample restricted to teachers whose gender could be imputed by name. Equation (1) simply calculates the raw wage gap. Equation (2) includes controls for teacher experience and training. Equation (3) adds controls for job specifics: principal status and subjects taught. Equation (4) includes a full set of control for the class of the school and dummies for the county of residence. Experience includes years of experience at all high schools. Subjects taught includes dummies for: math, science, social science, languages (including English), home economics, agricultural science, education, professional (e.g., typing, bookkeeping), history, and music. Training (e.g., college, normal school, and high school) is equal one if years in the given school is greater than zero. High school represents teachers with a high school degree as a terminal degree. "College" and "Normal School" categories indicated teachers who attended only that type of school. "College and Normal School" indicate a teacher who attended both types of school. Normal School is the omitted term. Experience squared and current school squared are divided by 1000 to produce meaningful output.

TABLE 11: WOMEN'S EMPLOYMENT IN SKILLED OCCUPATIONS, TENNESSEE, 1900 TO 1930, BY PERCENT EMPLOYED

Occupation	1900	1910	1920	1930
Accountants and auditors			0.95	0.43
Actors and actresses		1.13		0.21
Artists and art teachers	2.86	0.56		
Draftsmen	1.43			
Clergymen		0.56		
Dentists		1.13		
Subject not specified-Professors and in			0.63	0.64
Editors and reporters		0.56	0.63	
Designers				0.21
Farm and home management advisors			0.32	0.21
Librarians		0.56	0.32	0.21
Musicians and music teachers	5.71	9.04	2.86	1.93
Photographers	1.43			
Nurses, professional		3.39	3.49	3.43
Religious workers			0.63	0.21
Social and welfare workers, except grou			0.63	0.64
Professional, technical and kindred wor	1.43			
Teachers	34.29	42.37	28.25	32.83
Testing-technicians		0.56		
Medical and dental-technicians			0.32	0.21
Postmasters			0.63	
Buyers and dept heads, store		1.13		0.43
Managers and superintendants, building	5.71	4.52	3.49	0.21
Officials and administratators (nec), p				0.21
Postmasters	1.43			0.43
Managers, officials, and proprietors (n	14.29			6.44
Agents				0.21

Occupation	1900	1910	1920	1930
Attendants, physicians and dentists off			0.63	0.21
Bank tellers				0.43
Bookkeepers	1.43	3.95	9.52	9.44
Cashiers	5.71	1.69	2.22	1.07
Collectors, bill and account			0.32	
Messengers and office boys			1.27	
Office machine operators			0.32	0.43
Shipping and receiving clerks			0.32	
Stenographers, typists, and secretaries	11.43	13.56	25.4	22.96
Telegraph operators		0.56	0.63	0.43
Telephone operators	4.29	4.52	6.03	4.51
Ticket, station, and express agents	1.43		0.32	0.21
Clerical and kindred workers	2.86	3.95	6.35	6.44
Bakers		0.56		0.43
Bookbinders	1.43		0.63	0.64
Carpenters			0.32	0.21
Compositors and typesetters	1.43	2.82	0.32	0.21
Machinists		1.69		
Pressmen and plate printers		0.56		
Foremen			1.27	1.93
Inspectors				0.21
Jewelers, watchmakers, goldsmiths, and				0.21
Shoemakers and repairers, except factor				0.21
Stone cutters and stone carvers				0.21
Tailors and tailoresses	1.43	0.56	0.95	0.21
Upholsterers				0.21

Source: IPUMS 1% Sample, 1900 - 1930 United States Census

Notes: Skilled is defined as those working in a professional, clerical, managerial, or craftsman capacity. Occupation definitions are the IPUMS occ1950 variable. Percent is unweighted percentage of skilled women in the sample who did a particular job. The high percentage of Managers in 1900 is very difficult to explain. This consists of 10 different women. They are across multiple industries but most involve working with clients. Real estate employs 4 of them.

TABLE 12: TEACHERS IN THE LABOR FORCE, 1920

	Northeast	Midwest	South	West	Tennessee
Percent of American men Living in this region	31.39%	37.37%	20.93%	10.19%	1.76%
Percent of men in the labor force	95.28%	94.33%	94.92%	94.12%	95.07%
Percent in the labor force involved in skilled labor	46.16%	37.45%	31.62%	38.17%	27.73%
Percent of men who are Teachers	0.32%	0.41%	0.41%	0.51%	0.50%
Percent of men in the labor force who are teachers	0.34%	0.43%	0.43%	0.54%	0.53%
Percent of skilled men in the labor force who are teachers	0.73%	1.16%	1.36%	1.42%	1.90%
Percent of American women Living in this region	33.03%	36.67%	21.14%	9.11%	1.86%
Percent of women in the labor force	26.67%	19.17%	15.62%	21.14%	13.92%
Percent in the labor force involved in skilled labor	39.62%	46.25%	39.93%	49.11%	37.86%
Percent of women who are Teachers	1.98%	2.25%	1.96%	2.40%	1.65%
Percent of women in the labor force who are teachers	7.39%	11.64%	12.49%	11.23%	11.72%
Percent of skilled women in the labor force who are teachers	18.65%	25.17%	31.27%	22.88%	30.96%

Source: IPUMS 1% Sample, 1920 United States Census.

Notes: Skilled is defined as those working in a professional, clerical, managerial, or craftsman capacity. Labor force and gender are taken from the labforce and sex variables in the IPUMS dataset. The sample includes data from the 48 contiguous US states. West Virginia is considered to be a Midwestern state. The South is defined as the Confederate states, Kentucky, and Maryland.

TABLE 13: MARITAL AND RESIDENTIAL STATUS OF TEACHERS, WHITES, TENNESSEE, 1920

Percent in the labor force who work as teachers

	Men	Women
All	0.1%	11.5%
Skilled	2.0%	28.9%

Percent of white teachers who are also single

	Men	Women
	26.8%	81.9%

Percent of single whites (25 to 65) who live with relatives

	Men	Women
All Tennessee	65.1%	86.0%
Rural Tennessee	72.8%	93.1%

Source: IPUMS 1% Sample, 1920 United States Census.

Notes: Figures omit household service workers. "Single" includes never married, divorced, and widowed.

TABLE 14: GENDER GAP AND MARKET POWER

	(1)	(2)	(3)	(4)	(5)
Female * Market Power	-0.282	-0.147	-0.152	-0.196	-0.214
	(0.080)	(0.087)	(0.073)	(0.075)	(0.069)
Female	-0.176	-0.242	-0.126	-0.104	-0.070
	(0.051)	(0.057)	(0.056)	(0.062)	(0.057)
Market Power	-0.084	-0.071	-0.058	0.040	
	(0.085)	(0.087)	(0.078)	(0.066)	
Controls for Experience/Tenure	N	Y	Y	Y	Y
Controls for College/School Attended	N	Y	Y	Y	Y
Controls for Subjects Taught	N	N	Y	Y	Y
Controls for Principal Status	N	N	Y	Y	Y
Controls for Class of School	N	N	N	Y	Y
Controls for County Characteristics	N	N	N	Y	N
Controls for County Fixed Effects	N	N	N	N	Y
N	1391	1391	1391	1381	1391
Adjusted R-square	0.130	0.225	0.261	0.324	0.349

Notes: Market Power Index creation is described in the text. The index takes values between 0 and 1 with 1 being the most market power a school district had in Tennessee and 0 representing the least market power a school district had. County characteristics comes from Haines (2010). Exact variable names used are included in parentheses. County characteristics include percent of the population which is black (negmtot + negftot/totpop), value of crops per farm (cropval/farms), and average manufacturing wage (mfgwages/mfgavear). Column (4) loses 10 observations because manufacturing wages are missing in Moore, Pickett, and Van Buren counties.

TABLE 15: SUMMARY OF COUNTERFACTUAL DISTRIBUTIONS BY DEGREE OF MARKET POWER EXERCISED BY THE SCHOOL BOARD

	Women	Women (Counterfactual)	Women (Counterfactual) Standard Error	Men
		Entire Sample		
25th Percentile	800	875	(35.2)	1000
Median	900	1055	(77.1)	1277.5
75th Percentile	1100	1325	(193.2)	1800
		Low Market Power Counties		
25th Percentile	880	916	(82.5)	990
Median	1000	1101	(214.1)	1260
75th Percentile	1200	1609	(262.1)	1800
		High Market Power Counties		
25th Percentile	725	851	(55.3)	1000
Median	900	924	(121.7)	1300
75th Percentile	1000	1070	(119.9)	1800

Notes: Counterfactual is performed using the method of Dinardo, Fortin, and Lemieux (1996). The first stage uses a logit regression to estimate the likelihood being male based on underlying characteristics. The characteristics in this regression are: principal status, high school teaching experience, high school teaching experience squared, tenure at the current school, tenure at the current school squared, educational background (college, normal school, high school, college and normal school), courses taught, school class, race of the school, and county characteristics (percent of the county that is black, average manufacturing wage, and value of the average farm's output). Women are then assigned a distribution of wages assuming their distribution of characteristics is the same as men's. Standard errors for the counterfactual are calculated using a bootstrap.

TABLE 16: COUNTY-LEVEL VARIABLES AND THE GENDER GAP

	(1)	(2)
Female * Market Power	-0.196	-0.231
	(0.075)	(0.071)
Female * Percent Black		-0.130
		(0.190)
Female * Value of Farm Output		-0.071
		(0.052)
Female * Manufacturing Wages		-0.216
		(0.127)
N	1381	1381
Adjusted R-square	0.324	0.324

Notes: Column (1) is taken from Column (4) in Table 14. Column (2) runs the same regression with controls added for the cross product of female and percent black, value of crops per farm, and average manufacturing wage. Market Power Index creation is described in the text. The index takes values between 0 and 1 with 1 being the most market power a school district had in Tennessee and 0 representing the least market power a school district had. County characteristics comes from Haines (2010). Exact variable names used are included in parentheses. County characteristics include percent of the population which is black (negmtot + negftot/totpop), value of crops per farm (cropval/farms), and average manufacturing wage (mfgwages/mfgavear). The latter two values are divided by 1000 to make the regression results easier to read.

TABLE 17: CHARACTERISTICS OF SCHOOL PRINCIPALS, BY GENDER

	Men					Women				
	1923	1926	1928	1932	1935	1923	1926	1928	1932	1935
Annual Salary (1928 Dollars)	$1,461.25	$1,666.05	$1,674.59	$2,020.23	$1,653.66	$1,129.65	$1,126.86	$1,165.32	$1,518.35	$1,297.38
Annual Salary (Nominal Dollars)	$1,461.25	$1,616.07	$1,674.59	$1,555.58	$1,273.32	$1,129.65	$1,093.05	$1,165.32	$1,169.13	$998.98
Experience	8.25	9.15	9.40	9.42	10.90	7.21	9.24	9.61	9.64	9.54
Tensure	3.17	3.61	4.03	4.65	5.89	3.21	3.24	4.05	5.50	5.06
High School	0.13	0.09	0.035	0.020	0.051	0.111	0.053	0.051	0.083	0.077
Normal School	0.17	0.13	0.104	0.032		0.200	0.289	0.128	0.083	
College	0.42	0.49	0.565	0.732	0.949	0.378	0.289	0.308	0.611	0.923
College and Normal School	0.28	0.29	0.297	0.216		0.311	0.368	0.513	0.222	
Principal of:										
First Class Schools	0.54	0.65	0.66	0.69	0.78	0.40	0.34	0.33	0.42	0.42
Second Class Schools	0.18	0.10	0.08	0.09	0.07	0.27	0.18	0.08	0.11	0.19
Third Class Schools	0.28	0.25	0.26	0.22	0.16	0.33	0.47	0.59	0.47	0.38
N	407	414	434	444	474	45	38	39	36	26

Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

Notes: Experience is defined as total years teaching at any high school. Tenure is years teaching at their current high school. Education levels and school class are mutually exclusive and represent the proportion of each gender the with given level of education or working as a principal at the given class of school.

TABLE 18: GENDER OF TENNESSEE HIGH SCHOOL PRINCIPALS, 1923-1935, REGRESSION ANALYSIS

	All Schools		First Class		Third Class
	(1)	(2)	(3)	(4)	(5)
Teachers at the School		-0.002		-0.0003	
		(0.003)		(0.002)	
Year = 1926	-0.012	-0.002	-0.026	-0.026	0.035
	(0.020)	(0.019)	(0.021)	(0.020)	(0.043)
Year = 1928	-0.014	-0.005	-0.031	-0.031	0.056
	(0.021)	(0.020)	(0.021)	(0.021)	(0.045)
Year = 1932	-0.028	-0.012	-0.026	-0.026	0.007
	(0.023)	(0.022)	(0.024)	(0.023)	(0.059)
Year = 1935	-0.046	-0.022	-0.038	-0.038	-0.014
	(0.023)	(0.023)	(0.026)	(0.025)	(0.063)
County Dummies	Y	Y	Y	Y	Y
Controls for Class of School	N	Y			
N	2357	2357	1521	1521	583

Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

Notes: Standard errors are clustered at the county level. Left hand side variable is a dummy for gender. Omitted year is 1923.

TABLE 19: GENDER WAGE GAP AMONG HIGH SCHOOL PRINCIPALS, 1923-1935

	(1)	(2)	(3)
Female	-0.200	-0.281	-0.192
	(0.097)	(0.068)	(0.064)
Year = 1926	0.154	0.163	0.124
	(0.036)	(0.039)	(0.034)
Year = 1928	0.173	0.176	0.144
	(0.033)	(0.035)	(0.027)
Year = 1932	0.325	0.343	0.314
	(0.034)	(0.036)	(0.029)
Year = 1935	0.081	0.096	0.049
	(0.041)	(0.045)	(0.038)
1926 * Female	-0.236	-0.148	-0.131
	(0.140)	(0.125)	(0.121)
1928 * Female	-0.210	-0.132	-0.109
	(0.094)	(0.084)	(0.087)
1932 * Female	-0.107	-0.022	-0.021
	(0.114)	(0.099)	(0.098)
1935 * Female	0.005	0.112	0.144
	(0.112)	(0.097)	(0.080)
High School	0.031	0.058	0.039
	(0.048)	(0.050)	(0.049)
College	0.188	0.153	0.053
	(0.031)	(0.033)	(0.030)
College and Normal School	0.225	0.195	0.088
	(0.030)	(0.029)	(0.028)
Experience	0.013	0.013	0.008
	(0.002)	(0.002)	(0.002)
Tenure	0.003	0.000	-0.005
	(0.003)	(0.004)	(0.003)
First Class			0.435
			(0.028)
Second Class			0.061
			(0.041)
County Fixed Effects	N	Y	Y
N	2133	2133	2133

Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

Notes: Standard errors are clustered at the county level. Normal School is the omitted category. 1923 is the omitted year. The education categories are 1/0 for having attended the given type of school and not number of years attended.

TABLE 20: CHARACTERISTICS OF FARMS IN TENNESSEE, 1920

	Tax in 1916	No Tax in 1916	p-value
Farmers as Share of Labor Force	0.311	0.352	0.001
Average Crop Value per Acre	15.696	15.038	0.684
Density (Population per Square Mile)	44.462	42.082	0.741
Average Farm Size (Acres)	77.313	81.221	0.582
Percent of All farms 500+ acres	0.008	0.008	0.657
Percent of All Farms 100 to 499 Acres	0.243	0.266	0.430
Cereals Produces (Percent of Total Crop Value)	0.452	0.494	0.112

Source: Haines (2010), Ruggles et al. (2010)

Notes: Average Crop Value is cropval variable, weighted by farm acreage in the county. Population Density is weighted by the total acreage of the county. The last four columns are produced by adding up nominators and denominators across all counties in a given grouping. P-Value is the result of a two-sided t-test.

TABLE 21: CHARACTERISTICS OF CHILDREN IN THE DATASET, 1910

	Tax in 1916	No Tax in 1916	P-Value
Father's Literacy Rates	79.74%	79.13%	0.875
Mother's Literacy Rates	85.89%	84.65%	0.717
Father is a Farmer (%)	65.64%	77.93%	0.004
Father is in White Collar Labor (%)	7.99%	4.29%	0.092
Percent of Children 6 - 12 in School, 1910	81.36%	72.03%	0.216
N	326	163	

Source: Collins and Wannamaker (2014)

Notes: Farmer is defined as farmer or farm laborer. Children in school is taken from the census. The question asks whether the person has been in school at all in the previous 7.5 months. The sample is further restricted to observations between 6 and 12 in the 1910 census, reducing the number of observations to 102 in counties with a tax and 50 in counties without. P-Value is the result of a two-sided t-test.

TABLE 22: RURAL CHILDREN WHO BECAME FARMERS AND WHITE COLLAR WORKERS, WHITES, TENNESSEE

	Unaffected Counties	N	Affected Counties	N	p-value
			Farmer		
Pre 1917	0.294	109	0.309	94	0.818
Post 1917	0.210	229	0.319	72	0.056
p-value	0.090		0.881		
			White Collar		
Pre 1917	0.156	109	0.191	94	0.506
Post 1917	0.105	229	0.056	72	0.211
p-value	0.179		0.010		

Source: Collins and Wannamaker (2014), Ruggles et al. (2010)

Notes: Numbers given are ratios. Post Law is after the 1917 law which required taxes for high schools. Affected County is a county with no tax rate prior the 1917 law. Standard errors are clustered at the county level. Farmers are both farm owners and farm workers. White workers are those involved in professional, clerical, or managerial jobs.

TABLE 23: IMPACT OF HIGH SCHOOL ACCESS ON CHILDREN'S OUTCOMES, SIMPLE REGRESSION, WHITES, TENNESSEE, 1930

	Farmer			White Collar		
Post*Affected	0.101	0.107	0.118	-0.079	-0.082	-0.094
	(0.093)	(0.094)	(0.091)	(0.058)	(0.059)	(0.058)
Post	-0.077	-0.082	-0.079	-0.057	-0.054	-0.056
	(0.056)	(0.056)	(0.056)	(0.044)	(0.044)	(0.044)
Affected	-0.004	-0.009	-0.013	0.037	0.040	0.045
	(0.073)	(0.074)	(0.072)	(0.044)	(0.045)	(0.044)
Controls for Father's Occupation	Y	Y	Y	Y	Y	Y
Controls for Father's Occscore	N	Y	Y	N	Y	Y
Controls for Parents' Literacy	N	N	Y	N	N	Y
N	489	489	489	489	489	489

Source: Collins and Wannamaker (2014), Ruggles et al. (2010)

Notes: Regressions are linear probability models. Post Law is after the 1917 law which required taxes for high schools. Affected County is a county with no tax rate prior the 1917 law. Standard errors are clustered at the county level. Farmers are both farm owners and farm workers. White workers are those involved in professional, clerical, or managerial jobs.

TABLE 24: IMPACT OF HIGH SCHOOL ACCESS ON CHILDREN'S OUTCOMES, WHITES, TENNESSEE, 1930

	Farmer			White Collar		
Post Law * Affected County	0.158	0.166	0.174	-0.076	-0.086	-0.099
	(0.101)	(0.101)	(0.099)	(0.079)	(0.081)	(0.079)
Turned 14: 1909	0.057	0.049	0.029	-0.122	-0.112	-0.084
	(0.158)	(0.157)	(0.160)	(0.173)	(0.172)	(0.158)
Turned 14: 1910	0.083	0.081	0.082	0.003	0.005	0.004
	(0.169)	(0.166)	(0.164)	(0.158)	(0.158)	(0.151)
Turned 14: 1911	0.073	0.082	0.078	-0.200	-0.210	-0.206
	(0.165)	(0.162)	(0.161)	(0.123)	(0.124)	(0.114)
Turned 14: 1912	0.127	0.132	0.156	-0.201	-0.206	-0.241
	(0.191)	(0.189)	(0.185)	(0.131)	(0.131)	(0.126)
Turned 14: 1913	-0.003	0.003	0.020	-0.050	-0.057	-0.081
	(0.168)	(0.166)	(0.163)	(0.157)	(0.158)	(0.147)
Turned 14: 1914	0.007	0.001	0.011	-0.174	-0.167	-0.180
	(0.157)	(0.153)	(0.150)	(0.141)	(0.141)	(0.136)
Turned 14: 1918	0.053	0.047	0.047	-0.121	-0.115	-0.114
	(0.159)	(0.156)	(0.154)	(0.149)	(0.149)	(0.142)
Turned 14: 1919	-0.003	0.002	0.018	-0.099	-0.105	-0.127
	(0.184)	(0.182)	(0.177)	(0.160)	(0.162)	(0.155)
Turned 14: 1920	0.088	0.081	0.089	-0.149	-0.142	-0.153
	(0.166)	(0.162)	(0.159)	(0.144)	(0.145)	(0.134)
Turned 14: 1921	-0.115	-0.131	-0.111	-0.239	-0.221	-0.248
	(0.165)	(0.163)	(0.161)	(0.137)	(0.137)	(0.130)
Turned 14: 1922	-0.136	-0.140	-0.136	-0.092	-0.088	-0.094
	(0.169)	(0.166)	(0.162)	(0.165)	(0.164)	(0.153)
Turned 14: 1923	-0.117	-0.118	-0.101	-0.177	-0.176	-0.198
	(0.180)	(0.177)	(0.175)	(0.130)	(0.131)	(0.125)
Turned 14: 1924	-0.096	-0.104	-0.094	-0.241	-0.232	-0.246
	(0.176)	(0.172)	(0.169)	(0.148)	(0.148)	(0.142)
Controls for Father's Occupation	Y	Y	Y	Y	Y	Y
Controls for Father's Occscore	N	Y	Y	N	Y	Y
Controls for Parents' Literacy	N	N	Y	N	N	Y
N	489	489	489	489	489	489

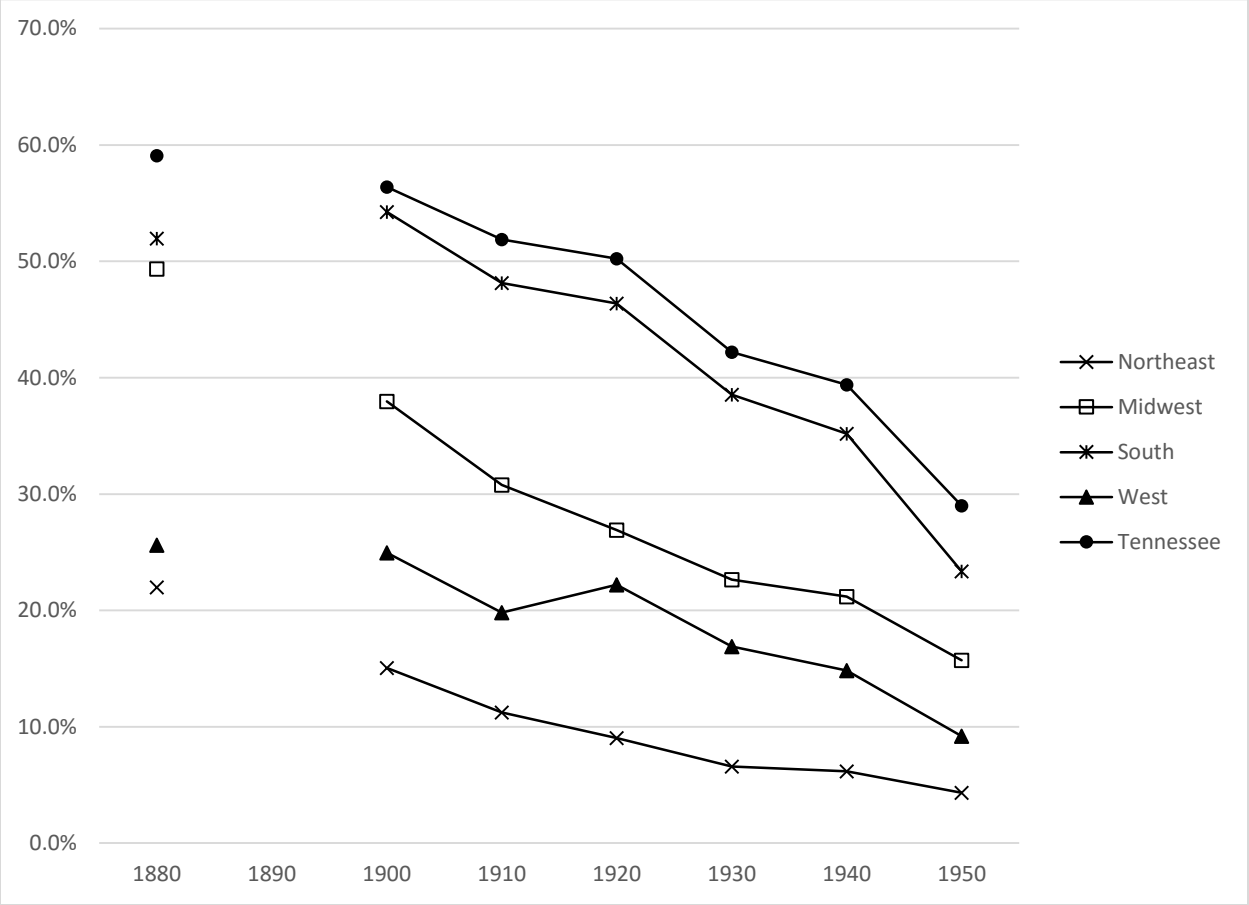
Source: Collins and Wannamaker (2014), Ruggles et al. (2010)

Notes: Regressions are linear probability models. Post Law is after the 1917 law which required taxes for high schools. Affected County is a county with no tax rate prior the 1917 law. Standard errors are clustered at the county

level. Farmers are both farm owners and farm workers. White collar workers are those involved in professional, clerical, or managerial jobs.

FIGURES

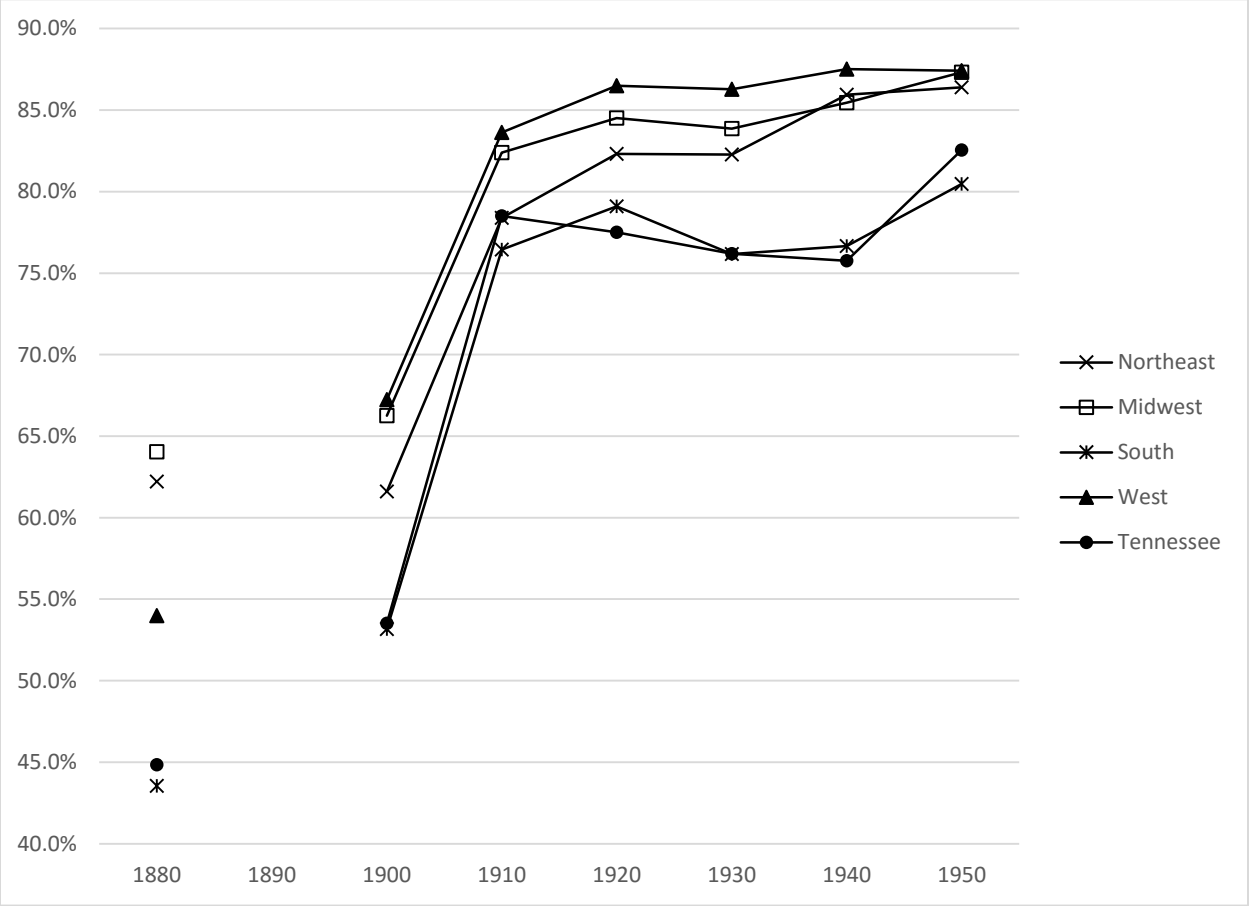
FIGURE 1: PERCENTAGE OF Adult Population Living ON FARMS, 1880 TO 1950, BY REGION



Source: Ruggles et al. (2010), 1% samples

Notes: Data consists of all people over 25 years of age. The sample includes data from the 48 contiguous US states. The South is defined as the Confederate states, Kentucky, West Virginia, and Maryland. Delaware is considered a northeastern state.

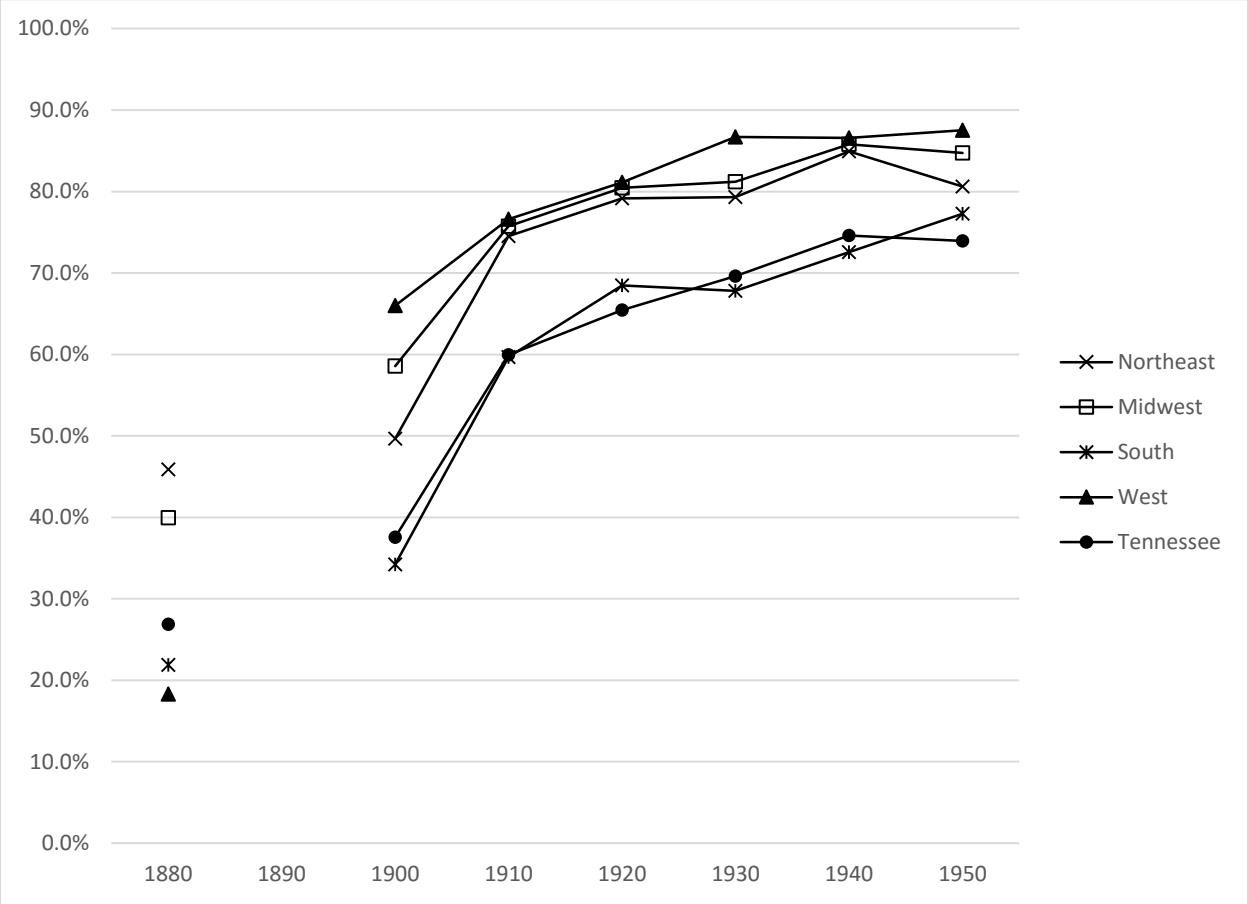
FIGURE 2: SCHOOL ATTENDANCE RATES, WHITE, AGE 6 TO 18, 1880 TO 1950, BY REGION



Source: Ruggles et al. (2010), 1% samples

Notes: Data consists of all people 6 to 18 years of age. The sample includes data from the 48 contiguous US states. The South is defined as the Confederate states, Kentucky, West Virginia, and Maryland. Delaware is considered a northeastern state.

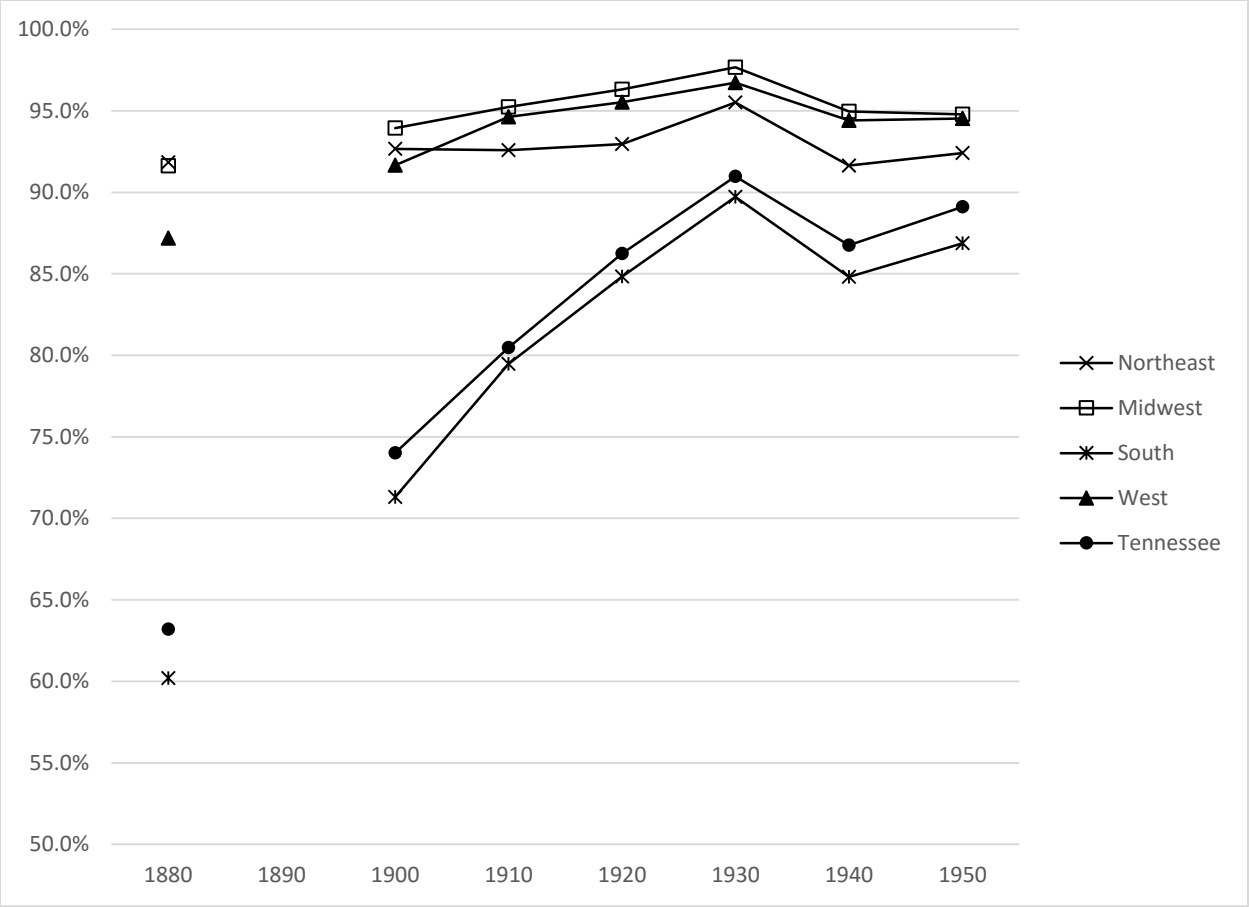
FIGURE 3: SCHOOL ATTENDANCE RATES, BLACK, AGE 6 TO 18, 1880 TO 1950, BY REGION



Source: Ruggles et al. (2010), 1% samples

Notes: Data consists of all people 6 to 18 years of age. The sample includes data from the 48 contiguous US states. The South is defined as the Confederate states, Kentucky, West Virginia, and Maryland. Delaware is considered a northeastern state.

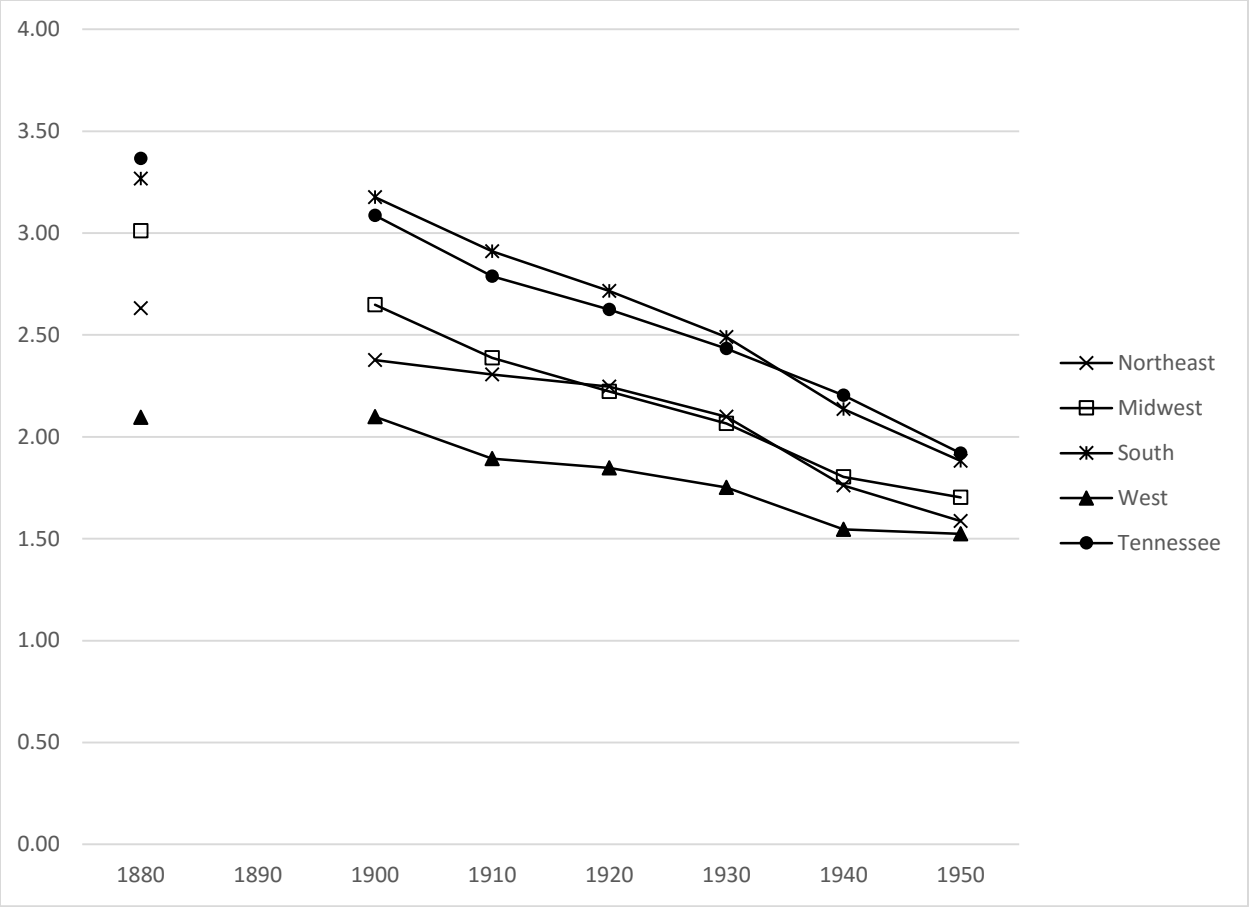
FIGURE 4: LITERACY RATES OF ADULT POPULATION, 1880 TO 1950, BY REGION



Source: Ruggles et al. (2010)

Notes: Data consists of all people over 25 years of age. The sample includes data from the 48 contiguous US states. The South is defined as the Confederate states, Kentucky, West Virginia, and Maryland. Delaware is considered a northeastern state. Literacy is taken from the census data prior to 1940. In 1940 and after, people with more than 2 years of schooling are assumed to be literate. This explains the drop in 1940.

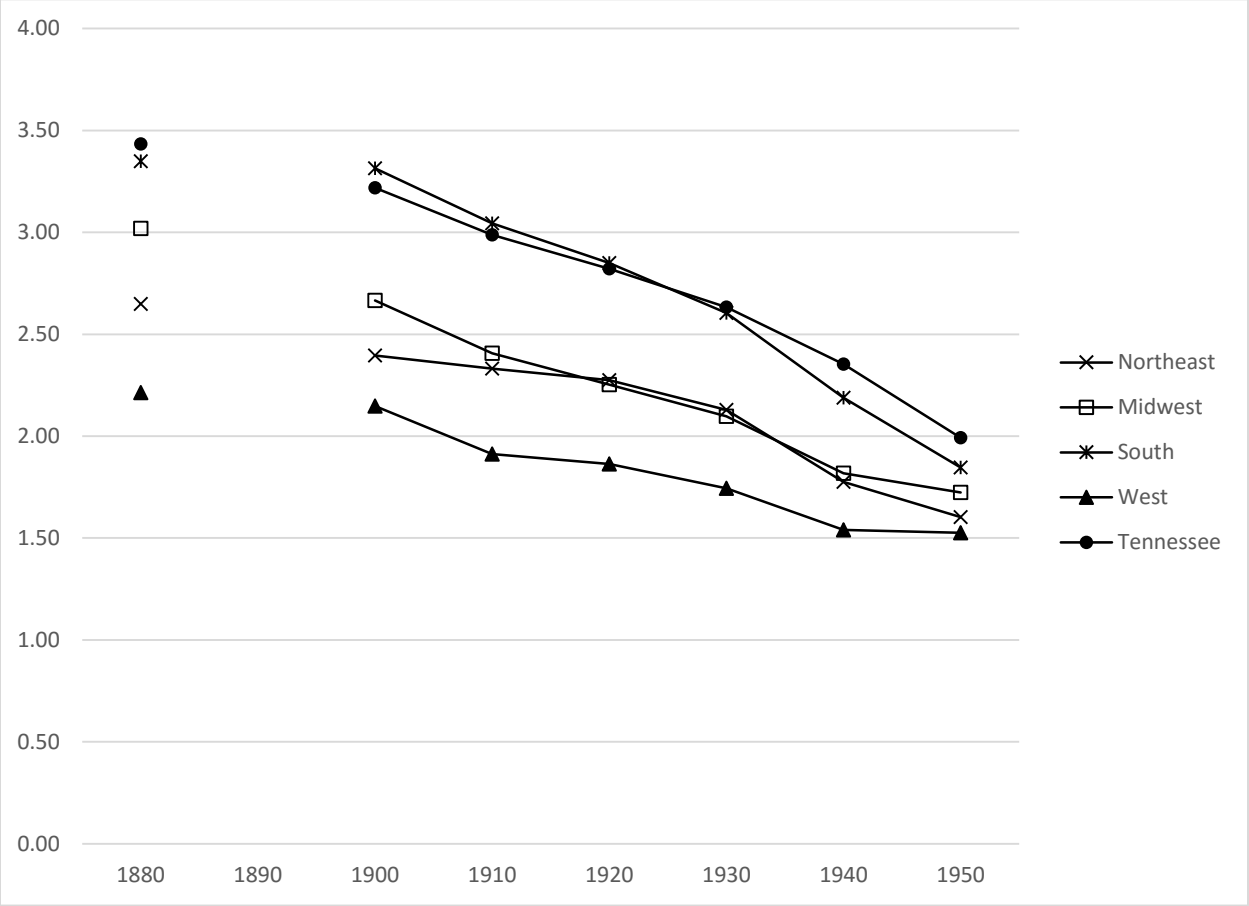
FIGURE 5: NUMBER OF CHILDREN PER HEAD OF HOUSE, 1880 TO 1950, BY REGION



Source: Ruggles et al. (2010)

Notes: Contains children for heads-of-house between 30 and 50 years of age. The sample includes data from the 48 contiguous US states. The South is defined as the Confederate states, Kentucky, West Virginia, and Maryland. Delaware is considered a northeastern state. Number of children is the nchild variable in IPUMS.

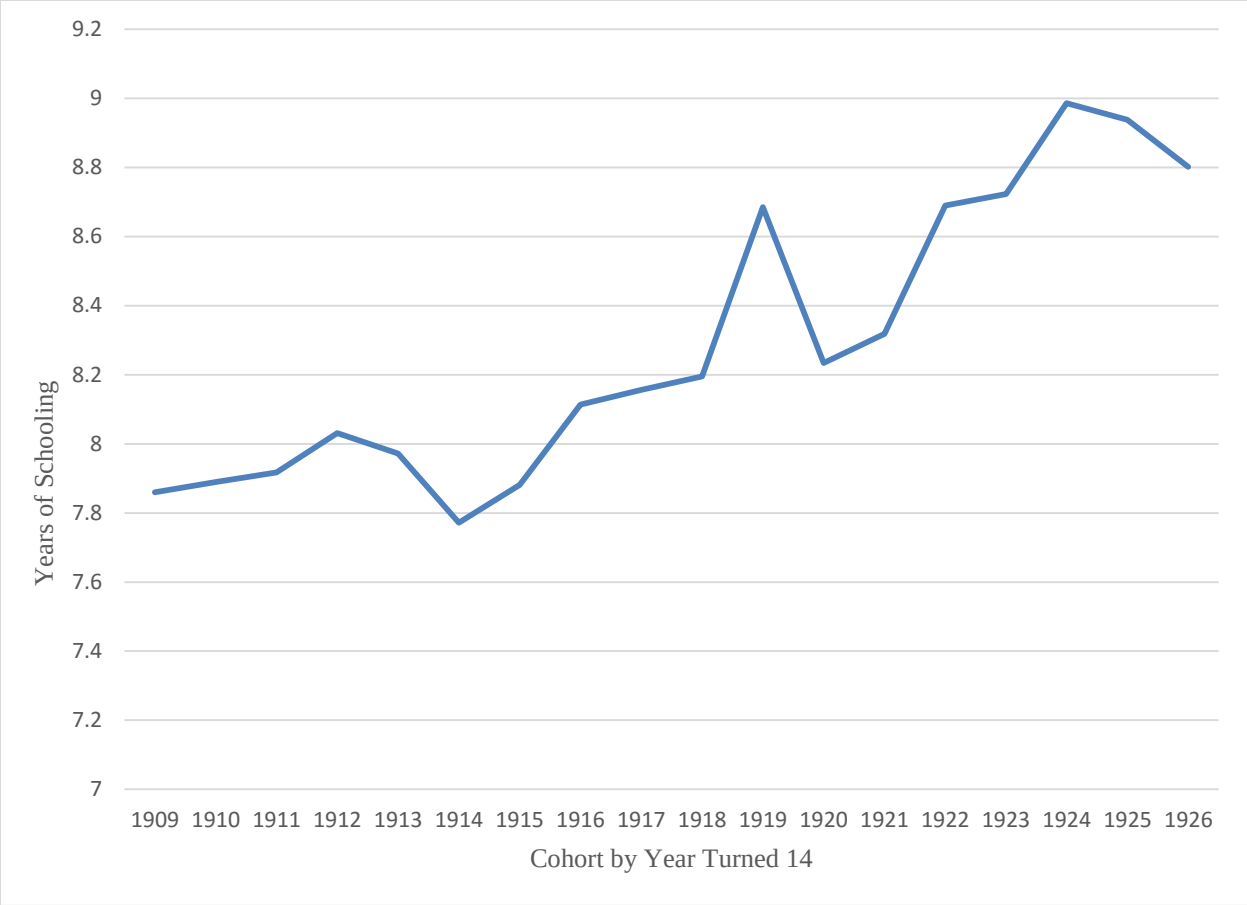
FIGURE 6: NUMBER OF WHITE CHILDREN PER HEAD OF HOUSE, 1880 TO 1950, BY REGION



Source: Ruggles et al. (2010)

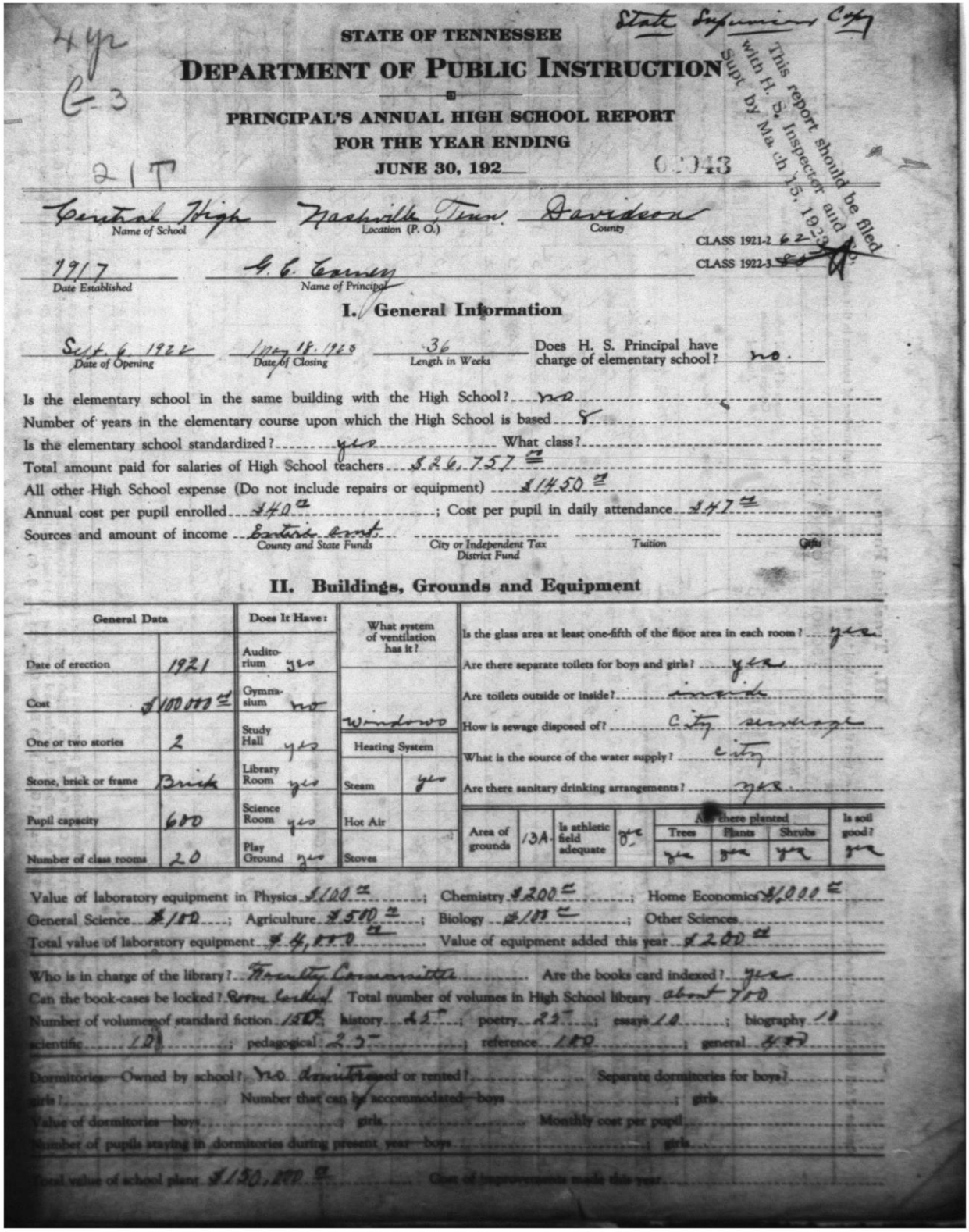
Notes: Contains children for heads-of-house between 30 and 50 years of age. The sample includes data from the 48 contiguous US states. The South is defined as the Confederate states, Kentucky, West Virginia, and Maryland. Delaware is considered a northeastern state. Number of children is the nchild variable in IPUMS.

FIGURE 7: YEARS OF SCHOOLING BY BIRTH COHORT, WHITES, TENNESSEE



Source: Ruggles et al. (2010)

FIGURE 8: EXAMPLE HIGH SCHOOL REPORT



STATE OF TENNESSEE

DEPARTMENT OF PUBLIC INSTRUCTION

PRINCIPAL'S ANNUAL HIGH SCHOOL REPORT

FOR THE YEAR ENDING

JUNE 30, 192

State Supervisor's Copy

This report should be filed with H. S. Inspector and Supt by March 15, 1922

4 yr

G-3

21 T

0043

Central High — Name of School; Nashville Tenn. — Location (P. O.); Davidson — County

CLASS 1921-2 6 2

CLASS 1922-3

1917 — Date Established; G. G. Tanner — Name of Principal

I. General Information

Sept. 6 1922 — Date of Opening; May 18. 1923 — Date of Closing; 36 — Length in Weeks; Does H. S. Principal have charge of elementary school? no.

Is the elementary school in the same building with the High School? yes

Number of years in the elementary course upon which the High School is based 8

Is the elementary school standardized? yes What class?

Total amount paid for salaries of High School teachers $26,757 ☑

All other High School expense (Do not include repairs or equipment) $1450 ☑

Annual cost per pupil enrolled $40 ☑; Cost per pupil in daily attendance $47 ☑

Sources and amount of income — County and State Funds: Entire Amt.; City or Independent Tax District Fund:; Tuition:; Gifts:

II. Buildings, Grounds and Equipment

General Data		Does It Have:		What system of ventilation has it?
Date of erection	1921	Auditorium	yes	
Cost	$100,000 ☑	Gymnasium	no	
One or two stories	2	Study Hall	yes	Windows
Stone, brick or frame	Brick	Library Room	yes	Heating System
Pupil capacity	600	Science Room	yes	Steam yes
Number of class rooms	20	Play Ground	yes	Hot Air
				Stoves

Is the glass area at least one-fifth of the floor area in each room? yes

Are there separate toilets for boys and girls? yes

Are toilets outside or inside? inside

How is sewage disposed of? City sewerage

What is the source of the water supply? city

Are there sanitary drinking arrangements? yes

Area of grounds	13A	Is athletic field adequate	yes	Are there planted: Trees	Plants	Shrubs	Is soil good?
				yes	yes	yes	yes

Value of laboratory equipment in Physics $100 ☑; Chemistry $200 ☑; Home Economics $1,000 ☑

General Science $150; Agriculture $500 ☑; Biology $100 ☑; Other Sciences

Total value of laboratory equipment $4,000 ☑ Value of equipment added this year $200 ☑

Who is in charge of the library? Faculty Committee Are the books card indexed? yes

Can the book-cases be locked? Doors locked Total number of volumes in High School library about 700

Number of volumes of standard fiction 150; history 45; poetry 25; essays 10; biography 10

scientific 10; pedagogical 25; reference 100; general 425

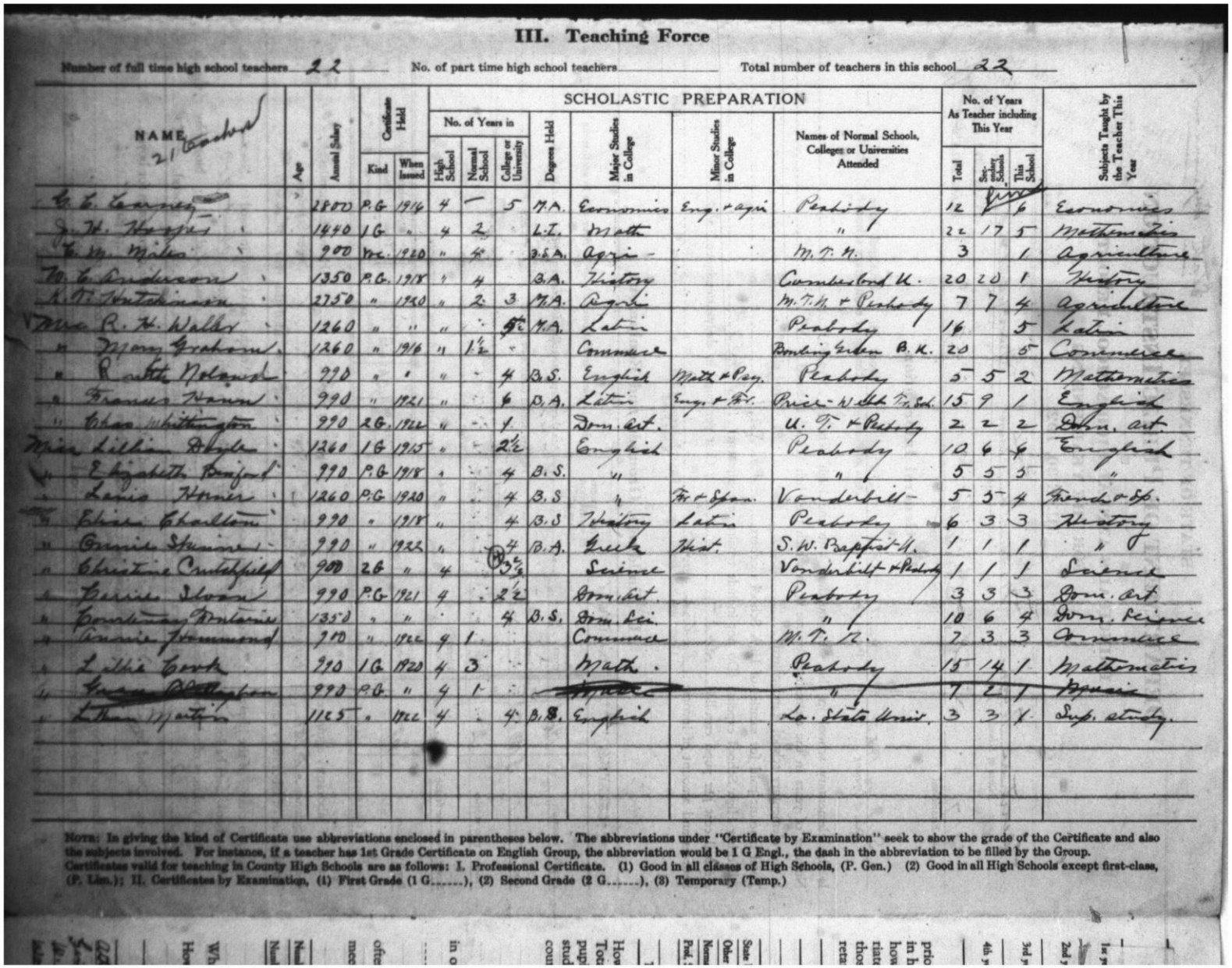
Dormitories—Owned by school? no dormitories Leased or rented? Separate dormitories for boys?

girls? Number that can be accommodated—boys; girls

Value of dormitories—boys girls Monthly cost per pupil

Number of pupils staying in dormitories during present year—boys; girls

Total value of school plant $150,000 ☑ Cost of improvements made this year



III. Teaching Force

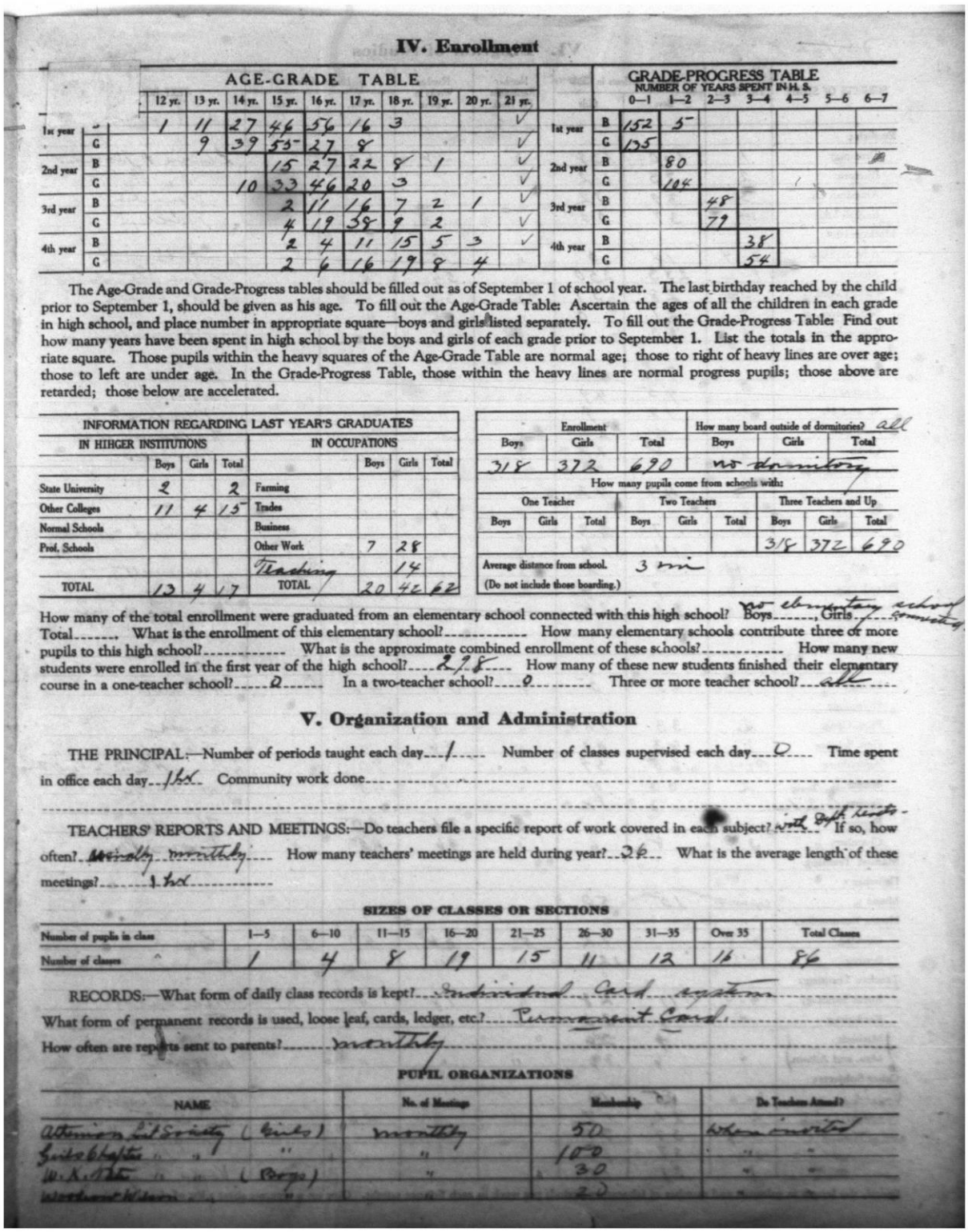
Number of full time high school teachers 22 No. of part time high school teachers _____ Total number of teachers in this school 22

NAME 21 Teachers	Age	Annual Salary	Certificate Held: Kind	Certificate Held: When Issued	SCHOLASTIC PREPARATION: No. of Years in High School	No. of Years in Normal School	No. of Years in College or University	Degrees Held	Major Studies in College	Minor Studies in College	Names of Normal Schools, Colleges or Universities Attended	No. of Years As Teacher including This Year: Total	No. of Years As Teacher including This Year: Since Last School	No. of Years As Teacher including This Year: This School	Subjects Taught by the Teacher This Year
E. E. Tearney		2800	P.G.	1914	4	—	5	M.A.	Economics	Eng. + Agri.	Peabody	12		6	Economics
J. H. Hodges		1440	1G	"	4	2		L.I.	Math		"	22	17	5	Mathematics
T. M. Miles		900	Mc.	1920	"	4		B.S.A.	Agri.		M. T. H.	3		1	Agriculture
W. C. Anderson		1350	P.G.	1918	"	4		B.A.	History		Cumberland U.	20	20	1	History
A. T. Tutchinson		2750	"	1920	"	2	3	M.A.	Agri.		M. T. H. + Peabody	7	7	4	Agriculture
Miss R. H. Walker		1260	"	"	"		5½	M.A.	Latin		Peabody	16		5	Latin
" Mary Graham		1260	"	1916	"	1½			Commerce		Bowling Green B. K.	20		5	Commerce
" Ruth Noland		990	"	"	"		4	B.S.	English	Math + Psy.	Peabody	5	5	2	Mathematics
" Frances Nunn		990	"	1921	"		4	B.A.	Latin	Eng. + Fr.	Price-Webb Tr. Sch.	15	9	1	English
" Chas. Whittington		990	2G.	1922	"		1		Dom. Art.		U. T. + Peabody	2	2	2	Dom. Art.
Miss Lillian Boyle		1260	1G	1915	"		2½		English		Peabody	10	6	4	English
" Elizabeth Benford		990	P.G.	1918	"		4	B.S.	"		"	5	5	5	"
" Lania Horner		1260	P.G.	1920	"		4	B.S.	"	Fr. + Span.	Vanderbilt	5	5	4	Friends + Sp.
" Elsie Chaillon		990	"	1918	"		4	B.S.	History	Latin	Peabody	6	3	3	History
" Annie Stevens		990	"	1922	"		4	B.A.	Greek	Hist.	S. W. Baptist U.	1	1	1	"
" Christine Critchfield		900	2G	"	4		3½		Science		Vanderbilt + Peabody	1	1	1	Science
" Lillie Sloan		990	P.G.	1921	4		2½		Dom. Art.		Peabody	3	3	3	Dom. Art.
" Courtenay Watkins		1350	"	"	"		4	B.S.	Dom. Sci.		"	10	6	4	Dom. Science
" Annie Drummond		900	"	1922	4	1			Commerce		M. T. H.	7	3	3	Commerce
" Lillie Cook		990	1G	1920	4	3			Math.		Peabody	15	14	1	Mathematics
" Eugene Pettigrew		990	P.G.	"	4	1			[illegible]		"	7	2	1	Music
" Lillian Martin		1125	"	1922	4		4	B.S.	English		La. State Univ.	3	3	1	Sup. study.

Note: In giving the kind of Certificate use abbreviations enclosed in parentheses below. The abbreviations under "Certificate by Examination" seek to show the grade of the Certificate and also the subjects involved. For instance, if a teacher has 1st Grade Certificate on English Group, the abbreviation would be 1 G Engl., the dash in the abbreviation to be filled by the Group. Certificates valid for teaching in County High Schools are as follows: I. Professional Certificate. (1) Good in all classes of High Schools, (P. Gen.) (2) Good in all High Schools except first-class, (P. Lim.); II. Certificates by Examination, (1) First Grade (1 G.....), (2) Second Grade (2 G.....), (3) Temporary (Temp.)

FIGURE 8: EXAMPLE HIGH SCHOOL REPORT (CONT)

FIGURE 8: EXAMPLE HIGH SCHOOL REPORT (CONT)



IV. Enrollment

AGE-GRADE TABLE

		12 yr.	13 yr.	14 yr.	15 yr.	16 yr.	17 yr.	18 yr.	19 yr.	20 yr.	21 yr.
1st year	B	1	11	27	46	56	16	3			✓
	G		9	39	53	27	8				✓
2nd year	B				15	27	22	8	1		✓
	G			10	33	46	20	3			✓
3rd year	B				2	11	16	7	2	1	✓
	G				4	19	38	9	2		✓
4th year	B				2	4	11	15	5	3	✓
	G				2	6	16	19	8	4	

GRADE-PROGRESS TABLE
NUMBER OF YEARS SPENT IN H. S.

		0—1	1—2	2—3	3—4	4—5	5—6	6—7
1st year	B	152	5					
	G	135						
2nd year	B		80					
	G		104					
3rd year	B			48				
	G			79				
4th year	B				38			
	G				54			

The Age-Grade and Grade-Progress tables should be filled out as of September 1 of school year. The last birthday reached by the child prior to September 1, should be given as his age. To fill out the Age-Grade Table: Ascertain the ages of all the children in each grade in high school, and place number in appropriate square—boys and girls listed separately. To fill out the Grade-Progress Table: Find out how many years have been spent in high school by the boys and girls of each grade prior to September 1. List the totals in the appropriate square. Those pupils within the heavy squares of the Age-Grade Table are normal age; those to right of heavy lines are over age; those to left are under age. In the Grade-Progress Table, those within the heavy lines are normal progress pupils; those above are retarded; those below are accelerated.

INFORMATION REGARDING LAST YEAR'S GRADUATES

IN HIHGER INSTITUTIONS	Boys	Girls	Total	IN OCCUPATIONS	Boys	Girls	Total
State University	2		2	Farming			
Other Colleges	11	4	15	Trades			
Normal Schools				Business			
Prof. Schools				Other Work	7	28	
				Teaching		14	
TOTAL	13	4	17	TOTAL	20	42	62

Enrollment			How many board outside of dormitories? all
Boys	Girls	Total	Boys / Girls / Total
318	372	690	no dormitory

How many pupils come from schools with:

One Teacher			Two Teachers			Three Teachers and Up		
Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
						318	372	690

Average distance from school. 3 mi
(Do not include those boarding.)

How many of the total enrollment were graduated from an elementary school connected with this high school? Boys ____, Girls ____ No elementary school in community Total ____. What is the enrollment of this elementary school? ____ How many elementary schools contribute three or more pupils to this high school? ____ What is the approximate combined enrollment of these schools? ____ How many new students were enrolled in the first year of the high school? 298 How many of these new students finished their elementary course in a one-teacher school? 0 In a two-teacher school? 0 Three or more teacher school? all

V. Organization and Administration

THE PRINCIPAL:—Number of periods taught each day 1 Number of classes supervised each day 0 Time spent in office each day 1hr Community work done ____

TEACHERS' REPORTS AND MEETINGS:—Do teachers file a specific report of work covered in each subject? with Dept. heads If so, how often? Weekly monthly How many teachers' meetings are held during year? 24 What is the average length of these meetings? 1 hr

SIZES OF CLASSES OR SECTIONS

Number of pupils in class	1—5	6—10	11—15	16—20	21—25	26—30	31—35	Over 35	Total Classes
Number of classes	1	4	8	19	15	11	12	16	86

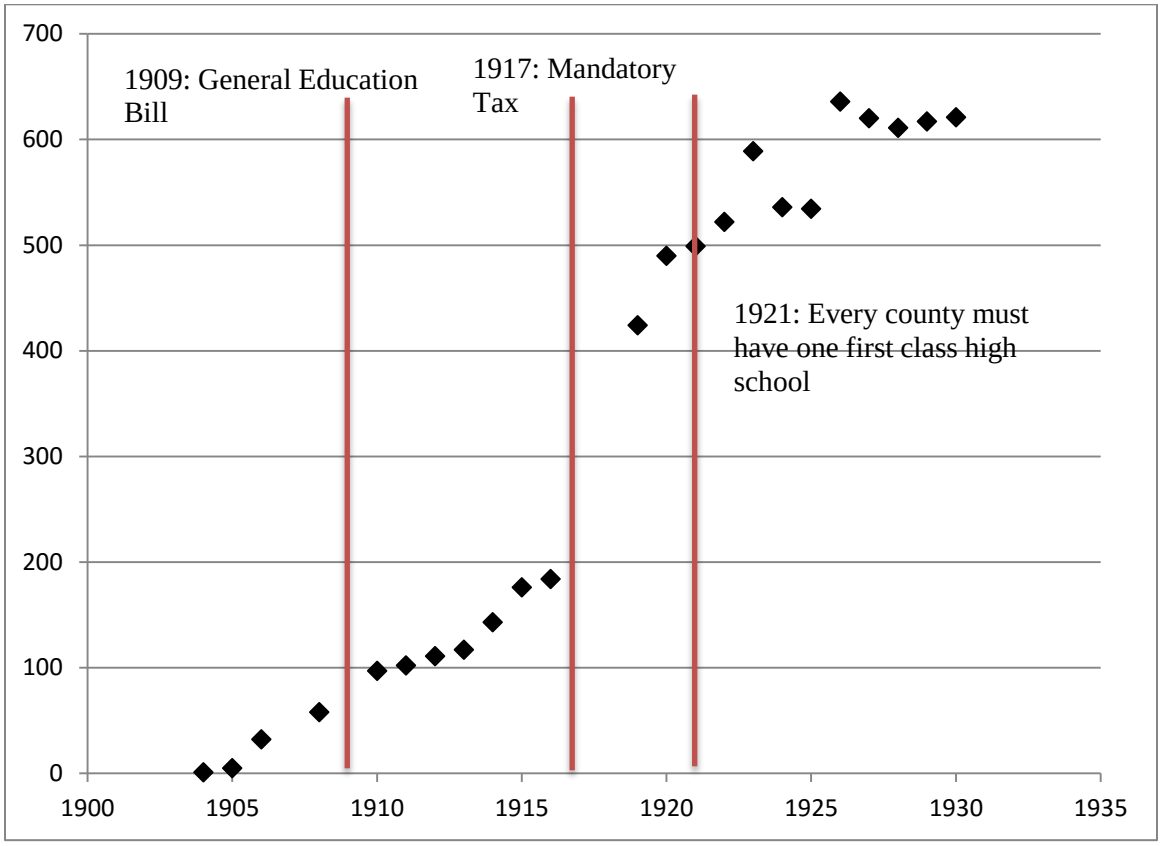
RECORDS:—What form of daily class records is kept? Individual card system
What form of permanent records is used, loose leaf, cards, ledger, etc.? Permanent Card
How often are reports sent to parents? monthly

PUPIL ORGANIZATIONS

NAME	No. of Meetings	Membership	Do Teachers Attend?
Athletic Ass't Society (Girls)	monthly	50	When invited
Girls' Chapter "	"	100	" "
W. A. A. (Boys)	"	30	" "
Merchant Widow "	"	20	" "

Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

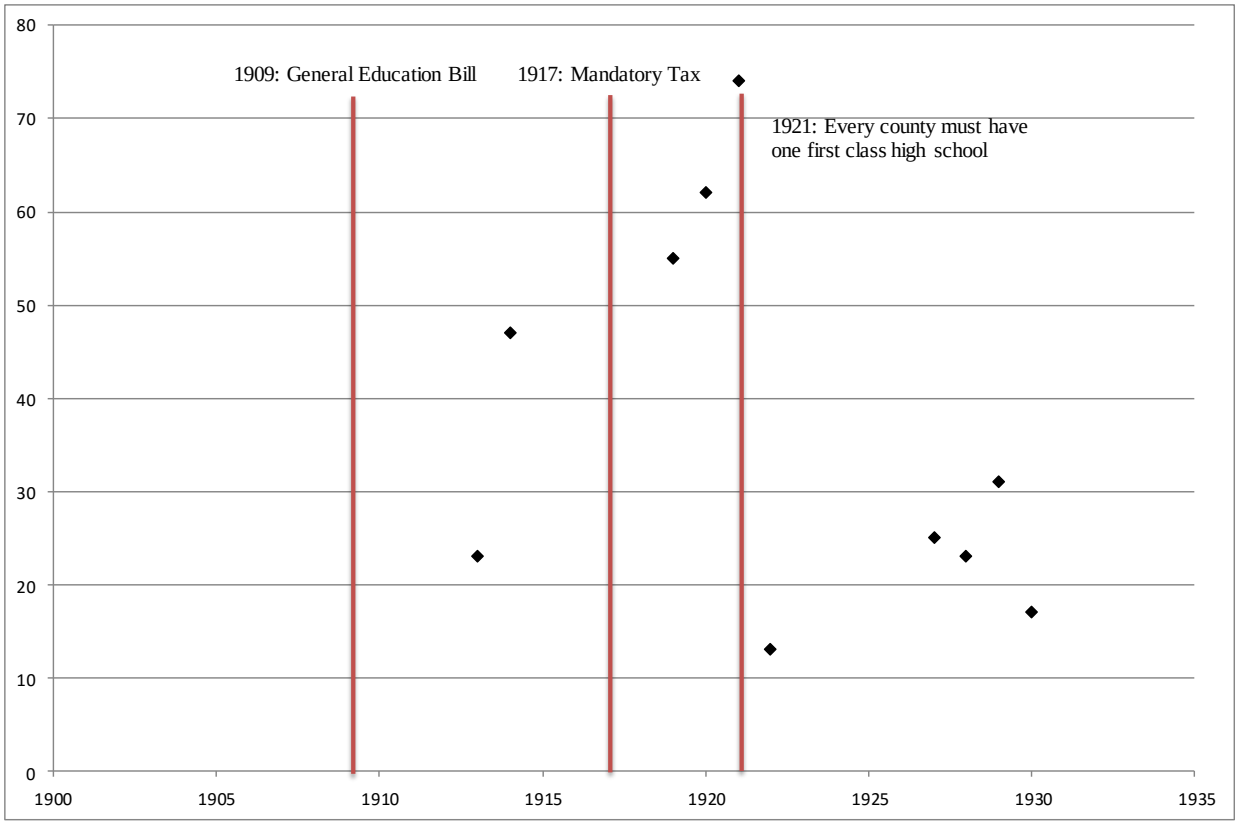
FIGURE 9: TENNESSEE COUNTY WHITE HIGH SCHOOLS BY YEAR, 1900 TO 1930



Source: Report of the State Superintendent of Public Education.

Notes: These numbers exclude city high schools. Missing points in the chart after 1904 represent missing data.

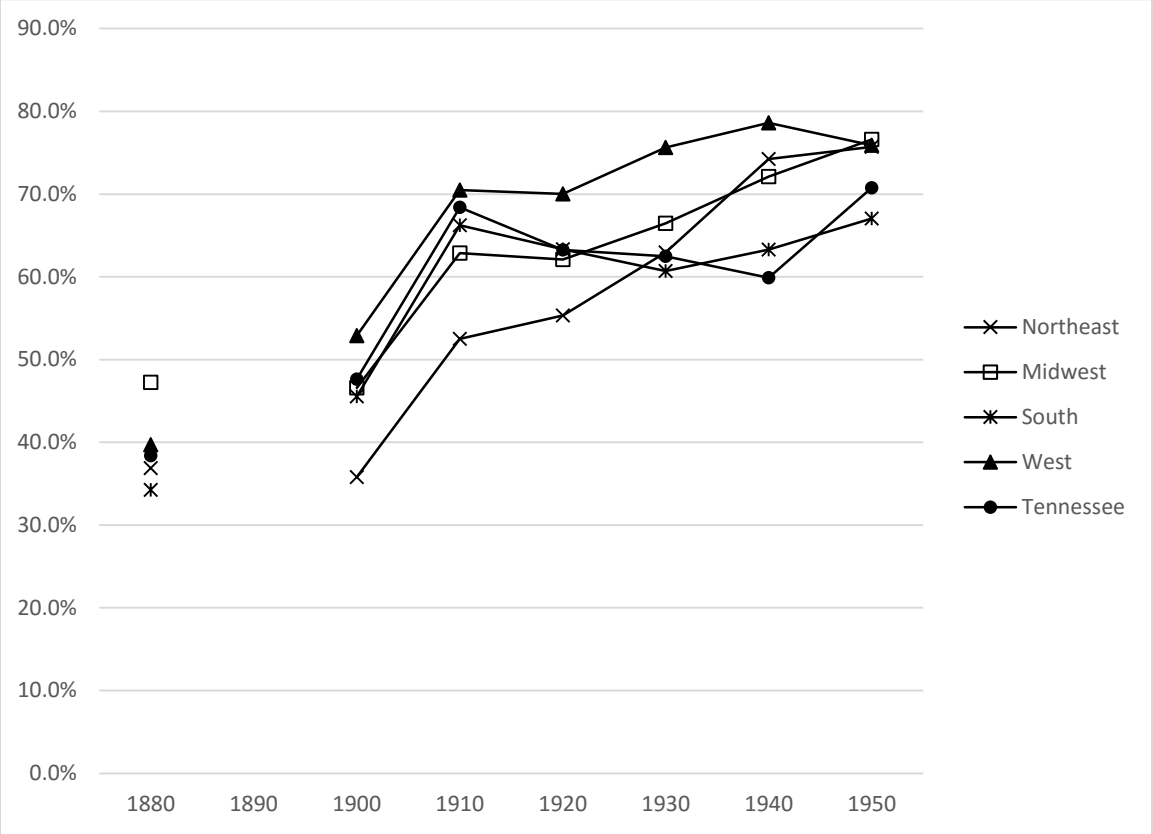
FIGURE 10: CITY WHITE HIGH SCHOOLS BY YEAR, 1900 TO 1930



Source: Report of the State Superintendent of Public Education.

Notes: These numbers exclude county high schools. Missing points in the chart after 1904 represent missing data.

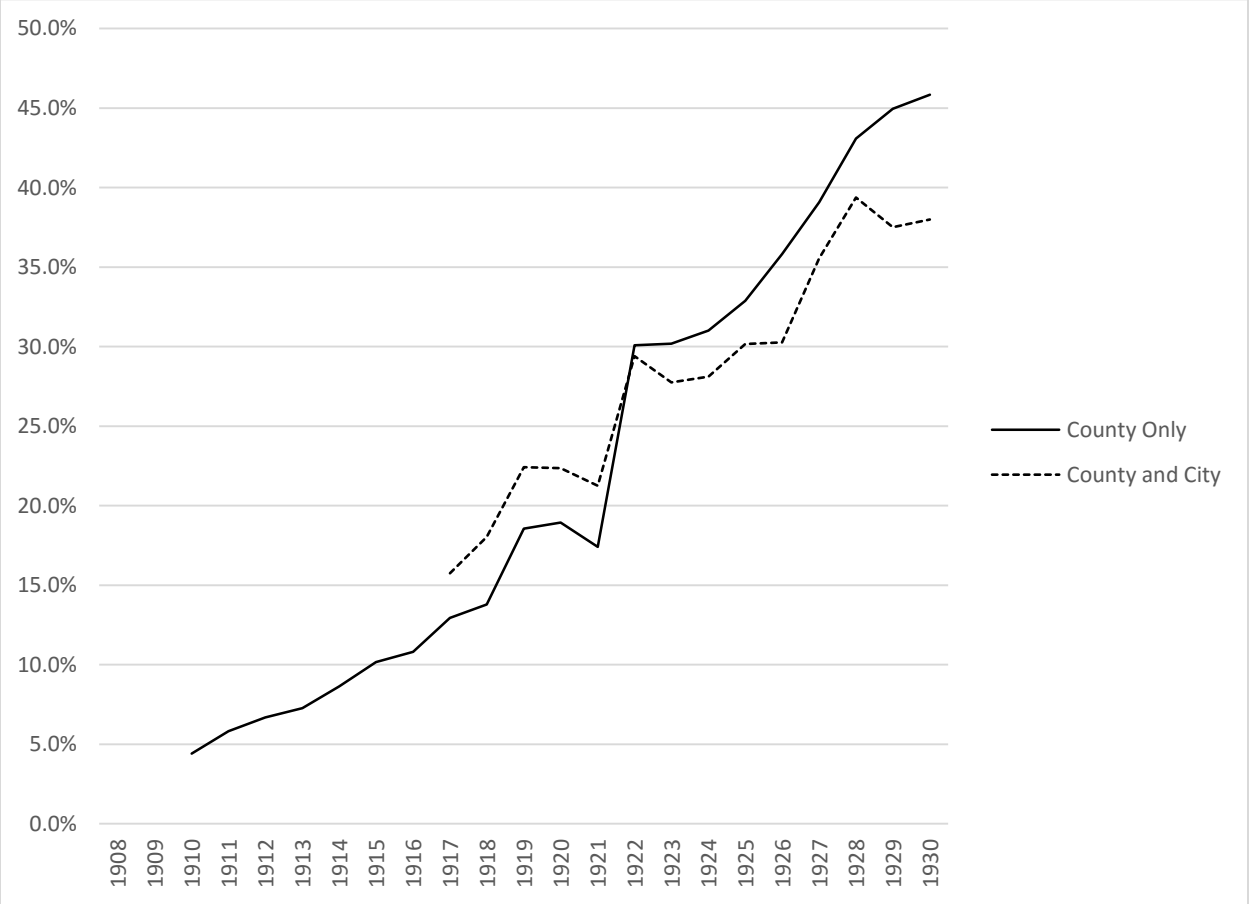
FIGURE 11: PERCENT OF WHITES 14 TO 18 ATTENDING SCHOOL, 1880 TO 1950, BY REGION



Source: Ruggles et al. (2010)

Notes: The sample includes data from the 48 contiguous US states. West Virginia is considered to be a Midwestern state. The South is defined as the Confederate states, Kentucky, and Maryland. School attendance is the school variable in IPUMS, which captures whether the respondent attended some school between a specified date. The critical dates vary by year. From 1880 to 1890, it was the 12 months ending June 1. For 1910, September 1 to April 15. For 1920, September 1 to January 1. For 1930 to 1950, September 1 to April 1.

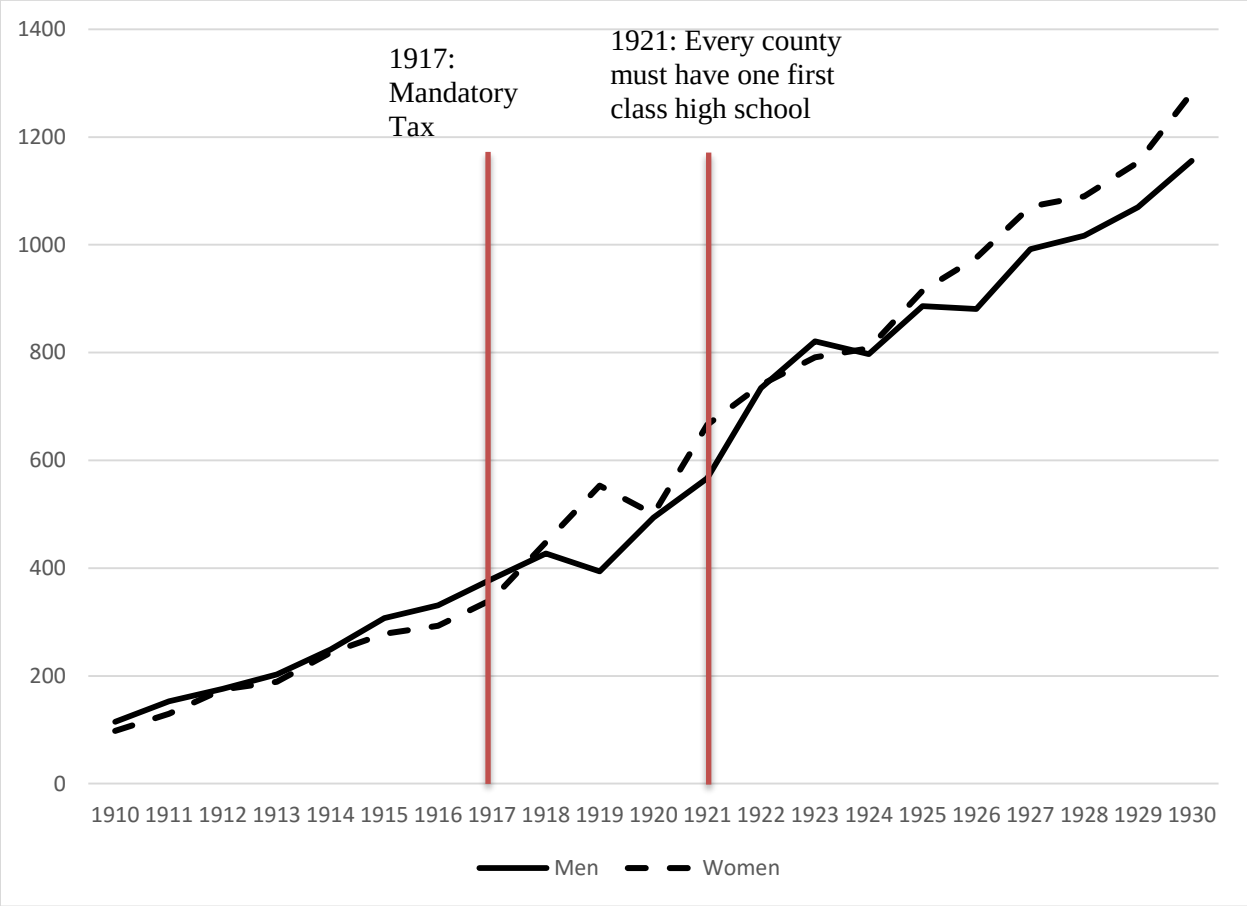
FIGURE 12: PERCENT OF WHITES IN TENNESSEE 14 TO 18 ATTENDING HIGH SCHOOL SCHOOL, 1908 TO 1930, COUNTY LEVEL AND TOTAL



Source: State Superintendent of Public Education (Tennessee) Annual Report.

Notes: Data points 1924 and before are calculated using children 6 to 21 as the basis. Data points after are calculated using children 6 to 18 as the basis. The basis is used to calculate the number of children over a 4 year cohort.

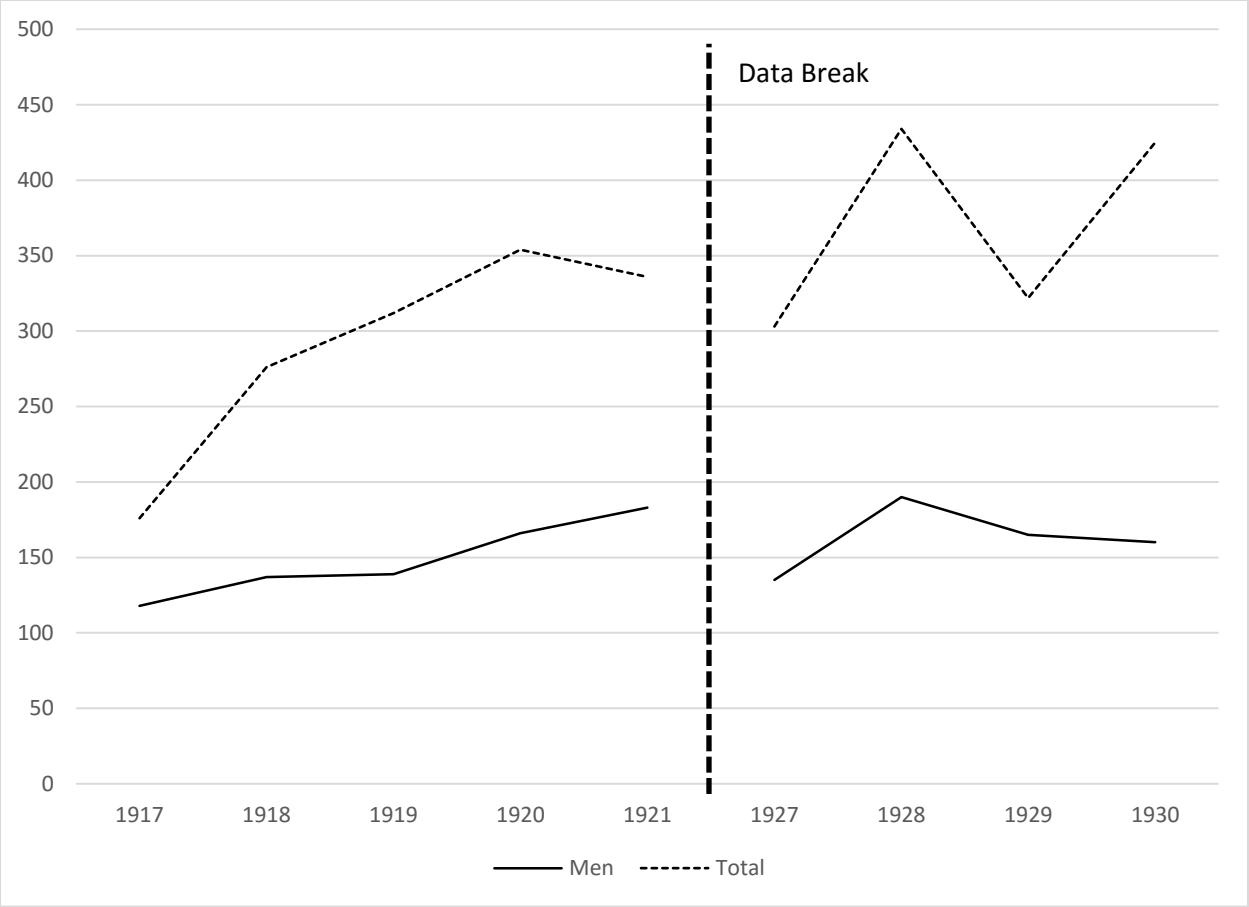
FIGURE 13: HIGH SCHOOL TEACHERS BY GENDER, 1910 TO 1930



Source: Report of the State Superintendent of Public Education.

Notes: These numbers exclude city high schools.

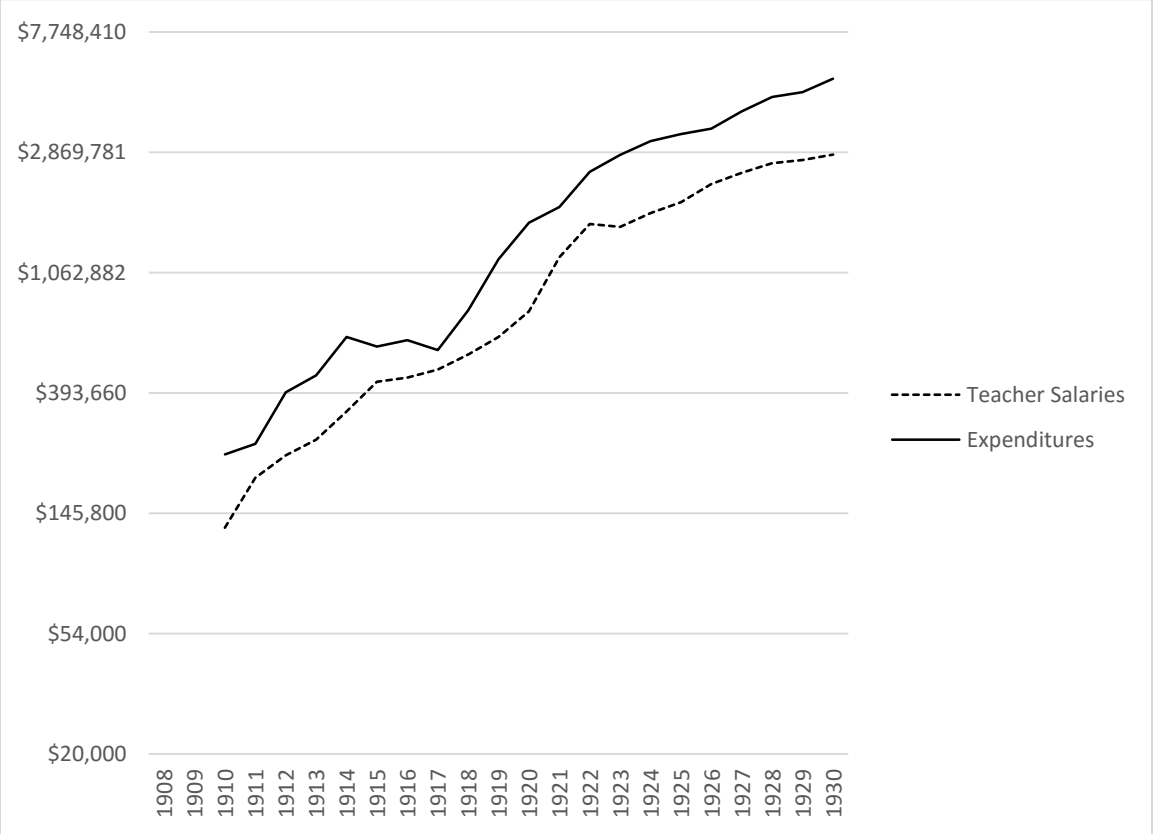
FIGURE 14: CITY HIGH SCHOOL TEACHERS BY GENDER, 1917 TO 1930



Source: Report of the State Superintendent of Public Education.

Notes: These numbers exclude county high schools.

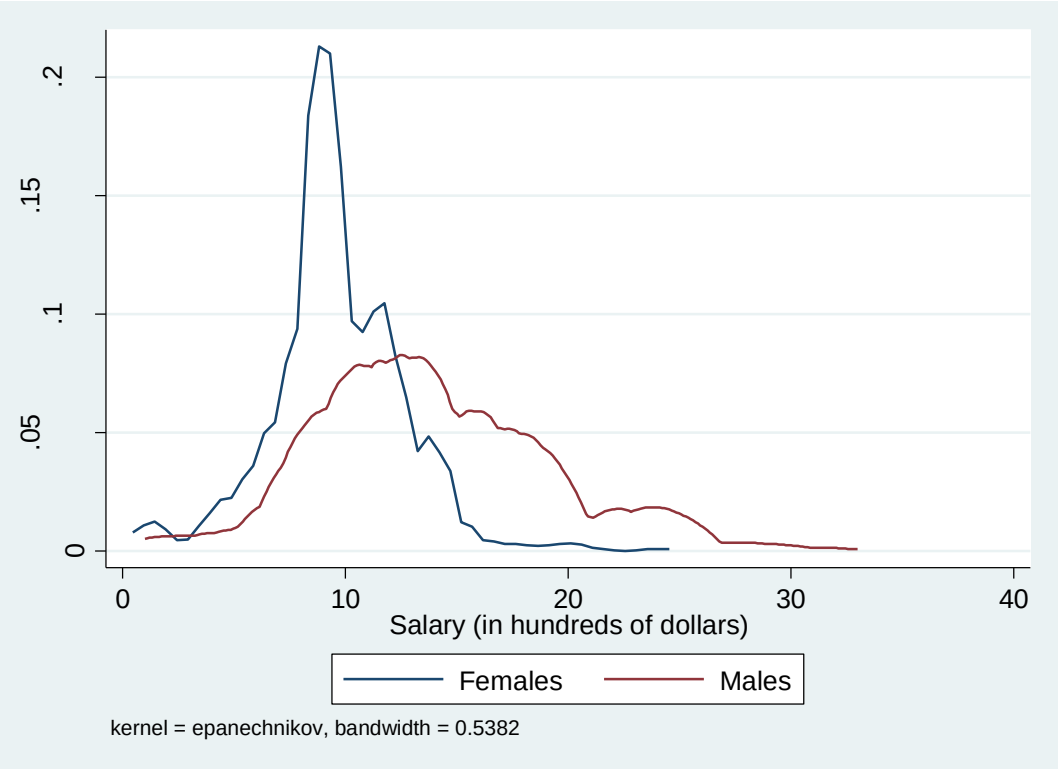
FIGURE 15: COUNTY TEACHER SALARIES AND TOTAL EXPENDITURES, 1908 TO 1930



Source: Report of the State Superintendent of Public Education.

Notes: These numbers exclude city high schools. The vertical axis has a natural log scale.

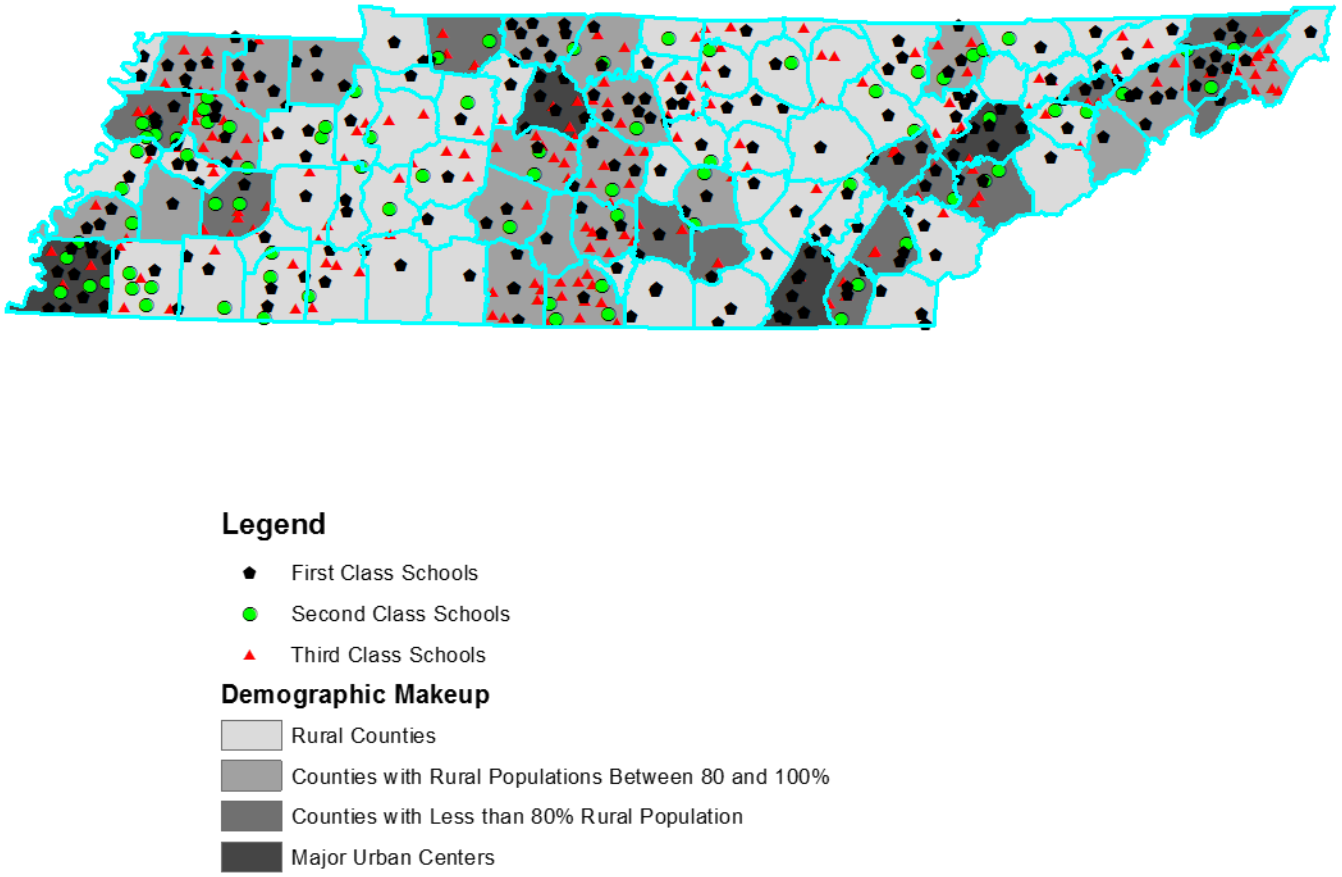
FIGURE 16: ESTIMATED DISTRIBUTIONS FOR MALE AND FEMALE SALARIES



Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

Notes: Densities are calculated using STATA's kdensity function. The distributions are based of 706 observations for men and 685 observations for women.

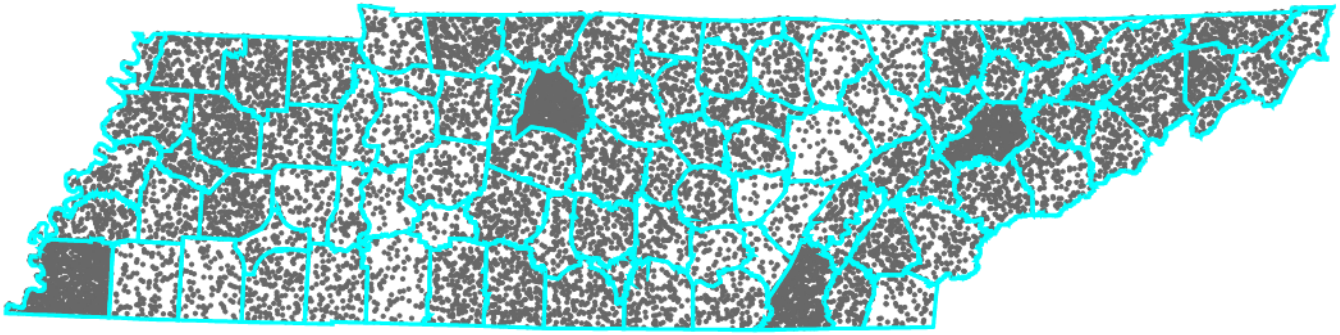
FIGURE 17: HIGH SCHOOLS IN TENNESSEE



Source: United States Board on Geographic Name database and Google Maps.

Notes: Historical location was used where available, then present school location (USBGN then Google if not available in USBGN), location of the city closest to the school (USBGN then Google), and name of the river closest to the school.

FIGURE 18: TOTAL POPULATION, AGES 14 TO 20



Source: Haines (2010).

Notes: Each dot represents 15 people.

FIGURE 19: A SAMPLE ENTRY FROM THE INDIVIDUAL SCHOOL REPORTS

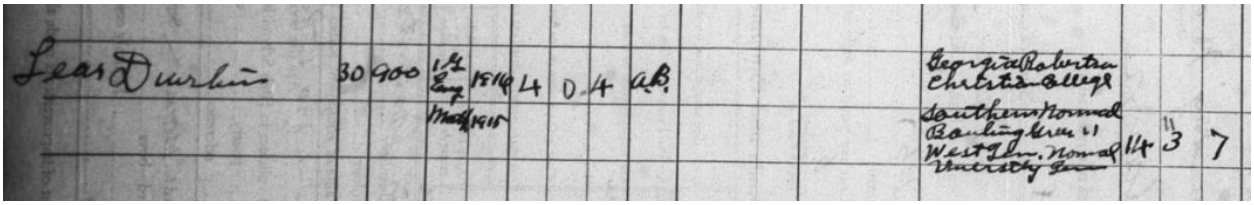
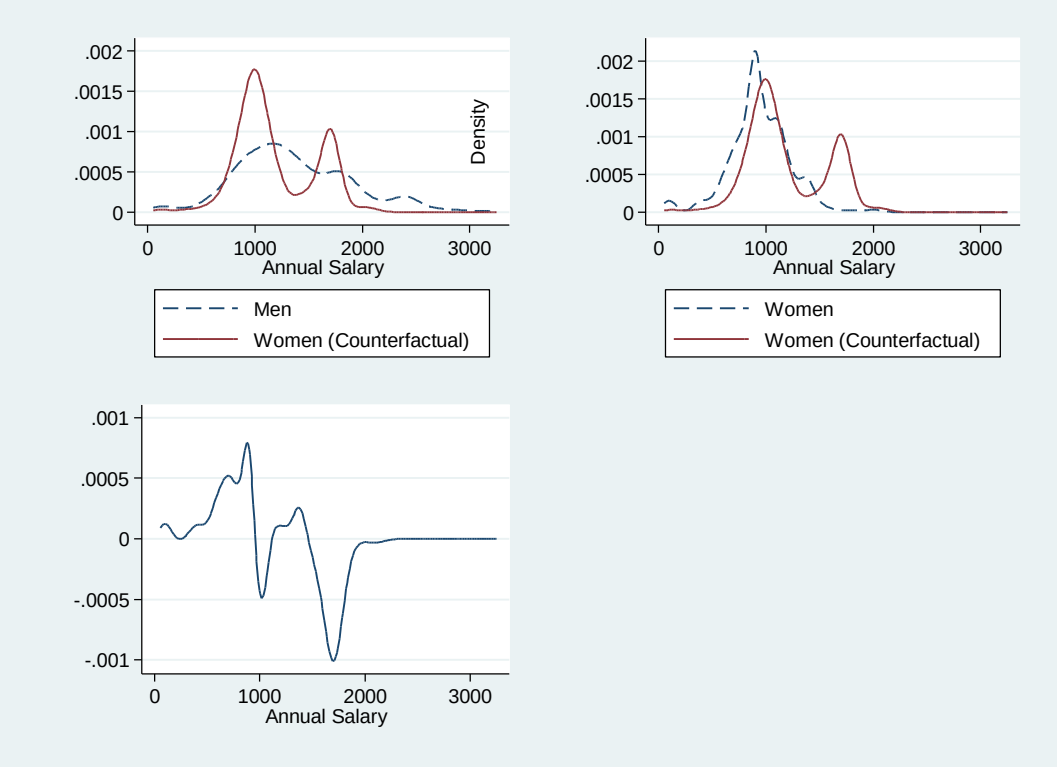
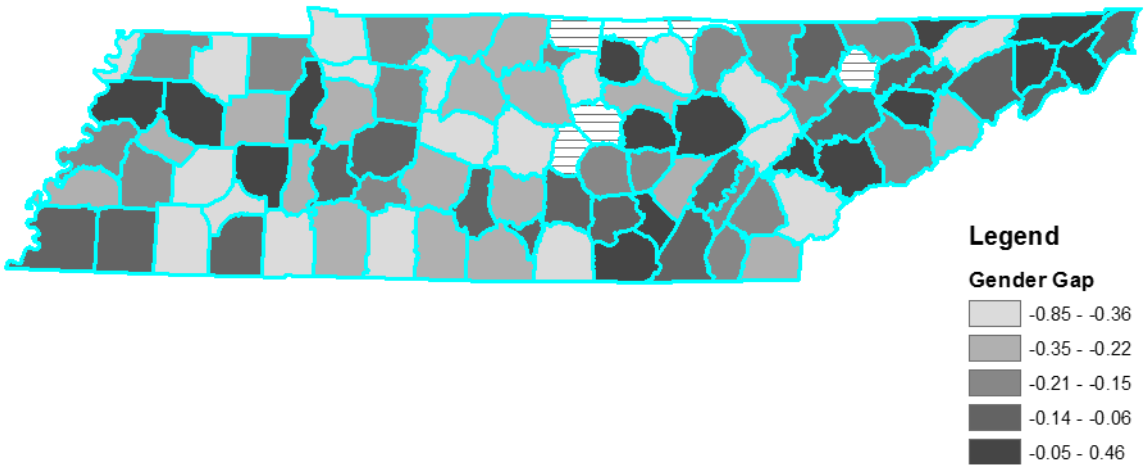


FIGURE 20: COUNTERFACTUAL DISTRIBUTION FOR WOMEN ASSUMING MALE CHARACTERISTICS



Notes: Counterfactual is performed using the method of Dinardo, Fortin, and Lemieux (1996). The first stage uses a logit regression to estimate the likelihood being male based on underlying characteristics. The characteristics in this regression are: principal status, high school teaching experience, high school teaching experience squared, tenure at the current school, tenure at the current school squared, educational background (college, normal school, high school, college and normal school), courses taught, and school class. Women are then assigned a distribution of wages assuming their distribution of characteristics is the same as men's. The top left graph shows the actual men's salary distribution (dashed line) and the counterfactual women's salary distribution (if they had a distribution of underlying characteristics that were the same as men's). The chart on the top right has the women's actual distribution and the same women's actual counterfactual as in the top left chart. Finally, the lower left chart shows the difference in the distributions in the top right chart.

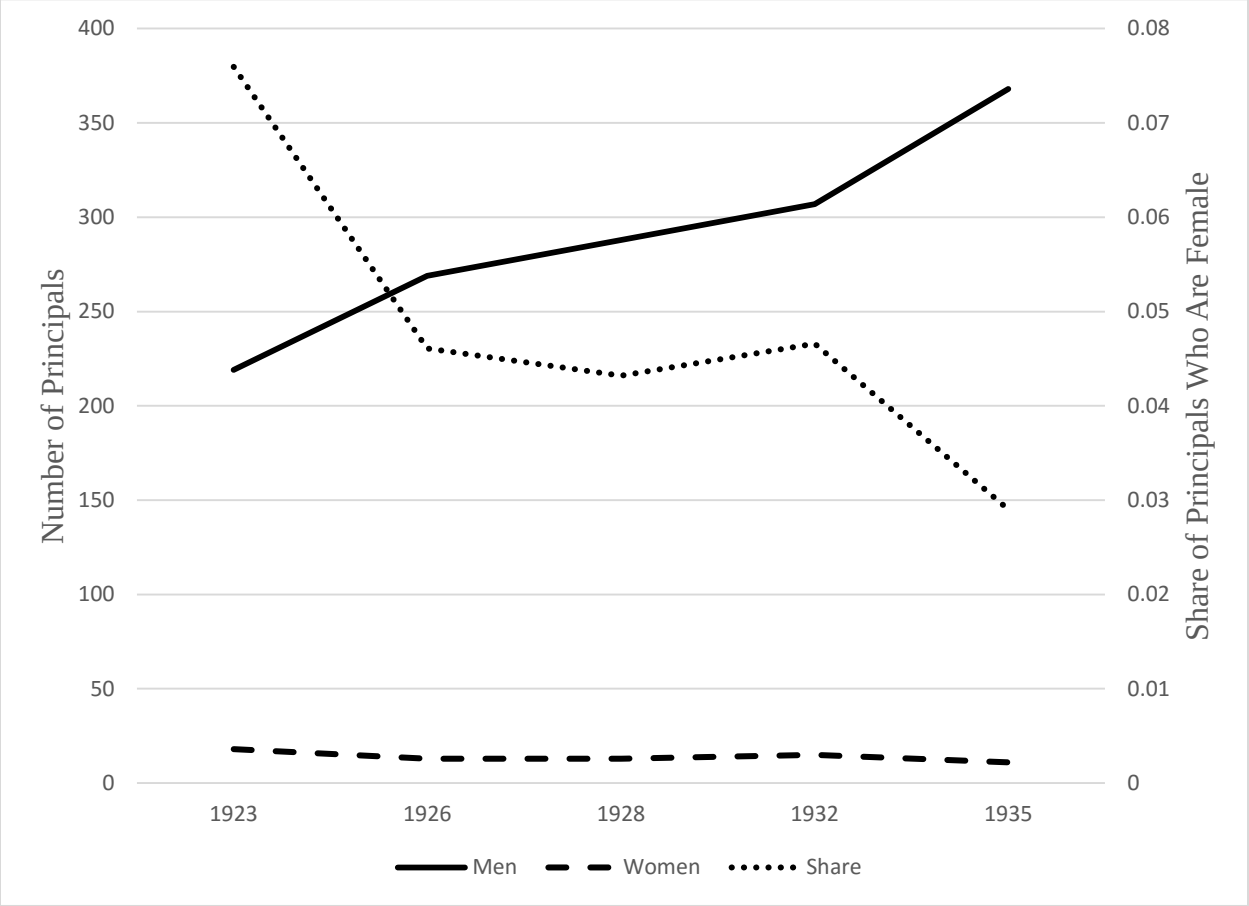
FIGURE 21: GENDER WAGE GAP, BY COUNTY



Source: Created using ArcGIS from data contained in the Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

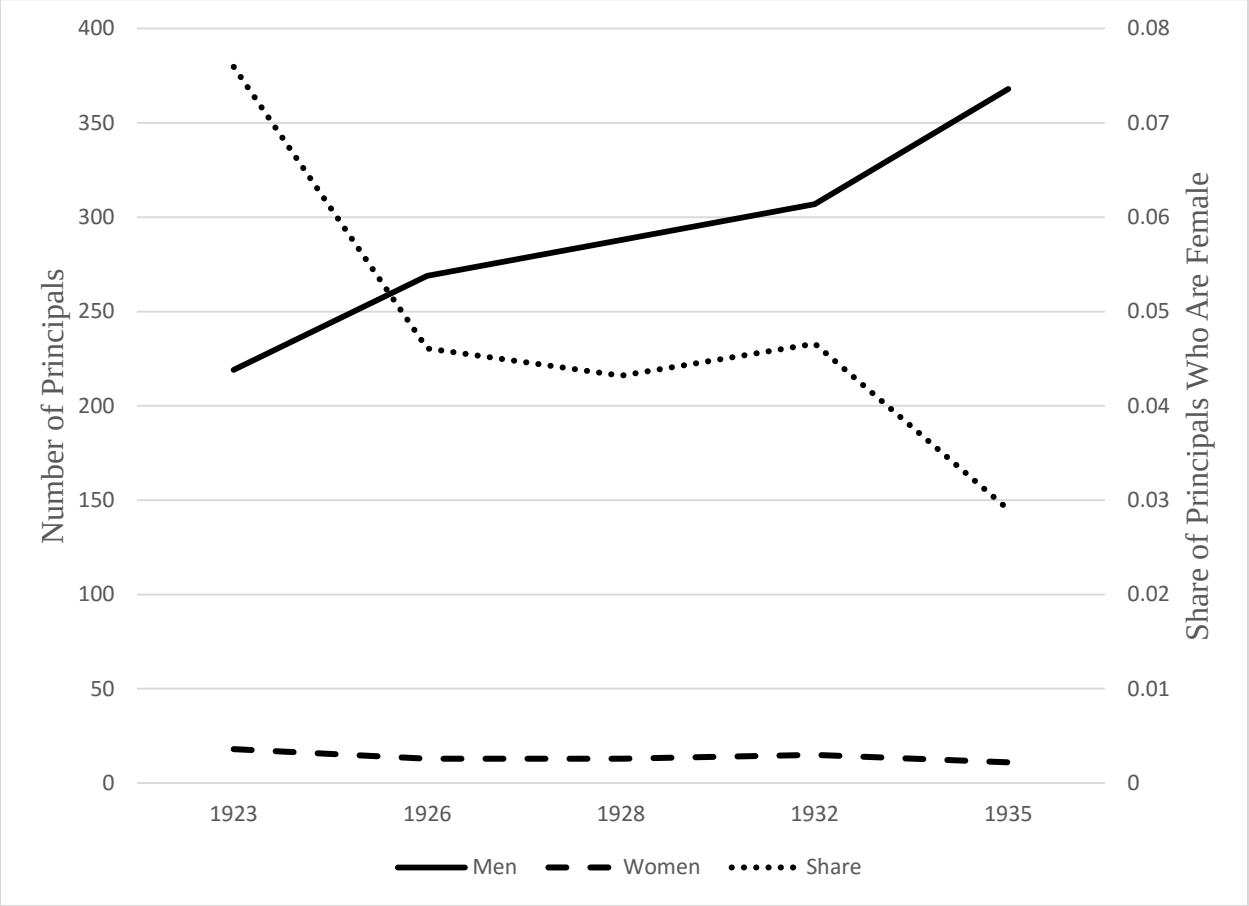
Note: The gender wage gaps above are the result of a regression of log wage on county fixed effects, gender × county, and the control variables in column (4) of Table 10. The coefficient on gender × county is plotted above. The greyscale represents different deciles of the calculated gender gaps. Counties with lines through them did not have enough data to calculate a gender gap. Using column (2) or (3) of Table 10 changes the numbers but not the general patterns.

FIGURE 22: NUMBER AND MEN AND WOMEN AND SHARE OF WOMEN SERVING AS PRINCIPAL, TENNESSEE, 1923-1935



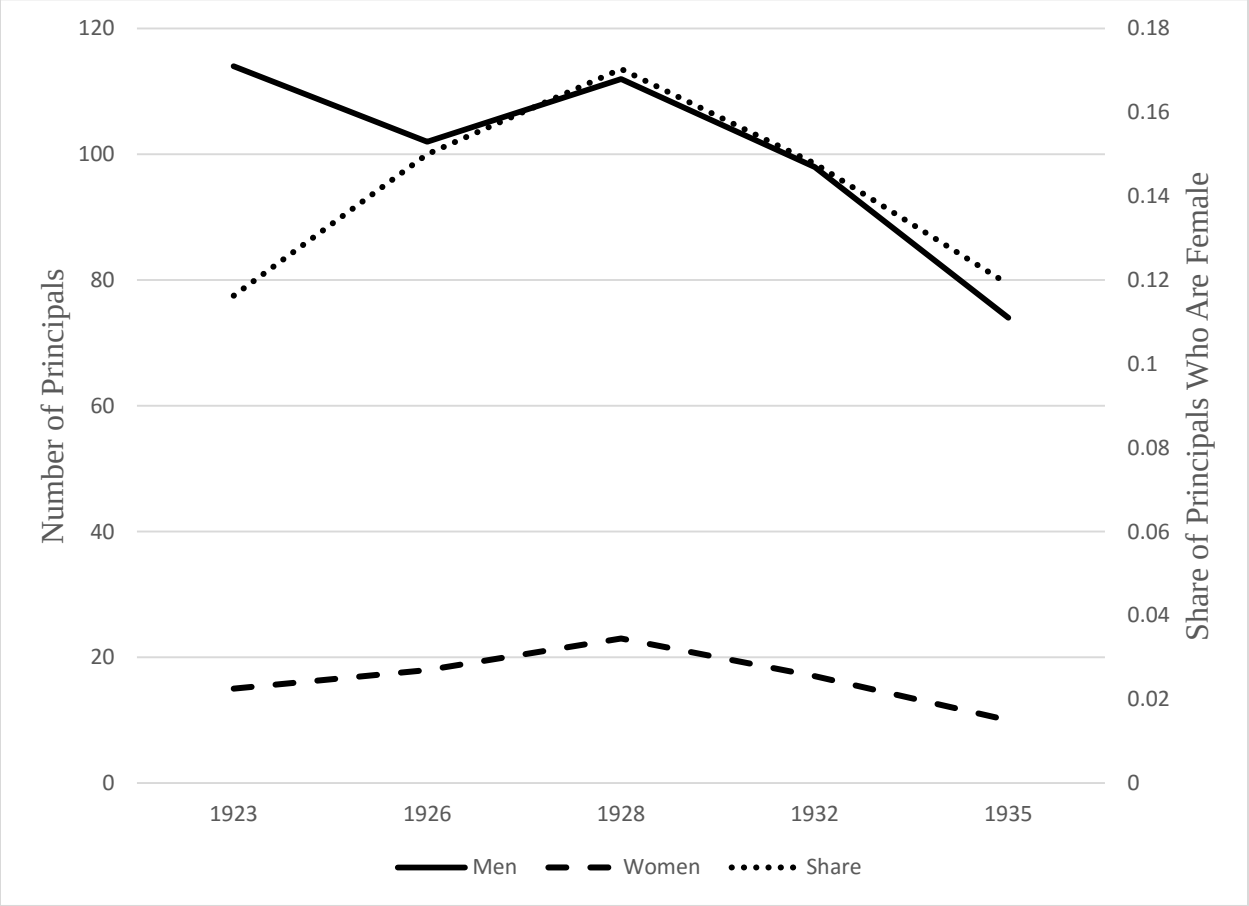
Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

FIGURE 23: NUMBER AND MEN AND WOMEN AND SHARE OF WOMEN SERVING AS PRINCIPAL, FIRST CLASS SCHOOLS, TENNESSEE, 1923-1935



Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

FIGURE 24: NUMBER AND MEN AND WOMEN AND SHARE OF WOMEN SERVING AS PRINCIPAL, THIRD CLASS SCHOOLS, TENNESSEE, 1923-1935



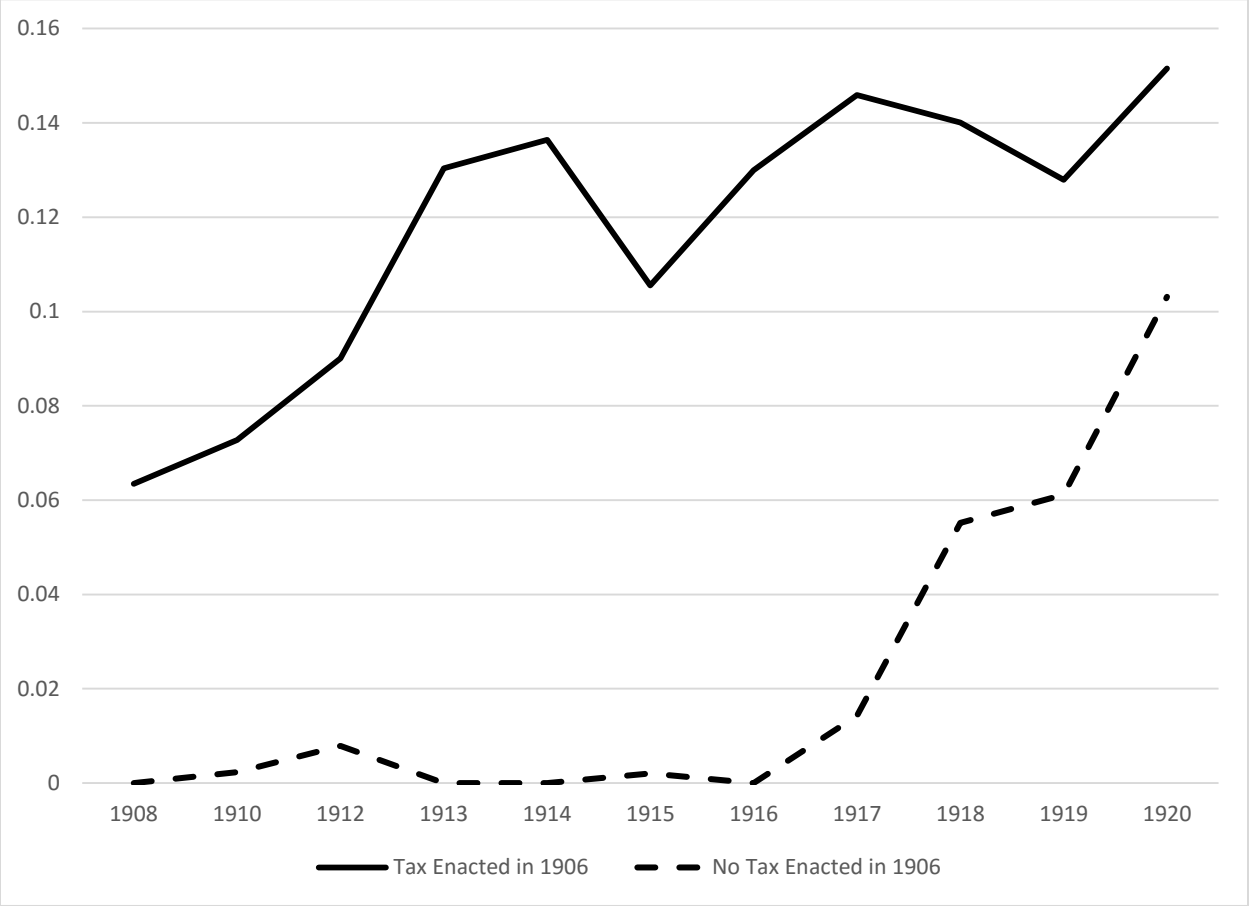
Source: Tennessee State Library and Archives. Tennessee Department Education Records, 1874-1984. Record Group 273, Box 64.

FIGURE 25: FARMERS AND WHITE COLLAR WORKERS IN TENNESSEE, BY EXISTENCE OF A HIGH SCHOOL TAX PRIOR TO 1917, 1900 TO 1940

Source: IPUMS (Ruggles et al. 2010)

Notes: Numbers represent ratio of Tennessee-born men in the labor force performing a job. Farmers are measured on the left hand access and skilled workers are measured on the right hand access. Farmers are defined as farm owner/operators. White collar workers are defined as those in either the professional, clerical, or managerial fields, as defined in the IPUMS 1920 1% sample. The four counties with major urban centers – Davidson, Hamilton, Knox, and Shleby – are omitted.

FIGURE 26: HIGH SCHOOL ENROLLMENT RATES, BY COUNTY, TENNESSEE, 1908 - 1920



Source: Tennessee State Superintendent of Public Education Annual Report

Notes: Enrollment rate is calculated as the number of students enrolled in high school divided by children 15-18 in a county. Since the reports only give children 6-21 in a county, children 15/18 is calculated as the 5/16 of children 6-21. Five values were missing in the records and were imputed using straight-line estimation from the two values immediately before and after the missing value. These were: Johnson County, 1914, Montgomery County, 1914, Sequatchie County, 1913, Tipton County, 1913, and Washington County, 1910. The four counties with major urban centers – Davidson, Hamilton, Knox, and Shleby – are omitted