The Birth of an Occupation:

Professional Nursing in the Era of Public Health*

Anthony Bald†

October 18, 2024

Click here for the latest version

Abstract

Over the past century, nursing was transformed from untrained menial labor into a professional, credentialed occupation. This paper studies the role of hospital-provided training in this transformation. In the early 20th century, hospitals founded training schools for nurses to meet the growing demand for medical care. Using linked census records and school openings as a source of variation, I find that white women who were geographically close to an opening in adolescence were more likely to become nurses in their twenties. Effects are largest for women from professional backgrounds, consistent with the standards that schools used to select students. Proximity to nurse training did not raise labor force participation or improve occupation-based measures of economic status. Furthermore, by their thirties, women proximate to training schools were less likely to work as physicians. These results paint a mixed picture: Nurse training elevated the status of nursing by changing the composition of nursing students, however, it also reinforced gender segregation within the medical field.

JEL Classification: I25, J24, J44, N31

*I thank Marcella Alsan, Christopher Avery, Peter Blair, Marianne Bitler, Amy Cross, Claudia Goldin, Brian Johnson, Samantha Batel Kane, Martin Karlsson, Savannah Noray, Bitsy Perlman, Kelsey Pukelis, Emma Rackstraw, Melissa Thomasson, Richard Zeckhauser, and participants in meetings of the Economic History Association, the Essen Health Conference, the Harvard Economic History Workshop, the Harvard Economics and Social Policy Workshop, and the NBER Summer Institute for thoughtful comments. This research was supported, in whole or in part, by the Institute of Education Sciences, U.S. Department of Education, through grant R305B150012 to Harvard University, and by the National Institute on Aging, through grant T32-AG000186 to the National Bureau of Economic Research. The opinions expressed here do not represent the views of the Institute, the U.S. Department of Education, the National Bureau of Economic Research, or Harvard University. All errors are my own.

†Harvard University, Cambridge, MA. Contact: abald@g.harvard.edu.

1 Introduction

> *No single change transformed the hospital's day-to-day workings more than the acceptance of trained nurses and nurse training schools, which brought a disciplined corps of would-be professionals into wards previously dominated by the values and attitudes of working-class patients.*
>
> - Rosenberg (1987), *The Care of Strangers*, pg. 344

Workforce training is a key driver of human capital growth and economic development (Ma et al. 2024). Much of this training is provided privately by firms, in the form of on-the-job training and apprenticeships. Training can be mutually beneficial to both workers and firms (Becker 1964; Acemoglu and Pischke 1998, 1999a). For the worker, training represents an investment in specific human capital and often comes with formal credentials. For the firm, training increases productivity and provides a mechanism for adapting its workforce in the face of technological change.

We have little evidence on how the establishment of firm-sponsored training affects the outcomes of individual workers in the long run. The effects of training can emerge over long time horizons, which poses a challenge for both data collection and the generalizability of short-term findings. Furthermore, causal effects of training are challenging to measure (Muehlemann and Wolter 2020). Take-up of training in any given occupation is determined by individual preferences and ability, as well as external factors like parental occupation (Greenberg et al. 2024).

This paper studies the birth of training for nurses in the United States. In 1870, nursing was considered low-status menial labor, employing primarily women from poor economic backgrounds. Inspired by Florence Nightingale's regimented approach to nurse training pioneered in England, the number of hospital-based training schools grew from 15 in 1880 to over 1,500 in 1920 (Figure 1). The establishment of these schools transformed the occupation of nursing and provides a canonical example of firm-sponsored training in U.S. history. Growth of training mirrored growth of nursing in the female labor force. By 1930, nursing had become the ninth largest occupation for young women, employing 4 percent of women ages 20 to 29. Today, nursing is the second most common four-year degree and one of the largest professional occupations.[1]

[1] Of the 1.2 million bachelor's degrees granted to women in 2021, 15 percent were in business and 11 percent were in registered nursing (NCES, 2021). In the 2022 American Community Survey, registered nurses made up 4.4 percent of employed women, the most common occupation code for this group. Labor force statistics throughout this paper are calculated from the census (Ruggles et al. 2022).

This paper asks how the spread of training affected selection into nursing and labor market outcomes for women. In an era when few professional careers were open to women, nurse training schools provided entrance into a certified occupation and work in the burgeoning field of public health.[2] However, several countervailing forces may have hampered the ability of women to benefit from training. Nursing, like many occupations, was not welcoming to married women in the early 20th century, and most trained nurses left the labor force upon marriage (Goldin 1991). Furthermore, contemporaries were concerned that nurse training did not guarantee adequate employment or compensation, especially during economic downturns (Reverby 1987; Whelan 2021).

To study the expansion of nurse training, I build a new dataset of nurse training schools that opened in the U.S. between 1870 and 1920. The dataset contains precise geographic location and year of organization, which I use to measure access to training for white women born between 1885 and 1900. These birth cohorts experienced a rapid increase in the availability of nurse training as schools opened across the country. To study outcomes for these women, I use newly released cross-census links (Buckles, Haws, et al. 2023). These links allow me to follow women from childhood to adulthood, including women who marry and change surnames. In particular, I link women observed in the 1900 census as children to the 1920 and 1930 censuses, when they were aged 20-35 and 30-45.

To identify the causal effects of training, I leverage geographic proximity to school openings and the age requirements of schools, which admitted women between the ages of 18 and 21. In a two-way fixed effects strategy, I compare women within five miles of a training school opening at ages 16 to 20 with comparable women who were slightly further away or slightly older. This approach allows me to answer two related questions. First, how did access to nurse training affect entry into nursing for women of different socioeconomic backgrounds? Second, what effects did training have on overall labor force participation, occupational sorting, and family outcomes?

First, I examine how access to training affected women observed in 1920. Women who were geographically close to a school opening in adolescence were more likely to become trained nurses (an effect size of 20 percent relative to the sample mean). The effect of a proximate opening is largest for women from relatively higher-status families, as proxied using father's occupation. This

[2] Some nurses spent entire careers in medicine and public health. Prominent figures include Clara Barton, Linda Richards, Lilian Wald, and Mary Eliza Mahoney.

aligns with Rosenberg (1987), suggesting that training schools elevated the status of nursing through their standards of admission. Notably, I find no evidence that training affected overall labor force participation or marriage rates in 1920.

Next, I examine how access to training affected occupational sorting and economic status. Sorting patterns reveal that women substituted away from a wide range of common occupations in the short-run, including teacher and bookkeeper. My analysis of economic status is limited by the lack of wage information in the historical census. However, using both absolute and relative measures of occupational status, I find largely null effects of access to training. This result is partially driven by the take-up of nursing by higher-status women, who plausibly had access to well-earning alternative occupations.

Finally, I examine persistence of effects for women observed in 1930 at ages 30-45. Effects on nursing are smaller in magnitude, consistent with the fact that over 70 percent of women in the sample were married and many no longer worked. However, women who were proximate to nurse training schools were both more likely to be married and more likely to work as nurses conditional on marriage. This result suggests that training increased attachment to nursing in mid-life. I observe negative effects on the likelihood of becoming a physician. While the overall rate of women working as physicians is low, the prevalence of nurse training seems to have further lowered entry into the role. Given that 95 percent of physicians in 1930 were men, I interpret this estimate as evidence that the expansion of nurse training reinforced existing gender segregation in the medical field.

This paper contributes to several strands of literature. First, my findings add to a broad literature on the history of women's work (Costa 2000; Goldin 2021). I provide evidence on how training availability affected work in a period when employment opportunities for women were highly limited (Goldin 1980, 1986b, 1990). While nurse training does not appear to have raised labor force participation significantly, it induced entry into nursing primarily for women from relatively higher-status backgrounds, consistent with efforts of schools to improve nursing's social standing (Rosenberg 1987). Exposure to nurse training increased the likelihood of married women working in nursing, which was uncommon in the 1930s (Goldin 1991). Additionally, this paper contributes to a growing number of papers that link women over time in the historical census to measure long-run outcomes (Withrow 2021; Eriksson et al. 2023; Buckles, Price, et al. 2023; Li 2023; Espín-Sánchez et al. 2023; Bailey and Lin 2024; Althoff et al. 2024; Feigenbaum and Gross 2024). I show limited

evidence that nurse training raised the economic status of employed women. This suggests that firm training was not an engine for upward mobility, consistent with the conclusion of Buckles, Price, et al. (2023) that single women were historically less upwardly mobile than single men.

Second, I contribute to a growing body of research on the effects of vocational education and firm-provided training (Muehlemann and Wolter 2020). Modern-day estimates suggest positive returns to vocational education and public training, especially in the healthcare sector (Grosz 2020; Silliman and Virtanen 2022; Ng and Riehl 2024; Xu et al. 2024).[3] We know less about whether these returns are also present in firm-sponsored settings, especially for new forms of training. This paper is the first to estimate the causal effects of access to firm-sponsored training in the early 20$^{\text{th}}$ century U.S. I employ distance to schooling, which is more commonly used as an instrument for general education (Card 1995; Rouse 1995; Kling 2001; Cameron and Taber 2004; Carneiro et al. 2011; Nybom 2017; Mountjoy 2022).[4] I find that substitution across occupations is key to understanding the effects of nurse training on economic status. This analysis is similar in nature to prior paper that examine selection and returns to field of study (Altonji et al. 2012; Altonji et al. 2016; Kirkeboen et al. 2016; Lovenheim and Smith 2022).

Third, I inform research on the causes and consequences of occupational sorting. Broadly, occupation is determined by many related factors, including perceived earnings, choice of college major, and non-pecuniary factors, all of which can differ by gender (Zafar 2013; Arcidiacono et al. 2020; Sloane et al. 2021). Few papers have studied exogenous variation in occupation (Emran and Shilpi 2011; Greenberg et al. 2024). I identify a setting in which the spread of firm training shifted women into nursing, and I use this variation to estimate effects on economic status and occupation, including the likelihood of becoming a physician. These findings shed light on the historical role of firms in patterns of occupational segregation by gender (Blau and Hendricks 1979; Goldin 1986a; Bertaux 1991; Card et al. 2016; Cortés et al. 2024). Such segregation has been shown to have broad consequences for the gender wage gap and the gap in meaningful work (Turner and Bowen 1999; Blau and Kahn 2017; Burbano et al. 2023, 2024).

[3] A broader literature examines the returns to general education at different points in history (Griliches 1977; Blundell et al. 1999; Goldin 1999; Kane and Rouse 1995; Feigenbaum and Tan 2020; Clay et al. 2021).

[4] My approach is similar to papers that estimate effects of access to general education on historical outcomes like mortality and the production of notable people (Connolly 2021; Doxey et al. 2022).

This paper proceeds as follows. Section 2 describes the history of nurse training schools and the growth of professional nursing.[5] Section 3 outlines the data sources and sample construction. Section 4 describes the empirical approach. Section 5 presents short-run results for the 1920 sample. Section 6 presents long-run results for the 1930 sample. Section 7 concludes.

2 Nurse Training in the Era of Public Health

In this section, I describe the rise of nurse training beginning in the 1870s.[6] I discuss the structure of nurse training schools and the employment prospects of trained nurses. I then provide historical context on certification for nurses and the broader growth of nurses in the labor force.

2.1 The Rise of Nurse Training

Before the widespread adoption of nurse training, hospitals relied on untrained stewards and attendants to care for patients (Whelan 2021). Work in the hospital wards was low-status and employed men and women from working-class backgrounds. Hospitals provided no formal training in the medical sciences, in part because worker attrition was substantial (Rosenberg 1987). In this period, the prevailing sentiment was that women were more suited to nursing than men, however, men still found work in hospitals and typically earned higher wages.[7]

The Civil War in the U.S. and the Crimean War in Europe revealed the inadequacies of hospital organization, resulting in calls for reform. Florence Nightingale's approach to nurse training, pioneered in England in the 1860s, quickly became known to American hospital reformists (Abel-Smith 1960). Nightingale's innovation was a year-long course in the art of nursing, holding trainees to high standards of discipline, cleanliness, and professionalism (Hawkins 2010). The first U.S. training schools that followed the Nightingale model were established in the 1870s, including the Bellevue Hospital Training School in New York City, the Connecticut Training School in New Haven,

[5] Throughout this paper, I use the terms "professional" and "trained" interchangeably.

[6] See Appendix A for details on the history of nursing. This paper draws on a large body of literature in history and nursing: Abel-Smith (1960) provides an overview of nursing in England. For more contemporary discussions on nursing in U.S. history, see Rosenberg (1987), Reverby (1987), Hine (1989), Kalisch and Kalisch (2004), D'Antonio (2010), and Whelan (2021).

[7] We have little systematic evidence on hospital wages by gender in the pre-professional era. Rosenberg (1987) reports that at New York Hospital in 1840, the Board of Governors raised wages for male nurses in an effort to attract men of higher quality, but kept wages for female nurses the same and forbid them from earning outside income.

and the Massachusetts General Hospital School of Nursing in Boston.[8] It wasn't until the 1890s—20 years later—that hospitals adopted nurse training schools in large numbers (Figure 1). This lag was partially due to capital costs (housing for trainees) as well as concerns over giving women autonomy in the hospital (Rosenberg 1987; Comin and Hobijn 2010).

The nurse training school was the product of two driving forces: the rising standards of modern medicine and the entrance of single women into the formal labor force. The modern hospital required a trained nursing staff to uphold hygienic practices and to care for the growing number of middle-class Americans frequenting hospitals (Thomasson and Treber 2008). Advances in medical technology increased the value of nurses by freeing physicians to learn new tools and perform more complex procedures. For instance, hospitals began stocking laboratory equipment to aid physician diagnoses in the 1880s.[9] During the Era of Public Health (1890-1936), large investments in infrastructure and public health further increased demand for nurses (Costa and Kahn 2006).

Hospitals marketed nurse training as an opportunity at a time when young women had few viable career options. In 1870, before the proliferation of nurse training, a single woman most likely found work in manufacturing or domestic service (Goldin 1980). Women graduated from high school at increasing rates in the early 1900s. Many of these women worked in the teaching and clerical sectors before dropping out of the labor force upon marriage (Goldin 1992; Goldin and Katz 2011). In comparison to a college degree, a nurse training program offered several years of paid employment, lectures in medicine, and the possibility of certification in nursing.[10] Demand for training was high, and prestigious training schools received more letters of interest than they had available seats.[11] The standards set by training schools in this period shifted the image of nursing from low-status domestic labor to a professional occupation for women.

---

[8] A few hospitals had structured training programs prior to the 1870s. The Philadelphia Lying-in Charity began training nurses alongside physicians as early as 1832. Credit for the early founding of the Philadelphia Lying-in Charity is given to Dr. Joseph Warrington, who organized the training of nurses in obstetrics and general medicine (Huzza 1887). In the 1860s, the Woman's Hospital in Philadelphia affiliated with Woman's Medical College to support the training of female physicians (Peitzman 2000).

[9] Physicians may have also substituted toward leisure. This 1873 excerpt from the *Boston Medical and Surgical Journal* is revealing:

> *There is no reason why the long list of surgical dressings, the application of splints, etc., should not be done by trained nurses, and the hurried surgeon be given the opportunity of eating his dinner in comfort and getting to the polls at election time.* - Putnam (1874)

[10] I provide more details comparing nursing with other common occupations in Appendix Section A.2.

[11] Bellevue Hospital Training School reportedly received 117 applications but only accepted 29 for its initial probationary period in 1875.

2.2 The Nurse Training School

> *The object of this Society is the training of nurses for the sick in order that women shall find a school for their education and the public shall reap the advantage of skilled and educated labor.*
>
> - Articles of Incorporation, Bellevue Training School for Nurses

Nurse training schools had two primary goals. The first was to educate young women in the medical sciences. Nurses attended lectures by training school staff and hospital physicians. Educational standards varied across schools, and the most prestigious programs included extensive courses in general medicine. For instance, in 1883, Massachusetts General Hospital provided lectures on diseases of women and children, surgical dressing, anatomy, anesthesia, and care of the newborn child. Such education would have been largely inaccessible to women outside of the nursing field. The professionalization of medical education during this period made it more difficult for women to train as physicians; Moehling et al. (2019) show that the proportion of female physicians fell from 1900 to 1940. Nurse training schools were therefore the primary venue for women to work in the medical sector. The second goal was to train student nurses to provide skilled nursing care to patients. To that end, student nurses spent much of their training in the wards of an affiliated hospital, sanitarium, or asylum. Hours were long, and as compensation, schools provided housing, uniforms, textbooks, and a small monthly stipend.

The Nightingale model fits well within the canonical framework of firm-sponsored training (Pigou 1912; Becker 1964). Much of nursing involves general skills (e.g., caring for bedridden patients, dressing wounds, and sterilizing equipment). Hospitals have no incentive to invest in such skills when the labor market is competitive. This is because they must pay trained nurses the market wage, and so there is no way to recoup the cost of training. However, labor market imperfections can make it worthwhile for hospitals to invest in students (Acemoglu and Pischke 1999a,b). In this setting, there is historical evidence of asymmetric information: hospitals privately observed the most productive students and hired them upon graduation (D'Antonio 2010). The rest of the graduates entered the market for hospital jobs or private nursing services. Another feature of this setting was the inability of student nurses to ensure high-quality training through contracting. Clinical exposure (i.e., on-the-job training) was a core component of the Nightingale model, so work in hospital wards was expected. However, hospitals frequently took advantage of student labor. For instance, Reverby

(1987) documents the practice of sending nurses into the community, where they would charge for their services and earn revenue for the hospital.

To ensure a minimum standard of care, training schools screened applicants based on educational background, physical health, and character.[12] Most schools required a short probationary period of one to three months. Several reports published by the Bureau of Education provide detailed information on nurse training school admissions. In the 1880s, the most common educational requirement was "a good common school education". High school was not explicitly mentioned, as this was decades before the high school movement took place (Goldin 1998). Schools in 1880 listed the age range that they would consider, most commonly ages 22 to 35 (Figure A.6). It is unclear how binding these conditions were without detailed records from individual training schools. Regulations on child labor were uncommon before 1900, and so the minimum age for training would have been at the school's discretion (Hindman 2002).

Over time, schools reported lower ages of admission and more precise educational requirements. At the median, the reported age minimum was 21 in 1896, falling gradually to 18 by 1927. To assess the actual working ages of nurses in training, I examine the age distribution of all women who received the 1950 occupation code for "student nurse". The distributions in each census from 1900 to 1930 are shown in Figure A.7. The majority of nurses were in training at ages 18 to 24, and I use this fact in my empirical strategy in Section 4. The left-ward shift in the distribution over time reflects the gradual adoption of lower age minimums. This shift was driven by several factors, including nurse shortages and school preferences (Reverby 1987). The Bureau of Education did not report educational requirements again until 1911, when 24 percent of schools required a common school education, 39 percent required a high school diploma, and the remainder fell in between. By 1918, all schools required at least an 8^{th} grade education. By 1927, the breakdown of required education was 2 percent requiring 8^{th} grade, 64 percent requiring some high school, and 31 percent requiring high school. Very few programs listed postsecondary education as a requirement.

After spending two to three years in training (Figure A.8), graduates of nurse training schools could enter a number of roles in private or public health. In Appendix B, I discuss the employment and marriage patterns of graduates from three schools with digitized archival records: the Bellevue Hospital Training School for Nurses in New York City, the Presbyterian Hospital School of Nursing

[12] Prospective nurses were asked detailed personal questions and often needed references (Figure A.2).

in Chicago, and the Massachusetts General Hospital School of Nursing in Boston. In 1900, the most common role for graduates was private duty nursing, which involved providing nursing care to sick individuals in their homes. As hospitals expanded during the early 20th century, nurses became more likely to work in hospitals or on the staff of training schools. Public health spending by states and cities led to the expansion of the public health nurse, infant nurse, and visiting nurse roles (Miller 2008; Hoehn-Velasco 2021). Finally, graduates of nurse training schools often became certified, which allowed them to practice as registered nurses under state law.

2.3 Certification Laws for Nurses

The rise of nurse training programs was concurrent with the growth of certification laws regulating titles such as "registered nurse" (R.N.).[13] The first state to enact a law regulating the practice of nursing was North Carolina (Shannon 1975). As of March 3, 1903, nurses in North Carolina were required to obtain a certificate to practice under the title of registered nurse. All states adopted some form of certification law for registered nurses between 1900 and 1940. Nurse certification laws from this period were permissive, i.e., nurses without certification could still practice so long as they did not use the title of R.N. It wasn't until the mid-20th century that states passed mandatory legislation, which prevented unlicensed nurses from working at all (Law and Marks 2013, 2017). In Appendix Section A.3, I provide more details on certification laws, including provisions contained in the laws and predictors of state passage.

Certification laws had implications for both training schools and the composition of the nursing profession. Certification laws often required training at an accredited or state-approved program, which put pressure on training schools to improve or face closure. D'Antonio (2010) uses the example of Georgia, which passed a law regulating nursing in 1907 and tightened standards throughout the 1920s. Small, rural hospitals, often run by black physicians, were vulnerable to rising standards.[14] Nurse training programs for nurses were distinct from medical schools, which

[13] I distinguish between "certification" laws, which were generally voluntary, and "licensing" laws, which were mandatory and legally barred unlicensed individuals from practicing in an occupation. Many laws before 1940 did not require a license, but rather required an individual to obtain specific qualifications or certifications to use a protected title. These laws were also distinct from "registration" laws that require an individual to register in an occupation, without proof of competency (Carollo 2024).

[14] Similar anecdotal evidence comes from Nutting (1912):

> *It appears from the statements of officers of various State societies that a number of small, private, or special hospitals or sanatoria which had been maintaining schools have been sufficiently affected*

were subject to scrutiny in the 1910s as a result of the Flexner report (Clay et al. 2023). Nursing was not a subject of Flexner-era reform, so efforts to improve the quality of nurse training were more diffuse (Whelan 2021).

Certification laws created barriers in nursing that disproportionately affected minority nurses (Whelan 2021). Healthcare provision at the turn of the 20th century was racially segregated, which was legal under *Plessy v. Ferguson* (1896). Once states adopted nurse certification, black graduate nurses in many southern states were not allowed to sit for certification exams or received a separate "Negro" certification to distinguish them from white nurses (Hine 1989). These barriers compounded such that, conditional on receiving state-recognized training, black nurses did not achieve professional recognition. Black nurses were denied membership in nursing associations, received lower salaries, and were often treated poorly by members of the medical profession. These historical accounts suggest that certification laws furthered the divide between white and black nurses.

2.4 Growth of Nurses in the Labor Force

In this section, I characterize the growth of nurses in the decennial census, both in raw numbers and as a share of the labor force from 1870 to 2015. Tracking the growth of trained nurses is challenging because the concept of a nurse did not stay constant over this long period. This resulted in changes over time in the subcategories assigned to nurses in the census. For instance, it was not until 1910 that practical (untrained) nurses received an occupation code to distinguish them from professional (trained) nurses. The 1950 coding of occupations allows us to partially overcome these challenges by partitioning nurses into subgroups based on industry and text from the original census enumeration. For the 1870-1940 period, I follow the 1950 coding and categorize nurses with an occupation code of 58 as professional nurses. For 1950-2015, I use the codes for professional nurse listed in Table C.2. For additional details on the samples and definitions used in this section, see Appendix Section C.2.

In 1870, the number of women coded as professional nurses ages 15-64 in the census was just 50. Following the growth of hospital-based medicine and the establishment of nurse training, professional female nurses in 1920 numbered over 100,000 (Table D.1). It is informative to measure

by the laws to lead them either to seek suitable affiliations or to close the schools and carry on their work by means of a salaried nursing staff.

growth of all nurses, including trained, student, and untrained nurses.[15] The total number of nurses increased from 9,000 to 230,000 during this same period. Figure 2 plots the percentage of the labor force employed in professional nursing from 1870 to 2015. From 1900 to 1930, professional nursing grew from less than 0.5 percent to 2 percent of the female labor force, and by 1930 the majority of nurses were counted as professional.[16] Overall growth slowed in the 1930s during the Great Depression, but growth has since been steady. In 2015, 4 percent of employed women worked in professional nursing.

The growth of male and minority nurses is a recent development in the nursing field (Munnich and Wozniak 2020). Prior to the era of professionalization, the field was more diverse, with male nurses and black nurses making up 5 and 20 percent of all nurses, respectively (Figure D.2, Panel (a)).[17] In Figure D.2, Panel (b), I form a ratio of the percent in nursing and the percent in the overall labor force to assess each group's level of representation. A prominent U-shaped pattern emerges: representation of black female nurses fell from 1880 to 1940 as white women entered the profession in large numbers. Notably, representation improved in the 1960s with the Civil Rights movement and the broadening of access to nursing programs at colleges and universities.

3 Data Sources

In this section, I describe the dataset of nurse training schools, the records available from the complete-count census, and cross-census linking. See Appendix C for details.

3.1 Nurse Training Schools

I construct a dataset of nurse training schools by digitizing publications of the U.S. Bureau of Education and the American Nurses Association (henceforth the Bureau and the ANA). The Bureau was organized in 1867 and began publishing statistical reports on American education that year.

[15] Women may have received two other occupation codes: hospital attendant or midwife. These occupations were small relative to nursing: In 1880, there were 19,000 female nurses (Table D.1), relative to 900 hospital attendants and 2,000 midwives.

[16] Figure D.1 shows alternative measures of the percent of women and men working in nursing. Panel (a) shows that the share of women categorized as any type of nurse grew steadily to 4 percent in 1930. Panel (b) shows the rising share of men working in nursing since 1960.

[17] Men made up more than half of trained nurses in 1870. However, in this period, the distinction between trained and untrained nurses was not well-developed. Men may have been more likely to be counted as professional on account of their gender. The share of professional nurses who are male has increased to 10 percent in 2015 (Figure D.2, Panel (c)).

Nurse training schools were educational in nature, therefore the Bureau took an early interest in them. The first report on training schools was published in 1879, six years after the first schools opened. The Bureau published yearly reports until 1906 (excluding 1883) and less frequently thereafter. I digitize all reports with school-level information, including the reports from 1879 to 1906, as well as reports from 1911, 1918, and 1927 (Nutting 1912; U.S. Bureau of Education 1919, 1928). The available information includes school name, city, year of organization, number of students (by gender), number of graduates, and the duration of training. In addition to the Bureau, the ANA was formed in 1896 (originally titled the Nurses Associated Alumnae of the United States). I digitize two ANA reports containing accredited nurse training schools with similar information on location and year of organization (American Nurses Association 1918, 1922).

I use the digitized reports to count the number of nurse training schools in each city and town over time. I construct separate counts using the Bureau and ANA reports, so that the two measures can be compared. To construct yearly counts from the Bureau reports, I first drop "post-graduate" schools, which were short supplemental courses for training school graduates. These were listed separately from fully-fledged schools and make up less than 3 percent of observations in a given year. I map the remaining schools to cities and towns in the Census Place Project (Berkes et al. 2023). One concern with historical location data is that names and spellings can change over time. The Census Place Project standardizes city and town names across census years so that all schools are linked to a consistently defined location (latitude and longitude). Finally, the Bureau recorded each school's year of organization up until 1905. I use this field in the data to measure the number of schools in each location and year from 1870 to 1904. After 1904, I count the number of schools recorded in each location and report year (1905, 1906, 1911, 1918, and 1927). To construct an analogous count using the ANA records, I first drop data from the 1918 report, which contains substantial missingness in year of organization. Using the 1922 report, I drop "post-graduate" affiliations, which are listed separately. The mapping to the Census Place Project is identical. I then use year of organization to count the total number of schools in each location and year from 1870 to 1920.[18] For the remainder of my analysis, I use the count of schools based on the ANA's

[18] By starting the panel in 1870, I drop several training programs that were founded earlier. These programs were not based on the Nightingale model of training schools. See Section 2.1. I end the panel in 1920, the last year that I can reliably measure school openings in the American Nurses Association (1922) report.

1922 report since it contains better coverage of opening year than the Bureau reports. I provide a comparison of these two sources in Appendix Section C.1.

I use these data to plot the growth of nurse training schools and students in Figure 1. The total number of schools peaked in the late 1920s at just under 2,000 schools. Figure 3 plots the geographic distribution of nurse training schools in 1900, 1910, and 1920. Each point is a location, where point size is proportional to the number of schools. The earliest schools were affiliated with large urban hospitals, which explains the early concentration of schools in the Northeast. By 1900, 76 percent of the 100 most populous cities had a training school, compared to 63 percent of the next 100 populous cities. While some cities had ten or more training schools, many towns had just one by 1920.

3.2 Census Records

I use census records to construct my main analysis sample. The complete-count census is enumerated every ten years and covers the entire U.S. population. I primarily use the census files from 1900 to 1930, available via IPUMS (Ruggles et al. 2020).[19] The census contains demographic, household, and geographic characteristics, such as state and county. City identifiers are not universally available in IPUMS data, especially for individuals in smaller cities and towns. My approach is to link individuals to the Census Place Project, which assigns a standardized city or town of residence, latitude, longitude, and 2016 county (Berkes et al. 2023). Note that this approach does not incorporate within-town variation, so all individuals in a given town are assigned the same geographic coordinates. The match rate of individuals to towns is approximately 95 percent for the time period under study.[20]

I construct several labor market outcomes available from the census, including labor force participation and occupation. Prior to 1940, the modern-day definition of the labor force (i.e., the distinction between "employed", "unemployed", and "not in the labor force") was not fully developed. I define labor force participation as whether an individual is recorded with any gainful occupation, following prior literature (Katz and Margo 2014). To study occupational sorting, I use each individual's assigned 1950 occupation code. The 1950 occupation codes harmonize occupation

[19] These records exclude the 1890 census, which was unable to be salvaged following a fire in 1921.

[20] Note that I use the term "town" loosely. The set of geocoded places can include cities, town, as well as suburbs or neighborhoods within a given metro area.

groups across census years. This is particularly useful in the case of nursing, since nurse subcategories changed over time (see Appendix Table C.2). To study family outcomes, I use marital status and construct spousal characteristics for women who were married at the time of the census.

Historical records pose two challenges for studying the effects of nurse training. First, the 1940 census is the first to collect an individual's level of education, and no historical census contains information on the type of education received. As a consequence, it is not possible to observe which women graduated from specific nurse training schools. Second, the 1940 census is the first to collect wage information, which poses a challenge for historical research on income and economic status.[21] I use several proxies for economic status based on observable characteristics in the census. My preferred measure is based on Song et al. (2020), which examines intergenerational mobility by constructing a relative measure of economic status. The authors construct a percentile ranking of occupations within each birth cohort based on average human capital (literacy, or years of education when available). Ward (2023) adjusts this methodology to allow an occupation's percentile rank to vary by race and region. Buckles, Price, et al. (2023) further allow this rank to vary by gender. I construct an "adjusted Song score" following the methodology of Buckles, Price, et al. (2023) (see Appendix Section C.2.3). To measure absolute income, I assign occupations to their 1950 occupational income score, or the occupation's overall median earnings in hundreds of 1950 dollars. I also use the LIDO score, which allows median 1950 earnings to vary flexibly by state, gender, race, age, occupation, and industry (Saavedra and Twinam 2020). These measures are problematic, especially so for nursing, because they do not account for relative changes in occupational earnings over time. I therefore focus my analysis of economic status on the adjusted Song score.

3.2.1 Cross-census links for women

To study the effects of exposure to nurse training on later-life outcomes, I link women from childhood to adulthood. This linkage is necessary to observe precise geographic location as children and labor

[21] Prior to 1940, census enumerators asked about literacy and schooling status as a binary indicator. Some earlier datasets do contain information on education and wages. The 1915 Iowa state census is the largest such dataset, which prior studies have used to estimate returns to education and economic mobility (Goldin and Katz 2000; Feigenbaum 2018).

market outcomes as adults. Without this linkage, the only information on birth location in the census is birth state (or country).[22]

I use cross-census links derived from the FamilySearch genealogy platform to link women from childhood to adulthood (Price et al. 2021; Buckles, Haws, et al. 2023).[23] The majority of women from this period married and changed surnames, which precludes the use of traditional linking methods that rely on first and last name as listed in the census (Abramitzky et al. 2021). FamilySearch users can draw on a variety of sources (birth, marriage, and death records, as well as private knowledge) to construct detailed family trees. Buckles, Haws, et al. (2023) use these trees to train an XGBoost machine learning algorithm and form additional links. By combining these links with other linked panels, such as the Multigenerational Longitudinal Panel, Buckles, Haws, et al. (2023) construct over 300 million links for women.[24] I use a subset of these links in my analysis, and I provide further details on the linking process in Appendix Section C.2.1.

3.3 Sample Construction

I identify women in the 1900 census who were born in the U.S. between 1885 and 1900. I use this sample for several reasons. First, the women in this sample was aged 15 or younger at the time of the 1900 census.[25] I can therefore use location in 1900 to measure proximity to nurse training school openings in childhood, before work or migration decisions are made.[26] Second, the number of nurse

[22] Prior studies have used state of birth and year of birth to estimate returns to schooling using state-level variation in compulsory schooling laws, without the need for cross-census links (Acemoglu and Angrist 2000; Lleras-Muney 2002; Stephens and Yang 2014; Clay et al. 2021). This approach will incorrectly assign laws for some individuals who move states during childhood. Alternatively, Goldin and Katz (2011) use contemporaneous high school enrollment to estimate effects of compulsory schooling laws. To identify location during schooling age, Moorthy et al. (2022) use longitudinal records from the Health and Retirement Study to match individuals to state of residence at age ten.

[23] As an alternative, Espín-Sánchez et al. (2023) and Althoff et al. (2024) link women over time using Social Security Administration (SSA) records matched to census. These records are primarily made up of Social Security Number (SSN) applications that record place of birth, date of birth, father name, and mother (maiden) name. While these linked samples have the advantage of containing place of birth, sample sizes using this method are generally smaller than comparable samples in FamilySearch.

[24] For more information, see here: https://censustree.org.

[25] The 1900 census was enumerated on June 1. The destruction of the 1890 census means that 1900 is the first year that we observe these cohorts in surviving census records.

[26] One challenge faced in prior literature is that precise location in childhood is not always observable. This can pose a problem for proximity-style instruments, sometimes termed "endogenous mobility bias". For instance, Currie and Moretti (2003) use mother's location at first birth to estimate effects of proximate college openings on health. If this location is endogenous (e.g., women move to areas with more openings, for educational or non-educational reasons), then instrumental variables estimates may be biased. To circumvent this issue, Connolly (2021) uses a combination of census and Social Security Numident death records to identify town of birth. The Numident records have poor coverage of individuals who died before 1970. Many of the women in my sample likely died before 1970, and so I do not employ this approach.

training schools grew from 1880 to 1920 (Figure 1). The 1885 to 1900 birth cohorts were exposed to the precise geographic and temporal variation in training that I use in my empirical strategy.

3.3.1 Linked 1900-1920 sample

I use the FamilySearch links to match the 1900 sample to the 1920 census (Buckles, Haws, et al. 2023). This link is the primary reason that I limit the sample to white women born in the U.S.[27] The 1900 sample contains 11,420,000 native-born white women. Of these, 6,331,000 (55 percent) are linked to the 1920 census via FamilySearch. These links include married women in 1920, who make up 60 percent of the linked sample. While this linkage is necessary to observe adult outcomes, it affects the interpretation of the results by limiting the sample. I follow Bailey et al. (2020) and construct propensity scores for the likelihood of linkage using observable characteristics in 1900. I use these propensity scores to assess the representativeness of the linked sample. Additionally, I use inverse propensity score weights in all regressions. See Appendix Section C.2 for details.

To obtain my main analysis sample, I calculate the straight-line distance from a woman's location in 1900 to nearby training school openings. I define a woman as "treated" if she was within five miles of a training school opening at ages 16 to 20. To form a comparable sample of women with relatively less access to nurse training, I restrict to women within ten miles of a training school opening at ages 16 to 25. By restricting the sample to women within ten miles, I drop women from small, rural areas that were not in close proximity to any nurse training schools. This approach is sometimes referred to as the "ring method" in studies that identify effects of spatial treatments (Linden and Rockoff 2008; Currie et al. 2015, 2024). Furthermore, by only including women who experienced openings at ages 16-25, the treatment group (ages 16-20) is compared to relatively older women (ages 21-25) who were less likely to be admitted to training. I provide more details in Section 4.2.

Table 1 reports descriptive statistics. Characteristics of childhood town, father, and mother are measured in the baseline year of 1900. Adult characteristics are measured in 1920. For reference, column (1) reports statistics for the full 1900-1920 linked sample, restricted to women with non-missing geocoded location (N = 6,004,123). Columns (2) and (3) report statistics for the final sample

[27] In principle, foreign-born women can be linked if they were already in the U.S. in 1900. I restrict to native-born women for ease of interpretation.

(N = 1,646,657), split by treatment status (e.g., whether a woman was less than five miles from a training school opening at ages 16-20). Relative to the full linked sample in column (1), treated women in column (2) are located in larger towns in the Northeast. They were also from families with a higher economic status in 1900, in terms of parental literacy and occupational income. By limiting the full sample to women within ten miles of an opening at ages 16-25, the sample of untreated women appears more similar to treated women (column (3)). The goal of the identification strategy in Section 4.2 is to control for any remaining differences.

The 1920 sample is observed at ages 20 to 35, which allows for the measurement of labor force and marriage outcomes early in life (when many women were still of training age). In 1920, the marriage rate for women in the sample was roughly 50 percent. Conditional on working, 2 percent of women in the sample worked in a nursing occupation. The use of the 1920 census is not without its limitations. It was enumerated in January, in the wake of World War I and the Spanish flu epidemic. Additionally, a few questions were not asked in 1920, including duration of marriage, number of children ever born, and children surviving. I use the number of children in the household to examine family size, but cannot include age at first marriage as an outcome for this sample.

3.3.2 Linked 1900-1930 sample

The construction of the 1900-1930 linked analysis sample is similar to that of the 1900-1920 sample. Of the 11,420,000 native-born white women in the 1900 sample, 5,530,000 (48 percent) are linked to the 1930 census via FamilySearch. This sample is then restricted to women within ten miles of an opening at ages 16-25 using the same set of openings as in 1920. That is, openings after 1920 are not considered because the data after 1920 is incomplete. Note that I do not restrict the 1930 sample to women observed in the 1920 sample. The primary reason is to improve power by maximizing sample size.

Table D.2 reports descriptive statistics. The final sample size is 1,326,420, slightly smaller than in 1920. The 1930 census captures outcomes for the linked birth cohorts at ages 30-45, after all education has taken place. This allows me to examine whether effects persist (or newly emerge) relative to 1920. The 1930 sample is largely similar to the 1920 sample in terms of observable characteristics, with the exception of marriage and employment. Whereas 40 percent of the sample

were employed in 1920, just 25 percent were employed in 1930. This drop in employment is consistent with the fact that over 70 percent of women in the 1930 sample were married.

4 Estimating Effects of Nurse Training School Proximity

In this section, I provide a conceptual framework for understanding the effects of proximity to nurse training. I discuss my identification strategy, followed by tests of identifying assumptions.

4.1 Conceptual Framework

I identify the impacts of nurse training school growth using variation in proximity to training. Access to nurse training schools in the early 20th century was determined by a number of factors, such as distance from home. Distance represented both a monetary travel cost and a non-monetary cost of locating farther from family and established social networks. Non-monetary costs were likely larger for women due to gender norms: Withrow (2021) shows that white female migrants in the 1920s and 1930s were less likely to migrate long distances than white men. Qualitative evidence suggests that distance mattered in the case of nurse training. Reverby (1987) examines the birthplaces of nurses training at three Boston-area training schools. She finds that for women who entered training in the 1900-1919 period, 25 percent were from greater Boston and 25 percent were from elsewhere in New England. The concentration of local women aligns with other literature that shows the relevance of distance for educational access in other settings (Card 1995; Long 2004; Doyle and Skinner 2016; Walters 2018; Acton et al. 2024).

The opening of a nearby nurse training school induced more women into training by lowering the costs associated with distance. Moreover, nurse training schools were cheaper than other forms of postsecondary schooling. In this period, a young woman could have attended a college or university (one that admitted women), which required tuition and boarding. Alternatively, nurse training schools provided housing and compensated students with a monthly stipend. The option to earn money while undergoing training was appealing, more so for women who were credit constrained (Withrow 2021). Nurse training may have even caused women to substitute away from careers with higher lifetime earnings but larger upfront costs (Hanushek et al. 2017). Motivated by this, I study how nurse training school access affected occupational sorting for women from different economic backgrounds.

Access to nurse training mattered differently for women of different age, race, and family background. In Section 2.2, I discuss how nurse training schools primarily admitted women ages 18 to 24 while they were still unmarried and without children. This made economic sense from the perspective of the hospital, but may have also made sense for the women themselves. For a woman older than 24, switching into nursing may have been costly, especially if she had already married or left the labor force. Black women were excluded from most nurse training schools regardless of background or age (Hine 1989). This leads me to examine results primarily for white women, and I discuss the sample selection further in Section 3.3. Finally, much of the rhetoric around nursing during the early 20^{th} century was focused on the profound shift in social status. The installment of structured training legitimized nursing as a profession and raised standards significantly. Women from higher-status families may have had an easier time meeting these standards. I therefore show results separately by father's occupation as a proxy for social status.

Along with nursing, I examine other outcomes, including overall labor force participation, entrance into alternative occupations, and economic status. Much of the allure of nursing was that it was a respectable path in which a woman could earn independent income. Training and obtaining a nursing credential may increase overall labor force attachment by shifting attitudes or connection to work (Fernández 2013). Regarding alternative occupations, I examine whether women upskilled by substituted away from untrained nursing and hospital attendant roles. I also examine substitution from occupations that were common (teaching and clerical work). One concern at the time was the overproduction of graduate nurses coming from unregulated training schools. While graduates of elite programs received high-quality training, other nurses found themselves in low-quality programs and lacking work as private duty nurses (Whelan 2021). The challenges faced by graduates may have lowered the average return to training. A possible modern parallel is the expansion of for-profit institutions offering vocational certification, where evidence on short-run returns is mixed (Deming et al. 2012; Gilpin and Stoddard 2017; Soliz 2018). I cannot estimate effects on short-run wage income in historical data, so I study occupation-based measures. Conditional on working, effects are determined by occupational sorting.

4.2 Identification Strategy

I use a two-way fixed effects strategy to measure causal effects of proximity to nurse training schools. The goal of this approach is to use exogenous changes in access to nurse training to remove bias from selection into nursing. Women who entered the nursing profession during this period were likely different from those who did not. For instance, women in nursing may have had a propensity for caregiving, which affects job choice or propensity to drop out of the labor force.

I leverage the fact that nurse training schools spread across the country beginning in the 1870s (Figure 1). I compare women who were in close geographic proximity to a new training school with women who were relatively further away. I also take advantage of the precise admissions criteria of training schools, which typically admitted women between the ages of 18 and 21. I compare women who were relatively young at the time of an opening with women who were relatively older, but had the same proximity to training. This approach is similar to papers that use treatment "intensity" at different ages as a source of identifying variation (Duflo 2001; Baker et al. 2020; Kose et al. 2021).

Formally, consider woman i born in year t and observed in location j in childhood. Let y_{ijt} be the outcome for woman i observed in adulthood (for instance, employment as a trained nurse). Define Near_j to be an indicator for whether location j was less than five miles from a nurse training school opening (where miles is geocoded using straight-line distance). Let Age_{jt} be an indicator for whether a woman was aged 16 to 20 at the time of the opening. I estimate the following specification:

$$y_{ijt} = \theta_j + \delta_{c(j)t} + \beta \text{Near}_j \times \text{Age}_{jt} + \Gamma \mathbf{X}_i + \varepsilon_{ijt}. \tag{1}$$

The treatment effect of interest is β, the effect of proximity to a training school opening (distance less than five miles) at ages 16 to 20. θ_j is a set of fixed effects for childhood town in 1900, which captures time-invariant differences across places. For instance, choice of occupation might have been driven by differences across towns in proximity to large cities or historical presence of specific industries. I control for these differences to the extent that they stay constant across birth cohorts. Recall that the town indicators from the Census Place Project are fine-grained and capture differences across small geographies, such as suburbs outside a major city (Berkes et al. 2023). $\delta_{c(j)t}$ is a set of birth year-by-county fixed effects, which controls for differences in outcomes across birth cohorts. This

term also captures county-level changes that affect cohorts differently, such as implementation of suffrage laws or local labor regulations. $\mathbf{X}_i$ is a vector of individual characteristics measured in 1900 for woman i that may determine labor market outcomes. The following characteristics are included: indicators for father/mother presence, indicators for father/mother literacy, father/mother occupation scores, indicators for father/mother employment in medicine, and family size.

As tests of identification, I examine whether the treatment of interest, $\text{Near}_j \times \text{Age}_{jt}$ in equation (1), predicts changes in baseline characteristics measured in 1900. This is akin to running placebo tests with the potential confounding variables as the outcome (Pei et al. 2019). These specifications retain all fixed effects (θ_j, $\delta_{c(j)t}$), but do not include individual controls $\mathbf{X}_i$. The results from these tests for the 1920 sample are shown in Table 2. The treatment measure does not predict significant differences in baseline family size, parental literacy, parental occupation score, or whether a parent is employed in a medical occupation.[28]

5 Short-run Results: 1920

In this section, I report results for the linked 1900-1920 sample, when the women were ages 20 to 35. I examine entrance into nursing, overall employment, substitution from other health occupations, and heterogeneity by observable characteristics. I then examine other outcomes that capture the economic effects of training.

5.1 Employment and Occupation: 1920

Table 3 reports results from equation (1), estimated using OLS with standard errors clustered at the town level. All estimates use inverse propensity score weights to adjust for the likelihood of linkage (Bailey et al. 2020). Coefficients have been multiplied by 100 and can be interpreted in percentage points. The outcome in column (1) is an indicator for whether a woman was employed in any nursing occupation in 1920. In Panel A, results are shown for the full analysis sample. Women who were within five miles of a training school opening at ages 16 to 20 were 0.13 percentage points more likely to work in nursing. This coefficient is statistically significant at the 1 percent level and is not sensitive to the addition of individual-level controls (column (2)). Relative to the overall sample

[28] In Table D.3, I run the same set of placebo tests for the 1930 sample. I find null results for all outcomes except mother's occupation score. The magnitude of this imbalance is small, and all main specifications control for background characteristics. Note also that p-values are unadjusted for multiple hypothesis testing.

mean of 0.83, this represents a 16 percent increase in the likelihood of entering nursing. Recall that these estimates can be considered "reduced form" effects of access to nurse training at ages 16-20, since the 1920 census measures occupation and not actual training school enrollment. The goal of training schools was to train women in professional nursing. As shown in column (3), women were 0.10 percentage points more likely to be observed as a trained nurse (19 percent relative to the sample mean of 0.54). The subcategory of practical nurse likely captures women working in either untrained or informal nursing roles. As shown in column (5), effects on practical nursing are small and insignificant. None of the estimates are sensitive to adding individual-level controls. This suggests that the entire effect in column (1) is driven by entrance into trained nursing, with little substitution away from practical nursing roles.

The estimates in Table 3, Panel A include a substantial number of women who were not working. In Table D.4, I examine whether my proximity to training measure causes an increase in gainful employment (having a recorded occupation in 1920). I do not find strong evidence of effects on this margin: the coefficient of 0.20 percentage points is positive but not statistically significant. I conclude that access to nurse training did not increase employment among women who otherwise might not have worked. However, the coefficient of 0.20 is large enough to include the overall increase in nursing employment observed in Table 3.

In the following analysis, I examine entrance into nursing conditional on gainful employment. While the effect on employment is statistically insignificant, conditioning on an outcome can cause endogenous selection bias (Elwert and Winship 2014). I report results with this caution in mind. Results are shown in Table 3, Panel B. Point estimates indicate that women were 0.45 percentage points more likely to work in a nursing occupation (column (1)), with most of the effect driven by trained nursing (column (3)). Relative to the sample mean of 1.37 percent, white women who were within five miles of a training school opening when they were ages 16 to 20 were 25 percent more likely to work as trained nurses. As in Panel A, the results for practical nursing conditional on employment are positive and not statistically significant, and the estimates do not change with the addition of individual controls.

I consider the effect of proximity to nurse training on entrance into other health occupations in Table D.5. First, hospital or physician attendant was one occupation that could have been affected by the opening of a nurse training school. This group was small (just 0.21 percent of employed

women) and I do not see significant effects of proximity (column (1)). While other health professions, such as physician, make up a small fraction of women, it is possible that nurse training acted as a gateway to higher-status health professions. Alternatively, nursing may have reinforced the pathway for women into medicine and diverted them from medical schools, especially during the medical school closures of the 1910s (Moehling et al. 2019; Clay et al. 2023). Column (2) shows that female physicians make up just 0.06 percent of employed women in the sample, and proximity to nurse training has no effect on becoming a physician (point estimates are negative). Results are similarly null for the broader group of professional health occupations: physicians, dentists, osteopaths, and pharmacists (column (3)). In column (4), I find an insignificant effect of proximity on the set of semi-professional health occupations that includes chiropractors, dietitians, medical technicians, therapists, and midwifes. Entrance into other health occupations may have been hindered by the spread of licensing regulations at the state level during this period. For instance, midwifery was a regulated occupation by the 1920s (Anderson et al. 2020).

5.1.1 Robustness: 1920

I examine the robustness of the coefficient on nursing in several ways, shown in Figure D.3. The main point estimate in Table 3, column (2) is shown in red. I show that the main coefficient is robust to specifying different end points for the birth cohorts I consider in my main analysis sample. Adding foreign-born women (N = 45,000) has no effect, however, note that this is a highly selected sample of foreign-born women who were present in the 1900 census while young. I assess whether the weighting scheme matters. The main coefficient is not sensitive to using an unweighted estimator or dropping the top 1 percent of weights. I also estimate an alternative set of weights that is based on an entirely female census sample (equivalent to fully interacting gender with all other predictors of linkage), and this does not change the main results.

5.1.2 Heterogeneity analysis: 1920

In this section, I estimate effects of proximity to nurse training for different groups. First, I examine selection into nursing by family background. Rosenberg (1987) highlights the interaction between nurse training and class background, writing how training filled the wards with "would-be

professionals".[29] Family characteristics are measured in 1900, and include parental occupation and occupational income. I proxy for class background by splitting the sample into three groups based on father's occupation. The groupings are sourced from the occupational scheme of Song et al. (2020): (1) professional/managerial occupations (e.g., doctors, lawyers, managers), and (2) routine/manual occupations (e.g., clerks, craftsmen, miners), and (3) farming occupations. The sample contains roughly 200,000 women in group 1, 1,000,000 women in group 2, and 200,000 women in group 3.

Table D.6 reports the estimates on entrance into nursing for each group. In Panel A, women from professional families are 0.51 percentage points more likely to work as a trained nurse (a 78 percent increase relative to the mean). This is a large increase in relative magnitude, surpassing the size of the coefficient for employed women. The effect of exposure does not appear to operate through any other healthcare profession (including practical nursing). Compared to Panel A, Panel B shows an increase in the likelihood of working in nursing, but the coefficient is much smaller for this group. Panel C shows that women from farming families were not entering nursing despite their close proximity. This suggests that nurse training was primarily an opportunity for relatively higher-status women. Note that these estimates are reduced-form in nature and I cannot examine who applied to nurse training schools. As discussed in Section 2.2, standards were high, and so women from non-professional backgrounds may have been screened out at the application stage. These results demonstrate that training schools were a mediating factor in elevating the status of nursing, which is evidence in line with Rosenberg (1987).

Next, I examine how effects vary by marital status. Nurse training schools often would not admit married women, choosing instead to admit young single women. Once a woman graduated from training school, she often continued working up until marriage.[30] The departure from nursing upon marriage may not have been a choice: Nurse training emerged during a period when married women were frequently fired (Goldin 1991). Given that half of the sample is married, I have

[29] Other historical sources have a similar theme. Nutting (1912) quotes a physician at Johns Hopkins, who writes:

> *As long as nurses were drawn from the lower classes and as long as the training was merely manual and not intellectual, there could be no profession of nursing. Now that nursing is offering ever new incentives, the position of the trained nurse has become more elevated, the educational standards are high, the time of nursing has become long enough to permit of thoroughness without overtaxing, the material position of nurses is being improved, the opportunities for higher careers in nursing are multiplied, we find the profession appealing more and more to the best class of woman.*

[30] In Appendix B, I conduct a case study of training school graduates and show that, conditional on marriage, employment among graduates was very low. For women in their twenties in 1910 and 1920, only 9 and 3 percent of graduates were working while married, respectively (Appendix Table B.2).

adequate power to assess differences by marital status. Note that I examine whether exposure to training affected marital status in Table D.4, and I find a positive and insignificant coefficient of 0.17 percentage points (Panel A). Additionally, I do not see strong effects on overall employment for single or married women (Panels B and C). Point estimates are positive for married women and negative for single women, but standard errors are relatively large.

Tables D.7 and D.8 report results for single and married women, respectively. For single women, the effect of proximity on employment in nursing is 2.5 times larger than the baseline estimate in Table 3. Women who were less than five miles from an opening at ages 16 to 20 were 0.34 percentage points more likely to work in nursing, an increase of 21 percent relative to the baseline mean (Panel A). Most of this effect is driven by entrance into trained nursing, whereas the coefficient on practical nursing is positive but not statistically significant. Conditional on employment, women were 0.53 percentage points more likely to work in nursing (an increase of 24 percent relative to the baseline mean). As with the full sample, I do not find evidence that other health occupations were affected (columns (4)-(7)). For married women, in Table D.8, Panel A, I find insignificant effects of proximity to nurse training on work in nursing. The overall rate of work is low (7.5 percent), therefore the gainfully employed sample in Panel B is small. Conditional on employment, the coefficient on nursing is 0.392 (statistically insignificant). The magnitude, however, is sizeable relative to the mean. Additionally, columns (3) and (4) suggest that some married women were working as practical nurses rather than attendants. I view this as suggestive evidence that trained nurses in 1920 had a limited ability to work in formal nursing roles while married, which aligns with historical evidence (Goldin 1991; Kalisch and Kalisch 2004).

5.2 Economic Status: 1920

In this section, I examine how access to nurse training affected measures of economic status. As discussed in Section 3.2, the census does not contain wage information before 1940, so I employ occupation-based income measures commonly used to study historical inequality and mobility. I focus on occupational income and the Song score, which captures an occupation's relative status based on human capital content (Song et al. 2020; Ward 2023; Buckles, Price, et al. 2023). Conditional on working, changes in economic status are determined by sorting across occupations.

In Table 4, I show effects of proximity to nurse training in adolescence on the log of occupational income score and LIDO score. In the full sample (Panel A), non-working individuals are assigned a score of zero. As shown in columns (1) and (2), the average effect of proximity is positive but statistically insignificant. Nursing lies in the middle of the occupational income distribution (Figure C.5), so we might not expect a strong positive effect unless women were entering employment or substituting from low-income occupations. In Panel B, I report results conditional on employment. I add the Song score as an outcome since this is a relative percentile measure. Across all measures, I find no significant positive effects. Similar to the LIDO score, the Song score for nurses relative to other common occupations is slightly lower (Figure C.6).

I conduct several exercises to understand the estimates in Table 4. First, I examine how proximity to nurse training affected substitution from other common occupations. In Table D.9, I report results for the occupations of teacher, clerk, bookkeeper, secretary, and laborer (including farm laborer). These five occupations cover approximately 50 percent of the employed women in the sample. In both the full and employed samples, I find mostly negative and insignificant coefficients.[31] This suggests that women were substituting into nursing and away from a wide range of occupations, rather than just one (such as teaching).[32] Second, I examine heterogeneity in the effects on economic status by family background in Table D.10. Conditional on employment, columns (1) and (2) report positive but insignificant effects on absolute measure of income. Despite the strong positive effect of proximity on entrance into nursing for women with professional fathers in Panel A, I estimate no gain in economic status. In column (3), the only significant increase in relative status (Song score) is for women with fathers in routine and manual occupations (Panel B). This coefficient is in percentiles and represents a less than 1 percent increase relative to the subgroup mean.[33] These results suggest that nursing did not result in meaningful gains in economic status, measured in either absolute or relative terms. Note the primary limitation is that real wage income is not observed in 1920.

[31] In results not shown, I examine heterogeneity in occupational substitution by marital status, and results are similar.

[32] In Appendix Section A.2, I discuss historical evidence that nurses worked previously in teaching, clerical, or health occupations. This only captures women who worked before starting nursing, whereas I capture substitution for all women, including those for whom nursing was their first job.

[33] In results not shown, I find no effect of proximity on employment and marital status in each subgroup. I also look at occupation score and LIDO score in the full sample and find insignificant results. Replicating Table D.9 for each subgroup, I find largely null results, however, women from professional backgrounds are less likely to become teachers relative to the other subgroups.

6 Long-run Results: 1930

In this section, I report results for the 1900-1930 linked sample, when the women were ages 30 to 45. I follow Section 5 and examine employment in nursing, occupational substitution, and economic status. The sample is smaller, but allows for more power to study married women, who make up 70 percent of the analysis sample. This analysis can inform whether availability of training played a role in increasing labor force participation, which was rising among married women during this period (Goldin 1983).

6.1 Employment and Occupation: 1930

Table 5 reports results for nursing in 1930, analogous to the 1920 results reported in Table 3. Effects on nursing in the full and employed samples (Panels A and B) are less than half the size of the effects in 1920 and no longer statistically significant. As in 1920, I examine whether women exposed to nurse training are more or less likely to be employed or married in 1930 (the extensive margin). These results are reported in Table D.11. Effects on employment are negative but not statistically significant with controls (column (2)). In column (3), women are 0.6 percentage points more likely to be married by 1930 (0.8 percent increase relative to the sample mean), which likely drives a portion of the negative coefficient observed on employment. The coefficient on marriage has several interpretations, according to related literature on history and nursing. For instance, Kalisch and Kalisch (2004) point out that with the lowering of training age from 1900 to 1920, it was easier for a woman to train and subsequently marry in her twenties.

I examine substitution among other health occupations in 1930, shown in Table D.12. The estimates in Panel A suggest that women exposed to nurse training are 0.02 percentage points less likely to become physicians (column (2)), with no significant effect on other occupations. The coefficient is sizeable given that only 0.04 percent of women were physicians in 1930. Note also that the effect on physician in 1920 is negative but not statistically significant. It is possible that 1920 was too early to fully capture effects on the physician occupation given that many women were still of training age. I interpret this result as evidence that the professionalization of nursing via training contributed to gender segregation within medicine.[34]

[34] Moehling et al. (2019) find reductions in medical school enrollment for women during this period.

6.1.1 Heterogeneity analysis: 1930

In Table D.13, I report results for married women in 1930.[35] In Panel A, column (1) shows that married women who were proximate to nurse training in adolescence were more likely to work in nursing. Much of this increase is driven by practical nursing (column (3)). I interpret this result as an increased attachment to the field of nursing, as it suggests these women were working in less formal nursing roles. In 1930, as in 1920, nurses who married were uncommon, so work in practical nursing might have been a more feasible option. Note also that married women were less likely to become physicians, and this pattern of results holds conditional on employment (Panel B). In Table D.11, I show that married women were not more likely to be gainfully employed, which suggests that the increase in nursing operates through the intensive margin.

6.2 Economic Status: 1930

In this section, I study effects of proximity on measures of economic status, as in Section 5.2. The approach for 1930 is the same as for 1920, and includes the same measures (occupation score, LIDO score, and Song score). In Table 6, women who were proximate to nurse training in adolescence have marginally lower occupation scores. In Panel A, all non-employed women receive a value of zero, so the negative coefficient on employment in Table D.11 is driving this result. In Panel B, results are insignificant and negative across all measures, which suggests that occupational sorting did not improve economic status for employed women in 1930. However, the earlier caveat with the 1920 sample remains that actual wage income is not observable.

These results are incomplete without a discussion of husband's characteristics. For the married women in the sample, nurse training may have affected economic status through the marriage market. I test whether married women exposed to training schools were more likely to marry relatively higher-earning men in 1930. These results are reported in Table D.14. In column (1), I do not find increases in husband's occupational income score. In column (2), I find increases in the likelihood of having a husband working in a medical occupation (e.g., physician, pharmacist), however, the coefficient is not statistically significant. I conclude that training has limited effects on measures of

[35] In results not shown, I estimate analogous results for single women. Since only 28 percent of the 1930 sample are single, the standard errors for these results are large. Results show increases in the likelihood of working as nurses (coefficient of 0.10) and decreases in the likelihood of working as physicians (coefficient of -0.05), but coefficients are not statistically significant.

economic status, either through the labor market for women who were employed, or through the marriage market for women who were married.

7 Conclusion

In this paper, I use the spread of nurse training schools over time to estimate how training affected the economic outcomes of women born at the turn of the 20$^{\text{th}}$ century. First, I show that women who had relatively more access to training were more likely to become nurses in their twenties. This effect is driven by women from professional families, which I interpret as evidence that nurse training school standards made training less accessible to women from non-professional backgrounds. Second, I find limited evidence that access to nursing improved economic status or increased labor force participation. While this analysis is limited by the lack of historical wage data, the results imply that the women most likely to become nurses had equally or higher-earning alternative occupations. Third, I show evidence of fade out in the effects of employment in nursing. By 1930, effect sizes are smaller given the propensity for women to marry and drop out of the labor force. However, I show some evidence that married women were more likely to stay working in nursing, which foreshadows the coming rise of married women in the labor force. Finally, women proximate to nurse training were more likely to be married and less likely to work as physicians, which I interpret as evidence that nurse training reinforced gender segregation within medicine.

This paper finds that structured training assisted in the transformation of nursing from menial labor for a professional occupation. Since the establishment of nurse training, nursing has grown into one of the largest occupations among women today. Over 1 in 25 women work as professional nurses, the largest single occupation group for working women. The supply of new nursing degrees is also large: nursing ranked second in terms of bachelor's degrees granted to women in 2021. In stark contrast to nursing in 1870, nursing today is highly respected. According to a Gallup poll from 2022, nursing was the highest ranked career in terms of honesty, a position it has held for the past 20 years.

This paper leaves several open questions. Black and foreign-born women may have benefited from the spread of new vocational training opportunities at the turn of the 20$^{\text{th}}$ century. Workforce training may have also generated positive spillovers during this period, even if private returns to training were not significant. In particular, the supply of nurses may have helped hospitals save

lives, especially during times of crisis. More broadly, this paper highlights the role of training in the professionalization of an occupation. In countries with nascent vocational education, new programs could shift the composition of occupations in ways that elevate social standing. More work is necessary to build an understanding of how firm-sponsored training affects the outcomes of workers and determines broader patterns in the labor market.

References

Abel-Smith, Brian (1960). *A History of the Nursing Profession*. London, UK: Heinemann.

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, James J. Feigenbaum, and Santiago Pérez (2021). "Automated Linking of Historical Data". *Journal of Economic Literature* 59.3, pp. 865–918.

Acemoglu, Daron and Joshua Angrist (2000). "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws". *NBER Macroeconomics Annual* 15, pp. 9–59.

Acemoglu, Daron and Jörn-Steffen Pischke (1998). "Why Do Firms Train? Theory and Evidence". *The Quarterly Journal of Economics* 113.1, pp. 79–119.

— (1999a). "Beyond Becker: Training in Imperfect Labour Markets". *The Economic Journal* 109.453, pp. 112–142.

— (1999b). "The Structure of Wages and Investment in General Training". *Journal of Political Economy* 107.3, pp. 539–572.

Acton, Riley, Kalena E. Cortes, and Camila Morales (2024). *Distance to Opportunity: Higher Education Deserts and College Enrollment Choices*. Working Paper.

Althoff, Lukas, Harriet Brookes Gray, and Hugo Reichardt (2024). *The Missing Link(s): Women and Intergenerational Mobility*. Working Paper.

Altonji, Joseph, Peter Arcidiacono, and Arnaud Maurel (2016). "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects". *Handbook of the Economics of Education*. Vol. 5. Cambridge, MA: Elsevier, pp. 305–396.

Altonji, Joseph G., Erica Blom, and Costas Meghir (2012). "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers". *Annual Review of Economics* 4.1, pp. 185–223.

American Nurses Association (1918). *A List of Schools of Nursing Accredited by the State Boards of Nurse Examiners*. Chicago, IL.

— (1922). *A List of Schools of Nursing Accredited by the State Boards of Nurse Examiners*. Chicago, IL.

Anderson, D. Mark, Ryan Brown, Kerwin Kofi Charles, and Daniel I. Rees (2020). "Occupational Licensing and Maternal Health: Evidence from Early Midwifery Laws". *Journal of Political Economy* 128.11, pp. 4337–4383.

Arcidiacono, Peter, V. Joseph Hotz, Arnaud Maurel, and Teresa Romano (2020). "Ex Ante Returns and Occupational Choice". *Journal of Political Economy* 128.12, pp. 4475–4522.

Autor, David H. and David Dorn (2013). "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market". *American Economic Review* 103.5, pp. 1553–1597.

Bailey, Martha J., Connor Cole, Morgan Henderson, and Catherine Massey (2020). "How Well Do Automated Linking Methods Perform? Lessons from U.S. Historical Data". *Journal of Economic Literature* 58.4, pp. 997–1044.

Bailey, Martha J. and Peter Z. Lin (2024). *Marital Matching and Women's Intergenerational Mobility in the Late-19th and Early-20th-Century U.S.* Working Paper.

Baker, Richard B., John Blanchette, and Katherine Eriksson (2020). "Long-Run Impacts of Agricultural Shocks on Educational Attainment: Evidence from the Boll Weevil". *The Journal of Economic History* 80.1, pp. 136–174.

Becker, Gary S. (1964). *Human Capital*. Chicago, IL: University of Chicago Press.

Berkes, Enrico, Ezra Karger, and Peter Nencka (2023). "The Census Place Project: A Method for Geolocating Unstructured Place Names". *Explorations in Economic History* 87.101477.

Bertaux, Nancy E. (1991). "The Roots of Today's "Women's Jobs" and "Men's Jobs": Using the Index of Dissimilarity to Measure Occupational Segregation by Gender". *Explorations in Economic History* 28.4, pp. 433–459.

Blau, Francine D. and Wallace E. Hendricks (1979). "Occupational Segregation by Sex: Trends and Prospects". *The Journal of Human Resources* 14.2, p. 197.

Blau, Francine D. and Lawrence M. Kahn (2017). "The Gender Wage Gap: Extent, Trends, and Explanations". *Journal of Economic Literature* 55.3, pp. 789–865.

Blazier, Florence A. (1924). *Investigation of Nursing as a Professional Opportunity for Girls*. Bloomington, Ind.

Bleemer, Zachary and Sarah Quincy (2022). *Characteristics and Consequences of College-Going in the Great Depression*. Working Paper.

Blundell, Richard, Lorraine Dearden, Costas Meghir, and Barbara Sianesi (1999). "Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy". *Fiscal Studies* 20.1, pp. 1–23.

Buckles, Kasey, Adrian Haws, Joseph Price, and Haley E.B. Wilbert (2023). *Breakthroughs in Historical Record Linking Using Genealogy Data: The Census Tree Project*. Working Paper 31671. National Bureau of Economic Research.

Buckles, Kasey, Joseph Price, Zachary Ward, and Haley E.B. Wilbert (2023). *Family Trees and Falling Apples: Historical Intergenerational Mobility Estimates for Women and Men*. Working Paper 31918. National Bureau of Economic Research.

Burbano, Vanessa, Olle Folke, Stephan Meier, and Johanna Rickne (2023). "The Gender Gap in Meaningful Work". *Management Science*, pp. 1–20.

Burbano, Vanessa, Nicholas Padilla, and Stephan Meier (2024). "Gender Differences in Preferences for Meaning at Work". *American Economic Journal: Economic Policy* 16.3, pp. 61–94.

Cameron, Stephen V. and Christopher Taber (2004). "Estimation of Educational Borrowing Constraints Using Returns to Schooling". *Journal of Political Economy* 112.1, pp. 132–182.

Card, David (1995). *Using Geographic Variation in College Proximity to Estimate the Return to Schooling*. Working Paper 4483. National Bureau of Economic Research.

Card, David, Ana Rute Cardoso, and Patrick Kline (2016). "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women". *The Quarterly Journal of Economics* 131.2, pp. 633–686.

Carneiro, Pedro, James Heckman, and Edward Vytlacil (2011). "Estimating Marginal Returns to Education". *American Economic Review* 101.6, pp. 2754–2781.

Carollo, Nicholas A. (2024). *Historical Data on Professional and Occupational Credentialing Requirements in the United States*. Working Paper.

Carollo, Nicholas A., Jason F. Hicks, Andrew Karch, and Morris M. Kleiner (2022). *The Origins and Evolution of Occupational Licensing in the United States*. Working Paper.

Clay, Karen, Jeff Lingwall, and Melvin Stephens Jr. (2021). "Laws, Educational Outcomes, and Returns to Schooling Evidence from the First Wave of U.S. State Compulsory Attendance Laws". *Labour Economics* 68.101935.

Clay, Karen, Grant Miller, Margarita Portnykh, and Ethan Schmick (2023). *Medical School Closures and Health Care Market Adjustment in the Flexner Report Era*. Working Paper.

Coles, Anna B. (1969). "The Howard University School of Nursing in Historical Perspective". *Journal of the National Medical Association* 61.2, pp. 105–118.

Comin, Diego and Bart Hobijn (2010). "An Exploration of Technology Diffusion". *American Economic Review* 100.5, pp. 2031–2059.

Committee for the Study of Nursing Education (1923). *Nursing and Nursing Education in the United States*. New York, NY: The Macmillan Company.

Connolly, Kevin (2021). *How Does Access to College Affect Long-Term Life Outcomes? Evidence from U.S. Openings of Two-Year Public Colleges*. Working Paper.

Cortés, Patricia, Ying Feng, Nicolás Guida-Johnson, and Jessica Pan (2024). *Automation and Gender: Implications for Occupational Segregation and the Gender Skill Gap*. Working Paper 32030. National Bureau of Economic Research.

Costa, Dora L. (2000). "From Mill Town to Board Room: The Rise of Women's Paid Labor". *Journal of Economic Perspectives* 14.4, pp. 101–22.

Costa, Dora L. and Matthew E. Kahn (2006). "Public Health and Mortality: What Can We Learn from the Past?" *Public Policy and the Income Distribution*, pp. 359–398.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker (2015). "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings". *American Economic Review* 105.2, pp. 678–709.

Currie, Janet, Bahadir Dursun, Michael Hatch, and Erdal Tekin (2024). *The Hidden Cost of Firearm Violence on Infants in Utero*. Working Paper.

Currie, Janet and Enrico Moretti (2003). "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings". *The Quarterly Journal of Economics* 118.4, pp. 1495–1532.

D'Antonio, Patricia (2010). *American Nursing: A History of Knowledge, Authority, and the Meaning of Work*. Baltimore, MD: Johns Hopkins University Press.

Deming, David J., Claudia Goldin, and Lawrence Katz (2012). "The For-Profit Postsecondary School Sector: Nimble Critters or Agile Predators?" *Journal of Economic Perspectives* 26.1, pp. 139–164.

Deming, Dorothy (1950). "Practical Nursing Then and Now". *American Journal of Nursing* 50.10, pp. 621–623.

Doxey, Alison, Ezra Karger, and Peter Nencka (2022). *The Democratization of Opportunity: The Effects of the U.S. High School Movement*. Working Paper.

Doyle, William R. and Benjamin T. Skinner (2016). "Estimating the Education-Earnings Equation Using Geographic Variation". *Economics of Education Review* 53, pp. 254–267.

Duflo, Esther (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment". *American Economic Review* 91.4, pp. 795–813.

Elwert, Felix and Christopher Winship (2014). "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable". *Annual Review of Sociology* 40, pp. 31–53.

Emran, M. Shahe and Forhad Shilpi (2011). "Intergenerational Occupational Mobility in Rural Economy: Evidence from Nepal and Vietnam". *Journal of Human Resources* 46.2, pp. 427–458.

Eriksson, Katherine, Jacqueline Craig, and Gregory T. Niemesh (2023). *Marriage and the Intergenerational Mobility of Women: Evidence from Marriage Certificates 1850-1910*. Working Paper.

Espín-Sánchez, José-Antonio, Joseph P. Ferrie, and Christopher Vickers (2023). *Women and the Econometrics of Family Trees*. Working Paper 31598. National Bureau of Economic Research.

Feigenbaum, James J. (2018). "Multiple Measures of Historical Intergenerational Mobility: Iowa 1915 to 1940". *The Economic Journal* 128.612, F446–F481.

Feigenbaum, James J. and Daniel P Gross (2024). "Answering the Call of Automation: How the Labor Market Adjusted to Mechanizing Telephone Operation". *The Quarterly Journal of Economics* 139.3, pp. 1879–1939.

Feigenbaum, James J. and Hui Ren Tan (2020). "The Return to Education in the Mid-Twentieth Century: Evidence from Twins". *The Journal of Economic History* 80.4, pp. 1101–1142.

Fernández, Raquel (2013). "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century". *American Economic Review* 103.1, pp. 472–500.

Gilpin, Gregory and Christiana Stoddard (2017). "Does Regulating For-Profit Colleges Improve Educational Outcomes? What We Know, What We Don't Know, and What We Need to Find Out". *Journal of Policy Analysis and Management* 36.4, pp. 942–950.

Goldin, Claudia (1980). "The Work and Wages of Single Women, 1870-1920". *The Journal of Economic History* 40.1, pp. 81–88.

— (1983). "The Changing Economic Role of Women: A Quantitative Approach". *The Journal of Interdisciplinary History* 13.4, pp. 707–733.

— (1986a). "Monitoring Costs and Occupational Segregation by Sex: A Historical Analysis". *Journal of Labor Economics* 4.1, pp. 1–27.

— (1986b). "The Economic Status of Women in the Early Republic: Quantitative Evidence". *The Journal of Interdisciplinary History* 16.3, pp. 375–404.

— (1990). *Understanding the Gender Gap: An Economic History of American Women*. Vol. 96. New York, NY: Oxford University Press.

— (1991). "Marriage Bars: Discrimination Against Married Women Workers from the 1920s to the 1950s". *Favorites of Fortune: Technology, Growth, and Economic Development since the Industrial Revolution*. Cambridge, MA: Harvard University Press, pp. 511–536.

— (1992). *The Meaning of College in the Lives of American Women: The Past One-Hundred Years*. Working Paper 4099. National Bureau of Economic Research.

— (1998). "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century". *The Journal of Economic History* 58.2, pp. 345–374.

— (1999). "Egalitarianism and the Returns to Education during the Great Transformation of American Education". *Journal of Political Economy* 107.S6, S65–S94.

— (2021). *Career & Family: Women's Century-Long Journey Toward Equity*. Princeton, NJ: Princeton University Press.

Goldin, Claudia and Lawrence Katz (2000). "Education and Income in the Early 20th Century: Evidence from the Prairies". *The Journal of Economic History* 60.3, pp. 782–818.

— (2011). "Mass Education and the State: The Role of State Compulsion in the High School Movement". *Understanding Long Run Economic Growth*. Chicago, IL: University of Chicago Press, pp. 275–311.

Greenberg, Kyle, Matthew Gudgeon, Adam Isen, Corbin L. Miller, and Richard W. Patterson (2024). *Intergenerational Transmission of Occupation: Lessons from the United States Army*. Working Paper 33009. National Bureau of Economic Research.

Griliches, Zvi (1977). "Estimating the Returns to Schooling: Some Econometric Problems". *Econometrica* 45.1, pp. 1–22.

Grosz, Michel (2020). "The Returns to a Large Community College Program: Evidence from Admissions Lotteries". *American Economic Journal: Economic Policy* 12.1, pp. 226–253.

Hanushek, Eric A., Guido Schwerdt, Ludger Woessmann, and Lei Zhang (2017). "General Education, Vocational Education, and Labor-Market Outcomes over the Lifecycle". *Journal of Human Resources* 52.1, pp. 48–87.

Hawkins, Sue (2010). *Nursing and Women's Labour in the Nineteenth Century*. Routledge.

Hindman, Hugh D. (2002). *Child Labor: An American History*. Armonk, NY: M. E. Sharpe, Inc.

Hine, Darlene Clark (1989). *Black Women in White: Racial Conflict and Cooperation in the Nursing Profession, 1890-1950*. Bloomington: Indiana University Press.

Hodson, Jane (1898). *How to Become a Trained Nurse*. New York, NY: William Abbatt.

Hoehn-Velasco, Lauren (2021). "The Long-term Impact of Preventative Public Health Programs". *The Economic Journal* 131.634, pp. 797–826.

Huzza, Harry (1887). "The Philadelphia Lying-in Charity". *Atlanta Medical and Surgical Journal* 25, pp. 360–364.

Kalisch, Philip A. and Beatrice J. Kalisch (2004). *American Nursing: A History*. Philadelphia, PA: Lippincott Williams & Wilkins.

Kane, Thomas J. and Cecilia E. Rouse (1995). "Labor-Market Returns to Two- and Four-Year College". *American Economic Review* 85.3, pp. 600–614.

Katz, Lawrence and Robert A. Margo (2014). "Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective". *Human Capital in History: The American Record*. University of Chicago Press, pp. 15–57.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad (2016). "Field of Study, Earnings, and Self-Selection". *The Quarterly Journal of Economics* 131.3, pp. 1057–1111.

Kling, Jeffrey R. (2001). "Interpreting Instrumental Variables Estimates of the Returns to Schooling". *Journal of Business & Economic Statistics* 19.3, pp. 358–64.

Kose, Esra, Elira Kuka, and Na'ama Shenhav (2021). "Women's Suffrage and Children's Education". *American Economic Journal: Economic Policy* 13.3, pp. 374–405.

Law, Marc T. and Sukkoo Kim (2005). "Specialization and Regulation: The Rise of Professionals and the Emergence of Occupational Licensing Regulation". *The Journal of Economic History* 65.3, pp. 723–756.

Law, Marc T. and Mindy S. Marks (2013). "From Certification to Licensure: Evidence from Registered and Practical Nurses in the United States, 1950-1970". *The European Journal of Comparative Economics* 10.2, pp. 177–98.

— (2017). "The Labor-Market Effects of Occupational Licensing Laws in Nursing". *Industrial Relations: A Journal of Economy and Society* 56.4, pp. 640–661.

Li, Sophie (2023). "The Effect of A Woman-Friendly Occupation on Employment: U.S. Postmasters Before World War Two".

Linden, Leigh and Jonah Rockoff (2008). "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws". *American Economic Review* 98.3, pp. 1103–1127.

Lleras-Muney, Adriana (2002). "Were Compulsory Attendance and Child Labor Laws Effective? An Analysis from 1915-1939". *Journal of Law and Economics* 45.2, pp. 401–35.

Long, Bridget T. (2004). "How Have College Decisions Changed over Time? An Application of the Conditional Logistic Choice Model". *Journal of Econometrics*. Higher Education (Annals Issue) 121.1, pp. 271–296.

Lovenheim, Michael F. and Jonathan Smith (2022). *Returns to Different Postsecondary Investments: Institution Type, Academic Programs, and Credentials*. Working Paper 29933. National Bureau of Economic Research.

Ma, Xiao, Alejandro Nakab, and Daniela Vidart (2024). "Human Capital Investment and Development: The Role of On-the-Job Training". *Journal of Political Economy Macroeconomics* 2.1, pp. 107–148.

Miller, Grant (2008). "Women's Suffrage, Political Responsiveness, and Child Survival in American History". *The Quarterly Journal of Economics* 123.3, pp. 1287–1327.

Moehling, Carolyn M., Gregory T. Niemesh, and Melissa A. Thomasson (2019). *Shut Down and Shut Out: Women Physicians in the Era of Medical Education Reform*. Working Paper.

Moorthy, Avinash, Theodore F. Figinski, and Alicia Lloro (2022). "Revisiting the Effect of Education on Later Life Health". *Finance and Economics Discussion Series* 2022.007, pp. 1–31.

Mountjoy, Jack (2022). "Community Colleges and Upward Mobility". *American Economic Review* 112.8, pp. 2580–2630.

Muehlemann, Samuel and Stefan C. Wolter (2020). "The Economics of Vocational Training". *The Economics of Education*. 2nd ed. Elsevier, pp. 543–554.

Munnich, Elizabeth and Abigail Wozniak (2020). "What Explains the Rising Share of U.S. Men in Registered Nursing?" *ILR Review* 73.1, pp. 91–123.

National Center for Education Statistics (2021). *Table 318.30: Bachelor's, Master's, and Doctor's Degrees Conferred by Postsecondary Institutions, by Sex of Student and Field of Study: Academic Year 2020-21*.

Ng, Kevin and Evan Riehl (2024). "The Returns to STEM Programs for Less-Prepared Students". *American Economic Journal: Economic Policy* 16.2.

Nutting, Adelaide (1912). *Educational Status of Nursing*. Washington, D.C.: Department of the Interior.

Nybom, Martin (2017). "The Distribution of Lifetime Earnings Returns to College". *Journal of Labor Economics* 35.4, pp. 903–952.

Olivetti, Claudia and M. Daniele Paserman (2015). "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940". *American Economic Review* 105.8, pp. 2695–2724.

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt (2019). "Poorly Measured Confounders Are More Useful on the Left than on the Right". *Journal of Business & Economic Statistics* 37.2, pp. 205–216.

Peitzman, Steven J. (2000). *A New and Untried Course: Woman's Medical College and Medical College of Pennsylvania, 1850-1998*. Rutgers University Press.

Pigou, Arthur C. (1912). *Wealth and Welfare*. London: Macmillan.

Pollitt, Phoebe and Wendy Miller (2010). "North Carolina, Pioneer in American Nursing". *American Journal of Nursing* 110.2, pp. 70–71.

Price, Joseph, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley (2021). "Combining Family History and Machine Learning to Link Historical Records: The Census Tree Data Set". *Explorations in Economic History* 80.101391.

Putnam, Chas. P. (1874). *Training Schools for Nurses in America*. American Social Science Association.

Reverby, Susan M. (1987). *Ordered to Care: The Dilemma of American Nursing, 1850-1945*. New York, NY: Cambridge University Press.

Rosenberg, Charles E. (1987). *The Care of Strangers : The Rise of America's Hospital System*. New York, NY: Basic Books.

Rouse, Cecilia E. (1995). "Democratization or Diversion? The Effect of Community Colleges on Educational Attainment". *Journal of Business & Economic Statistics* 13.2, pp. 217–224.

Ruggles, Steven, Catherine A. Fitch, Ronald Goeken, Josiah Grover, J. David Hacker, Matt A. Nelson, Jose Pacas, Evan Roberts, and Matthew Sobek (2020). *IPUMS Restricted Complete Count Data: Version 2.0 [Dataset]*. Minneapolis: University of Minnesota.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek (2022). *IPUMS USA: Version 12.0 [Dataset]*. Minneapolis, MN: IPUMS.

Saavedra, Martin and Tate Twinam (2020). "A Machine Learning Approach to Improving Occupational Income Scores". *Explorations in Economic History* 75.101304.

Shannon, Mary L. (1975). "Nurses in American History: Our First Four Licensure Laws". *The American Journal of Nursing* 75.8, p. 1327.

Shen, Zejiang, Ruochen Zhang, Melissa Dell, Benjamin Lee, Jacob Carlson, and Weining Li (2021). "LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis". *International Conference on Document Analysis and Recognition*, pp. 131–146.

Silliman, Mikko and Hanna Virtanen (2022). "Labor Market Returns to Vocational Secondary Education". *American Economic Journal: Applied Economics* 14.1, pp. 197–224.

Sloane, Carolyn M., Erik G. Hurst, and Dan A. Black (2021). "College Majors, Occupations, and the Gender Wage Gap". *Journal of Economic Perspectives* 35.4, pp. 223–248.

Soliz, Adela (2018). "The Effects of the Expansion of For-Profit Colleges on Student Enrollments and Outcomes at Community Colleges". *Educational Evaluation and Policy Analysis* 40.4, pp. 631–652.

Song, Xi, Catherine G. Massey, Karen A. Rolf, Joseph P. Ferrie, Jonathan L. Rothbaum, and Yu Xie (2020). "Long-Term Decline in Intergenerational Mobility in the United States since the 1850s". *Proceedings of the National Academy of Sciences* 117.1, pp. 251–258.

Stephens, Melvin and Dou-Yan Yang (2014). "Compulsory Education and the Benefits of Schooling". *American Economic Review* 104.6, pp. 1777–1792.

Thomasson, Melissa A. and Jaret Treber (2008). "From Home to Hospital: The Evolution of Childbirth in the United States, 1928–1940". *Explorations in Economic History* 45.1, pp. 76–99.

Turner, Sarah E and William G Bowen (1999). "Choice of Major: The Changing (Unchanging) Gender Gap". *ILR Review*.

U.S. Bureau of Education (1907). *Report of the Commissioner of Education for the Year Ending June 30, 1905*. Washington, D.C.: Department of the Interior.

— (1919). *Nurse Training Schools 1917-18*. Washington, D.C.: Department of the Interior.

— (1928). *Statistics of Nurse Training Schools 1926-1927*. Washington, D.C.: Department of the Interior.

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

Walters, Christopher (2018). "The Demand for Effective Charter Schools". *Journal of Political Economy* 126.6, pp. 2179–223.

Ward, Zachary (2023). "Intergenerational Mobility in American History: Accounting for Race and Measurement Error". *American Economic Review* 113.12, pp. 3213–3248.

Wesley, Nathaniel (2010). *Black Hospitals in America: History, Contribution, and Decline*. NRW Associates Publications.

Whelan, Jean C. (2021). *Nursing the Nation: Building the Nurse Labor Force*. Rutgers University Press.

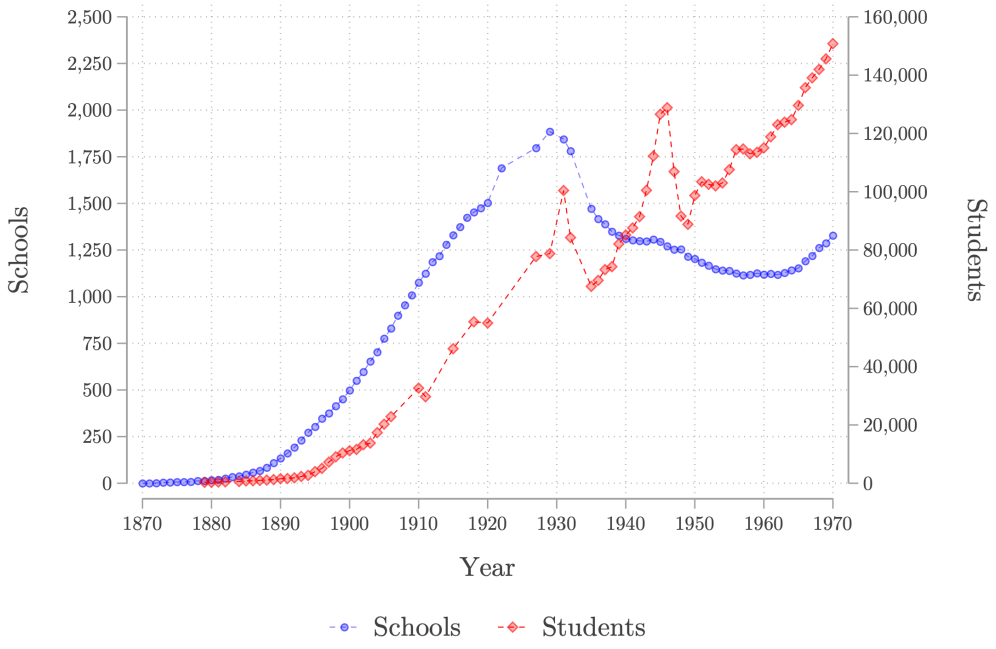
Withrow, Jennifer (2021). *The Farm Woman's Problems: Farm Crisis in the U.S. South and Migration to the City, 1920-1940*. Working Paper.

Xu, Di, Kelli A. Bird, Michael Cooper, and Benjamin L. Castleman (2024). *Noncredit Workforce Training, Industry Credentials, and Labor Market Outcomes*. Annenberg Institute at Brown University.

Zafar, Basit (2013). "College Major Choice and the Gender Gap". *Journal of Human Resources* 48.3, pp. 545–595.

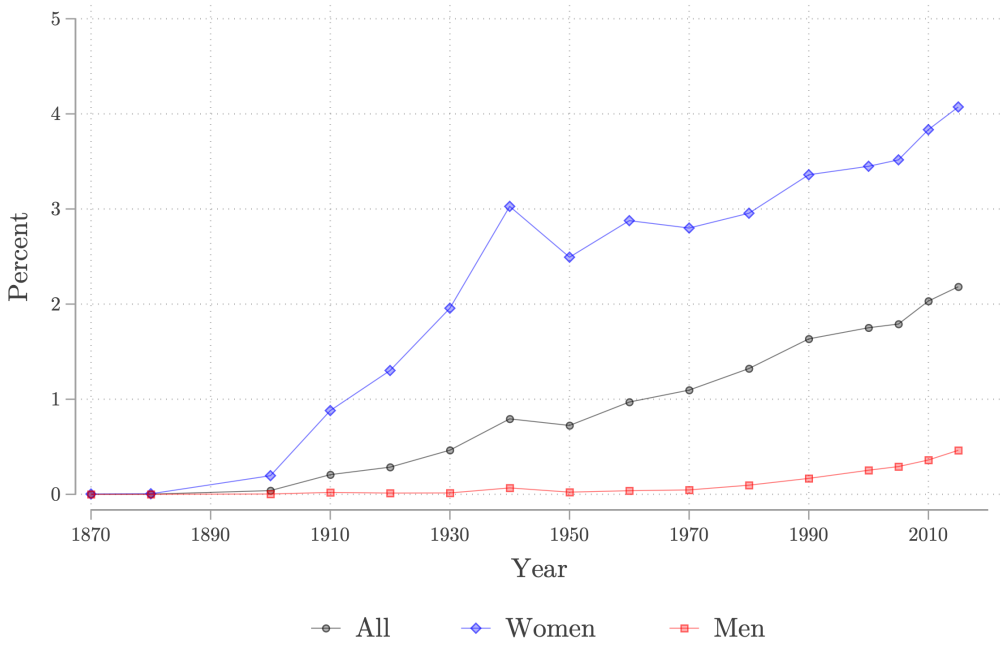
8 Figures

Figure 1: Growth of nurse training schools and students



Sources: Reports of the U.S. Bureau of Education, the American Nurses Association, and the Census Bureau. See Appendix C.
Notes: Figure plots the growth of nurse training schools (left axis) and nursing students (right axis) from 1870 to 1970.

Figure 2: Percent of the labor force in professional nursing



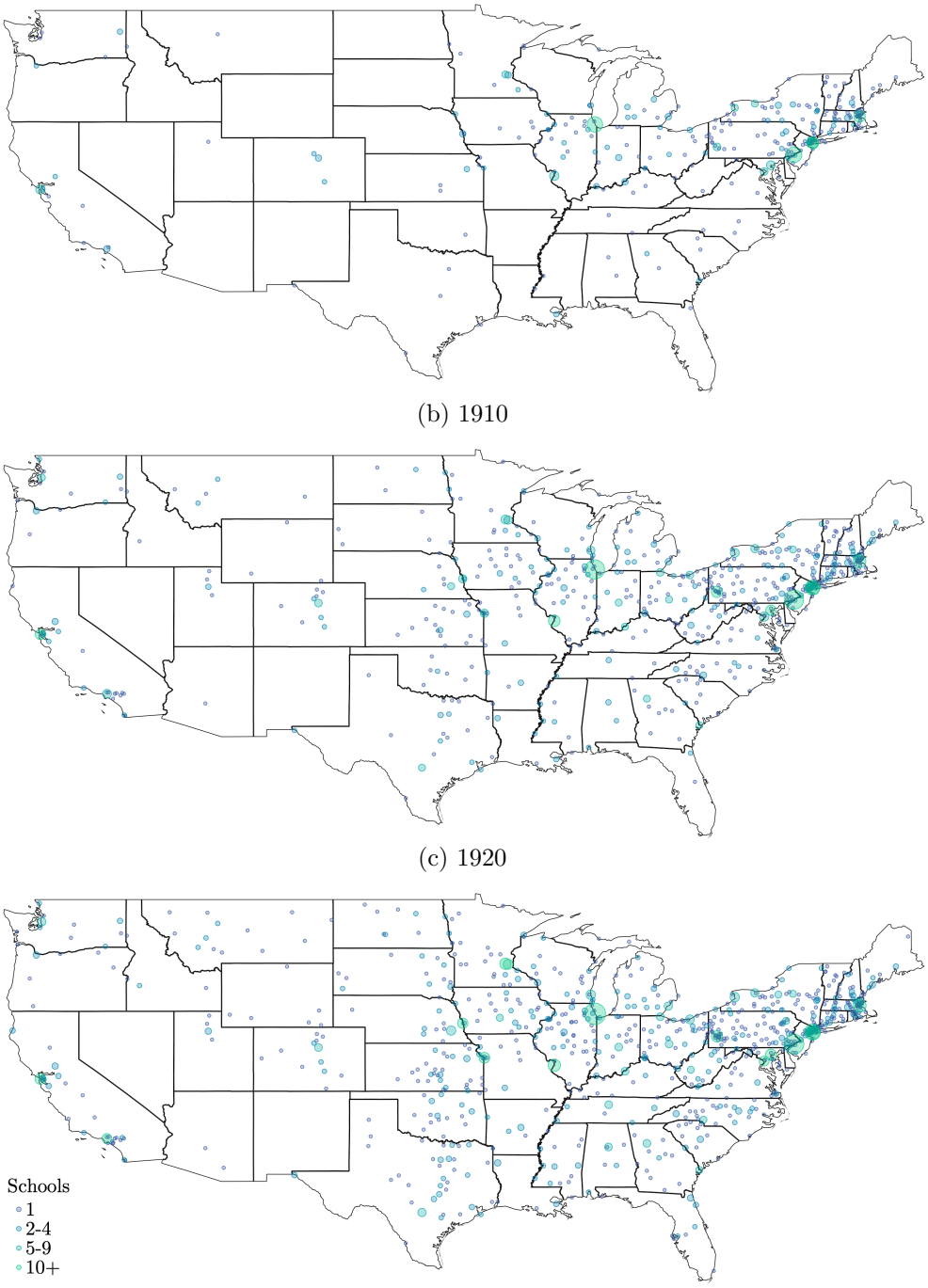
Sources: Complete-count census for 1870-1940; 1% samples for 1950 and 1970; 5% samples for 1960, 1980, 1990, and 2000; American Community Survey for 2005, 2010, and 2015. Person weights are used for all samples.
Notes: Figure shows professional nursing as a percent of the total labor force from 1870 to 2015. Percent in nursing by gender is also reported. The labor force is defined in Appendix Section C.2.2. Prior to 1950, "professional nurses" are individuals with a 1950 occupation code of 58. For 1950 onward, professional nurses are defined according to the codes in Table C.2.

Figure 3: Locations of nurse training schools

(a) 1900

(b) 1910

(c) 1920



Sources: Reports of the U.S. Bureau of Education, the American Nurses Association, and the Census Bureau. See Appendix C.
Notes: Figure plots the locations of nurse training schools in 1900, 1910, and 1920. Counts are aggregated to the town level, and point size is proportional to the number of schools (see legend).

9 Tables

Table 1: Descriptive statistics, linked 1920 sample

	(1) Full sample	(2) Treated = 1	(3) Treated = 0
Childhood town, 1900			
Ln(population)	8.36	11.74	9.90
Region = North	0.28	0.41	0.53
Region = South	0.42	0.41	0.28
Region = Midwest	0.25	0.13	0.15
Region = West	0.05	0.05	0.04
Father characteristics, 1900			
Literate	0.93	0.96	0.95
Employed	0.91	0.87	0.89
Occupation score | Employed	21.55	27.78	24.58
Mother characteristics, 1900			
Number of children	4.17	3.93	4.08
Literate	0.93	0.95	0.93
Employed	0.04	0.04	0.04
Occupation score | Employed	15.57	17.92	16.95
Adult characteristics, 1920			
Married	0.60	0.50	0.53
Employed	0.29	0.41	0.37
Nurse | Employed	0.03	0.02	0.02
Occupation score | Employed	21.88	23.02	22.60
Unique places	33,217	1,907	4,581
Observations	6,004,123	972,710	673,947

Notes: Table reports descriptive statistics for the 1920 linked sample, described in Section 3.3.1. Statistics are weighted using inverse propensity scores to adjust for the likelihood of linkage. Column (1) reports means for the full sample. Column (2) reports means for treated women (e.g., women who were within five miles of an opening at ages 16-20). Column (3) reports means for the remainder of the final analysis sample. Father and mother characteristics are for women with an identifiable father or mother in the 1900 census. Nurse is an indicator for work in any nursing occupation. Occupation score and nurse variables are conditional on gainful employment.

Table 2: Placebo tests with father and mother characteristics, 1920 sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Father characteristics			Mother characteristics			
	Literate	Occ score	Medicine	N children	Literate	Occ score	Medicine
Opening <5 mi. at ages 16-20	0.052	0.063	0.019	−0.012	0.046	−0.029	−0.007
	(0.114)	(0.057)	(0.036)	(0.009)	(0.119)	(0.020)	(0.005)
Controls	No	No	No	No	No	No	No
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	95.62	23.32	0.65	3.99	94.60	0.68	0.04
Observations	1,526,425	1,526,425	1,526,425	1,585,412	1,585,412	1,585,412	1,585,412

Notes: Table reports results for equation (1), estimated using OLS. Placebo outcomes are father and mother characteristics from the 1900 census: *Number of children* (as reported by the mother), *Literate* (father/mother can read and write), *Occ score* (father/mother occupation score), and *Medicine* (father/mother works in medicine). Coefficients on binary variables (*Literate* and *Medicine*) are multiplied by 100 and can be interpreted in percentage points. The sample is described in Section 3.3.1. Sample is restricted to women with an identifiable father (columns (1)-(3)) or mother (columns (4)-(7)). All specifications include town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 3: Effects of proximity on entrance into nursing, 1920 sample

	(1)	(2)	(3)	(4)	(5)	(6)
			Subcategories			
	Nurse		Trained nurse		Practical nurse	
Panel A: Full sample						
Opening <5 mi. at ages 16-20	0.126***	0.126***	0.103***	0.103***	0.023	0.023
	(0.048)	(0.048)	(0.037)	(0.037)	(0.029)	(0.029)
Controls	No	Yes	No	Yes	No	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	0.83	0.83	0.54	0.54	0.28	0.28
Observations	1,646,657	1,646,657	1,646,657	1,646,657	1,646,657	1,646,657
Panel B: Employed						
Opening <5 mi. at ages 16-20	0.447***	0.449***	0.342***	0.342***	0.105	0.106
	(0.151)	(0.151)	(0.116)	(0.116)	(0.094)	(0.094)
Controls	No	Yes	No	Yes	No	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	2.08	2.08	1.37	1.37	0.71	0.71
Observations	655,027	655,027	655,027	655,027	655,027	655,027

Notes: Table reports results for equation (1), estimated using OLS. The dependent variable in column (1) is an indicator for working in a nursing occupation in 1920. Columns (3) and (5) reports effects for the subcategories of trained and practical nurse. Columns (2), (4), and (6) report results with individual controls added. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.1. Panel A reports results for the full sample. Panel B reports results for women who are gainfully employed. All specifications include town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table 4: Effects of proximity on occupational income, 1920

	(1) Ln(occ. score)	(2) Ln(LIDO score)	(3) Song score
Panel A: Full sample			
Opening <5 mi. at ages 16-20	0.009	0.006	
	(0.007)	(0.006)	
Controls	Yes	Yes	
Fixed effects	Yes	Yes	
Mean outcome	1.25	1.01	
Observations	1,646,657	1,560,844	
Panel B: Employed			
Opening <5 mi. at ages 16-20	−0.003	−0.003	0.280
	(0.003)	(0.004)	(0.229)
Controls	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Mean outcome	3.13	2.77	74.79
Observations	655,027	569,214	653,990

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are the natural log of occupation score, the natural log of LIDO score, and Song score (see Appendix Section C.2.3). The sample is native-born white women, as described in Section 3.3.1. Panel A reports results for the full sample. Panel B reports results for women who are gainfully employed. All specifications include individual controls, town fixed effects, and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table 5: Effects of proximity on entrance into nursing, 1930

	(1)	(2)	(3)	(4)	(5)	(6)
			Subcategories			
	Nurse		Trained nurse		Practical nurse	
Panel A: Full sample						
Opening <5 mi. at ages 16-20	0.058	0.059	0.035	0.036	0.023	0.023
	(0.051)	(0.051)	(0.039)	(0.039)	(0.033)	(0.033)
Controls	No	Yes	No	Yes	No	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	0.82	0.82	0.56	0.56	0.26	0.26
Observations	1,326,420	1,326,420	1,326,420	1,326,420	1,326,420	1,326,420
Panel B: Employed						
Opening <5 mi. at ages 16-20	0.371	0.373	0.186	0.187	0.184	0.185
	(0.231)	(0.231)	(0.175)	(0.175)	(0.155)	(0.155)
Controls	No	Yes	No	Yes	No	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	3.16	3.16	2.16	2.16	1.00	1.00
Observations	346,121	346,121	346,121	346,121	346,121	346,121

Notes: Table reports results for equation (1), estimated using OLS. The dependent variable in column (1) is an indicator for working in a nursing occupation in 1930. Columns (3) and (5) reports effects for the subcategories of trained and practical nurse. Columns (2), (4), and (6) report results with individual controls added. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.2. Panel A reports results for the full sample. Panel B reports results for women who are gainfully employed. All specifications include town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table 6: Effects of proximity on occupational income, 1930

	(1) Ln(occ. score)	(2) Ln(LIDO score)	(3) Song score
Panel A: Full sample			
Opening <5 mi. at ages 16-20	−0.011*	−0.005	
	(0.007)	(0.006)	
Controls	Yes	Yes	
Fixed effects	Yes	Yes	
Mean outcome	0.82	0.68	
Observations	1,326,420	1,292,775	
Panel B: Employed			
Opening <5 mi. at ages 16-20	−0.002	−0.004	−0.115
	(0.005)	(0.004)	(0.295)
Controls	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Mean outcome	3.13	2.81	74.87
Observations	346,121	312,476	344,313

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are the natural log of occupation score, the natural log of LIDO score, and Song score (see Appendix Section C.2.3). The sample is native-born white women, as described in Section 3.3.2. Panel A reports results for the full sample. Panel B reports results for women who are gainfully employed. All specifications include individual controls, town fixed effects, and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

A Historical Appendix

A.1 Descriptive Statistics of Nurses

In this section, I characteristics the population of nurses in the complete-count census from 1870 to 1930. In Table A.1, I examine statistics for all nurses. From 1870 to 1930, nurses grew to become the largest occupation within the medical sector. The shift from private home nursing to hospital nursing is pronounce, especially between 1900 and 1910. By 1930, two-thirds of nurses were categorized as professional, and 40 percent were working in a hospital setting. Though nurses were 91 percent female in 1870, this percentage increased to 98 percent by 1930. We see an increase in the share of nurses who were white and native-born. The age distribution of nurses shifted notably during this period. Before 1900, just 20 percent of nurses were ages 18 to 24. This share rose to 33 percent by 1930. Much of this shift is likely due to both the spread of nurse training and the rise in high school graduation rates (Goldin 1998). Finally, I report the share of nurses currently married and with children. Note that this sample includes a small number of men as well as practical nurses. The share of nurses who were currently married did not change sharply during this period, however, the share married among trained female nurses was substantially lower.

I use a linked sample of nurses to describe who became a nurse from 1870 to 1930. This analysis uses FamilySearch links published by Buckles, Haws, et al. (2023). I take a backward-linking approach, matching nurses in the complete-count census backward to childhood. This allows me to study the background characteristics of nurses in the population and to draw comparisons with other historical occupations. Table A.2 reports match rates across census years for female nurses ages 15 to 64. Match rates are higher in later censuses. Native-born women are matched at higher rates than foreign-born women, in part because foreign-born women may not be present in the earlier census. In more recent census years, I still observe match rates of 10 to 30 percent for foreign-born nurses. In Figure A.3, I plot the percent of nurses (ages 18 to 25) from different backgrounds, proxied by father's occupation in childhood. Young nurses became increasing selected from families with professional fathers. This share rose from less than 10 percent of young nurses in 1870 to 15 percent of young nurses in 1930. Women also became less likely to come from manual or farming backgrounds.

A.2 Alternatives to Nursing

Exploring patterns of occupational sorting is key to understanding the effects of nurse training on economic status. Many women likely had other careers in mind, yet there are few sources that characterize the types of work women would have done had they not trained as nurses. In a report sponsored by the Indiana University School of Education, Blazier (1924) surveyed 250 trained nurses and asked about prior occupation. Only 35 percent reported a prior occupation, half of whom worked as a teacher and one third worked in the clerical sector. A far more extensive survey was conducted by the Committee for the Study of Nursing Education, sponsored by the Rockefeller Foundation in 1918 (Committee for the Study of Nursing Education 1923). Of the 15 training

schools surveyed with records of prior work for their students, the most common industries were teaching, clerical, or healthcare (*e.g.*, children's nurse, dental assistant). I explore the effect of proximity to nurse training on entrance into alternative occupations in Section 5.2.

A.3 Regulation of Nursing

Nursing became a regulated occupation in the first decade of the 20$^{\text{th}}$ century. In North Carolina, Mary Lewis Wyche, a nurse, is thought to have initiated the push for regulation by founding North Carolina's State Nurses Association in 1902. In 1903, North Carolina passed a certification law requiring that all nurses who wished to practice under the title of "registered nurse" (R.N.) pass an examination administered by the state nursing board. Successful applicants could practice as R.N. in any county in the state, and would be included in the county clerk's register of trained nurses. Use of the title of R.N. without proper paperwork was punishable by a $50 fine or no more than 30 days imprisonment (Pollitt and Miller 2010).

Within one year, New Jersey, New York, and Virginia followed the example of North Carolina. Figure A.4 plots the cumulative adoption of laws concerning registered nurses in the U.S. from 1890 to 1940, sourced from Carollo (2024). All states passed laws regulating nurses by 1940. By contrast, midwifery was regulated by 20 percent of states in 1900 and 60 percent of states in 1940. Figure A.5 plots the year a certification law for registered nurses was adopted in each state. Early adopters were more likely to be mid-Atlantic states, and late adopters were more likely to be states that recently joined the Union. In Table A.3, I identify the state characteristics that predict certification law adoption. Column (1) reports the results from a regression of adoption year on characteristics measured in 1900. Population and literacy emerge as statistically significant predictors, with larger and more literate states adopting earlier. In columns (2) and (3), I separately identify predictors of state adoption in the 1900s and the 1910s. The estimated coefficients largely align with column (1), though only literacy is statistically significant in these specifications. These findings broadly align with Carollo et al. (2022), who find that larger labor markets and the presence of state associations predict the passage of licensing laws in many occupations.

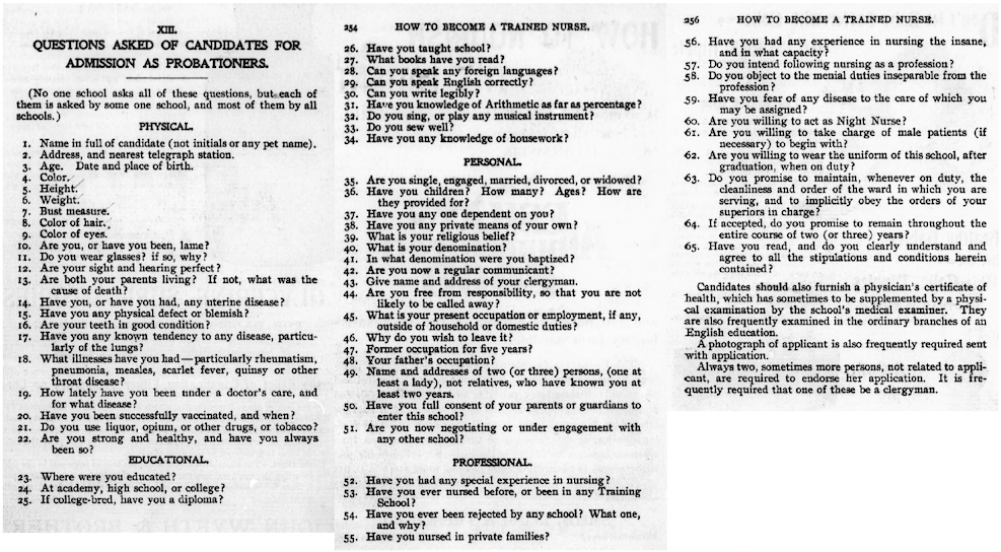
Laws from this period were permissive: Individuals who didn't obtain certification could still practice as a nurse under a less formal title like "practical nurse" or "nursing aide" (Law and Kim 2005; Law and Marks 2017). It was common for the laws to include provisions for older nurses who did not have access to their same standards of formal training. For example, Massachusetts passed its certification law in 1910 and stipulated that nurses with at least five years of nursing experience could become registered without a formal examination. Nurses who moved across state lines could remain certified, so long as the standards of their prior state were acceptable in their new state. Though certification was voluntary, advocates for reform viewed them as an initial step furthering the professionalization of nurses (Shannon 1975). Opposition to these laws was common among untrained nurses and graduates of low quality schools, who were fearful of losing work. Additionally, physicians were concerned that the recognition of nursing as a regulated profession would threaten their authority and embolden nurses to push for further reform (Rosenberg 1987).

As the medical field developed, new nurse categories emerged and states passed stronger regulation. Licensed practical nurses (LPNs, also known as licensed attendants) emerged out of a growing need to supply basis nursing services during the interwar period (Deming 1950). States began regulating practical nursing in the 1920s (Carollo 2024). In the 1940s and 1950s, states shifted from permissive to mandatory nurse licensure, and literature suggests that this shift increased nurse wages and didn't reduce minority participation (Law and Marks 2017). The certified nursing assistant (CNA) formalized as an occupation during the mid to late 20$^{\text{th}}$ century, when community colleges and vocational schools began offering CNA certification alongside other nursing programs.

Figure A.1: Student nurses in Indianapolis, IN

Source: Blazier (1924).

Figure A.2: Questions for admission to nurse training programs in 1898



XIII.

QUESTIONS ASKED OF CANDIDATES FOR ADMISSION AS PROBATIONERS.

(No one school asks all of these questions, but each of them is asked by some one school, and most of them by all schools.)

PHYSICAL.

1. Name in full of candidate (not initials or any pet name).
2. Address, and nearest telegraph station.
3. Age. Date and place of birth.
4. Color.
5. Height.
6. Weight.
7. Bust measure.
8. Color of hair.
9. Color of eyes.
10. Are you, or have you been, lame?
11. Do you wear glasses? if so, why?
12. Are your sight and hearing perfect?
13. Are both your parents living? If not, what was the cause of death?
14. Have you, or have you had, any uterine disease?
15. Have you any physical defect or blemish?
16. Are your teeth in good condition?
17. Have you any known tendency to any disease, particularly of the lungs?
18. What illnesses have you had—particularly rheumatism, pneumonia, measles, scarlet fever, quinsy or other throat disease?
19. How lately have you been under a doctor's care, and for what disease?
20. Have you been successfully vaccinated, and when?
21. Do you use liquor, opium, or other drugs, or tobacco?
22. Are you strong and healthy, and have you always been so?

EDUCATIONAL.

23. Where were you educated?
24. At academy, high school, or college?
25. If college-bred, have you a diploma?

26. Have you taught school?
27. What books have you read?
28. Can you speak any foreign languages?
29. Can you speak English correctly?
30. Can you write legibly?
31. Have you knowledge of Arithmetic as far as percentage?
32. Do you sing, or play any musical instrument?
33. Do you sew well?
34. Have you any knowledge of housework?

PERSONAL.

35. Are you single, engaged, married, divorced, or widowed?
36. Have you children? How many? Ages? How are they provided for?
37. Have you any one dependent on you?
38. Have you any private means of your own?
39. What is your religious belief?
40. What is your denomination?
41. In what denomination were you baptized?
42. Are you now a regular communicant?
43. Give name and address of your clergyman.
44. Are you free from responsibility, so that you are not likely to be called away?
45. What is your present occupation or employment, if any, outside of household or domestic duties?
46. Why do you wish to leave it?
47. Former occupation for five years?
48. Your father's occupation?
49. Name and addresses of two (or three) persons, (one at least a lady), not relatives, who have known you at least two years.
50. Have you full consent of your parents or guardians to enter this school?
51. Are you now negotiating or under engagement with any other school?

PROFESSIONAL.

52. Have you had any special experience in nursing?
53. Have you ever nursed before, or been in any Training School?
54. Have you ever been rejected by any school? What one, and why?
55. Have you nursed in private families?

56. Have you had any experience in nursing the insane, and in what capacity?
57. Do you intend following nursing as a profession?
58. Do you object to the menial duties inseparable from the profession?
59. Have you fear of any disease to the care of which you may be assigned?
60. Are you willing to act as Night Nurse?
61. Are you willing to take charge of male patients (if necessary) to begin with?
62. Are you willing to wear the uniform of this school, after graduation, when on duty?
63. Do you promise to maintain, whenever on duty, the cleanliness and order of the ward in which you are serving, and to implicitly obey the orders of your superiors in charge?
64. If accepted, do you promise to remain throughout the entire course of two (or three) years?
65. Have you read, and do you clearly understand and agree to all the stipulations and conditions herein contained?

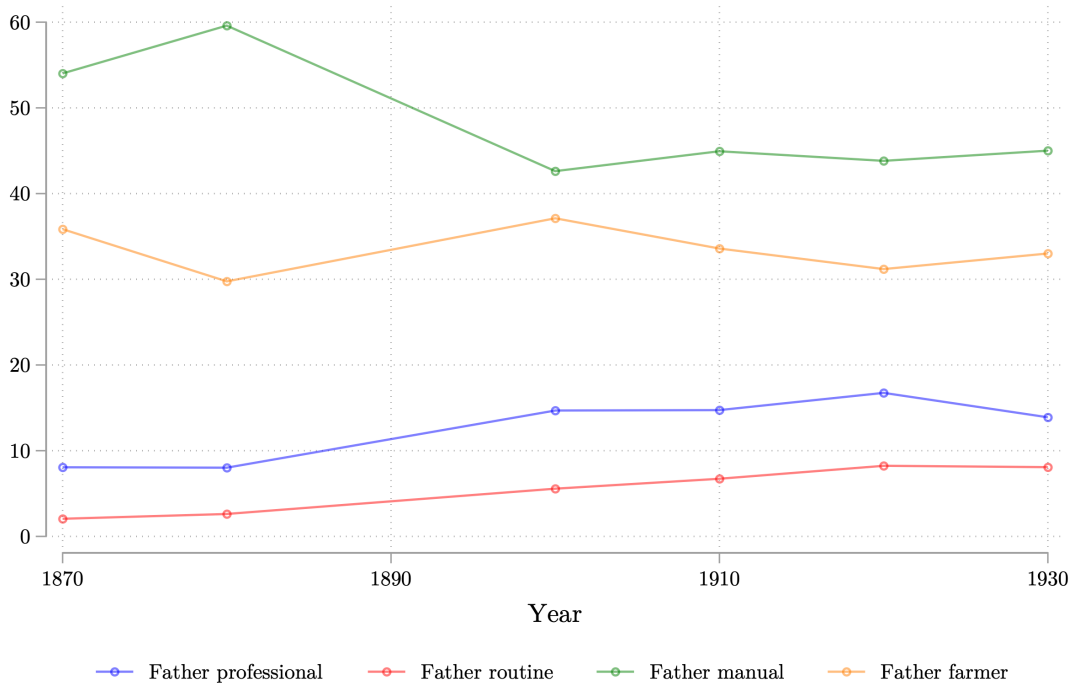
Candidates should also furnish a physician's certificate of health, which has sometimes to be supplemented by a physical examination by the school's medical examiner. They are also frequently examined in the ordinary branches of an English education.

A photograph of applicant is also frequently required sent with application.

Always two, sometimes more persons, not related to applicant, are required to endorse her application. It is frequently required that one of these be a clergyman.

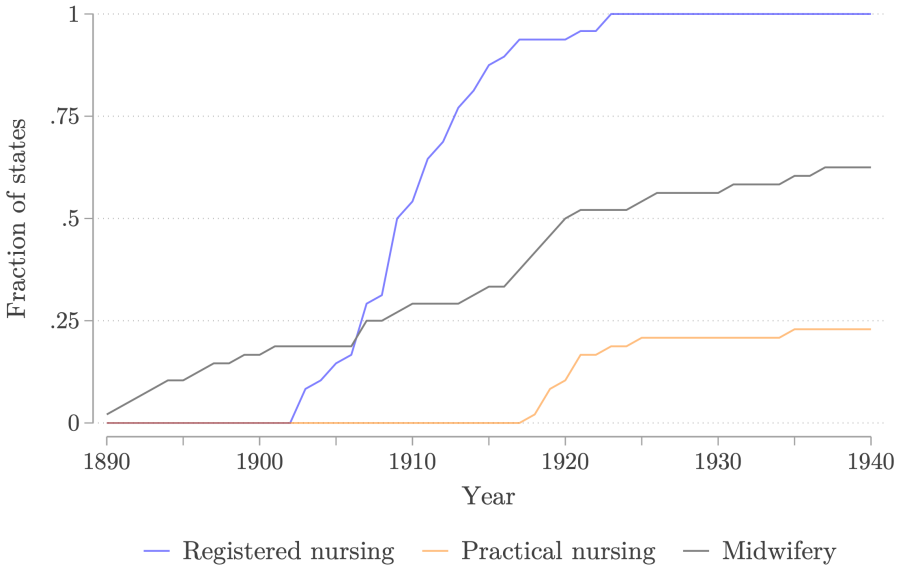
Source: Hodson (1898).

Figure A.3: Background of nurses over time, 1870 to 1930



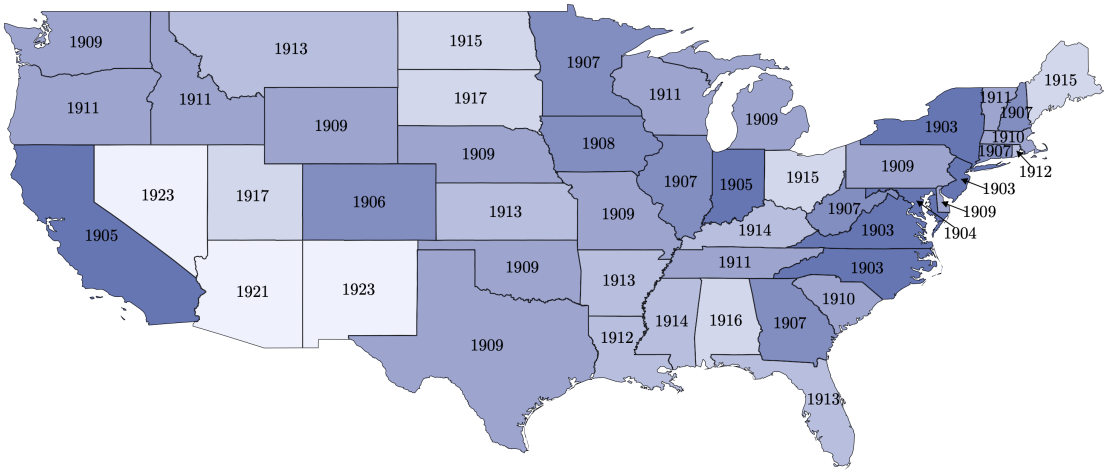
Notes: Figure reports the background of nurses in the census from 1870 to 1930. Nurses are linked backward to their childhood census using FamilySearch crosswalks (Buckles, Haws, et al. 2023). Background is proxied using father's occupation, as observed in the most recent childhood census at ages 0 to 15. Father's occupation is grouped into four categories using the occupational scheme of Song et al. (2020): professional, routine (non-manual), manual, and farming. The sample is limited to all nurses at ages 18 to 25.

Figure A.4: Adoption of certification laws for registered nurses



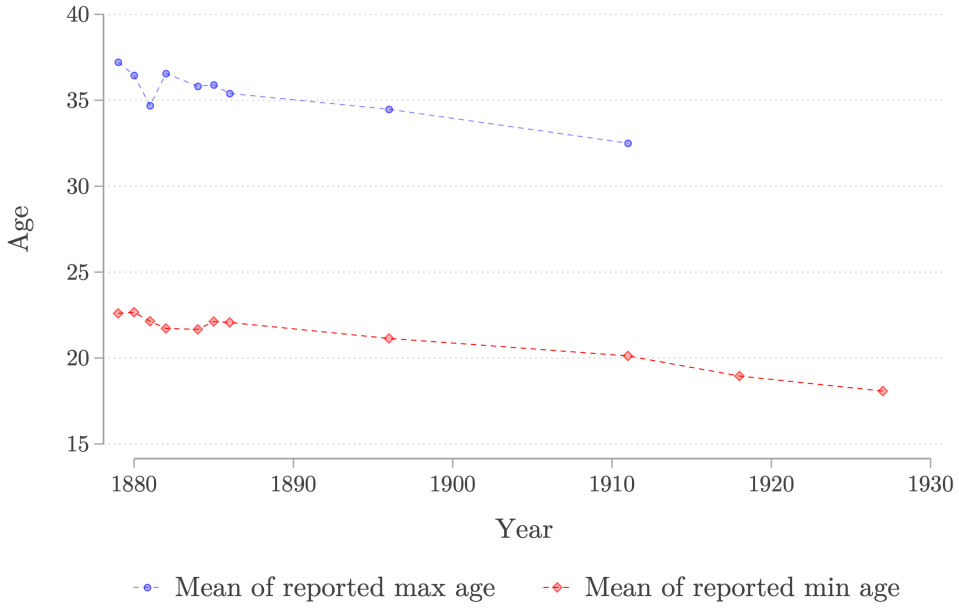
Notes: Figure reports the fraction of states with any law regulating the practice of registered nursing, licensed practical nursing, or midwifery from 1890 to 1940. Laws regulating registered and practical nurses are sourced from Carollo (2024). Laws regulating midwives are sourced from Anderson et al. (2020). Alaska, the District of Columbia, and Hawaii are excluded.

Figure A.5: Timing of certification laws for registered nurses



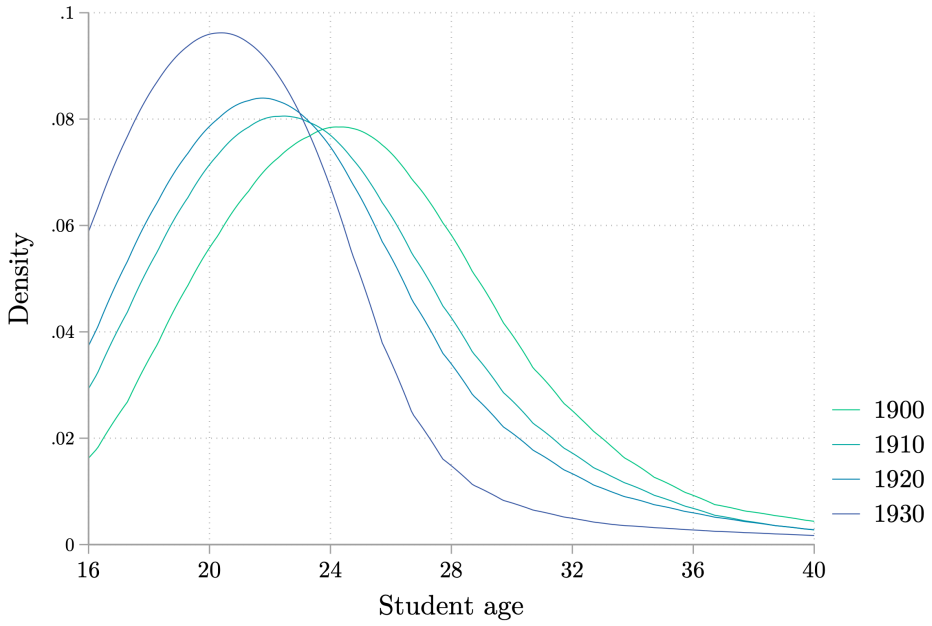
Notes: Figure plots state adoption of certification laws for registered nurses. Year of first adoption is reported, with lighter colors indicating later adoption years. Laws regulating registered nurses are sourced from Carollo (2024). Alaska, the District of Columbia, and Hawaii are excluded.

Figure A.6: Reported minimum and maximum ages for student nurses



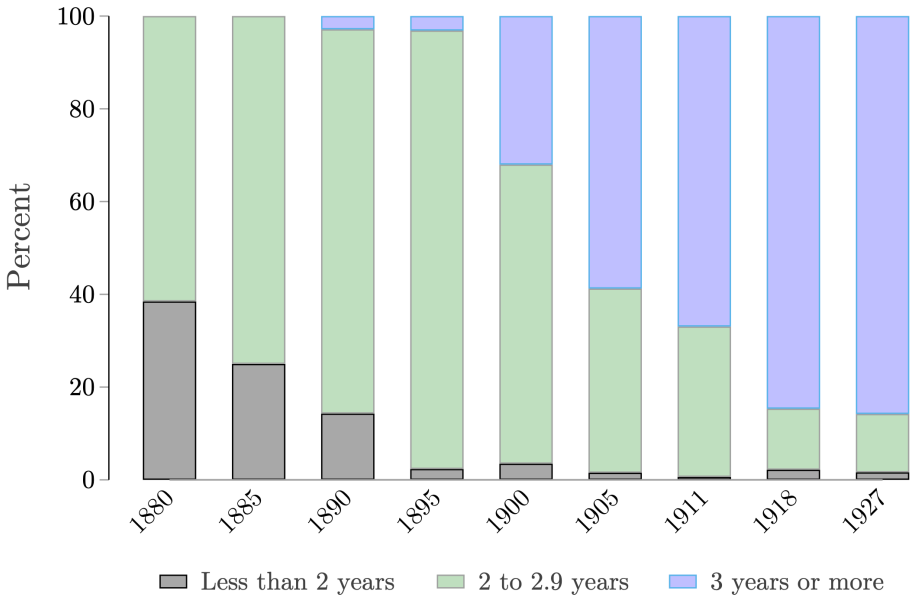
Sources: U.S. Bureau of Education. See Appendix Section C.1.
Notes: Figure shows the minimum and maximum ages that nurse training schools accepted over time. The mean in each year is shown. Maximum ages were not reported after 1911.

Figure A.7: Age distributions of student nurses, 1900-1930



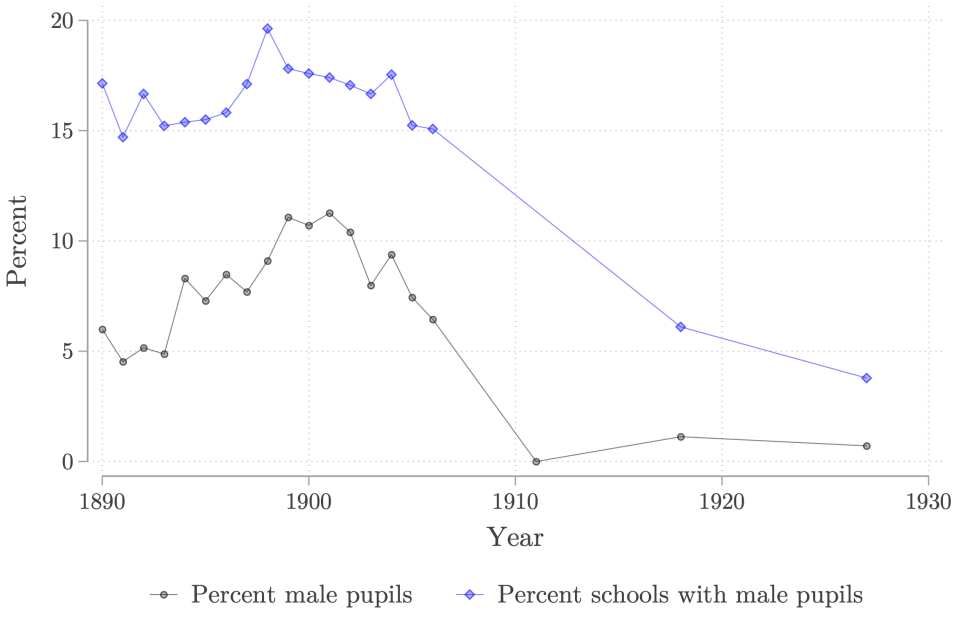
Notes: Figure shows the age distribution of student nurses in each complete-count census from 1900 to 1930. Sample is limited to white, female, native-born student nurses (1950 occupation code of 59). Ages under 16 and over 40 are binned at 16 and 40, respectively. The distribution is plotted separately for each census, estimated using kernel density with bandwidth of 3.

Figure A.8: Length of nurse training over time



Sources: U.S. Bureau of Education. See Appendix Section C.1.
Notes: Figure shows the reported length of nurse training over time. Length of training is grouped into three bins: Less than 2 years, 2 to 2.9 years, and 3 years or more. The percent of schools that fall in each bin is reported in each year.

Figure A.9: Male nursing students over time



Sources: U.S. Bureau of Education. See Appendix Section C.1.
Notes: Figure shows in black the percent of students who were male in each year, aggregated across all training schools. The percent of schools in each year that had any male students is shown in blue.

Table A.1: Descriptive statistics of nurses in the census

	(1)	(2)	(3)	(4)	(5)	(6)
	Census year					
	1870	1880	1900	1910	1920	1930
Nurses (count)	11,058	23,406	88,346	178,018	250,563	417,344
Percent of medical workers	11.8	16.4	31.9	44.3	50.2	59.2
Subcategories (percent)						
Trained nurses	1.4	1.7	12.3	43.8	51.8	65.7
Practical nurses	98.6	98.3	87.7	56.2	48.2	34.3
In hospital	1.2	1.2	2.4	23.4	29.3	38.9
In private home	95.6	98.3	86.3	48.3	38.3	34.3
Demographics (percent)						
Female	91.0	94.8	92.4	94.3	96.1	97.7
White/non-Hispanic	71.8	69.7	82.3	87.7	93.1	94.2
Foreign-born	28.6	25.4	23.4	22.0	16.8	13.8
Ages 15-17	16.5	14.0	7.0	4.3	1.9	1.3
Ages 18-24	18.7	20.1	22.8	24.6	26.3	33.5
Ages 25-34	15.4	14.6	27.2	29.8	29.5	24.4
Ages 35-44	15.4	14.6	16.8	18.0	18.8	17.5
Ages 45-65	34.2	36.7	26.1	23.2	23.4	23.3
Currently married	7.7	13.1	12.9	14.2	13.8	17.4
Has children	23.1	24.2	18.5	16.5	13.8	13.3

Notes: Table reports descriptive statistics (means) for nurses in the complete-count census from 1870 to 1930. Sample includes all nurses aged 15 to 65. Nurses (count) includes trained and practical nurses. Medical workers includes nurses as well as the medical professions (dentist, optometrist, osteopath, pharmacist, physician).

Table A.2: Match rates of female nurses across census years

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Age eligible									
					Linking year				
N	Base year	1850	1860	1870	1880	1900	1910	1920	1930
8,730	1870	0.216	0.204						
19,498	1880	0.265	0.254	0.299					
74,037	1900	0.242	0.277	0.324	0.361				
149,997	1910	0.223	0.252	0.291	0.343	0.392			
228,740	1920	.	0.277	0.296	0.339	0.410	0.449		
362,517	1930	.	.	0.294	0.330	0.374	0.422	0.444	
413,525	1940	.	.	.	0.341	0.366	0.399	0.447	0.513
Panel B: Age eligible U.S.-born									
					Linking year				
N	Base year	1850	1860	1870	1880	1900	1910	1920	1930
6,294	1870	0.253	0.213						
14,538	1880	0.305	0.276	0.321					
56,691	1900	0.313	0.335	0.390	0.432				
116,786	1910	0.287	0.311	0.354	0.421	0.440			
190,889	1920	.	0.334	0.360	0.411	0.460	0.484		
311,902	1930	.	.	0.358	0.403	0.436	0.469	0.479	
377,860	1940	.	.	.	0.414	0.421	0.440	0.472	0.525
Panel C: Age eligible foreign-born									
					Linking year				
N	Base year	1850	1860	1870	1880	1900	1910	1920	1930
2,436	1870	0.136	0.181						
4,960	1880	0.177	0.202	0.233					
17,346	1900	0.045	0.106	0.138	0.152				
33,211	1910	0.031	0.048	0.088	0.112	0.226			
37,851	1920	.	0.018	0.047	0.078	0.171	0.275		
50,615	1930	.	.	0.029	0.042	0.116	0.156	0.225	
35,665	1940	.	.	.	0.020	0.076	0.138	0.199	0.386

Notes: Table reports match rates of female nurses aged 15-64 across census years. Each base year is linked to prior census years using the FamilySearch crosswalks (Buckles, Haws, et al. 2023). In Panel A, the sample includes nurses old enough in the base year to be linked (for example, nurses aged 20 and up for the link from 1940 back to 1920). In Panel B, I restrict the sample in Panel A to native-born nurses. In Panel C, I restrict the sample in Panel A to foreign-born nurses. I report total nurses in the base year in the left-most column. Each cell uses a different subset of age-eligible observations.

Table A.3: Predictors of nurse certification law adoption

	(1)	(2)	(3)
	Year adopted	Adopted 1901-1910	Adopted 1911-1920
Ln(population)	−2.115***	0.143	0.246
	(0.667)	(0.089)	(0.163)
Percent black	−0.006	−0.018	0.023
	(0.137)	(0.016)	(0.020)
Percent urban	0.012	0.001	0.001
	(0.064)	(0.008)	(0.004)
Percent foreign-born	0.032	−0.011	0.010
	(0.105)	(0.015)	(0.022)
Percent literate	−0.276***	0.015	0.040**
	(0.075)	(0.014)	(0.012)
Percent women employed	−0.116	0.037	−0.024
	(0.265)	(0.035)	(0.027)
Ln(physicians per capita)	3.149	−0.233	−0.780*
	(2.663)	(0.307)	(0.367)
Ln(nurses per capita)	−1.034	0.086	−0.216
	(1.044)	(0.155)	(0.247)
Training schools per 100k	−0.998**	0.085	0.044*
	(0.383)	(0.063)	(0.021)
Women's suffrage law	−1.541	0.072	−0.178
	(2.181)	(0.295)	(0.227)
Region FE	Yes	Yes	Yes
Mean outcome	1911	0.54	0.86
F-statistic	16.048	6.926	46.532
Observations	48	48	22

Notes: This table reports the predictors of nurse certification law adoption at the state level. Each column shows results from a multivariate regression estimated using OLS, where the dependent variable is the year in which a state first adopted a nurse certification law (column (1)), an indicator for adoption from 1901 to 1910 (column (2)), and an indicator for adoption from 1911 to 1920 (column (3)). Explanatory variables are measured in 1900 for columns (1) and (2) and 1910 for column (3). Columns (1) and (2) include all states (excluding Alaska, Hawaii, and the District of Columbia), and column (3) restricts to states that had not adopted a law by 1910. Percent and per capita variables are measured using the entire state population. Percent women employed is measured using women aged 15-64, where employment is defined in Appendix Section C.2.2. Training schools is measured as the count of schools per 100,000 women aged 15-64. The mean of the outcome is reported, along with the F-statistic from a joint test of significance for the variables shown. All specifications include region fixed effects. Heteroskedasticity-robust standard errors in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

B Case Study of Nurse Training School Alumnae

Institutional archives are a rich source of data and contain information not found in other sources. Prior to 1940, the decennial census did not record information on educational attainment for adults. In 1940, individuals reported years of education, but didn't provide information on type of degree or institution. In addition, the decennial nature of the census makes it difficult to capture short-run labor market dynamics.

B.1 Archival Records from Three Training Schools

In this section, I access and digitize the archival records of nurse training schools. Training schools kept updated lists of graduates, similar to colleges and universities.[B.1] These records were published in annual reports that listed both current graduates and information on prior alumnae. An example of the original records is shown in Figure B.1. From these data, I observe a nurse's maiden name, married name, graduating cohort, location, and occupation. By comparing alumnae information in published reports that are updated over time, I can answer questions like: *What are the rates of marriage for graduates over time? What share of graduate nurses work in private homes or in hospitals?* This approach relies on graduates updating their information, which they frequently did.[B.2] One reason is that new graduates frequently took jobs as private or visiting nurses. The training school could refer work to these women using its register of graduates, and so it was important for nurses to keep their contact information up to date. A second reason is that each school had an alumni association, which mailed announcements as their primary form of communication.

I collect and digitize records from three nurse training schools: the Bellevue Hospital Training School for Nurses in New York City, the Presbyterian Hospital School of Nursing in Chicago, and the Massachusetts General Hospital School of Nursing in Boston (henceforth Bellevue, Presbyterian, and MGH). These schools were chosen because their records were previously scanned and made public. These schools have the added benefit of being prominent, geographically diverse institutions: Bellevue and MGH were two of the earliest and most prestigious schools, both founded in 1873. Presbyterian was founded in 1903 as part of Rush Medical College. The reports contain graduates of Bellevue from 1875 to 1919, graduates of Presbyterian from 1906 to 1932, and graduates of MGH from 1875 to 1939.

To generate usable data from the scanned records, I use the functionality provided by LayoutParser (Shen et al. 2021). Optimal character recognition is used to extract the text from each page. I use Faster R-CNN to match text to fields in the portions of the records that have cleanly delineated columns. The remaining text is sorted into fields based on page location and regular expressions. I combine the three schools to form a dataset of nurse training school graduates over time, where the length of time I can see a given graduating cohort is limited by the available

[B.1] See Bleemer and Quincy (2022) for an example of historical records from California universities.
[B.2] Only 5 percent of women report no information after graduation, and half update their information at least once.

data. For instance, the Bellevue records end in 1919, so I only observe the graduating class of 1919 once. Table B.1 summarizes the processed records for each school and includes the dates of hospital and nurse training school founding. The full dataset contains 44,429 observations of 5,499 graduates. Without a broader data collection effort, it is difficult to know how these records compare to the average training school. Given that Bellevue and MGH were understood to be high-quality institutions, I expect the following results to be positively selected in terms of career outcomes.

From these data, I construct several outcomes. *Married* is an indicator for whether a woman was listed with a change of surname (for example, in Figure B.1, "S. W. Adams . . . Now Mrs. F. W. Conrad"). I allow this indicator to change over time, for instance if a woman is widowed and returns to work. I therefore also define *Ever married* as an indicator for whether a woman has ever been married by a given age (with *Never married* as the converse). *Employed* is an indicator for whether a woman is listed with an occupation. This measure includes military service and excludes women recorded as attending school or at home. I impute age and birth cohort using graduation year, assuming all women graduated at 22 years of age (Committee for the Study of Nursing Education 1923). While this assumption is consistent with some historical records, I may be underestimating age, especially for early cohorts.

B.2 Patterns of Marriage and Employment

I characterize the marriage patterns of nurses relative to broad trends in the female labor force. In this exercise, I cannot explicitly subset by race or nativity. However, nurses were fairly homogeneous during this period: 90 percent of professional nurses were white women (Figure D.2), and roughly 25 percent were foreign-born. Note that "deceased" is an outcome listed in the records, so statistics are not conditional on survival. I first collapse the data by 10-year age bin and decade observed. For instance, a woman who graduated from Bellevue in 1880 would appear in 1890 with an imputed age range of 30-39. Table B.2, Panel A reports marriage rates for graduate nurses observed in each decade from 1890 to 1940.

In 1890, 13 percent of graduate nurses ages 20 to 29 were recorded as married. By 1940, this rate has increased to 39 percent, which reflects the decline in age at first marriage documented by Goldin (2021). It is informative to plot the fraction of graduates never married, to allow for a more direct comparison with prior literature. As shown in Figure B.2, the fraction never married falls across all age groups from the 1870 birth cohort to the 1890 birth cohort. Even though the main cohorts I observe are from a slightly earlier period, the downward trend is consistent with the cohorts studied by Goldin (2021) (page 34). Note also that there are several differences that make a direct level comparison difficult, including that these women are not college educated and may have different preferences for marriage.

In Table B.2, Panel B, I report employment among currently married women. In the archival records, these women were listed with a married name but also reported an occupation. Rates of employment while married are very low. I calculate that 50 percent of ever married women exited the labor force entirely upon marriage, and roughly 30 percent never worked and married soon

after graduation. Prior to 1940, rates of employment among married women are typically higher at older ages. In 1940, the decline reverses as more married nurses remain in the workforce. These patterns line up with Goldin (1990), who reports a participation rate of roughly 20 percent among married women in the 1906-1915 birth cohorts (page 121). I additionally calculate that 50 percent of graduates exit the labor force entirely upon marriage

Table B.2, Panel C shows employment among women never recorded as married. I observe very high rates of participation in the labor force for the 20-29 age range (between 65 and 85 percent). This is not surprising given that this sample of women received two to three years of medical training that was persistently in high demand from hospitals and households. These numbers are almost uniformly higher than Goldin (1990), which reports participation rates of 53 and 68 percent for similar women in 1890 and 1920 (page 18). This evidence suggests that even among a group of medically trained and highly selected women, we see the harbingers of broader changes in the labor market. Graduates born between 1860 and 1870 have very low marriage rates, but beginning with the 1870 cohort, a rising number of graduates are married in their twenties. A portion of this change will reflect selection into nursing: As nursing schools spread and working hours were lowered, nursing became a more attractive option and less tied to its laborious, monastic origins (Rosenberg 1987). By the 1900 birth cohort, the fraction of unmarried graduates at age 40 had fallen from 0.70 to 0.45.

B.3 Types of Employment

Next, I investigate what fraction of nurses worked in different occupations. I construct indicators for employment in private homes and businesses, hospitals, and the military. I also identify women working in a managerial role (such as training school superintendent). These categories are not defined to be mutually exclusive. Conditional on reporting any occupation, Figure B.3 plots the fraction of nurses aged 20-39 in these categories over time. Before 1900, the majority of graduates entered work as a private duty nurse in either homes or businesses. With the rise of larger, more modern hospitals, the profession shifted from private to hospital employment. Between 10 and 20 percent of young nurses took on managerial roles, which reflects the need for leadership during a period when training schools were expanding. Bellevue and MGH were considered top-tier programs, so more of their graduates took on leadership roles within nursing relative to Presbyterian. The majority of records in the 1930s come from Presbyterian, which explains the drop in managers observed in these years. Finally, we see the impact of World War I on nursing employment in the military. In the MGH alumni records, nearly 40 percent of the MGH graduating classes of 1915-1917 spent time in the military, primarily as a nurse in the Army Nurse Corps (Kalisch and Kalisch 2004).

B.4 Locations of Graduates

Finally, I characterize the locations of graduates. Note that I do not observe a graduate's hometown, just reported location after graduating from nurse training school. Between 50 and 70 percent of graduates remained in the state where they trained, a percentage that remains relatively stable as

the cohorts age. City of residence isn't always recorded, but the MGH records contain the best coverage of this variable. Over time, between 25 and 45 percent of MGH alumnae were located in the city of Boston proper, not including the surrounding towns. This finding is in line with my main results that demonstrate the effect of training schools on local nurse supply. It is also consistent with prior literature that shows how cities offered women better independent work and marriage prospects than small towns (Withrow 2021).

Figure B.1: Sample of Bellevue Training School records

LIST OF GRADUATES FROM 1875 TO 1886

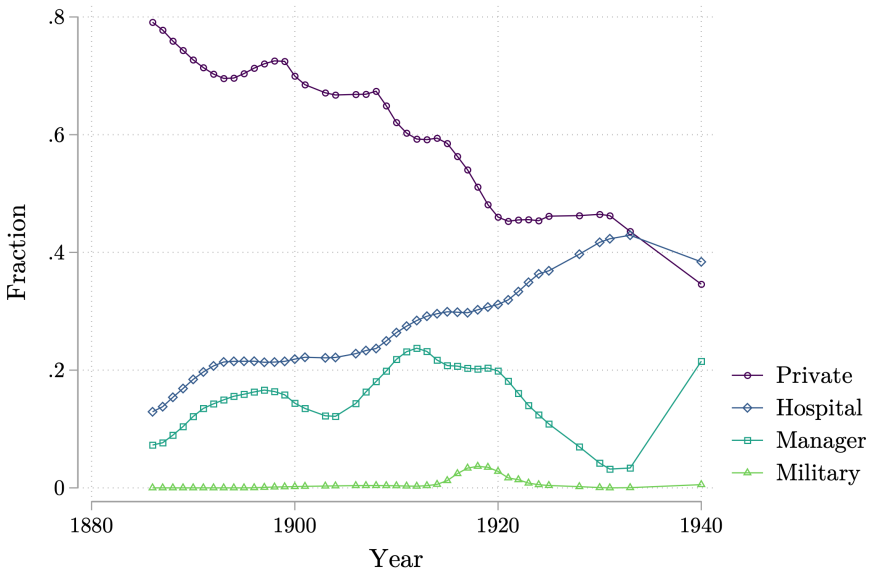
Name.	Graduated.	Occupation.	Residence.
Miss A. E. Andrews	1878	Matron, City Hospital	Lawrence, Mass.
" J. J. Angell	1879	Private nurse	New York City.
" E. A. Aldrich	1880	District nurse	" "
" S. E. Allen	1881	Matron, lying-in hospital	London, Eng.
" S. W. Adams	1882	Now Mrs. F. W. Conrad, at home	Santa Barbara, Cal.
" A. M. Alton	1884	Private nurse	New York City.
" Jennie Arthur	1884	At home	Brynmaur, Pa.
" H. M. Alford	1882	Now Mrs. Devine, at home	New York City.
" Lucretia Buck	1875	Private nurse	" "
" Annie Brennan	1876	" "	" "
" Mary Brady	1876	Now Mrs. Cummings, private nurse	" "
" Mary Bestow	1876	At home	Coolville, Ohio.
" R. G. Belt	1878	Private nurse	New York City.

Source: Annual Report, Bellevue Hospital Training School for Nurses, 1887.

Figure B.2: Fraction of graduates never married by birth cohort

Notes: This figure shows the fraction of women measured as never married in training school alumnae records. The available data cover the 1860 to 1910 birth cohorts, where birth year is estimated using graduation year. The available time range is smaller for older age bins because the records end in 1940, which censors more recent cohorts. For details on the specific years and cohorts available, see Table B.1. Data have been smoothed with a bandwidth of 0.5.

Figure B.3: Fraction of graduates employed in private, hospital, managerial, or military roles



Notes: This figure shows the fraction of graduates employed in private, hospital, managerial, or military roles. Categories are not mutually exclusive. The sample is conditional on any listed employment among graduate nurses aged 20-39. For details on the specific years available, see Table B.1. Data have been smoothed with a bandwidth of 0.2.

Table B.1: Statistics of nurse training schools with available records

	Hospital founded	First nursing graduates	Mean cohort size	Total graduates	Observations
Bellevue[1]	1812	1875	33.5	1,249	12,166
Presbyterian[2]	1883	1906	54.7	1,478	8,899
MGH[3]	1811	1875	51.2	2,772	23,364
Total				5,499	44,429

Notes: This table reports statistics for Bellevue, Presbyterian, and MGH training schools for nurses.
[1]Data on Bellevue come from annual reports published starting in 1875. Graduating classes of 1875-1919 are compiled from reports published in 1886-93, 1895-1901, 1906-08, 1916-17, and 1919.
[2]Data on Presbyterian come from annual reports published starting in 1884. Graduating classes of 1906-32 are compiled from reports published in 1907-16, 1920-25, 1928, 1930, 1931, and 1933.
[3]Data on MGH come from annual reports published starting in 1888. Graduating classes of 1875-1939 are compiled from reports published in 1888-95, 1899, 1903-04, 1906-20, 1922, and 1940.

Table B.2: Marriage and employment of graduates by age bin and decade

Married				
	20-29	*30-39*	*40-49*	*50-59*
1890	0.13	0.27	.	.
1900	0.11	0.25	0.29	.
1910	0.18	0.32	0.32	0.30
1920	0.22	0.42	0.38	0.34
1930	0.28	0.58	0.62	.
1940	0.39	0.55	0.54	0.50
Employed | Currently married				
	20-29	*30-39*	*40-49*	*50-59*
1890	0.11	0.29	.	.
1900	0.05	0.06	0.23	.
1910	0.09	0.05	0.06	0.18
1920	0.03	0.03	0.06	0.02
1930	0.08	0.02	0.02	.
1940	0.38	0.24	0.16	0.17
Employed | Never married				
	20-29	*30-39*	*40-49*	*50-59*
1890	0.77	0.70	.	.
1900	0.65	0.58	0.41	.
1910	0.83	0.64	0.47	0.28
1920	0.69	0.66	0.54	0.32
1930	0.69	0.43	0.52	.
1940	0.86	0.78	0.68	0.49

Notes: This table reports marital status and employment conditional on marital status for women in each decade and age bin from 1890 to 1940. Married is an indicator for listing a married name. Widows and divorcees will be captured if the women reverts to her maiden name. Employed is an indicator for any occupation listed in the year, including military service. Some age-year cells are missing due to varying coverage of the training school records (Table B.1).

C Data Appendix

C.1 Nurse Training Schools

I collect data on nurse training schools from a variety of historical sources. Soon after its founding in 1867, the U.S. Bureau of Education (the Bureau) began surveying industrial and vocational training programs, including nurse training schools. I therefore begin my data collection by digitizing school-level information contained in Bureau reports for the years 1879-1882 and 1884-1906. I digitize three additional Bureau reports on nurse training schools to obtain data for 1911, 1918, and 1927 (Nutting 1912; U.S. Bureau of Education 1919, 1928). I also digitize reports published by the American Nurses Association (ANA) in 1918 and 1922 (American Nurses Association 1918, 1922). Available fields include city, year of organization, number of pupils, and school information such as program length and pay (Figures C.1 and C.2). Note that in some years, school-level statistics were not published, but the Bureau still collected data on total training schools, students, and graduates (see U.S. Bureau of the Census (1975), Series B 286-290). I add these data to the series in Figure 1.

Table C.1 provides summary statistics on training schools contained in the Bureau reports. The reports are largely similar across years, with a few exceptions. First, from 1898 onward, the Bureau listed schools affiliated with insane asylums separately. Between 5 and 10 percent of schools in a given year were affiliated with asylums. I include these schools in my analysis since the training structure was similar to schools affiliated with general hospitals. Second, in 1898, 1899, 1905, 1906, and 1911, the Bureau of Education identified post-graduate (supplemental) programs. These were typically 3 to 6 month programs meant for nurse graduates to supplement their training. Very few of these programs existed (1 to 3 percent of observations in a given year). I drop these programs from my analysis since they did not cater to new student nurses.

C.1.1 Schools for male nurses

The majority of nurse training programs from this era admitted exclusively white women, however, programs for (white) male nurses did exist. I identify schools admitting male nurses using reported pupils by gender (available starting in 1889). I flag schools that trained male nurses exclusively: these schools report no female pupils and can often be identified by name (e.g., City Hospital Training School for Male Nurses). In Table C.1, the fraction of schools with any male pupils falls over time from 17 percent in 1890 to 4 percent in 1927. Over 40 percent of schools with at least one male student were located at state psychiatric hospitals. About two to five programs in a given survey year were exclusively male. In 1927, the data contain one identifiable male-only training school.

C.1.2 Schools for black nurses

The Bureau and ANA reports did not record nurse pupils by race. Nursing in the early 20th century was racially segregated, and black nurses typically trained at black-founded and black-owned

establishments (Hine 1989). To better understand the numbers of black relative to white training schools, I turn to other sources. Hodson (1898), a manual for nurses, describes black-admitting and black-only schools. This source reports five schools for black nurses and six schools that admitted black probationers, but were not necessarily black-only. Multiple editions of *The Negro Yearbook*, published from 1912 to 1952, list hospitals and nursing schools for black individuals. The edition from 1912 lists 12 training schools for black women. Finally, Wesley (2010) gathered data on hundreds of historically black hospitals from numerous sources. The author notes that substantial effort was made to capture facilities that would have housed black trainee nurses. However, the information on founding date and presence of a nurse training school is incomplete.

C.1.3 Defining proximity to nurse training

School location is consistently reported as a state and town name (Figure C.1). I match school location to the Census Place Project, which contains consistently defined historical place names, latitude, and longitude (Berkes et al. 2023). I am able to match 95 percent of cities and towns directly. For the observations that do not match, I manually code the matched location to the nearest town.[C.1]

I use each school's year of organization to measure training school growth from 1870 to 1920. This is easiest to accomplish using the 1922 ANA report, which contains the year of opening for all schools through 1922. One drawback to relying on this source is that it only contains accredited schools as of 1922, not unaccredited schools or schools that have closed. The downside to relying on the Bureau reports alone is that, to my knowledge, 1905 is the last year that the Bureau of Education published the date of opening in its statistical reports. Additionally, I find evidence of a lag in reporting, especially in the earliest years of surveying when the Bureau was still in its infancy (Figure C.3, Panel (a)). For instance, the Freedmen's Hospital School of Nursing opened in 1894, but did not appear in the Bureau reports until 1896 (Coles 1969). If a school is not surveyed until several years after opening, this approach will add significant measurement error. My approach is to use the ANA report as my primary source of openings, and to validate school counts over time using reports from the Bureau. In Figure C.3, Panel (b), I show that the two sources are well-aligned. The difference in 1918 may be partially due to the inclusion of post-graduate programs in the 1918 Bureau report.

C.2 Census Records

I use complete-count census records from 1870 to 1940 (Ruggles et al. 2020). These records exclude the 1890 census, which was unable to be salvaged following a fire in 1921. For my analysis of the labor force after 1940, I add census samples from 1950 to 2015. Complete-count data for these years

[C.1] This might occur if, for instance, a school location is listed as "Flatbush, NY", which I relabel as "New York, NY". In some cases, a school's location will map to multiple places of the same name within a state (for instance, a town and an unincorporated place). I choose the place with the largest fraction of the population across census years, as defined in Berkes et al. (2023). This affects 12 percent of observations in the training schools dataset.

are preliminary (in the case of 1950) or not yet released. I use the following samples: 1 percent samples for 1950 and 1970; 5 percent samples for 1960, 1980, 1990, and 2000; American Community Survey for 2005, 2010, and 2015. Data are sourced from IPUMS (Ruggles et al. 2022).

C.2.1 Linked census samples

Linking individuals across census years is an area of substantial methodological research (Abramitzky et al. 2021). The latest innovation in this area is using data from the FamilySearch genealogical platform to link large samples of men and women over time (Buckles, Haws, et al. 2023). I use the FamilySearch data to construct two linked samples: a 1900-1920 linked sample and a 1900-1930 linked sample. One inherent drawback to this approach is that the set of linked individuals will differ in some respects from the entire U.S. population. Bailey et al. (2020) provide a partial solution: Construct inverse propensity score weights that adjust for the likelihood of linkage. To link individuals from 1900 forward to 1920 and 1930, I take the following approach:

1. Append the linked sample to the complete 1900 census.
2. Construct a set of 1900 characteristics that may predict linkage, including:
    - Age in years
    - Length of first, middle, and last names
    - Indicator for any middle name
    - Commonness of last and first names, measured as the natural log of the number of people with that name
    - Number of siblings and indicator for any siblings
    - Indicators for race and gender
    - Length of father's first name and mother's first name
    - Indicators for state of birth
    - Indicator for foreign-born
    - Polynomials (quartic) in age, number of siblings, length of first/last name, length of father/mother name
    - Interaction of gender with race and first/last name commonness
    - Missingness indicators for first/last name, age, siblings, and father/mother first name
3. Estimate a probit with an indicator that takes on a value of 1 for the linked sample and 0 for the reference population (1900).
4. Calculate the inverse propensity score as: $\frac{(1-p)}{p} \times \frac{q}{(1-q)}$, where p is the propensity score, or the probability of being in the linked sample, and q is the linkage rate.

I plot the distribution of propensity scores for the reference population and linked samples in Figure C.4. While there is a substantial amount of overlap in the propensity score distributions, individuals with very low propensity scores (less than 0.2) may not be well-represented in the linked sample. In Tables C.3 and C.4, I examine the characteristics of the sample of white women born in the U.S. between 1885 and 1900. The linked samples are somewhat selected on characteristics that predict linkage (such as having more siblings or a middle initial). Reweighting the linked sample in column (3) aligns these observables more closely with the full sample.

C.2.2 Defining the labor force

I frequently construct samples that identify individuals who are "in the labor force". The measurement of employment and occupation has changed over time, therefore the definition of the labor force differs across census years. I adopt a definition of the labor force that is in line with prior literature (Katz and Margo 2014). I restrict to individuals aged 15 to 64 with a valid occupation code (*occ1950* between 0 and 970, dropping armed forces code 595). For years 1870, 1880, 1900, and 1920, I restrict to individuals in the labor force (*labforce* = 2). For 1910 and 1930, I restrict to employed individuals (*empstat* = 1). For 1940 onward, I restrict to employed individuals (*empstat* = 1) and drop members of the armed forces (drop if *empstatd* = 13, 14, or 15). To ensure the 1950-2015 samples are representative, I apply person weights (*perwt*). However, these weights do not account for the intensive margin of total labor supplied (Autor and Dorn 2013). I therefore report additional results weighted by the product of person weights and total labor supplied during the year (weeks worked multiplied by usual hours per week).

C.2.3 Measuring occupation and economic status

The census began collecting income data in 1940. For the pre-1940 period, prior literature relies on occupation to examine trends in economic status and intergenerational mobility. Occupation can be mapped to economic status in a number of ways, some of which I discuss here. A common approach is to measure economic status using the variable *occscore*, constructed by IPUMS as the median income level of an occupation in 1950 (Ruggles et al. 2022; Olivetti and Paserman 2015). This variable will not capture changes in occupational earnings over time before 1950 (for instance, if incomes in healthcare grew relatively quickly from 1900 to 1950, then *occscore* will overstate economic standing in 1900). This is a fundamental issue that prior literature has tried to address. In particular, Song et al. (2020) use an occupation ranking procedure that weights by relative education level within each birth cohort and assigns occupations to a percentile rank (0-100). This measure is meant to capture the relative economic status of individuals in a given occupation over time. Ward (2023) constructs an "adjusted Song score" that allows occupational rank to vary by race and region. Buckles, Price, et al. (2023) partition the Song score by race, region, and gender.

For my analysis of economic status, I start with the standard *occscore* assigned to each occupation. I use a modified version of the *occscore*, the lasso-adjusted industry, demographic,

and occupation (LIDO) scores (Saavedra and Twinam 2020). This version allows occupation score to vary more flexibly, but does not solve the fundamental challenge of mapping economic status back in time. I follow Buckles, Price, et al. (2023) and estimate adjusted Song occupation scores, stratifying by cohort, region, gender, and race. While adjusted Song scores are relative and not absolute measures, they capture shifts in each occupation's human capital content prior to 1950. To avoid the challenge of shifting occupational categories, these scores are defined over "microclass" occupational groupings (for instance, nurses fall under "health semi-professionals", see Song et al. (2020)). In practice, I show results separately for each of these different measures.

Figure C.6 plots the distribution of adjusted Song scores for women in each of the four census regions. The scores are shown for white women born from 1895 to 1904. While some variation exists across regions, nurses score relatively lower than clerks and teachers, and relatively higher than housekeepers. Absolute measures of economic status, such as *occscore*, show a similar pattern. The *occscore* for nurses (in hundreds of 1950 dollars) is 21, compared to 27 for teachers and 25 for clerical workers. Similarly, the LIDO scores for white women, aggregated over industry, state, and age, are 15 for nurses, 16 for teachers, and 18 for clerical workers (Figure C.5).

C.3 Additional Data Sources

C.3.1 Licensing and certification laws

State laws regulating registered nurses and licensed practical nurses are sourced from Carollo (2024). The data include all state statutes passed since the late 1800s concerning the registration, certification, and licensing of nurses. State laws regulating midwives are sourced from Anderson et al. (2020). States are coded as having a law in the first year that the law went into effect, which sometimes differs from the year of enactment. See Carollo (2024) and Anderson et al. (2020) for more details on the historical sources of these laws.

Figure C.1: Excerpt from the 1905 Report of the Commissioner of Education

TABLE 3.—*Statistics of training schools for nurses, for the year 1904–5.*

	Location.	Name of institution.	Capacity of hospital (beds).	Year nurse school first opened.	Superintendent of nurses.	Session closes (about).	Nurse pupils. Women.	Men.	Graduated in 1905.	Years in the course.	Are pupils sent to nurse out-patients.	Pupils receive free board, (l) laundry, (u) uniforms, (t) text-books.	Monthly allowance to pupils. First year.	Second year.	Third year.	Estimated value of grounds and buildings of hospital.
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	ALABAMA.															
1	Birmingham	Hillman Hospital	80	1903	Jane W. Barry	Feb. 28	11	....	7	3	No.	b, l	$7	$7	$9	$100,000
2	....do	St. Vincent's Hospital	150	1900	Sister Chrysostom	June 12	35	....	12	3	Yes	b, l	5	5	5	200,000
3	Mobile	Providence Infirmary	50	1902	Sister Agatha	June 1	19	....	6	3	Yes	b, l	5	5	5	150,000
4	Montgomery	St. Margaret's Hospital	75	1903	Sister Margaret	Jan. —	14	....	0	3	Yes	b, l				150,000
5	Normal	M. French Sheldon Hospital of Agricultural and Mechanical College.	.....	1895	Mrs. F. M. Innis	May 25	36	0	7	2	Yes					900
6	Tuskegee Institute	Tuskegee Institute Hospital	25	1892	John A. Kenney	....do	18	6	5	3	Yes	b, l	0	0	0	2,000
	ARKANSAS.															
7	Hot Springs	St. Joseph's Infirmary	100	1905	—— Revelle		5	....	....	2		b, l	5	5		
8	Little Rock	Pulaski County Hospital	200	1902	Kate M. Illing	Apr. 20	14	0	7	3	Yes	b, l	8	10	10	75,000
	CALIFORNIA.															
9	Eureka	Sequoia Hospital and Sanitarium.	30	1903	D. L. Anderson	Dec. 29	9	0	4	3	Yes	b	7	7	7	30,000
10	Los Angeles	California Hospital	200	1889	Helen W. Kelly	June 5	70	....	26	3	No.	b	6	9	12	300,000
11	....do	County Hospital	350	1895	N. Estelle Woods	June 1	25	4	11	2	No.	b, l	5	10		
12	....do	Deaconess Hospital*	40	.....	Anna Du Sold		30	1	....	3			8	10	12	90,000
13	....do	Emergency and General Hospital.	150	1903	Nora E. Brown	(a)	35	3	20	3	Yes	b, l	6	6	8	70,000
14	....do	Good Samaritan Hospital*	100	1897	Harriet W. Pahl	June —	37	0	10	3			5	8	10	150,000
15	....do	Los Angeles Hospital	12	1900	Lydia Wismar		4	....	4	2	No.	b	10	12		
16	....do	Los Angeles Infirmary	200	1900	Sister Regis	June 15	60	2	19	3	Yes	b, l	5	5	5	
17	Oakland	Fabiola Hospital	100	1887	Katharine Fitch	June —	59	4	12	3	Yes	b, l	9	9	9	110,000
18	....do	Providence Hospital	100	1904	Sister Mary Alphonse	Apr. 5	30	0	0	3	Yes	b, l	5	5	5	

* In 1903–4.

a No definite session.

Source: U.S. Bureau of Education (1907).

Figure C.2: Excerpt from the 1922 Report of the American Nurses Association

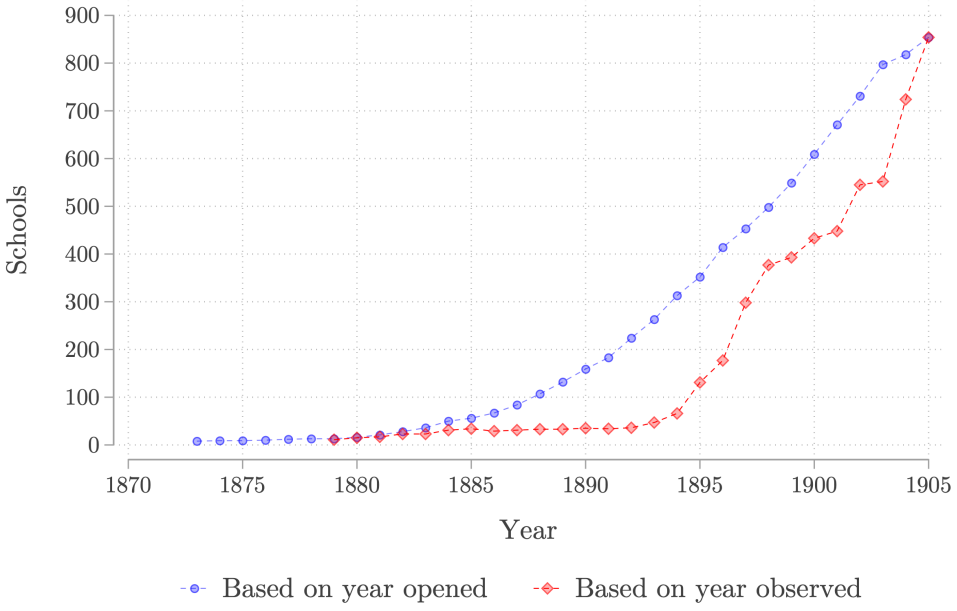
MASSACHUSETTS

Name of Hospital Infirmary or Sanitarium with which School is Connected	City	Year School Established	Denomination	Bed Capacity	Daily Average No. of Patients, 1921	Nursing Staff: Number of Graduates	Number of Full Time Instructors	Number of Students	Entrance Age	Entrance Education High School	Division of Service: Medical	Surgical	Obstetrical	Children's	Contagious	Mental	Length of Course Months	Weekly Number of Hours on Duty	Total Number of Graduates	Superintendent of Nurses
Arlington Training School, Arlington Health Resort & Ring Sanatorium	Arlington Heights	1907	N.S.	60	46	2	0	29	18	2 yr.	"	"	"	"	....	"	36	56	59	Grace L. Reilly, R. N.
Sturdy Memorial Hospital	Attleboro	1913	N.S.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	22	....
Beverly Hospital	Beverly	1893	N.S.	85	52	8	1	27	18	2 yr.	"	"	"	"	0	0	36	56	92	Frances P. West, R. N.
Boothby Surgical Hospital	Boston	1898	N.S.	55	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
*Boston City Hospital	Boston	1878	N.S.	....	850	40	3	85	19	2 yr.	"	"	"	"	"	"	36	58	1422	Ellen C. Daly, R. N.
*Boston Lying-In Hospital	Boston	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
Boston State Hospital	Boston	1899	N.S.	1944	1890	7	2	....	18	1 yr.	"	"	"	"	"	"	36	60	214	Mary A. McMahon, R. N.
Carney Hospital	South Boston	1892	C.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	438	....
*Children's Hospital	Boston	1889	N.S.	156	129	10	1	93	18	H.S.	"	"	"	"	0	0	36	52	314	Mary L. Wakefield, R. N.
*Infants Hospital	Boston	1881	N.S.	50	37	4	....	....	....	....	....	....	....	....	....	....	....	62	....	M. Edith Philpott, R. N.
Long Island Hospital	South Boston	1900	N.S.	317	260	18	1	7	18	H.S.	"	"	"	"	"	"	36	54	345	Mary A. Morris, R. N.
*Mass. Charitable Eye & Ear Hosp.	Boston	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
*Massachusetts General Hospital	South Boston	1874	N.S.	355	315	25	3	192	19	H.S.	"	"	"	"	0	0	36	52	1454	Sally Johnson, R. N.
*Massachusetts Homeopathic Hospital	Boston	1885	N.S.	541	418	30	2	107	19	H.S.	"	"	"	"	"	"	36	56	777	Ethel A. Humphrey, R. N.
Massachusetts Women's Hospital	Boston	1896	N.S.	44	35	3	....	24	20	H.S.	"	"	"	"	0	....	....	....	121	Zillah MacLauchlin, R. N.
New England Baptist Hospital	Boston	1896	P.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	172	....
New England Deaconess Hospital	Boston	1896	P.	74	63	7	1	57	19	H.S.	"	"	"	"	0	0	36	52	154	Adelina A. Betts, R. N.
*Peter Bent Brigham Hospital	Boston	1912	N.S.	201	188	24	2	93	19	H.S.	"	"	"	"	....	....	36	54	184	Carrie M. Hall, R. N.
St. Elisabeth's Hospital	Brighton	1895	C.	190	163	14	1	80	18	2 yr.	"	"	"	"	....	....	36	56	230	Sr. M. Florence, R. N.
Brockton Hospital	Brockton	1896	N.S.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	124	....
Goddard Hospital	Brockton	1914	N.S.	38	32	5	....	21	18	2 yr.	"	"	"	"	"	....	....	48	18	Blanche Hyslop, R. N.
*Women's Free Hospital	Brookline	1878	....	63	30	1	....	....	21	2 yr.	....	....	....	....	....	....	....	58	....	Hannah J. Ewing, R. N.
*Cambridge Hospital	Cambridge	1904	N.S.	113	92	5	1	52	18	1 yr.	"	"	"	"	0	0	36	57	170	....
Charlesgate Hospital	Cambridge	1907	N.S.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
Rufus S. Frost Hospital	Chelsea	1892	N.S.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	96	....
Clinton Hospital	Clinton	1893	P.	60	42	4	....	22	18	1 yr.	"	"	"	"	0	....	36	55	143	Mrs. E. M. Simpson, R. N.
Danvers State Hospital	Danvers	1889	N.S.	1552	1502	8	0	9	18	1 yr.	"	"	"	"	....	"	36	60	241	....
*Whidden Memorial Hospital	Everett	1897	N.S.	34	20	2	2	16	21	H.S.	"	"	"	"	....	....	....	58	71	Hannah F. Leavey, R. N.
Truesdale Hospital	Fall River	1912	N.S.	44	34	5	0	38	18	H.S.	"	"	"	"	0	0	....	52	48	Helen J. Leader, R. N.
Union Hospital	Fall River	1890	N.S.	140	....	9	1	42	18	1 yr.	"	"	"	"	....	....	36	48	321	Jessie M. Cann, R. N.
Burbank Hospital	Fitchburg	1894	N.S.	125	97	9	1	32	18	1 yr.	"	"	"	"	....	....	36	53	157	Mary J. MacKay, R. N.
Franklin Co. Public Hospital	Greenfield	1895	N.S.	61	42	4	0	15	18	1 yr.	"	"	"	"	....	....	....	52	118	Annie S. Barclay, R. N.
Stephen Henry Gale Hospital	Haverhill	1903	N.S.	....	83	9	....	40	18	2 yr.	"	"	"	"	....	....	....	52	65	Margaret I. Nicholson, R. N.
The Hale Hospital	Haverhill	1898	N.S.	59	37	7	1	12	18	2 yr.	"	"	"	"	....	....	....	56	120	Emma A. Mortimer, R. N.
Holyoke City Hospital	Holyoke	1893	N.S.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	195	....
Adams Nervine Hospital	Jamaica Plains	1897	N.S.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	129	....
*Faulkner Hospital	Jamaica Plains	1903	N.S.	70	59	6	0	39	18	1 yr.	"	"	"	"	....	....	36	56	90	Frances P. Ladd, R. N.
Lawrence General Hospital	Lawrence	1882	N.S.	111	77	9	1	51	18	1 yr.	"	"	"	"	"	....	36	54	277	Jessie E. Catton, R. N.
Leominster Hospital	Leominster	1912	N.S.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	15	....
Lowell General Hospital	Lowell	1893	N.S.	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	178	....
Lowell Hospital	Lowell	1904	N.S.	91	63	7	2	28	18	2 yr.	"	"	"	"	0	0	....	54	170	M. H. Kelly, R. N.
St. John's Hospital	Lowell	1867	C.	144	60	7	2	30	18	1 yr.	"	"	"	"	....	....	36	54	7	Helen F. O'Rourke, R. N.

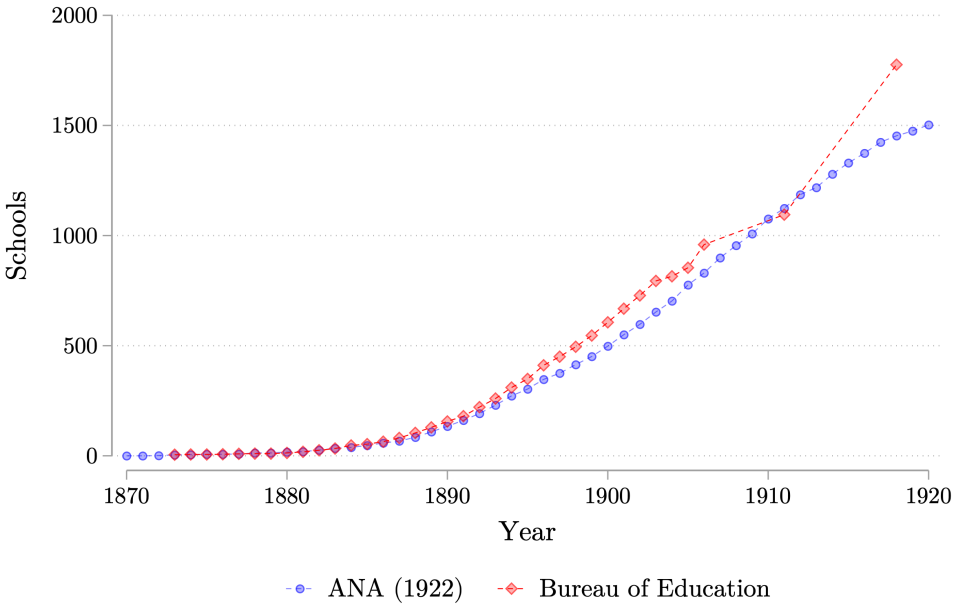
Source: American Nurses Association (1922)

Figure C.3: Count of training schools over time from different sources

(a) Bureau reports



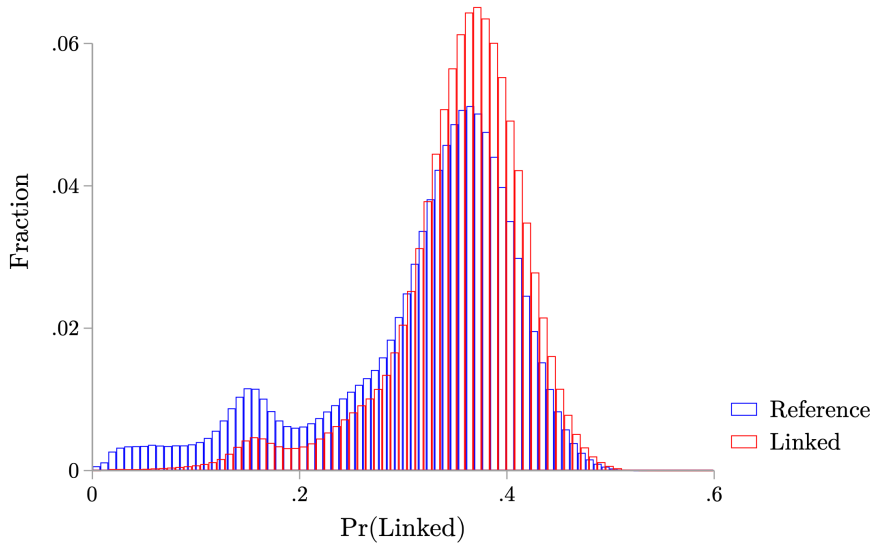
(b) Comparison of ANA and Bureau reports



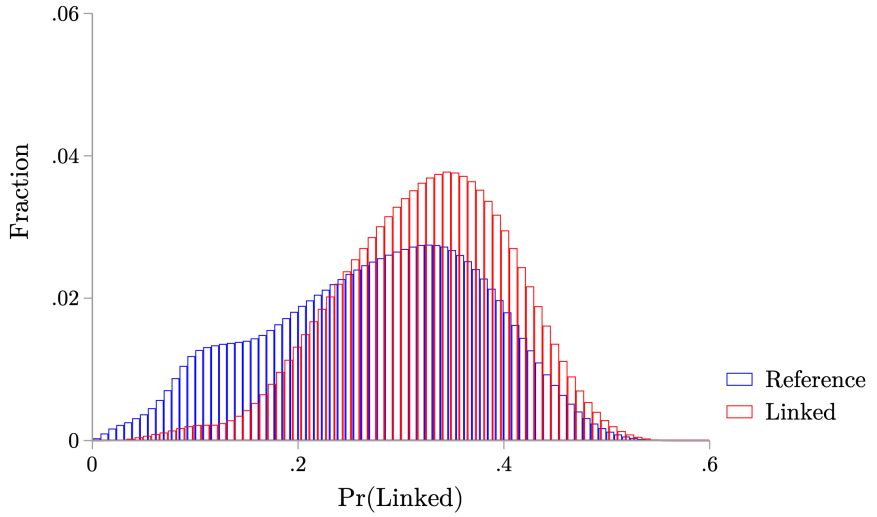
Notes: Figure reports the count of nurse training schools over time from several sources described in Appendix Section C.1. Panel (a) reports counts constructed from reports of the U.S. Bureau of Education, 1879-1905. Counts are shown based on year opened and year observed. Panel (b) compares counts in American Nurses Association (1922) and the Bureau reports from 1870 to 1920.

Figure C.4: Propensity scores, 1900 reference population

(a) 1900-1920 linkage

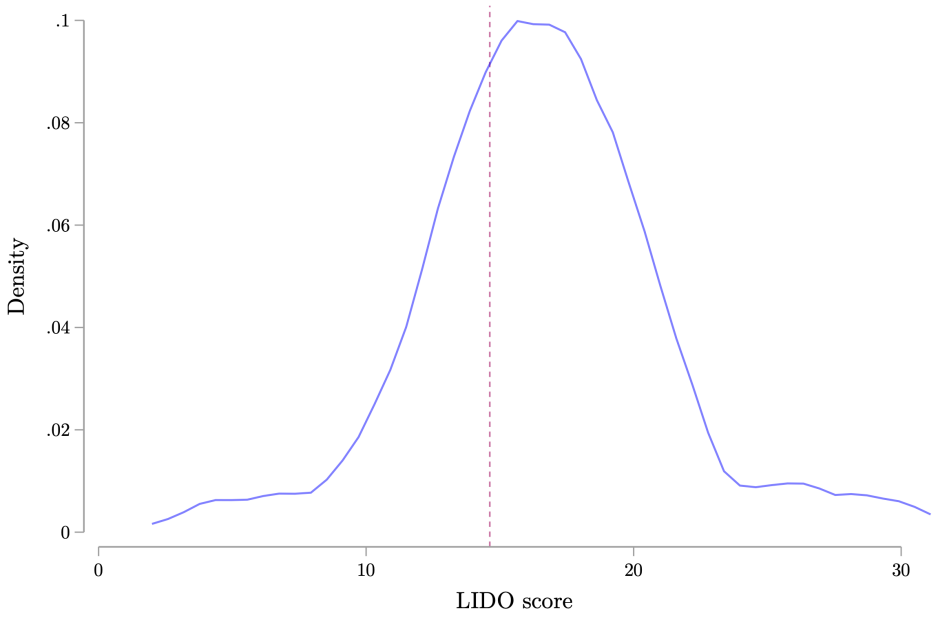


(b) 1900-1930 linkage



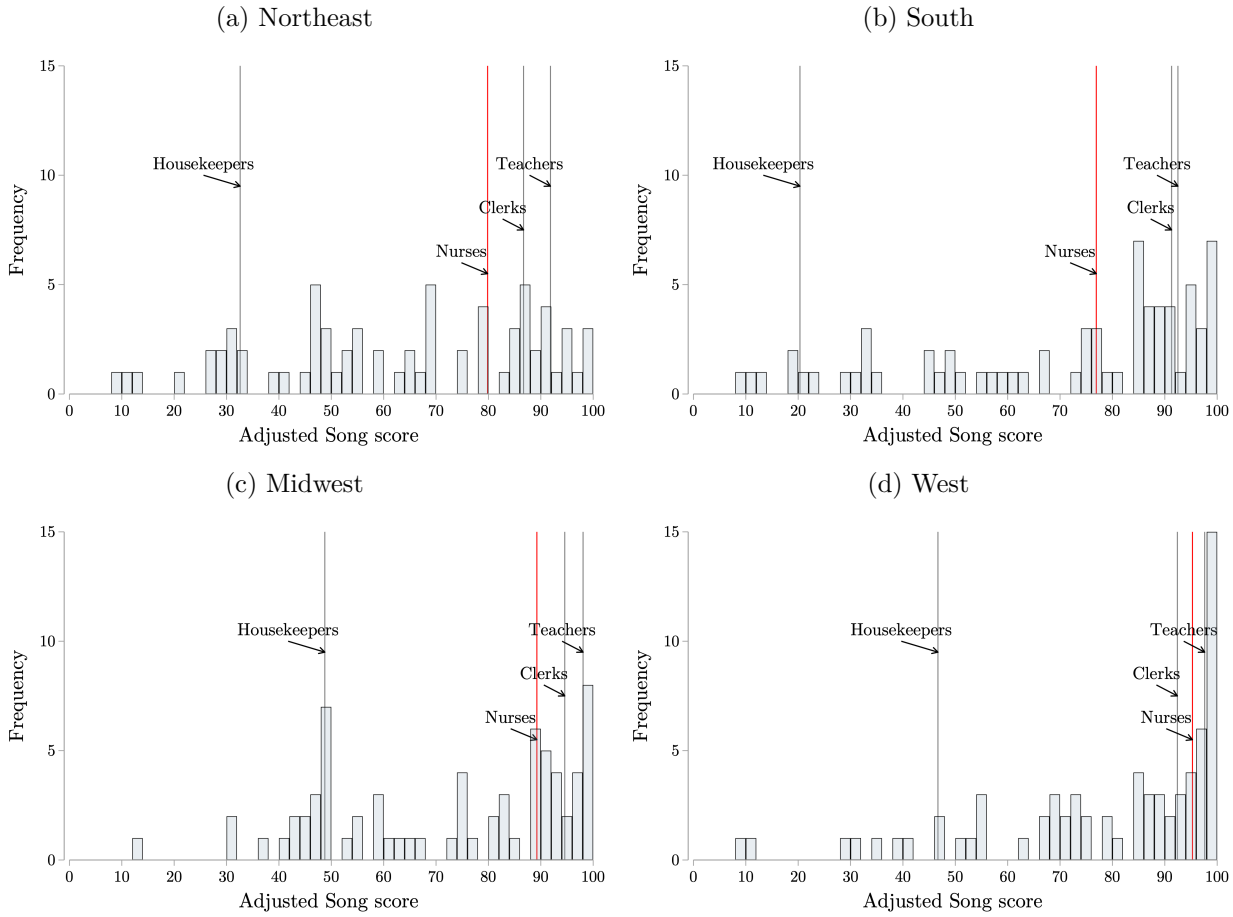
Notes: Figure reports propensity score distributions constructed according to Appendix Section C.2.1. In Panel (a), propensity scores are shown for the 1900-1920 census linkage. Plotted in blue are propensity scores for the complete-count 1900 census (the reference population). Plotted in red are propensity scores for the sample linked to 1920. Panel (b) reports analogous propensity scores for the 1900-1930 census linkage.

Figure C.5: LIDO scores, white women at age 25



Notes: Figure reports the distribution of LIDO income scores for white women at age 25. LIDO score is constructed from the 1950 census (Saavedra and Twinam 2020). The red vertical line indicates the LIDO score for professional nurses.

Figure C.6: Adjusted Song scores for white women born 1895-1904



Notes: Figure reports histograms of adjusted Song scores for white women born from 1895 to 1904. Each panel (a)-(d) is a census region. Observations are micro-class occupation groups following Song et al. (2020). Vertical lines denote the locations of the micro-class groups containing nurses, clerks, teachers, and housekeepers. The line for nurses is shown in red. See Appendix Section C.2.3 for details.

Table C.1: Statistics of nurse training schools from the Bureau of Education

Year	(1) Schools	(2) Students	(3) Graduates	(4) Average length	(5) Average students	(6) Percent any men	(7) Percent asylum	(8) Percent postgrad
1879	11	298	141	1.7	27.1	.	.	.
1880	15	323	157	1.7	24.8	.	.	.
1881	17	414	133	1.8	25.9	.	.	.
1882	23	475	124	1.9	25.0	.	.	.
1884	31	579	221	1.8	23.2	.	.	.
1885	34	793	218	1.8	24.8	.	.	.
1886	29	837	349	1.8	28.9	.	.	.
1887	31	989	335	1.9	31.9	.	.	.
1888	33	1,093	421	2.0	34.2	.	.	.
1889	33	1,255	443	1.9	38.0	9.1	.	.
1890	35	1,552	471	1.9	44.3	17.1	.	.
1891	34	1,613	527	1.9	47.4	14.7	.	.
1892	36	1,862	582	1.9	51.7	16.7	.	.
1893	47	2,338	786	2.0	50.8	15.2	.	.
1894	66	2,710	970	2.0	41.7	15.4	.	.
1895	131	3,994	1,489	2.0	31.0	15.5	.	.
1896	177	5,093	1,773	2.1	28.8	15.8	.	.
1897	298	7,254	2,498	2.1	24.3	17.1	.	.
1898	388	9,072	3,223	2.2	23.4	19.1	7.0	2.8
1899	405	10,387	3,328	2.2	25.6	17.3	8.9	3.0
1900	433	11,164	3,539	2.3	25.8	17.6	9.7	.
1901	448	11,599	3,712	2.4	25.9	17.4	10.0	.
1902	545	13,231	3,998	2.4	24.3	17.1	9.7	.
1903	552	13,788	4,201	2.5	25.0	16.7	9.8	.
1904	724	17,364	5,232	2.5	24.0	17.5	7.7	.
1905	862	20,280	5,805	2.6	23.6	15.1	6.7	0.9
1906	974	22,831	6,574	2.6	23.5	14.8	7.0	1.5
1911	1,118	29,685	7,707	2.7	26.9	.	6.3	2.1
1918	1,776	55,374	13,791	2.9	32.8	6.1	4.7	.
1927	1,797	77,767	18,628	2.9	43.3	3.8	4.2	.

Sources: U.S. Bureau of Education, 1879-1882, 1884-1906, 1911, 1918, 1927. See Appendix Section C.1.
Notes: Table reports descriptive statistics for nurse training schools in each year that data is available from the U.S. Bureau of Education. Columns (1), (2), and (3) report the total schools, students, and graduates. Columns (4) and (5) report the average years of study and the average number of students. Columns (6), (7), and (8) report the percent of schools with any male students, the percent affiliated with an asylum, and the percent classified as postgraduate schools. Percents are missing in years where student breakdown by gender, asylum status, and postgraduate status are not reported.

Table C.2: Occupation codes for nurses in the census

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
								Census year								
Job Title	1870	1880	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2005	2010	2015
Professional nurse	46[1,2,4]	46[1,2,4]	46[1,2,4]	872[1,2]	872[1,2]	216[1,2]	19[1]	58[1]	150[1]	75[1]	95[1]	95[1]	313[1]	3130[1]	3255[1], 3256[1], 3258[1]	3255[1], 3256[1], 3258[1]
Student nurse							19[1]	59[2]	151[2]	923[2]						
Midwife	44[3]	44[3]	44[3]	934[3]	934[3]	257[3]	760[4]	772[3]	840[3]	924[3]						
Practical nurse				936[4]	936[4]	258[4]	760[4]	781[4]	842[4]	926[4]	207[4]	207[4]	350[4]	3500[4]	3500[4]	3500[4]
Health administrator										212[1]	15[1]	15[1]	35[1]	350[1]	350[1]	350[1]
Physician assistant											106[1]	106[1]	311[1]	3110[1]	3110[1]	3110[1]

Notes: Table reports census occupation codes for nurses from 1870 to 2015. Titles include professional nurse (e.g., trained nurse or registered nurse), student nurse, midwife, practical nurse (e.g., licensed practical or vocational nurse), health administrator (e.g., medical and health services manager), and physician assistant. Blank cells indicate that the job title did not receive a code in that census year's classification. Superscripts indicate the occupation codes that are included in the 1950 classification of (1) professional nurse (*occ1950* = 58), (2) student nurse (*occ1950* = 59), (3) midwife (*occ1950* = 772), and (4) practical nurse (*occ1950* = 781). The 1870-1900 occupation code for professional nurse likely captures all nurses except midwives. Professional and student nurses were combined under one category in 1940, as were midwives and practical nurses. Prior to 1940, practical nurses were classified under the title "untrained nurse". All codes are sourced from IPUMS (Ruggles et al. 2022).

Table C.3: Descriptive statistics, 1900-1920 linked sample

	(1) 1900 sample	(2) Linked 1920	(3) Linked 1920
Age	7.15	6.86	7.13
Number of siblings	3.07	3.17	3.16
Region = North	0.27	0.25	0.26
Region = Midwest	0.40	0.43	0.41
Region = South	0.28	0.28	0.28
Region = West	0.05	0.05	0.05
Length of first name	5.36	5.38	5.36
Length of last name	6.35	6.36	6.37
Middle initial	0.37	0.41	0.38
First name commonness	10.80	10.85	10.74
Last name commonness	7.81	7.99	7.78
Father present	0.90	0.94	0.92
Father age	40.30	40.08	40.25
Father literate	0.92	0.93	0.93
Father employed	0.91	0.92	0.92
Mother present	0.94	0.96	0.96
Mother age	35.56	35.39	35.60
Mother literate	0.91	0.93	0.93
Mother employed	0.05	0.04	0.04
Reweighted	No	No	Yes
Observations	11,420,519	6,331,497	6,331,497

Notes: Table reports descriptive statistics for the 1900-1920 linked sample. Column (1) reports means for the population of white women born in the U.S. between 1885 and 1900 and observed in the 1900 census. Column (2) reports means for the subset that can be linked to the 1920 census via FamilySearch. Column (3) reports means for the linked subset in column (2), reweighted using the inverse propensity score reweighting approach described in Appendix Section C.2.1.

Table C.4: Descriptive statistics, 1900-1930 linked sample

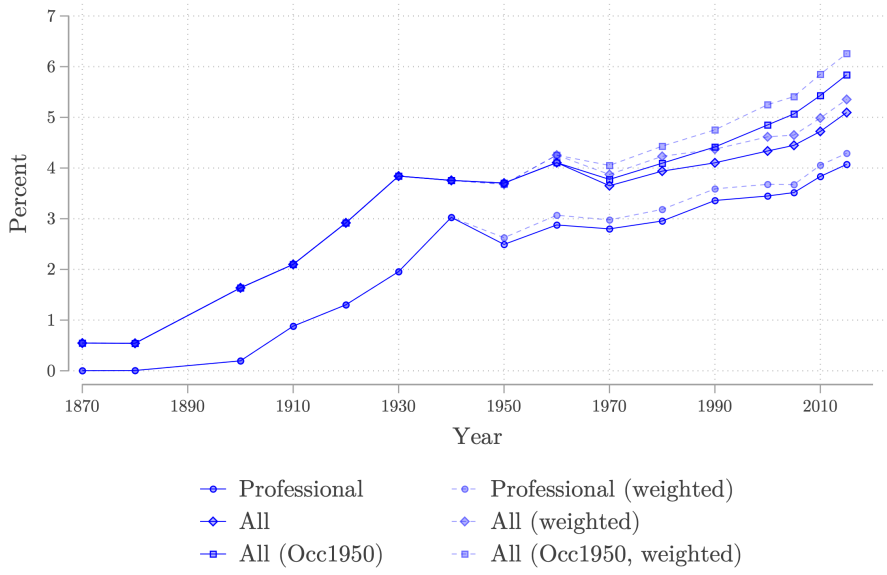
	(1)	(2)	(3)
	1900 sample	Linked 1930	Linked 1930
Age	7.15	7.07	7.52
Number of siblings	3.07	3.21	3.23
Region = North	0.27	0.23	0.25
Region = Midwest	0.40	0.44	0.41
Region = South	0.28	0.29	0.29
Region = West	0.05	0.05	0.05
Length of first name	5.36	5.36	5.35
Length of last name	6.35	6.35	6.36
Middle initial	0.37	0.42	0.38
First name commonness	10.80	10.82	10.68
Last name commonness	7.81	8.00	7.79
Father present	0.90	0.94	0.94
Father age	40.30	40.28	40.63
Father literate	0.92	0.93	0.93
Father employed	0.91	0.92	0.92
Mother present	0.94	0.96	0.97
Mother age	35.56	35.53	35.92
Mother literate	0.91	0.93	0.93
Mother employed	0.05	0.04	0.04
Reweighted	No	No	Yes
Observations	11,420,519	5,533,161	5,533,161

Notes: Table reports descriptive statistics for the 1900-1930 linked sample. Column (1) reports means for the population of white women born in the U.S. between 1885 and 1900 and observed in the 1900 census. Column (2) reports means for the subset that can be linked to the 1930 census via FamilySearch. Column (3) reports means for the linked subset in column (2), reweighted using the inverse propensity score reweighting approach described in Appendix Section C.2.1.

D Additional Figures and Tables

Figure D.1: Percent of the labor force in nursing by gender

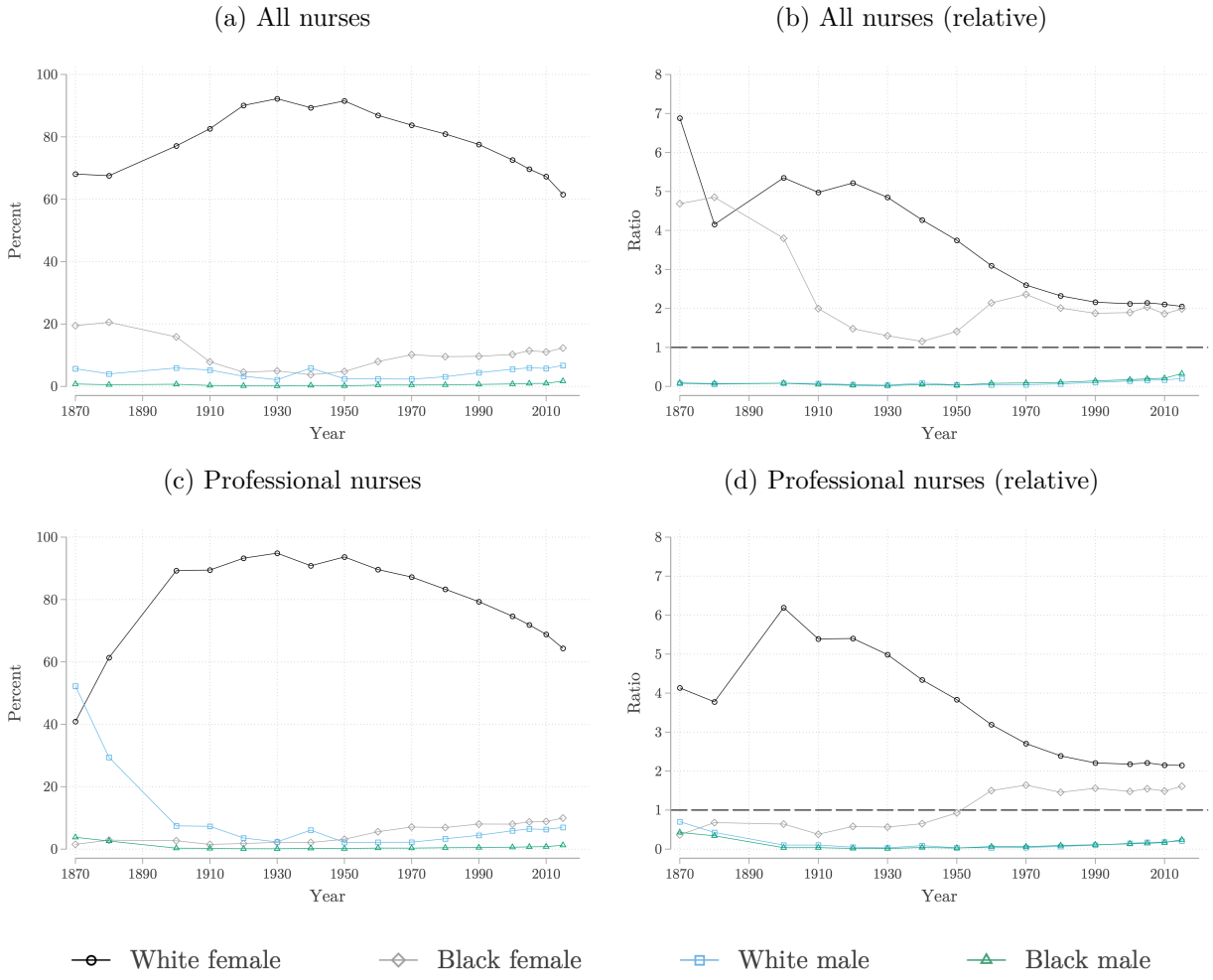
(a) Women



(b) Men

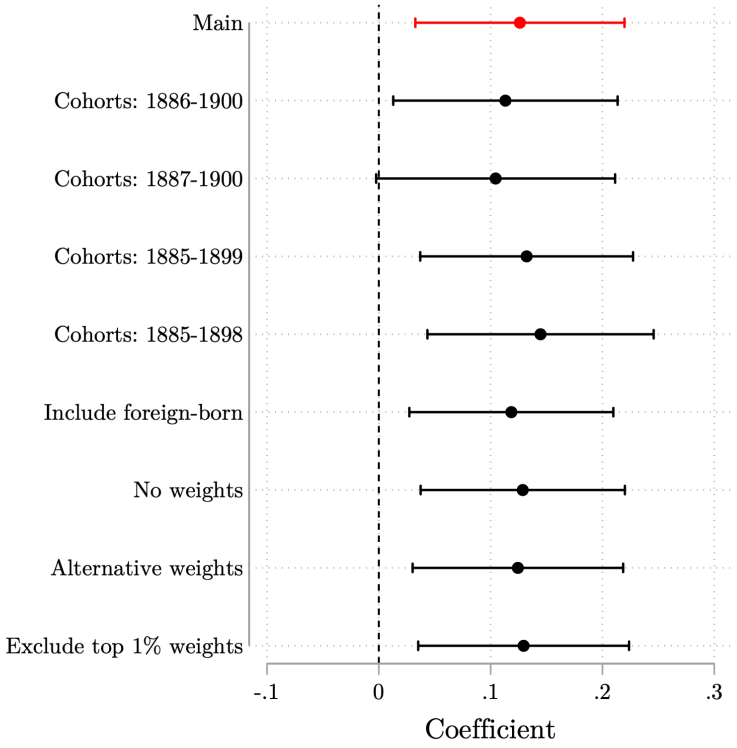
Notes: Figure reports the percent of the labor force in nursing by gender. The labor force is defined in Appendix Section C.2.2. Prior to 1950, the "Professional" series includes individuals with a 1950 occupation code of 58. For 1950 onward, the series includes professional nurses defined according to the codes in Table C.2. The "All" series includes professional, student, and untrained (practical) nurses defined according to the codes in Table C.2. The "All (Occ1950)" series counts all nurses according to the 1950 classification. Data come from the complete-count census for 1870-1940; 1% samples for 1950 and 1970; 5% samples for 1960, 1980, 1990, and 2000; American Community Survey for 2005, 2010, and 2015. The "Professional", "All", and "All (Occ1950)" series use person weights. The "weighted" series report results for years 1950-2015 where individuals are additionally weighted using annual labor supply.

Figure D.2: Demographic composition of nursing



Notes: Figure shows the demographic composition of nurses from 1870 to 2015. Panel (a) reports the percent of all nurses that are white female, black female, white male, and black male. Panel (b) transforms the percents in Panel (a) into ratios that represent the disproportionate representation of each group relative to the overall labor force (for instance, percent of all nurses who are black female divided by percent of the labor force that is black female). The labor force is defined in Appendix Section C.2.2. Panels (c) and (d) report the same statistics for professional nurses. Prior to 1950, "professional nurses" are individuals with a 1950 occupation code of 58. For 1950 onward, professional nurses are defined according to the codes in Table C.2. Data come from the complete-count census for 1870-1940; 1% samples for 1950 and 1970; 5% samples for 1960, 1980, 1990, and 2000; American Community Survey for 2005, 2010, and 2015. Person weights are used for all samples.

Figure D.3: Robustness checks for employment in nursing



Notes: Figure reports point estimates and 95 percent confidence intervals for equation (1), estimated using OLS. Robustness checks are described in Section 5.1.1. The dependent variable is an indicator for employment in nursing in 1920. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, described in Section 3.3.1. All specifications include individual controls, town fixed effects and birth year-by-county fixed effects. Standard errors clustered at the town level.

Table D.1: Count of nurses in the census by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Professional nurses			All nurses		
	All	Women	Men	All	Women	Men
1870	132	56	76	9,352	8,730	622
1880	347	227	120	20,486	19,498	988
1900	9,650	8,898	752	79,330	74,037	5,293
1910	68,204	62,943	5,261	159,142	149,997	9,145
1920	106,059	102,054	4,005	237,180	228,740	8,440
1930	189,366	184,677	4,689	371,177	362,517	8,660
1940	356,188	333,227	22,961	441,232	413,525	27,707
1950	396,225	387,064	9,161	590,689	575,050	15,639
1960	565,300	550,200	15,100	808,980	785,100	23,880
1970	806,700	785,400	21,300	1,054,900	1,024,100	30,800
1980	1,247,340	1,195,400	51,940	1,659,260	1,593,200	66,060
1990	1,828,769	1,727,121	101,648	2,237,115	2,109,349	127,766
2000	2,195,959	2,027,042	168,917	2,756,707	2,548,766	207,941
2005	2,356,045	2,150,184	205,861	2,968,567	2,719,877	248,690
2010	2,705,442	2,455,822	249,620	3,320,370	3,024,677	295,693
2015	3,113,789	2,767,984	345,805	3,905,085	3,462,736	442,349

Notes: Table reports the number of nurses aged 15-64 by gender for the years 1870-2015. Columns 1-3 report professional nurses only. Columns 4-6 report all nurses (professional, practical, and student nurses). Counts from 1870 to 1940 use the 1950 classification of occupations. Counts from 1950 onward use the codes in Table C.2. Data come from the complete-count census for 1870-1940; 1% samples for 1950 and 1970; 5% samples for 1960, 1980, 1990, and 2000; American Community Survey for 2005, 2010, and 2015. Person weights are used for all samples.

Table D.2: Descriptive statistics, 1930

	(1) Full sample	(2) Treated = 1	(3) Treated = 0
Childhood town, 1900			
Ln(population)	8.18	11.61	9.53
Region = North	0.26	0.40	0.49
Region = South	0.43	0.41	0.30
Region = Midwest	0.26	0.14	0.16
Region = West	0.06	0.05	0.05
Father characteristics, 1900			
Literate	0.93	0.96	0.95
Employed	0.92	0.87	0.90
Occupation score | Employed	21.21	27.78	24.21
Mother characteristics, 1900			
Number of children	4.23	3.97	4.14
Literate	0.93	0.95	0.94
Employed	0.04	0.04	0.04
Occupation score | Employed	15.48	17.96	16.82
Adult characteristics, 1930			
Married	0.80	0.71	0.74
Employed	0.21	0.27	0.25
Nurse | Employed	0.04	0.03	0.03
Occupation score | Employed	22.08	23.43	22.95
Unique places	33,210	1,903	4,577
Observations	5,232,365	777,766	548,654

Notes: Table reports descriptive statistics for the 1930 linked sample, described in Section 3.3.2. Statistics are weighted using inverse propensity scores to adjust for the likelihood of linkage. Column (1) reports means for the full sample. Column (2) reports means for treated women (e.g., women who were within five miles of an opening at ages 16-20). Column (3) reports means for the remainder of the final analysis sample (see Section 3.3). Father and mother characteristics are for women with an identifiable father or mother in the 1900 census. Nurse is an indicator for work in any nursing occupation. Occupation score and nurse variables are conditional on gainful employment.

Table D.3: Placebo tests with father and mother characteristics, 1930 sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Father characteristics			Mother characteristics			
	Literate	Occ score	Medicine	N children	Literate	Occ score	Medicine
Opening <5 mi. at ages 16-20	0.016	0.075	0.020	−0.013	−0.020	−0.044**	−0.004
	(0.118)	(0.061)	(0.037)	(0.011)	(0.123)	(0.018)	(0.006)
Controls	No	No	No	No	No	No	No
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	95.76	23.18	0.69	4.02	95.07	0.69	0.04
Observations	1,227,495	1,227,495	1,227,495	1,275,292	1,275,292	1,275,292	1,275,292

Notes: Table reports results for equation (1), estimated using OLS. Placebo outcomes are father and mother characteristics from the 1900 census: *Number of children* (as reported by the mother), *Literate* (father/mother can read and write), *Occ score* (father/mother occupation score), and *Medicine* (father/mother works in medicine). Coefficients on binary variables (*Literate* and *Medicine*) are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, described in Section 3.3.2. Sample is restricted to women with an identifiable father (columns (1)-(3)) or mother (columns (4)-(7)). All specifications include town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.4: Effects on employment and marriage, 1920

	(1)	(2)	(3)	(4)
	Employed		Married	
Panel A: Full sample				
Opening <5 mi. at ages 16-20	0.196	0.219	0.175	0.168
	(0.213)	(0.213)	(0.221)	(0.222)
Controls	No	Yes	No	Yes
Fixed effects	Yes	Yes	Yes	Yes
Mean outcome	39.78	39.78	50.86	50.86
Observations	1,646,657	1,646,657	1,646,657	1,646,657
Panel B: Married				
Opening <5 mi. at ages 16-20	0.127	0.129		
	(0.165)	(0.165)		
Controls	No	Yes		
Fixed effects	Yes	Yes		
Mean outcome	7.57	7.57		
Observations	837,445	837,445		
Panel C: Single				
Opening <5 mi. at ages 16-20	−0.202	−0.174		
	(0.340)	(0.338)		
Controls	No	Yes		
Fixed effects	Yes	Yes		
Mean outcome	73.11	73.11		
Observations	809,212	809,212		

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables in columns (1) and (3) are indicators for gainful employment and marriage in 1920, respectively. Columns (2) and (4) report results with individual controls added. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.1. Panel A reports results from the full sample. Panels B and C report results for women who are currently marriage or single, respectively. All specifications include town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table D.5: Effects of proximity on entrance into healthcare, 1920

	(1)	(2)	(3)	(4)
	Attendant	Physician	Health professional	Health semi-pro.
Panel A: Full sample				
Opening <5 mi. at ages 16-20	0.006	−0.004	0.004	0.005
	(0.013)	(0.008)	(0.011)	(0.007)
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Mean outcome	0.08	0.03	0.05	0.02
Observations	1,646,657	1,646,657	1,646,657	1,646,657
Panel B: Employed				
Opening <5 mi. at ages 16-20	0.016	−0.006	0.015	0.009
	(0.043)	(0.026)	(0.035)	(0.023)
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Mean outcome	0.21	0.06	0.14	0.06
Observations	655,027	655,027	655,027	655,027

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are indicators for attendant, physician, health professional (physician, dentist, optometrist, osteopath, or pharmacist), and health semi-professional (chiropractor, dietitian, medical technician, therapist, or midwife) in 1920. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.1. Panel A reports results for the full sample. Panel B reports results for women who are gainfully employed. All specifications include individual controls, town fixed effects, and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.6: Effect of proximity on entrance into healthcare by father's occupation, 1920

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Nurse subcategories		Other health occupations			
	Nurse	Trained	Practical	Attendant	Physician	Professional	Semi-pro.
Panel A: Father - professional							
Opening <5 mi. at ages 16-20	0.533**	0.513***	0.020	−0.017	0.005	0.096	−0.022
	(0.212)	(0.179)	(0.118)	(0.047)	(0.059)	(0.072)	(0.034)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	0.93	0.65	0.28	0.08	0.07	0.13	0.04
Observations	194,546	194,546	194,546	194,546	194,546	194,546	194,546
Panel B: Father - routine/manual							
Opening <5 mi. at ages 16-20	0.161**	0.102*	0.059	−0.012	−0.000	0.005	0.008
	(0.072)	(0.055)	(0.047)	(0.018)	(0.010)	(0.014)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	0.79	0.51	0.28	0.08	0.02	0.04	0.02
Observations	959,326	959,326	959,326	959,326	959,326	959,326	959,326
Panel C: Father - farmer							
Opening <5 mi. at ages 16-20	−0.008	−0.021	0.013	0.033	0.001	−0.007	−0.010
	(0.099)	(0.083)	(0.060)	(0.032)	(0.013)	(0.017)	(0.013)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	0.72	0.47	0.24	0.07	0.01	0.02	0.01
Observations	194,665	194,665	194,665	194,665	194,665	194,665	194,665

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are indicators for nurse, trained nurse, practical nurse, attendant, physician, health professional (physician, dentist, optometrist, osteopath, or pharmacist), and health semi-professional (chiropractor, dietitian, medical technician, therapist, or midwife) in 1920. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.1. Panels A, B, and C report results for women with fathers in professional, routine/manual, and farming occupations, respectively. All specifications include individual controls, town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.7: Effect of proximity on entrance into healthcare - single women, 1920

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Nurse subcategories		Other health occupations			
	Nurse	Trained	Practical	Attendant	Physician	Professional	Semi-pro.
Panel A: Single							
Opening <5 mi. at ages 16-20	0.338***	0.257***	0.081	0.032	−0.001	0.014	−0.004
	(0.115)	(0.089)	(0.069)	(0.030)	(0.018)	(0.024)	(0.016)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	1.59	1.06	0.53	0.15	0.04	0.09	0.04
Observations	809,212	809,212	809,212	809,212	809,212	809,212	809,212
Panel B: Single and employed							
Opening <5 mi. at ages 16-20	0.532***	0.412***	0.119	0.041	0.002	0.020	−0.004
	(0.171)	(0.132)	(0.104)	(0.046)	(0.028)	(0.037)	(0.024)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	2.17	1.45	0.72	0.21	0.06	0.12	0.06
Observations	591,625	591,625	591,625	591,625	591,625	591,625	591,625

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are indicators for nurse, trained nurse, practical nurse, attendant, physician, health professional (physician, dentist, optometrist, osteopath, or pharmacist), and health semi-professional (chiropractor, dietitian, medical technician, therapist, or midwife) in 1920. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.1. Panel A reports results for single women. Panel B reports results for single women who are gainfully employed. All specifications include individual controls, town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.8: Effect of proximity on entrance into healthcare - married women, 1920

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Nurse subcategories		Other health occupations			
	Nurse	Trained	Practical	Attendant	Physician	Professional	Semi-pro.
Panel A: Married							
Opening <5 mi. at ages 16-20	0.014	0.002	0.012	−0.015**	−0.004	0.001	0.004
	(0.021)	(0.013)	(0.017)	(0.007)	(0.006)	(0.008)	(0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	0.09	0.04	0.05	0.01	0.01	0.02	0.01
Observations	837,445	837,445	837,445	837,445	837,445	837,445	837,445
Panel B: Married and employed							
Opening <5 mi. at ages 16-20	0.392	0.041	0.351	−0.301**	−0.053	−0.081	0.014
	(0.348)	(0.218)	(0.272)	(0.119)	(0.077)	(0.106)	(0.079)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	1.19	0.58	0.61	0.19	0.15	0.28	0.09
Observations	63,402	63,402	63,402	63,402	63,402	63,402	63,402

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are indicators for nurse, trained nurse, practical nurse, attendant, physician, health professional (physician, dentist, optometrist, osteopath, or pharmacist), and health semi-professional (chiropractor, dietitian, medical technician, therapist, or midwife) in 1920. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.1. Panel A reports results for married women. Panel B reports results for married women who are gainfully employed. All specifications include individual controls, town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.9: Effects of proximity on entrance into common occupations, 1920

	(1) Teacher	(2) Clerk	(3) Bookkeeper	(4) Secretary	(5) Laborer
Panel A: Full sample					
Opening <5 mi. at ages 16-20	−0.048	−0.039	−0.026	0.149	−0.053
	(0.099)	(0.090)	(0.073)	(0.107)	(0.045)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes
Mean outcome	3.70	4.80	3.18	7.66	0.76
Observations	1,646,657	1,646,657	1,646,657	1,646,657	1,646,657
Panel B: Employed					
Opening <5 mi. at ages 16-20	−0.122	−0.199	−0.336	−0.022	−0.016
	(0.299)	(0.253)	(0.232)	(0.303)	(0.140)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes
Mean outcome	9.31	12.06	7.99	19.25	1.91
Observations	655,027	655,027	655,027	655,027	655,027

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are indicators for working in other common occupations in 1920: teacher, clerk, bookkeeper, secretary, and laborer (including farm laborer). Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.1. Panel A reports results for the full sample. Panel B reports results for women who are gainfully employed. All specifications include individual controls, town fixed effects, and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table D.10: Effects of proximity on occupational income by father's occupation, 1920

	(1) Ln(occ. score)	(2) Ln(LIDO score)	(3) Song score
Panel A: Father - professional			
Opening <5 mi. at ages 16-20	0.008	0.008	0.188
	(0.010)	(0.011)	(0.646)
Controls	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Mean outcome	3.18	2.81	82.28
Observations	73,613	64,214	73,522
Panel B: Father - routine/manual			
Opening <5 mi. at ages 16-20	0.006	0.004	0.639**
	(0.004)	(0.005)	(0.296)
Controls	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Mean outcome	3.14	2.78	73.76
Observations	419,935	363,885	419,560
Panel C: Father - farmer			
Opening <5 mi. at ages 16-20	0.001	0.013	0.664
	(0.012)	(0.013)	(0.717)
Controls	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Mean outcome	2.99	2.60	69.74
Observations	41,641	36,327	41,222

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are the natural log of occupation score, the natural log of LIDO score, and Song score (see Appendix Section C.2.3). The sample is native-born white women, as described in Section 3.3.1. Panels A-C report results for employed women whose fathers work in professional, routine/manual, and farming occupations, respectively. All specifications include individual controls, town fixed effects, and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.11: Effects on employment and marriage, 1930

	(1)	(2)	(3)	(4)
	Employed		Married	
Panel A: Full sample				
Opening <5 mi. at ages 16-20	−0.354*	−0.345	0.647***	0.642***
	(0.211)	(0.212)	(0.210)	(0.211)
Controls	No	Yes	No	Yes
Fixed effects	Yes	Yes	Yes	Yes
Mean outcome	26.09	26.09	72.30	72.30
Observations	1,326,420	1,326,420	1,326,420	1,326,420
Panel B: Married				
Opening <5 mi. at ages 16-20	0.031	0.033		
	(0.164)	(0.163)		
Controls	No	Yes		
Fixed effects	Yes	Yes		
Mean outcome	9.34	9.34		
Observations	959,031	959,031		
Panel C: Single				
Opening <5 mi. at ages 16-20	−0.019	−0.020		
	(0.511)	(0.511)		
Controls	No	Yes		
Fixed effects	Yes	Yes		
Mean outcome	69.82	69.82		
Observations	367,389	367,389		

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables in columns (1) and (3) are indicators for gainful employment and marriage in 1930, respectively. Columns (2) and (4) report results with individual controls added. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.2. Panel A reports results from the full sample. Panels B and C report results for women who are currently marriage or single, respectively. All specifications include town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.12: Effects of proximity on entrance into healthcare, 1930

	(1)	(2)	(3)	(4)
	Attendant	Physician	Health professional	Health semi-pro.
Panel A: Full sample				
Opening <5 mi. at ages 16-20	0.007	−0.022**	−0.009	0.010
	(0.012)	(0.011)	(0.014)	(0.012)
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Mean outcome	0.07	0.04	0.07	0.06
Observations	1,326,420	1,326,420	1,326,420	1,326,420
Panel B: Employed				
Opening <5 mi. at ages 16-20	0.018	−0.100**	−0.046	0.053
	(0.057)	(0.051)	(0.063)	(0.052)
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Mean outcome	0.29	0.17	0.27	0.21
Observations	346,121	346,121	346,121	346,121

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are indicators for attendant, physician, health professional (physician, dentist, optometrist, osteopath, or pharmacist), and health semi-professional (chiropractor, dietitian, medical technician, therapist, or midwife) in 1930. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.2. Panel A reports results for the full sample. Panel B reports results for women who are gainfully employed. All specifications include individual controls, town fixed effects, and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.13: Effect of proximity on entrance into healthcare - married women, 1930

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Nurse subcategories		Other health occupations			
	Nurse	Trained	Practical	Attendant	Physician	Professional	Semi-pro.
Panel A: Married							
Opening <5 mi. at ages 16-20	0.052*	0.008	0.043**	0.000	−0.020*	−0.006	0.004
	(0.029)	(0.022)	(0.020)	(0.008)	(0.011)	(0.013)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	0.19	0.11	0.09	0.03	0.02	0.04	0.02
Observations	959,031	959,031	959,031	959,031	959,031	959,031	959,031
Panel B: Married and employed							
Opening <5 mi. at ages 16-20	0.640*	0.136	0.504**	−0.023	−0.331**	−0.203	0.052
	(0.346)	(0.255)	(0.240)	(0.088)	(0.134)	(0.159)	(0.094)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome	2.09	1.18	0.91	0.28	0.24	0.42	0.19
Observations	89,616	89,616	89,616	89,616	89,616	89,616	89,616

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are indicators for nurse, trained nurse, practical nurse, attendant, physician, health professional (physician, dentist, optometrist, osteopath, or pharmacist), and health semi-professional (chiropractor, dietitian, medical technician, therapist, or midwife) in 1930. Coefficients are multiplied by 100 and can be interpreted in percentage points. The sample is native-born white women, as described in Section 3.3.2. Panel A reports results for married women. Panel B reports results for married women who are gainfully employed. All specifications include individual controls, town fixed effects and birth year-by-county fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table D.14: Effects of proximity on husband's occupation

	(1) Husband ln(occ. score)	(2) Husband in health occ.
Panel A: Married women		
Opening <5 mi. at ages 16-20	0.002	0.093
	(0.005)	(0.071)
Controls	Yes	Yes
Fixed effects	Yes	Yes
Mean outcome	3.18	1.69
Observations	920,933	920,933
Panel B: Married women, husband employed		
Opening <5 mi. at ages 16-20	−0.003	0.095
	(0.002)	(0.074)
Controls	Yes	Yes
Fixed effects	Yes	Yes
Mean outcome	3.37	1.80
Observations	867,878	867,878

Notes: Table reports results for equation (1), estimated using OLS. The dependent variables are husband log occupational income score (column (1)) and an indicator for whether husband works in a professional health occupation (column (2)). Coefficient in column (2) is multiplied by 100 and can be interpreted in percentage points. The sample is native-born married white women, as described in Section 3.3.2. All specifications include individual controls, town, state-by-cohort, and urban group-by-cohort fixed effects. Inverse propensity score weights are applied to adjust for the likelihood of linkage. Standard errors clustered at the town level in parentheses. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.