UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Labor Demand with Market Imperfections

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

Ryan Boone

2021

ABSTRACT OF THE DISSERTATION

Essays on Labor Demand with Market Imperfections

by

Ryan Boone
Doctor of Philosophy in Economics
University of California, Los Angeles, 2021
Professor Adriana Lleras-Muney, Chair

My first chapter examines whether tacit collusion occurs in the market for BigLaw associates. Many large firms across the U.S. offer the exact same associate salaries despite substantial heterogeneity on both sides of the labor market. I show that empirical market dynamics are difficult to reconcile with competitive labor markets. I then provide evidence for an alternative explanation – tacit collusion. A few firms act as price leaders and set maximum salaries. Some smaller firms are excluded from punishment to maintain cartel stability, and firms strategically communicate compensation decisions to align on decisions. Tacit collusion is facilitated by communication and standardization. Many of these practices originated in historical explicit collusion. This research highlights the potential for collusion in labor markets and the need for further scrutiny of other markets.

My second chapter is joint work with Anna Aizer, Adriana Lleras-Muney, and Jonathan Vogel, and we study the role of WWII in reducing occupational discrimination against Black men. The 1940s witnessed substantial reductions in the Black-white earnings gap. We show that domestic WWII defense production played an important role. In labor markets with more war production contracts, Black workers were more likely to be upgraded into skilled occupations and receive higher wages. War spending also led to an increase in the high

school graduation rate of Black children, suggesting important inter-generational spillovers. These results are attributable to the interaction between tight labor markets and federal prohibition against discrimination for war contractors. Using a structural model, we show that WWII defense production generated substantial improvements in national labor-market outcomes by decreasing discrimination for Black workers.

My third chapter looks at how firm acquisitions affect working conditions. I focus on acquisitions in a narrowly defined industry, nursing homes, to allow direct comparison across acquisitions and working conditions. I find that focusing only on the limited average effect on wages would miss more significant effects on benefits (6% decrease) and on workload (3-4% increase). Most importantly, working conditions in the acquired facility quickly converge towards those of the acquiring firm. This dynamic creates substantial heterogeneity in the effect of acquisitions on acquired facilities based on working conditions relative to the acquirer. Finally, I provide suggestive evidence that firm behavioral factors (e.g., managerial inertia) play an important role in the standardization of working conditions.

The dissertation of Ryan Boone is approved.

Till von Wachter

Jonathan Vogel

John Asker

Adriana Lleras-Muney, Committee Chair

University of California, Los Angeles

2021

To Andie. This work is as much yours as it is mine.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am extremely grateful for everyone who has helped me through this journey - family, professors, and of course my fellow students and colleagues. First and foremost, I would like to thank my wife, Andie Wheatley. While the last five years have held many milestones, including this dissertation, first and foremost in my heart is all of our moments together. Next, I would like to thank my family, especially my parents, Dan and Anne, for always providing unconditional support. My uncle, Paul Winters, ignited and supported my initial interest in economics. Finally, my grandparents have been my role models and inspiration throughout my life.

I am indebted to Adriana Lleras-Muney for her guidance and care over the years, especially during personally difficult times. I do not know if I could have completed this research without her help. I have also benefited from the advice and feedback from many other professors; including (but not limited to) Jon Vogel, John Asker, Katherine Meckel, Till von Wachter, Martin Hackmann, and Anna Aizer. I would also like to thank participants of the UCLA Applied Microeconomics Proseminar and the Applied Reading Group for invaluable feedback. Chapter 3, "Discrimination, the Racial Wage Gap and the Schooling of the Next Generation: Evidence from WWII" was on possible thanks to my great co-authors, Anna Aizer, Adriana Llera-Muney, and Jon Vogel. Finally, I would also like to thank my fellow students and friends, especially Tomas Guanziroli and Fernanda Rojas-Ampuero. Rustin Partow deserves special recognition for coming up with the idea to examine whether the market for BigLaw associates is collusive. The library staff at the UCLA Law Library were also indispensible resources. Ronda Fox provided crucial editing assistance and any remaining mistakes are my own. My time at UCLA was also supported by the Graduate Research Mentorship and Graduate Summer Research Mentorship fellowships.

VITA

2009-2013	B.A. Economics, Claremont McKenna College.
2013	Chairman's Award of Merit, Claremont McKenna College.
2013	Phi Beta Kappa and Omicron Delta Epsilon
2013-2016	Associate Consultant, Bain & Company.
2015-2016	Researcher, Living Standards Measurement Study, World Bank.
2017-2021	Research and Teaching Assistant, Economics Department, UCLA.
2018	Graduate Summer Research Mentorship fellowship, UCLA.
2018-2019	Graduate Research Mentorship fellowship, UCLA.
2017-2018	TA Award for Outstanding Performance (Fall '17, Winter '18), Economics Department, UCLA.
2020-2021	TA Award for Outstanding Performance in a Pandemic (Spring '20, Fall '20, Winter '21), Economics Department, UCLA.

PUBLICATIONS

Cash Transfer Programs and Agricultural Production: The Case of Malawi (joint with Katia Covarrubias, Benjamin Davis, and Paul Winters), Agricultural Economics, 44(3): 365-378, May 2013

CHAPTER 1

Tacit Collusion in Labor Markets: The Case of BigLaw

1.1 Introduction

In 2018, about 20% of new law school graduates received a starting salary of exactly $190,000 (Figure 1.1). This fact is puzzling. Price dispersion has been recognized as the norm since at least Stigler (1961), who stated that "dispersion is ubiquitous even for homogeneous goods." A large empirical literature has confirmed this observation (Baye, Mogan and Scholten, 2006). In the case of lawyers, there is significant heterogeneity on both sides of the market - lawyers are highly differentiated (e.g., quality, location, preferences, outside options) as are law firms (e.g., prestige, amenities, promotion rates, location).

This exact uniformity is even stranger because it is a recent phenomenon. Figure 1.2 shows the distribution of starting salaries in 1996 and 2000. In 1996 there is no clear mass point, but then it suddenly appears in 2000. Moreover, almost no lawyers receive salaries above the modal salary, so it is effectively the maximum salary as well. Jobs at the modal salary are typically in "BigLaw" firms. These firms are the largest law firms, which can have thousands of lawyers. The immediate question is what causes this phenomenon?[1] I propose a simple explanation – tacit collusion.

I begin by documenting key empirical facts about the market for BigLaw lawyers.

[1]Several commentators have noted this fact and informally proposed explanations. Kevin Drum, of *Mother Jones*, attributed it to "weird cultural collusion" (see motherjones.com/kevin-drum/2013/11/starting-salaries-attorneys-are-pretty-weird/), Peter Turchin, a professor at the University of Connecticut, attributed it to "extreme competition" (see peterturchin.com/cliodynamica/bimodal-lawyers-how-extreme-competition-breeds-extreme-inequality/), and Andrew Sorkin, of *The New York Times*, attributed it to a desire to secure "bragging rights" (see nytimes.com/2007/12/02/business/02deal.html).

Significantly more firms began paying the exact same salary in major cities starting in 2000. The frequency of salary adjustments has declined, but when they do occur they are larger. For example, between 1996-2000 the median (nominal) salary offer increased by over 70% but then did not change again until 2006. Salaries have grown much slower than profit per equity partner. Salary increases and associate employment are positively correlated, which suggests that the wage increases are not primarily due to labor supply shocks.

These key empirical facts seem inconsistent with competitive markets. For example, pre-tax salaries are equalized across major cities despite substantially different tax regimes, amenities, and local price indices. Major salary jumps are not accompanied by corresponding increases in productivity or decreases in markups.

I next present the case for tacit collusion. A few firms consistently lead compensation decisions, most prominently Cravath, Swaine & Moore (hereafter Cravath). Other firms wait for its decisions and will sometimes match its offers within an hour of the initial announcement. These price leaders set a maximum salary for the market. They have established a reputation of immediately matching or exceeding any competing offers that exceed this salary cap. This policy severely limits th short-term payoffs to deviating from the collusive equilibrium.

Price leaders' strategies extend beyond a simple matching strategy. Some small firms are excluded from the strategic set and allowed to exceed the maximum salary to ensure cartel stability (Bos and Harrington, 2010). Firms strategically communicate about compensation to avoid miscommunication; for example, some firms might publicly communicate that their higher salary will be offset by lower bonuses so they are not exceeding the market maximum. Price leaders will also slightly exceed competing offers, both to increase punishment and to prevent cheating firms from obtaining reputational benefits.

Tacit collusion is facilitated by communication and standardization. Communication occurs through trade publications, trade groups, and private informal settings. It allows firms to monitor competitor choices and align on compensation expectations. Associate job ladders and pay scales are standardized. The standardization of associate job ladders and pay scales allows “apples to apples” comparisons of salaries across firms which facilitates monitoring. The standardization of salaries within job levels (i.e., no “discounting”) also reduces the ability of firms to cheat. Some of these facilitating practices began due to explicit collusion.

Finally, I provide some suggestive evidence of efficiency costs. Large salary jumps are only possible if associates were previously paid significantly less than their marginal product. This distortion could lead to more productive firms being inefficiently small because it allows “fringe” firms to enter. There is anecdotal evidence that large salary increases occur at least in part to eliminate competition with less productive “fringe” firms. This distortion could also alter relative input usage. I measure how “constrained” firms are by calculating how much their ratio of associates to partners decreased in the years prior to large salary jumps. More “constrained” firms increase associate hiring by significantly more than other firms after the large salary jumps. This result suggests that these firms wanted to hire more associates at the market salary prior to the jump, but they were unable to.

The market for BigLaw associates has almost always involved collusion. Explicit collusion began soon after the origin of the modern BigLaw firm. Associate salaries were set at an annual luncheon of the major New York law firms from 1927-1968. In 1968, Cravath unilaterally broke with the “luncheon” cartel by raising salaries by 50%. At the same time, Cravath publicly invited other leading New York law firms to follow their pricing strategy. After 1968, regional legal markets were imperfectly competitive or tacitly collusive. By the late 1990s, law firms had transformed from regional to national. At that time, a localized tech boom shock caused California firms to raise compensation. They also increased salaries in their New York satellite offices. These New York increases exceeded the local collusive maximum and temporarily broke down collusion. Leading New York firms responded by matching or exceeding any salary increases to establish a regime of national price leadership. This effort led to significant salary increases in 2000, 2006, and 2007. These increases are often referred to by commentators as “salary wars”.

This research shows that labor markets are not immune to collusion. In some ways, labor markets might be more susceptible to collusion than many product markets. For example, firms can easily monitor cheating in collusive no-poach agreements. The Department of Justice and Federal Trade Commission have recently increased their scrutiny of collusion in labor markets and issued guidance to Human Resources professionals.[2] There have been significant recent cases, for example against technology and animation companies for entering into no-poach agreements (Adobe, Apple, eBay, Google, Intel, Intuit, Lucasfilm, and Pixar). The DOJ also recently brought its first criminal cases for labor market collusion.[3]

This paper contributes to several strands of literature. First, it contributes to the large literature on detecting collusion. Detection approaches vary significantly across empirical settings. One approach is to look for behavior inconsistent with competition or for structural breaks in behavior. For example, Kawai and Nakabayashi (2020) show that Japanese procurement auction bids are inconsistent with competition. Alternatively, competitive markets can be used as direct benchmarks. Porter and Zona (1999) find that the bids of colluding firms in school milk auctions decrease with distance, unlike competitive firms. Finally, studies can directly test for whether structural models of collusion or competition best fit the data, like Bresnahan (1987), showing that the auto price war of 1955 is best explained by a breakdown of collusion. This paper is most closely related to papers that test whether behavior is inconsistent with competition; the most similar paper is Knittel and Stango (2003), which shows that the excess clustering of credit card rates at statutory ceilings is due to tacit collusion.

This paper also contributes to the separate literature on the operation of collusive cartels (Asker, 2010). Most closely related from this literature are Byrne and de Roos (2019) and Genesove and Mullin (2001). Byrne and de Roos (2019) use detailed retail gasoline pricing data to show how firms used dynamic pricing strategies to establish a mutual understanding

[2]U.S. Dep't of Justice Antitrust Div. and Federal Trade Commission. 2016. "Antitrust Guidance for Human Resource Professionals."

[3]Nina Beck. 2021. "DOJ brings First Criminal Antitrust Charges for No-Poach Agreement Between Employers." *The National Law Review*, January 15. https://natlawreview.com/article/doj-brings-first-criminal-antitrust-charges-no-poach-agreement-between-employers.

to achieve a tacitly collusive equilibrium. Genesove and Mullin (2001) use extensive notes from cartel meetings to detail the operation of the Sugar Institute cartel. I combine elements of both approaches by using both time series data on compensation decisions and qualitative narrative evidence to more fully illustrate firm strategies.

This paper contributes to both of these literatures in several ways. Existing papers typically focus on output markets, especially settings involving bidding or commodities, such as retail gasoline markets. There are only a few papers looking at cartels in input commodity markets with a focus on agricultural inputs (e.g., Huang, 2020). There is limited research on collusion in service markets and to my knowledge there is no existing research looking at examples of collusion in labor markets. Further, I provide a new case study on the origin and maintenance of price leadership under tacit collusion.

In addition, this paper is related to a growing literature on monopsony power in labor markets (Card et al., 2018). Such research has demonstrated that labor markets are far from perfectly competitive with firms typically facing labor supply elasticities around 1.5-2 (Sokolova and Sorensen, 2020). There is also a separate literature on non-compete agreements (e.g., Ashenfelter and Krueger, 2018). Most existing papers have focused on unilateral monopsony power from concentration (Arnold, 2020) or search frictions (Manning, 2003) rather than monopsony power deriving from coordinated action. This paper provides evidence that concerted action can contribute to monopsony power, even when markets are unconcentrated according to conventional measures.

Section 2 provides some background on BigLaw firms and introduces the data. Section 3 begins by showing key empirical facts about the labor market for BigLaw associates. It then assesses whether these key empirical facts are consistent with collusion. Section 4 presents the case for collusion. Section 4.1 shows potential indicators of collusion. Section 4.2 identifies price leaders. Section 4.3 goes through the history of price leadership and how it originated. Section 4.4 outlines the strategy of colluding firms. Section 4.5 discusses practices that facilitate collusion. Section 4.6 ends with analysis suggesting potential efficiency costs. Section 5 concludes and presents topics for future research.

1.2 Background on BigLaw

> However influence and power are measured – whether in raw economic terms or in subtler, political ones – these firms remain the leaders of the bar.
>
> - Anthony Kronman, Professor of Law at Yale University, on BigLaw firms (1993)[4]

BigLaw is a colloquial term referring to major law firms. There is no specific definition, but it generally means firms that feature on prominent industry lists, such as *The National Law Journal's* list of largest firms ("NLJ200"), *Vault's* list of most prestigious firms ("Vault 100"), or *The American Lawyer's* list of highest revenue firms ("AMLAW100"). These firms are typically large, with at least 100 attorneys; the largest, Baker McKenzie, had 4,720 attorneys in 2019. While these firms are large, the BigLaw market is significantly less concentrated than many other white-collar service markets; for example, there is no law firm equivalent to the "Big Four" accounting firms.

Firms are primarily composed of associates, partners, and support staff. They typically serve large corporate clients on their most complex legal matters, such as mergers. This allows them to charge premium rates, with many partners charging significantly in excess of $1,000 per billable hour. Most firms charge by the billable hour, although a handful of firms use other billing practices, such as fixed fees.

The operation and structure of BigLaw firms are heavily indebted to Paul Cravath who developed the "Cravath system" at the firm Cravath in the early 20th century. Most major firms still follow a version of the Cravath system. The system emphasizes hiring associates straight from elite law schools, "lockstep" compensation, and "up or out" promotion to partner.

There is a strong emphasis on hiring the best candidates. For law firms, this means hiring the best students directly from elite law schools. Almost all new associate recruiting occurs

[4] Anthony Kronman. 1993. *The Lost Lawyer: Failng Ideals of the Legal Profession.* Cambridge, MA: Harvard University Press.

directly from law school, rather than from already practicing lawyers in other industries. Historically, major firms did not poach associates from each other, but this norm has disappeared.

Lockstep compensation means that associates are paid a uniform salary based on their years of BigLaw experience. Most major firms use lockstep salaries for associates, although a handful of firms use this system only for the first few associate years. Many firms also pay associates lockstep bonuses, but there is a relatively greater use of individualized bonuses, the most important criteria typically being billable hour thresholds.

Historically, many top firms also used lockstep compensation for partners; however, this practice is becoming increasingly rare as most firms have switched to individualized partner compensation. Firms set partner compensation on a yearly basis. A major determinant of compensation levels is origination credits, which partners receive when they bring in new business. In recent years, a divide has grown between two types of partners – equity and non-equity. Non-equity partners receive compensation that is less dependent on firm profits and are typically paid significantly less.

Firms also rely on "up or out" systems. They accept large associate classes with the expectation that most associates will not make partner. According to *The Wall Street Journal*, "Most associates know their chances of making partner at the big firms is less than 5%."[5] Some associates will transfer and become partners at smaller firms. Many others will end up working in other sectors, primarily in-house corporate law departments or governmental positions.

Associates are typically required to work long hours. According to the *National Association for Law Placement*, the average minimum billable hours requirement for large firms was 1,918 in 2015. Associates are typically expected to bill 2,000 to 2,100 hours, and associates who want to become partners can bill closer to 2,300 to 2,400 hours. Actual hours

[5] Cameron Stracher. 2006. "Cut My Salary, Please!" *The Wall Street Journal*, April 1. https://www.wsj.com/articles/SB114384471634713946.

worked are significantly higher since typically less than 80% of hours worked are billable.[6]

Law firms have changed significantly over time. Prior to 1900, even the largest firms were small and even in 1933 the largest firms only had around 70 lawyers. These firms were local, with significant offices usually only in one city. In the early 1960s, less than forty firms had more than 50 lawyers nationwide. Lawyers were forbidden by The Canons of Professional Ethics to advertise or solicit clients until a 1977 Supreme Court ruling struck down these prohibitions. Around this same time, law firms began to grow significantly faster (Galanter and Palay, 1991). The largest firms have continued to grow and have increasingly become national or even international organizations (Galanter and Henderson, 2010).

BigLaw is relatively unique because of the large amount of publicly available data on private firms. Much of this data comes from various trade organizations that rank law firms, such as *The National Law Journal*, *Vault*, and *The American Lawyer*. These groups collect data in order to create their rankings. Several market intelligence firms or consulting groups also collect data. Examples include Legal Compass by ALM Intelligence and Citi Private Banking's Law Firm Group. Finally, there are many trade publications that cover law firms, such as *Above the Law's* coverage of firms' compensation announcements. I use primarily three sources of data – *The National Law Journal* rankings for employment and associate salaries, *The American Lawyer* rankings for revenue and partner compensation, and *Above the Law* coverage for compensation announcement timing.

The National Law Journal (NLJ) focuses on ranking firms by the number of attorneys and extends back to 1978. The number of firms it covers has increased over time; in 1978 it covered the top 200 firms, whereas the most recent editions cover the top 500. I focus on the top 200 firms ("NLJ200") throughout the paper.[7]

---

[6]Lateral Link. 2012. "Law Firm Hours - The Real Story." *Above the Law*, July 24. https://www.abovethelaw.com/career-files/law-firm-hours-the-real-story/.

[7]This data is for fiscal years, which typically end in June. Generally the distinction between fiscal year and calender year does not create confusion, so I simply refer to both by year. However, when discussing compensation announcements I also use calendar years, which creates some slight differences. For example, in June 2016 firms announced raises, but they were for the fiscal year 2017 which began in July 2016.

NLJ data also include the minimum and maximum offered starting associate salary. Salaries are almost always uniform within a given office since most firms use lockstep compensation. Many firms also use a uniform starting salary across offices – for example, 83% of firms reporting salaries in 2019 did not report separate minimum and maximum salaries.[8] However, the share of firms that offer uniform salaries across offices does vary across years. Unless stated otherwise, I report the maximum salary because it is the salary likely to prevail in the firm's main offices.[9] For some firms and years (1989-2014), data on minimum and maximum billing rates is also available. The billing rate data is mainly used in supplementary analysis because the sample coverage varies.

The American Lawyer's (AMLAW) rankings focus on firm financial data, so they collect more information on revenue, cost, and partner compensation. The first full year of available data I have is for 1986. Like the NLJ rankings, the list of firms expanded over time from 100 to 200 firms in terms of revenue. For some pieces of analysis, I restrict the sample to the top 100 firms ("AMLAW100").

The final main source of data is the trade publication *Above the Law*. It was founded in 2006 and covers BigLaw firms from an associate's perspective. Most importantly, it collects salary and bonus memos from major firms, allowing me to reconstruct the timing of compensation decisions.

A few supplementary data sources are used. The firm BigLaw Investor provides financial advice to associates and has helpfully compiled the "Cravath scale" salaries and bonuses by associate tenure and year starting in the early 2000s. Data on the salary of newly qualified solicitors (the UK equivalent of first year associates) at major UK firms comes from the trade publication *The Lawyer*. Finally, I use quotes and background information from a variety of

---

[8]Firms often even apply these salaries to international offices using recent quarterly exchange rates, which can lead to large salary fluctuations from quarter to quarter.

[9]For most years only 5-10% of observations are missing salary information; however, both the first few and last few sample years have higher rates of missing observations (20-30%). Both high and low salary firms have similar missing rates; for example, in 2019 I cannot reject the null hypothesis that missing and non-missing firms were equally likely to have paid the top salary of $160,000 in 2006. However, to create growth rates I also use within firm growth rates to deal with any potential issues due to changing samples.

publications covering the legal industry, such as *The New York Times* and the *ABA Journal*.

1.3 Key empirical facts and consistency with competition

1.3.1 Empirical facts

It's very odd.

- Brackett Denniston, former GE general counsel, on associate salaries (2016)[10]

First I document key empirical facts about the market for BigLaw associates. Next, I test whether these empirical facts are consistent with competitive labor markets.

Empirical Fact #1: Exact nominal salary uniformity across major firms and cities, especially after 2000.

As shown in the introduction, many law school graduates earn the exact same starting salary. This uniformity is due to BigLaw firms. Figure 1.3 shows the share of NLJ 200 firms that match the modal salary. The rate is roughly constant at 15-20% until 2000, when it jumps to 40-60%. Figure 1.4 plots the 10th, 25th, 50th, 75th, and 90th percentiles of the starting salary distribution across firms. Before 2000, the 50th, 75th, and 90th percentiles are distinct while after 2000 they are generally the same value.

This fact is true across major cities. Table 1.2 shows the matching rate by the firm's home office city and year.[11] Before 2000, almost all non-New York firms paid less than the national modal rate, which was also the New York modal rate. Matching rates increase significantly after 2000 from nearly zero for many major cities, especially Washington D.C., Chicago,

[10] Casey Sullivan. 2016. "Is Following Cravath's Lead the Best Way to Set Salaries?" *Bloomberg Law*, June 7. https://www.news.bloomberglaw.com/business-and-practice/is-following-cravaths-lead-the-best-way-to-set-salaries.

[11] Home office is defined as the city with the most firm employees.

Boston, Los Angeles, and San Francisco.[12] This evidence is consistent with anecdotal industry observations that these firms pay the same salary across all major legal markets and typically only vary salaries for smaller markets such as Charleston or Salt Lake City.

Empirical Fact #2: Large salary jumps (70% from 1996-2000, 28% from 2005-2007, 19% from 2017-2019) followed by periods of stagnant nominal salaries.

Table 1.1 provides data on the mean, median, and modal salaries. It also shows the mean and median salary change from the previous year. There are several key takeaways. First, there are many years when the median and modal salaries do not change. This is also true for individual firms; there are long periods where the median salary change is zero dollars. Because these salaries are nominal, real salaries often decrease for several years in a row.

Second, these periods are followed by large salary jumps. Between 1991-1996, the median change was always zero dollars and the median salary increased only from $66,000 in 1990 to $73,000 in 1996. Then the median salary jumped from $73,000 to $125,000 (a 71% increase) between 1996 and 2000. Salaries stagnated again and the median salary did not increase from 2000 to 2005. Another jump occurred from $125,000 to $160,000 (28%) between 2005 to 2007. Then another period of stagnation followed, with the median change of zero dollars every year until 2017 to 2019 when median salaries jumped by 19%. Thus, there have been long periods of stagnant nominal (and declining real wages) followed by periods of rapid growth.

Empirical Fact #3: Real wages have not grown since 2000 despite large real growth in profits per partner and other metrics.

Figure 1.5 plots the growth in real associate salaries and profit per equity partner from 1986-2020 (values are indexed to 1986 levels). All previous salary figures have not been

[12] These rates are for the maximum salary, so these firms could be matching salaries only in New York, but results are similar if the analysis is restricted to only firms that reported the same minimum and maximum salary.

adjusted for inflation. Adjusting for inflation shows there has been little real growth in starting associate salaries.[13] Almost all real salary growth since 1980 occurred between 1997 and 2000. Salary increases since 2000 have mainly been to reset real compensation to 2000 levels. On the other hand, profit per equity partner has increased substantially throughout the entire period.

Associate billing rates have also increased much more rapidly than associate salaries (see Panel A of Appendix Figure A.7). Potentially, profit per equity partner is due to the increasing use of non-equity partners, but average partner compensation, including non-equity partners, has also increased substantially faster than associate salaries (see Panel B of Appendix Figure A.7). Therefore, stagnant associate compensation does not seem to be due to a general negative shock to BigLaw firms.

Empirical Fact #4: Salary increases and associate employment are positively correlated.

Figure 1.6 plots the change in the log salary from $t-2$ to t versus the change in the log number of associates from $t-1$ to t. The two variables are strongly positively correlated. Therefore, short-run salary changes seem to be driven more by shifts in labor demand along the supply curve. The number of law school graduates also does not substantially change around salary increases. Instead, changes in the number of graduates lag behind salaries by three to four years (law school programs take three years; see Appendix Figure A.5).

These are the key empirical facts, but there are a few other trends to note. While the share of firms matching the national modal salary increased sharply in 2000, in many cities a large share of firms previously matched the local modal salary. Table 1.2 shows the share of firms headquartered in each city that pay the same salary. In Boston almost two-thirds of large firms paid the same starting salary prior to 2000. A separate trend is the general evolution of law firms into national, rather than local, firms. Firms began expanding across

[13] Appendix Figure A.4 plots the real total compensation (including bonuses) for firms following the "going rate." Inflation-adjusted bonuses were higher in 2000, which means real compensation has decreased.

states in the late 1980s. On average, in 1978 each firm had 94% of associates in the same state, while this number had fallen to only 71% by 2000 (Table 1.3).

Additionally, I always focus on the starting salary for a new associate in the previous figures. A natural question is whether these trends apply to more experienced associates. Since at least 2000, the salaries for all associate levels have always changed at the same time (see Appendix Figure A.2). Moreover, salaries increase proportionately across associate experience levels. In 2000, eighth-year associates were paid 1.80 times first-year associates, while in 2019 the same ratio was 1.79. Bonuses also tend to increase at the same time across associate experience levels (see Appendix Figure A.3). Salary increases are not offset by decreases in bonuses.

1.3.2 Basic model

> The laws of supply and demand dictate that thousands of entry-level associates now command the princely sum of $160,000 per year.
>
> - Galanter and Henderson (2010)

I begin with the most basic model of firm decision making. A law firm in labor market m at time t produces legal services, Q_t, using associates, N_t, and partners, K_t. The production function is given by $Q_t = A_t f(N_t, K_t)$. Output markets may be imperfectly competitive, so revenue is given by $P_t(Q_t)Q_t$. There are time-varying shocks to both firm productivity, A_t, and the pricing function $P_t(\cdot)$. The firm treats associate salary, w^n_{mt}, and partner compensation, w^k_{mt}, as exogenous. The firm's problem at time t is then:

$$\max_{N_t, K_t} P_t(Q_t(N_t, K_t))Q_t(N_t, K_t) - w^n_{mt}N_t - w^k_{mt}K_t$$

The firm's first order condition for associates is given by:

$$w_{mt}^n = P_t \frac{\partial Q_t}{\partial N_t} \frac{1}{\mu_t}$$

where $\mu_t = \frac{\varepsilon_t^{QP}}{1+\varepsilon_t^{QP}}$ is the markup of price over marginal cost. The right hand side is the marginal revenue product of associates. Taking the log of both sides gives:

$$\ln w_{mt}^n = \underbrace{\ln P_t \frac{\partial Q_t}{\partial N_t}}_{\text{Productivity}} - \underbrace{\ln \mu_t}_{\text{Markup}}$$

The first term on the right-hand side is the marginal physical productivity times the output price (for simplicity, I will refer to this as "productivity"), andthe second term is the markup. Any differences between wages across markets or time should be related to differences in productivity or markups.

Now, there are many potential reasons why this equation might not perfectly hold for any given firm in the data. Let $\ln \psi_t$ be the wedge (i.e., residual gap) between marginal revenue productivity of associates and wages. Then we have

$$\ln w_{mt}^n = \underbrace{\ln P_t \frac{\partial Q_t}{\partial N_t}}_{\text{Productivity}} - \underbrace{\ln \mu_t}_{\text{Markup}} + \underbrace{\ln \psi_t}_{\text{Input Wedge}}$$

$\ln \psi_t$ might be non-zero for many different reasons. For example, in the static formulation there might be measurement error or shocks that are realized after making input decisions. Or there might also be dynamic considerations due to adjustment costs, deferred compensation, or firm-specific human capital. Appendix Section A.1 discusses some stylized alternative dynamic models and show that the primary determinants of wages are the above factors plus future expected input wedges. Therefore, we should generally expect wage differences to be accompanied by changes in productivity or markups.

1.3.3 Consistency with competition

> In our conversations with firm leaders, many express bafflement as to why so many firms adopted the increases when their productivity and profitability results couldn't support them.
>
> - Gretta Rusanow, Head of Advisory Services at Citi Private Bank (2018)[14]

The goal is to determine whether the evolution of relevant market fundamentals is approximately consistent with competitive input markets. Because uncertainty about the true model of labor markets is a first order concern, I focus on market fundamentals rather than directly modeling and estimating labor market dynamics. The key empirical facts are:

1. Empirical Fact #1: Exact nominal salary uniformity across major firms and cities, especially after 2000.

2. Empirical Fact #2: Large salary jumps (70% from 1996-2000, 28% from 2005-2007, 19% from 2017-2019) followed by periods of stagnant nominal salaries.

1. Exact nominal salary uniformity across major firms and cities, especially after 2000: In a competitive labor market, observed salary equalization across regions i and j can be due either to labor supply or demand factors.[15] Labor supply factors can equalize salaries if regions i and j are in a common market ($i, j \in m$) or if they share a common outside option. Labor demand factors can equalize salaries across regions if they share common production fundamentals (e.g., constant marginal productivity across regions or common output market).

There are several arguments against salary equalization due to labor supply factors. First, in a perfectly competitive market, utility offers (and not nominal salaries) should be

[14] Debra Weiss. 2017. "Some law firm leaders question associate pay hikes amid tepid year." *ABA Journal*, February 15.

[15] There can also be "knife's edge" cases where various differences across regions exactly cancel out, but given the lengthy time period involved this seems unlikely.

equalized (Rosen, 1986). There is substantial variation in local tax rates, amenities, and local price indices (especially housing) across cities. Table 1.4 provides the after-tax earnings (on a salary of $190,000) in 2019 as well as estimated amenities and local prices (as of 2000) relative to New York from Diamond (2016).

Tax rates can vary significantly across cities since some states do not have state income taxes. Columns (1) and (2) show after-tax earnings. For example, $190,000 is worth $129,000 after tax in New York, whereas it is worth $140,000 in Dallas or Houston. Columns (3) and (4) provide amenities and local prices values in log salary equivalent terms (relative to New York).[16] Amenities and prices vary significantly across cities. For example, San Francisco is substantially more expensive than Los Angeles while also offering lower amenities.

Column (5) gives the total compensation relative to New York, and column (6) gives the equivalent New York salary. The uniform nominal salary of $190,000 hides substantial variation in "real" salary across locations. Associates in Los Angeles earn the equivalent of $264,000 in New York, while associates in Philadelphia earn the equivalent of $141,000. The point of this exercise is not to definitively determine the real salary across locations; rather it is to highlight that reasonable estimates imply that the utility value of $190,000 nominal salary varies widely across cities.

Second, even suppose that associates care only about nominal pre-tax income when comparing job offers. Compensation varied across major cities prior to 2000, which suggests they were separate markets. Potentially, there was a structural change around 2000 that increased competition and merged markets. For example, maybe the internet significantly reduced geographic search frictions. One way of testing this theory is to look at markets for lawyers in other countries. Figure 1.7 plots the distribution of starting salaries for newly qualified solicitors in London. There is substantial dispersion of salaries even within London. Therefore, any structural change would have to be powerful enough to equalize salaries across

[16]Local prices are primarily based on housing prices, with a scaling factor to adjust for the impact of housing prices on locally purchased goods. Amenity values are specific to college-educated workers. See Diamond (2016) for more detail.

regions in the United States, but not significantly affect the United Kingdom.

Common production fundamentals are also an unlikely explanation. Figure 1.8 shows that billing rates vary widely across firms. Firms also differ significantly in relative input usage across regions. For example, New York firms have a significantly higher ratio of associates to partners when compared to other regions (see Appendix Figure A.10).

2. Large salary jumps (70% from 1996-2000, 28% from 2005-2007, 19% from 2017-2019) followed by periods of stagnant nominal salaries: Salaries increased by 70% from 1996-2000 from but then did not change significantly until 2006. They jumped again in both 2006 and 2007 for a total increase of 28% but then again stagnated until 2017. Then in 2017 and 2019 they rose again for a total increase of 19%. In a competitive market we should expect these salary jumps to be accompanied by large increases in associate productivity or declines in markups. Note that these salary increases are unlikely to have been driven by labor supply shocks since they were accompanied by large increases in associate employment (see Empirical Fact #4)[17]

The key outcomes are associate salaries, productivity, and markups. I focus on the years immediately around major wage changes in 1997-2000, 2006-07, and 2017-19. For each outcome Y_{it}, I run the regression:

$$\ln Y_{it} = \gamma_i + \gamma_t + \epsilon_{it}$$

where γ_i and γ_t are firm and year fixed effects. I then plot the γ_t estimates to show how the mean outcome Y_{it} changes over time. Standard errors are clustered at the firm level and the regressions are weighted by the number of employed attorneys. The next question is how to measure productivity and markups.

For productivity: As an approximation, assume that $Q_{it} = A_{it} N_{it}^{\alpha_i} K_{it}^{\beta_i}$. Because the

[17]There might be supply shocks that affect local markets; case studies will be discussed later.

production function is Cobb-Douglas we have $\frac{\partial Q_{it}}{\partial N_{it}} = \alpha_i \frac{Q_{it}}{N_{it}}$. This gives us:

$$\ln \frac{P_{it}Q_{it}}{N_{it}} = -\ln \alpha_i + \ln P_{it}\frac{\partial Q_{it}}{\partial N_{it}}$$
$$\ln \frac{P_{it}Q_{it}}{N_{it}} = \gamma_i + \gamma_t + \epsilon_{it}$$

where γ_t captures the average log of productivity. Intuitively, we can measure changes in average log of productivity by looking at changes in the revenue per associate.

For markups: Let VC_t be total variable costs. Because the production function is homogeneous of degree $\alpha_i + \beta_i$ we have $MC_{it} = \frac{1}{\alpha_i+\beta_i}\frac{VC_{it}}{Q_{it}}$ from the firm's cost minimization problem.[18] Since $\mu_{it} = \frac{P_{it}}{MC_{it}}$:

$$\ln \frac{P_{it}Q_{it}}{VC_{it}} = -\ln(\alpha_i + \beta_i) + \ln \mu_{it}$$
$$\ln \frac{P_{it}Q_{it}}{VC_{it}} = \gamma_i + \gamma_t + \epsilon_{it}$$

where γ_t captures the average markup. We can measure changes in markups by looking at changes in the ratio of total revenue to total variable costs. I view these estimates for productivity and markups as approximations within an order of magnitude. In Appendix Section A.1, I discuss alternative approaches to measuring productivity and markups.

The estimated $\hat{\gamma}_t$ are plotted in Figure 1.10. For all three major discrete salary jumps, there is little evidence of any accompanying discrete changes in productivity or markups. Additionally, the employment of associates (both in absolute terms and relative to partners) significantly increases (Appendix Figure A.12), which is inconsistent with large labor supply shocks. Therefore, there seems to be a sudden and large change in the input wedge, $\ln \psi_{it}$.

One major potential concern is that there might be a shock to the productivity of associates relative to partners (α_i). Law firms provide a straightforward way to test this concern since there are separate billing rates for associates and partners. While many inputs

[18] Note that this means we can rely on a much weaker assumption than Cobb-Douglas production if we want to identify changes in markups and not the level.

are used to produce a billable associate hour (e.g., human resources), associates are the most important input. If there is a large shock to the relative productivity of associates then it should be reflected in relative billing rates.[19] Appendix Figure A.12 shows that there is almost no change in relative billing rates around the salary jumps.

These results are also inconsistent with many standard models of imperfect competition in input markets. Bertrand competition could explain uniform salaries. But it is not clear why pure Bertrand competition would occur given that firms face significant short-run capacity constraints. Firms are small relative to the market, it is difficult to recruit new partners, and most existing attorneys work very long hours, so firms have little excess capacity. Even if Bertrand competition does occur, it is not clear how it would explain the sudden changes in $\ln \psi_{it}$ followed by periods of stagnation. A basic Cournot model suffers from a similar issue. Models of differentiated firms (Card et al., 2018) imply firm-specific labor supply curves, which would result in differing salaries across firms. These empirical facts are also inconsistent with the large literature on the importance of firm-specific wage effects (Abowd, Kramarz and Margolis, 1999).

Models of wage bargaining would also have trouble explaining exact salary uniformity across firms given substantial differences in firm productivity and profitability. They would also have difficulty explaining the large salary jumps without discrete changes in bargaining power. Finally, they have the additional issue that real salaries have not increased with profitability and prices (Empirical Fact #3).

A final issue is that there exist significant quality differences among associates. Uniform salaries are even more difficult to explain if associates vary in quality. The focus that firms place on recruiting top students from elite schools suggests that quality does matter. It seems reasonable that firms know the quality of experienced associates. Law firms are

[19]An alternative shock would be an increase in the number of hours worked relative to partners, but it seems unlikely that short-run changes in hours worked can explain the magnitude of the salary jumps. However, increased work intensity might be an important factor in any long-run salary growth. Real hourly wages for associates have likely declined by more than total salary since 2000 because total hours worked have generally increased for associates.

highly sophisticated when it comes to setting individualized compensation; they typically adjust individual partner compensation on a year-to-year basis.

With sufficient degrees of freedom, there likely exists a model that could reconcile both salary uniformity across regions and large wage jumps with perfect input market competition (or with standard models of imperfect competition). For example, a model could incorporate internal fairness constraints, regime changes in the output market, nominal wage rigidity, and more. However, in the next section I will suggest a simpler alternative – tacit collusion – and provide significant supporting evidence.

1.4 Investigating for tacit collusion

1.4.1 Collusion indicators

> Masters are always and everywhere in a sort of tacit, but constant and uniform combination, not to raise the wages of labour above their actual rate. To violate this combination is everywhere a most unpopular action, and a sort of reproach to a master among his neighbours and equals. We seldom, indeed, hear of this combination, because it is the usual, and one may say, the natural state of things which nobody ever hears of.

- Adam Smith

Collusion occurs when firms in a market coordinate to restrain competition. From an economic perspective, it is an equilibrium in a repeated game with profit above the static equilibrium outcome. Collusion can either be explicit, with mutual understanding through direct communication, or tacit, with mutual understanding through indirect means.[20]

[20]From a legal standpoint, tacit collusion can be further distinguished by how the mutual interpretation of indirect means is established. Indirect actions need to be mutually interpretable by participants in order to establish mutual understanding. In the case where participants are exogenously "endowed" with a mutual understanding, then their behavior is "conscious parallelism." If a participant works to establish a mutual interpretation (e.g., through dynamic pricing strategies (Byrne and de Roos, 2019) or company statements (Bourveau, She and Zaldokas, 2020)), then there is potentially concerted action.

Collusion can lead to significant price distortions (Connor and Bolotova, 2006), so the detection of potential collusion is an important economic issue.

Section 1.3 provided evidence that market behavior is not consistent with competition, but this does not necessarily imply collusion. There are no universal markers of collusion, but Harrington (2005) does provide some common indicators.[21] At least several of these markers are present in the market for BigLaw associates. Under certain conditions, price variance is lower (prices are more stable) under collusion. As Empirical Fact #2 shows, there are limited salary adjustments. Under certain conditions, there is stronger positive correlation between firm prices. We observe exact matching of any salary changes for many firms (Empirical Fact #1). Harrington (2005) also discusses that high price levels or margins are not good indicators for collusion, but that sharp changes in these outcomes might be markers. For example, sharp changes in prices could indicate pricing wars due to punishment. Empirical fact #2 shows that there are sharp changes in salary levels for associates. In fact, many industry observers even refer to these sharp increases as "salary wars" (e.g., Galanter and Henderson, 2010). These markers suggest that potential collusion deserves further investigation.

Figure 1.9 plots starting associate salaries versus the log of the revenue per lawyer for fiscal year 2016.[22] Initially, there is a strong positive relationship between revenue per lawyer and associate salaries. However, the relationship completely breaks down once salaries reach the maximum of $160,000. It appears that firms could be tacitly colluding on a maximum salary. Colluding on a maximum salary is consistent with theory because there might not exist a symmetric price that raises profits for all firms if they are heterogeneous. Harrington (2016) proves that heterogeneous firms can always find a minimum output price that increases

[21] Almost all the literature on collusion applies to output markets; for now I will assume that similar results hold for labor markets, but an interesting avenue of future research would be to extend these results to input markets and see how they differ. Additionally, the underlying models typically assume explicit collusion, but many of the same indicators were previously proposed for tacit collusion by Posner (1968).

[22] This year was chosen since it was immediately before a salary increase to $180,000 – i.e., the maximum salary would be most binding.

profits for all colluding firms.[23]

Ideally, I would model the labor market with reasonable assumptions and directly test for collusion (e.g., Miller, Sheu and Weinberg, 2021). However, there are two issues. First, the "Folk Theorem" asserts that nearly any set of payoffs in a repeated game is feasible for sufficiently low discount rates. Therefore, detailed empirical analysis is first needed to understand firm strategies. Second, labor markets are more complicated than many product markets since they are matching markets with unobserved quality and important dynamic considerations. Observations are limited because labor markets might be national and salaries are updated only yearly, which reduces my power to distinguish between models from quantitative data alone.[24] Therefore, I focus on combining quantitative data with qualitative evidence, such as direct quotes from industry participants, to build a case for tacit collusion.

First, I document evidence of price leadership. Certain firms, most notably Cravath, consistently announce salary scales and bonuses first. Next, I show how these firms developed price leadership. Initially, Cravath explicitly asked other top firms to follow its salary decisions. Cravath then built a reputation of matching or exceeding any compensation offer over decades of decisions. Firms employ a more complicated strategy than simply matching each other's offers. They exclude some smaller firms from the strategic set. They communicate about their expectations and to avoid misinterpretation. The salary leaders will also slightly exceed any deviating offers in order to maintain their reputation.

Next, I outline key facilitating practices. Firms have close to perfect information about the choices of other firms, and "products" are standardized without discounts, facilitating monitoring, which limits cheating. Finally, I show some suggestive evidence that there are efficiency costs. Lower productivity firms are able to compete for top talent, and relative input use is distorted.

[23]Setting minimum output prices is conceptually similar to setting maximum input prices (salaries).

[24]In contrast, gasoline stations potentially update prices daily and have highly localized markets.

The proposed simple narrative is as follows: prior to 1968, the market was explicitly collusive. Cravath replaced explicit with tacit collusion in 1968 by raising New York salaries. After this shock, local market equilibriums were often tacitly collusive. In the late 1990s, a localized tech boom shock to California firms led them to increase compensation. This increase included their satellite offices in New York, breaking down the local collusive equilibrium. Top New York firms then switched to national price leadership with significant 2000 and 2006 salary increases. These increases established their reputation of matching or exceeding any compensation increase. The market is currently tacitly collusive at the national level.

1.4.2 Price leadership

> Cravath raises, we raise.
>
> - Scott Edelman, chair of Milbank, Tweed, Hadley & McCloy (2016)[25]

Even if firms have mutual beliefs that any price changes might be matched, they might not achieve collusive prices if there is not an agreement about who will lead and when (Harrington, 2017). Establishing price leadership has been recognized as an important step in establishing tacit collusion since at least Stigler (1947).

Industry observers believe that certain firms consistently lead compensation decisions.[26] Again, the most prominent firm is Cravath, with the typical BigLaw associate salary scale often informally referred to as the "Cravath scale." Industry observers even know that Cravath's announcements typically come on Monday after its weekly lunchtime partner meeting.

To show price leadership, I focus on the timing of compensation announcements. Table

[25]Sullivan, Casey. 2016. "Is Following Cravath's Lead the Best Way to Set Salaries?." *Bloomberg Law*, June 7. https://www.news.bloomberglaw.com/business-and-practice/is-following-cravaths-lead-the-best-way-to-set-salaries.

[26]See Appendix Figure A.8 for a satirical guide to a law firm partner's compensation choices.

1.5 catalogues major bonus announcements from 2007-2021.[27] The "First" firm is the initial firm to announce bonuses. The "Standard" firm is the firm whose bonus scale is most commonly matched (if it is a different firm than the "First" firm). Most notably, out of 17 listed bonuses, only three firms have set the standard bonus scale: Cravath (13 times), Davis Polk (3 times), and Milbank (1 time). Milbank's leading scale even comes with caveats. It proposed the exact same bonus scale that Cravath had set the prior year, and no major firm matched it until Cravath did. Therefore, effectively only two firms, Cravath and Davis Polk, have set the prevailing bonus scales for at least the last 15 years.

This dynamic is not due to these firms systematically announcing bonuses at an earlier date. Cravath has announced bonuses as early as October 29th or as late as December 7th and other firms have still waited for them. Many firms wait for them even if another firm announces first. Some firms have matched within hours of Cravath's initial announcement. Take the example of the 2015 bonus: Cravath internally circulated a memo with its bonus scale at 3:15 PM on December 7th. By 3:47 PM (on the same day), Milbank sent a memo to its employees matching Cravath.[28]

When other firms have tried setting bonus scales, Davis Polk or Cravath have typically not just matched but instead exceeded them. Interestingly, Davis Polk has only topped the first movers other than Cravath, despite Cravath leading the majority of bonus announcements. In each of these cases, the first mover eventually ended up increasing their initial offer to match the higher bonus scale.

Similar dynamics exist for salary announcements. Figure 1.11 shows the timing of firm announcements for the 2016, 2018, and 2021 raises. In 2016, Cravath raised starting salaries to $180,000 (from $160,000) on June 6th (Panel A). A significant number of matching announcements quickly followed. In 2018, Milbank further raised starting salaries to $190,000

[27]The primary data source for announcement timings is the industry site *Above the Law*.

[28]Lat, David. 2015. "Associate Bonus Watch: The First Cravath Match - Wow, That Was Fast!" *Above the Law*, December 7. https://www.abovethelaw.com/2015/12/associate-bonus-watch-the-first-cravath-match-wow-that-was-fast/.

on June 4th; however, significantly fewer firms initially matched their scale (Panel B). Instead, many firms waited until the next week when Cravath announced their salary scale. Cravath matched Milbank's starting associate salaries but slightly exceeded Milbank's salary scale for experienced associates. Significantly more firms matched Cravath within a week of its announcement than had matched Milbank. All of the firms that had initially matched Milbank later increased their own offers to match Cravath (including Milbank). A similar pattern was repeated in 2021. Millbank initially raised salaries, and then Davis Polk issued higher raises. Most firms waited to match until Cravath issued its announcement that it was matching Davis Polk. It appears that most major firms wait for Cravath's announcement.

There seems to be a clear pattern of price leadership. A handful of firms in New York City typically set the prevailing maximum compensation scale for major firms nationwide. Cravath is the most prominent, but a few other firms participate; most notably Davis Polk and Simpson Thacher (lead role in 2006-07 salary increases).

1.4.3 Establishing price leadership

> Has Cravath ever not been at the norm?
>
> -Joshua Holt, of Biglaw Investor (2021)[29]

The next question is why are these firms the price leaders? They are not the largest firms; in 2019 Cravath employed 519 attorneys (90th largest US firm), Davis Polk employed 982 attorneys (30th largest), and Simpson Thacher employed 964 attorneys (34th largest). Therefore, their employment decisions do not inherently have broad impacts on the industry and there is no reason to believe that they are especially well-informed about market fundamentals.[30] We need to look further back in history to understand why they are price leaders.

[29]Roy Strom. 2021. "Should We Still Say 'Cravath Scale' If Other Firms Pay More?" *Bloomberg Law*, April 8. https://www.news.bloomberglaw.com/business-and-practice/should-we-still-say-cravath-scale-if-other-firms-pay-more.

[30]There are alternate justifications for price leadership that do not involve collusion, most prominently if

Pre-1968 explicit collusion: Prior to 1968 large law firms did not need a price leader. Instead, they simply explicitly colluded to set maximum salaries. As related in Smigel (1969),

> Starting salaries at the largest New York firms were uniform; the "going rate" was fixed at a luncheon, attended by managing partners of prominent firms, held annually for this purpose. Salaries rose from $4,000 in 1953 to $7,500 in 1963.

This luncheon ran for over forty years. It originated when a partner at Root Clark, Emory Buckner, believed that associates had demonstrated poor judgement because they were too responsive to salary differences. As recounted in Galanter and Palay (1991), he wrote in a letter that he created "the 'big employers' trust... I called twenty firms to lunch – knowing someone in each – and we made an effort to stabilize the situation." More informal collusion might have even pre-dated the luncheon.

1968 raise: Cravath established price leadership by breaking with the "luncheon" cartel in 1968. Evidently, Cravath was unhappy with the scheduled 'going rate' increase to $10,500. Instead, Cravath unilaterally raised starting salaries by almost 50% to $15,000. Cravath's short-term payoff to deviation would be maximized if other firms stuck to the "going-rate." Instead, Cravath took the long view. The issue was not collusion, but rather that the agreed-upon rate was too low. The managing director of Cravath, George Gillespie, publicly issued an invitation to other firms: "We are very hopeful that similar New York law firms will adopt a similar salary policy."[31] The largest and most profitable New York firms ended up matching the increase.

This increase affected legal markets nationwide. Prior to 1968, firms in other markets benchmarked their salaries to the "New York rate." In 1968, prior to Cravath's increase,

---

there is a dominant firm (van Damme and Hurkens, 2004) or if some firms act as "barometers" due to being better informed about market conditions (Cooper, 1997). These firms are concentrated in New York City, so it is not clear why they would be especially well-informed about national markets.

[31] Jack Tate. 1968. "Law Firm Offers $15,000". *Harvard Law Record*, February 15. https://www.iiif.lib.harvard.edu/manifests/view/drs:45687744$4i

over 100 firms had sent their standard letters to the Placement Office of Harvard with some version of "starting salaries will, of course, be competitive with those paid by major New York Law firms."[32] After Cravath's unexpected raise, firms were left to scramble and most markets ended up only partially matching the raise.[33] This also highlights the rationale for Cravath's deviation – they wanted to increase New York salaries relative to other regions because they were losing associates to firms in other cities. A member of the *Harvard Law Review* stated, "New York law firms are suffering in their ability to attract the talent they want. The living conditions in New York are worse than those in Washington and San Francisco." Cravath called the salary increase a "subsidy for New York costs."[34]

1986 raise: New York maintained a salary premium, and years of stable increases to keep pace with inflation followed. In 1985, top New York salaries were around $50,000. This salary figure is equivalent to about $16,000 in 1968 dollars, so associates had seen little real wage growth since 1968. In the mid-1980s, major New York law firms faced increasing losses of associates to investment houses and consulting firms. Cravath responded in 1986 by raising salaries to $65,000.[35] Large New York firms again exactly matched this raise. In general, about 40% of New York firms matched the modal salary each year during the 1980s and 1990s (see Appendix Figure A.1).

In 1980, most firms had large offices in only one state (Table 1.3). Therefore many cities had local price leaders that determined the salary differential with New York. Table 1.2 shows the average share of firm-year observations that match the local modal salary across years. The rates vary across cities, but some are very high. About 50-60% of Boston firms

---

[32]Jack Tate. 1968. "Law Firm Offers $15,000". *Harvard Law Record*, February 15. https://www.iiif.lib.harvard.edu/manifests/view/drs:45687744$4i

[33]See Appendix Figure A.17 for the resulting salary differential by region. Even the federal government had to significantly adjust salaries – it changed policy to allow new lawyers with outstanding academic records to begin at level GS-12, which paid $12,174, instead of GS-11, which only paid $10,203.

[34]Jack Tate. 1968. "Law Firm Offers $15,000". *Harvard Law Record*, February 15. https://www.iiif.lib.harvard.edu/manifests/view/drs:45687744$4i

[35]Tamar Lewin. 1986. "At Cravath, $65,000 to start." *The New York Times*, April 18. https://www.nytimes.com/1986/04/18/business/at-cravath-65000-to-start.html.

offered the same salary in any given year prior to 2000.[36] The existence of local price leaders was well-known. For example, reporters covering Seattle's salary increases wrote: "Many in the field here are watching for moves by Seattle's largest firms, particularly Preston Gates & Ellis LLP and Perkins Coie LLP. The two firms tend to set the standard for salaries."[37] However, during the 1980s and 1990s, firms increasingly began to expand across markets. These conditions led to the breakdown of local markets and a salary war during 1998-2000.

2000 raise: The New York salary differential eroded during the 1990s, especially compared to California. California firms were dealing with local associate supply shocks as technology companies began poaching associates. Between 1989 and 1997, average salaries at San Francisco firms increased from $60,000 to $79,000 ($19,000 increase). In comparison, the average salary of New York firms increased only from $78,000 to $84,000 ($6,000 increase) in the same period (see Appendix Figure A.15). In response, New York firms began raising salaries in 1998. This led to a breakdown of the local equilibrium – the share of New York firms exactly matching the modal salary dropped sharply in the late 1990s (see Appendix Figure A.1).

This process culminated in 2000, when the small Silicon Valley firm, Gunderson Dettmer Stough Villeneuve Franklin & Hachigian raised starting salaries to $125,000. Soon after, three other small Silicon Valley firms matched. All of these firms were small; none were in the largest 250 US firms in 1999. That changed when San Francisco's Brobeck, Phleger & Harrison (27th largest firm) announced that all new associates would receive $125,000 because ,"We want to make it harder for people to leave us for clients." Importantly, this firm had a New York office with about 50 attorneys, so their salary substantially exceeded the previous New York maximum. A partner at a New York firm said, "You now have California setting a trend, this has never happened before."[38]

[36]The modal salary is also often the local maximum.

[37]George Erb. 2000. "Dot-Coms bid up pay at law firms." *Puget Sound Business Journal*, February 13. https://www.bizjournals.com/seattle/stories/2000/02/14/story3.html.

[38]David Leonhardt. 2000. "Law Firms' Pay Soars to Stem Dot-Com Defections." *The New York Times*, February 20. https://www.nytimes.com/2000/02/02/business/law-firms-pay-soars-to-stem-dot-com-

Soon after, Davis Polk made the dramatic decision to raise salaries by $25,000 to $125,000 demonstrating a willingness to match any salary raise. The size of the increase generated significant media coverage and publicized the new salary point.[39] The other major New York firms soon followed. Notably, many other large firms outside of New York also matched the new salary, especially those in California, Chicago, Boston, and Washington D.C.. [40]

2006 raise: The collapse of the Dot-Com bubble in 2000 significantly reduced the incentive for California firms to raise salaries. The maximum salary remained at $125,000 until 2005 before some California firms again tried to exceed it. In 2005, the local Los Angeles firm Irell & Manella raised salaries to $135,000. Another local firm, Quinn Emanuel Urquhart & Sullivan, matched.

First, Gibson, Dunn & Crutcher matched the raise. They are a national firm but were headquartered in Los Angeles. Simpson Thacher then decided to exceed the California raises and increased starting salaries to $145,000. Davis Polk soon matched them and Cravath followed. After these firms matched, most other large firms followed suit. Prior to this increase, it was not clear to the Los Angeles firms that they would trigger a response from the larger New York firms. William Urquhart, head of Quinn Emanuel Urquhart & Sullivan, specifically stated that his firm "hoped bigger firms wouldn't follow so we could separate ourselves, but they did."[41] This episode helped solidify the belief that larger firms would match or exceed most increases.

As shown previously, Cravath and Davis Polk continued enforcing price leadership in the 2016, 2018, and 2021 raises. Therefore, these firms, especially Cravath, have established a long reputation of being salary leaders. This reputation of leading New York increases

defections.html.

[39]For example, there were several articles covering the raise in the New York Times and the Wall Street Journal. The rise of associate message board groups also helped publicize salary announcements – see Taras and Gesser (2003) for a discussion of this phenomenon.

[40]David Leonhardt. 2000. "Law Firms See a Bill Come Due." *The New York Times*, May 22. https://www.archive.nytimes.com/nytimes.com/library/tech/00/05/biztech/articles/22neco.html.

[41]Ellen Rosen. 2006. "For New Lawyers, the Going Rate Has Gone Up." *The New York Times*, September 1. https://www.nytimes.com/2006/09/01/business/01legal.html.

has its origin in Cravath's break from the "luncheon" cartel in 1968. The reputation was solidified through leading years of New York raises. These firms then extended this reputation nationwide by showing a commitment to match or exceed any raise through the 2000, 2006-07, 2016, 2018, and 2021 raises.

1.4.4 Strategic Behavior

> Like our bigger competitors, we've paid our first-year lawyers $160,000 until now. But we're growing and doing more exciting work, and we want to attract even more top-level talent.
>
> - John Zavitsanos, founding partner of AZA (2013) after raising salaries to $170,000[42]

A closer examination of firms' strategic behavior confirms that their strategy extends beyond simple matching. Price leaders exclude some smaller firms from the strategic set. Firms communicate about their plans to align expectations and reduce miscommunication. Price leaders slightly exceed competing offers to maintain their reputation.

Strategic set: Firms form beliefs about what firms are in the relevant strategic set. If large firms committed to matching all salary raises, then any small firm could force large salary increases. Therefore, certain smaller firms are allowed to offer above-market compensation. Table 1.6 shows some firms that offered above-market compensation in 2012.[43] These are typically smaller or boutique firms that are focused on litigation. While they are smaller, it is worth remembering that Cravath is not particularly large. For example, Boies, Schiller & Flexner and Williams & Connolly together are larger than Cravath. Some of these firms offer significantly higher compensation. Famously, Wachtell Lipton sometimes offers bonuses equal to associate salaries, implying a doubling of the "going rate" compensation.

---

[42] David Lat. 2013. "The $160K-Plus Club Welcomes A New Member" *Above the Law*, January 11. https://www.abovethelaw.com/2013/01/the-160k-plus-club-welcomes-a-new-member/.

[43] Compiled from Above the Law references. There might be additional small firms with unavailable data.

For example, when Cravath raised salaries to \$65,000 in 1986, the small firm Reboul MacMurray re-raised to \$70,000. Large New York firms did not match this rate, instead "The larger firms, by and large, have tended to dismiss the Reboul MacMurray action as an aberration by a non competitor. Cravath's pay offer is viewed as the new standard."[44] In these cases, a simple matching strategy would have pushed salaries significantly higher and reduced the gains to tacit collusion. Instead, firms were more strategic about who to match. The exclusion of some smaller firms from a stable cartel is consistent with theoretical predictions for heterogeneous firms in Bos and Harrington (2010).[45]

An interesting case is the 2008 year-end bonus after the onset of the Great Recession. The firm Skadden initially announced that bonuses would remain at \$35,000, the same as 2007 (less a one-time special bonus). The next day Cravath issued its bonus memo, which instead offered \$17,500. This bonus memo was not issued as typical after Monday meeting. Potentially, Cravath wanted to intervene before other firms began matching Skadden. If other firms matched Skadden and Cravath made a below-market offer then it might lose credibility. Instead, Cravath could portray Skadden as an outlier, and, in fact, other firms ended up matching Cravath.[46] Again, a simple policy of matching the highest firm would predict that at least some firms match Skadden.

Firms have also tried to unsuccessfully exclude firms that increase salaries from the strategic set. For example, during Cravath's 1986 increase, Milbank said it would refuse to follow Cravath and hoped it would deter other firms from "going over the cliff." A few other firms tried making similar announcements, but ultimately they had to renege on their promises. As Roseman Colin said, "It was our hope that a substantial number of other large firms would also choose this course for themselves. This failed to occur and accordingly, we

[44]Gary Hengstler. 1986. "If I Can Make It There..." *ABA Journal*, August 1.

[45]The fact that some of these firms have salaries slightly above the collusive maximum is also consistent with Bos and Harrington (2010).

[46]Also, Cravath offered significantly more explanation in the bonus memo than usual. This might be due to having to explain a bonus reduction to employees, but Cravath also included forward guidance, such as "they may receive significantly reduced or no year-end bonuses next year." Bonuses were in fact reduced again the following year.

have determined to adjust our associate salaries so as to remain competitive."'[47]

Communicating compensation: Firms strategically communicate about their compensation to set expectations and avoid misinterpretation. Cravath's large raise in 1986 was implemented as a new $12,000 "housing allowance" due to high New York costs. The goal might have been to discourage firms outside of New York from matching the raise. Other New York firms structured their compensation in slightly different ways but, as Alexander Forger of Milbank put it, "no matter how you slice it, we'll all be in the same ballpark."[48]

The "going rate" in 2000 was $125,000, but Skadden offered salaries of $140,000. Again, a simple matching strategy would mean other firms should raise their salary to match. Instead, *The New York Times* reported that Skadden's leader, Robert C. Sheehan, believed that "because Skadden will pay relatively small bonuses, its partners do not consider the raise to be an attempt to top the other firms' salaries."[49] Therefore, other firms did not feel the need to match Skadden's salary offer. There is little justification for this statement in a competitive market. Imagine the lowest priced retail firm publicly stating "that if you consider all-in costs then we are not a better value than our competitors." Skadden's bonuses were in fact $15,000 lower to exactly offset the $15,000 higher salary.

Similar events occurred during the 2006 raise. Before Simpson Thacher raised salaries to $145,000, another major firm, Sullivan & Cromwell, had raised them to $145,000. However, firms did not feel the need to match. Why? The head of Sullivan & Cromwell's associates committee, Benjamin Stapleton III, said in an interview, "Total compensation this year could be more, less, or the same as last year." Confirming, the actual memo said that the "increase represents a shift of that amount from the 2006 year-end bonus which you would otherwise

---

[47]Gary Hengstler. 1986. "If I Can Make It There..." *ABA Journal*, August 1.

[48]Gary Hengstler. 1986. "If I Can Make It There..." *ABA Journal*, August 1.

[49]David Leonhardt. 2000. "And Let the Lawyers Sing: 'Glory to the Salary King'." *The New York Times*, February 4. https://www.archive.nytimes.com/nytimes.com/library/financial/020400law-salaries.html.

receive."[50] A participant described Cravath's next partner meeting as, "Everybody looked around the room ... nobody seemed to care very much about it because it was clear that Sullivan was just moving money around." That changed when Simpson Thacher released its raise to $145,000 and stated, "The bonus portion of your compensation will be announced at year-end as in prior years." This was interpreted to mean Simpson Thacher's increase (unlike Sullivan & Cromwell's) would not be offset by decreased bonuses. Soon the major firms matched Simpson Thacher's raise.[51]

Firms have also tried to strategically communicate about compensation to allow them to exceed the market. An example is the structure of Cahill Gordon's 2010 year-end bonus. *Above the Law* reported in 2010, "A tipster believes that Cahill Gordon intends to double the Cravath bonus. But not all at once. Cahill doesn't want to look like it's breaking the market ... Partners said the associates would be paid one bonus in December and another in January (assuming to make it look like they're just matching the market)."[52] Despite the effort, the major firms ended up issuing special "Spring bonuses" with a similar value.

Punishment: What is the payoff to firms that deviate? With a higher salary, they can increase both the number and quality of associates. If firms deviate, then the maximum salary will be raised to either match or exceed their offer. Therefore, deviating does not improve the salary offer relative to other major firms. Firms might still have limited payoff to deviating since not all firms or competing industries will match. Competing industries act as a competitive fringe that disciplines the rents that can be extracted from collusion. Most significant raises have been triggered at least partially by competition from other industries, especially finance firms in New York or technology firms in California.

[50]The stated concern was that some associates were living paycheck to paycheck, so shifting compensation from bonuses to salary would reduce cash flow issues.

[51]Anna Schneider-Mayerson. 2006. "Sullivan Bonus Babies Get Lift in Salaries As White Shoes Tap." *Observer*, February 20. https://www.observer.com/2006/02/sullivan-bonus-babies-get-lift-in-salaries-as-white-shoes-tap/

[52]Elie Mystal. 2010. "Associate Bonus Speculation: Will Cahill Double the Market In Secret?". *Above the Law*, 2010. https://www.abovethelaw.com/2010/12/associate-bonus-speculation-will-cahill-double-the-market-in-secret/.

The fact that the price leaders typically exceed any deviations is important for two reasons. First, it increases the punishment. Second, firms receive a reputation benefit from being the "salary leader," which they use when recruiting. Price leaders prevent deviators from getting this benefit by exceeding their offer. For example, Simpson Thacher raised bonuses in 2014, but the bonus is called the "Davis Polk bonus" since Davis Polk exceeded Simpson Thacher's offer.[53] Since 2009, every major firm, other than Cravath and Davis Polk, has been exceeded any time they try to raise compensation.

The next question is why do so many firms exactly match the maximum when they could offer lower salaries? There is the standard incentive that firms want to increase the quantity and quality of associates, but it appears that both potential associates and potential clients view salary as a strong signal of unobserved quality. This creates a strong incentive for firms to match the leading compensation exactly since there is a discrete change in perceived quality. Many commentators have issued similar statements to this one from a law firm consultant: "For years, Cravath has set the bar for what it pays its associates, and other law firms follow them like lemmings to avoid any negative inference about their financial strength; and they need to reassure their associates that they're a top-tier firm."[54]

1.4.5 Facilitating practices

> It's a funny phenomenon that it's all very public.
>
> -Steven J. Steinman, a partner at Fried, Frank, Harris, Shriver & Jacobson. (2006)[55]

There are a variety of practices that facilitate tacit collusion in this setting. First, trade publications, private associations, and informal forums allow significant communication.

[53]Lat, David. 2015. "Where Are The Biglaw Bonuses? Associate Bonus Watch, Day 3." December 2. https://www.abovethelaw.com/2015/12/where-are-the-biglaw-bonuses-associate-bonus-watch-day-3/.

[54]Peter Lattman. 2012. "Cravath Sets the Tone for Law Firm Bonuses." *The New York Times*, November 26. https://www.dealbook.nytimes.com/2012/11/26/cravath-announces-bonuses-for-its-associates/.

[55]Ellen Rosen. 2006. "For New Lawyers, the Going Rate Has Gone Up." *The New York Times*, September 1. https://www.nytimes.com/2006/09/01/business/01legal.html.

Communication allows firms to monitor each other's choices and set expectations. Experimental evidence shows that communication significantly eases collusion (Cooper and Kuhn, 2014). Antitrust authorities also emphasize the important role of monitoring in sustaining collusion: "A market typically is more vulnerable to coordinated conduct if each competitively important firm's significant competitive initiatives can be promptly and confidently observed by that firm's rivals." (DOJ and FTC, 2010) Second, associate positions and compensation are standardized since most major firms use lockstep structures. Product standardization in output markets has been shown to facilitate collusion (Harrington, 2018).

Communication: There are a significant number of publications that cover BigLaw firms. These organization disseminate information on compensation and also allow public communication between firms. For example, the site *Above the Law* publishes bonus and salary announcements, often within hours of the initial internal memo (the site relies on associates to anonymously share memos). This allows firms to match or exceed compensation decisions, sometimes issuing memos within an hour of the initial announcement. Firms can even retroactively increase salaries or bonuses if necessary. Therefore, firms can perfectly monitor each other's actions with a very limited delay.[56] Punishment has to be simple and monitoring has to be near perfect and rapid because of the large number of law firms.[57]

Firms also publicly communicate to align on expectations for increases. After major raises in 1986 and 1987, large New York firms did not want further increases. Early in 1988, a Sullivan & Cromwell partner stated that he had "no reason to think there will be an increase." A Davis Polk partner commented that he did not anticipate any increases beyond

[56]There is near perfect monitoring of choices, but not necessarily of firm fundamentals, which means the decisions of price leaders could violate participation constraints and trigger price wars.

[57]There are other cases of tacit collusion with a large number of firms. NASDAQ market makers specify "bid" and "ask" prices in terms of eighths of dollars. Christie and Schultz (1994) noticed that market makers almost never specified odd eighths which guaranteed a minimum spread of $0.25. Economists hired by NASDAQ argued that collusion was inconceivable due to the high number of market makers (over 400) and low barriers to entry. However, the DOJ investigated and reached a $1 billion settlement because the regularity disappeared after the paper and there was evidence of intimidating phone calls and refusal to deal with violators. Collusion was sustainable because the market maker reaction time was measured in minutes, meaning there is limited payoff to cheating (Harrington, 2018).

$72,000. Of course, these announcements come with a caveat. A Skadden partner said he did "not expect any increase in the new people's salaries there unless another firm raises its level." A senior partner at Weil, Gotshal & Manges similarly opined, "Unless someone else makes a startling announcement, then we don't expect our increases will be anything more than modest. And the rumor and gossip on the Street is that no one will be making a major increase."[58]

There are many private associations where law firm leaders come together that allow private discussion, for example at Citigroup Private Bank meetings.[59] Because of the private nature of these meetings, it is unclear to what extent discussion occurs. However, there is suggestive evidence that it does happen. According to the *ABA Journal*, one firm leader "believes year-end bonuses will remain unchanged, although special bonuses will be eliminated. He adopted that view based on the discussion during a meeting of law firm managing partners hosted by Citigroup Private Bank in August."[60] Interestingly, these are exactly the bonus levels adopted by Skadden in 2008 before being undercut by Cravath. After the 2016 salary increase, the head of advisory services at Citi Private Bank wrote, "In our conversations with firm leaders, many express bafflement as to why so many firms adopted the increases when their productivity and profitability results couldn't support them."[61] These quotes suggest that firms do privately discuss these matters.

Firms might also discuss compensation in more informal private forums. An anonymous associate allegedly overhead a private conversation between a Cravath partner and a partner at another law firm: "Susan Webster of Cravath [while attending] a meeting introduces

[58]Stephen Labaton. 1988. "Business and the Law; Young Lawyers' Salaries Stabilize". *The New York Times*, June 13. https://www.nytimes.com/1988/06/13/business/business-and-the-law-young-lawyers-salaries-stabilize.html.

[59]Citi offers "law firm advisory services" and has ongoing relationships with "over 700 prominent law firms."

[60]Debra Weiss. 2008. "Will Bonuses Be Cut? Cravath and S&C Key to the Answer." *ABA Journal*, November 18. https://www.abajournal.com/news/article/will_bonuses_be_cut_cravath_and_sc_key_to_the_answer.

[61]Debra Weiss. 2017. "Some law firm leaders question associate pay hikes amid tepid year." *ABA Journal*, February 15.

herself to another BigLaw partner. [The o]ther partner says, 'Oh thanks for the bonus, [it] really was great.' Susan smiles and says, 'Yeah I know.' Then she complains about people wanting spring bonuses."[62]

Standardization: Standardization can facilitate collusion by simplifying monitoring and communication. For example, Genesove and Mullin (2001) show that the Sugar Institute cartel explicitly colluded on standard business practices to facilitate implicit price collusion. Among BigLaw firms, associate positions and compensation are standardized. Most BigLaw firms use a lockstep system, which means associates advance based on years of experience. Therefore, a fourth-year associate who transitions between firms remains a fourth-year associate. Firms use posted prices and do not offer "discounts," i.e., they typically do not offer individualized compensation or signing bonuses.[63] It is especially interesting that firms do not rely on signing bonuses, since they are common in many other industries.[64]

The practice of standardized compensation actually has its origin in explicit collusion. Robert Swaine (1947) wrote in the official history of Cravath, Swaine, & Moore:

> Adoption by other city offices of many of the same principles on which the "Cravath system" is based led, about 1910, to competitive bidding for the highest-ranking men of the leading law schools. This gave a few men inordinately high beginning salaries, sometimes double those of the generally applicable scale. The discrimination among the men just coming out of law school became unfair and

[62]Elie Mystal. 2011. "What Do Cravath Partners Say About The Bonuses To Other Biglaw Partners When They Think Nobody Is Listening?" *Above the Law*, December 2011. https://www.abovethelaw.com/2011/12/what-do-cravath-partners-say-about-the-bonuses-to-other-biglaw-partners-when-they-think-nobody-is-listening/. When asked by Above the Law about the overheard conversation, Susan Webster did reply "The characterization of your report is inaccurate."

[63]The absence of signing bonuses might be due to historical norms against poaching workers. While these norms have broken down at most firms, they used to be widespread. For example, in the 1950s and 1960s, "The firms will not pirate an employee from another law office, and they maintain a gentleman's agreement to pay the same beginning salary." (Smigel, 1969)

[64]There are some signs that this standard is currently changing with increasing reports of signing bonuses in 2021. See Casey Sullivan and Jack Newsham. 2021. "Kirkland & Ellis has offered up to $250,000 signing bonuses to young lawyers amid nonstop M&A and capital-markets work." *Insider*, May 19. https://www.businessinsider.com/kirkland-ellis-offered-junior-lawyers-signing-bonuses-250k-big-law-2021-5.

> made the initial salary offered too important a criterion in the choice of offices. Within a few years the evils of the practice were admitted by the offices and strongly objected to by the faculties of the law schools; on their suggestion it was abandoned after World War 1, following a conference among the managing partners of the larger offices. Beginning salaries thereafter tended to become uniform...

Therefore, the practice of offering uniform salaries was reached by agreement. For output markets, there are non-collusive rationales for posted pricing without discounts if they significantly reduce search costs or the cost of selling (in this case "buying" labor). However, if these costs are not significant, then posted pricing without discounts is inimical to competition (Harrington, 2011).

1.4.6 Efficiency costs

> No doubt we will be raising as will other firms; this market is very efficient in that way.

-John Quinn, co-founder of Quinn Emanuel Urquhart & Sullivan (2016)[65]

There are substantial reasons to believe that these behaviors meaningfully distort associate compensation ("markdown"). First, there have been significant jumps in compensation that are only feasible if associates were previously earning substantially less than their marginal product.

Second, partners react as if the decisions of salary leaders are meaningful. I have already discussed a few examples previously, but there are many others. For example, after Cravath kept 2010 bonuses at the low 2009 rates an anonymous partner succinctly responded, "Oh,

[65]"Law Firm Cravath Raising Starting Salaries to $180,000." 2016. https://www.consultzg.com/ideas-and-insights/news-mentions/law-firm-cravath-raising-starting-salaries-to-180000/ (accessed on 6/21/2021)

thank God."[66] After the 1986 increase, Joseph Bainton, a partner at Reboul MacMurray, said "I hope we don't get into a wage spiral. After all, it's coming out of my salary and my partners."[67]

Even BigLaw clients react as if these decisions matter.[68] For example, Bank of America's top lawyer sent an email to law firms after Cravath's 2016 salary raise saying that the pay raises were "unjustified" and that the bank would not help firms absorb increased costs.[69] After Cravath cut bonuses in 2008, the firm's head, Evan Chessler, stated "I've got to tell you, and I don't want to name any names, but I have gotten calls from a half dozen clients this morning thanking me."[70]

Countering this view, some believe that new associates are paid more than their marginal product. Many of these objections are normative statements about what associates should make. For example, news articles will cite the fact that associates are paid more than Supreme Court Justices. Or when referring to pay increases they use language such as "princely sum" or "exorbitant." However, as discussed in Harrington (2005), price levels are poor measures of collusion.

A different critique focuses on the fact that clients often complain that new associates do not provide sufficient value to justify their billing rate. However, the value to the law firm is based on the gap between the billing rate and hourly compensation. Billing rates and hourly compensation are not necessarily the same. For example, Figure 1.13 shows that associate billing rates (relative to partners) increased even while associate compensation (relative to

[66]2009. "Cravath Bonuses Hold at 2009 Rates." *New York Law Journal*, November 23. https://www.law.com/newyorklawjournal/almID/1202475234897/.

[67]Gary Hengstler. 1986. "If I Can Make It There..." *ABA Journal*, August 1.

[68]It is a recurring practice for industry new publications to produce articles about unhappy clients after any major salary or bonus increase.

[69]Sara Randazzo. 2016. "Corporate Clients Push Back After Law Firms Hike Starting Salaries". *The Wall Street Journal*, June 15. https://www.wsj.com/articles/companies-push-back-at-law-firms-starting-salary-hikes-1466029554.

[70]Aric Press. 2008. "Cravath Cuts Bonuses, Hints at 2008 Financials." *The AmLaw Daily*, November 21. https://www.amlawdaily.typepad.com/amlawdaily/2008/11/cravath-cuts-bo.html.

partners) decreased. In 2014, a new associate in a major law firm might have a salary of $160,000, bonus of $10,000, and work 2000 billable hours; then the associate earns $85 for every billable hour.[71] In comparison, the average minimum associate billing rate was $300 for firms paying the "going rate."[72]

Quantifying the markdown (which differs by firm) and the total efficiency costs of collusion is beyond the scope of this paper. However, I do provide some suggestive evidence that efficiency costs exist. The two likeliest sources of efficiency losses are if productive firms are inefficiently small or if relative input usage is distorted.

Collusion can lead to some firms being inefficiently small. The OPEC cartel raises oil prices, allowing marginal producers, with higher marginal costs, to produce (Asker, Collard-Wexler and Loecker, 2019). Productive firms are unable to raise compensation to expand without losing some of the rents in the collusive equilibrium. Some of the workers end up at less productive "fringe" firms that are able to compete due to the artificially depressed input costs. There is anecdotal evidence that large price jumps occur partly to remove these firms. Flood (1989) said the goals of the 1986 Cravath increase were "to persuade associates to stay longer on average than they had been doing hitherto and to exterminate a stratum of law firms that would find it difficult to compete for the most highly qualified law school graduates." Commentators will sometimes use the fact that some firms will be put out of business as an argument that associate salaries cannot be raised. However, the continued existence of too many unproductive "fringe" firms is potentially a symptom of an unhealthy market.

Galanter and Palay (1991) find a structural break in firm growth rates around when Cravath broke with the "luncheon" cartel in 1968. Potentially, these higher salaries allowed more productive firms to expand. Table 1.7 looks at which firms hired more

[71] The hourly pay of associates is actually significantly lower, since only 70-80% of hours worked are billable. These estimates are also conservative; many associates at top firms work for significantly more than 2,000 billable hours.

[72] Other inputs are used to produce an associate billable hour (e.g., support staff), but it is clear law firms would earn profit on each associate billable hour even with significant additional costs.

associates after the salary increase in 2006-07 (conditional on paying the previous maximum salary of $125,000). More productive (profitable) firms expanded more after the increase. Continued collusion might be one reason (among many) that the legal market is significantly less concentrated than many other white-collar service markets, such as accounting or management consulting.

A second potential inefficiency is if the salary distortions alter input usage. For example, firms might want to hire more associates but are unable to. Firms typically target specific ratios of associates to partners. If this ratio declines for firms, then it might signify that firms are unable to recruit sufficient associates. Table 1.7 shows that firms that saw the larger declines in the ratio of associates to partners from 2002-05 (when salaries were frozen) hired relatively more associates after the salary increases in 2006-07. Therefore, it seems these firms were unable to maintain their desired number of associates at the lower salary levels and they hired more associates once they could.

Lower associate salaries also means that firms want to have higher rates of associates to partners, since they earn additional rents. This dynamic could force firms to use stricter "up or out" policies to maintain their high associate to partner ratio. New York firms, which might have the largest gap between productivity and compensation, also have the highest ratios of associates to partners and strictest "up or out" policies.

There could be other efficiency losses. If law firms cannot adjust salaries to compete for associates, then they might adjust on other margins such as recruiting earlier.[73] Earlier recruitment might have real efficiency costs since law firms have noisier signals of potential ability, reducing match quality. Artificially low associate salaries might also reduce the incentive to invest in alternative production technologies. For example, firms might experiment less with the use of alternatives to elite law school graduates, such as paralegals or graduates from lower-ranked institutions. It could also decrease the long-run supply of lawyers or partners. The loss of high-quality associates to other industries due to low

[73]Recruitment now begins as early as the winter of the first year, when firms have only one semester of grades to observe.

salaries also reduces the long-run supply of high quality partners.

1.5 Conclusion

The market for BigLaw associates has a long history of collusion. Explicit collusion began shortly after the introduction of the "Cravath system" in the early 20th century with the "luncheon" cartel and other agreements between major law firms and lasted for at least forty years. Explicit collusion was replaced with tacit collusion by Cravath in 1968 when it explicitly invited other leading New York City firms to match its salary rates. Collusion even might be considered the natural state of the market. It also might have existed long enough that many market participants do not even realize that their behavior has its roots in both explicit and tacit collusion.

Price leaders' strategy extends beyond the simplest matching strategy. They make strategic decisions about which firms to exclude from the strategic set. They communicate about total compensation to reduce misinterpretations and set expectations. They preserve price leadership by slightly exceeding initial competing offers to reduce reputational gains to cheating. These strategies are enabled by a variety of facilitating practices. Communication, in trade publications, industry organizations, and private settings, allows firms to set expectations and perfectly monitor each other. Associates levels and compensation are standardized and firms do not offer "discounts" (individualized salaries), reducing the ability of firms to cheat and simplifying monitoring. Some of these facilitating practices, such as standardized salaries, have their origin in explicit collusion.

The objective of this paper is to answer a straightforward question: what explains uniform salaries in the market for BigLaw associates? The answer is also straightforward: tacit collusion. An interesting question for future research is to model and quantify the effect of this collusion. The large salary jumps suggest that collusive markdowns could be significant in some years, and there is suggestive evidence that collusion does have efficiency costs. Quantifying the effect of collusion requires a model that accounts for quality differences, effects on the supply of lawyers, firm entry and exit, and many additional dynamics. Given the long-run nature of collusion, there might even be important effects on the structure of production and the development of production technology. It is also interesting to consider how the rents from collusion are divided between other input providers (e.g., partners) and the consumers of legal services.

Economic collusion is not necessarily illegal from a legal perspective since it could fall into the category of "conscious parallelism." However, detailed legal analysis is beyond the scope of this paper and the qualifications of the author. A separate interesting question is whether there are potential remedies, but it is difficult to identify clear remedies for tacit collusion that does not rely on direct communication (Turner, 1962). I also do not take a stance on firm motivation. For example, Cravath partners could seek price leadership for reputational benefits (e.g., they derive utility from being seen as the "best" firm) without explicitly desiring tacit collusion.

Another interesting strand of future research would be to examine other labor markets for potential collusion. The absence of uniform salaries in a market does not mean tacit collusion does not occur; law firms have to rely on relatively simple techniques, such as exact salary uniformity, due to the large number of firms. In 2005 the accounting firm KPMG had to pay significant penalties for creating fraudulent tax shelters. While KPMG was litigating this case, the other three "Big Four" accounting firms unofficially agreed to not poach workers from KPMG.[74] Other professional service industries, such as consulting or finance, could face similar issues. Medical professions are also a good focus for research, since they have been a frequent target of the initial labor market antitrust cases. Meatpacking companies also have a history of collusion (Huang, 2020) and poultry processors currently face an active wage-fixing suit that alleges firms depressed pay through illegal data exchanges and secret meetings at industry conventions.[75] Universities are also a potential target of investigation.

[74] Bill Carlino. 2005. "The Big 4: A Growing Concern." *Accounting Today*, October 9.

[75] Mike Leonard. 2021. "Tyson, Pilgrim's, Hormel to Face Poultry Worker Wage-Fixing Suit." *Bloomberg Law*, March 21. https://www.news.bloomberglaw.com/antitrust/tyson-pilgrims-hormel-to-face-poultry-worker-wage-fixing-suit.

The heads of top economics departments used to agree on pay and teaching requirements at the Annual Meeting of the American Economic Association (Krueger, 2017); potentially similar practices still occur in other fields.

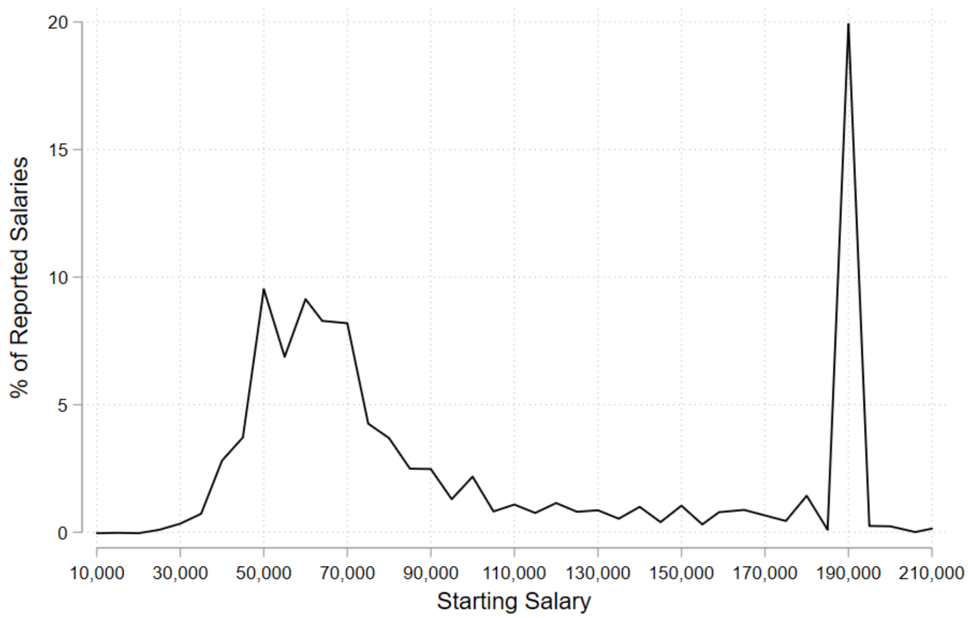
Finally, there might be many localized cases where collusion can even span across industries. While there are few current detailed case studies of local labor markets in economics, some historical studies find that informal no-poach agreements are pervasive. For example, Myers and MacLaurin (1943) followed 1500 workers at thirty-seven companies in one city over six years. They found that "gentleman's agreements" not to "pirate" each other's employees were a significant barrier to worker mobility, and Reynolds (1951) found a similar result for another city. This evidence aligns with the widespread use of non-compete agreements, including notably for Jimmy John's sandwich makers (Ashenfelter and Krueger, 2018).

We currently do not understand how widespread these problems are because meaningful regulatory enforcement is a recent phenomenon. But it is clear there are important policy implications; if collusion is common, then mergers should not be scrutinized for coordinated effects just in output markets (Asker and Nocke, 2021), but also in input markets. For example, many animation studios had collusive no-poach agreements. Several participants in this scheme, including Pixar and Lucasfilm by Disney, have since been acquired, potentially facilitating future collusion. The former head of the Antitrust Division, Makan Delrahim, recently said (in regard to no-poach agreements), "In the coming couple of months you will see some announcements, and to be honest with you, I've been shocked about how many of these there are, but they're real."[76] These examples all suggest that collusion in labor markets deserves increased attention from researchers and regulatory authorities.

[76] Matthew Perlman. 2018. "Delrahim Says Criminal No-Poach Cases Are in the Works." *Law360*, January 19, https://www.law360.com/articles/1003788/delrahim-says-criminal-no-poach-cases-are-in-the-works.

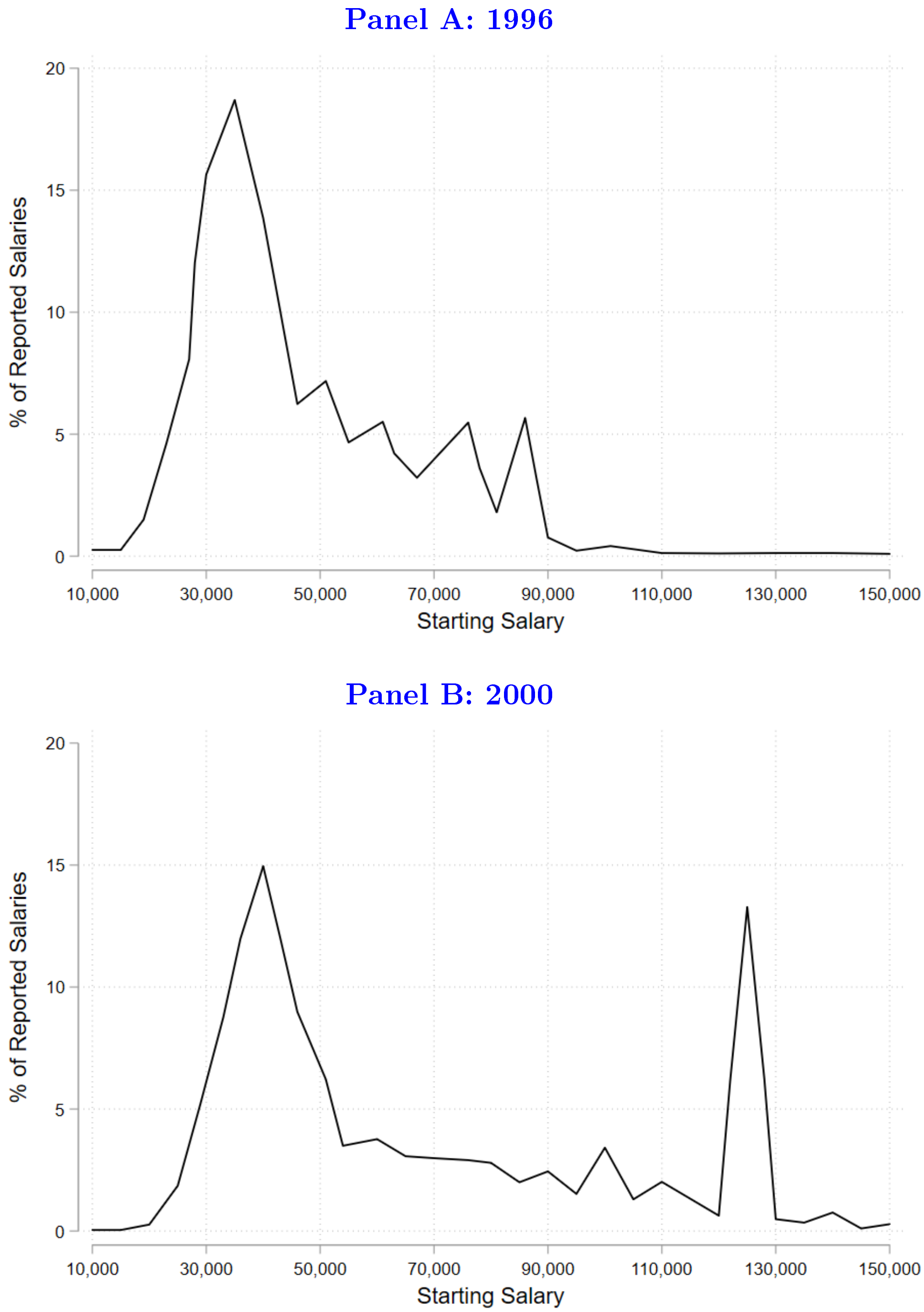
1.6 Tables and figures

Figure 1.1: Distribution of starting salaries for new law school graduates (2019)



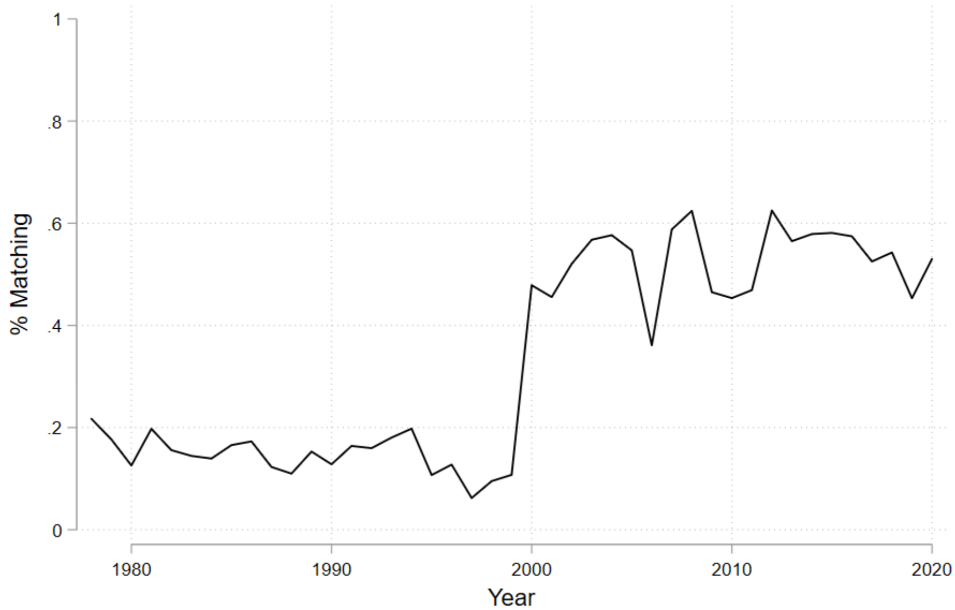
Note: Conditional on finding a job. From published National Association for Law Placement figures (2019 dollars)

Figure 1.2: Distribution of starting salaries for new law school graduates (1996 and 2000)



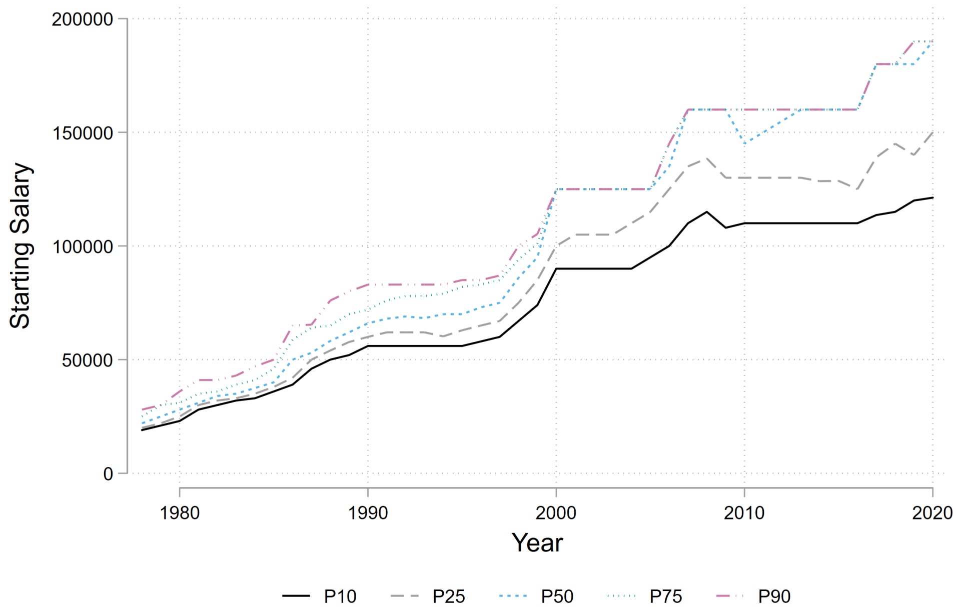
Note: Conditional on finding a job. From published National Association for Law Placement figures (nominal dollars)

Figure 1.3: Share of firms matching modal starting salary (NLJ 200)



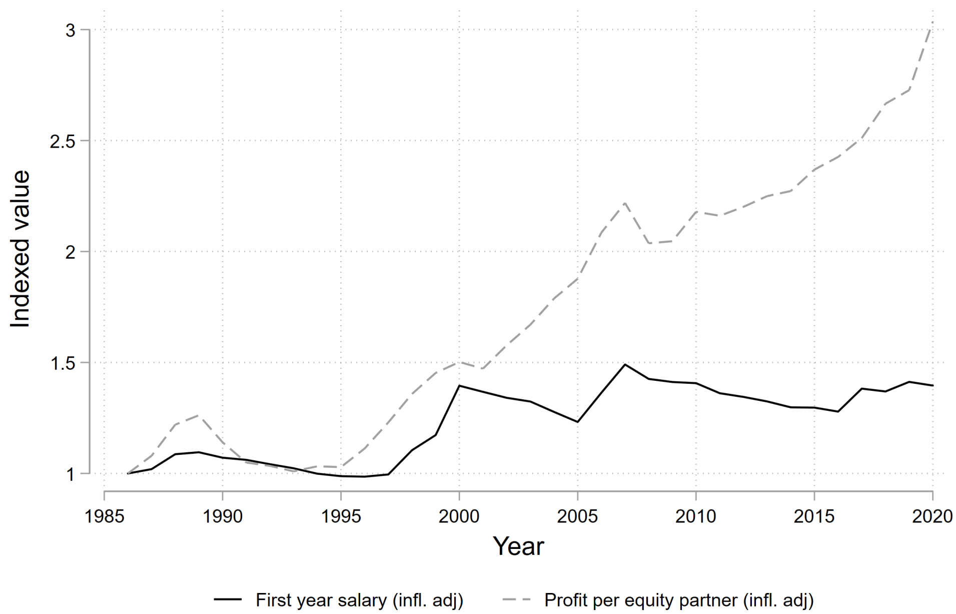
Note: Figures are conditional on firm's reporting starting salaries in NLJ surveys. Starting salaries represent the highest reported starting salary for each firm.

Figure 1.4: Distribution of starting salaries by year (NLJ 200)



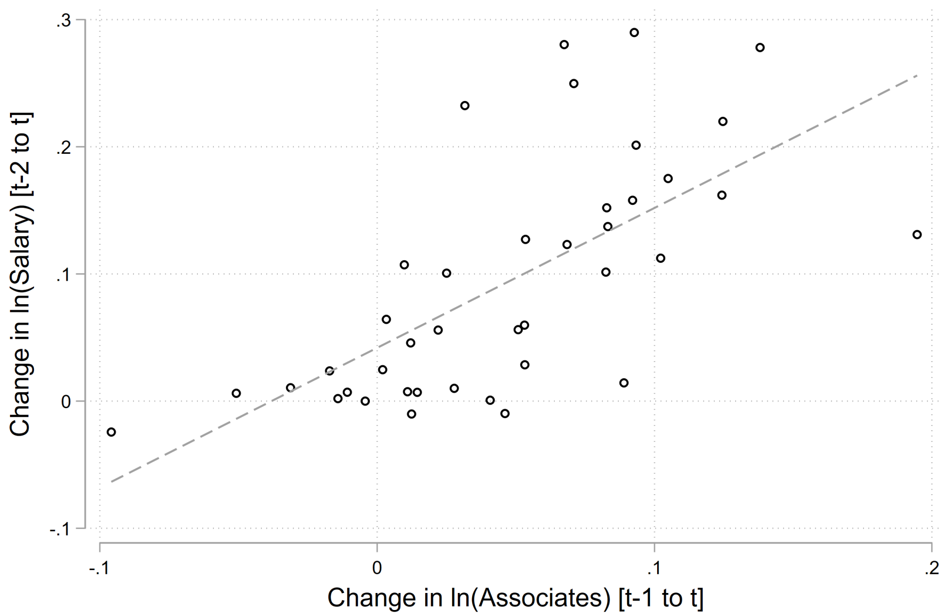
Note: Figures are conditional on firm's reporting starting salaries in NLJ surveys. Starting salaries represent the highest reported starting salary for each firm.

Figure 1.5: Growth in real starting salaries and profit per equity partner (NLJ 200)



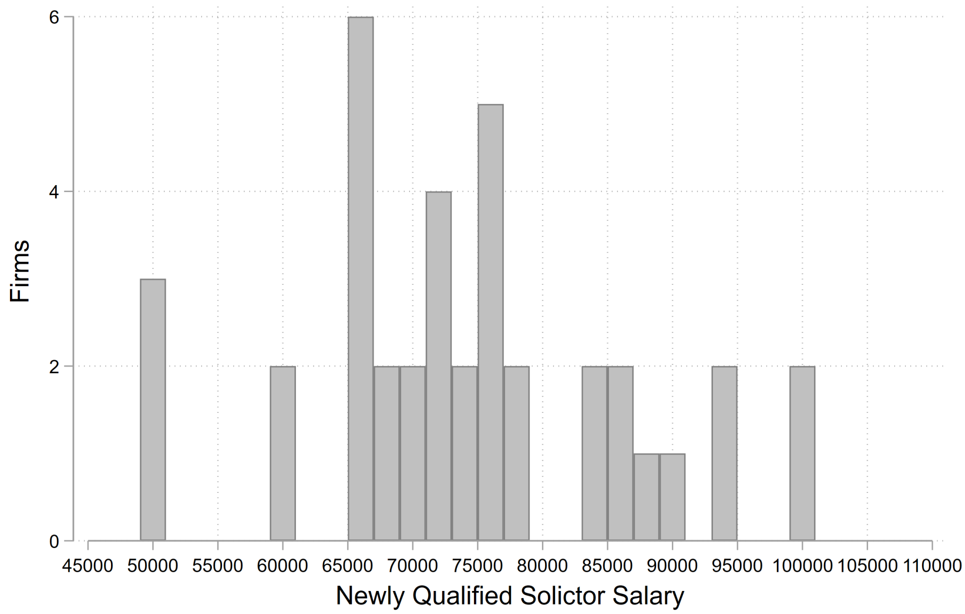
Note: Figures are conditional on firm's reporting starting salaries in NLJ surveys and featuring in the AMLAW100. Starting salaries represent the highest reported starting salary for each firm. Values are weighted by the total number of employed attorneys. All values are adjusted to 2020 dollars using the CPI and indexed to 1986 levels.

Figure 1.6: Growth by year in log of starting salaries versus log of associate employment (NLJ 200)



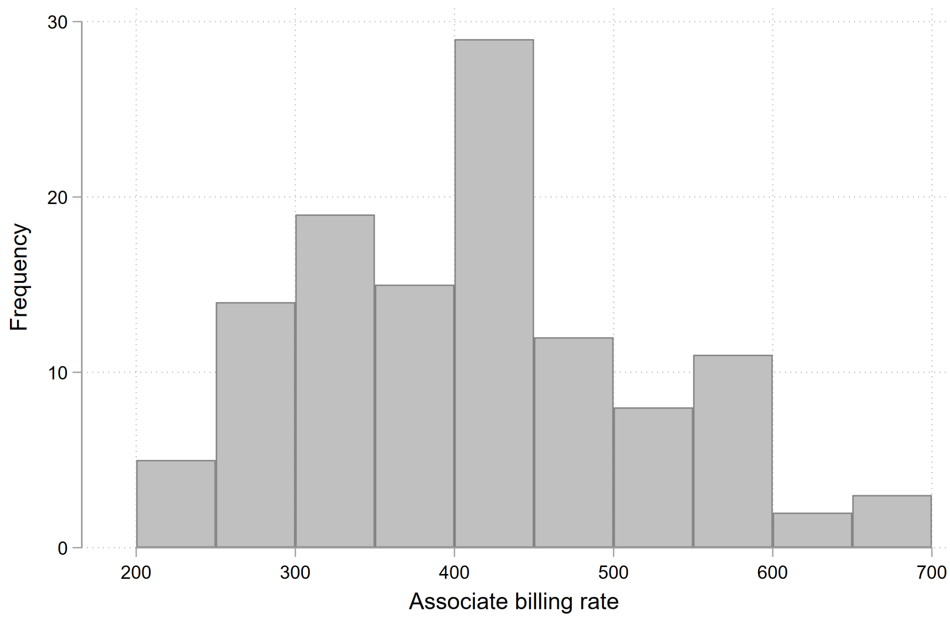
Note: Each point represents average growth for one year, created by weighting firm-level growth by lagged associate employment. For years 1980-2020.

Figure 1.7: Distribution of newly qualified solictor salaries in London (2020)



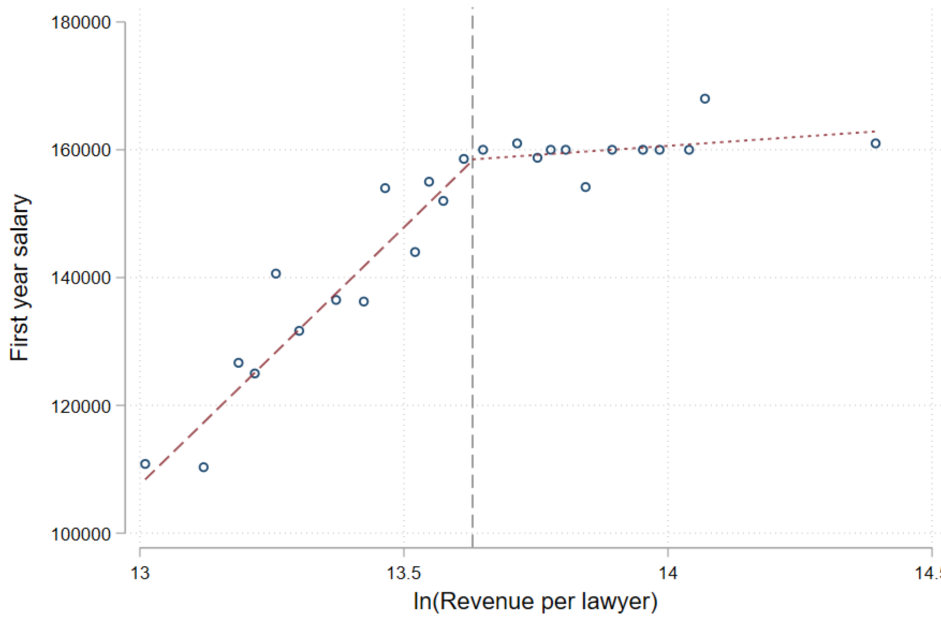
Note: For the Top 75 UK firms by revenue with available salary data. Observations are binned in 2000 increments. Salaries are in nominal British Pounds.

Figure 1.8: Distribution of minimum associate billing rates (2014)



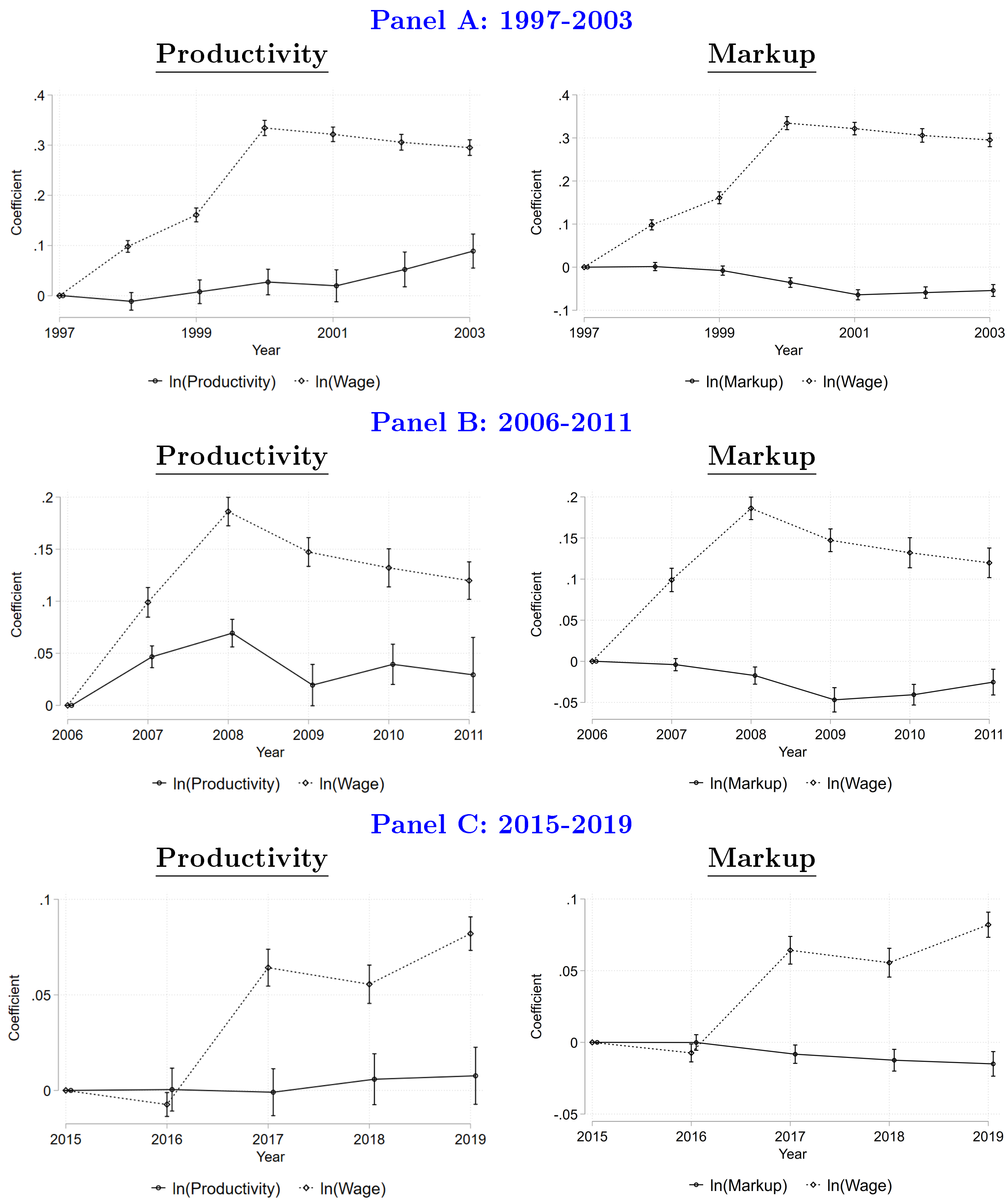
Note: For NLJ200 firms with billing rate data.

Figure 1.9: Binned scatterplot of starting salaries versus log of revenue per lawyer (NLJ 200, 2016)



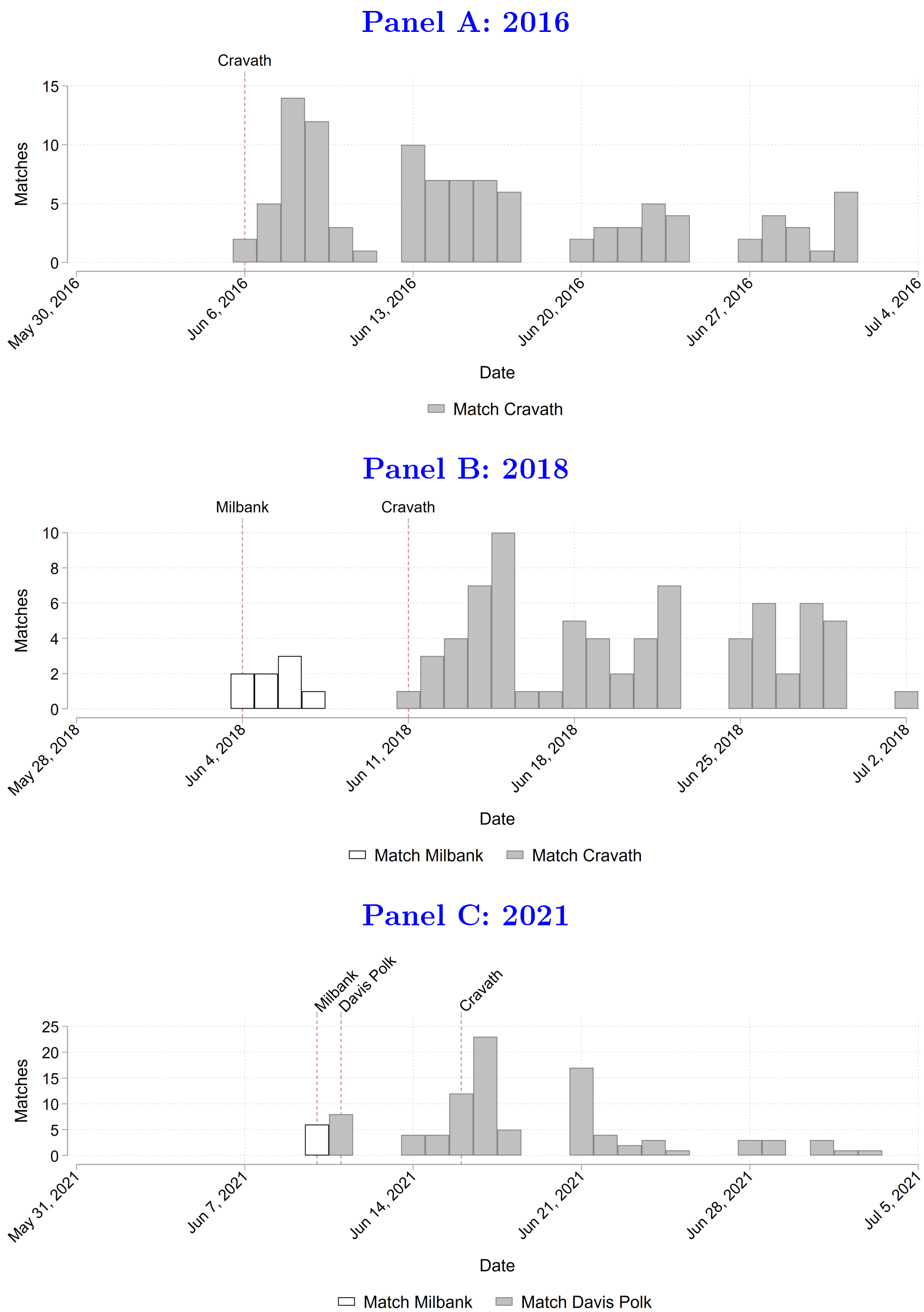
Note: For NLJ 200 firms with reported starting salaries and AMLAW revenue data. 2016 data is used since it is a year prior to 2017 salary increases. Salaries are in nominal dollars.

Figure 1.10: Salaries, productivity, and markups around major salary increases (NLJ 200)



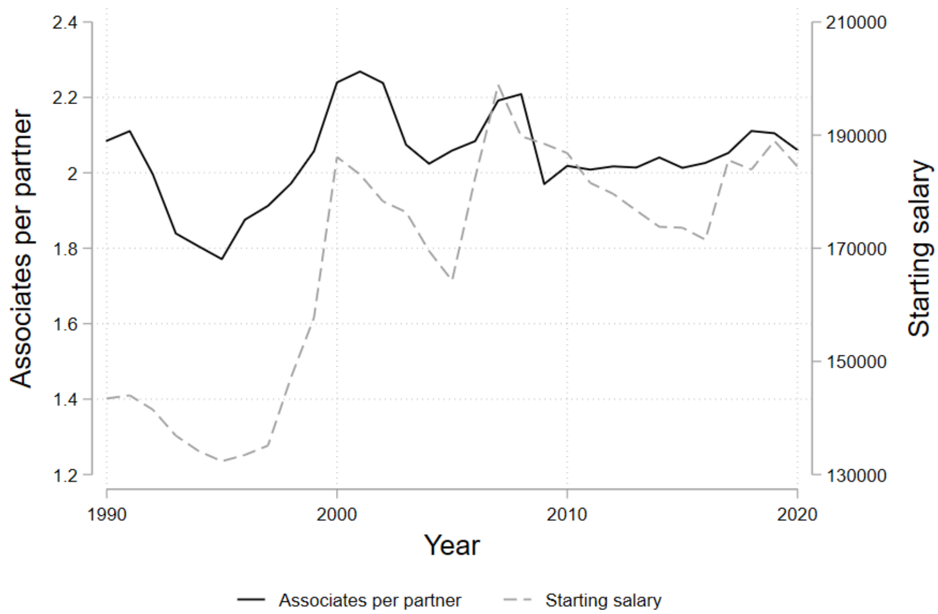
Note: For NLJ 200 firms with reported starting salaries and AMLAW data. Coefficients are for year fixed effects; the regression includes firm fixed effects. Regressions are weighted by initial attorney employment. Standard errors are clustered at the firm level. Bands represent 95% confidence intervals.

Figure 1.11: Announcement timings for salary increase matches



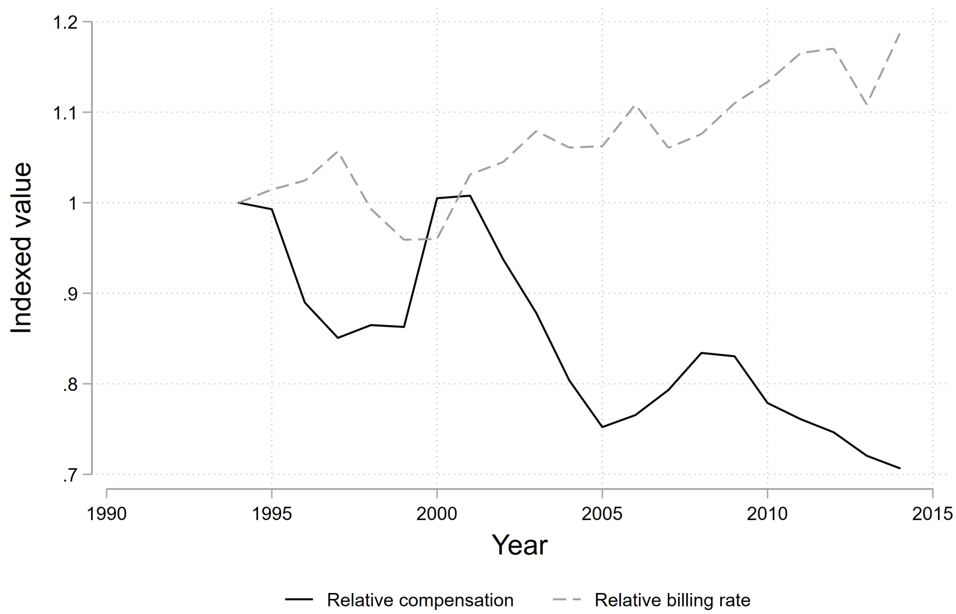
Note: For firms with reported matches on AboveTheLaw.com. In 2018, all firms that initially matched Milbank later matched Cravath's higher salary figures. In 2021, all firms that initially matched Milbank later matched Davis Polk's higher salary figures.

Figure 1.12: Ratio of associates to partners and starting salaries (NLJ 50)



Note: Salaries are in 2020 dollars

Figure 1.13: Ratio of associate to partner compensation and billing rates (NLJ 200)



Note: For firms reporting billing rates. Rates are indexed to 1994. Billing rates are based on minimum billing rates for associates and partners.

Table 1.1: Salary increases by year (NLJ200, $1000s)

	Mean Salary	Median Salary	Modal Salary	Mean Change	Median Change	Share Increasing	Reported Salaries
1979	25	25	30	2.6	2.0	92.2	130
1980	28	28	28	3.4	3.0	89.9	137
1981	32	31	30	3.7	4.0	92.7	162
1982	34	34	33	2.7	3.0	86.1	133
1983	37	35	35	1.5	1.0	67.8	165
1984	39	38	37	2.2	2.0	89.6	155
1985	42	40	40	2.7	2.5	87.8	159
1986	50	50	50	8.0	8.0	89.5	183
1987	55	53	65	4.9	4.3	73.1	167
1988	60	58	50	6.0	6.0	89.9	168
1989	64	62	60	3.4	4.0	78.7	196
1990	67	66	70	3.0	3.0	72.5	195
1991	69	68	70	1.5	0.0	36.1	195
1992	69	69	70	0.4	0.0	14.4	191
1993	69	68	83	0.3	0.0	22.6	194
1994	69	70	83	0.6	0.0	23.3	197
1995	71	70	70	1.2	0.0	41.8	196
1996	72	73	85	1.5	0.0	43.7	186
1997	75	75	87	2.6	2.0	67.0	187
1998	84	86	90	8.4	8.0	91.0	185
1999	92	95	90	8.0	7.0	86.8	180
2000	113	125	125	21.2	22.0	96.6	185
2001	114	125	125	1.9	0.0	24.7	188
2002	115	125	125	-0.6	0.0	9.8	192
2003	116	125	125	1.8	0.0	19.7	192
2004	116	125	125	0.6	0.0	15.2	187
2005	117	125	125	1.4	0.0	24.1	172
2006	130	135	145	13.9	12.0	91.3	175
2007	146	160	160	15.8	15.0	92.2	177
2008	148	160	160	2.2	0.0	19.0	180
2009	144	160	160	-4.6	0.0	2.6	155
2010	143	145	160	0.0	0.0	10.1	172
2011	144	150	160	0.1	0.0	6.9	162
2012	-	-	-	-	-	-	-
2013	144	160	160	-	-	-	154
2014	145	160	160	0.5	0.0	12.2	152
2015	145	160	160	1.3	0.0	12.4	148
2016	146	160	160	1.2	0.0	17.8	138
2017	159	180	180	13.7	20.0	76.9	139
2018	161	180	180	2.4	0.0	20.1	139
2019	166	180	190	6.3	10.0	63.6	139
2020	170	190	190	2.6	0.0	21.9	132

Note: Sample is NLJ200 firms with reported salaries. Salaries are the highest reported salaries for new associates. Salaries are in nominal dollars ($1000s). NLJ salary data for 2012 is not available. Reported salaries is the number of NLJ200 firms with non-missing salary data.

Table 1.2: Local and NY modal starting salary matching rates by headquarters city (NLJ200, $1000s)

City		1980-84	1985-89	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
NYC	Match city mode	0.35	0.41	0.58	0.25	0.82	0.83	0.80	0.81
	Match NY mode	0.35	0.41	0.58	0.25	0.82	0.83	0.80	0.81
	Firm-year obs	168	207	216	208	195	181	169	144
D.C.	Match city mode	0.36	0.34	0.30	0.28	0.76	0.78	0.67	0.82
	Match NY mode	0.00	0.00	0.01	0.03	0.76	0.75	0.67	0.82
	Firm-year obs	53	67	98	104	113	99	85	65
Chicago	Match city mode	0.34	0.42	0.56	0.41	0.78	0.75	0.68	0.70
	Match NY mode	0.01	0.02	0.11	0.07	0.78	0.72	0.68	0.70
	Firm-year obs	96	96	117	97	86	75	69	71
Philadelphia	Match city mode	0.35	0.64	0.28	0.29	0.32	0.40	0.38	0.39
	Match NY mode	0.00	0.00	0.06	0.04	0.26	0.28	0.29	0.33
	Firm-year obs	26	42	54	56	57	50	48	49
Boston	Match city mode	0.60	0.56	0.49	0.60	0.67	0.81	0.79	0.91
	Match NY mode	0.00	0.06	0.06	0.00	0.65	0.81	0.79	0.91
	Firm-year obs	43	34	49	43	46	37	38	34
Los Angeles	Match city mode	0.57	0.33	0.26	0.32	0.83	0.84	0.71	0.73
	Match NY mode	0.00	0.00	0.12	0.06	0.83	0.84	0.71	0.73
	Firm-year obs	49	51	42	34	41	44	38	37
San Francisco	Match city mode	0.44	0.52	0.25	0.38	0.72	0.60	0.75	0.88
	Match NY mode	0.00	0.00	0.02	0.06	0.72	0.60	0.75	0.88
	Firm-year obs	27	50	65	53	54	42	28	26
Dallas	Match city mode	0.39	0.26	0.26	0.29	0.29	0.57	0.65	0.70
	Match NY mode	0.00	0.00	0.00	0.00	0.13	0.35	0.65	0.70
	Firm-year obs	18	35	42	35	38	23	23	23
Houston	Match city mode	0.57	0.32	0.22	0.35	0.40	0.80	0.96	0.92
	Match NY mode	0.00	0.00	0.09	0.16	0.33	0.80	0.96	0.83
	Firm-year obs	28	38	32	31	30	30	28	24
Atlanta	Match city mode	0.81	0.72	0.65	0.50	0.42	0.54	0.44	0.39
	Match NY mode	0.00	0.00	0.00	0.00	0.42	0.54	0.40	0.39
	Firm-year obs	26	29	26	30	31	28	25	28

Note: Sample is NLJ200 firms with reported starting salaries. Firms are classified based on the city of their largest office. Matching rates are the share of firm-year observations (within a five-year band) that match the modal salary within the city-year.

Table 1.3: Firm expansion by year (NLJ200)

Year	Avg. # of Attorneys	Share in Home State	Avg. # of Branches	Avg. # of Large Branches	Avg. Branch Size	Any Large NY Office	Any Intl. Branch
1978	102	0.94	1.5	0.1	7.4	0.26	0.19
1980	119	0.92	2.0	0.1	7.4	0.26	0.23
1982	144	0.89	2.6	0.3	9.4	0.26	0.25
1984	178	0.89	2.9	0.4	10.2	0.25	0.25
1986	216	0.86	3.3	0.7	14.3	0.28	0.27
1988	260	0.84	4.0	1.0	17.6	0.29	0.28
1990	302	0.81	4.6	1.4	20.9	0.32	0.32
1992	299	0.79	5.1	1.5	19.8	0.31	0.43
1994	299	0.78	5.7	1.6	19.4	0.34	0.43
1996	322	0.77	6.0	1.7	20.7	0.37	0.48
1998	365	0.74	6.6	2.2	23.9	0.37	0.48
2000	438	0.71	7.2	2.8	28.8	0.42	0.48
2002	497	0.68	8.0	3.3	31.9	0.46	0.50
2004	514	0.66	8.6	3.5	31.3	0.48	0.51
2006	560	0.64	9.3	4.0	32.7	0.54	0.48
2008	625	0.60	10.6	4.7	33.1	0.58	0.54
2010	586	0.59	11.3	4.6	29.7	0.57	0.51
2012	587	0.57	12.0	4.7	28.6	0.57	0.54
2014	622	0.55	13.5	5.3	28.0	0.59	0.56
2016	622	0.54	14.1	5.4	27.5	0.60	0.56
2018	645	0.52	14.7	5.6	28.0	0.61	0.55
2020	684	0.51	15.7	5.9	28.8	0.62	0.54

Note: Home state is defined as the state containing the largest share of the firm's lawyers. Each branch is a city with an office (excluding the city containing the firm's headquarters). A large office or branch contains at least 25 lawyers. Any large NY office is the share of firms with a large NY office (includes if the firm's headquarters is in New York). Data is from NLJ reports.

Table 1.4: Value of $190K pre-tax salary across cities in 2019

	$190K After-Tax Salary ($1000s)	ln(After-Tax Salary)	Relative to NYC Amenities	Local Prices	Total Compensation	NYC Pre-Tax Equivalent ($1000s)
NYC	129	0.00	0.00	0.00	0.00	190
D.C.	127	-0.01	0.22	-0.03	0.24	243
Chicago	131	0.01	-0.05	-0.08	0.05	200
Philadelphia	134	0.04	-0.49	-0.15	-0.30	141
Boston	131	0.02	0.16	0.04	0.13	217
Los Angeles	129	0.00	0.28	-0.04	0.32	264
San Francisco	129	0.00	0.16	0.15	0.01	192
Dallas	140	0.08	-0.05	-0.25	0.28	253
Houston	140	0.08	-0.35	-0.32	0.05	200
Atlanta	130	0.01	0.12	-0.14	0.27	250

Note: After tax salary calculated using NBER TAXSIM calculator for unmarried workers with no other income. Data on amenity and local prices adustments comes from Diamond (2016) estimates for college educated workers in 2000.

Table 1.5: Bonus announcements and leaders

Bonus	First	Date	Standard	Date	Description
2021 Spring	Willkie Farr	3/19	Davis Polk	3/22	\$7,500 → \$12,000
2020 Year End	Baker McKenzie	11/11	Cravath	11/23	\$15,000 → \$22,500
2020 Covid	Cooley	9/14	Davis Polk	9/15	\$2,500 → \$7,500
2019 Year End	Milbank	11/7	-	-	\$15,000; No other matches until Cravath 11/11
2018 Year End	Cravath	11/19	-	-	\$15,000
2017 Year End	Cravath	11/27	-	-	\$15,000
2016 Year End	Cravath	11/27	-	-	\$15,000
2015 Year End	Cravath	12/7	-	-	\$15,000
2014 Year End	Simpson Thacher	11/21	Davis Polk	11/25	\$15,000; Davis Polk raised for older classes
2013 Year End	Cravath	12/2	-	-	\$10,000
2012 Year End	Cravath	11/26	-	-	\$10,000
2011 Year End	Cravath	11/28	-	-	\$7,500
2011 Spring	Sullivan & Cromwell	1/21	Cravath	1/31	\$2,500; Cravath raised for older classes
2010 Year End	Cravath	11/22	-	-	\$7,500
2009 Year End	Cravath	11/2	-	-	\$7,500
2008 Year End	Skadden	11/19	Cravath	11/20	\$35,000 →\$17,500
2007 Year End	Cravath	10/29	-	-	\$35,000

Notes: For major law firms. Standard is the firm whose bonus scale is most commonly matched (if it is a different firm from the first firm). Bonus amounts are for first-year associates. Second listed bonus amount is from the eventual standard. Data from AboveTheLaw.com announcements.

Table 1.6: Example firms exceeding market compensation (2012)

Firm	Lawyers	Salary	Bonuses	Notes
Cravath	**453**	**$160,000**	**$10,000**	
Bickel & Brewer	43	$185,000	N/A	Litigation
Boies Schiller & Flexner	258	$174,000	$25,000+	Litigation
Desmarais	60	$180,000	N/A	Litigation
McKool Smith	180	$165,000	$12,500	Litigation
Susman Godfrey	90	$170,000	$40,000+	Litigation
Wachtell Lipton	249	$165,000	Above market	M&A
Williams & Connolly	271	$180,000	$0	Litigation

Notes: Data is for 2012 first-year associates. Data from AboveTheLaw.com article.

Table 1.7: Relationship between change in log of associates (2005-2008) and firm constraints

	(1) Profit	(2) Leverage	(3) Both
ln(Profit per partner) in 2005	0.0809**		0.104***
	(0.0388)		(0.0364)
Change in ln(Associates/Partners) 2002-05		-0.221**	-0.309***
		(0.109)	(0.113)
Observations	79	79	79
R-squared	0.049	0.032	0.108

Notes: For NLJ 200 firms with data on revenue from Amlaw that also matched the 2000 salary increase. Regressions are weighted by employment in 2002. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

CHAPTER 2

Discrimination, the Racial Wage Gap and the Schooling of the Next Generation: Evidence from WWII

2.1 Introduction

Racial disparities in labor market outcomes in the U.S. are persistent and pervasive. Today the unemployment rate of Black men is roughly double that of white men, and among the employed, Black men earn 27% lower wages. The racial gaps in wages have contracted little since 1980; however, there were two previous periods of rapid convergence: the 1940s and the 1960s.[1] These two periods of rapid convergence in adult wages by race also witnessed significant gains in the educational achievement of Black youth (see Figure 2.1).

This paper focuses on the period of the 1940s and examines the causes and consequences of the decline in the racial wage gap over this period. We focus on the role played by government contracts to private firms as part of the war production effort.[2] The federally funded domestic war production effort, totaling \$3.1 trillion (2014\$), combined significant increases in labor demand among private firms with federal requirements of anti-discriminatory policies in hiring in all firms with war contracts. We exploit variation across local labor markets in the allocation of war contracts to identify the effect of the war production effort on the wages of workers by race 1940-1970, as well as spillover effects on the educational attainment of the

[1] For a detailed analysis of long-run trends in Black-white wage gaps see Bayer and Charles (2018).

[2] For the 1960s, researchers have attributed the shrinking of the gap to economic expansion, increasing educational attainment among Black individuals (Smith and Welch, 1989; Card and Krueger, 1992), affirmative action and the rise of anti-discrimination policies (e.g., Donohue and Heckman, 1991; Miller, 2017), and the rise in the minimum wage (Derenoncourt and Montialoux, 2019).

next generation.

We find that in areas with more war contract spending, the wages of Black workers rose nine log points more than in areas with fewer war contracts; this finding is unaffected by controlling for the local area draft rate. Importantly, the gains for Black workers persist until at least 1970, with only modest fading. By contrast, we find no (negative or positive) effects of war contract expenditures on the average outcomes of white men or of women of any race by 1950. The null effects for women are consistent with previous work showing that women's gains in the labor market during WWII as a result of mobilization were temporary (Acemoglu, Autor and Lyle, 2004; Goldin, 1991). Our estimated effects are not driven by endogenous allocation of war contracts or more skilled Black workers moving to areas with more contracts. But we do document that WWII contracts attracted a large number of Black workers to the receiving areas.

Two exercises provide evidence that the effects we estimate are driven by the interaction between the increase in labor demand caused by the war contracts and the requirements of non-discrimination. First, we show that the war contracts improve Black worker outcomes only in areas with tight labor markets during the war. We interpret this as evidence that without labor market shortages, requirements of non-discrimination were insufficient to improve long-run labor-market outcomes for Black workers. Second, we take advantage of the fact that the war contracts created demand for inputs that were supplied by firms that were not required to hire in a non-discriminatory manner. Using the input-output table, we find that the improvement in Black labor market outcomes occurred in markets where government contracts raised demand only if the requirements of non-discrimination applied. We interpret this as evidence that without the requirements of non-discrimination, labor market demand alone was insufficient to improve long-run labor-market outcomes for Black workers, consistent with (Collins, 2001). Finally, we provide some evidence that the effects of these policies persisted because they resulted in an institutional change that resulted in lower discrimination in the labor market after the war: The desegregation of unions.

We then consider the consequences of these policies with respect to intergenerational spillovers. We find that high school graduation rates increased more over this period for Black children, particularly for boys, in areas with greater war contract expenditures. Previous work on racial gaps in education has typically focused on changes in school access during this period. The quality and quantity of schooling increased substantially for Black children in the South (Card and Krueger, 1992; Collins and Margo, 2003; Aaronson and Mazumder, 2011; Carruthers and Wanamaker, 2017). Turner and Bound (2003) show that the GI Bill (created for WWII veterans) also improved Black educational levels. We document another reason why the education level of Black children increased in the post-war years: improved labor market outcomes of their parents.[3] We show that war expenditures did not affect the returns to education, nor did they change school expenditure levels or residential segregation in affected cities. Thus the most plausible mechanism for the increases in education we document is the increase in family incomes associated with declines in labor market discrimination. This is consistent with the large body of literature that shows that the single most important determinant of educational outcomes is parental socioeconomic status (Duncan and Magnuson, 2012).

WWII defense contracts and the associated requirement of nondiscrimination significantly improved the labor market outcomes of Black workers and the educational attainment of their offspring. But how much of the aggregate improvements for Black workers and families observed during this period can the WWII defense effort explain? How much of these improvements can be ascribed to changes in discrimination? And finally, how important is migration for driving these aggregate impacts? To answer these questions, we develop a general equilibrium model building on the canonical macroeconomic model of discrimination in Hsieh et al. (2019). Discrimination is a wedge between the wage an employer pays a Black worker and the actual value of the marginal product of that worker (potentially caused by incorrect statistical discrimination, overt racism, etc.), which we allow to vary freely across regions, occupations, and industries. By appropriately extending Hsieh

[3]Although Margo (1993) hypothesized that improving labor market outcomes for Black workers during the war ultimately improved the education of Black children, to our knowledge we are the first to document that war expenditures raised Black high school graduation rates.

et al. (2019) to include regions, (defined as metropolitan areas or commuting zones), in addition to industries, migration, and trade, we are able to leverage exogenous variation in expenditures across regions to identify the impact of war production on discrimination. Based on our model, we find that WWII defense production generated substantial improvements in national labor-market outcomes for Black workers, with no negative effects for white workers (e.g., explaining 25 percent of the reduction in the racial wage gap observed over this period), that migration played an important role in amplifying these otherwise local effects (explaining almost a third of this decline), and that the majority of the aggregate effects arise from reductions in discrimination.

Our work makes three important contributions to our understanding of the causes and consequences of significant gains in Black worker wages during the 1940s. First, it builds upon existing research examining the factors responsible for those gains by identifying the underlying causes and mechanisms. It was already established that skill upgrading (Margo, 1995) and (Collins, 2000) and Black worker migration from the South to the North over this period (Boustan, 2009) were important drivers of the closing of the racial wage gap over this period. In our work, we are able to identify and quantify two important causes behind both: Increased demand for labor from expansions in domestic production, which generated tight labor markets, particularly in the North, combined with requirements of non-discrimination in hiring. This is consistent with the work of Collins (2001), showing that the executive orders limiting discrimination resulted in greater employment of Black workers in defense industries, and of Ferrara (2020), showing that reductions in labor supply due to mortality from WWII increased skill-upgrading among Black workers. Our work extends theirs by showing that neither alone was sufficient but that the combination of tight labor markets and requirements of non-discrimination was needed for skill-upgrading and wage gains. Moreover, by focusing on increases in labor demand, not supply, we identify a factor that can be more easily influenced by policy and therefore replicated in other settings.

Our second contribution is to show that declines in discrimination are responsible for the lasting effects of the domestic war production effort on the skill-upgrading and wages of

Black workers. Our work is the first attempt to link well-identified evidence of declines in labor market discrimination with changes in aggregate racial gaps using a structural model, as suggested by Lang and Spitzer (2020). Our final and perhaps most important contribution is to document the persistent effects of these policies on the labor market outcomes of adults and the human capital of their children. While Margo (1995); Smith and Welch (1989) showed that educational gains by Black workers played a role in explaining the closing of the racial gap in wages in the 1940s, we provide evidence of the reverse relationship as well: That the closing of the racial gap in adult wages increased the subsequent educational attainment of Black children.

These results have important implications for current economic and social policy. Our findings suggest that tight labor markets and consistent enforcement of rules against discrimination are likely to improve labor market outcomes of Black and other minority workers in the short run, with persistent effects generated from further declines in discrimination. Moreover, reductions in current labor market disparities will generate spillover effects, reducing gaps in the educational attainment and future earnings of the next generation. Efforts to address racial gaps in educational achievement should therefore consider policies that reduce racial disparities in the labor market.

2.2 Background

2.2.1 Labor market conditions for Black men in 1940

In 1940, despite similar overall rates of employment, prime-age Black men earned half as much yearly income as their white counterparts. The source of this difference was not primarily due to differential pay within occupation or job, as many pre-war surveys found that Black and white men in the same jobs within the same firms typically earned similar wages (Billips, 1936; Frazier and Perlman, 1939). Rather, the main source of the difference in earnings was the disproportionate concentration of Black workers in lower-paid industries and, within those industries, lower-paid occupations (Wright, 1986). In 1940, 78.6% percent of Black men were employed in unskilled occupations or as farmers, compared to only 37.5% of white men.[4]

Locational or educational differences cannot fully explain this occupational segregation. In Appendix Section B.2.2, we conduct an exercise where we compare the actual distribution of Black workers across occupations to the expected distribution if workers were randomly allocated to jobs conditional on education and location. Following Margo (1995), we adjust reported years of education downward for Black men born in the South to account for school quality differences. Black workers were significantly underrepresented in occupations such as engineers, salesmen, managers, skilled blue-collar workers, and foremen.[5] Conversely, we would expect about 45,000 Black janitors and porters in metropolitan areas conditional on education and location. Instead we observe 145,000.[6]

Overt discrimination of firm owners played a role in the occupational segregation of Black workers. At the onset of WWII, 51% of war manufacturers reported they did not and would not employ Black workers.[7] Unions also discriminated, with Black men barred from joining many unions or forced into segregated "Jim Crow" locals which prevented them from obtaining jobs in many skilled blue collar professions. AFL-affiliated craft unions, such as the Machinists' union, were especially known for these policies.[8] Limits on promotion

[4]Unskilled occupations are defined as all occupations falling under "Laborers," "Farm Laborers," or "Service workers" occupational categories in the 1950 Census Bureau classifications. All occupations which are not unskilled or farmers are labeled as skilled for our purposes. Because our focus is on metro areas, whether or not we include farmers as skilled occupations does not meaningfully change our results. See Appendix Table B.1 for occupational shares by category.

[5]Appendix Table B.3 shows the most over and underrepresented occupations. We use five education groups and multiply years of education by 0.85 for Black men born in the South with less than 15 years of education to account for school quality differences. Appendix Figure B.2 shows average wages and education by occupation; most occupations with excess Black men are occupations with the lowest average wage and education levels.

[6]With only 1.27 million Black men employed in metro areas, the excess number of janitors and porters alone is about 8% of all employed Black men. Conditioning on education could understate segregation given that even jobs with more Black men purely due to segregation will also tend to have lower average educational levels since Black men have less education on average.

[7]President's Committee on Fair Employment Practice. "First Report, July 1943-December 1944." Washington, D.C., 1945.

[8]See Appendix Table B.2 for example unions with discriminatory membership policies.

affected hiring into entry-level positions: firms had less incentive to hire Black workers since they could not be promoted even if they showed talent and Black men had less incentive to provide effort since they could not be promoted (Sundstrom, 1994). All evidence suggests that widespread discrimination and occupational segregation characterized the labor market for Black men on the eve of WWII.

2.2.2 The effects of WWII on the labor market

WWII transformed the American labor market in three main respects. First, the dramatic increase in federal expenditures greatly increased labor demand. The U.S. spent roughly $3.1 trillion (2014 $) on war-related production, roughly 40 percent of GDP each year in 1943, 1944, and 1945, creating the largest increase in expenditure in U.S. history (Appendix Figure B.3). This was four times larger than "New Deal" expenditures meant to alleviate the Great Depression (Fishback and Kachanovskaya, 2015). Military equipment contracts accounted for 85% of this spending; new production facilities accounted for the rest. The Stabilization Act of 1942 limited changes to prices, wages, and salary levels during the war to prevent inflation.

Second, military enlistments dramatically decreased labor supply. About 15.8 million working-age men—equivalent to 40% of the male labor force in 1940—served in the military during WWII.[9] Additionally, about half a million men died during the war, permanently reducing the labor force by 1.3 percent. Previous work has shown that mobilization resulted in an increase in female labor force participation (Goldin, 1991; Acemoglu, Autor and Lyle, 2004; Rose, 2018). Large increases in labor demand and decreases in labor supply led to large labor shortages in many industries and cities.[10]

Third, the government enacted several important anti-discrimination measures to ensure maximum labor force utilization. President Roosevelt issued Executive Order 8802 (1941)

[9] Although some 350,000 women served in the military, they accounted for a small fraction of the total.

[10] We show later in Table 2.1 that higher war expenditures and enlistment rates are associated with labor shortages

asserting: "I do hereby reaffirm the policy of the United States that there shall be no discrimination in the employment of workers in defense industries or government because of race, creed, color, or national origin." The order also established the Committee on Fair Employment Practice (FEPC) to encourage industries receiving government contracts to hire minorities. Executive Order 9346 (1943) made the FEPC an independent commission and established regional offices, increasing its reach. The FEPC had little power to directly punish violating companies, but it could publicly shame employers and provide advice on integration. Collins (2001) shows that the share of Black workers employed in defense industries increased in places with more FEPC intervention, suggesting that the executive orders were effective.[11]

We focus on the effects of federal war contracts with private firms to produce all goods related to the war effort. War contracts were allocated across the U.S. primarily based on existing industrial capacity and not based on political considerations (Rhode, Jr. and Stumpf (2018)) nor were they targeted to places with more available labor (Brunet (2018)). Moreover, despite its scope and scale, WWII spending did not significantly affect local per capita economic development (Fishback and Cullen (2013), Li and Koustas (2019), Brunet (2018), Jaworski (2017), Lewis (2007)), though it did increase local populations. Brunet (2018) finds a small state-level fiscal multiplier for WWII expenditures of 0.25-0.30. Garin (2019) finds persistent positive effects of large new manufacturing plants when located in smaller communities. Building on this work, we investigate how WWII contracts to existing private firms affected the occupational upgrading and wages of Black workers. Importantly, we measure these effects up to 25 years after the conclusion of the war.

[11] Additionally, some firms receiving defense contracts also received some form of management training. Bianchi and Giorcelli (2020) find that human resource practice training had the largest effect on firm performance after the war.

2.3 Data and empirical approach

2.3.1 Data

WWII contract expenditures. Data on war contract expenditures by county come from the War Production Board's Major War Supply Contracts and Major War Facilities Projects.[12] War contracts (excluding food and food processing) worth over $50,000 are assigned to the county of the primary production plant. The data covers all contracts awarded from June 1940 until September 1945. Electrical machinery, transportation, automotive, and iron/steel production account for 61 percent of expenditures.

War contracts were disbursed to locations that already produced these goods. Figure 2.2 shows the spatial distribution of WWII expenditures. War contracts were less likely to be distributed to the South and more likely to be distributed to the Northeast, Midwest, and West coast. These expenditures typically went to urban rather than rural areas.[13] In our baseline analysis we focus on metropolitan areas, aggregating spending up across counties and calculating cumulative spending per capita for each metropolitan area. In 1950 these 146 metro areas covered 55 percent of the U.S. population, and 50 percent of the Black population. More than 90 percent of the Black population living outside of the South lived in metro areas. In our robustness checks we replicate our results using two alternative geographic aggregations (states and commuting zones) so as to include the entire population.

The average war contract spending per capita in a given metropolitan area was about $1,830 per person in 1940 dollars, with a standard deviation $1,715 across metro areas. For comparison, GDP per capita in 1940 was only $779. Figure 2.3 shows a very skewed distribution of expenditures. All metro areas received at least some war contracts but there is significant variation in size: 50% received less than $1,280 per capita, while almost 10%

[12]These data are available from the 1947 County Data Book, available through ICPSR 02896 (Haines and ICPSR, 2010).

[13]Metro areas are based on Census definitions. They consist of groups of counties, and the primary qualification is that the county grouping contains a city of at least 50,000 people.

received $4,000 or more per capita (Panel A). This substantial variation in expenditures persists even after we condition on 1940 city-level characteristics, including percent employed in manufacturing, percent Black, and the predicted enlistment rate (Panel B).

Enlistment rates. We construct metro-level enlistment rates using individual-level data from the WWII Army Enlistment Records. Unfortunately these records include information only on 9 million of the 16 million individuals who served.[14] We supplement the data by digitizing select tables from Selective Service System (1956), which contain information on total voluntary and draft enlistments by month and service branch.

There are several complications with interpreting observed enlistment rates. First, voluntary enlistments will be highly endogenous to local conditions. Second, draft regulations provided various exemptions for individuals, such as deferments based on age or marital status. Most importantly for our work, men who worked in industries deemed essential for the war effort could obtain exemptions. Third, areas with net positive in-migration after 1940 (e.g., due to war expenditures) will appear to have higher enlistment rates. Fourth, individual records are missing in batches related to geographic areas. To address these measurement and endogeneity issues, we construct a predicted enlistment rate to isolate the major exogenous source due to demographic characteristics of the local population pre-war (in 1940). See Appendix B.1.2 for details. Reassuringly, our results are robust to controlling for either the observed or the predicted draft rate.

Labor market and education data. Our primary outcomes of interest come from 1920-1970 individual-level census data from IPUMs (Ruggles et al., 2020) aggregated to the race-sex-metro-year level. The individual data contains information on occupation and school enrollment for all census years. Employment is available starting in 1930. Education and wage earnings are reported starting in 1940. In the 1950 1% sample, some of these outcomes were asked only of sample-line persons and are available for a small share of the population.

[14]The primary reason for missing data is that the files contains only Army records and excludes other branches, such as the Navy (the Air Force was still part of the Army during WWII). The secondary reason for missing service members is poor quality scans or missing records.

To improve metro-level variables derived from the smaller 1950 census data, we digitized metro-level aggregates from the 1950 Census Volumes on occupation and income distributions by race and gender.

We define skilled occupations as individuals reporting an occupation falling under "Profession, Technical"; "Managers, Officials, and Proprietors"; "Clerical and Kindred"; "Sales workers"; "Craftsmen"; or "Operatives" categories.[15] We refer to the share of employed persons in these occupations as the "share skilled" throughout the paper. Prime-age employment is defined as the percentage of men ages 25-54 who are currently employed. The yearly wage is total wage earnings in the previous year for people who were wage-earning employees at the time of the Census.[16]

Other data. We digitized reports on the extent of labor shortages during WWII from the monthly Labor Market Reports compiled by the War Manpower Commission (1945). Data on employment in war-related industries during WWII comes from ES-270 reports. Data on local education expenditures at the city level collected in County and City Data Books and available through the ICPSR. Public expenditures on education at the city level are available for years 1940 and 1947-48 for cities with a population of at least 25,000.[17] Finally, we use neighborhood segregation measures at the city level from Cutler, Glaeser and Vigdor (1999). More details on the data sources are provided in the Appendix Section B.1.3.

[15]Skilled occupations include occupational codes 000-093, 200-690 under the 1950 IPUMS occupational coding scheme. Given our focus on metro areas, our results are not sensitive to whether farm owners are defined as skilled. Technically, some of these jobs could be considered "semi-skilled" jobs, but we grouped them all into a skilled category for conciseness throughout the paper. Occupational upgrading by Black workers was mainly from unskilled to semi-skilled jobs during this time.

[16]In 1940 individuals were only asked about wage income, so self-employed or business income is excluded. The majority of individuals excluded by the wage-earning employee restriction are farmers.

[17]These data are available through ICPSR: County and City Data Book [United States] Consolidated File: City Data, 1944-1977 (ICPSR 7735). Of the 146 metropolitan areas in our analysis sample, we have education data for approximately 83 percent of the sample residing in those metropolitan areas. The data include total expenditures for the city and do not distinguish between different schools within the city. We cannot, therefore, identify spending in schools attended predominantly by a specific race.

2.3.2 Understanding the determinants of WWII expenditures

To investigate which areas received more war contracts based on 1940 characteristics, we regress expenditures per capita in each of the 146 metro areas on 1940 characteristics and predicted enlistment rates in those areas. To facilitate comparison, we standardize all variables to have mean zero and a standard deviation of one. The results in column 1 Table 2.1 show that, as expected, manufacturing is the main determinant of war expenditures. Both the share employed in manufacturing and the (log of) manufacturing output per capita in 1940 are positive and statistically significant predictors of war production expenditures.

Importantly, other measures of economic activity, like the share unemployed or the share employed in skilled occupations, do not predict expenditures. Neither does the share of the population that is Black. But as expected, the draft rate negatively and significantly predicts expenditures.[18] Overall our findings are consistent with those reported previously in the literature. In our analysis, we address this non-random allocation of contracts in multiple ways, including extensive controls and location fixed effects as well as IV methods, as described below.

2.3.3 Empirical approach

We estimate the impact of war expenditures on the labor market outcomes of Black and white workers separately. We assess whether outcomes changed differentially between 1940 and 1950 in areas that received greater WWII expenditure (per capita) relative to those that received less, conditional on covariates. This strategy is a difference-in-difference approach where the treatment varies in intensity. Importantly, we measure outcomes in 1940 and 1950, five years after war production ceased. In so doing, we estimate the impact of the war production effort on the employment and wages of workers three to five years after the contracts ended and the requirements of nondiscrimination were rescinded. Specifically we

[18]As does the predicted share. See Appendix Table B.4. Individuals working in key war production sectors were more likely to be exempted from the draft. Additionally, reducing labor supply could have resulted in smaller contracts.

estimate the following equation separately by race:

$$Y_{rt} = \beta_1 WarExp_r \times Post_t + \beta_2 Draft_r \times Post_t + Post_t + \gamma_r + X_{rt}\rho + \varepsilon_{rt} \tag{2.1}$$

where the outcome of interest for a given metro area r in census year t is one of three measures: the share of workers employed in skilled occupations, (log of) the average wage, and the prime-age employment rate. Regressions are weighted by the population of the relevant race. The standard errors are corrected for heteroskedasticity.[19] The main independent variable of interest, $WarExp_r \times Post_t$, is the total cumulative war contract spending per capita in metro area r ($WarExp_r$) interacted with a post war indicator that equals one in 1950 ($Post_t$). We also include metro fixed effects (γ_r), the post-war indicator ($Post_t$), and metro area characteristics in 1940 interacted with a ($Post_t$) indicator (X_{rt}) described below.

WWII represents a large exogenous shock that differentially affects metropolitan areas based on the composition of their existing manufacturing base. We are essentially comparing places with similar manufacturing employment shares that differ based on how easily their manufacturing base could be converted into war production. The identification assumption is that conditional on manufacturing (and other baseline) covariates, the areas that received greater WWII expenditures would have been on the same trajectory as those receiving smaller amounts. This assumption would be violated if the pre-existing trends differ (if WWII expenditures went to places that were on different trajectories) or if areas that received higher expenditures were affected by other factors that are correlated with expenditures. We conduct several checks to confirm that our identification assumption is reasonable and also present estimates that use an instrumental variable approach. Specifically, we create a Bartik instrument based on the underlying geographic variation in the industrial composition across markets and the national industrial composition of the war contracts.

[19] Our results are the same if we cluster the standard errors at the metro level instead. Because we only have two time periods, one before and one after, adjusting for heteroskedasticity is appropriate (Bertrand, Duflo and Mullainathan, 2004).

2.4 Empirical results on wages, occupations, and employment

2.4.1 Effects of expenditures on labor markets during WWII

Local war contract expenditures led to labor shortages and to greater defense industry employment during the war. Column 2 of Table 2.1 shows that a standard deviation increase in war contract expenditures in a local metropolitan area is predicted to result in 0.469 standard deviations increase in the number of months of severe labor shortages, conditional on the predicted enlistment rate, which is also independently associated with labor shortages. As expected, greater war contract expenditures increased the employment of Black and white workers in the defense industry during the war; see Appendix Table B.5.

2.4.2 Short-term effect of expenditures 1940-1950

The large increase in expenditures that took place during the war decreased substantially and immediately upon the war's conclusion. In 1945 military expenditures amounted to 39 percent of GDP, falling to less than 10 percent by 1947.[20] Requirements of non-discrimination also ceased at the end of the war. In this section we estimate whether wartime spending had impacts on the labor market outcomes of Black and white workers that persisted after the war, to 1950.

Preliminary evidence of the effects of war expenditures is presented in Figure 2.4. Changes in the share of skilled employment are presented in Panel A for Black men and white men separately. The blue diamonds indicate the 1940-1950 changes. Metro areas that received higher expenditures saw a larger, positive and statistically significant increase in the share of Black workers employed in skilled occupations. The same is not true of white workers; the share of white workers employed in skilled occupations grew at the same rate in areas with more and less war expenditures. We also show the changes prior to the war from 1930-1940

[20] There was another increase in expenditures associated with the Korean War, but expenditures at their highest rose only to 15 percent of GDP in 1953 and the Korean War began after the 1950 Census. See Appendix Figure B.3.

(in dark circles). War expenditures were not associated with increases in the share employed in skilled occupations for Black workers prior to the war. This provides preliminary evidence for the validity of our identifying assumption of parallel pre-trends.

Panel B shows that expenditures were also associated with large and statistically significant increases in wages among Black workers from 1940 to 1950. White workers also appeared to benefit from expenditures, though the association for them is weaker. There is no Census wage question prior to 1940 to examine wage pre-trends.

Panel C shows that there was substantial Black (and to a lesser extent white) migration during this period related to WWII expenditures. While Black workers had started migrating North earlier in the century (after WWI), the figure shows that WWII expenditures redirect Black migrants towards cities with large war contracts. These cities were not receiving disproportionate numbers of migrants between 1930 and 1940, but saw very large increases in their Black populations in the 1940s.[21] For example Detroit, which received a very large share of WWII contracts and was known as "the arsenal of democracy," saw its non-white population rise from 150,790 in 1940 to 213,345 in 1944.[22]

These preliminary results are confirmed in Table 2.2, where we present the results from estimating equation (2.1). Regression analysis allows us to control for census-region-specific time trends and to weight by the relevant population of interest. We find positive and statistically significant effects of expenditures on the share employed in skilled occupations, on wages, and on the log of population for Black workers. These results hold even when we control for baseline characteristics, including the predicted draft rate and the manufacturing share, both of which predict expenditures (column 2). The effects on white workers in columns 3 and 4 are smaller in magnitude and there is no effect on occupational upgrading.[23]

---

[21] This result is consistent with Boustan (2009) and Derenoncourt (2019) —see Appendix Section B.3.5

[22] The 1944 figures come from a special census and are reported by the UAW-CIO research department in "Discrimination against Negroes in Employment 1942-7" Box 9, Folder 9-24, UAW Research Department, Archives of Reuther Library.

[23] One interpretation is that occupational segregation did not decrease and white men upgraded *within* the skilled category. We show in Appendix Table B.6 that war expenditures decreased occupational segregation

In all cases we can reject (at the 5 percent level) the null hypothesis that the effects are the same for Black and white workers. Thus, war expenditures reduced racial gaps in wages and skilled employment shares in metro areas with more expenditures.

The magnitudes of these changes for Black workers are economically meaningful. The share of Black workers in skilled occupations increased by 4.8 percentage points more in metro areas at the 90th percentile of expenditures compared with metro areas at the 10th percentile, representing a 14.7% percent increase relative to the mean in 1940. Similarly, the wage gains for Black workers were 9.3 log points (9.7 percent higher wages relative to the mean in 1940), which is higher than the estimated effects of an additional year of school at the time.[24]

Not surprisingly, war spending did significantly increase migration to these areas for white workers (1.8 log points per 1000 in war spending) and more so for Black workers (4.2 log points). Thus the war effort appears to have improved the labor market outcomes of Black workers in two ways. First, by increasing the wages of Black workers already residing in these metropolitan areas and second by inducing migration into these areas. We consider the relative importance of these two factors in the quantification exercise presented later.

The last panel of the table shows that expenditures were not associated with significant changes in prime-age employment rates for either Black or white workers. Figure 2.5 also shows that areas receiving more WWII contracts did not have greater employment shares in defense sectors by 1950. Although the defense employment of all groups rose significantly during the War, these effects did not persist to 1950. Consistent with previous work, we confirm these contracts were a temporary labor demand shock and did not result in different per capita economic trajectories for the receiving cities.

---

indexes – therefore occupational distributions did become more similar even at more granular levels. This result is consistent with limited wage increases for white men. Appendix Table B.9 shows Black increases are concentrated in "Operatives" and "Craftsmen" occupational categories.

[24]OLS estimates of the returns to schooling at the time from for all men range from 5 percent (Goldin and Katz, 2000) to 8 percent (Clay, Lingwall and Stephens, 2012). Returns to schooling were typically lower for Black men in this time period.

Threats to identification. Our main identification assumption is that cities receiving large contracts would be on similar trends as cities receiving small contracts. A first check is to test the sensitivity of the findings to adding other baseline covariates interacted with a post WWII indicator (post). These include the predictors of expenditures, namely manufacturing and predicted enlistment rate, and a vector of other controls from the 1940 census (share of men employed in agriculture, share Black, and average years of education) interacted with $Post_t$. Figure 2.6 presents the coefficient on $WarExp_r \times Post_t$ for a number of alternative specifications, separately for white and Black men. The coefficients are not very sensitive to the inclusion of any control. One might worry that these controls do not capture all regional differences. The figure shows that results are not driven by the North or the South; they hold within region.

A second check is to examine pre-trends in the outcomes. The coefficients in Figure 2.6 show that, as was the case in the previous figures, war expenditures do not predict changes in outcomes from 1920 to 1930 or 1930 to 1940. Unfortunately, we cannot conduct this test for wages, data for which was first collected in the 1940 census.

IV approach. We adopt a second identification strategy and instrument for WWII expenditures. We make use of firm-level data on war contracts collected by Li and Koustas (2019) to predict expenditures at the city level, based on detailed industrial composition at the local level and national expenditures by industry. This standard Bartik approach relies on a different identification assumption. Following the logic of Borusyak, Hull and Jaravel (2019), it assumes that war expenditures across industries at the national level during WWII are as good as randomly assigned, conditional on shock-level observables.[25] In the first column of Appendix Table B.7, we reproduce our main results for Black workers for reference. The second column shows the OLS results again, but using all the additional controls that this new IV approach requires (See Borusyak, Hull and Jaravel (2019)). These results are very similar to those in column 1. Column 3 shows the IV results. The F-statistic at the bottom of the table shows that predicted expenditures are a strong predictor of actual

[25]See Appendix Section B.3.3 for more detailed discussion.

expenditures. The IV coefficients are statistically indistinguishable from the DD/OLS coefficients for the share skilled and wages. The IV estimates are larger for the effect on the log of the male population.[26] We interpret this as further evidence in favor of the validity of our baseline identification assumption and strategy.

The role of worker composition. Given the increase in the Black population in areas receiving WWII contracts, a natural question is whether the occupation and wage effects we observe are due to changes in the composition of workers. Our evidence suggests this is not the case. Columns 2 and 4 of Table 2.3 show that the results are very similar when we exclude potential interstate migrants.[27] We also show in Appendix Figure B.5 that war expenditures are not associated with statistically significant changes in the share of (Black or white) prime-age men with high school degrees between 1940 and 1950.[28] Finally, Figure 2.6 shows that occupational upgrading and wage increases also occurred within education and age levels. Altogether these results suggest that migration and the possible changes in composition it may have generated are not the main sources of the wage and occupation changes we observe. Rather, war expenditures increased the wages of workers already residing where money flowed (and also wages of migrants). Of course, whereas migration does not drive our difference-in-difference empirical results, it may play an important role in generating aggregate effects; we investigate this possibility in Section 2.7.

Alternatively, Black men might have gained additional experience during the war. To investigate whether work experience could explain the results, we focus on the outcomes of

---

[26]For both the share skilled and wages, we cannot reject the null hypothesis of exogeneity for war expenditure per capita. Therefore, we prefer the efficient OLS estimator. The coefficients for wages for 1940-1950 are small, but there are large standard errors, likely due to the very small number of Blacks workers in the 1950 census, which recorded wages only for sample line persons (less than 0.5% of the population). The results for 1940-1960, which use a larger 5 percent sample with wages show that the wage results are very similar on the IV specification.

[27]We define a potential interstate migrant as any individual who is living in a different state than their state of birth and who is not living with a child born in their current residence state before April 1, 1942. The regression is at the individual level, clustered at the metro-year level, with additional controls for age, marital status, and whether born in the South.

[28]Prime-age men is defined as men ages 25-54. Most would not have been affected by the increase in high school graduation associated with war expenditures that we document below.

workers who were too young to have gained significant experience during the war. When looking at changes from 1940 to 1950, we focus on comparing men ages 18-24 in 1940 to men ages 18-24 in 1950. When looking at changes from 1940-1960, the age range is 18-34. In both cases the point estimates in Table 2.4 are very similar to our main results, suggesting that our results cannot be explained only by additional experience gained during the war. Rather, Black workers without war experience were also able to access higher paying occupations after the war.[29]

Oaxaca-Blinder decompositions can also be informative about the extent to which changes in the distribution of workers across regions, occupations, and industries affected the wage gap. The results in Appendix B.2.3 are consistent with those reported in Margo (1995) and suggest that occupational upgrading—along with wage compression across education groups and occupations—are the main sources of relative wage increases for Black workers during this period.

Other robustness checks. We also consider alternative weighting schemes and levels of geographic aggregation. Figure 2.6 shows that the choice of weights does not affect the estimated coefficients, though the standard errors are larger if we do not use any weights. We also consider an alternative level of geographic aggregation because city level data excludes rural areas and does not cover 100 percent of the population. We reproduce our results at the commuting zone (CZ) level and the state level in Appendix Table B.8. The results are robust to these alternatives.

Additional results. Figure 2.6 shows that WWII expenditures did not affect women in the long run, consistent with the findings by Goldin (1991) and Rose (2018).[30] This result may be surprising given that the employment of women substantially increased during the war

[29]These results are consistent with Collins (2000)'s findings using the retrospective Palmer Survey for six cities, who writes "it was not the training that made war-industry jobs in 1944 especially valuable to blacks later in the decade, but rather it was the continued access to the high wages associated with continuous employment in such industries relative to the wages in the other industries in which blacks were likely to work."

[30]Other work that leverages variations in the draft rate finds some evidence of persistent effects for women, for example Acemoglu, Autor and Lyle (2004).

(Appendix Table B.5). There might be several reasons for the difference between women and Black male workers. First, the executive orders banned discrimination on "race, creed, color, or national origin" but not on gender. Second, the large baby boom that occurred after the war resulted in many women exiting the labor force. Finally, historical accounts show that firms and unions gave preference to hiring returning soldiers in the post WWII period for both patriotic reasons and because many of the jobs in manufacturing were seen as traditionally male and the entrance of women into these jobs had caused considerable controversy. Historian Stephen Meyer writes that in the car industry "after World War II ended, gender solidarity prevailed over racial solidarity when managers and white workers accepted black men and purged white women from American auto plants."[31]

We also estimate results separately for defense and non-defense industries. It is unclear ex-ante whether non-defense industries would be affected by WWII contracts: they were indirectly affected by demand shocks (a point we return to below) but were not directly subject to the anti-discrimination policies attached to the contracts. Figure 2.6 shows that the results are larger in magnitude within the defense industry. This finding is consistent with Collins (2000)'s evidence from the Palmer survey showing that Black progress was mostly concentrated in defense industries.

2.4.3 Long-term effects of expenditures 1920-1970

We now investigate the effect of expenditures over the longer term. To do so, we estimate the following regression, after stacking the data for all census years 1920 to 1970:

$$Y_{rt} = \sum_{r \neq 1940} \beta_r WarExp_r \times \mathbb{I}_{t=j} + \gamma_r + \alpha_t + X_{rt}\rho + \varepsilon_{rt} \tag{2.2}$$

[31] Meyer, Stephen. 2004. "The Degradation of Work Revisited: Workers and Technology in the American Auto Industry, 1900-2000." *Automobile in American Life and Society.* http://www.autolife.umd.umich.edu/Labor/L_Overview/L_Overview6.htm (accessed June 29, 2021). In her book Kesselman (1990) writes "Research has demonstrated that while the wartime labor shortage created opportunities for women, lasting change was inhibited by the government, unions, and media, and management." For a discussion of how unions treated women after the war see Loos (2005) and references therein.

where Y_{rt} is either the log of average wages or the share in skilled occupations, $WarExp_r$ are total war expenditure per capita for metro r and they are now interacted with a dummy for each decade other than 1940, which serves as the reference category. To account for changing metro definitions over time, we use metro boundaries based on commuting zones in order to maintain a uniform definition of labor market over time.[32] All other controls are defined as before and interacted with Census year indicators.

The estimated coefficients for each decennial census year are presented in Figure 2.7. War expenditures were negatively correlated with the share of employment in skilled occupations among Black workers in 1920 and 1930, though not statistically significantly so. However, the effect of war expenditures on skilled employment becomes positive and significant in 1950 and 1960. It is still positive (though insignificant) in 1970 (Panel A). The effect on wages (Panel B) is also positive and significant from 1950 through 1970 (recall that the Census did not ask about wages before 1940). Expenditures have small and statistically insignificant effects for white workers for both measures in all years.

In sum, the results suggest that during the war, war expenditures led to labor shortages and increased employment of Black workers, many of whom moved to cities where firms received large contracts. Black workers appear to have gained access to previously unavailable skilled occupations in these industries and earned higher wages as a result. These positions and their associated higher wages remained available to Black workers for many decades after the war ended. We investigate any intergenerational effects by examining whether war spending affected the schooling of the next generation, before turning to an examination of underlying mechanisms.

[32]For example, metro areas expanded due to "white flight" to suburban counties in the 1950s and 1960s. See Boustan (2010) for evidence on the effect of the Great Migration on "white flight."

2.5 Intergenerational effects on human capital

2.5.1 Effects of war expenditures on school attendance and completion

Why would war expenditures affect the schooling of the next generation? With the increase in labor demand brought about by the war production effort, one might expect school enrollment and high school completion to decline given the additional labor market opportunities for 16-18 year olds. In fact, enrollment did temporarily decline during the war. But income gains for Black workers were permanent, reducing the need for offspring to work at young ages. Other reasons we might expect schooling to increase include possible increases in local school spending and reductions in residential segregation by race that might have increased Black families' access to better-resourced schools.

To explore this, we first estimate the effects of war expenditures on school enrollment among 16-18 year olds across metropolitan areas between 1940 and 1950. We estimate our main DD equation (2.1), but with school attendance at the individual level as the outcome of interest and with errors clustered at the metro-year level. We focus on 16-18 year old children because almost all children in metropolitan areas, including Black children, report being enrolled in school at ages 14-15 in 1940.[33]

We find that the school attendance of Black boys increased more in areas with greater war contract expenditures (Appendix Table B.11), as evidenced by the positive and statistically significant coefficient on $WarExp_m \times Post_t$ for Black boys. The results are positive for Black girls, though about half the size and not statistically significant. The results also hold if we exclude the South (Panel B), which we do because of existing evidence showing significant improvements in the quality of schools serving Black children during this period (Card and Krueger, 1992). There is no effect on white boys or girls in any specification.[34]

[33] 88% of Black children ages 14-15 attended school in 1940. In comparison, only 60% in the age group 16-17 attended school.

[34] In Appendix Table B.12 we repeat this exercise for 1940 to 1960 and also find positive impacts on the school attendance of Black boys. As a falsification exercise, we repeat the analysis for 1930 to 1940 and find no effects (as expected) in Appendix Table B.13.

However, this analysis suffers two limitations. The first is that they pertain to only the few cohorts included and the second is that we are limited by the small sample in 1950 since only sample-line persons were asked about schooling. Thus, we present results from an alternative approach that uses completed schooling reported in the 1960 Census, where we have a 5% sample in which all individuals were asked about years of education. We then estimate the following equation,

$$Y_{irgc} = \sum_{r \neq 1939} \beta_r WarExp_r \times \mathbb{I}_{c=j} \times \mathbb{I}_{g=Black} + \sum_{r \neq 1939} \gamma_r WarExp_r \times \mathbb{I}_{c=j} + \gamma_{rg} + \gamma_{tg} + X_{irgc}\rho + \varepsilon_{irgc} \tag{2.3}$$

where Y_{irgc} is a dummy equal to one if individual i of race g and graduation cohort c living in metropolitan area r graduated from high school.[35] WWII expenditures are interacted with a dummy for Black race and with cohort dummies. A cohort is defined as a three-year age group, based on expected high school graduation year. Individuals graduating high school during 1939-1941 serve as the excluded baseline category. We restrict the sample to non-Southern metropolitan areas and drop individuals who have moved between states in the previous five years.[36]

We estimate this regression separately by gender and plot the estimated coefficients in Figure 2.8. Note that the coefficient for $WarExp_r \times \mathbb{I}_{c=j} \times \mathbb{I}_{g=Black}$ identifies the impact of war expenditures on Black children relative to white children. There is a clear increase in the share of Black boys graduating high school and a similar increase for Black girls, albeit noisier, starting with the 1942-44 graduating classes, but not before.[37]

---

[35]The main concern with these estimates is migration. There are two concerns: first, individuals who migrated to a metropolitan area after completing school might be counted as more (or less) treated than they actually were, which would attenuate estimates. Secondly, war expenditures could have differentially attracted more educated migrants for younger cohorts. We include a dummy for whether an individual was born in the South and interact with race dummies as well as a full set of cohort indicators.

[36]We restrict the sample to non-Southern metropolitan areas since Southern metropolitan areas have substantially higher rates of within-state migration, making it more difficult to determine where someone likely received their education.

[37]Appendix Figure B.8 uses the 1940 Census to look at pre-trends across cohorts prior to WWII and finds little evidence of pre-trends for Black boys, though again finds noisier results for Black girls.

We conclude that war expenditures are associated with increases in the high school graduation rates of Black children relative to white children. The magnitude of the coefficients suggests the effects are not trivial with the mean coefficient on the interaction term for boys (girls) being 2.1% (1.6%). The high school graduation rate for Black boys (girls) in metropolitan areas at the 90th percentile of expenditures is 6.3 (4.8) percentage points higher than the graduation rate in metropolitan areas at the 10th percentile of expenditures.

Importantly, the results show that war expenditures were not associated with greater high school graduation rates for the Black or white cohorts that graduated before 1940. The effects we find are only for cohorts graduating after 1941. In contrast to the estimated effects for Black boys, higher war expenditures are associated with slightly lower high school graduation rates during WWII for white boys.[38] The estimates for white boys remain negative but become statistically insignificant after 1947. The results for Black and white girls (shown in Panel B) are similar to those for boys.

2.5.2 Why did schooling increase?

We consider four potential mechanisms behind the positive impact of war spending on the schooling of Black children: i) changes in the returns to schooling, ii) changes in public spending on schools, iii) reductions in residential segregation and iv) increases in parental income. First, we investigate whether war expenditures affected the returns to school using a standard Mincerian wage equations where we interact whether a Black individual completed at least some high school with $WarExp_m \times Post_t$.[39] Appendix Table B.14 shows this triple

[38] We verified that high school enrollment and graduation rates decreased nationally during WWII using data provided by Claudia Goldin, which comes from the Biennial Reports of the Commissioner of Education.

[39] Specifically we estimate the following equation:

$$Y_{irt} = \beta_1 E_{irt} + \beta_3 E_{irt} \times Post_t + \beta_2 Exp_r \times Post_t + \beta_4 E_{irt} \times Exp_r + \beta_5 E_{irt} \times Exp_r \times Post_t + Post_t + \gamma_r + X_{irt}\rho + \varepsilon_{irt}$$

where E_{irt} is an indicator for whether the individual completed at least some high school. We use whether the individual completed at least some high school so we can directly compare with model predictions. Results are similar if years of education is used instead.

difference is statistically insignificant: War expenditures did not increase returns to schooling (column 1). This is consistent with the fact that returns to school declined during the "Great Compression" period (Goldin and Margo, 1992) and not because of changing selection into schooling (Bishop, 1989). The table also shows that there were no changes in the extent to which education allowed Black children to access high skilled occupations (column 2). Thus, higher returns to school do not explain the increased investment in school that we document.

Next, we examine school expenditures. The fact that white children are not positively affected by WWII contracts suggests there were no major changes in education policy or expenditures in cities with greater expenditures. We verify this by estimating equation (2.1), but replacing the outcome with the log of education expenditures per capita. The results in Table 2.7 show there were no significant increases in education expenditures in cities with more war expenditures.[40]

To investigate if there were changes in residential segregation, we look at whether war expenditures affected two indices of segregation: the dissimilarity index and the isolation index from Cutler, Glaeser and Vigdor (1999).[41] We observe no declines in residential segregation associated with war expenditures using either one, as shown in Table 2.7.

Overall, the most plausible mechanism appears to be the change in family income. Previous work has shown that parental income remains the most important predictor of children's educational achievement, even more so than parental education (Reardon, 2011). Recent analysis of the strong association between racial segregation and racial achievement gaps concludes that the gap is completely accounted for by racial differences in poverty rates (Reardon et al., 2019). This is true even after years of increasing public expenditures on schools that serve lower income students (Lafotune, Rothstein and Schanzenbach, 2018).

Given this, it should not be surprising that declines in workplace discrimination that led

---

[40] Unfortunately there is no data at the sub-city level that would allow us to investigate whether expenditures or quality of school increase in Black neighborhoods.

[41] The index of dissimilarity is defined for metropolitan area r as $Dissim_{rt} = \frac{1}{2}\sum_{i=1}^{N} \left| \frac{Black_{irt}}{Black_{rt}} - \frac{White_{irt}}{White_{rt}} \right|$ where i is a residential area. The isolation index is defined as $Isol_{rt} = \frac{\sum_{i=1}^{N} \frac{Black_{irt}}{Black_{rt}} \frac{Black_{irt}}{Pop_{irt}} - \frac{Black_{rt}}{Pop_{rt}}}{\min(\frac{Black_{rt}}{Pop_{irt}}, 1) - \frac{Black_{rt}}{Pop_{rt}}}$.

to substantial increases in the earned income of Black families would result in increases in the educational achievement of their children. A move from the 10th to the 90th percentile of war expenditures is associated with an absolute (not relative to whites) increase in wages of 9.4% and an absolute (not relative to whites) increase in the share of Black boys graduating high school of 3.6%.[42] If we assume all of the increases in schooling are due to greater incomes, then this implies an elasticity of 1.0 (for Black girls 0.5). This is broadly consistent with analysis based on more contemporary data of an outsized role of parental income in explaining educational outcomes of children.[43]

2.6 Why did labor market outcomes improve and persist?

2.6.1 Historical accounts

Evidence from historical accounts suggests that a combination of civil rights activism, following the Executive order, and severe labor shortages led firms to hire Black workers into more highly skilled, higher paying occupations. To illustrate, we begin with a case study of Boeing, a recipient of roughly $10 billion (2020 dollars) in WWII contracts.[44] At the start of the war, Boeing employed 8,500 workers, none of whom were Black. By the end of the war, Boeing employed 1,600 Black workers. The growth in the Black workforce at Boeing was achieved through a combination of pressure from civil rights organizations and labor shortages. When initially confronted by Civil Rights activists about the lack of Black workers among the 29,000 employed in 1941, Boeing argued that the union's refusal

[42] The share of Black boys graduating high school for the 1942-1959 cohorts was 38.6%, excluding the South. For Black girls it was 44.8%.

[43] Existing work based on more recent data has generated estimates of parental income elasticities with respect to years of completed schooling of their children (not high school completion) that range from 3 to 80 percent (Taubman, 1989). Our estimates are on the higher end, which may be due to (1) the extremely low levels of schooling at this time among Black families, (2) the different definition of the outcome (high school completion) or (3) the effect of aggregate income shocks possibly differing from family-specific income shocks (for example, by generating peer effects).

[44] This account is based on the work of Meyers (1997) and Davenport (2006).

to issue work permits to Black workers was responsible. The NAACP subsequently brought a complaint against Boeing to the FEPC, which resulted in the hiring of 329 Black workers, representing only 1% of its workforce in 1943.

The hiring of Black workers remained controversial and was met with complaints and work stoppages by white workers. Boeing considered many alternatives to hiring more Black workers, including negotiating war deferments for its workforce and hiring women. Ultimately, facing a severe labor shortage of 9,000 workers and under pressure from Civil rights activists, Boeing hired 1,600 Black workers in 1944, all of whom joined the union, guaranteeing them continued access to skilled, higher paying jobs.

The integration of the automobile industry is another such example. In the heavily unionized automobile industry, Black workers made up 4 percent of the workforce before the war, and 15 percent by the end of the war (where it remained through 1960). Regarding this integration, historian Sugrue writes: "Civil rights activism alone did not, however, open workplaces to blacks. Corporate leaders, facing a desperate shortage of workers because of wartime mobilization and the draft, opened their doors to black workers for the first time."[45]

But why would these effects persist long after the contracts ended and the requirement of non-discrimination was rescinded? These accounts suggest a likely explanation. Prior to the war, unions explicitly denied Black workers entry into skilled jobs by refusing to grant membership, but labor shortages combined with civil rights activism led to the integration of the unions. Some unions integrated voluntarily, as it was in their own self-interest to do so. Abel (2011), explaining the factors driving integration of aviation unions in Texas, writes that "white aircraft workers had little choice but that their welfare and the strength of their union depended on maintaining such color-blind economic principles as seniority and equal work." Other unions were forced to integrate by the courts. An additional channel could have been that white union members and firm owners may have learned during the

[45]Throughout the nation the integration of Black workers into the mostly white labor force was difficult. In 1943 there were at least 242 racial incidents in 47 cities due to racial frictions caused by housing shortages, employment conflicts, or outright racism (Sitkoff, 1971).

war that Black workers were more productive than previously thought and updated their beliefs. This labor market "hysteresis" due to changes in discrimination is consistent with other research.[46] Once Black workers were admitted to unions, either voluntarily or due to legal pressure, their access to higher paying jobs within the industry remained. In the next section, we provide further empirical support for this explanation.

2.6.2 Additional empirical evidence

Using additional data, we quantify the extent to which the Executive order and associated civil rights activism and/or severe labor shortages led firms to hire Black workers into more highly skilled, higher paying occupations. First we explore the role of labor shortages by incorporating information on months of labor market shortages by metropolitan area as determined and recorded by the US employment services during the war. Specifically, we re-estimate our main equation of interest, but interact WWII contracts with the number of months of severe labor shortages in each city. Column 1 of Table 2.8 reproduces our main result: WWII contracts led to higher employment in skilled occupations. In column 3 we interact WWII contracts with months of labor shortages. The coefficient on WWII contracts by itself becomes insignificant, but the interaction with shortages is positive and statistically significant. This suggests that WWII contracts had no impact on outcomes in places where there were no shortages. As one historian of this period concludes "These manpower shortages gradually forced white employers and workers to forget their prejudices, if only temporarily, and accept Black employees. By the end of the war the quantity and quality of jobs open to Afro-Americans had increased dramatically" (Wynn, 1976).

Next we explore whether the executive orders and the associated civil rights activism

[46] Whatley (1990) finds that Cincinnati manufacturing firms during WWI exhibited state dependence – once they hire a Black worker they are more likely to hire Black workers in the future. Miller and Segal (2012) show that the effect of affirmative action quotas on police hiring persists even after the quotas are no longer mandatory. Miller (2017) shows similar evidence for private employers after they are no longer subject to federal contracting affirmative action policies. Finally Saez, Schoefer and Seim (2019) find that subsidies for youth employment increased youth employment even after ending the policy due to a permanent decline in discriminatory job postings.

also mattered. We do so by comparing the effects of direct and indirect increases in demand due to war contracts. The war contracts increased demand among the firms that received contracts (direct effects). But they also increased demand for goods and services for the firms that supplied the inputs to the industries receiving defense contracts (indirect effects). For example, the direct demand for a B-17 bomber generates significant indirect demand for aluminium. However, the indirect suppliers were not bound by the executive orders barring discrimination in hiring and could not easily be targeted by civil rights activists and courts. Thus, we can compare the impact of direct and indirect demand for labor on the outcomes of Black workers to learn whether the executive order and the requirement of nondiscrimination played an important role in closing the racial wage gap.

We measure the direct demand shock to each region as the sum across industries of the value of national industry wartime contracts, weighted by regional employment shares across industries. We measure the indirect demand shock similarly, but additionally incorporating input-output linkages across sectors to measure the total industry demand shock induced by government contracts.[47] To do so, we use historical Leontief input-output tables. Crucially, the direct and indirect demand shocks measure the importance of demand increases induced by government expenditures, but only the direct shock is subject to the executive orders and corresponding civil rights activism.

Direct shocks (column 1 of Table 2.9) and indirect shocks (column 2) are both associated with larger shares of Black workers in skilled jobs, when we consider their effects separately. But when we include them together (column 3) we find that only direct demand shocks led to improvements for Black workers. The coefficient on indirect shocks is substantially smaller in magnitude than the coefficient on direct shocks and it is not statistically significant. These results imply that labor demand shocks were not sufficient. Instead, the executive orders played an important role in converting labor demand shocks into improved labor market outcomes for Black workers, consistent with the industry narratives.

[47] See Appendix B.3.4 for details of both measures.

Finally, we investigate the role of unions. From 1900 to 1935 unions had explicitly or implicitly discriminated against Black workers and barred them from obtaining well paid jobs and promotions in high paying manufacturing jobs. As a result black workers very frequently participated in strike breaking, accessing jobs when employers wanted to break the unions, further fostering animosity between black workers and white union members. This decade-long stalemate started to turn around in 1935 with the Wagner Act of 1935, which ended the advantages of strike-breaking (Whatley, 1991). The events that took place during WWII resulted in the acceptance of black workers into unions. The research department of the UAW-CIO in 1944 reports, "In 1936 the Negro membership of trade unions was 150,000. Today there are upwards of three-quarters of a million Negroes organized into trade unions."[48] These drastic increases in unionization were tied to WWII contracts.

To document this empirically, we take advantage of newly collected data by Farber et al. (2021), who document that rising unionization rates led to the post war decline in inequality in the U.S.. By collating data from multiple sources, Farber et al. compute unionization rates by state, year, and race from 1937 onward. We replicate their findings and show in Appendix Figure B.7 that areas that received more WWII contracts had higher unionization rates among workers.[49] Moreover, these increases were larger among Black workers than among white workers. Altogether, these results are consistent with the historical narratives and suggest a real decline in institutional prejudice in the form of union membership increases, which can explain why the effects of the war contracts persisted for so many decades.

[48]UAW Research Department, Box 9, Folder 9-24. Discrimination Against Negroes in Employment 1942-7. Archives of the Reuther Library.

[49]This is consistent with other historical evidence. Unions grew substantially among firms receiving WWII grants, including in the United Auto Worker union (UAW), which was already large before the war (Troy, 1965).

2.7 Quantification

We have shown that WWII contracts and the associated requirement of nondiscrimination played a significant role in improving the labor market outcomes of Black workers and the educational attainment of their offspring by comparing changes in outcomes across regions. But how much of the aggregate changes in this period can WWII expenditures account for? And how much of these improvements can be ascribed to changes in discrimination? And finally how important is migration for driving these aggregate impacts? To answer these questions, we develop a general equilibrium model.

2.7.1 Model

We build on a canonical macroeconomic model of labor-market discrimination introduced in Hsieh et al. (2019), extending the model to match the details of our empirical setting and estimating parameters to match our empirical evidence. In our model, the allocation of labor groups (e.g. the intersection of education and race) across regions, the allocation of labor groups within regions across industries and occupations, and the average wages of labor groups across regions are endogenous.

At time t there is a continuum of workers indexed by $z \in \mathcal{Z}_t$, each of whom inelastically supplies one unit of labor.[50] Workers are exogenously divided into a finite number of labor groups, indexed by g. The set of workers in group g is given by $\mathcal{Z}_{gt} \subseteq \mathcal{Z}_t$, which has mass N_{gt}. Workers choose in which region (indexed by r) to live and in which industry (indexed by i) and occupation (indexed by o) to work in order to maximize utility. Labor is the only factor of production. All markets are perfectly competitive and all factors are freely mobile across occupations, industries, and regions. We index by $\mathcal{Z}_{rgt}$ and $\mathcal{Z}_{riogt}$ the endogenous sets of workers in group g who choose to live in region r and who choose to live in region r and work in industry-occupation io at time t.

[50]Given our reduced-form evidence showing no clear impact of wartime spending on employment shares, we abstract from endogenous labor supply. Incorporating endogenous labor supply will leave our baseline results largely unchanged given well-identified estimates of the labor-supply elasticity.

Production. Final good output is produced locally and is not traded, so that its consumption equals its production, both of which are denoted by C_{rt}. This final good is produced combining the services of industries according to a Cobb Douglas production function

$$C_{rt} = \prod_i C_{rit}^{\mu_i} \tag{2.4}$$

where $C_{rit} \geq 0$ is region r's consumption of industry i, $\mu_i \geq 0$, and $\sum_i \mu_i = 1$.[51] Consumption of industry i in region r is itself an aggregation across consumption of industry i purchased from all regions and is given by

$$C_{rit} = \left(\sum_j \mu_{jit}^{1/\rho} C_{jrit}^{(\rho-1)/\rho} \right)^{\rho/(\rho-1)} \tag{2.5}$$

where C_{jrit} is consumption of industry i in region r purchased from region j, $\mu_{jit} \geq 0$ is a demand shifter for industry i output produced in region j, and $\rho \geq 0$ is the elasticity of substitution across regions (which is common across industries and time).

Output of industry i in region r is given by

$$Y_{rit} = \left(\sum_o \mu_{riot}^{1/\eta} Y_{riot}^{(\eta-1)/\eta} \right)^{\eta/(\eta-1)} \tag{2.6}$$

where Y_{riot} is the output of occupation o used in the production of industry i in region r at time t, $\mu_{riot} \geq 0$ is a demand shifter for this occupation output, and $\eta \geq 0$ is the elasticity of substitution across occupations (which is common across industries and time). Occupation o output supplied in industry i is the sum of efficiency units, L_{riogt}, provided by all groups employed therein

$$Y_{riot} = \sum_g L_{riogt}. \tag{2.7}$$

[51] During the war, most output of war industries is purchased by the government. We use the model to quantify the impact of government expenditures between 1940 and 1950, years in which government national defense expenditure shares were relatively low at 2.7% and 7.6% of GDP in 1940 and 1950 respectively. Expenditures had peaked at 43.3% of GDP in 1944.

A worker $z \in \mathcal{Z}_{rgt}$ supplies $T_{riogt}\varepsilon_{ziot}$ efficiency units of labor if employed in industry-occupation pair io in region r at time t, so that

$$L_{riogt} = \int_{\mathcal{Z}_{riogt}} T_{riogt}\varepsilon_{ziot}dz, \tag{2.8}$$

The parameter T_{riogt} is the systematic component of net productivity (productivity combined with a discriminatory wedge). A high value of T_{riogt} represents a combination of a high productivity and/or a low discriminatory wedge of group g in region r at time t within industry-occupation io. In what follows, we often refer to T_{riogt} as a "net productivity" for brevity. The parameter ε_{ziot} is the idiosyncratic component of productivity. Each worker is associated with a vector of ε_{ziot}, one for each io pair, allowing workers within $\mathcal{Z}_{rgt}$ to vary in their relative productivities across io pairs. We assume that each ε_{ziot} is drawn independently from a Fréchet distribution with cumulative distribution function $G(\varepsilon) = \exp\left(-\varepsilon^{-\theta}\right)$, where a higher value of $\theta > 1$ implies lower within-worker dispersion of efficiency units across io pairs.[52]

Worker choices. We take as given the supply of worker types at the aggregate level and model their allocation across space and across industry-occupation pairs within each location. The utility of a worker z living in region r and working in industry-occupation io is the product of an amenity from living in region r times an amenity from working in io times the worker's real wage. The amenity from residing in region r is given by the product of a systematic component, U_{rgt}, and an idiosyncratic preference shock, ε^{U}_{zr}, which is distributed Fréchet with shape parameter $\nu > 1$. The amenity from working in io within region r is given by A_{riog}. We assume that each worker first draws her preference shocks across regions and chooses her region, and then draws her productivity shocks across industry-occupation pairs and chooses her industry-occupation.

Market clearing and trade. Goods markets and labor markets clear. We assume that

[52] In the Quantitative Appendix we test and confirm a central prediction of the Fréchet distribution for changes in average wages; see a short description in footnote 58.

occupation output and final goods are not traded. We assume that industrial output is traded freely across regions and that trade is balanced.

Discussion of modeling assumptions. Following Hsieh et al. (2019), we model the impact of labor-market discrimination on occupation allocations and wages as a "wedge" between wages and marginal products in an otherwise competitive labor market. This wedge, embedded within T_{riogt}, reduces the perceived benefit to firms of employing Black workers; it is a reduced-form proxy consistent with a range of theoretical formulations of discrimination. For example, the wedge captures the fact that Black workers' productivity was reduced by threats and acts of violence.[53] In the next section, we allow this wedge to be affected by local government wartime expenditure.

We recognize that labor markets are not perfectly competitive and that this was especially so for the labor market confronting Black workers in the 1940s.[54] Nevertheless, given our goal of providing an internally consistent framework for evaluating the macroeconomic effects of our reduced-form findings, we view building on the canonical macroeconomic model of discrimination—Hsieh et al. (2019)—to be the best choice. Data to estimate specific models of discrimination are not available.[55]

2.7.2 Parametrization

Mapping theory to data. We map industries and occupations in the model, i and o, to the two aggregate industries—defense and non-defense—and the two aggregate occupations—skilled and unskilled—defined in our empirical work above. We map labor groups in the model, g, to four labor groups in the data defined by the intersection of two education levels

[53] Detroit alone lost three million hours of work in the first six months of 1943 due to strikes over the hiring and promotion of Black workers (Wynn, 1976).

[54] For example, unions restricted hiring practices, and firm owners and workers were subject to threats and violence for deviating from norms.

[55] For example, in order to estimate a model in which government wartime spending affects taste-based discrimination, one would need survey data on white perceptions of Black workers both before and after WWII across all regions. Such data do not exist.

(at least some high school and no high school)[56] and two races (Black workers and others, referred to as "white workers"). We map regions in the model, r, to the 146 metropolitan areas used in our empirical exercises, and time, t, to the years 1940 and 1950.

Calibration. While we estimate the key novel aspects of our theory (the impact of government spending), we calibrate four structural elasticities: θ, ρ, η, and ν. The parameter θ determines the elasticity of labor supply across io pairs within a region to changes in wages per efficiency unit. We set $\theta = 1.5$, in line with estimates in Burstein, Morales and Vogel (2019), Galle, Rodriguez-Clare and Yi (2018), and Hsieh et al. (2019).[57] The parameter ν determines the elasticity of population across regions to changes in real wages. We set $\nu = 1.5$, in line with a large literature estimating geographic labor mobility; see e.g., the review in Fajgelbaum et al. (2019). The parameter ρ determines the trade elasticity across regions. We set $\rho = 4$, in line with a large literature both in international and intra-national trade; see e.g., the review in Head and Mayer (2014). Finally, the parameter η determines the elasticity of substitution between the skilled and unskilled occupation within each industry. We set $\eta = 11$, which allows us to match closely our difference-in-difference empirical results on wages and occupation upgrading by race in regions receiving more relative to less government spending when feeding in all estimated shocks; see the Quantitative Appendix, Section B.4.3.2, for details.

Parametrizing anti-discriminatory effects of government spending. In Appendix B.4.3.3 we parameterize net productivities, T_{riogt}, and amenity values, U_{rgt}, as time-varying functions of government expenditure per capita, G_r.[58] Given this parametrization, in

[56] The majority of Black workers and many white workers did not have any high school education in 1940. We do not assume that a given reported education implies an equivalent productivity across races; see e.g. Boustan (2009) and Carruthers and Wanamaker (2017).

[57] Burstein, Morales and Vogel (2019) and Galle, Rodriguez-Clare and Yi (2018) estimate the equivalent of our parameter θ leveraging exogenous variation in labor demand across occupations (Burstein, Morales and Vogel, 2019) and industries (Galle, Rodriguez-Clare and Yi, 2018) using exposure to computerization and the China shock, respectively. Hsieh et al. (2019) estimate the equivalent of our parameter θ to match the dispersion of wages.

[58] In the Quantitative Appendix we allow the amenity associated with working in io within region r, A_{riog},

equilibrium, we obtain a simple reduced-form relationship between government wartime spending per capita, G_r, and the share of group g's labor in industry-occupation io within region r at time t, denoted by π^L_{riogt}, which can be estimated with a transformed version of our data:

$$\ln \pi^L_{riogt} = \gamma_{rgt} + \gamma_{riot} + \gamma_{riog} + \gamma_{iogt} + \beta_3 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g + \beta_4 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g \mathbb{I}_o + \iota_{riogt} \tag{2.9}$$

The parameter γ_{riot} in (2.9) captures the price of output in industry-occupation io in region r at time t as well as the common impact of G_r on the productivities of white and Black workers employed therein; we refer to these common changes in net productivity for Black and white workers as the compositional components of wartime spending.[59] The remaining γ parameters are fixed effects capturing, among other things, any changes at the national level in productivities of each labor group in each io across time (γ_{iogt}). The variables $\mathbb{I}_t$, $\mathbb{I}_i$, $\mathbb{I}_o$, and $\mathbb{I}_g$ are, respectively, indicator functions that equal one if the year is 1950 ($\mathbb{I}_t$), the industry is defense ($\mathbb{I}_i$), the occupation is skilled ($\mathbb{I}_o$), or the group (which is defined both by education and race) is Black ($\mathbb{I}_g$). This regression can be estimated by transforming the data into cells defined by region, industry, occupation, group, and time cells.

In this regression if $\beta_3 > 0$ then G_r reduces racial discrimination in the unskilled defense occupation, and if $\beta_4 > 0$ it reduces racial discrimination even more within the skilled defense occupation. β_3 and β_4 are identified as differential changes between 1940 and 1950 in the allocations of Black and white workers across io pairs across regions receiving different amounts of government contracts. In other words, the model predicts that if the share of Black workers with a particular level of education that is employed in occupation o within the defense industry grows more within regions receiving more government monies (conditional

to also vary over time with G_r. This allows us to match the relationship between G_r and changes in average wages within each $riog$. However, in our calibration we set A_{riog} constant across time since observed changes in wages within $riog$ cells match our model's predictions (that arise from the assuming that idiosyncratic productivities are distributed Fréchet).

[59]While these changes in primitives are common for Black and white workers, they can have differential effects across races because of the different initial compositions of Black and white workers across regions and across industries and occupations within regions.

on a set of fixed effects), then this is because government spending reduced discrimination in these labor markets within that occupation.

The main identification assumption allowing us to identify β_3 and β_4 is that—conditional on region-group, region-industry-occupation-time, and industry-occupation-labor-group-time fixed effects—changes in productivities of Black workers relative to white workers in region-industry-occupation triplets that *would have occurred in the absence of government spending* are uncorrelated with government spending. These fixed effects control for—among other things—national changes in each group g's employment patterns across industry-occupation pairs (for example, national skill-biased technical change within the skilled occupation in the defense industry) and local changes in the demand for and productivity of each industry-occupation pair (for example, factories built with government monies raising the productivity of the skilled occupation in defense). We interpret these parameters β_3 and β_4, as changes in discrimination—as opposed to changes in relative productivities—by assuming that government spending does not raise the primitive productivities of Black workers relative to white workers in industry-occupation pairs in regions receiving more government spending, conditional on the large set of fixed effects.

One threat to identification that is not captured by the model would be if the unobserved characteristics of Black workers improved relative to white workers in regions receiving more spending. In our reduced-form analysis, we showed that there is no similar pattern of Black occupational upgrading in non-defense industries (Figure 2.6), that the same patterns of occupational upgrading occur within the set of non migrants (Figure 2.6 and Table 2.3), and that the average share of prime-age men who have a high school degree did not change between 1940 and 1950 as a result of government spending (Figure B.5). Each of these facts suggests that migration did not affect the relative unobserved abilities of Black and white workers across regions, consistent with our identification assumption. Another threat to identification is that government spending raised the primitive productivities of Black relative to white workers, perhaps through wartime training. However, we also showed in our reduced-form analysis that the same patterns of occupational upgrading occur within the set of workers too young to have benefited from wartime training (Table 2.4), consistent with our identification assumption. A final threat to identification that is not captured by the model or our reduced-form analysis would be if the unobserved characteristics of Black workers improved relative to white workers in regions receiving more spending at a more disaggregate level: *within* either the unskilled or skilled occupation in the defense industry. Columns 2 and 3 of Table B.15 test this hypothesis explicitly by estimating

$$Y_{riogt} = \gamma_{rgt} + \gamma_{iogt} + \gamma_{riog} + G_r \mathbb{I}_t \mathbb{I}_i \big(\beta_1 + \beta_2 \mathbb{I}_o + \beta_3 \mathbb{I}_g + \beta_4 \mathbb{I}_g \mathbb{I}_o\big) + \iota_{riogt} \tag{2.10}$$

defining Y_{riogt} as either the average years of education (column 2) or average age (column 3) within each $riogt$ cell. Each coefficient is small and statistically insignificant in both columns of Table B.15, consistent with our identification assumption.

Table 2.5 presents results of estimating (2.9). In column 1, we find no evidence that greater government expenditure per capita decreases discrimination in the unskilled occupation in the defense industry: the coefficient $\beta_3 = 0.014$ is not significantly different from zero. On the other hand, we find statistically significant evidence that greater government expenditure per capita reduces racial discrimination within the skilled occupation in the defense industry. This increase is economically large. The point estimate of coefficient $\beta_4 = 0.115$ implies that a metropolitan area at the 90th percentile of government expenditure per capita experiences an increase of approximately 29% percent in the net productivity of Black (relative to white) workers within the skilled (relative to unskilled) occupation in the defense industry, relative to a metropolitan area at the 10th percentile.[60]

Similarly, in equilibrium we obtain a simple relationship between government wartime spending per capita, G_r, and the share of group g living in region r at time t, denoted by π^N_{rgt}. Together with our parametrization, the allocation of group g to region r at time t can

[60] The reduced-form parameter $\beta_4 = 0.115$ is related to the structural parameter β_4^T through $\beta_4^T = \beta_4/\theta$. Together with $\theta = 1.5$, we have $\beta_4^T \approx 0.078$. Government expenditure per capita (in \$1000s), G_r, is 3.95 and 0.25 at the 90th and 10th percentiles. Hence, $0.29 \approx 0.078 \times (3.95 - 0.25)$.

be expressed as

$$\frac{1}{\nu}\ln \pi^{N}_{rgt} - \ln Wage_{rgt} = \gamma_{rg} + \gamma_{gt} + \gamma_{rt} + \beta^{U}_{2} G_r \mathbb{I}_t \mathbb{I}_g + \iota^{U}_{rgt} \qquad (2.11)$$

where $Wage_{rgt}$ is the average wage of workers in group g in region r at time t and ν is the elasticity of migration to real wages. The right-hand-side of (2.11) represents the amenity value for group g of living in region r at time t (plus a constant across regions), which must be high in region r at time t if the share of group g living in region r at time t is high relative to the wage the group receives there, conditional on the labor supply elasticity across regions, ν, which is the left-hand-side of (2.11). For instance, we would interpret an increase in the share of more-educated Blacks living in region r between 1940 and 1950 (relative to what would be predicted by the change in their wages) as an increase in the amenity value of this group living in region r over time.

According to (2.11), differential changes between 1940 and 1950 in allocations of Black and white workers across regions receiving different amounts of government wartime spending (relative to that predicted by the observed changes in wages) identify the anti-discriminatory effects of government spending on amenities. The identification assumption is that, conditional on fixed effects, the differential changes in amenities of Blacks relative to whites in regions receiving more relative to less government wartime spending per capita that *would have occurred in the absence of government spending* are uncorrelated with government spending.

Column 4 of Table 2.5 presents the results of estimating (2.11), which is estimated under our baseline assumption that $\nu = 1.5$. We find statistically significant evidence that greater wartime expenditure raises the amenity value of a metropolitan area relatively more for Black than white workers. The coefficient $\beta^{U}_{2} = 0.054$ implies that a metropolitan area at the 90th percentile of exposure experiences an increase in its amenity value for Black relative to white workers of approximately 20 percent $(0.2 \approx 0.054 \times (3.95 - 0.25))$ relative to a metropolitan area at the 10th percentile.

In summary, we find that government spending during World War II had substantial anti-discriminatory effects. Regions receiving greater government expenditure per capita experience a substantial reduction in racial discrimination in the labor market, although only narrowly in skilled occupations within defense industries. Similarly, regions receiving greater government expenditure per capita become more attractive places for Black workers to live in (relative to white workers), conditional on the wages that Black and white workers receive.

2.7.3 Aggregate results

In this section we use the model to quantify the total impact of government wartime spending on aggregate patterns; the extent to which these effects of wartime spending are driven by reductions in discrimination; and the importance of migration for magnifying the impact of government spending on Black workers. To do this, we calibrate our model to match 1940 data and feed into our model the estimated changes in net productivities and amenities.

Upon feeding in these shocks, we solve the model for the 1950 equilibrium, holding all other parameters at their 1940 levels. Given the 1950 equilibrium, we then measure the aggregates of interest. Finally, in order to quantify the importance of migration, we revisit these exercises in a restricted version of the model in which workers cannot reallocate across space.

Table 2.6 reports our results. The first column reports the change in the share of Black relative to white workers in skilled occupations and the percent (ln) change in the wage of Black relative to white workers in the actual data between 1940 and 1950 aggregated across the 146 metropolitan areas. The second column reports the changes in these outcomes caused by wartime expenditure according to the model. Wartime spending causes a 2.7 percentage point decline in the difference between the share of white and Black workers employed in skilled jobs between 1940 and 1950, which is about a third of the total decline of 8.2 percentage points in the data. Wartime spending causes a 5.6 percent decline in the relative wage of white to Black workers between 1940 and 1950, which is a quarter of the total decline of 22.6 percent in the data. The third column reports the changes in these outcomes caused by the anti-discriminatory impacts of wartime expenditure. Almost all of the aggregate effects of wartime spending on the Black-white wage and skilled-employment gaps are caused by the anti-discriminatory effects of wartime spending. Finally, the final columns report the changes in these outcomes caused by wartime expenditure if there were no migration between 1940 and 1950. While migration has only a small impact on the change in the relative share of Black workers in skilled occupations, it has a first-order effect on the aggregate contraction of the Black-white wage gap. In particular, without migration the impact of government spending on the wage gap would have been smaller by a third. The intuition is straightforward. Government spending improved real wages and amenities for Black workers initially living in metro areas receiving more spending per capita. Migration spread these benefits more widely, as workers initially living elsewhere migrated towards regions receiving more spending. Thus, migration increased the aggregate impact of spending on Black labor-market outcomes.

2.8 Conclusion

Black workers experienced unprecedented improvements in their absolute and relative status in the 1940s (Brouillette, Jones and Klenow, 2021). We document that WWII contracts and their associated anti-discriminatory requirements were responsible for a substantial part of these increases, leveraging both local labor-market comparisons and a structural model. These contracts increased labor demand, at a time when labor supply fell due to the draft, generating substantial labor shortages. In addition, the president's Executive orders created a political and legal framework that allowed civil rights activist to demand employment and promotion of Black workers among firms receiving government contracts. As Abel (2011) summarizes, historians are divided regarding the relative role of these forces. Most acknowledge the importance of shortages and dismiss the role of the FEPC, while others give greater importance to the executive orders and the efforts of civil rights activists. This paper argues based on new evidence that these two forces complemented each other.

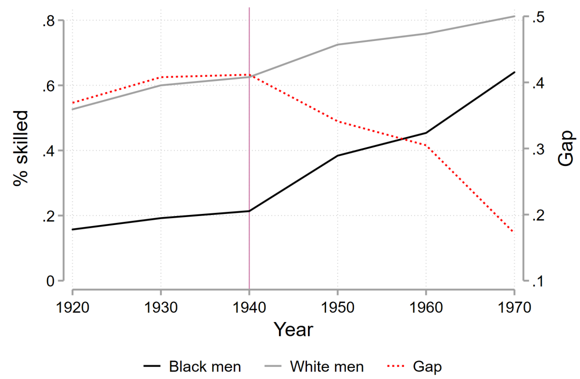
We also show that the effects of these WWII events persisted into the 1970s – in cities that received more contracts, Black workers earned greater wages and worked in higher skilled occupations. The persistence of these gains is likely due in great part to the fact that the events during WWII led unions to desegregate. Indeed, we estimate using our structural model that most of the effects of WWII contracts were driven by declines in discrimination. Our findings suggest that neither tight labor markets nor anti-discriminatory policy alone may be sufficient to generate permanent improvements in the labor market outcomes of minorities. However, in combination these forces appear to generate lasting improvements, in part because these forces led to institutional changes that lowered discrimination in the labor market.

Finally we also document that these labor market gains translated into higher education achievement among Black children. The evidence suggests that greater income among the parents is the most likely mechanism for these increases. This suggests that labor market policies can have important inter-generational effects. While much of the existing literature and policy debate on racial gaps in schooling focuses on disparities in educational inputs for Black and white children, our findings suggest that efforts to reduce the racial gap in schooling should also consider interventions that address existing discrimination in the labor market faced by Black families.

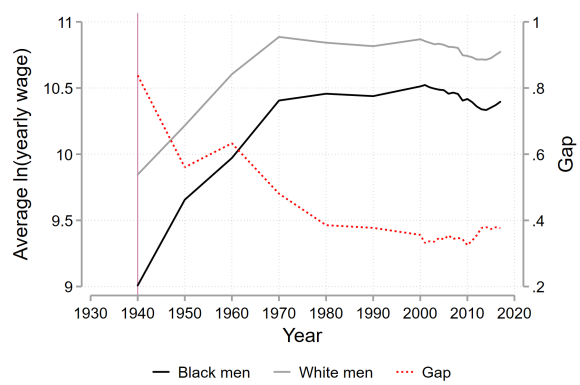
2.9 Tables and Figures

Figure 2.1: Long-term trends in Black-White gaps

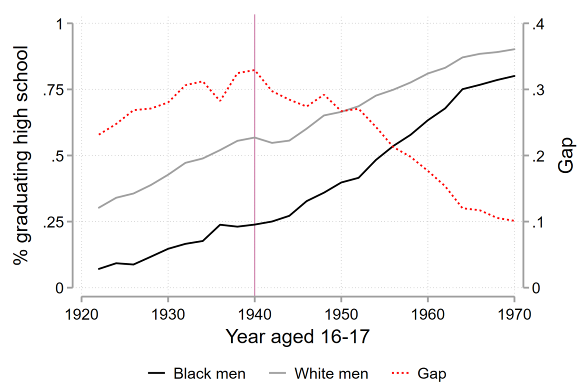
Panel A: Share in skilled occupations



Panel B: Average (log of) wages



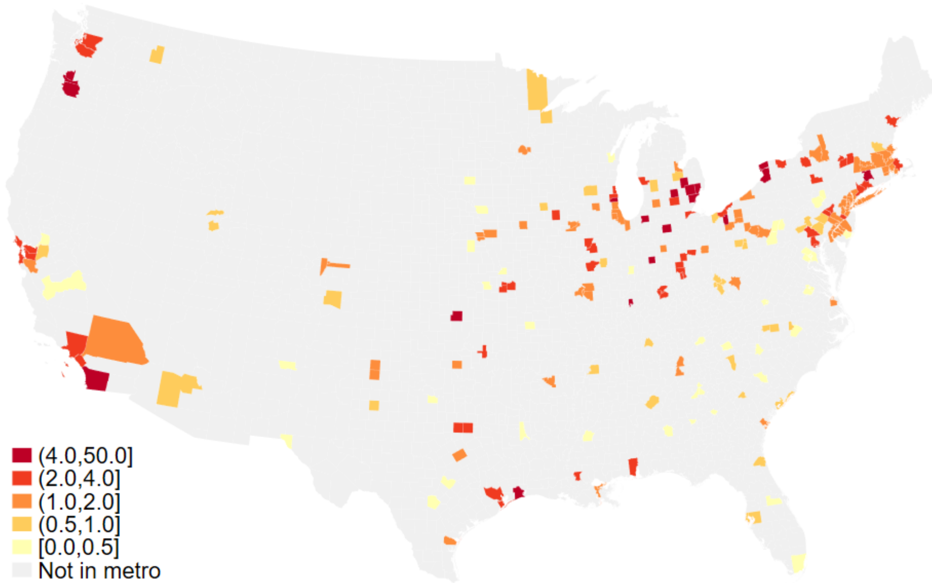
Panel C: Share graduating HS



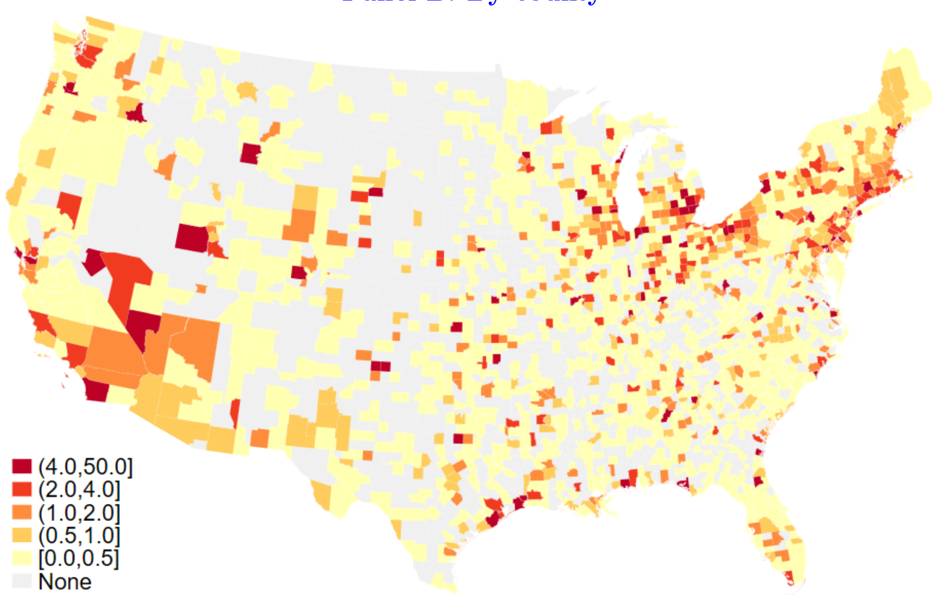
Note: Share skilled is the share of employed men who are not farmers, laborers, or service workers. Wages are total wage earnings (2017 dollars) in the previous year for men ages 25-54 who are currently employees. Share graduating high school is based on share completing at least twelve years of school by age 35. Data from Census and ACS samples for 1920-2017 accessed from IPUMS (Ruggles et al., 2020).

Figure 2.2: WWII expenditures per capita ($1000s, 1940)

Panel A: By metro



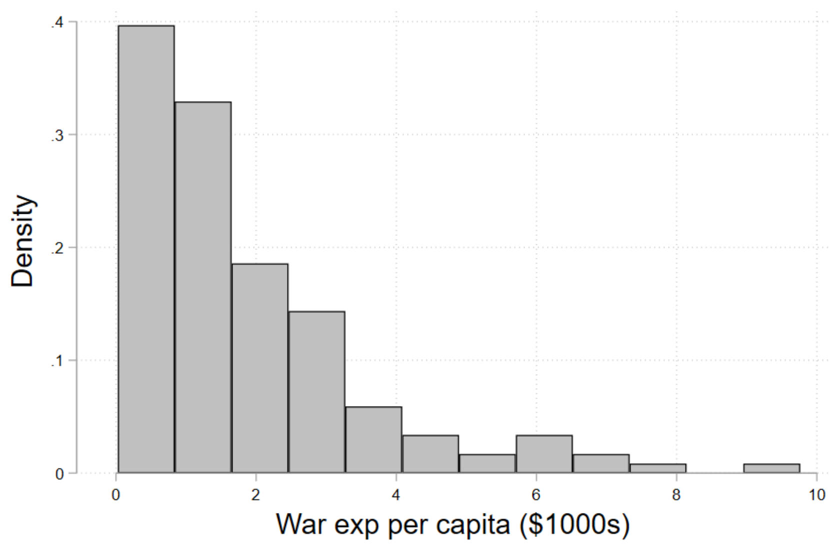
Panel B: By county



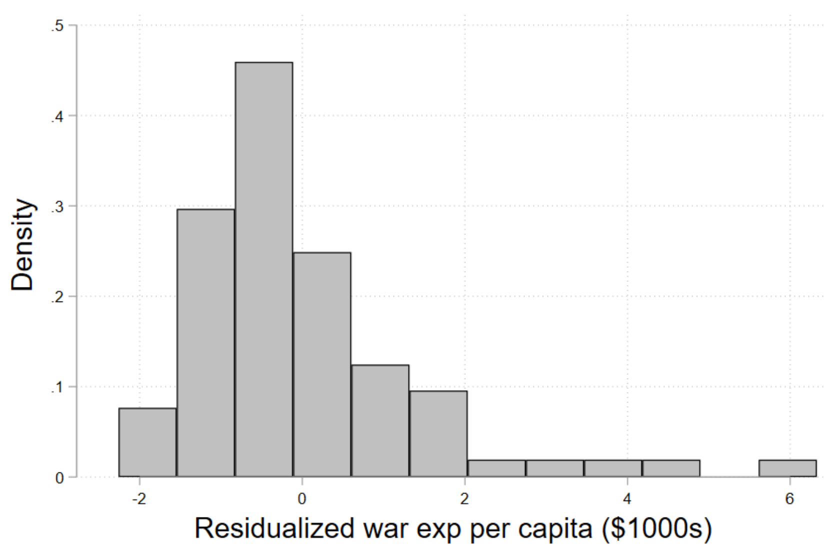
Note: Includes 146 metropolitan areas, which are county groupings based on 1950 Census definitions. The primary qualification is containing a city with population above 50,000. 55% of the population live in metropolitan areas in 1950. War expenditures per capita are total war expenditures divided by the 1940 population. Total war expenditures comes from the 1947 County Data Book. The mean war expenditure across metropolitan areas is $1,831 with standard deviation of $1,715 (1940 dollars).

Figure 2.3: Distribution of WWII expenditures per capita by metropolitan area ($1000s, 1940)

Panel A: Raw distribution

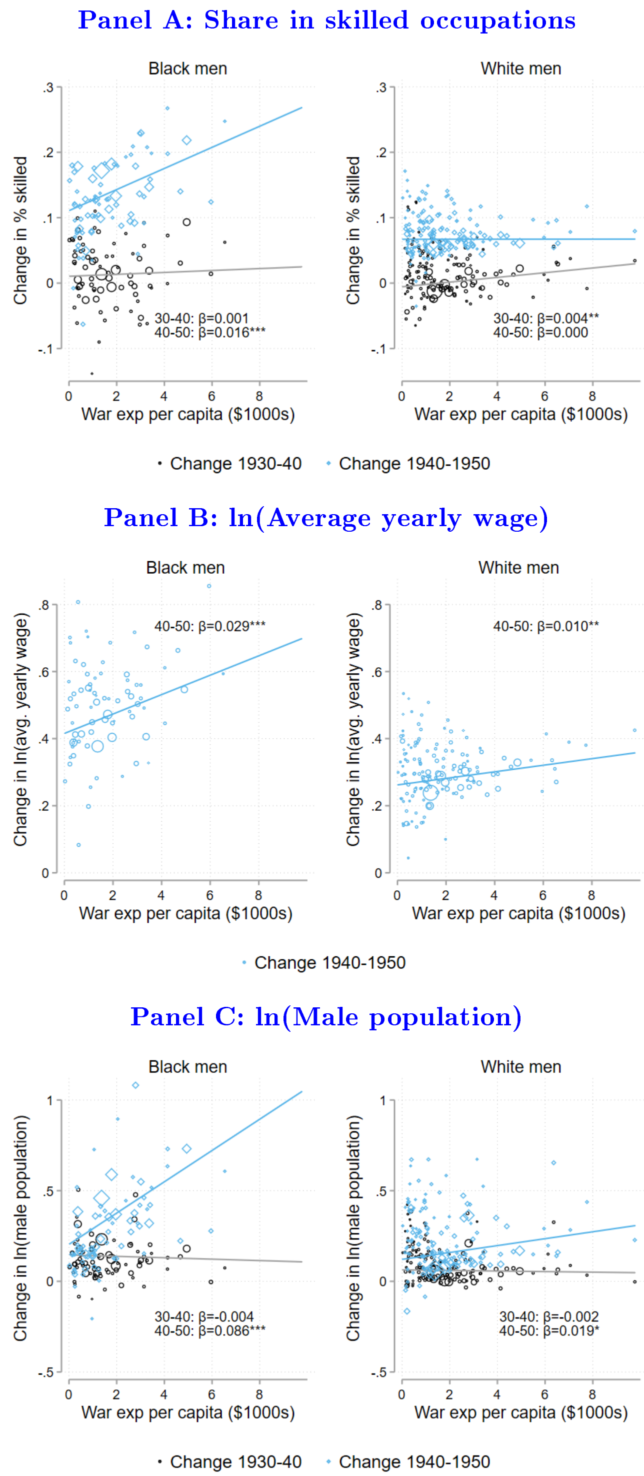


Panel B: Residualized



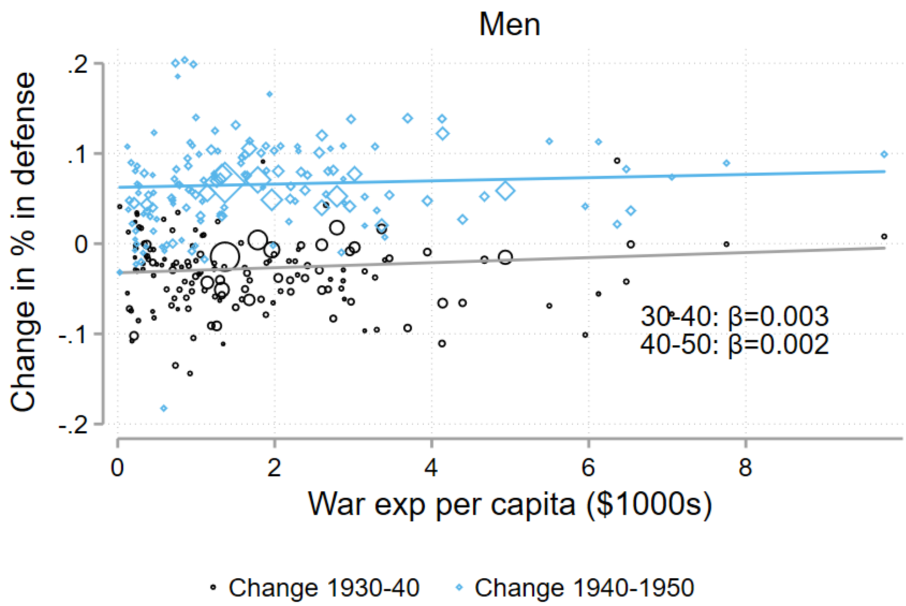
Note: Includes 146 metropolitan areas, which are county groupings based on 1950 Census definitions. War expenditures per capita are total war expenditures divided by the 1940 population. Controls include region fixed effects, share of employed men in manufacturing, in agriculture, share Black, and predicted draft rate based on demographics. Total war expenditures comes from the 1947 County Data Book. The mean war expenditure across metropolitan areas is $1,831 with standard deviation of $1,715 (1940 dollars).

Figure 2.4: Raw changes in outcomes by metro area



Note: Each point represents a metropolitan area. Metro areas with a relevant population of less than 2500 are omitted for visual clarity but are included in regressions. There is no wage data in the 1930 Census. Data is from 1930 Census (5%), 1940 Census (100%), and 1950 Census (1%) samples. Regressions are weighted by the relevant population, and robust SEs are used. *p<.1; **p<.05; ***p<.01

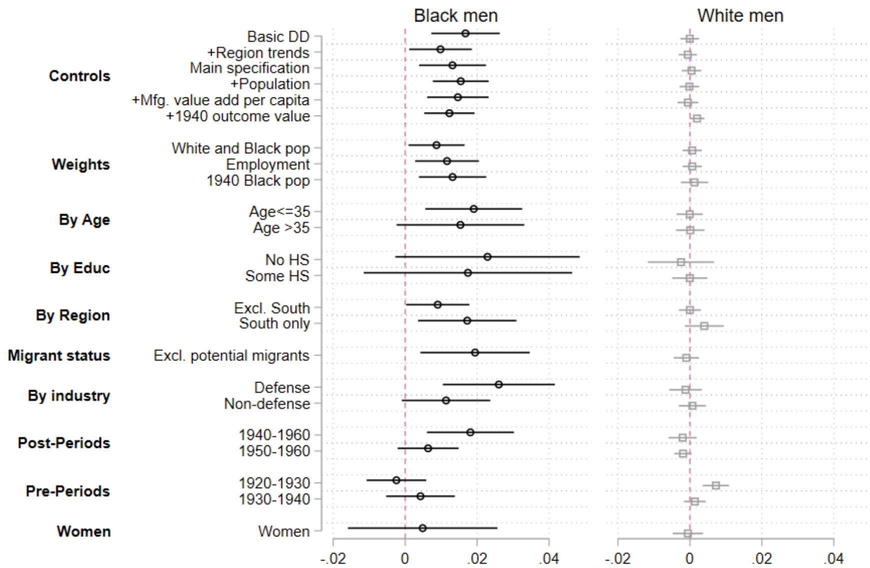
Figure 2.5: Raw changes in share of employed men in defense industries by metro area



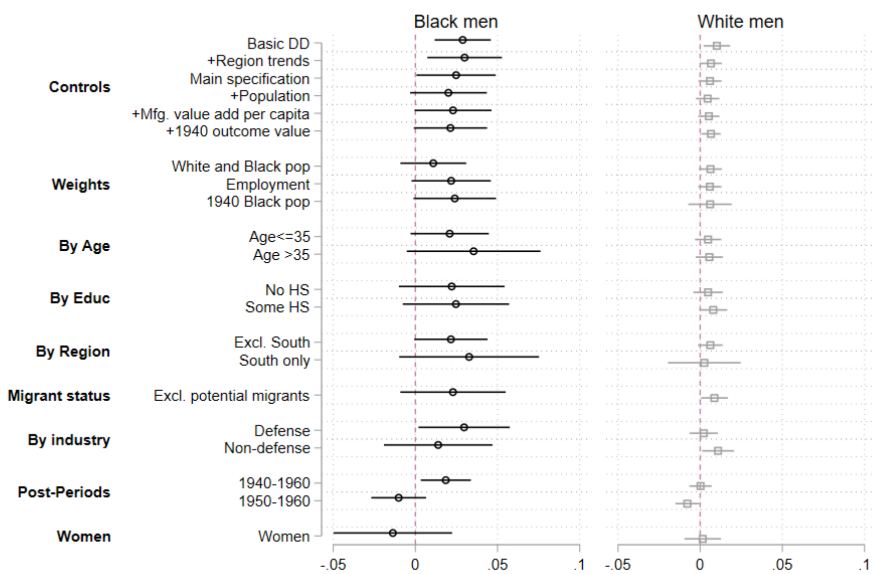
Note: Each point represents a metropolitan area. Defense industries include Mining, Manufacturing, Transportation, and Government. Data is from 1930 Census (5%), 1940 Census (100%), and 1950 Census (1%) samples. Regressions are weighted by population, and robust SEs are used. *p<.1; **p<.05; ***p<.01

Figure 2.6: Robustness of effects of war expenditures on main outcomes

Panel A: Share in skilled occupations



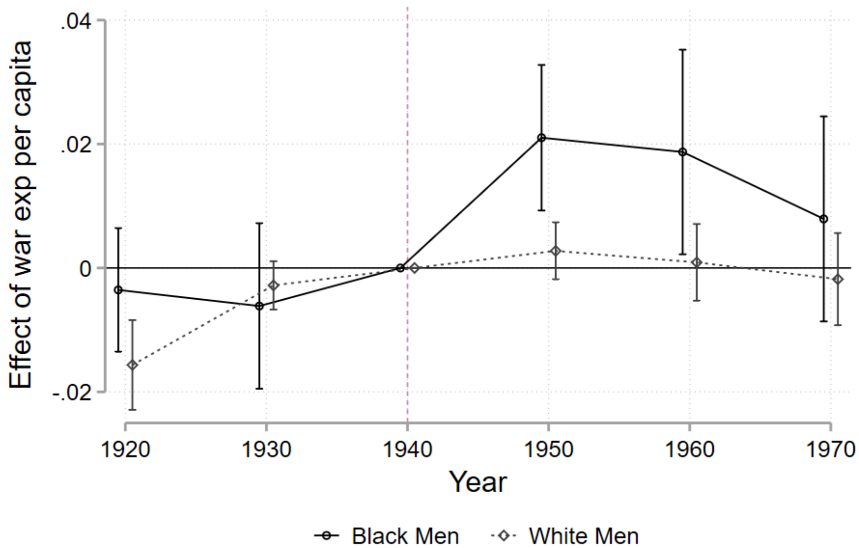
Panel B: ln(Average yearly wage)



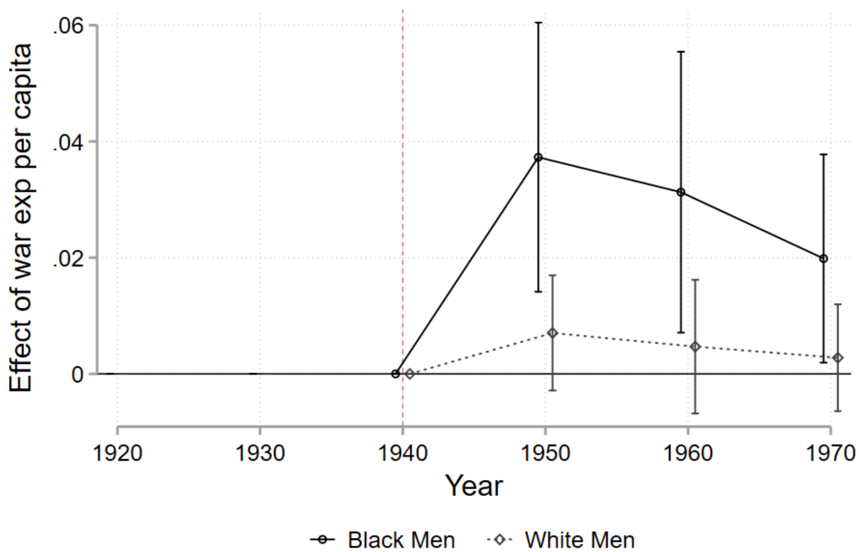
Note: See equation 2.1 for the basic specification. Intervals are 95% confidence intervals. All controls are interacted with an indicator for post. "Main specification" is our standard specification with controls for region, average years of education, share in manufacturing, share in agriculture, share Black, and predicted draft rate. "+Population" adds controls for the (log of) total population and Black population in 1940. "+1940 outcome value" adds controls for 1940 share employed, share skilled, and (log of) average yearly wage. "Excl. potential migrants" means excluding individuals in 1950 who were not born in their current state of residence and are not living with a child eight years or older born in the current state of residence. There are 146 metropolitan areas, and data comes from the 1920-1960 Census samples.

Figure 2.7: Long-term impacts of war expenditures (1920-1970)

Panel A: Share in skilled occupations



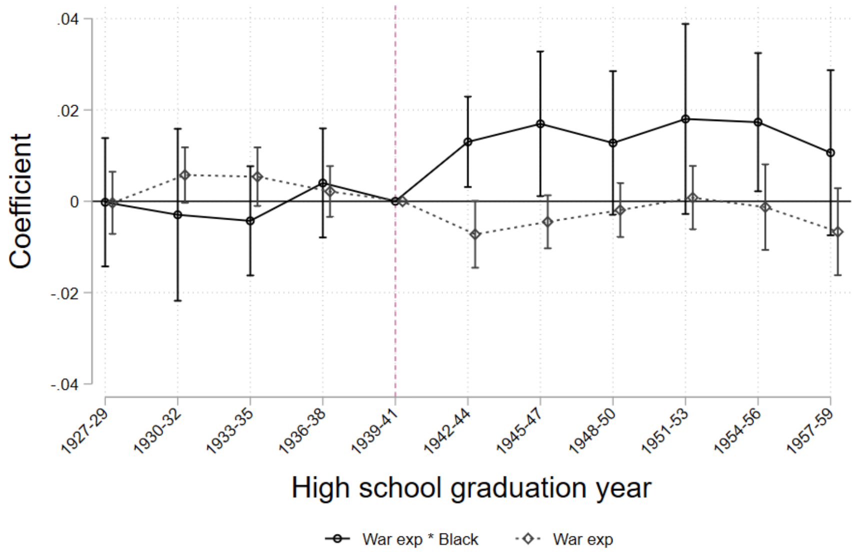
Panel B: ln(Average yearly wage)



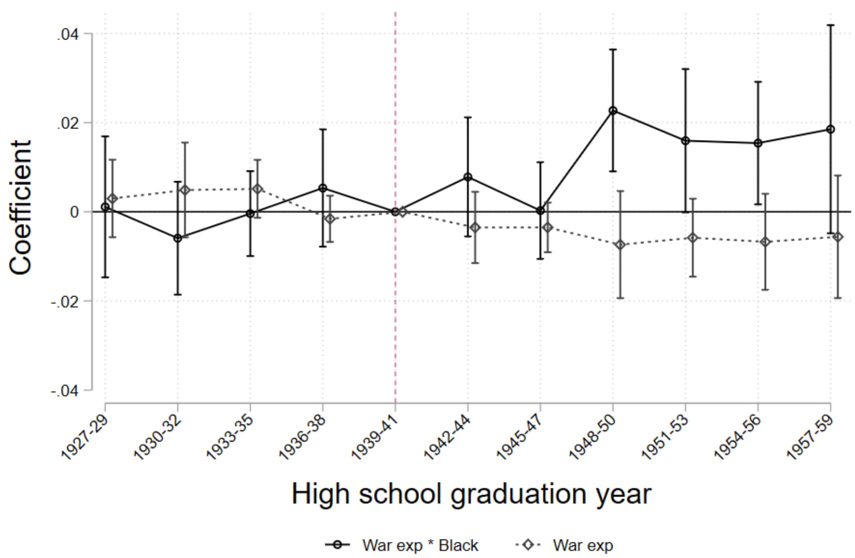
Note: See equation 2.2 for the basic specification; regressions are run separately for Black and white men, but the coefficients are plotted on the same graph. Controls include aggregate Census division, share employed in manufacturing, share employed in agriculture, share Black, and years of education in the first available year; each is interacted with a full set of year indicators; omitting the base year (1940). Commuting zone boundaries for metropolitan areas are used instead of 1940 and 1950 metropolitan area definitions due to changing metropolitan area boundaries over time. This results in some metropolitan areas being combined or dropped, leaving 135 commuting zones. Data comes from the 1920-1970 Census samples.

Figure 2.8: Effects of war expenditures on high school graduation rates

Panel A: Boys



Panel B: Girls



Note: See equation 2.3 for the estimating equation. Intervals are 95% confidence intervals. Cohorts are grouped by expected graduation year, and the sample excludes the South and individuals who are not living in metropolitan areas or who have moved to a state other than their birth state in the previous five years. Graduating high school is defined as having completed 12 years of schooling in 1960. Fixed effects include metro-race FE and cohort-race FE. Other controls interacted with race include indicators for whether born in the South interacted with race and cohort indicators. Results are similar if controls for veteran status are included. Data comes from the 1960 Census (5% sample).

Table 2.1: Predictors of per capita war expenditure

	(1) War exp per capita	(2) Months of labor shortages 1942-44
War exp per capita		0.469***
		(0.103)
Predicted draft rate	-0.174***	0.181**
	(0.064)	(0.090)
ln(Avg yearly wage)	0.012	-0.168
	(0.140)	(0.127)
% Agriculture	0.021	-0.237*
	(0.141)	(0.136)
% Government	0.078	0.277***
	(0.153)	(0.071)
% Manufacturing	0.397***	0.135
	(0.139)	(0.107)
ln(Mfg. value added per capita)	0.176**	0.104
	(0.081)	(0.099)
% Skilled	0.151	-0.067
	(0.176)	(0.147)
% Unemployed	-0.042	-0.149*
	(0.089)	(0.080)
% Black	0.039	0.146
	(0.075)	(0.115)
ln(Population)	-0.027	0.060
	(0.077)	(0.090)
Northeast	0.021	0.008
	(0.095)	(0.106)
Midwest	0.034	0.080
	(0.146)	(0.112)
West	0.111	0.365***
	(0.097)	(0.101)
R2	0.33	0.46
N	146	132

Note: An observation is a metro area, and all variables are as of 1940 and have been standardized to have $\mu = 0$ and $\sigma^2 = 1$. The denominator for percentage variables is the number of employed men except for the % unemployed for which it is the number of men in the labor force. Omitted aggregate Census division category is the South. War expenditure per capita in 1940 dollars. Months of labor shortages are percentage of months 1942-1944 with acute labor shortages according to Labor Market Reports. Only 132 of the 146 metro areas are identified in these reports. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table 2.2: Effect of war expenditures (1940-1950)

	(1)	(2)	(3)	(4)
	Black Men		White Men	
	Basic	Controls	Basic	Controls
Panel A: Share skilled				
War exp per capita * Post	0.010**	0.013***	-0.001	0.000
	(0.004)	(0.005)	(0.001)	(0.001)
Mean Y - 1940	0.33	0.33	0.77	0.77
Mean Y - 1950	0.48	0.48	0.83	0.83
Panel B: ln(Average yearly wage)				
War exp per capita * Post	0.030**	0.025**	0.007*	0.006*
	(0.011)	(0.012)	(0.003)	(0.004)
Mean Y - 1940	6.59	6.59	7.30	7.30
Mean Y - 1950	7.09	7.09	7.58	7.58
Panel C: ln(Male population)				
War exp per capita * Post	0.047***	0.042**	0.014**	0.018***
	(0.016)	(0.018)	(0.006)	(0.006)
Mean Y - 1940	10.80	10.80	13.25	13.25
Mean Y - 1950	11.22	11.22	13.36	13.36
Panel D: Prime-age employment rate				
War exp per capita * Post	-0.004	-0.005	-0.004**	-0.002
	(0.004)	(0.005)	(0.001)	(0.002)
Mean Y - 1940	0.80	0.80	0.87	0.87
Mean Y - 1950	0.84	0.84	0.92	0.92
Metro areas	146	146	146	146
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro FE	X	X	X	X
Division-Year FE	X	X	X	X
Baseline controls	-	X	-	X
Draft control	-	X	-	X

Note: Sample is 146 metro areas. See equation 2.1 for the basic specification. War expenditure is $1000s per capita. Share skilled is the share of employed men who are not farmers, laborers, or service workers. Wages are total wage earnings in the previous year for men who are currently employees. Prime-age employment is the share of men ages 25-54 who are employed. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Metro area definitions are based on 1940 and 1950 Census Bureau definitions. All values are in 1940 dollars. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table 2.3: Effect of war expenditures excluding potential interstate migrants (1940-1950)

	(1)	(2)	(3)	(4)
	Black Men		White Men	
	All	Excl. potential migrants	All	Excl. potential migrants
Panel A: Skilled occupation				
War exp per capita * Post	0.016***	0.022***	0.000	-0.001
	(0.003)	(0.005)	(0.001)	(0.001)
Mean Y - 1940	0.33	0.33	0.77	0.77
Mean Y - 1950	0.48	0.47	0.83	0.83
N - 1940	1,266,428	1,266,428	878,830	878,830
N - 1950	24,346	12,843	244,073	180,184
Panel B: ln(Yearly wage)				
War exp per capita * Post	0.026**	0.018*	0.002	0.003
	(0.010)	(0.011)	(0.003)	(0.003)
Mean Y - 1940	6.27	6.27	6.91	6.91
Mean Y - 1950	7.32	7.25	7.83	7.84
N - 1940	994,843	994,843	679,658	679,658
N - 1950	5,163	2,331	53,641	36,900
Panel C: Prime-age employment				
War exp per capita * Post	-0.003	-0.007**	-0.002	-0.002*
	(0.003)	(0.003)	(0.001)	(0.001)
Mean Y - 1940	0.80	0.80	0.87	0.87
Mean Y - 1950	0.84	0.86	0.92	0.93
N - 1940	1,164,169	1,164,169	703,935	703,935
N - 1950	20,597	10,134	184,570	135,342
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro FE	X	X	X	X
Division-Year FE	X	X	X	X
Baseline controls	X	X	X	X
Individual controls	X	X	X	X

Note: Regression at the individual level and only includes men living in one of 146 metro areas. See equation 2.1 for the basic specification. War expenditure is $1000s per capita. Excluding potential interstate migrants means excluding individuals in 1950 who were not born in their current state of residence and are not living with a child eight years or older born in the current state of residence. For employed men, a skilled occupation is defined as all occupations except farmers, laborers, or service workers. Wages are total wage earnings (1940 dollars) in the previous year for men who are currently employees. Prime-age employment is whether men ages 25-54 are employed. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Individual controls include a cubic in age, whether born in the South, and whether married. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. All values are in 1940 dollars. Regressions weighted by sampling weights. Standard errors clustered at the metro-year level. *p<.1; **p<.05; ***p<.01

Table 2.4: Effect of war expenditures on occupational upgrading for younger cohorts

	(1)	(2)	(3)	(6)	(6)	(6)
		Black Men			White men	
	1940-50	1940-50	1940-60	1940-50	1940-50	1940-60
	All	Ages 18-24	Ages 18-34	All	Ages 18-24	Ages 18-34
War exp per capita * Post	0.017***	0.020**	0.014***	0.000	0.000	-0.003
	(0.003)	(0.008)	(0.004)	(0.001)	(0.002)	(0.002)
Mean Y - 1940	0.33	0.30	0.32	0.77	0.71	0.77
Mean Y - 1950	0.48	0.46	0.51	0.83	0.80	0.82
N - 1940	1,266,428	174,103	527,464	878,830	120,470	348,777
N - 1950	24,346	3,697	45,740	244,073	31,266	165,113
Mean war exp per capita	1.83	1.83	1.83	1.83	1.83	1.83
Metro FE	X	X	X	X	X	X
Division-Year FE	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X
Draft control	X	X	X	X	X	X
Indiv controls	X	X	X			

Note: Regression at the individual level and only includes men living in one of 146 metro areas. See equation 2.1 for the basic specification. For employed men, a skilled occupation is defined as all occupations except farmers, laborers, or service workers. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Individual controls include a cubic in age, whether born in the South, and whether married. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites), 1950 (1%), and 1960 (5%; 40% sub-sample for whites) Census samples. Regressions weighted by sampling weights. Standard errors clustered at the metro-year level. *p<.1; **p<.05; ***p<.01

Table 2.5: Estimation of model shocks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Labor	Adj. income		Adj. labor		Adj. income		Adj. labor
	$\ln \pi^L_{riogt}$	riot	rit	r	r	riot	rit	r
β_1			0.004				0.005	
			(0.010)				(0.030)	
β_2		0.010				0.004	0.009	
		(0.029)				(0.029)	(0.030)	
β_3	0.014					-0.030	0.011	
	(0.035)					(0.034)	(0.039)	
β_4	0.115***					0.117***	0.103***	
	(0.040)					(0.038)	(0.039)	
β^U_1					-0.018**			-0.017**
					(0.007)			(0.007)
β^U_2				0.054**				0.042**
				(0.022)				(0.021)
Observations	3,494	3,622	3,622	3,622	3,622	3,622	3,622	3,622
R-squared	0.991	0.990	0.988	0.995	0.992	0.990	0.988	0.992
γ_{riog}	X	X	X	X	X	X	X	X
γ_{iogt}	X	X	X	-	-	X	X	-
γ_{riot}	X	-	-	-	-	-	-	-
γ_{rgt}	X	-	-	-	-	-	-	-
γ_{rit}	-	X	-	-	-	X	-	-
γ_{gt}	-	-	-	X	X	-	-	X
γ_{rt}	-	-	-	X	-	-	-	-
Draft control	-	-	-	-	X	-	-	X

Note: An observation is an r, i, o, t, g cell and only includes individuals living in metro areas. Columns 1 estimates (2.9) using employment to measure allocations across io within r. Columns 2 and 6 estimate (B.42) and (B.41); the column heading riot refers to the dependent variable being an adjusted measure of labor income at the *riot* level. Columns 3 and 7 estimate (B.45) and (B.44); the column heading rit refers to the dependent variable being an adjusted measure of labor income at the *rit* level. Column 4 estimates (2.11), column 5 estimates (B.39), and column 8 estimates (B.38). Predicted draft rate is based on 1940 demographics. Primary data sources are 1940 (100%) and 1950 (1%) Census samples. Regressions weighted by cell population. Standard errors clustered at the metro-year level. *p<.1; **p<.05; ***p<.01

Table 2.6: Ability of war expenditure shocks to explain aggregate changes in race gaps

	(1)	(2)	(3)	(4)	(5)
	Actual	Model		Model - No migration	
	Total	All shocks	Anti-discrim. only	All shocks	Anti-discrim. only
Change in black-white gap in share skilled:					
Δ1940-50	-0.082	-0.027	-0.026	-0.023	-0.023
% explained		*32.6%*	*32.1%*	*28.5%*	*28.3%*
Change in black-white gap in ln(avg. yearly wage):					
Δ1940-50	-0.226	-0.056	-0.058	-0.038	-0.040
% explained		*25.0%*	*25.5%*	*16.9%*	*17.8%*

Note: Sample are men living in one of 146 metro areas. Column 1 is the actual change in the gap between Black and white men. Column 2 gives the change in the gap due to all war expenditure shocks. Column 3 gives the change in the gap due to anti-discriminatory shocks. % explained is the percent of the actual change that can be explained by the given shocks. Columns 4 and 5 repeat columns 2 and 3 except with the migration channel removed. Share skilled is the share of employed men who are not farmers, laborers, or service workers. Wages are total wage earnings (1940 dollars) in the previous year for men who are currently employees. Primary data sources are 1940 (100%) and 1950 (1%) Census samples.

Table 2.7: Effect of war expenditures on school expenditures and residential segregation (1940-1950)

	(1)	(2)	(3)	(4)
	Education		Residential segregation	
	ln(Exp. per capita)	$\ln(\frac{Blackstudents}{Blackteachers})$	Dissimilarity	Isolation
Panel A: All metros				
War exp per capita * Post	0.010	0.022	-0.003	0.001
	(0.009)	(0.047)	(0.007)	(0.024)
N	242	198	86	86
Panel B: Excluding South				
War exp per capita * Post	0.011	0.095	-0.008	-0.022
	(0.010)	(0.072)	(0.006)	(0.026)
N	172	98	60	60
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro FE	X	X	X	X
Division-Year FE	X	X	X	X
Baseline controls	X	X	X	X
Draft control	X	X	X	X

Note: Full sample is 146 metro areas; educational expenditures is available for only 121 metro areas. See equation 2.1 for the basic specification. Residential segregation indices are from Cutler et al. (1999) and are only available for 43 of our metro areas. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Regressions weighted by Black population; results are similar if unweighted estimates are used. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table 2.8: Labor shortages, war expenditures, and the share of workers in skilled occupations (1940-1950)

	(1)	(2)	(3)	(4)	(5)	(6)
		Black men			White men	
War exp per capita * Post	0.013**		-0.001	-0.000		0.003**
	(0.005)		(0.008)	(0.001)		(0.001)
Labor shortage % * Post		0.062***			-0.007	
		(0.012)			(0.006)	
War exp per capita * Labor shortage % * Post			0.016**			-0.003**
			(0.007)			(0.001)
Observations	252	252	252	264	264	264
Mean war exp per capita	1.93	1.93	1.93	1.93	1.93	1.93
% of months with labor shortage	0.21	0.21	0.21	0.21	0.21	0.21
Metro FE	X	X	X	X	X	X
Division-Year FE	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X
Draft control	X	X	X	X	X	X

Note: The outcome is share of employed Black men in skilled occupations. Share skilled is the share of employed men who are not farmers, laborers, or service workers. Sample is 132 metro areas with data on labor shortages. Months of labor shortages are percentage of months 1942-1944 with acute labor shortages according to Labor Market Reports. War expenditure is $1000s per capita. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Metro area definitions based on 1940 and 1950 Census Bureau definitions. All values are in 1940 dollars. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table 2.9: Skill upgrading and direct vs. indirect expenditure (1940-1950)

	(1) Direct demand	(2) Indirect demand	(3) Both
Direct value add * Post	0.078***		0.073***
	(0.021)		(0.022)
Indirect value add * Post		0.070	0.028
		(0.043)	(0.043)
Mean Y - 1940	0.33	0.33	0.33
Mean Y - 1950	0.48	0.48	0.48
Metro areas	146	146	146
Mean war exp per capita	1.83	1.83	1.83
Metro FE	X	X	X
Division-Year FE	X	X	X
Controls	X	X	X

Note: The outcome is share of employed Black men in skilled occupations. Share skilled is the share of employed men who are not farmers, laborers, or service workers. See Appendix Section B.3.4 for a discussion of how the direct and indirect value added measures are created. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Metro area definitions based on 1940 and 1950 Census Bureau definitions. All values are in 1940 dollars. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

CHAPTER 3

Acquisitions, Working Conditions, and Convergence: Evidence from Nursing Homes

3.1 Introduction

There were over 14,000 mergers and acquisitions in the US during 2019.[1] Despite this, there is limited research on the effect of mergers on workers. Existing research typically finds mixed or inconclusive average effects (Li, 2012). However, there is growing research that worker outcomes vary systematically across firms (e.g., Card et al., 2018). There is also growing evidence that firms frequently standardize policies across locations (DellaVigna and Gentzkow, 2019). This work suggests that the effect of acquisitions might vary significantly across acquiring firms and acquired facilities.

There has been increasing consolidation through mergers and acquisitions in healthcare markets (Gaynor, Ho and Town, 2015). These acquisitions might harm stakeholders such as workers or patients; for example, Eliason et al. (2020) shows that dialysis patient outcomes significantly worsen post acquisitions. The effect of acquisitions on working conditions is an important questions not only for workers but also for patients since there is a strong relationship between working conditions and quality of care (Halm, 2019).

Specifically, I look at how acquisitions of California nursing homes affect working conditions. Nursing homes are a key healthcare sector, with spending of $166 billion in 2017 (Hackmann and Pohl, 2018). I use a leads and lags approach to identify the changes

[1] According to the consultancy PwC.

in working conditions after acquisitions. This approach controls for any time invariant characteristics of the facility and allows for visual inspection of pre-treatment trends. My data allows the examination of a wide range of outcomes beyond wages – such as benefits, workload, turnover, prices, quantities, and patient outcomes.

First, I look at the average effects of the acquisitions. I find that there is almost no effect on average wages after adjusting for shifts in occupational composition. However, focusing solely on average wages hides heterogeneity across occupations and outcomes. Low wage occupations see slight decreases on average (1-2%), while administrative and management positions have substantial wage increases (about 5%). Additionally, benefits are reduced by almost 6%, nursing workload increases by 3-4%, and, most significantly, turnover increases by 15%. Moreover, these effects do not dissipate with time. Patient outcomes worsen after acquisitions, even conditional on changes in the aggregate patient mix. I show that the negative effect on patients is primarily due to the deterioration of quality in facilities that initially peformed relatively better than the acquiring firm. On the other hand, patient outcomes do not deteriorate in lower-performing facilities and there is suggestive evidence that these facilities might improve slightly with fewer deficiency citations.

There is no reason to believe that most acquisitions have close to average treatment effects. Instead, there is substantial reason to believe that there is significant heterogeneity. I hypothesize that the working conditions in acquired facilities will converge to the acquiring firm, either due to standardization of production processes or due to behavioral factors. The results confirm that working conditions quickly converge toward those of the acquiring firm. For most outcomes at least 50% of the initial gap between the acquired facility and acquiring firm disappears within the first two years. Moreover, these estimates likely understate convergence due to measurement error and compositional differences. This means that the effect of an acquisition depends significantly on relative working conditions.

I combine the actual gaps in working conditions prior to acquisition with the estimated degree of convergence to calculate the distribution of expected treatment effects. This exercise shows there is substantial variation in the effect of acquisitions. Take the example of benefits. The average effect is a 6% decrease, but workers at the 10th percentile of expected facility treatment effects would expect a 25% decrease in benefits while workers at the 90th percentile would expect a 20% increase. Combining expected changes in wages, benefits, and workload implies that workers at the 10th percentile of acquisitions expect to see losses equivalent to a 12% wage cut while those at the 90th percentile see gains equivalent to a 7-8% increase in wages.

Finally, I provide suggestive evidence that behavioral factors that lead to policy standardization might play a role in the convergence of working conditions. This evidence is consistent with a growing literature showing that behavioral factors and managerial inertia play an important role in firms' pricing decisions (e.g., DellaVigna and Gentzkow, 2019).

Identification rests on a standard parallel trends assumption – in the absence of acquisition the facility would have experienced a similar change in outcomes as the control facilities. I use non-acquired facilities from the same state (California) and industry (nursing homes) as the control group. Additionally, I use county-year fixed effects, so results are based on comparisons within county and year. Visual inspection of the leads and lags reveals no significant pre-treatment trends or "dips" in the outcome variables. I also use a cross-validated Least Absolute Shrinkage and Selection Operator ("LASSO") to identify any levels or trend variables that are strong predictors of treatment; the only consistently strong predictor of treatment is the level (and not the trend) of the markup over variable cost. Therefore, it seems that the parallel trends assumption is reasonable. Results are also robust to conditioning on any pre-existing trends in key outcome variables.

These results show that acquisitions represent significant risks to workers. Some acquisitions result in large gains, whereas others result in large losses. While workers have the option of leaving the job after an acquisition, there is a long literature showing that job loss can have substantial negative long-term consequences for workers (e.g., Schmieder, von Wachter and Heining, 2019). Additionally, there might be efficiency losses if the threat of acquisitions reduces the ability to form implicit contracts (Shleifer and Summers, 1988). Finally, it is important to understand the effects of acquisitions on working conditions in healthcare since they affect patient care paid for by government payors.

This paper primarily contributes to two strands of literature. There is a small literature on the effect of acquisitions on workers. The existing studies typically focus only on average wage and employment effects and have found mixed results; some have found negative effects (Lichtenberg and Siegel, 1990; Li, 2012; He and le Maire, 2020), others positive effects (McGuckin, Nguyen and Reznek, 1998) or no consistent effects (Brown and Medoff, 1988). The most directly related paper, Currie, Farsi and Macleod (2005), looks at California hospital acquisitions and finds little impact on wages but an increase in nurse workload. I show that the focus on average wages misses other important margins of adjustment – such as benefits, workload, and turnover. There is substantial convergence in working conditions, resulting in significantly different impacts across acquisitions and potentially explaining the mixed results in the literature. To my knowledge, no previous studies have documented this convergence. The nursing home setting is important to this result since the production process is highly standardized across facilities, allowing more direct comparison of working conditions.

There is also a separate literature that focuses on healthcare acquisitions. There is mixed evidence on whether hospital acquisitions reduce costs (Schmitt, 2017). Eliason et al. (2020) find that acquired dialysis facilities replace nurses with technicians, increase patient loads, and have worse patient outcomes. They also find that acquired facilities quickly adopt acquiring firm strategies, for example their usage rate of the drug Epogen. Two recent working papers have focused more specifically on acquisitions of nursing homes by private equity firms. Gupta et al. (2021) find that private equity acquisitions substantially increase patient mortality, potentially due to decreased nursing staffing. Gandhi, Song and Upadrashta (2021) find that PE acquisitions makes facilities more sensitive to competition with larger RN staffing increases in more competitive markets. I add direct evidence on the impact on working conditions, including the importance of relative working conditions. This paper also adds to the growing evidence that acquisitions do not typically improve patient outcomes.

The rest of the article proceeds as follows. Section 2 discusses the existing literature on acquisitions and why they might affect workers. Section 3 provides additional context on nursing homes, introduces the data, and looks at what type of nursing home is acquired. Section 4 introduces the empirical approach and provides the average effects. Section 5 discusses heterogeneity, the importance of convergence, and suggestive evidence for the mechanism. Section 6 concludes. The Appendix contains additional information on the data, evidence on the importance of unobserved worker quality, estimates for the marginal willingness to pay for workplace conditions, and additional robustness checks and descriptive statistics.

3.2 Potential effects of acquisitions

There is a long literature on the effect of acquisitions on firm output. This research typically finds that firm-level revenue productivity increases (e.g., Siegel and Simons, 2010); however, the reason for revenue productivity increases is less clear. Several recent papers have suggested that revenue increases are not due to increases in plant-level technical productivity.[2] Instead Braguinsky et al. (2015) find that managerial improvements (such as capacity utilization and demand management) explain revenue productivity increases in their sample of Japanese textile firms. Blonigen and Pierce (2016) find that markups increase in their sample of US manufacturing plants but find little evidence for improvements in plant-level productivity. Davis et al. (2014) find that most TFP improvements from private equity acquisitions are due to closures of low productivity plants or the opening of new higher productivity plants rather than within-plant improvements.

Another important strand of research is on what type of firms are acquired. David (2020) documents several empirical facts: (1) acquirers are relatively larger and more profitable than their targets, (2) there is positive assortivity with higher productivity firms buying more productive targets, (3) acquirers are larger and more profitable, but the average target

[2] By technical productivity I mean productivity conditional on operating.

firm is similar to the median firm. These facts are generally consistent with others in the literature (Eckbo, 2014).

There is less research on the impact on workers, but there are at least several reasons to believe that acquisitions might effect workers. First, acquisitions might change productive technology. This technology change could be factor neutral or it could be factor biased – for example, a firm could implement new equipment allowing them to more effectively use lower skill workers. Factor neutral improvements would likely weakly increase compensation and employment while factor biased changes could either improve or worsen working conditions based on the firm's incentives to adjust quality.

Second, behavioral factors might be important. Managerial inertia due to the costs of choosing optimal policies could result in standardized working conditions within a firm. Or fairness concerns might lead to internal pay equality. There is recent evidence that firms often use uniform pricing schemes that do not respond to local shocks; for example, retail firms often rely on uniform pricing strategies across locations despite significant revenue costs (DellaVigna and Gentzkow, 2019) and grocery stores do not adjust prices after the entrance of Wal-Mart, despite revenue declines of 16% for stores within one mile (Arcidiacono et al., 2020). DellaVigna and Gentzkow (2019) also show that acquired grocery stores immediately adjust their prices to track typical prices in the acquiring firm. There is some evidence that similar dynamics might exist in labor markets; for example, minimum wage shocks to the headquarters of multinationals are transmitted to foreign establishments (Hjort, Li and Sarsons, 2020). The effect of an acquisition will then depend significantly on whether the acquirer's standard policies are better or worse than pre-existing working conditions.

Third, acquisitions might allow for rent transfers or managerial discipline. Shleifer and Summers (1988) hypothesize that a significant portion of the returns to hostile acquisitions could be due to rent transfers. They theorize that firms install credible managers who are able to make implicit contracts with workers (for example, compensation to losers of promotion tournaments). Another alternative is that managers might share rents with workers to enjoy the "quiet life" (Bertrand and Mullainathan, 2003). In either case, acquisitions allow the acquiring firms to appropriate rents. He and le Maire (2020) find that Danish firms with more generous managers are more likely to be acquired and reduce wages, supporting these theories. This mechanism predicts that acquisitions would worsen working conditions.

Fourth, acquisitions can affect input market power. There is recent evidence that mergers that lead to significant increases in concentration decrease wages (Arnold, 2020). For example, Prager and Schmitt (2021) find that hospital acquisitions that increase local concentration decrease wages for workers with industry-specific skills. The predicted effects would be negative for workers. Since most acquisitions in this setting are small, it is unlikely that individual acquisitions affect input market power in this setting.[3]

Both the changes to productive technology and the behavioral factors have ambiguous predictions about the effect on workers. These predictions are consistent with the mixed results in the limited existing evidence; however, they also suggest that there might be important heterogeneity in the effect of acquisitions. The key insight of both potential mechanisms is that the impact on workers likely depends on the working conditions in the acquired facility relative to the acquiring firm.[4] Working conditions likely converge to those of the acquiring firm. An important advantage of my setting is that nursing homes have standardized production processes, allowing for a more direct comparison of working conditions across facilities within a firm.

The goal of this paper is not to definitively distinguish between these mechanisms, but I will provide some suggestive evidence. The rent transfer story is primarily about working conditions in the acquired firm – for high rent facilities we should expect working conditions to deteriorate and for others we would expect little change. Therefore, significant convergence of working conditions in both directions suggests that changes to productive technology or behavioral factors are important.

[3]This assertion does not mean that there is not input market power, just that the impact of each individual acquisition is likely very small. Most facilities are located in large metro areas with a significant number of alternative facilities. The largest chain has less than 10% of total facilities.

[4]For the case of productive technology changes, the assumption is that the acquired facility adopts production technologies similar to that used by other facilities in the acquiring firm.

It is more difficult to distinguish between changes to productive technology and behavioral factors since it requires strong assumptions about what the "optimal" choice is. However, I will offer several pieces of suggestive evidence that behavioral factors might be important. First, if behavioral factors are important we might expect stronger convergence for choices where decisions are made less frequently.[5] For examples, facilities hire RNs much less frequently than nursing assistants so facilities might be less aware of RN market compensation. Second, we might be likelier to see mixed cases – where some working conditions improve and others worsen.[6] Third, the wages at one facility in a firm will be affected by local labor market conditions at other facilities even if they are geographically separate markets. None of these tests are conclusive, but they do suggest that behavioral factors play a role.

3.3 Background on nursing homes

3.3.1 Introduction to nursing homes

Nursing homes, or skilled nursing facilities, provide 24-hour skilled nursing care to residents. They employ licensed professionals and provide significantly more medical care and assistance than assisted living facilities. According to the CDC, there was an average of 1.3 million nursing home residents during 2015 and 1 million employees. The typical resident is over 65 and is recuperating from illness/surgery or is chronically ill and needs regular nursing care. For example, almost 50% of patients were diagnosed with either dementia or Alzheimer's. Most patients are admitted directly from hospitals. There is significant variation in lengths of stay; 43% stay less than 100 days, but there are many patients who will spend the rest

[5] The idea is that there might be some fixed cost to making a decision. For infrequent decisions there is a low benefit to paying the fixed costs. The fact that the decision is made less frequently might also mean the agent has less information that can be used to make the optimal choice.

[6] For a generic increase or decrease in quality we would expect convergence in the same direction. But split convergence can still be explained by productive function changes; for example, the firm could have a relative cost advantage in providing benefits so they consistently shift compensation from wages to benefits.

of their lives in nursing homes. Nursing homes have standardized production processes that typically provide similar services, with the exception of some specialized facilities. I exclude specialized facilities from my main analysis, including sub-acute care facilities, intermediate care facilities, and multi-level retirement communities.

Nursing homes are generally large, with an average size of over 100 beds and 120 employees. These facilities are usually privately operated – 88% of California facilities were for-profit in 2015. However, facilities rely on public payors, with nearly all facilities accepting both Medicare and Medicaid patients. Medicaid pays for 60% of patient days and Medicare pays for another 15% of days. Since 1998, Medicare has used a prospective payment system that adjusts reimbursements for the patient case mix and local wage indices. Medicaid reimbursement varies across states. Prior to 2005, California used a flat-rate reimbursement system that only varied with coarse geographical groupings. Since 2005, California has relied on a complex formula to derive facility specific rates that depend on historical facility costs with cost caps based on peer groups.

Medicare reimburses at a significantly higher rate than Medicaid (3-4x); however, it fully covers only the first 20 days and stops all coverage after 100 days. Medicaid will cover any length of stay for qualifying patients. Over one stay, a patient might use multiple payors; for example, they can begin on Medicare, then transition to private pay, then switch to Medicaid once their assets are low enough to qualify. Nursing homes have a strong incentive to target Medicare patients due to their high reimbursement rates. Nursing homes cannot legally discriminate between patients based on payor type; however, there is evidence that rationing occurs, especially when facilities are capacity constrained (Hackmann and Pohl, 2018; Gandhi, 2019).

The market is localized since patients usually choose nearby facilities. Facilities have limited ability to compete on price since most patient days are covered by government payors at fixed rates. Therefore, facilities primarily compete on quality of care. Many studies have shown a robust relationship between staffing and quality (Harrington et al., 2016). For example, Friedrich and Hackmann (2021) showed that a 12% decrease in RNs led to a 13%

increase in mortality for patients over 85. Nursing homes increase RN/LVN staffing when Medicaid reimbursement rates are increased in order to compete for patients (Hackmann, 2019). However, this competition is not strong enough to prevent the median home from having staffing levels that are below recommended standards.

Another important determinant of quality is staffing turnover. Turnover is extremely high in nursing facilities; Gandhi, Yu and Grabowski (2021) find a mean turnover rate of 128% and median of 94% for nursing staff nationwide. Loomer et al. (2021) find a strong correlation between turnover and infection control citations. Antwi and Bowblis (2018) instruments for turnover with local unemployment rates and finds that a 10 ppt increase leads to a 20% increase in deficiency citations. Finally, worker effort is an important element of quality. Higher wages or benefits might incentivize effort through standard efficiency wage arguments (Shapiro and Stiglitz, 1984).

3.3.2 Nursing home workforce

The core of the workforce (about 65% of total hours worked) is the nursing staff. The nursing staff is composed of three groups; registered nurses (RNs), licensed vocational nurses (LVNs), and nursing assistants (NAs). RNs typically have 2-4 years of training and LVNs 2 years. NAs complete short courses or receive on-the-job training. This skill differentiation is reflected in wages and worker demographics (see Appendix Table C.1). Other important occupations include food preparation (10% of total hours), cleaning and laundry (7%), and administrativon(8%). The remaining hours are from a mix of jobs, such as running social activities, maintenance, training, management (primarily directing nurses), and physical therapy.

Nursing assistants are responsible for most nursing care, with 60-70% of total nursing hours worked. Immigrants and women are both highly overrepresented among nursing assistants. Barriers to entry are limited, and many nursing assistants do not even have high school degrees. Their typical outside options include jobs as retail workers or home health aides. The work is labor intensive and primarily consists of assisting patients with activities of daily living (ADLs). For example: helping patients eat, repositioning bedridden patients, and assisting with personal hygiene. Appendix C.2 shows that there are limited returns to observable and unobservable measures of worker quality. Therefore, workers seem relatively homogenous in terms of fixed quality. These findings are consistent with the fact that facilities typically have standardized pay scales for workers.[7] Compensation differences across facilities might still reflect variation in quality due to endogenous effort. Finally, job separations are typically worker-initiated; Castle et al. (2007) find only 6% of separations are involuntary.

Working conditions other than wages are also important to nursing assistants. Many surveys have documented that nursing assistants care about the quality of care they provide (Castle et al., 2007). The most frequently reported reasons by nursing assistants for wanting to leave their job include poor pay, bad working conditions, having too many residents to care for, and low benefits (Squillace et al., 2008). In the same survey, the most frequently reported reasons for disliking the job involved co-workers (30% of respondents), workload (26%), and supervisors (23%).

In later analysis, I use two different measures of workload. First, I use the total hours per resident day (HPRD) of nursing. Fewer nursing hours per patient implies a higher workload for the nursing staff. Second, I use a version that is adjusted for patient severity. Some patients require significantly more time than others, e.g., if a resident is unable to feed themselves. I divide the required hours of nursing based on patient severity by the actual hours provided to get a workload measure.[8] Note that there are limited ways for facilities to improve their technical productivity without sacrificing quality – the technology of feeding or bathing patients does not vary significantly across facilities. Therefore, the primary way to improve efficiency is to reduce quality by omitting care. So facilities with lower staffing

[7]Based on conversations with an industry expert.

[8]Required hours are from CMS calculations based on studies of the standard nursing time provided to different types of patients.

require either more work per hour of nursing or more care omissions, both of which are disamenities to workers.

Finally, turnover is a good measure of the overall job quality in a simple search framework. In Appendix C.1, I document that lower pay and higher workload are associated with higher turnover. Based on these facts, I will examine the effect of acquisitions on wages, workload, benefits, and turnover. In contrast most other studies are limited to looking at wages. These other margins of adjustment will end up being empirically significant.

3.3.3 Nursing home chains

The nursing home industry is significantly more fragmented than many other healthcare sectors. Since there is very limited construction of new facilities, the primary way for chains to grow is to acquire existing facilities. Figure 3.2 shows that about 40-50 nursing homes change ownership every year in California (3-5% of the total).[9] Despite these acquisitions, most nursing homes are still part of smaller chains, as seen Figure 3.1.

Nursing home chains potentially offer advantages of economies of scale, service standardization, knowledge transfers, management expertise, and risk sharing. Chains might also benefit from the ability to conduct active marketing campaigns to attract patients (Harrington et al., 2011). Despite these potential advantages, there is limited evidence that chain nursing homes actually lower costs and there is suggestive evidence that they have worse patient outcomes (Harrington et al., 2012). One potential issue is that nursing homes have relatively homogenous production processes with limited scope for technological improvements.

Nursing home chains frequently standardize policies, such as working conditions. Figure 3.3 shows the interquartile range for working conditions across facilities for some of the largest California chains. There are noticeable differences across chains, with some chains

[9] This figure is consistent with Maksimovic and Phillips (2001), who find that about 4% of large manufacturing plants change ownership each year.

offering significantly higher staffing rates than others. Patient outcomes are also correlated across facilities in major chains; however, chain owners argue that these facilities were already having issues when they were acquired.

Facilities often obscure their true ownership. Owners separate property and facility management into separate LLCs and rely on management companies. Larger chains typically have a complex network of LLCs to own and operate their facilities. This structure is for legal reasons and not operational ease – the goal is to protect company assets from litigation (Harrington et al., 2011). Many chains do not even use branded names across their facilities or disclose on their websites which facilities they own. As an example, Figure 3.4 shows the 2015 ownership network of the largest California chain, Brius. The ownership network is a complex web of holding companies and even the naming conventions are not consistent. These issues extend to data reporting. In federal OSCAR / CASPER data only 52% of facilities in CA self-report being part of a chain in 2015. Detailed inspection of California data suggests that the true share is at least 75%.

Nursing home regulation typically focuses on individual facilities rather than chains.[10] For example, the federal website Five Star Quality has quality metrics only for individual facilities and does not provide information about related faciliites. Many quality measures are noisy; if these outcomes and inputs are strongly correlated within chains, then providing information on related facilities can improve consumer choice. Some chains are also repeat offenders or systematically misreport their quality metrics. These significant quality of care issues have received increased attention due to the COVID crisis. These issues have prompted sensational reports, such as a recent New York Times article entitled "Maggots, Rape and Yet Five Stars: How U.S. Ratings of Nursing Homes Mislead the Public".[11]

[10]According to the Sacramento Bee newspaper in 2015, "But in California, the agency charged with overseeing these skilled-nursing facilities, the Department of Public Health, makes no effort to measure quality of care throughout a chain, or determine whether corporate policies and practices are contributing to any patterns."

[11]https://www.nytimes.com/2021/03/13/business/nursing-homes-ratings-medicare-covid.html

3.3.4 Data and predictors of acquisition

One key advantage of studying nursing homes is the extremely detailed data available from regulatory filings. First, each California nursing home has to file yearly financial reports, which are audited by the state. These reports include data on revenues and patient days by payor, balance sheets, detailed expenses (including benefits), hours and salary by occupation, ownership stakes, and much more. These reports are for each individual facility and not for each firm. Individual facility filings allow me to track facilities before and after acquisition. I downloaded the individual reports for 1996-2019 from the California Office of Statewide Health Planning and Development's website and standardized the data across years. The working condition outcomes – hourly wages, staffing, benefits, and turnover – come primarily from this data. I use the information on ownership to identify acquisitions, which I will discuss in more detail later. This data is supplemented with facility utilization reports that provide data on patient admissions and discharges.

The other main set of data comes from national regulatory filings as compiled and standardized by Brown's LTCFocus group.[12] Their data comes primarily from: the Minimum Data Set, which is data from individual resident assessments; Medicare claims data; OSCAR/ CASPAR, which contains data from annual certification visits as well as cited deficiencies; and Nursing Home Compare / Five Star Quality, which are quality guides intended for consumer use. I supplement this data with cited deficiencies from California regulators.

There is substantial misreporting in ownership. Additionally, sometimes the nominal owners of the facility are not the true operators. In order to identify true ownership changes and firms I rely on a combination of manual matching and flags. First, I manually classify ownership for all facilities in 2015. Next, for the three largest chains, I identify when each facility was acquired.[13] I then flag major changes in a set of variables. These variables

[12]LTCFocus is sponsored by the National Institute on Aging (1P01AG027296) through a cooperative agreement with the Brown University School of Public Health.

[13]For simplicity, I refer to ownership changes as acquisitions. Ownership changes could also be due to mergers; however, mergers do not seem to be common in this setting.

include whether the names of owners, directors, facility, or report filing contact changes and whether the zip code of the parent organization or contact phone number changes. In any given year, a few of these variables might change; if enough of these change at once then I flag it as an acquisition. This procedure performs very well when compared against the largest three chains that I manually classified; however, any misclassified acquisitions would likely bias results downward. While this approach theoretically identifies all acquisitions, I have complete firm networks only for a subset of firms and years.

Table 3.1 presents summary statistics for facilities by whether they were acquired during the period 1998-2018. In general, these two sets of facilities seem similar in 1998. These findings are consistent with David (2020), who found that the average acquired firm is similar to the average firm.

A separate question is what attributes predict the timing of acquisitions. Acquisition timing is especially important since my empirical strategy is a leads and lags approach that relies on assumptions about the timing of treatment. There is a large set of potential regressors that could predict acquisitions at time t, especially considering both the levels and trends of each variable. I use the Least Absolute Shrinkage and Selection Operator ("LASSO") to identify which variables predict acquisition. LASSO techniques take advantage of sparsity (where many potential regressors have no effect) to choose a subset of "best" predictors for a given tuning parameter. I use a rolling cross-validation technique to select the tuning parameter.[14] Essentially, this procedure is identifying if there are any variables that strongly predict the timing of treatment.

I use a wide variety of predictors and include both the levels (at $t-1$) and trends (change from $t-4$ to $t-1$). These variables include the hourly wages by occupation, benefits, several turnover measures, log of hours worked by occupation, staffing levels per patient day, employment by aggregate occupation, patient days by payor, payor shares,

[14] The model is estimated on a sets of years and validated based on how well it performs in future years. The tuning parameter is selected based on out-of-sample performance. The tuning parameter acts as a penalty for including additional variables; a higher tuning parameter means fewer variables will be included.

patient discharges, occupancy, markup of revenue over variable costs, and various balance sheet measures per patient bed (total assets, current assets, cash on hand, total liabilities, current liabilities). Finally, I partial out year fixed effects to control for any general time variation.

The results are found in Table 3.2. The only variable that the LASSO regression selects as a consistent predictor is the markup of revenue over variable costs at time $t-1$. This finding is consistent with Gandhi, Song and Upadrashta (2021), who finds that facilities acquired by private equity firms are not substantially different than non-acquired facilities. It is also consistent with the broader acquisition literature, which finds acquired firms are similar on average to non-acquired firms. The fact that there are no trend variables that consistently predict acquisition is especially reassuring for my identification strategy. Acquisitions could be driven by idiosyncratic factors, such as the desire of an owner to sell for personal reasons, or by other factors, such as proximity to existing facilities in the chain.

3.4 Average effects of acquisitions

3.4.1 Specification and identification

First, I look at the average effect of acquisitions on workers. Then I examine how these acquisitions affect facility output in terms of prices, quantities, and quality (patient outcomes). Later I will examine heterogeneity and show that the effect of acquisitions vary significantly. To measure the effect of acquisitions on facility-level outcomes I rely on a leads and lags approach. The first specification I estimate is (for each facility i in year t):

$$Y_{it} = \sum_{j=-5}^{+5} \beta_j T_{it}^{j} + \mu_i + \theta_t + \varepsilon_{it}$$

where Y_{it} is the outcome of interest, μ_i are facility fixed effects, θ_t are year fixed effects[15] and ε_{it} is an error term. I cluster standard errors at the facility level because treatment is assigned at the facility level (Abadie et al., 2017) and due to potential serial correlation (Bertrand, Duflo and Mullainathan, 2004). The variables of interest are the T_{it}^j values. Following the advice of Schmidheiny and Siegloch (2019), these are defined by:

$$T_{it}^j = \begin{cases} \sum_{s=t+5}^{2019} d_{is} & \text{if } j = -5 \\ d_{i,t-j} & \text{if } -5 < j < 5 \\ \sum_{s=1996}^{t-5} d_{is} & \text{if } j = -5 \end{cases}$$

where d_{it} is a dummy that indicates whether an acquisition occured for facility i at time t. Therefore β_j captures the treatment effect of acquisitions after j years. In practice, I omit $j = -1$ so all treatment effects are relative to the year prior to the acquisition. Note that the total number of acquisitions occurring 5 or more years in the past are summed together (as are the number of acquisitions occurring 5 or more years in the future). This specification also allows facilities to be treated multiple times since the time period is long.[16]

The second specification is closer to a basic difference-in-differences. I bin $-4 \leq j \leq -1$ into the "pre" period and $1 \leq j \leq 4$ into the "post" period:

$$Y_{it} = \beta_{-5} T_{it}^{-5} + \beta_{\text{Pre}} T_{it}^{\text{Pre}} + \beta_0 T_{it}^0 + \beta_{\text{Post}} T_{it}^{\text{Post}} + \beta_5 T_{it}^5 + \mu_i + \theta_t + \varepsilon_{it}$$

[15]For my main specifications I use county-year fixed effects, but results are similar using only year fixed effects

[16]About 20% of facilities change ownership more than once during the time period.

where T_{it}^j is defined as:

$$T_{it}^j = \begin{cases} \sum_{s=t+5}^{2019} d_{is} & \text{if } j = -5 \\ \sum_{s=t+1}^{t+4} d_{is} & \text{if } j = \text{Pre} \\ d_{i,t-j} & \text{if } j = 0 \\ \sum_{s=t-4}^{t-1} d_{is} & \text{if } j = \text{Post} \\ \sum_{s=1996}^{t-5} d_{is} & \text{if } j = -5 \end{cases}$$

In practice I omit the $j = \text{Pre}$ so everything is relative to the four years prior to acquisition. There are two reasons for using the specification. First, it provides a single point estimate of the treatment effect. The full leads and lags specification shows that most outcomes adjust quickly after acquisition. Secondly, it increases power. For my main estimates power is not an issue, but for some sub-samples it is difficult to precisely estimate a full set of leads and lags coefficients. The downside is it is more difficult to visually assess the validity of the identification assumptions, which is why I initially present the full leads and lags specification for the main results.

3.4.2 Identification

The identification assumption is the standard parallel trends assumption – i.e. in the absence of treatment the outcomes of acquired ("treated") facilities would evolve the same way as for control facilities. The control group are facilities that have never been treated or whose treatment is occurring 5 or more years in the past or future.[17]

While it is impossible to directly validate this assumption, there are three key supporting pieces of evidence. First, as seen earlier in Table 3.1, characteristics of acquired and non-acquired facilities were generally similar in levels. The second piece of evidence is the fact that the cross-validated LASSO results show that few variables (either levels or changes) consistently predict treatment timing. Finally, and most importantly, visual inspection of

[17] The main results are also robust to dropping all facility-year observations with non-zero T_{it}^5; i.e. using only units that have never been treated or have yet to be treated as the control group.

the leads and lags for the key outcomes do not show any significant pre-treatment trends.

The primary remaining threat to identification would be if there was a concurrent shock that caused the acquisition. An example shock might be if prending lititgation forces the facility to either undergo bankruptcy or be acquired.[18] Then the true comparison group would be facilities that undergo bankruptcy and are not acquired. There are no sudden changes prior to acquisition in key financial indicators, such as cash on hand, that would suggest sudden financial distress.[19]

Finally, an alternative approach would be matching. Other papers that look at the effect of acquisitions frequently use matching. For example, Davis et al. (2014) match firms based on industry, age, size, and single/multi-unit establishment and Arnold (2020) match based on state, industry, size, and earnings. I am already focused on a narrow industry (nursing homes) and geography (California). Additionally, these facilities have similar sizes, ages, and prices; plus I am using county-year fixed effects so I am comparing treated facilities only with facilities in the same county.

A separate question from identification is how to interpret changes in working conditions. Changes in working conditions could be due either to cuts to existing workers, compositional changes in the quality of workers, or changes in required effort. Workload is shared among workers. Benefits are typically standardized. There is significant evidence that pay scales is not differentiated among nursing assistants within a facility (see Appendix C.2). Therefore, any compensation changes likely affect all workers. Potentially any compensation cuts could be offset by changes in required effort; however, I will show later that the large increases in turnover suggest that they are not perfectly offset. Therefore, I view the changes in

[18]The concern here is similar to the job-training literature – there is some shock that causes the worker to seek job training in the first place that biases the estimates.

[19]Even suppose that many acquisitions are caused by bankruptcy. If the primary goal of bankruptcy is re-structuring financial obligations or avoiding legal liability then it is not immediately clear that the counterfactual facility going through bankruptcy would have a substantially different evolution of workplace conditions than other facilities. Pensions or other long term financial obligations to workers might be the exception, but those are not a major factor in this setting. For example, Graham et al. (2019) find no significant effects of bankruptcy on the wages of firm stayers (although they do find negative effects overall for workers due to displacement).

working conditions as real changes that affect existing workers who can either accept the changes or quit (firing nursing assistants is uncommon in this setting). Either way, negative compensation changes represents losses for workers, especially considering the large literature on the negative effects of job loss (e.g., Schmieder, von Wachter and Heining, 2019).

3.4.3 Average effects of acquisition

Working conditions: The first outcome of interest is wages. I create two measures of facility level hourly wages. For the first measure, I take total salary and divide by total hours to get the average hourly wage. For the second measure I adjust for occupational composition. I calculate the average hourly wage by occupation and the average share of total hours for each occupation in each facility over the entire sample. I then use these weights to create the average hourly wage holding occupational composition fixed.

Panel A of Figure 3.7 shows the results. Average hourly wages increase by around 2% post acquisition. However, after adjusting for occupational composition there are very small or no effects on wages.[20] These results are consistent with the small literature on acquisitions, which often finds small or mixed effects on wages.

The limited effect on average iverakk wages hides heterogeneity by occupation. First, Panel B of Figure 3.7 shows the change in the log of average hourly wages by occupations. Administrative and management wages increase by 4-7%, while the wages of low wage workers (NAs, cleaning / laundry, and food preparation) decrease by 1-2%. RN/LVN wages do not change significantly. Therefore, the effects of acquisitions seem to vary with occupation.

The next outcome is the log of benefits per employee. Benefits include expenditures such as healthcare or paid time off / vacation time. Panel A of Figure 3.6 shows that average benefits decrease by about 6% post acquisition. Since expenditures on benefits are about 25-30% of expenditures on salaries, this change is approximately equivalent to a 2% wage

[20]In Appendix Figure C.1 I show that the hours of administrative workers increase after acquisitions (higher wage) and the hours of cleaning workers decrease (lower wage).

cut for all workers.

Another important workplace amenity is the workload. Here I focus on the nursing workload since it is the most comparable across facilities. It also has important implications for the quality of patient care. Panel B of Figure 3.6 shows that nursing workload increases by 3-4% after acquisition. This increase is driven primarily by an increase in the number of patients as well as a slight increase in patient severity. This result is consistent with Currie, Farsi and Macleod (2005), who find an increase in nurse workload after hospital acquisitions. Staffing is also the working condition most strongly associated with patient outcomes so any increase in workload is concerning.

Finally, I look at nursing turnover.[21] Turnover is a good measure of the overall desirability of working conditions, especially since most turnover is voluntary in this setting. Here we see the most dramatic effects in Panel C of Figure 3.6. Turnover spikes by 40% in the year of the acquisition and remains about 15% higher. The effect does not seem to fade with time. Appendix C.6 shows suggestive evidence that this increase in turnover is larger than can be explained purely by the changes in wages, benefits, and workload. Therefore, there might be other ways in which workplace conditions worsen. For example, the workplace culture could change, high turnover might weaken social relationships between workers, shift scheduling could change, or management could be less responsive to worker concerns.

Overall, nursing home acquisitions have a relatively muted effect on wages, with some heterogeneity across occupations. However, a focus on wages would miss other substantial effects on benefits, workload, and turnover.

Patient outcomes: Next I look at the effect on patient prices. Table 3.5 shows the effect on prices (revenue per patient day). Overall prices rise, but prices conditional on payor type change only slightly or not at all. Table 3.6 shows that the increase in overall price is primarily due to compositional effects – there is an increase in the share of patients on Medicare, which has significantly higher reimbursement rates. All of this (combined with

[21]Results are similar if I use overall turnover rates instead.

the increased workload and lower benefits) leads to an increase in the markup over variable costs.

The final set of outcomes are patient health outcomes. Potentially, acquiring firms could have implemented new technologies or management approaches that allowed them to reduce staffing or worker effort (by reducing compensation) without harming patients. However, if patients see significant harm, then it is unlikely that acquirers significantly improved technical efficiency. I look at several quality outcomes: the total deficiency citation score, the number of patients discharged to hospital or death per patient day, and the share of patients experiencing declines in their ability to perform Activities of Daily Living (ADL).

The count of deficiency citations comes from state inspections.[22] I create the deficiency citation score using the points assigned to each type of deficiency citation in the Five Star Quality rating system. This scoring system places additional weight on more severe deficiencies. The number of patients discharged to hospital or death per patient day could be due to poor patient care, or it could be part of a facility effort to remove low-value patients. The share of patients experiencing declines in their ability to perform ADLs captures the change in patient outcomes during a stay. I standardize all of these variables within year to have a mean of zero and a standard deviation of one. I standardize the variables due to reporting changes over time.

Table 3.7 gives the results. Panel A shows the main results in columns (1)-(3). Deficiency scores do not change, but there are significant increases in discharges to hospital or death and the share of patients experiencing declines in their ADL ability. One potential concern is that the patient mix changes post acquisition. Panel B adds additional controls for the aggregate patient severity, which does not significantly affect the results. Patient outcomes worsen for both of these variables by about 0.175 to 0.229 standard deviations.

Columns (4) and (5) look at two alternate measures of patient outcomes that adjust for patient characteristics using micro-data but are available only for a smaller sample of years.

[22]These inspections need to occur at least every 15 months, or they can be triggered by serious complaints.

The adjusted re-hospitalization within 30 days adjusts for the expected re-hospitalization rate for each individual based on 30 MDS variables. The expected successful discharge rate is the rate of successful discharges back into the community; it is also adjusted for individual factors such as end-stage prognosis. Patient outcomes worsen by a similar amount for these adjusted variables, by about 0.269 and 0.170 standard deviations.

To summarize, on average working conditions seem to worsen, especially for low-wage workers. A focus solely on the average wage effects would miss the key margins of adjustment. Facilities serve more Medicare patients, but quality seems to decrease. One potential explanation for the increases in Medicare patients despite quality declines is that acquiring firms might be more effective at recruiting Medicare patients or more aggressively ration spots for non-Medicare patients. However, there is no reason to believe that most acquisitions have very close to "average" effects. Instead, there might be significant heterogeneity.

3.5 Heterogeneity and convergence

3.5.1 Specification

The primary source of heterogeneity is if firms implement uniform production processes or behavioral factors lead to uniform policies. Take the example of staffing. A simple approach is to take the gap between staffing at the acquired facility and typical staffing at the acquiring firm. If the acquisition is at time t, then I take the median staffing at the acquiring firm's facilities at time $t-1$. I split acquisitions by whether the acquired facility initially had higher or lower staffing. I stack the acquisitions relative to the acquisition year and plot the median gap by year. Figure 3.7 shows clear convergence by the acquired facilities towards the staffing levels of the acquiring firm immediately after acquisitions. Acquired facilities with relatively lower staffing see increases, while those with relatively higher staffing see decreases.[23]

[23] Appendix Figure C.4 shows a similar picture for benefits convergence.

Next, I look at regression evidence. I use a slightly different approach than before. For each year t, I create a dataset containing all facilities that were acquired that year plus all facilities that did not undergo an acquisition in the four years prior or following. I then take the change in outcomes, ΔY_{it}.[24] I stack the datasets and run the regression:

$$\begin{aligned}\Delta Y_{it} =& \beta \text{Acquired}_{it} + \beta^{\text{Above}} \Delta \text{Above Acquirer}_{i,t-1} + \beta^{\text{Below}} \Delta \text{Below Acquirer}_{i,t-1} \\ &+ X_{it}\gamma + \theta_t + \varepsilon_{it}\end{aligned}$$

ΔAbove Acquirer$_{i,t-1}$ is the gap (if positive) in outcome Y_{it} between the acquired facility and the acquiring firm at time $t-1$. ΔBelow Acquirer$_{i,t-1}$ captures the gap if it is negative (e.g., if the acquired facility has lower wages than the acquiring firm). If convergence is important then we would expect β^{Above} to be negative and β^{Below} to be positive. βAcquired$_{it}$ allows for general effects of acquisitions that do not depend on the initial gap. X_{it} includes any controls; the primary concern is potential regression to the mean. I add controls for the gap between $Y_{i,t-1}$ of the facility and the median of the sample (ΔAbove Median and ΔBelow Median).

A few quick notes on the data: I define the gap variables only if the acquiring firm has at least 3 facilities at time $t-1$. Additionally, for some acquisitions I have not identified the acquiring firm, only that ownership changed. In both cases I drop these acquisitions.[25] Finally, instead of using the required vs. actual staffing I use the simpler measure of hours of nursing per patient day (HPRD). Higher values correspond to a lower workload (since there are more nursing hours per patient). I use this measure for two reasons. First, there is evidence that firms set HPRD standards. Second, the required staffing variable

[24] I take either the change from $t-1$ to $t+2$ or the change in the average of $t-4$ through $t-1$ to $t+1$ through $t+4$.

[25] I can identify acquisitions based on change in ownership; however, it is harder to link facilities by ownership in the cross-section since ownership is inconsistently reported. Therefore, I rely on the manual classification of firms in 2015 as well as the manual identification of all acquisitions by the three largest firms.

is measured with significant noise, which is important now that it is being used to construct an explanatory variable.[26] In Appendix Table C.7 I show that this specification produces similar estimated average treatment effects to the leads and lags approach in the previous section.

3.5.2 Results

Working conditions: Table 3.8 presents the results for the primary working condition outcomes. I focus on working conditions for nurses since they are the largest part of the workforce and I have measures of workload and turnover. As expected, across all outcomes we see significant convergence toward the working conditions of the acquiring firm. Convergence is lowest for wages, but is substantial for all other outcomes. Around half the initial gap is closed for both benefits and staffing. This convergence is true both for working conditions that are initially above the acquiring firm and for working conditions that are initially below. There is also not a consistent asymmetry between convergence from above versus convergence from below.

These estimates likely underestimate the actual degree of standardization of working conditions across facilities for several reasons. First, there is potentially measurement error in working conditions, which means the gap variable is measured with error. Second, I use the median facility conditions in the acquiring firm. However, the acquiring firm might standardize working conditions within geographies or other groupings of facilities. Third, workforce composition differences might lead to gaps remaining even when policies are fully standardized. Finally, the gap measure is fixed at time $t-1$, but firm policy will continue to evolve over time. Therefore, the acquired facility might perfectly match working conditions at time $t+1$ and on and still not fully converge by this measure.[27] Therefore, I view these

[26]Classical measurement error in the outcome variable will not affect estimates but will lead to attenuation for explanatory variables.

[27]However, I prefer using the fixed measure at time $t-1$ to be conservative since firm choices post-acquisition might be affected by the acquisition.

estimates as conservative estimates of the degree of convergence.

I can take these results one step further to get the distribution of expected acquisition effects. Multiplying the gaps by the estimated convergence coefficients gives the expected treatment effect for each acquisition,[28] i.e., for each outcome:

$$\text{Expected } \Delta Y_{it} = \hat{\beta}\text{Acquired}_{it} + \hat{\beta}^{\text{Above}}\Delta \text{ Above Acquirer}_{it} + \hat{\beta}^{\text{Below}}\Delta \text{ Below Acquirer}_{it}$$

This measure is more informative than simply looking at the estimated coefficients. If all facilities have very similar working conditions, then the effect of convergence is limited. Figure 3.8 presents the distribution of treatment effects. The graphs show substantial heterogeneity. The variation is smallest for wages; the 10th percentile sees wage losses of 6% while the 90th percentile sees wage gains of 3%. Benefits decrease by 25% at the 10th percentile, while they increase by 20% at the 90th percentile. Staffing decreases by 6% at the 10th percentile but increases by 4% at the 90th percentile.

Another potential concern is that these effects are not independently distributed. For example, a facility with relatively lower wages might have higher benefits. Any increase in wages could be offset by decreases in benefits. Or, working conditions might be positively correlated which creates larger effects. One way of addressing this issue is to combine the estimated treatment effects into one measure using estimates of workers' willingness to pay for working conditions.

In Appendix C.1 I conduct a simple exercise to calculate the marginal willingness to pay for changes in workload and benefits. I follow the methodology of Gronberg and Reed (1994), who show that changes in separation rates can be used to identify the marginal willingness to pay for amenities in a simple search framework. The intuition is that workers will vary on-the-job search effort based on the utility of the job. With this approach, I estimate that

[28] Likely the average convergence is composed of some firms completely converging and others only partially or not at all converging. An alternative approach would be to assume either facilities fully converge or not at all and use the point estimates to identify the share that fully converge. This would likely increase the expected heterogeneity in the treatment distribution, so I use my current approach to be conservative.

workers are willing to decrease wages by 1% to decrease workload by 2%. I also estimate that workers are willing to decrease wages by 1% to increase benefits by 5%.[29] The combined treatment effect is then:

$$\begin{aligned}\text{Expected }\Delta\text{Compensation}_{it} = 1 * \text{Expected }\Delta ln(Wage)_{it} + 0.5 * \text{Expected }\Delta ln(Workload)_{it}\\ + 0.2 * \text{Expected }\Delta ln(Benefits)_{it}\end{aligned}$$

Where Expected ΔCompensation$_{it}$ is the total change in wage-equivalent compensation. Figure 3.9 presents the results. Workers at the 10th percentile of acquisitions see losses equivalent to a 12% decrease in wages, while workers at the 90th percentile have gains of 7-8%. The median acquisition has no effect on workers.

I could also estimate unobserved amenities using changes in turnover as a proxy and include them in the estimated expected treatment effect. The changes in turnover are much larger than can be easily explained by changes in wages, benefits, and staffing, which would imply very significant unobserved amenities. Therefore, the results are very sensitive to the inclusion of turnover and so I omit turnover to be conservative.

One potential concern with combining these estimates is that there might be offsetting effects. For example, if an acquired facility has both high wages and benefits then an acquiring firm might reduce both by less. In Appendix Table C.9, I show that the estimated coefficients do not change significantly when the other working condition gaps are included.[30] The large increase in turnover also suggests that changes to one working condition are not offset by changes in other working conditions.

Table 3.9 shows the wage convergence for individual occupations.[31] As a reminder, on average wages increased for Administrative and Management positions, did not change

[29]The estimate for benefits is slightly less than if a worker valued a dollar in benefits equally to a dollar in salary.

[30]The distribution of expected treatment effects is similar if the cross-yrtmd are included, so I omit them for simplicity. Results are available upon request.

[31]The measures of workload, benefits, and turnover by occupation are much more coarse or unavailable.

for RN/LVN, and decreased slightly for NAs, Cleaning, and Food Preparation workers. None of these effects are still statistically significant except for the management effect. However, across all occupations there is significant convergence in wages. The convergence is significantly stronger for the higher wage occupations that are hired less frequently (Administration, Management, and RNs) than for lower wage occupations (LVN, NA, Cleaning, and Food Preparation).

Patient outcomes: Next, I turn to patient outcomes. It is less clear why patient outcomes might converge. While patients might be similar across time within a facility, there is more reason to believe that patients might significantly vary across locations. Therefore I use a simple indicator for whether the acquired facility is above or below the median value of the acquiring firm rather than the difference. Note that being initially "above" on these measures means the facility was initially worse.

As seen in Table 3.10, there is not a clear pattern of convergence in both directions. Acquired facilities that have more deficiencies see a reduction (which is an improvement in quality). On the other hand, acquired facilities that have fewer deficiencies see no change. Therefore, it seems as though acquiring firms with lower deficiency rates do reduce deficiencies for high deficiency facilities.

On the other hand, acquired facilities that do relatively worse on discharges to hospitals/death and ADL decline do not see improvements after being acquired. Instead, acquired facilities that did relatively well on these metrics see significant declines. Therefore, it seems as though acquiring firms worsen patient outcomes in facilities that were previously doing well. This finding raises concerns about acquisition of facilities that are initially doing a good job caring for patients. There is less concern about the acquisition of facilities that already have worse patient outcomes since there might be slight improvements due to a reduction in deficiencies.

3.5.3 Mechanism

To summarize, there is a clear convergence towards the working conditions of the acquiring firm. The next question is what causes the convergence. At the start, there are at least several potential mechanisms for how acquisitions might affect working conditions: (1) production function changes, (2) behavioral factors, (3) rent transfers, and (4) input market power.

It would be difficult to explain convergence due to changes in input market power. Rent transfers could explain convergence from above as rents are removed, but it does not explain convergence from below. For most estimates, the convergence from below is just as strong or stronger than the convergence from above.

That leaves productivity changes or behavioral factors. First, the potential story for how productivity changes might explain convergence. There is limited scope for differences in technical productivity since most work, such as changing patients or assisting them with the toilet, is highly manual and standard. However, there might be managerial improvements, similar to the Japanese textile firms in Braguinsky et al. (2015). For example, some firms might be better or worse at matching staffing levels with patient demand. Or a firm might have a systematic advantage at recruiting, which affects the optimal level of turnover.

Secondly, behavioral factors, such as managerial inertia, could explain convergence. It might be costly to discern the optimal wages, benefits, and working conditions for every location and occupation. Therefore, managers might rely on standardized policies across facilities. Or, internal fairness concerns might lead firms to use consistent policies.

Ultimately, it is difficult to definitively distinguish between productivity and behavioral explanations. However, there are three pieces of evidence that suggest behavioral concerns might play at least some role. First, as seen in Table 3.9, convergence is stronger for "thinner" markets. Firms hire the most NAs, followed by Cleaning, Food Preparation, and LVNs. All of these workers see relatively lower rates of wage convergence. On the other hand, Administrative, Management, and RN have substantially higher rates of convergence.

Secondly, there is convergence within each of the working condition categories, and the convergence is not uniformly in the same direction. I re-run the same convergence regressions looking only at facilities where the convergence of wages and benefits are in opposite directions – i.e. we expect one to increase and the other to decrease. The results in Table 3.11 show that the rate of convergence is similar even in these cases. There are potential alternative explanations, such as if a firm has cost advantages for supplying different types of working conditions, but the most straightforward answer is uniform policies due to behavioral factors.

Finally, there is evidence that wages at a facility are affected by local labor market conditions near other facilities within the same firm. I regress log of hourly wages for NAs, LVNs, and RNs on local wages (defined as average wage for other facilities within 5 kilometers) and on the average local wages at other facilities within the same firm. I restrict the sample to 2015, when I have a complete picture of facility ownership networks. For each facility, I construct a leave-out measure of average local wages at other facilities within the firm by excluding facilities in the same county. Table 3.12 shows that facility wages are correlated with wages in other markets within the same firm even conditional on local wages. Moreover, the coefficient on local wages and local wages at other facilities within the same firm are almost equal for RNs. Again, there are other potential explanations; for example, if markets are imperfectly competitive then high wage firms will increase local wages, which will create a correlation between facility wages and local wages at other facilities. However, a simpler explanation is uniform policies due to behavioral factors

Ultimately, the results suggest that changes to the production process or behavioral factors must be important to explain convergence. There is suggestive evidence that behavioral factors are important, although it is not conclusive. While there has been research on the importance of fairness constraints, the role of managerial inertia in labor markets might be an interesting avenue for future research.

3.6 Conclusion

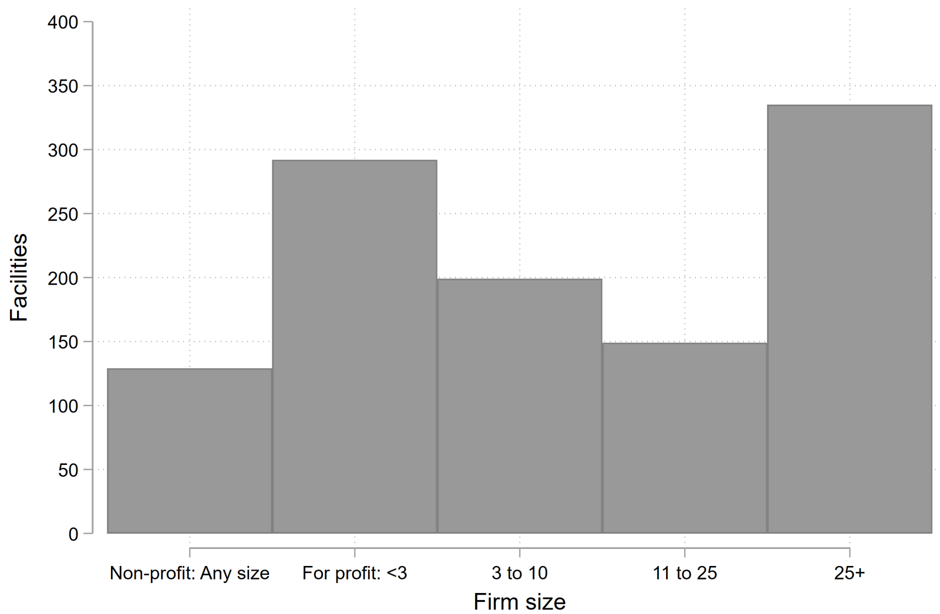
This paper studies the effects of acquisitions on working conditions, specifically looking at nursing home acquisitions. I use a leads and lags approach to identify the effect of acquisitions. There are several key findings: adjustments to working conditions other than wages are important with significant declines in average benefits, increase in workload, and increase in turnover. These changes are persistent. Patient outcomes worsen after acquisitions, primarily in facilities that had previously been doing better. Most importantly, there is substantial convergence across working conditions. This convergence creates significant risks for workers; some acquisitions result in large improvements in working conditions while others impose large costs. There is suggestive evidence that behavioral factors might be important in explaining this convergence.

There are several potential avenues for future work. First, it would be helpful to extend these results to other empirical settings. While nursing homes are an important sector, they are also relatively unique. Second, while this paper shows some suggestive evidence on behavioral factors, it is important to understand the role of behavioral factors in employment decisions. As discussed in DellaVigna and Gentzkow (2019), there are potentially broad implications to managerial inertia. For example, managerial inertia can dampen labor market adjustments to local shocks. National increases in employer concentration could then decrease the responsiveness of labor markets to local conditions. Finally, the large increases in turnover suggest that there are other important effects of acquisitions that are being missed. The increases are persistent, which suggests they are due not only to a temporary shock and they are larger than can be easily explained by changes to wages, benefits, and workload. Potential causes could be changes in other working conditions, such as company culture or the disruption of personal relationships. For example, interviews by The Guardian with Whole Foods workers after they were acquired by Amazon suggested that workers were unhappy with the "pressured environment and the erosion of Whole Foods' corporate culture."[32] If these other factors are important, then they might play an overlooked role in labor market decisions and be a fruitful topic for future research.

[32] Michael Sainato. 2019. "Whole Foods workers say conditions deteriorated after Amazon takeover." *The Guardian*, July 16.

3.7 Tables and figures

Figure 3.1: Number of facilities by chain size (CA 2015)



Note: For California nursing homes; chain size excludes nursing homes outside of California or non-nursing home businesses

Figure 3.2: Number of nursing homes changing ownership (CA 1998-2018)

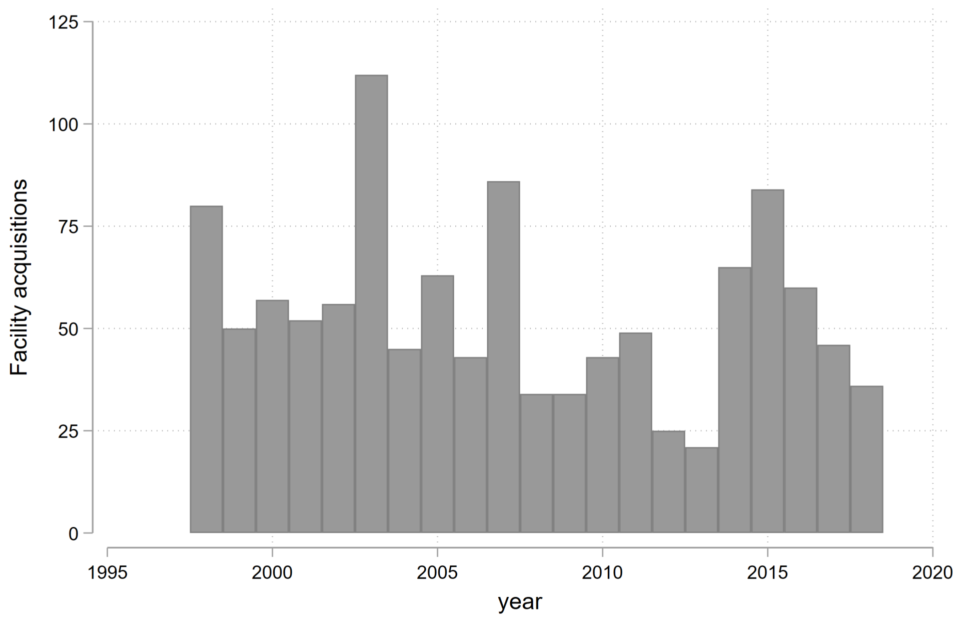


Figure 3.3: Interquartile range of working conditions across facilities for four large nursing home chains (CA 2015)

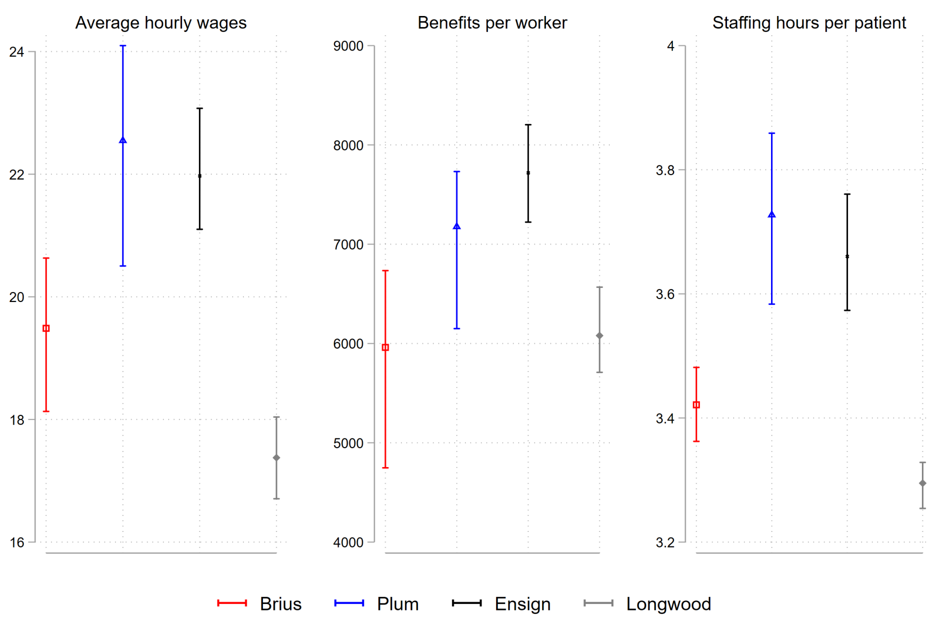
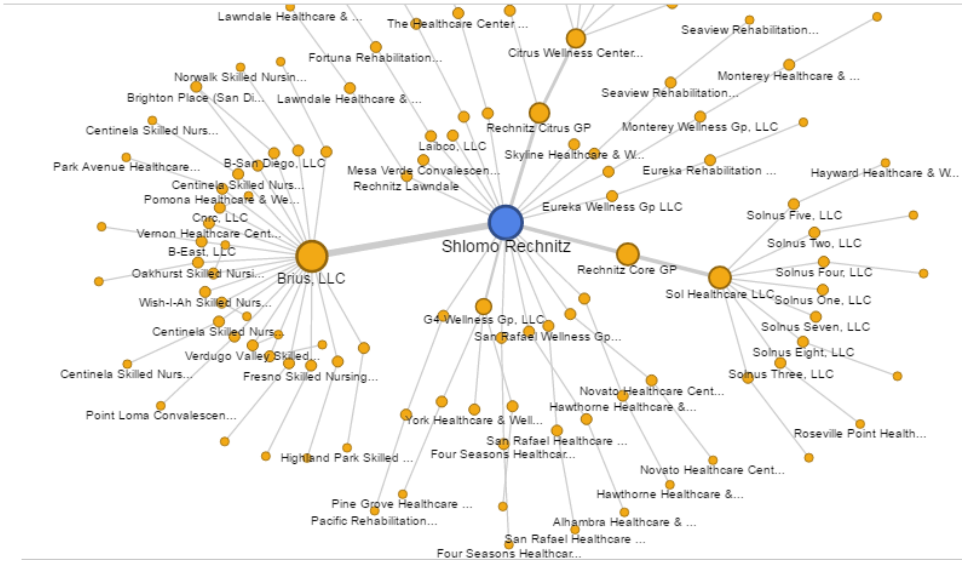
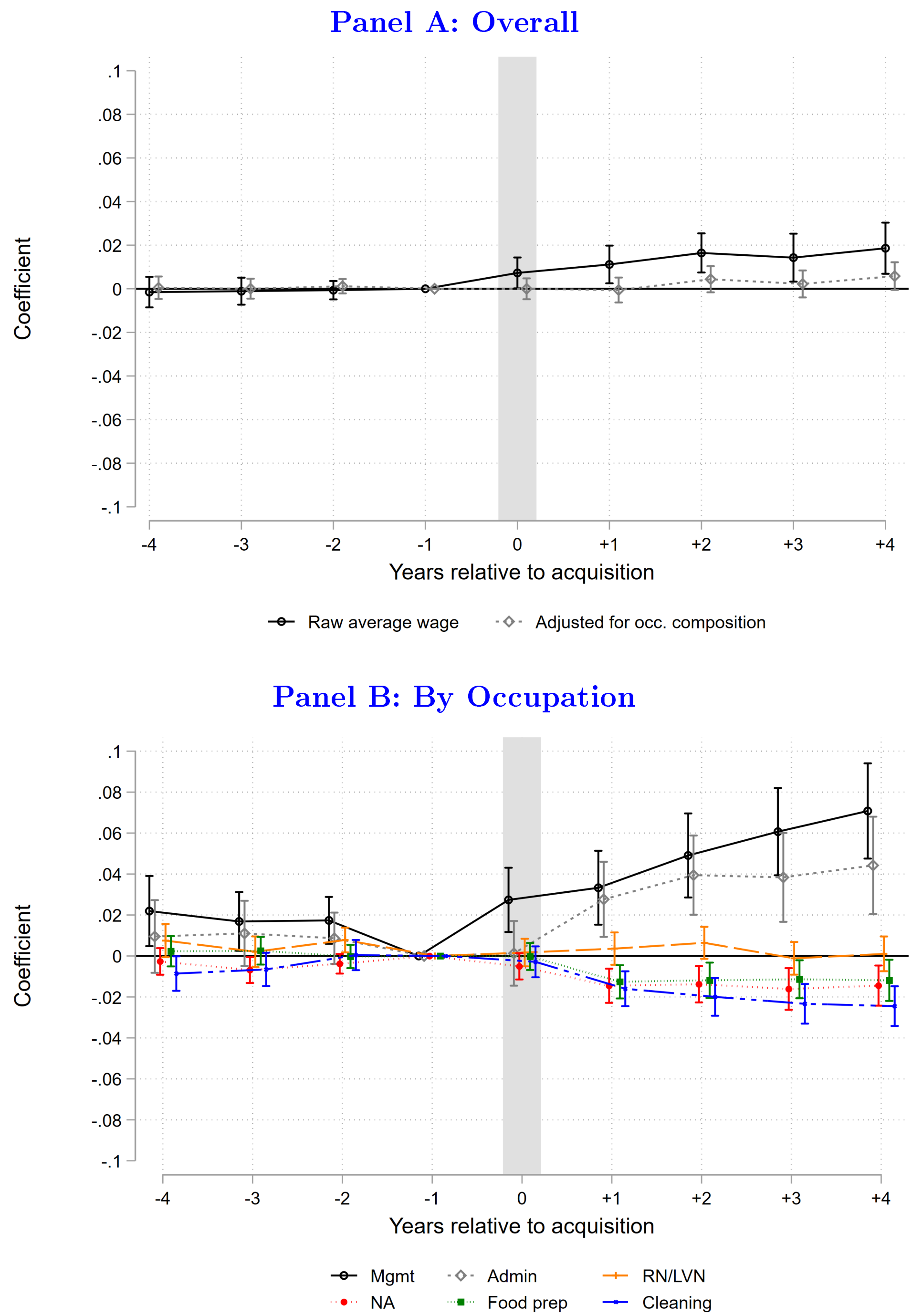


Figure 3.4: Brius ownership network (as of 2015)



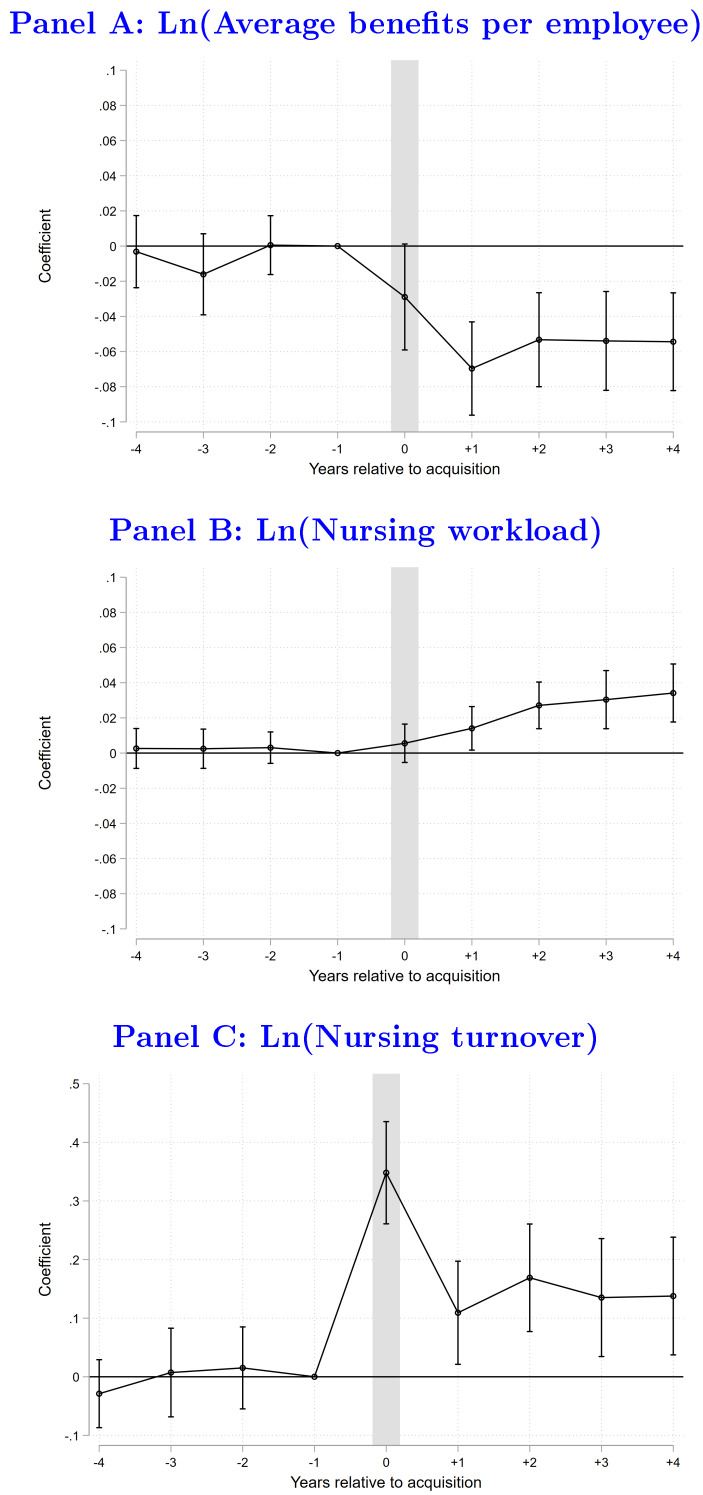
Note: Image from Sacramento Bee (see http://media.sacbee.com/static/sinclair/Nursing2/)

Figure 3.5: Effect of acquisitions on log of hourly wages



Note: For California for-profit nursing home facilities (1997-2019). Includes county-year and facility FE. Adjusted means holding occupation composition fixed based on facility averages over the entire sample period. Ranges are 95% confidence intervals.

Figure 3.6: Effect of acquisitions on other working conditions



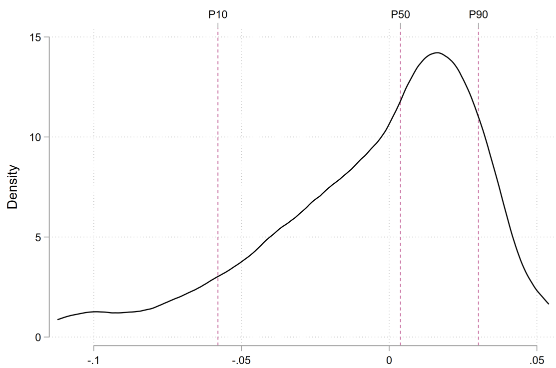
Note: For California for-profit nursing home facilities (1997-2019). Includes county-year and facility FE. Ranges are 95% confidence intervals. Benefits exclude WC and UI payments.

Figure 3.7: Gap in log of staffing between acquired facility and acquiring firm by year relative to acquisition

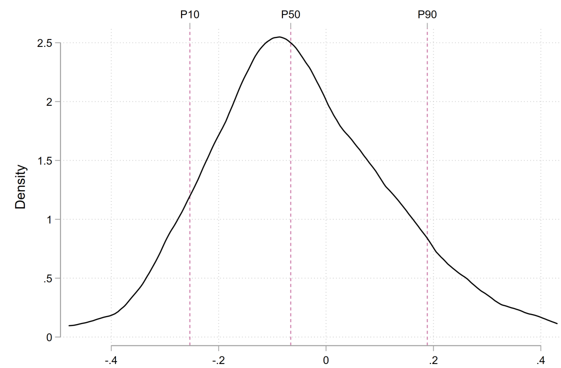
Note: For California for-profit nursing home facilities (1997-2019) that were acquired by an identified chain with at least three establishments. Gaps are relative to acquiring chain's median value at time t-1 and have been adjusted for general time trends

Figure 3.8: Distribution of expected acquisition treatment effects

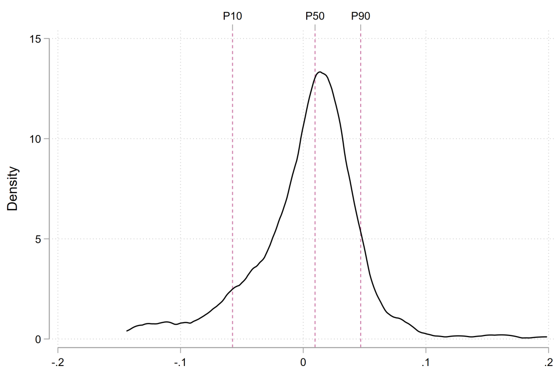
Panel A: ln(Nursing wage [Adj.])



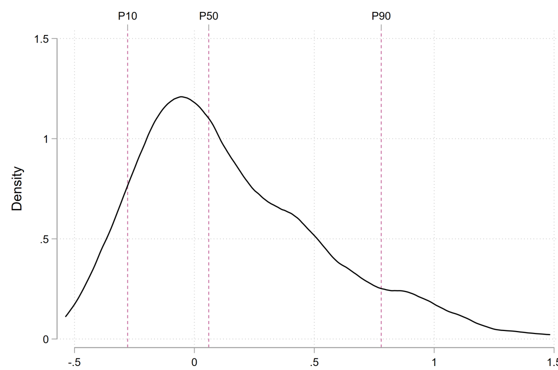
Panel B: ln(Nursing Benefits)



Panel C: ln(Staffing)

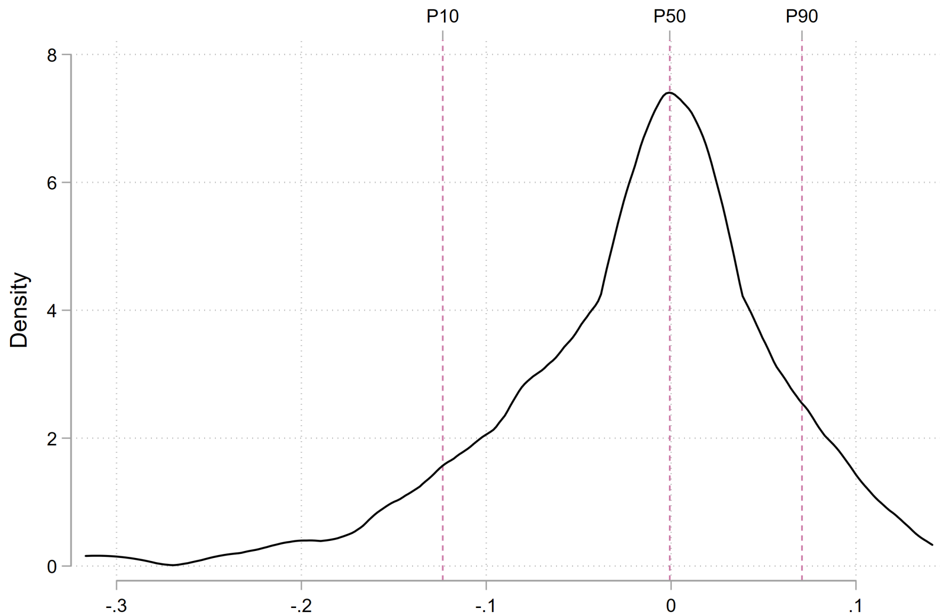


Panel D: ln(Nursing Turnover)



Note: For California for-profit nursing home facilities (1997-2019).

Figure 3.9: Distribution of expected willingness to pay for working condition changes



Note: For California for-profit nursing home facilities (1997-2019).

Table 3.1: Facility characteristics by whether ever acquired

	1998		2018	
	Never acquired	Ever acquired	Never acquired	Ever acquired
Facilities	275	638	246	621
Avg hourly wages - All	11.87	11.65	21.72	22.39
Avg hourly wages - NAs	7.59	7.74	15.54	15.72
Nursing hours per resident day	2.81	2.89	3.82	3.76
Average benefits	3,254.61	3,489.98	7,749.70	7,256.15
Turnover - Nursing	0.76	0.80	0.55	0.55
Employees - All	94.31	97.03	124.51	132.95
Employees - NA	37.48	39.63	51.26	53.63
Occupancy rate	0.89	0.88	0.89	0.88
Medicare share	0.06	0.07	0.15	0.17
Revenue per patient day	120	125	315	336
$\frac{Revenue}{VariableCosts}$	1.39	1.36	1.46	1.49

Note: For California nursing homes. Sample excludes specialized facilities (ICF and SAC).

Table 3.2: Acquisiton predictors selected by cross-validated LASSO

	(1)	(2)
$\frac{Revenue}{VariableCosts}_{t-1}$	-0.0950***	-0.0961***
	(0.00805)	(0.00824)
Observations	16,983	16,981
R-squared	0.018	0.022
Year FE	X	X
County FE	-	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.3: Effect of acquisitions on working conditions

	(1) ln(Avg. Hourly Wage)	(2) ln(Benefits)	(3) ln(Workload)	(4) ln(Nursing Turnover)
Acquired * Post	0.0152***	-0.0538***	0.0238***	0.137***
	(0.00503)	(0.0115)	(0.00751)	(0.0354)
Observations	16,891	20,141	12,290	20,175
R-squared	0.933	0.777	0.736	0.387
County-Year FE	X	X	X	X
Facility FE	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Controls include an indicator for the year of acquisition and the binned number of acquistions at least five years before and at least five years after. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.4: Effect of acquisitions on log of average hourly wages by occupation

	(1) All	(2) All - Adj.	(3) Mgmt.	(4) Admin	(5) RN/LVN	(6) NA	(7) Cleaning	(8) Food prep
Acquired * Post	0.0159***	0.00565*	0.0429***	0.0419***	-0.00131	-0.0103**	-0.0142***	-0.00956**
	(0.00517)	(0.00318)	(0.0101)	(0.0108)	(0.00376)	(0.00463)	(0.00454)	(0.00442)
Observations	16,443	17,577	17,413	17,571	17,561	17,566	14,551	17,040
R-squared	0.930	0.949	0.769	0.680	0.887	0.922	0.908	0.891
County-Year FE	X	X	X	X	X	X	X	X
Facility FE	X	X	X	X	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Controls include an indicator for the year of acquisition and the binned number of acquistions at least five years before and at least five years after. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.5: Effect of acquisitions on revenue per patient day by payor type

	(1) ln(Avg. Price)	(2) ln(Medi-Cal Price)	(3) ln(Medicare Price)	(4) ln(Private Price)
Acquired * Post	0.0582***	0.00335	0.0175**	0.0195
	(0.00670)	(0.00406)	(0.00813)	(0.0134)
Observations	20,288	19,434	18,724	17,840
R-squared	0.919	0.908	0.693	0.602
County-Year FE	X	X	X	X
Facility FE FE	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Controls include an indicator for the year of acquisition and the binned number of acquistions at least five years before and at least five years after. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.6: Effect of acquisitions on patient quantities by payor type

	(1) ln(Total days)	(2) Occupancy rate	(3) Share Medi-Cal	(4) Share Medicare	(5) Share private
Acquired * Post	0.0161**	0.0133***	-0.00456	0.0320***	-0.0199***
	(0.00636)	(0.00398)	(0.00717)	(0.00321)	(0.00611)
Observations	20,290	20,158	20,290	20,290	20,290
R-squared	0.944	0.561	0.744	0.730	0.684
County-Year FE	X	X	X	X	X
Facility FE	X	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Controls include an indicator for the year of acquisition and the binned number of acquistions at least five years before and at least five years after. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.7: Effect of acquisitions on normalized patient outcomes

	(1) Total Deficiency Score	(2) Discharge to Hosp/Death per pday	(3) ADL Decline	(4) Adjusted Re-hospitalization Rate	(5) Adjusted Discharge Success
Panel A: Basic DD					
Acquired * Post	0.0109	0.175***	0.229***	0.269***	-0.170*
	(0.0433)	(0.0418)	(0.0527)	(0.0755)	(0.100)
Panel B: Severity Controls					
Acquired * Post	0.0197	0.136***	0.179***	0.282***	-0.166*
	(0.0411)	(0.0427)	(0.0502)	(0.0752)	(0.0947)
Observations	16,627	13,302	8,834	5,279	3,884
County-Year FE	X	X	X	X	X
Facility FE	X	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. All outcomes have been standardized within each year to have a mean of 0 and standard deviation of 1. Controls include an indicator for the year of acquisition and the binned number of acquistions at least five years before and at least five years after. Severity controls include the average Resource Utilization Group Nursing Case Mix Index in April and for all admits, the average ADL score in April and for all admits, and the share of admits from acute hospitals. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.8: Effect of acquisitions on nursing working conditions by whether above or below acquiring firm

	(1) Δ ln(Wage [Adj.])	(2) Δ ln(Benefits)	(3) Δ ln(Staffing)	(4) Δ ln(Turnover)
Acquired	0.0127**	-0.0475**	0.0139*	0.0276
	(0.00585)	(0.0220)	(0.0228)	(0.0496)
Δ Above Acquirer	-0.356***	-0.461***	-0.551***	-0.472***
	(0.0585)	(0.0792)	(0.0992)	(0.113)
Δ Below Acquirer	0.168**	0.654***	0.380***	0.837***
	(0.0766)	(0.0717)	(0.0807)	(0.0846)
Δ Above Median	0.0114	-0.00110	-0.105***	-0.111***
	(0.00937)	(0.0123)	(0.0174)	(0.0212)
Δ Below Median	0.0648***	0.0503***	0.113***	0.0486**
	(0.0111)	(0.0159)	(0.0392)	(0.0222)
Observations	12,104	12,015	12,050	12,006
R-squared	0.560	0.071	0.138	0.072
Year FE	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.9: Effect of acquisitions on log of hourly wage for each occupation by whether above or below acquiring firm

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ ln(Hourly wage)	Admin	Mgmt.	RN	LVN	NA	Cleaning	Food prep.
Acquired	0.0109	0.0657***	0.00803	0.0112*	0.00671	-0.0102	-0.0133
	(0.0157)	(0.0209)	(0.00950)	(0.00650)	(0.00739)	(0.0129)	(0.00944)
Δ Above Acquirer	-0.514***	-0.698***	-0.523***	-0.336***	-0.295***	-0.211***	-0.209**
	(0.0765)	(0.113)	(0.0775)	(0.0479)	(0.0777)	(0.0706)	(0.0919)
Δ Below Acquirer	0.643***	0.495***	0.661***	0.260***	0.185**	0.226*	0.450***
	(0.0808)	(0.0980)	(0.123)	(0.0661)	(0.0769)	(0.115)	(0.0804)
Δ Above Median	-0.124***	-0.0771***	-0.0585***	-0.0111	-0.00270	-0.0143	-0.0353*
	(0.0230)	(0.0215)	(0.0173)	(0.0144)	(0.00863)	(0.0124)	(0.0196)
Δ Below Median	0.102***	0.0993***	0.118***	0.0469***	0.0705***	0.101***	0.0613***
	(0.0177)	(0.0198)	(0.0213)	(0.0129)	(0.0129)	(0.0229)	(0.0169)
Observations	11,468	11,864	11,865	12,087	12,082	9,113	10,841
R-squared	0.113	0.094	0.331	0.445	0.493	0.294	0.185
Year FE	X	X	X	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.10: Effect of acquisitions on patient outcomes by whether above or below acquiring firm

	(1) Δ Total Deficiency Score	(2) Δ Discharge to Death/Hosp per day	(3) Δ ADL Decline
Above Acquirer	-0.253***	0.0474	0.0274
	(0.0871)	(0.104)	(0.126)
Below Acquirer	0.0690	0.326***	0.595***
	(0.0691)	(0.0797)	(0.105)
Above Median	0.0342*	0.123***	0.127***
	(0.0195)	(0.0215)	(0.0308)
Observations	8,521	6,095	3,726
R-squared	0.005	0.012	0.020
Year FE	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.11: Convergence of working conditions by whether convergence of wages and benefits are in different directions

	(1)	(2)	(3)	(4)
	ln(Nursing wage)		ln(Benefits)	
	All	Split signs	All	Split signs
Acquired	0.0126**	0.00491	-0.0477**	-0.0239
	(0.00586)	(0.00750)	(0.0228)	(0.0368)
Δ Above Acquirer	-0.356***	-0.310***	-0.460***	-0.488***
	(0.0585)	(0.0756)	(0.0792)	(0.180)
Δ Below Acquirer	0.170**	0.558***	0.655***	0.573***
	(0.0766)	(0.196)	(0.0717)	(0.158)
Δ Above Median	0.0122	0.0112	0.000204	-0.00371
	(0.00933)	(0.00944)	(0.0123)	(0.0125)
Δ Below Median	0.0651***	0.0693***	0.0498***	0.0539***
	(0.0110)	(0.0108)	(0.0159)	(0.0166)
Observations	11,952	11,722	11,246	11,016
R-squared	0.563	0.563	0.071	0.054
Year FE	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Split signs means that the expected convergence is in opposite direction for benefits vs. wages; one is expected to increase and the other decrease. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table 3.12: Correlation of wages with nearby facilities and within firm (2015)

	(1)	(2)	(3)	(4)	(5)	(6)
ln(Hourly wage)	NA		LVN		RN	
Avg. local wages	0.863***	0.835***	0.787***	0.751***	0.355***	0.297***
	(0.0297)	(0.0294)	(0.0348)	(0.0346)	(0.0647)	(0.0625)
Firm avg. local wages (in other counties)	0.221***	-0.350***	0.332***	-0.251**	0.246**	-0.0518
	(0.0585)	(0.0987)	(0.0752)	(0.126)	(0.110)	(0.0842)
Firm avg. wages (in other counties)		0.560***		0.515***		0.638***
		(0.0779)		(0.0816)		(0.0726)
Multi-county firm	-0.525***	-0.498***	-1.037***	-0.817***	-0.827**	-2.038***
	(0.151)	(0.146)	(0.244)	(0.232)	(0.390)	(0.356)
Observations	766	766	766	766	757	757
R-squared	0.583	0.605	0.453	0.479	0.068	0.141

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Local wages are defined as the average wage at facilities within 5 km. Firm avg. local wages (in other counties) is the average local wage across the firm's facilities in other counties. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

APPENDIX A

Appendix Materials for Chapter 1

A.1 Supplementary analysis

A.1.1 Estimating market wage for equity partners

Equity partners earn compensation from directly providing labor, but they also earn compensation by providing capital. Non-lawyers are typically barred from owning law firms, so the expected return on equity could be higher than that of the economy overall. Therefore, it is not immediately clear how to separate wage labor and capital returns for equity partners. Note that for some analysis, I care only about the total compensation that is directly available.

In order to impute the market wage for equity partners, I rely on the wages for non-equity partners. I take three separate approaches, which require differing assumptions. Let w_{it}^{NE} be the yearly compensation for non-equity partners in firm i at time t and w_{it}^{E} be the same for equity partners. In the first approach , I assume that the market wage for equity partners is a firm-specific multiple of the non-equity partner wage, i.e. $r_i = \frac{w_{it}^{E}}{w_{it}^{NE}} \implies w_{it}^{E} = r_i w_{it}^{NE}$. The drawback of this approach is that not all firms have non-equity partners. Therefore, in the second approach, I rely on the predicted compensation for non-equity partners. I run the regression:

$$\ln w_{it}^{NE} = \beta \ln \bar{R}PL_i + \gamma_t + \varepsilon_{it}$$

for the sample of firms with w_{it}^{NE} where $\ln \bar{R}PL_i$ is the firm's average log of revenue per lawyer (after conditioning on year fixed effects). In practice, I allow β to vary by year. I then predict $\hat{w_{it}^{NE}}$ for the entire sample of firms and again assume that equity partner wage is a firm-specific multiple of the predicted non-equity partner wage. For calculating the total variable costs, I make the additional assumption that $r_i = 1$.

The third and final approach is used to estimate the time series variation in the market wage for equity partners. I assume that the growth rate in the market wage for equity partners is the same as the growth rate in the wage for non-equity partners. Starting in 1994, I take the average within firm growth rate in non-equity partner compensation (weighted by the number of non-equity partners at time t-1). I then apply this year by year to create the growth rate in the market wage for equity partners, indexed to the starting year of 1994.

A.1.2 Alternate models of production

I presented the most basic static model of firm production in the text. My focus is on the first order condition for associates. The key alternatives are cases where, conditional on the quantity of partners chosen, K_t, there are frictions that affect the quantity of associates, N_t. The two most obvious cases are (1) if there are adjustment frictions or (2) if new associates are imperfect substitutes for experienced associates (e.g., if associates develop firm-specific human capital) and firms cannot charge separate wages. I consider basic models of both cases and show that the first order conditions still depend primarily on current productivity and wages plus future expected input wedges.

(1) Adjustment costs: For expositional simplicity I focus on the case of symmetric convex (quadratic) adjustment costs where there is one input, associates (N_t). Adjustment costs are given by $\frac{b}{2}(N_{t+i} - N_{t+i-1})^2$. The firm's revenue production function is $F(A_t, N_t)$. The firm's expected discounted profit at time t is given by:

$$\Pi_t = \mathbb{E}_t[\sum_{i=0}^{\infty} \beta^i (F(A_{t+i}, N_{t+i}) - w_{t+i}N_{t+i} - \frac{b}{2}(N_{t+i} - N_{t+i-1})^2]$$

where $F(A_t, N_t)$ is the revenue production function and A_t are exogenous productivity shocks. Then the Euler equation is given by:

$$F_N(A_t, N_t) = w_t + b(N_t - N_{t-1}) - \beta b\mathbb{E}_t[N_{t+1} - N_t]$$

The term $b(N_t - N_{t-1}) - \beta b\mathbb{E}_t[N_{t+1} - N_t]$ creates a wedge between the marginal revenue productivity and the wage. If $(N_t - N_{t-1})$ and $(N_{t+1} - N_t)$ are positively correlated, then they will tend to offset each other and reduce the size of the wedge.

We can proceed further by rearranging to obtain $N_t - N_{t-1} = \frac{1}{b}F_N(A_t, N_t) - w_t + \beta\mathbb{E}_t[N_{t+1} - N_t]$. Iteratively substituting out $N_{t+1} - N_t$ in the Euler equation gives:

$$F_N(A_t, N_t) = w_t + b(N_t - N_{t-1}) - b\mathbb{E}_t[\sum_{i=1}^{\infty}(\beta)^i(F_N(A_{t+i}, N_{t+i}) - w_{t+i})]$$

In the data, wage increases are typically positively correlated with $N_t - N_{t-1}$. This correlation means that we would expect the marginal revenue productivity to increase by more than the increase in wage in the presence of adjustment costs (holding $\mathbb{E}_t[\sum_{i=1}^{\infty}(\beta)^i F_N(A_{t+i}, N_{t+i}) - w_{t+i}]$ fixed). The only other alternative is a discrete change in $\mathbb{E}_t[\sum_{i=1}^{\infty}(\beta)^i F_N(A_{t+i}, N_{t+i}) - w_{t+i}]$; which is a discounted sum of future input wedges. However, this means there is a large jump in the expected future gap between marginal revenue product and wages. Take the simplest case where the expected gap at time r between the marginal revenue product and wages for all future periods is some constant ψ_t. Then we have that the change in the Euler equation (from time $t-1$ to t) is given by:

$$\Delta F_N(A, N) = \Delta W + b(N_t - N_{t-2}) - \frac{\beta}{1-\beta}\Delta\psi$$

Discrete wage increases (ΔW) that are not accompanied by changes in $\Delta F_N(A, N)$ require large changes in future expected input wedges ($\Delta\psi$). Future productivity shocks affect current wages only if they affect the future wedge between marginal revenue productivity and wages.

(2) Firm-specific human capital: An alternative formulation is if new associates are required to gain firm-specific human capital before being productive. This dynamic means that associates in other firms are imperfect substitutes for a firm's own associates. For intuition, I will consider a simple model where experienced associates can be "grown" only within the firm. At time t, firms hire new associates N_t who produce no output in the current period. The associates then become experienced associates, E_t, at time $t+1$. After time $t+1$ they leave the firm, so they are productive only for one period. The firm's expected discounted profit at time t is given by:

$$\Pi_t = \mathbb{E}_t[\sum_{i=0}^{\infty} \beta^i (F(A_{t+i}, E_{t+i}) - w^n_{t+i} N_{t+i} - w^e_{t+i} E_{t+i})]$$

$$\text{where } E_{t+i} = N_{t+i-1} \; \forall i$$

The first order condition for N_t is given by:

$$\beta \mathbb{E}_t[F_E(A_{t+1}, N_t) - w^e_{t+1}] = w^n_t$$

so the firm will hire new associates until w^n_t equals the gap between the expected marginal revenue productivity and wages for experienced associates at time $t+1$. Therefore, an increase in w^N_t needs to be offset by either an increase in expected $F_E(A_{t+1}, N_t)$ or by a decrease in expected w^e_{t+1}. w^e_{t+1} typically increases at the same time as w^n_t, so the latter option seems unlikely.

A.1.3 Productivity and markups with general CES functions

In this section I will discuss more generalized measures of productivity and markups.

Productivity: In the main text, I discussed how to measure productivity when the production function is Cobb-Douglas. In this section I discuss extensions to the more general CES case. Take the general production function $Q_{it} = A_{it}(\alpha_i N^\rho_{it} + (1-\alpha_i) K^\rho_{it})^{\nu/\rho}$. Then we have:

$$\begin{aligned}
P_{it}\frac{\partial Q_{it}}{\partial N_{it}} &= \nu\alpha_i P_{it} N_{it}^{\rho-1} A_{it}(\alpha_i N_{it}^{\rho} + (1-\alpha_i)K_{it}^{\rho})^{\nu/\rho-1} \\
&= \nu\alpha_i P_{it}(\frac{Q_{it}}{N_{it}})(\frac{Q_{it}^{1/\nu}}{N_{it}})^{-\rho}
\end{aligned}$$

Consider the case where $\nu = 1$, i.e., constant returns to scale (P_{it} can still be decreasing in quantity). Then the above simplifies to:

$$\begin{aligned}
P_{it}\frac{\partial Q_{it}}{\partial N_{it}} &= \alpha_i(\frac{P_{it}Q_{it}}{N_{it}})^{1-\rho})P_{it}^{\rho} \\
\ln P_{it}\frac{\partial Q_{it}}{\partial N_{it}} &= \ln\alpha_i + (1-\rho)\ln\frac{P_{it}Q_{it}}{N_{it}} + \rho\ln P_{it}
\end{aligned}$$

Cobb-Douglas is the special case where $\rho = 0$. If we do not have a Cobb-Douglas production function, then we need to estimate ρ. Taking the ratio of the first order conditions for N_{it} and K_{it} from the profit maximization problem gives:

$$\begin{aligned}
\frac{w_{it}^{n}}{w_{it}^{k}} &= \frac{\alpha_i}{1-\alpha_i}(\frac{K_{it}}{N_{it}})^{\rho-1} \\
\ln\frac{w_{it}^{n}}{w_{it}^{k}} &= \ln\frac{\alpha_i}{1-\alpha_i} + (\rho-1)\ln\frac{K_{it}}{N_{it}}
\end{aligned}$$

Therefore, we can estimate ρ by looking at how relative employment and wages vary. There is limited cross-sectional variation in wages, so this ratio will be identified from time-series variation. I restrict the sample to firms that employ non-equity partners and assume the market compensation of partners is a constant firm-specific multiple of non-equity partner compensation. I include firm fixed effects. I estimate several variations: with average partner compensation or with average non-equity partner compensation and with and without fixed effects for five-year groups.

The results of this regression are in Appendix Table A.1. The estimated values of ρ are significantly greater than 0 and are around 0.7 and 0.9. ρ can be converted into the elasticity of substitution (σ) since $\sigma = \frac{1}{1-\rho}$. In all four specifications, $\sigma > 1$.

The next step is reproducing productivity estimates for various values of ρ. The outcome variable is $(1-\rho)\ln\frac{P_{it}Q_{it}}{N_{it}} + \rho \ln P_{it}$. I use the minimum hourly associate billing rate as a proxy for P_{it}. I estimate the regression using $\rho = 0, 0.5, 0.7, 0.9$ and plot the year fixed effects for each value of ρ. I omit the confidence intervals for visual clarity. The results are presented in Appendix Figure A.13. Using a CES production function generally decreases productivity estimates. Therefore, it is even more difficult to explain the wage jumps with more general CES production functions.

Markups: Suppose the total variable cost function ($VC(Q)$) is homogeneous of degree d. Then we have the relationship:

$$VC(\lambda Q) = \lambda^d VC(Q)$$

Taking the derivative of both sides with respect to λ and then setting $\lambda = 1$ gives:

$$\begin{aligned}
\frac{\partial VC(Q)}{\partial Q} Q &= dVC(Q) \\
\frac{\partial VC(Q)}{\partial Q} &= d\frac{VC(Q)}{Q} \\
MC(Q) &= d\frac{VC(Q)}{Q}
\end{aligned}$$

i.e., the marginal cost is equal to d times the average variable cost. The markup, μ is given by $\mu = \frac{P}{MC}$. Therefore we have:

$$\ln\mu = -\ln d + \ln\frac{PQ}{VC}$$

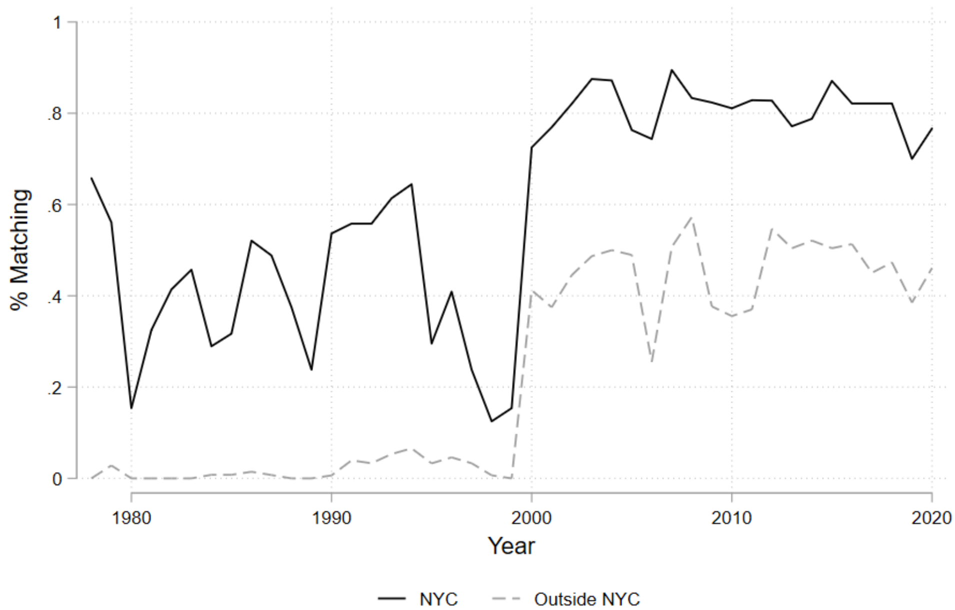
Accordingly, we can measure the change in markups using the ratio of revenue to variable costs if the variable cost function is homogeneous of degree d. A sufficient condition for the variable cost function to be homogeneous of degree d is if the production function is homogeneous of degree $\frac{1}{d}$. A major issue is classifying costs as either fixed or variable. Most costs for law firms are plausibly variable since they primarily consist of workers and rented office space, as opposed to large fixed capital investments. There are also minimal research and development or marketing costs compared to most industries. Therefore, total operating costs seem to be a reasonable estimate of variable costs.

A final issue is how to measure partner compensation. Equity partners earn compensation from directly providing labor, but they also earn compensation by providing capital. They are also the owners of the firm and will receive any direct profit. I use three alternative approaches. In the main analysis I impute the market wage of partners by using the observed non-equity partner wages in similar firms. In this section I try two alternative approaches: first, excluding all partner compensation; second, using only firm observations with non-equity partner wages (and excluding all firms without non-equity partners). I assume the market wage of equity partners is similar to that of non-equity partners within a firm.

I repeat the main analysis of looking at changes in markups around major wage changes with the alternative markups. The results are presented in Appendix Figure A.14. The alternative approaches to measuring markups do not change the main results significantly.

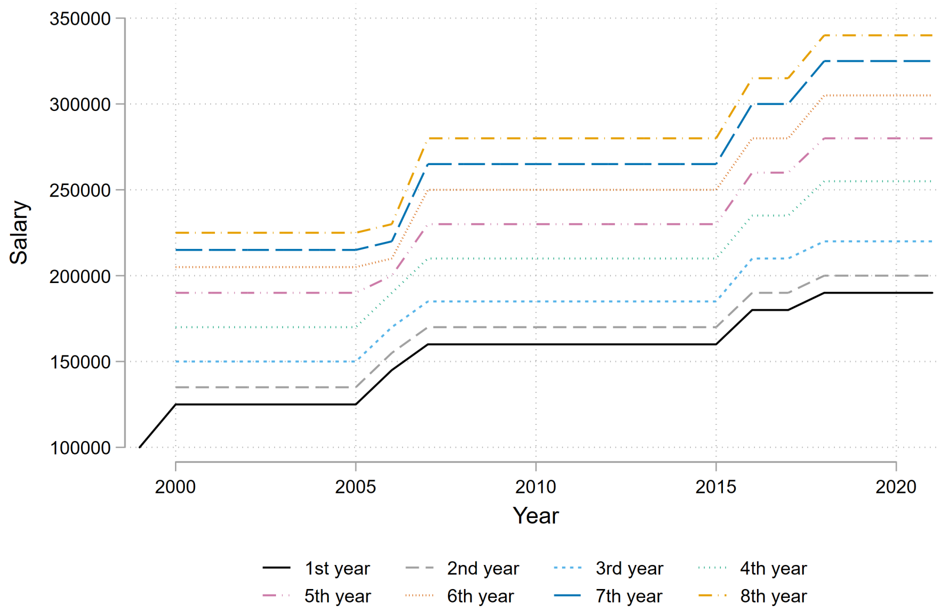
A.2 Appendix tables and figures

Figure A.1: Share of firms matching modal NY starting salary (NLJ 200)



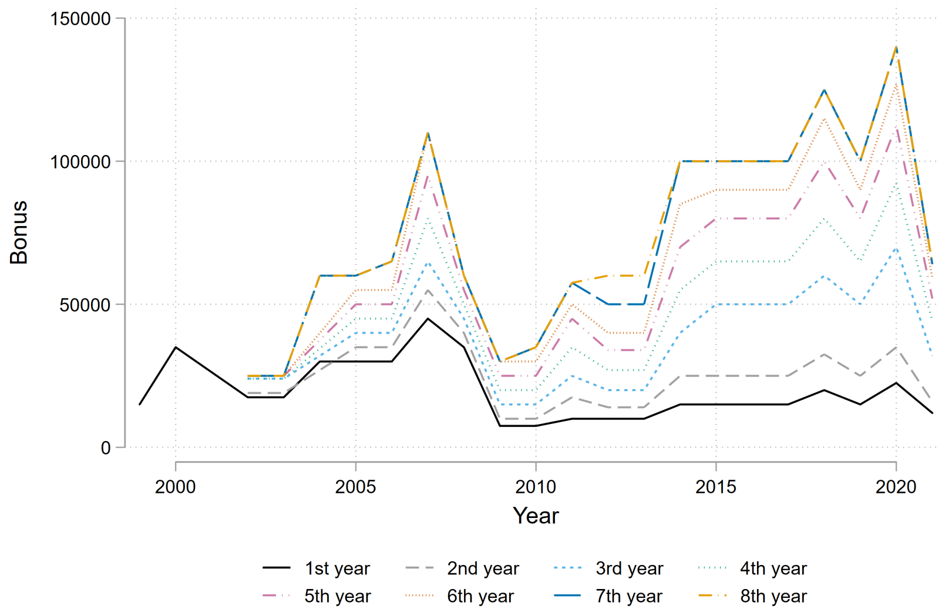
Note: Figures are conditional on firm's reporting starting salaries in NLJ surveys. Starting salaries represent the highest reported starting salary for each firm. New York firms are firms with headquarters in New York City.

Figure A.2: "Cravath" scale salaries by year and associate experience



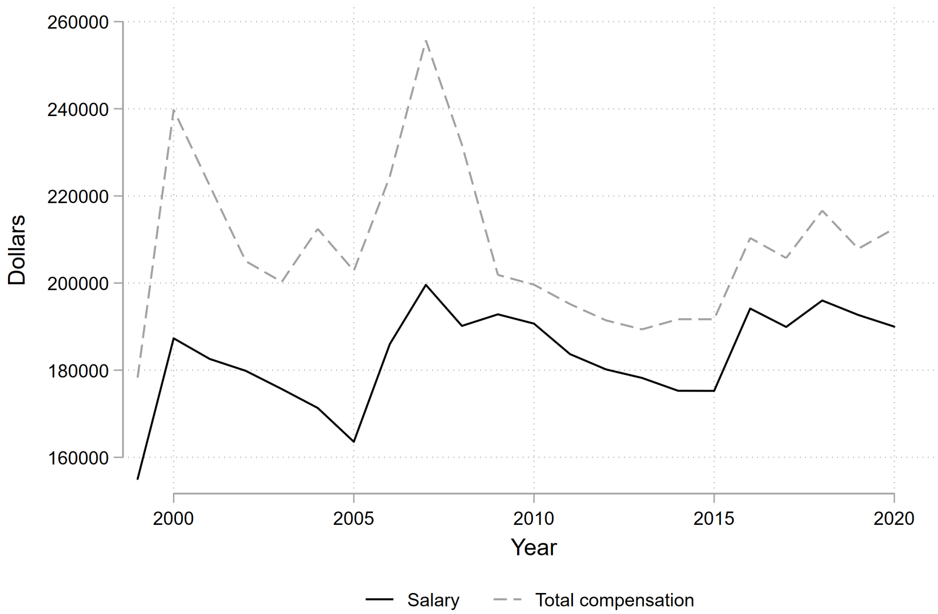
Note: Salaries are in nominal dollars. Source is BigLawInvestor.com.

Figure A.3: "Cravath" scale bonuses by year and associate experience



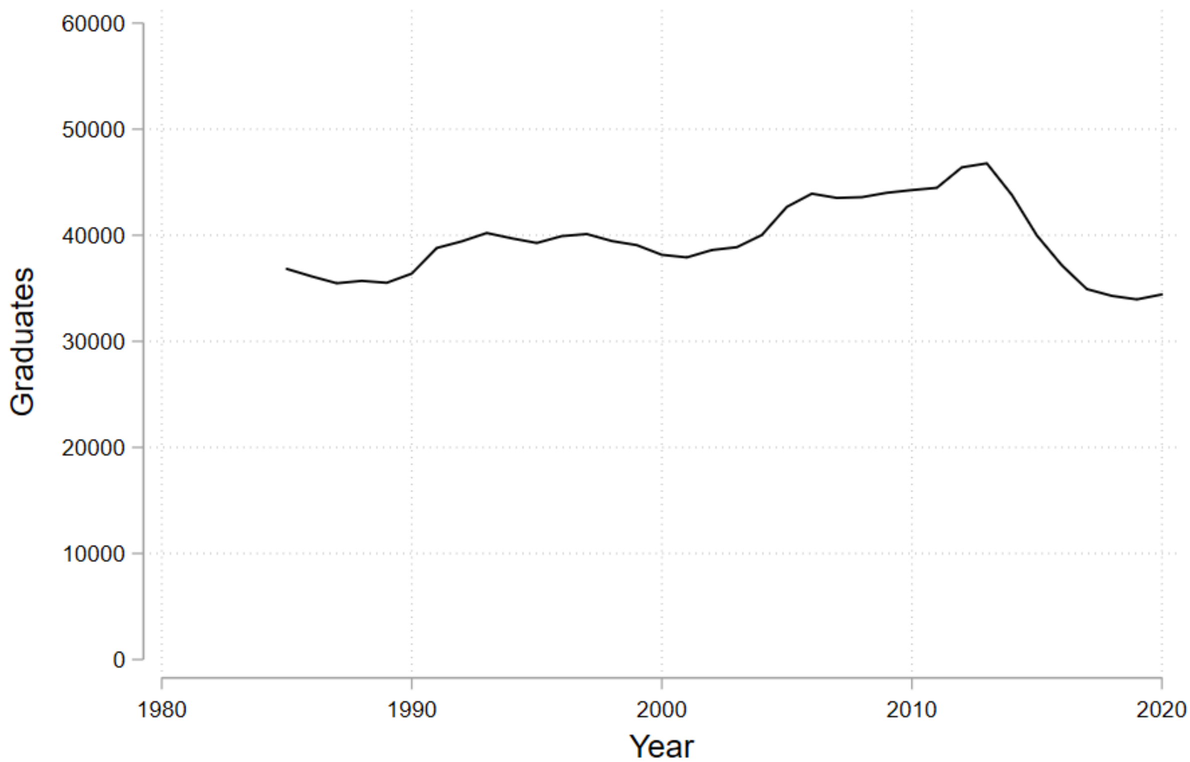
Note: Bonuses are in nominal dollars. Source is BigLawInvestor.com.

Figure A.4: "Cravath" scale total salary and total compensation for new associates (adjusted for inflation)



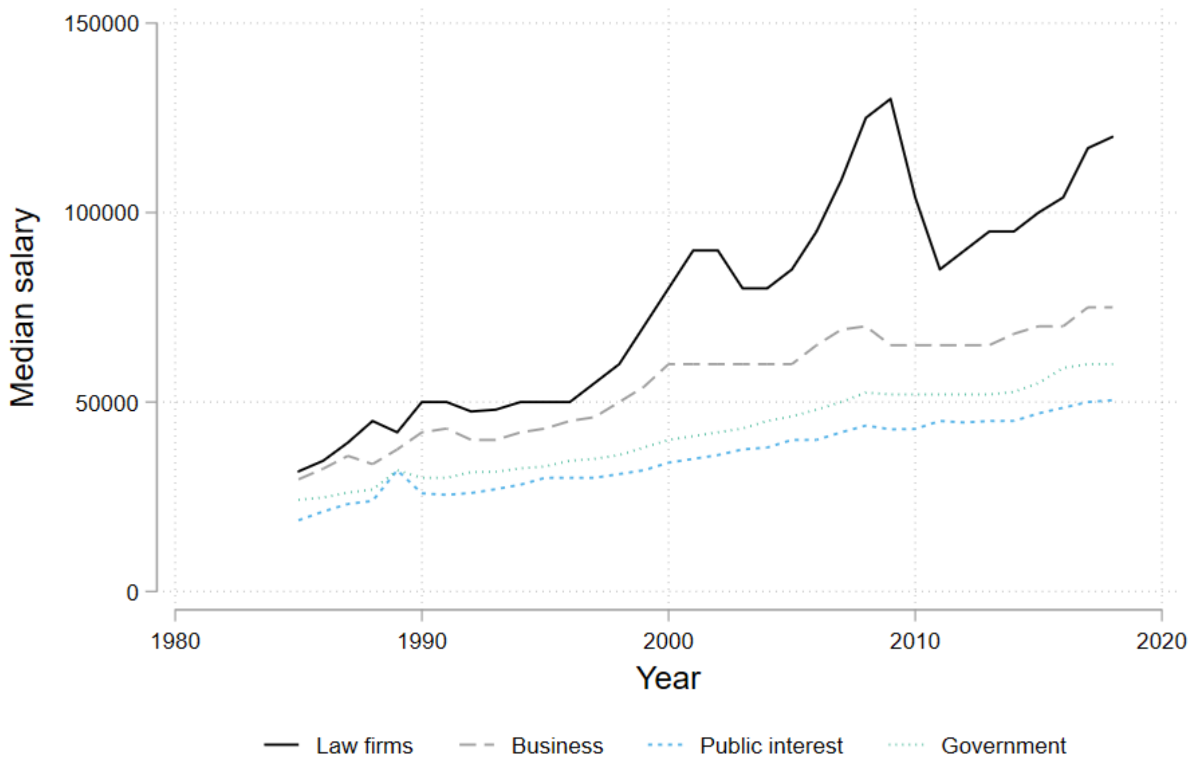
Note: Source is BigLawInvestor.com. Values are in 2020 dollars.

Figure A.5: Law school graduates by year



Note: Source is NALP figures.

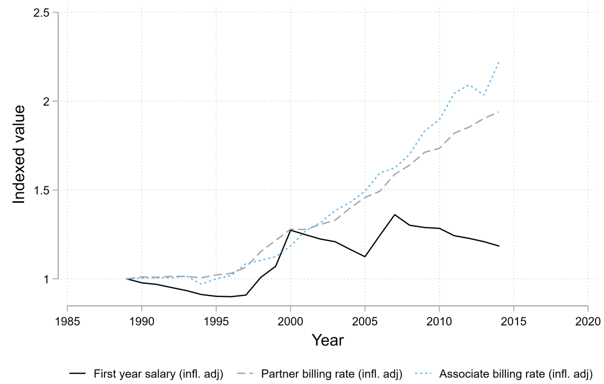
Figure A.6: Median earnings by initial industry for new law school graduates



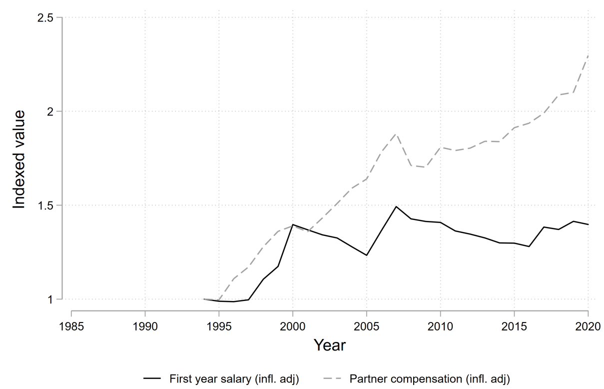
Note: Source is NALP figures. Values are in nominal dollars.

Figure A.7: Growth in real starting salaries relative to other firm metrics (AMLAW100)

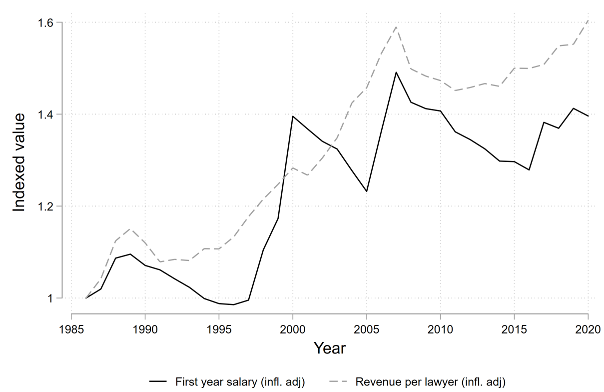
Panel A: Billing rates



Panel B: Partner compensation (incl. non-equity partners)

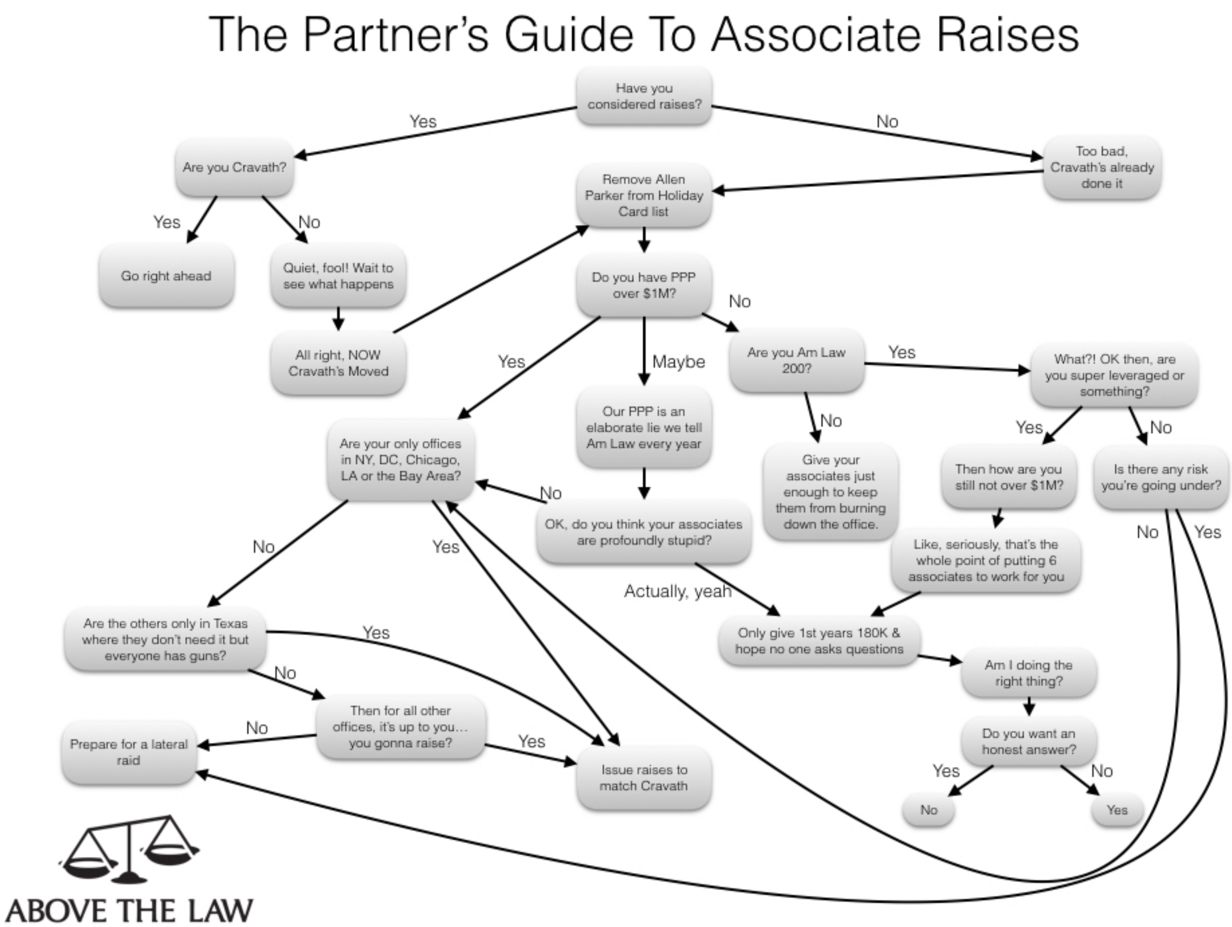


Panel C: Revenue per lawyer



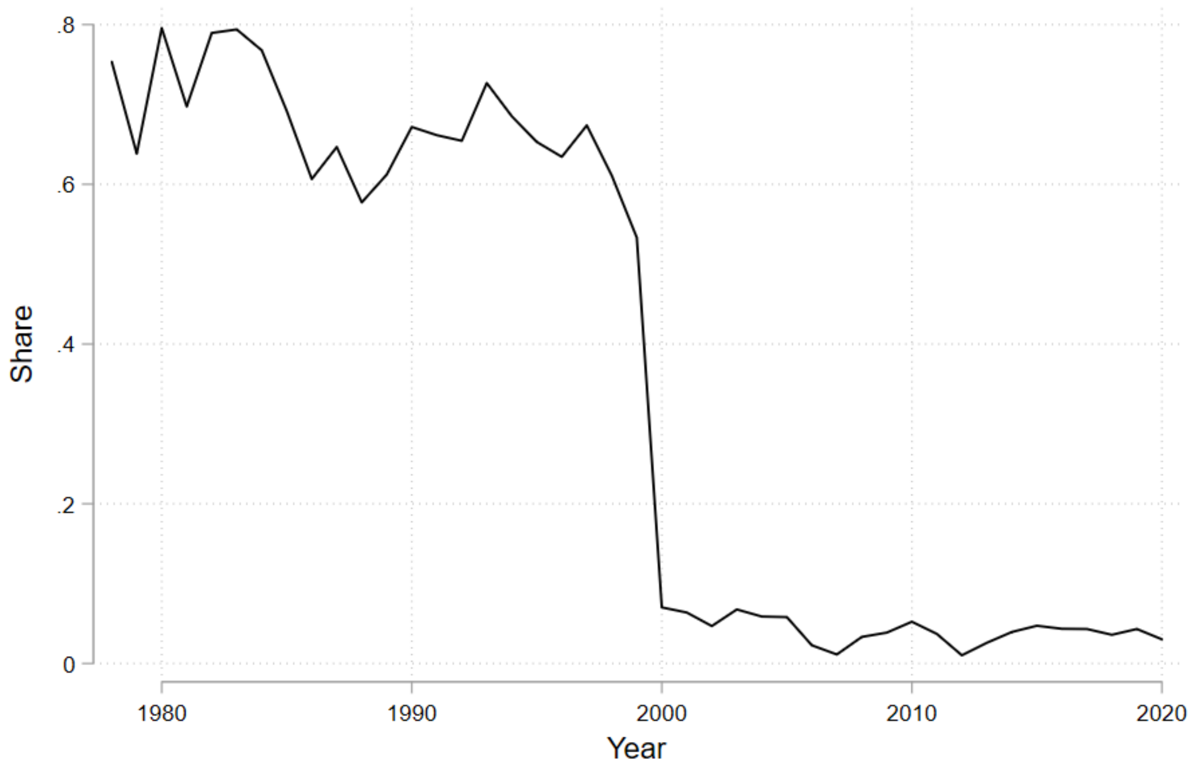
Note: Figures are conditional on firm's reporting starting salaries in NLJ surveys and featuring in the AMLAW100. Starting salaries represent the highest reported starting salary for each firm. Values are weighted by the total number of employed attorneys. Billing rates are the lowest reported partner and associate rates for each firm. For billing rates, growth is calculated by taking the average growth rate for each firm that reports rates in consecutive years due to higher rates of non-reported data in some years. Realized billing rates may be lower than reported rates. All values are adjusted to 2020 dollars using the CPI and indexed to 1986 levels.

Figure A.8: AboveTheLaw satirical guide to a firm's compensation decisions



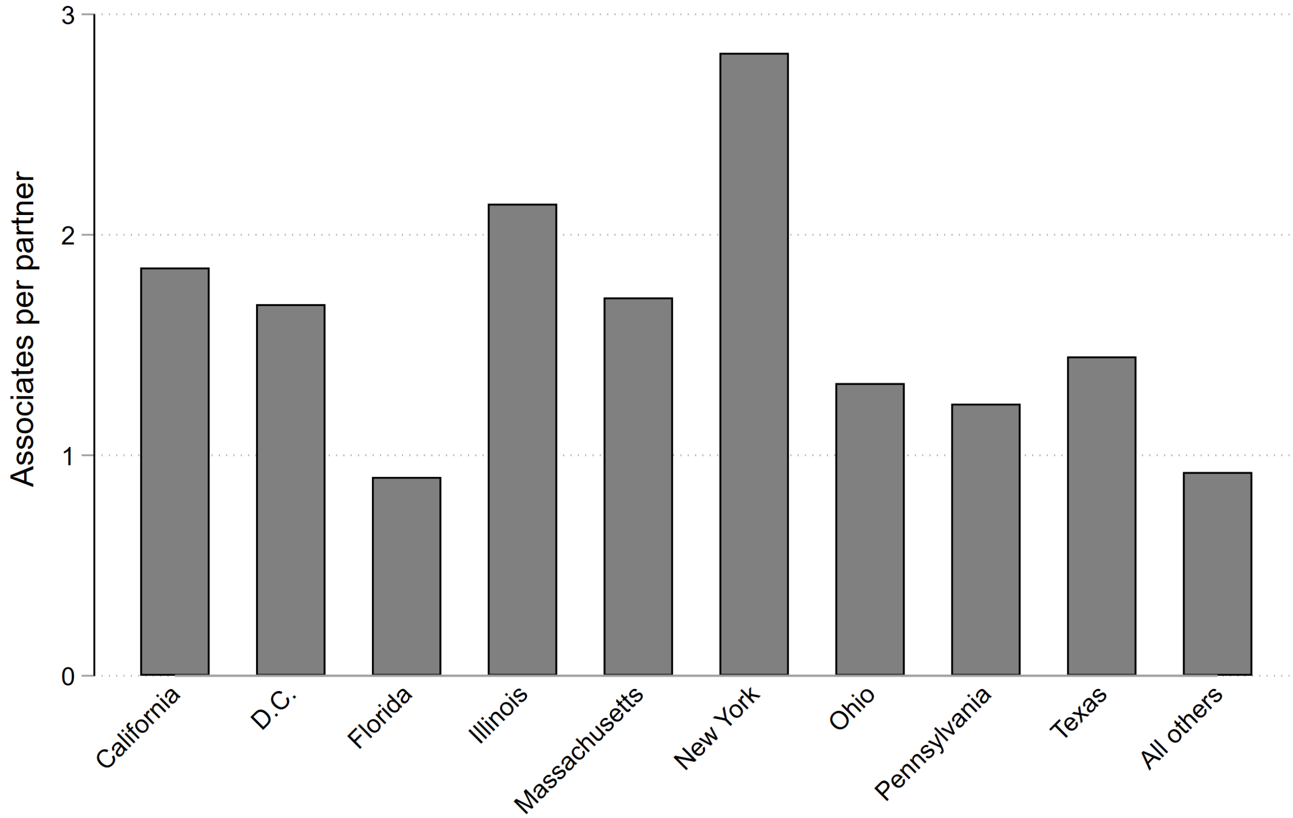
Note: Source is AboveTheLaw.com

Figure A.9: Share of firms reporting salaries that are not $5000 increments (NLJ200)



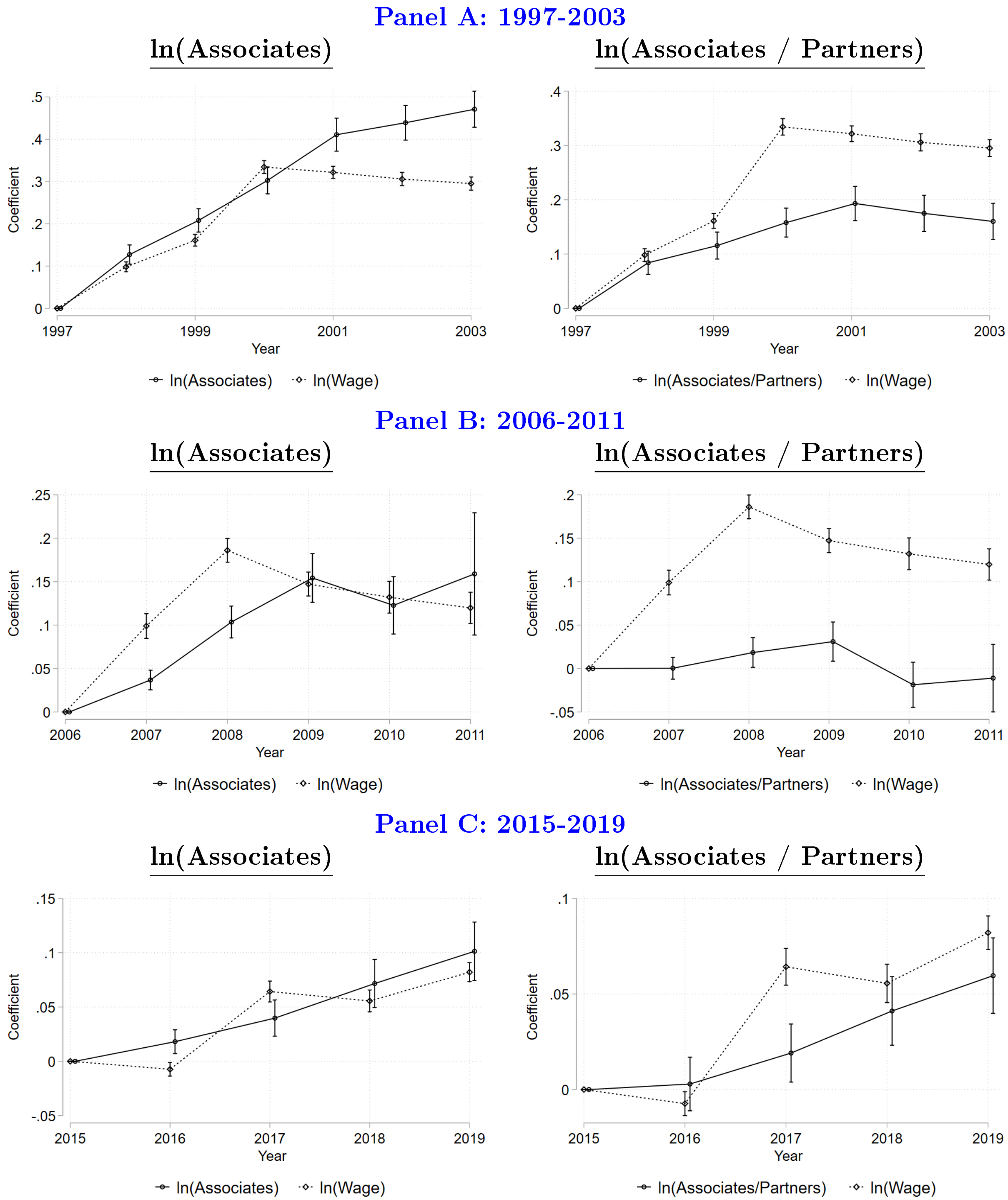
Note: Figures are conditional on firm's reporting starting salaries in NLJ surveys.

Figure A.10: Associates per partner by headquarter state (NLJ200, 2014)



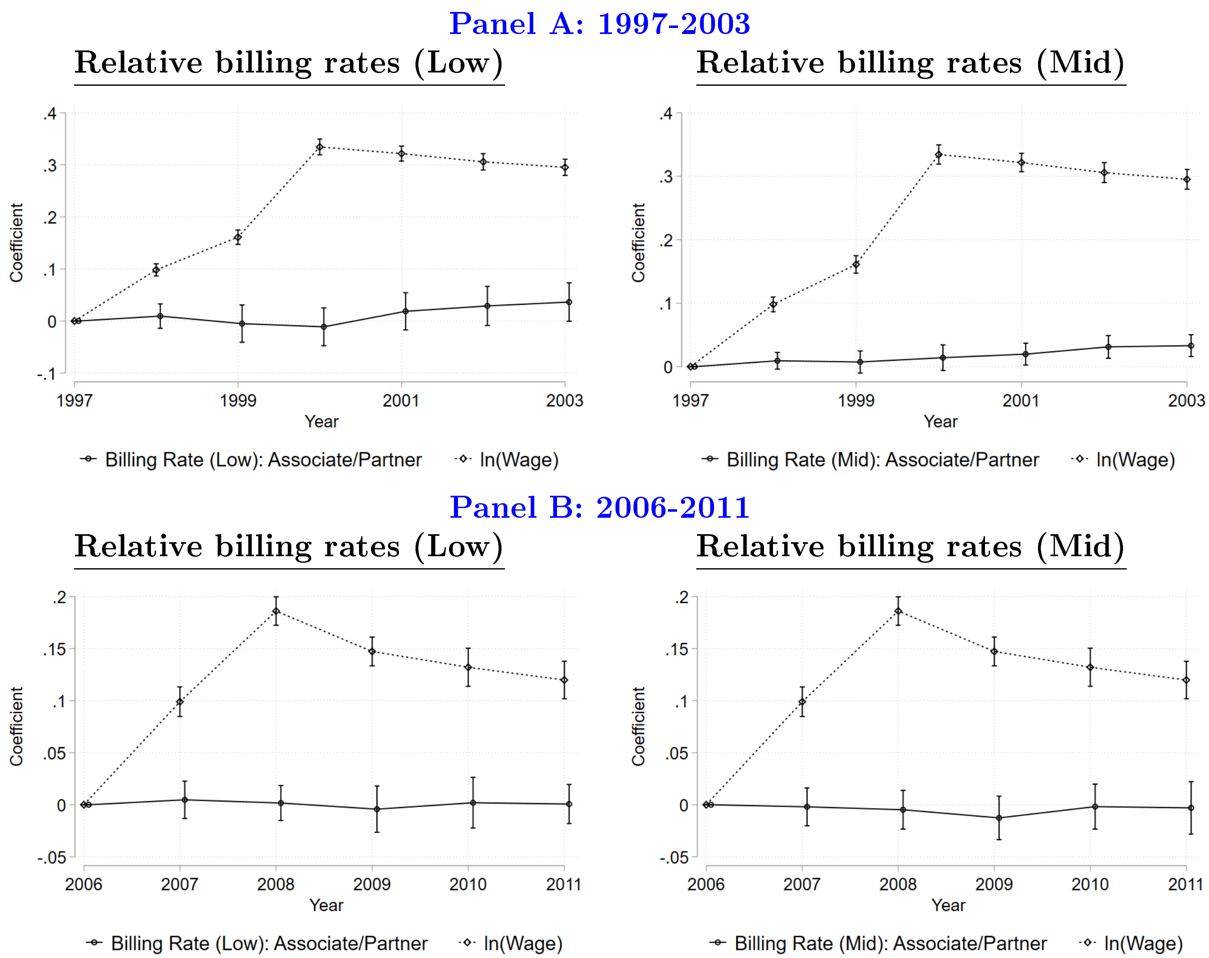
Note: Source is NLJ reports.

Figure A.11: Employment around major salary increases (NLJ 200)



Note: For NLJ 200 firms with reported starting salaries. Coefficients are for year fixed effects; the regression includes firm fixed effects. Regressions are weighted by initial attorney employment. Standard errors are clustered at the firm level. Bands represent 95% confidence intervals.

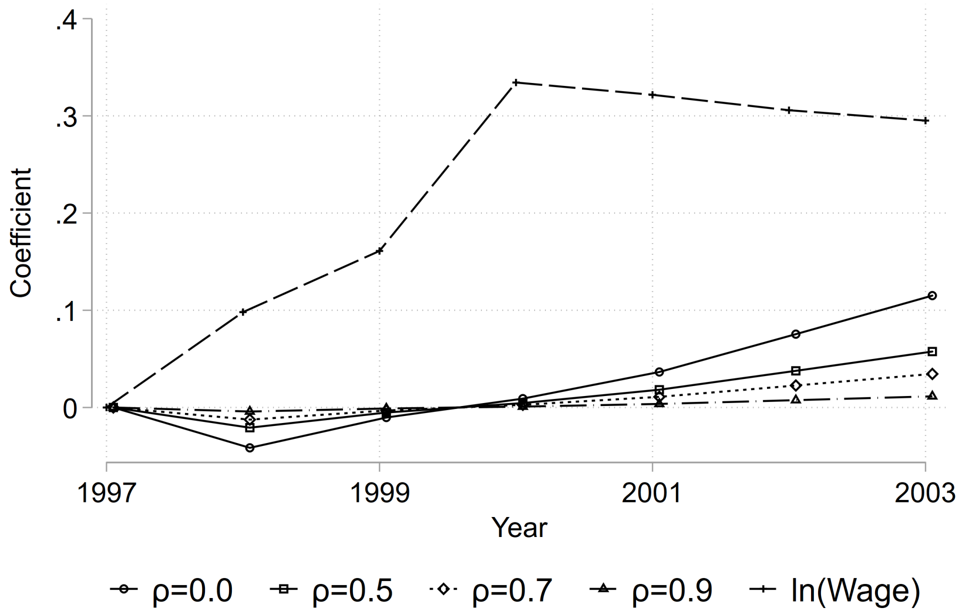
Figure A.12: Associate billing rates relative to partner billing rates around major salary increases (NLJ 200)



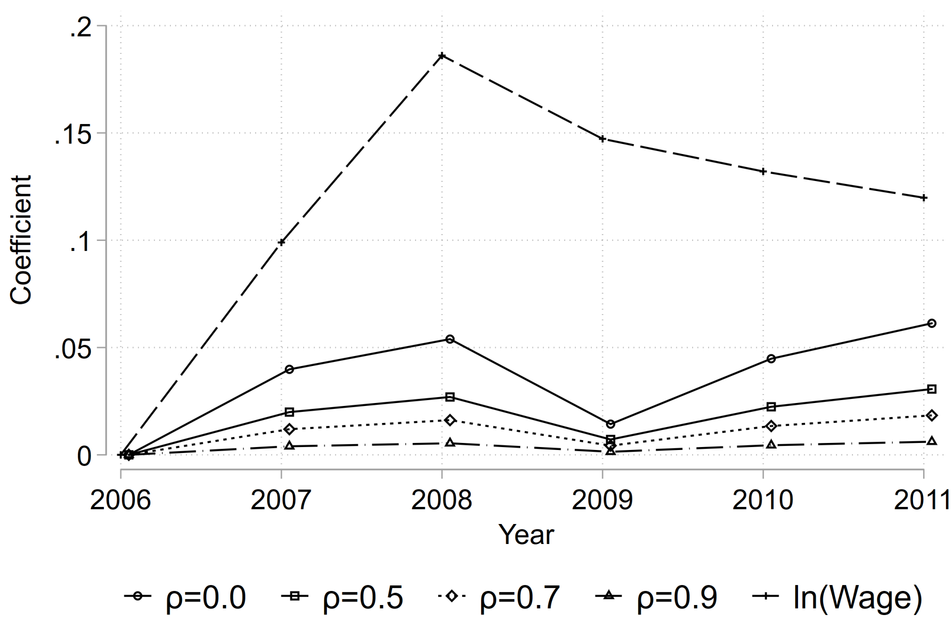
Note: For NLJ 200 firms with reported starting salaries and billing rates. Coefficients are for year fixed effects; the regression includes firm fixed effects. Regressions are weighted by initial attorney employment. Standard errors are clustered at the firm level. Bands represent 95% confidence intervals.

Figure A.13: Alternate productivity estimates (NLJ 200)

Panel A: 1997-2003



Panel B: 2006-2011

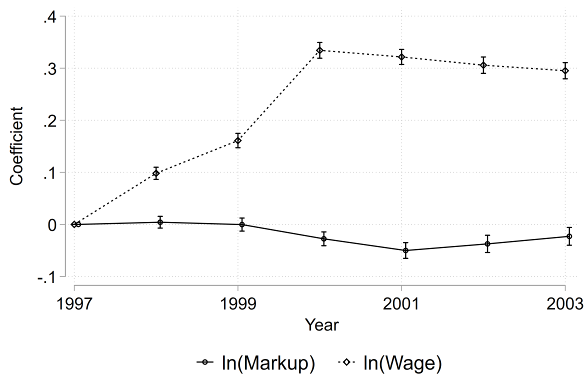


Note: For NLJ 200 firms with reported starting salaries and AMLAW data. Coefficients are for year fixed effects; the regression includes firm fixed effects. Regressions are weighted by initial attorney employment. Standard errors are clustered at the firm level. Bands represent 95% confidence intervals.

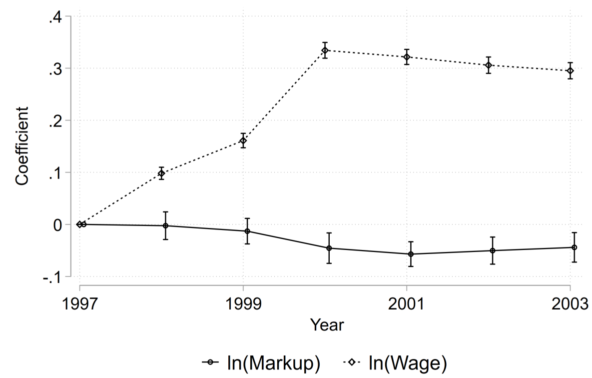
Figure A.14: Alternative markup estimates around major salary increases (NLJ 200)

Panel A: 1997-2003

Excluding equity partner compensation

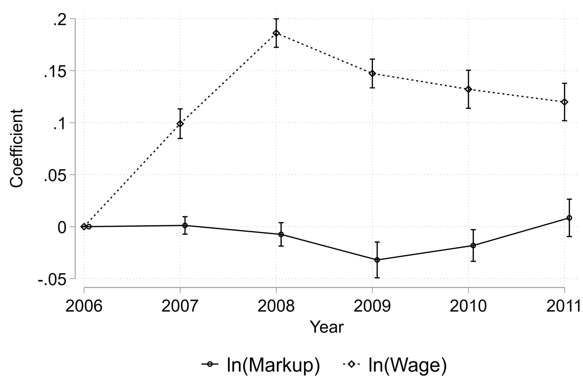


Only firms with non-equity partners

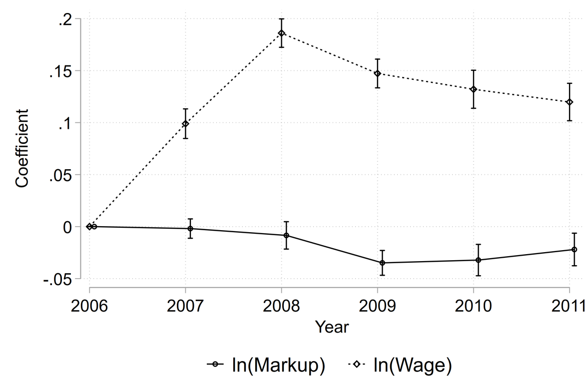


Panel B: 2006-2011

Excluding equity partner compensation

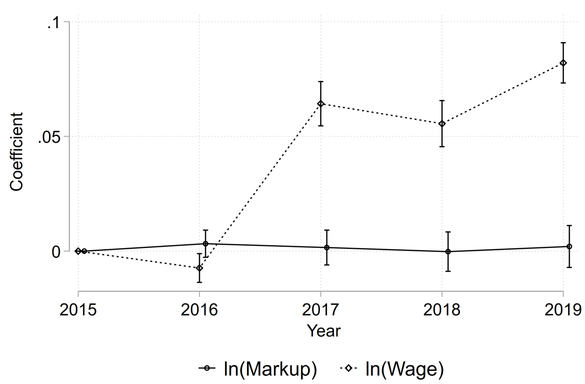


Only firms with non-equity partners

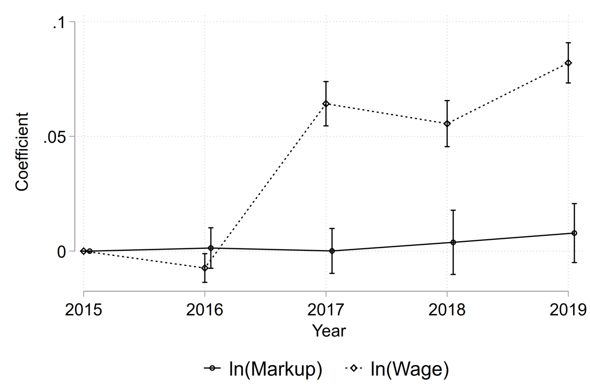


Panel C: 2015-2019

Excluding equity partner compensation

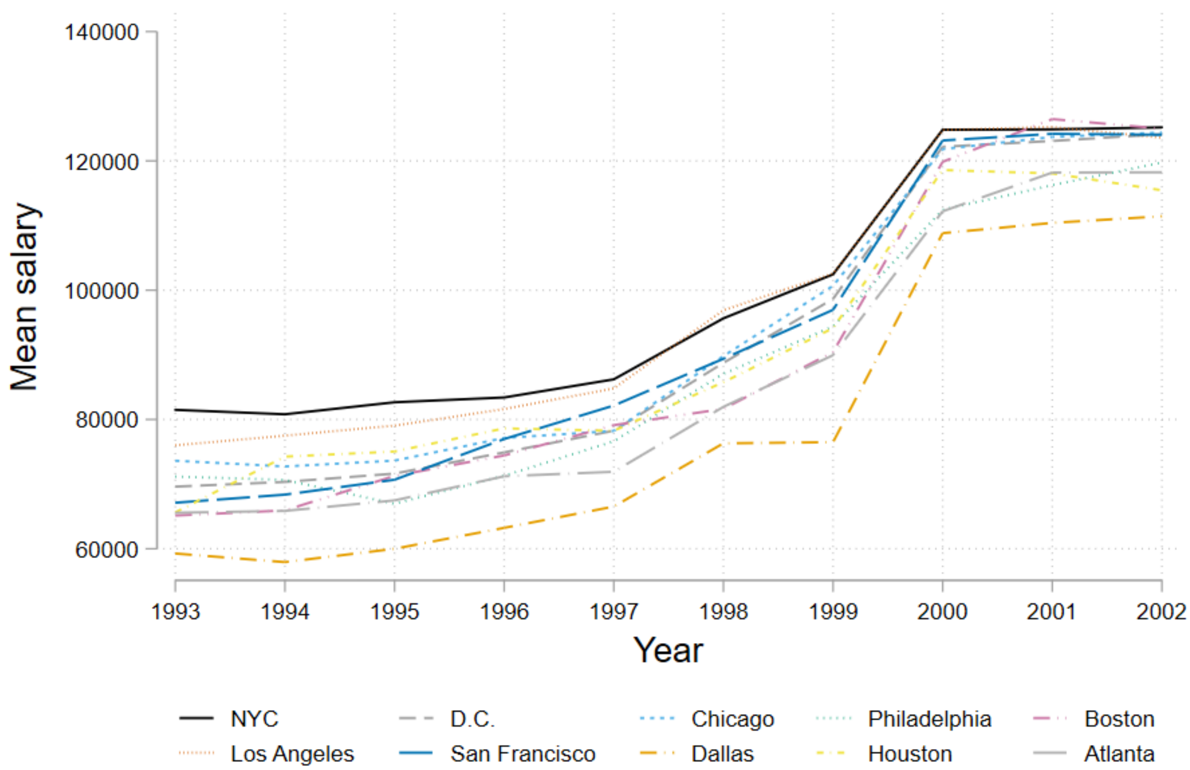


Only firms with non-equity partners



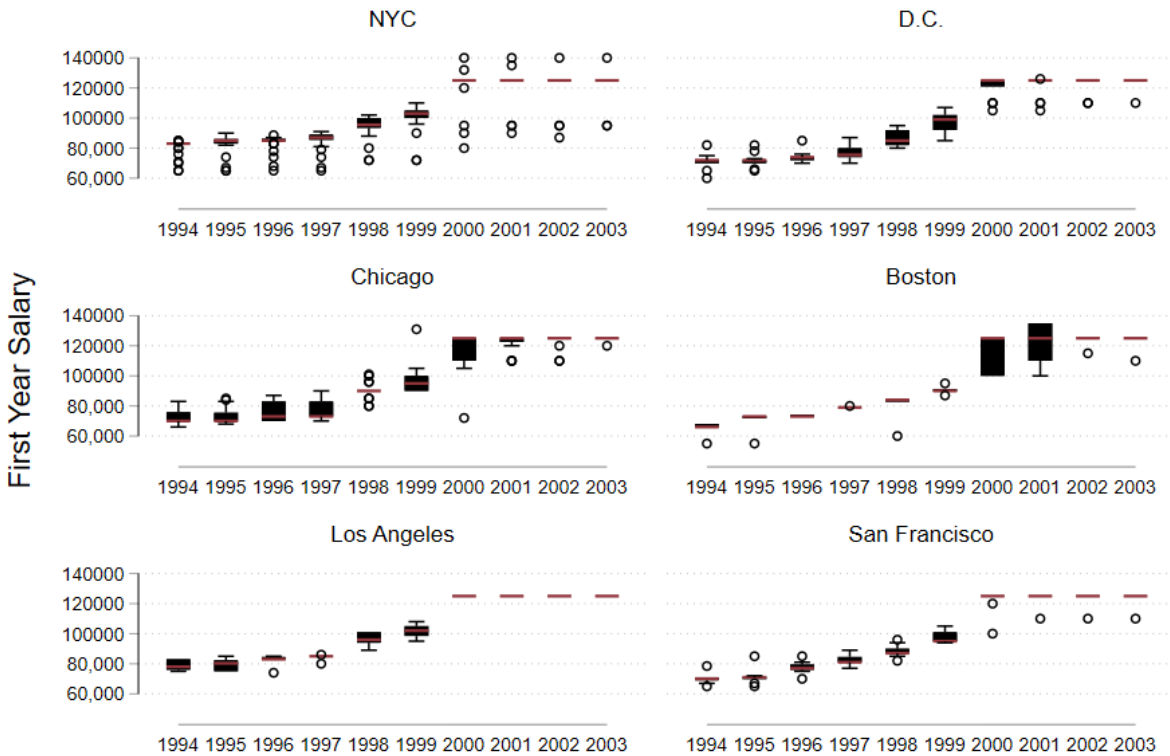
Note: For NLJ 200 firms with reported starting salaries and AMLAW data. Coefficients are for year fixed effects; the regression includes firm fixed effects. Regressions are weighted by initial attorney employment. Standard errors are clustered at the firm level. Bands represent 95% confidence intervals.

Figure A.15: Mean salary by city for 1993-2002 (NLJ200)



Note: Figures are conditional on firm's reporting starting salaries in NLJ surveys. Firms are assigned to cities based on the location of their headquarters.

Figure A.16: Interquartile salary range by city and year for 1993-2002 (NLJ200)



Note: Figures are conditional on firm's reporting starting salaries in NLJ surveys. Firms are assigned to cities based on the location of their headquarters. Only includes firms reporting salaries in at least seven of the years.

Figure A.17: Maximum salaries by city post-Cravath raise in 1968

SALARY SURVEY*

LAW FIRMS:

New York City	$13,000-15,000
Washington, D.C.	$13,500
Milwaukee	$13,500
Los Angeles	$12,000-13,500
San Francisco	$12,000-13,500
Chicago	$10,500-13,000
Cleveland	$12,500
Boston	$12,000-12,500
Detroit	$12,000
Grand Rapids	$12,000
Hartford	$12,000
Roanoke	$10,500-12,000
Atlanta	$10,800-11,000

FEDERAL GOVERNMENT:

GS-7 (Before admission to the Bar)	$ 6,981
GS-9	$ 8,462
GS-11	$10,203
GS-12 (Outstanding Academic Record)	$12,174

*Conducted and compiled by the Placement Office under the direction of Miss Eleanor Appel, Director. Note: In most instances these salaries represent the "going rate: which is actually the top salary." Many firms offer much lower salaries.

Note: As compiled by the Harvard Placement Office.

Table A.1: Relationship between change in log associates relative to partners and relative compensation

	(1)	(2)	(3)	(4)
	$\ln(\frac{\text{Associate Salary}}{\text{Avg. partner comp}})$		$\ln(\frac{\text{Associate Salary}}{\text{Avg. non-equity partner comp}})$	
ln(Associates / Partners)	-0.295***	-0.298***	-0.0938*	-0.123**
	(0.0541)	(0.0348)	(0.0499)	(0.0513)
Observations	4,711	4,711	3,081	3,081
R-squared	0.854	0.908	0.680	0.714
ρ	0.7	0.7	0.9	0.9
Elasticity of Subst.	3.4	3.4	10.7	8.1
Firm FE	X	X	X	X
Year Group FE	-	X	-	X

Notes: For NLJ 200 firms with data on revenue from Amlaw. Regressions are weighted by employment. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX B

Appendix Materials for Chapter 2

B.1 Data appendix

B.1.1 Census data

Individual Census records. The primary source of Census data are individual Census records from Ruggles et al. (2020). We use the 100% 1920, 5% 1930, 100% 1940, 1% 1950, 5% 1960, and 1% 1970 (metro) samples. Due to data and processing considerations, we take a 2% random sample of whites and 20% sub-sample of Blacks from the 1920 Census. Similarly, for the 1940 Census we use a 5% random sample of whites and for the 1960 we use a 40% sub-sample for whites (resulting in a 2% sample). Individuals in institutional group quarters are excluded.

Metro areas: Metro areas are based on the IPUMS variable "metaread." These are county based measures. Definitions vary slightly over time, but the basic qualification is the county must contain a city of at least 50,000 people or integrated with another county containing a qualifying city. Metro areas could expand or contract over time. Counties are identified in the 1940 and earlier samples, so a consistent county based definition is applied. Metro definitions were relatively unchanged between 1940 and 1950. There were more significant changes in the 1960 and 1970 Censuses. Several metro areas that were split between 1950 and 1960 are re-aggregated to maintain comparability. For long-term analysis (1920-1970), consistent metro definitions are imposed by using 1990 commuting zone boundaries. Earlier geographic boundaries (counties for 1940 and earlier, SEAs for 1950, PUMAs for 1960, and 1970 county groups for 1970) are crosswalked to commuting zones based on Eckert, Gvirtz and Peters (2018). Observations are weighted based on the geographic overlap between their geographic region and the commuting zone of interest.

Employment: Employment status is based on the IPUMS variable "empstat." This variable is not available in 1920. The reference period varies slightly across Censuses. In 1930 an individual is counted as employed if they were working on the most recent regular working day. In 1940 and later, an individual was counted as employed if they worked at all during the reference week. Prime-age employment is used as an outcome measure due to concerns about how the labor force and unemployment are measured across years. Prime-age workers are defined as individuals ages 25-54. Prime-age male employment is defined as the share of men in this age range who are employed.

Occupation and industry: Occupation and industry are coded using the 1950 Census coding system. Skilled occupations are defined as occupations falling in the following categories: "Professional, Technical"; "Managers, Officials, and Proprietors"; "Clerical and Kindred"; "Sales workers"; "Craftsmen"; or "Operatives" categories. This corresponds to occupational codes 000-093, 200-690 under the 1950 IPUMS occupational coding scheme. Semiskilled or skilled blue-collar occupations are a sub-category of skilled occupations that fall under the "Craftsmen" or "Operatives" categories. Occupational shares are constructed using currently employed individuals who are aged 14+.[1]

Defense industry is defined as mining, manufacturing, transportation, and government industries, following Collins (2001). These industries were most likely to be included in the War Manpower Commissions defense industry employment reports. These industries correspond to 1950 IPUMS industry codes 203-239, 306-499, 506-568, 906-946. Following Acemoglu, Autor and Lyle (2004), key defense industries are defined as durable goods manufacturing industries, and these correspond to IPUMS 1950 industry codes 326-388.

Employee: A worker is defined as an employee based on the IPUMS variable "classwkrd." Employees are defined as individuals who are currently in the labor force and have classwkrd

[1] Results are robust to using share of all individuals in the labor force instead.

codes 20-28, which corresponds to categories "Works on Salary," "Wage/salary, private," and "Wage/salary, government."

Wage income: Yearly wage income is created using the IPUMS variable "incwage". This variable comes from the Census question asking for each person's total pre-tax wage and salary income. This question was first introduced in 1940. Yearly wage income is specifically payments for work done as an employee; it excludes self-employment income or personal business income. This restriction is especially relevant for farmers. Unfortunately, the 1940 Census did not ask for information on business or other sources of income. The wage income sample is restricted to individuals who are (1) are employees at the time of the Census, (2) are employed at the time of the Census, and (3) their primary occupation is not farmer or unpaid family farm laborer. Only sample line respondents were asked about wage income in the 1950 Census.

An additional issue is how to deal with top-coded values or implausibly low earnings totals. We follow Goldin and Margo (1992) by multiplying top-coded values by 1.4 and recoding as missing values that are less than 1/2 the minimum weekly wage. This corresponds to weekly earnings below \$6 in 1940, \$8 in 1950, \$20 in 1960 and \$28 in 1970.

Education: Years of education is created based on the IPUMS variable "educ". Individuals with five or more years of college are all coded as having seventeen years of education. Individuals with twelve years of completed schooling are assumed to have completed high school. This question was first asked in the 1940 Census. Only sample line respondents were asked about highest completed grade in the 1950 Census.

The IPUMS variable "school" is used to classify whether a child is currently attending school. The question changed slightly across Census years but was relatively consistent from 1940 to 1950. The main changes across Censuses are (1) length of retrospective reference period and (2) qualifying educational institutions. The retrospective reference periods are: previous four months in 1920, previous six months in 1930, previous month in 1940, and previous two months for 1950 and on. Qualifying educational institutions are: any type of school in 1920, any school or night school in 1930, any school and night school/extension programs if part of a regular school system in 1940 and 1950, any school that advances a person towards high school or college degree in 1960. Across all years the respondent has to indicate only whether the person has attended a qualifying institution in the reference period; they do not need to regularly attend. Only sample line respondents were asked about school attendance in the 1950 Census.

Census aggregates. Census aggregates are taken from published Census volumes. County population totals by age and race are taken from ICPSR 02896 (Haines and ICPSR, 2010). Manufacturing output and value added from the Census of Manufactures are taken from the same source. We also digitized new metro-level data from the 1950 Census. The only individual level data available for 1950 is the 1% sample. This means we limited Black observations for metro areas with small Black populations. From Volume II of the 1950 Census of Population we digitize the following: Table 77, which has total employment for each metro by race-sex-occupation, Table 83, which has total employment for each metro by race-sex-industry, and Table 87, which has total counts for each metro by race-sex-income bin, as well as the median income by race and sex. We use this data rather than totals from individual counts whenever possible. For heterogeneity analysis (e.g., by age or education), we rely on the individual Census data.

B.1.2 Draft rate

Creating a predicted draft rate. We use a predicted draft rate rather than actual draft or enlistment rate. A predicted draft rate is created for each metro area by using draft records to identify national draft rates by group and then applying these draft rates to the baseline demographics for each metro in 1940.

The drawback of the predicted draft measure as a control is that it will not control for all sources of variation in draft rates. For example, some areas might have had stricter draft boards or a higher share of individuals who did not meet minimum military standards. However, we believe we are capturing the largest source of exogenous variation in draft rates.

If the predicted draft rate does not affect our estimates, then it is less likely these smaller sources of variation would meaningfully alter our results.[2]

Our primary source of draft data are the WWII Army Enlistment Records provided by the National Archives and Records Administration. This data series contains the records of about nine million men and women who enlisted in the U.S. Army. The records typically contain the serial number, name, place of residence, place and date of enlistment, education, occupation, marital status, and race of the enlistee. There are several gaps in the records. First, the data is only for the U.S. Army, so it excludes other service branches, such as the Navy (although the Air Force was still part of the Army during WWII). Second, some records are known to be missing. Finally, some of the scanned records are unusable due to poor scans.

The secondary source of data if from the Selective Service System (1956). We digitized tables reporting total inductions and enlistments by service branch, month, and race. This data identifies how many inductions are not captured in the individual enlistment data for each month. We re-weight the individual observations by the number of records missing in their enlistment month. For example, if the individual records cover half of total inductions in a given month then the observed inductions will be given double the weight. Implicitly, this also assumes that individuals drafted into the Navy in any given month have similar characteristics to individuals drafted into the Army in the same month, conditional on race. The reason we re-weight the observations within a month is that draft eligibility and probabilities changed throughout the war. For example, initially individuals younger than 21 were not eligible for the draft, but later in the war the minimum age eligibility was reduced to 18. We condition on race because there is evidence the Army was much more willing to accept Black men than the Navy.[3]

Only records of enlistments between January 1940 and December 1945 are included. Individuals who are younger than 17 or older than 45 at time of enlistment are dropped. We

[2]Our results are also robust to using the actual draft rate. Results available upon request.

[3]Black men served almost exclusively in mess units for much of the war in the Navy.

restrict the sample to individuals who were drafted based on their serial codes. Serial codes that start with three or four indicate that the individual was drafted.

We next find the total number of individuals drafted each year by demographic group. We create demographic cells using race, year of birth, nativity, and marital status. All of these variables were important determinants of draft probabilities.

The next step is to create draft rates by demographic group and year. We use 1940 Census data to determine the population in each demographic cell. Most of the characteristics are time-invariant, except for marital status. Marital status was one of the key determinants for whether someone was drafted. There is also significant variation across metro areas in typical age at marriage and marriage rates. We create marriage hazard rates using marriage rates across cohorts in 1940. We then follow each cohort from 1940 to 1945; first applying the marriage hazard rate and then subtracting the number of individuals who enlisted each year. For the next year, we apply the marriage hazard rate to the remaining single individuals and repeat the process. Therefore, for each year we have the number of married and single individuals for each demographic group in each cohort. To create the draft rate for that year and demographic group, we then take the total number of individuals drafted and divide by the population in each given demographic group.

The final step is to apply these draft rates to each metro. We follow a similar process to the above to generate the number of married and single individuals for each demographic group in each year, using metro specific marriage hazard rates for whites and a national rate for Blacks. We use a national rate for Blacks because some metro areas have small Black populations, making it difficult to calculate metro specific hazard rates. Once we have the number of individuals in each demographic group for each year we apply the national draft rates for that group and aggregate to create the total number of individuals drafted in each metro. We then divide by the male population ages 15-64 in 1940.

We also create an actual draft rate measure for comparison and to see if our predicted draft rate does predict the actual draft rate. When calculating the actual draft rate, we scale the denominator by an estimate of the population growth between 1940 and 1943. We do this to account for the fact that a large number of people migrated during the first part of the war. Without this correction, using 1940 population as a denominator would lead to higher draft rates in areas with higher net in-migration. Since war expenditure increase migration (see Appendix Section B.3.5) this would create a positive correlation between draft rate and war expenditures.

Results. Appendix Table B.4 shows that our predicted draft rate is a strong predictor of the actual draft rate. Therefore, it seems as though it is a valid measure of draft intensity. Table B.4 shows that both the predicted and actual draft rate are negatively correlated with war expenditures. Moreover, the predicted draft rate is positively correlated with labor shortages, even conditional on war expenditures. This correlation provides evidence that our predicted draft variable has the expected consequence on labor supply.

B.1.3 Other data

We supplement the above data with several additional data sources.

Labor shortage data. We digitized reports on the extent of labor shortages during WWII by month. These reports were from the monthly Labor Market Reports compiled by the War Manpower Commission. These classified labor markets by whether they were facing labor shortages. Labor shortages were defined based on comparing expected hiring to the number of people expected to be looking for work, combined with subjective adjustments by government officials.

We create our measure of labor market shortages by taking the percentage of months between 1942 and 1944 that the labor market experienced severe labor shortages (on the map this corresponded to labor markets with completely shaded circles). About 20% of metro-month observations were coded as severe labor shortages.

Defense industry employment during WWII. The War Manpower Commission regularly surveyed employers in war industries or critical labor markets on their employment. These surveys were ES-270 reports. These reports did not cover the entire labor market but did cover a large share of war industry employment. For more details and examples of usage in other research, please see Collins (2001) or Rose (2018).

B.2 Labor market context

B.2.1 Occupational distribution and changes 1940-50

First, some notes on occupational categories. The aggregate occupational categories are "Profession, Technical"; "Farmers"; "Managers, Officials, and Proprietors"; "Clerical and Kindred"; "Sales workers"; "Craftsmen"; "Operatives"; "Domestic Service"; "Service"; "Farm Laborer"; and "Laborer". Appendix Figure B.2 shows the average wages and education for white men by occupation. The occupations are colored based on which aggregate occupational category they belong to. It is clear that occupations in the "Domestic Service"; "Service"; "Farm Laborer"; and "Laborer" pay significantly less on average and also employ workers with lower education levels.

Appendix Table B.1 shows the occupational distribution for white and Black men in 1940 and 1950. Several facts are immediately clear. First, Black and white men have very different occupational distributions, with Black men being concentrated in unskilled occupations. As seen in Appendix Figure B.2, these are the occupations with the lowest pay and lowest education. Secondly, the occupational distribution for Black men significantly changed between 1940 and 1950, with large increases in the "Craftsmen" and "Operatives" categories. These observations are consistent with Collins (2000). The occupational distribution for white men changed as well but to a much lesser extent. These results are consistent with the finding of occupational upgrading for Black men in Collins (2000). These facts provide preliminary motivation for our focus on the impact of WWII expenditures on occupational upgrading for Black men.

B.2.2 Occupational segregation

An immediate question is to what extent these occupational differences between Black and white men can be explained by differences in education or location. For example, Black men were much more likely to live in the South and less likely to live in metro areas and had significantly less education on average. However, there are plenty of examples of explicit discrimination, for example Appendix Table B.2 lists a number of unions with explicit or effective bars on Black membership. There are two interesting questions to ask: first, which occupations seem to be most segregated, and second, which metro areas seem to be most segregated?

First, we compare segregation across occupations by looking at the expected number of Black workers, based on random allocation within education group and region, and compare it to the actual number of workers. We restrict the sample to men living in metro areas who are employed at the time of the Census. We define education groups as 0-5 years, 6-8 years, 9-11 years, 12-15 years, and 16+ years. Following Margo (1995), we account for school quality differences by multiplying years of education by 0.85 for Black men born in the South with less than 15 years of education. This adjustment roughly corresponds to the difference in average school term length between Blacks and whites in segregated Southern schools during the 1920s. Occupations are defined using the occupation and industry categories in Table 77 of the state breakouts in the 1950 Census Volume II. The number of expected Black workers is calculated by:

$$Expected_o = \sum_r \sum_e \left(\frac{Black_{re}}{Pop_{re}} * Positions_{ore}\right)$$

Where $\frac{Black_{re}}{Pop_{re}}$ is the share of Black men within region r and educational group e and $Positions_{ore}$ is the number of positions in occupation o held by men in region r and educational group e. To get a measure of the gap for each occupation we then divide by the actual number of Black men observed in occupation o. Appendix Table B.3 reports the occupations with the top fifteen largest and smallest ratios of expected vs. actual employment.

A second question of interest is comparing occupational segregation across regions. A natural index to measure occupational segregation is the Duncan index (Duncan and Duncan, 1955). The Duncan index is defined as:

$$Duncan_r = \sum_o |\frac{Black_{or}}{Black_r} - \frac{White_{or}}{White_r}|$$

Fundamentally, this index is a measure of "evenness," i.e., how evenly are Black men distributed across occupations. There are two related issues with this metric. First, if Black men are a small percentage of the population or many occupations have few positions then there will be substantial deviations from evenness due to pure chance as noted in Carrington and Troske (1997). Secondly, this metric does not distinguish between differences due to education versus occupational segregation. While occupational segregation can cause educational differences, our focus here is on occupational segregation conditional on education.

We can adjust the Duncan index by estimating the expected Duncan index, $E[Duncan_r]$, if workers are allocated randomly across jobs conditional on education and calculating the adjusted index:

$$Duncan_r^{Adj} = \frac{1}{2}\frac{Duncan_r - E[Duncan_r]}{1 - E[Duncan_r]}$$

We calculate $E[Duncan_r]$ by simulating fifty random occupational distributions for each metro area where the number of Black individuals in each occupation and education group is simulated using binomials where the probability of "success" is the share of Black men within the relevant education group.

Appendix Figure B.1 displays a map where the shading corresponds to the value of the adjusted Duncan index. The primary results are that there is substantial occupational segregation and that the segregation is not limited to the South.

B.2.3 Oaxaca-Blinder decomposition

Another way of examining the labor market context is to decompose wage differences in an Oaxaca-Blinder framework. We can decompose the aggregate wage gap into the portion that can be explained by differences in observables and the portion that cannot be explained by observables. The cross-sectional regression of (log of) wage on observables is:

$$Y_i = \beta X_i + \varepsilon_i$$

Evaluating the OLS estimate at the mean values gives:

$$\bar{Y} = \hat{\beta}\bar{X}$$

The difference between Black and white outcomes can be decomposed into:

$$\bar{Y}^{Wh} - \bar{Y}^{Bl} = \underbrace{\beta^{\hat{W}h}(\bar{X}^{Wh} - \bar{X}^{Bl})}_{\text{Observables}} + \underbrace{(\hat{\beta}^{Wh} - \hat{\beta}^{Bl})\bar{X}^{Bl}}_{\text{Unobservables}}$$

where the first term gives the portion of the wage gap that can be explained by observable differences and the second portion cannot be explained by observable characteristics.[4] For 1940 and 1950, we regress (log of) wages on a set of variables for education (years of education interacted with region of birth, whether graduated high school, whether graduated college), occupation (indicators for eleven aggregate occupation categories), industry (indicators for twelve aggregate industry categories), region, and a cubic in age. The resulting decompositions are given in the first two columns of Appendix Table B.10. We restrict the sample to native born men living in metro areas.

There are several key results. First, there is a large wage gap, but it declines significantly between 1940 and 1950 - declining from 0.63 log points in 1940 to 0.38 log points in 1950.

[4]We evaluate the gap at the coefficient values for white men. We could have evaluated the gap at the coefficients for Black men or some combination of the two, but alternative approaches do not change our qualitative findings.

Second, education and occupation differences are the most important observable factors. Third, there is still a large portion of the gap that cannot be explained by observable characteristics.

We can take this decomposition a step further and decompose the changes between 1940 and 1950. We focus on decomposing the change in the explained gap into a "price" effect and "pure" effect. The price effect is due to changing coefficient values that benefit one race relatively more than the other (changes in β). The pure effect is due to relative changes in observables (changes in $\bar{X}$).

$$\Delta_{40-50}\beta^{\hat{Wh}}(\bar{X}^{Wh} - \bar{X}^{Bl}) = \underbrace{\beta^{\hat{Wh},50}(\Delta_{40-50}\bar{X}^{Wh} - \Delta_{40-50}\bar{X}^{Bl})}_{\text{Pure}} + \underbrace{\Delta_{40-50}\hat{\beta}^{Wh}(\bar{X}^{Wh,40} - \bar{X}^{Bl,40})}_{\text{Price}}$$

The results are given in the last three columns of Appendix Table B.10. Overall, education and occupation changes explain most of the decline in the wage gap due to observables. The change in the gap due to education is almost entirely due to price effects (lower returns to education). On the other hand, the change in the gap due to occupation is due to both price effects (relatively higher returns for occupations with more Black men) and pure effects (Black men changing occupation). The price effects are consistent with the finding in Margo (1995) of wage compression across education groups and occupations that relatively benefited Black men. However, we also observe meaningful changes in the wage gap due to changes in the occupational composition of Black men, which is consistent with occupational upgrading.

B.3 Robustness and supplementary analysis

B.3.1 Impact of war expenditures on labor market outcomes during WWII

ES-270. Our main analysis looks at changes between 1940 and 1950, but it is also instructive to look at how war expenditures impacted employment during the war. The first way we can analyze the impact during the war is to look at the impact on employment in war industries using the ES-270 reports. For more discussion on the data, please see Appendix Section B.1.3. We use our standard difference-in-differences strategy, and our outcome is the share employed in defense industries. There are several issues with this outcome variable. First, not all establishments are included in the ES-270 reports. Second, we do not have a concurrent estimate of the employed population. Finally, the ES-270 data is split by race or gender but not by race and gender.

Despite these issues, it is still useful to look at the impact of war expenditures on outcomes during the war. First, our hypothesized mechanism requires employment changes during the war so if we do not see concurrent effects then we might question our results. Secondly, it is useful to compare the impact on whites and Blacks during the war. If there are effects on both during the war but only on Blacks after the war, then it strengthens the hypthesis that it is due to changes in discrimination rather than experience gained during the war.

Appendix Table B.5 has the results. Higher war expenditure is strongly associated with higher defense industry employment for Blacks and whites. Therefore, it seems war expenditures did affect both Blacks and whites during the war.

B.3.2 Short-term labor market outcomes

Geographic unit of analysis. Another concern is that our results might be dependent on the geographic unit of analysis. We repeat our main Table (Table 2.2) but for states and commuting zones. The results for states are presented in Appendix Table B.8. In both cases our findings are similar to our main results at the metro level.

Occupational segregation. Another interpretation consistent with our results on occupational upgrading is that white men changed occupations within the skilled occupation group and Black men moved into those vacated occupations. In this scenario there is no decrease in occupational segregation. Therefore we want to check to see if the occupational distribution of Black men and white men became more similar in areas with higher war expenditures at a more granular level. We use two measures for occupational segregation. First, we use a Duncan index to measure deviations from evenness. Second, we use an adjusted Duncan index that is deviations from the expected evenness after accounting for randomness. We use our standard difference-in-differences approach:

$$Y_{rt} = \beta_1 WarExp_r \times Post_t + \beta_2 Draft_r \times Post_t + Post_t + \gamma_r + X_{irt}\rho + \varepsilon_{irt} \tag{B.1}$$

The results are presented in Appendix Table B.6. Higher war expenditures are associated with lower occupational segregation for both measures. Therefore, it does seem as though the occupational composition for white and Black men became more similar in places with higher expenditures.

Excluding likely migrants. Our results could potentially be explained by selective migration. Black men with better skills and/or education could have migrated to metropolitan areas with higher war expenditures. They then stayed in these metropolitan areas after the war, which could explain higher wages and occupational upgrading. In order to test this theory, we re-run our main results but exclude potential interstate migrants in 1950. We define potential interstate migrants as anyone who was born in a different state than their state of residence and does not have a child eight years or older born in their current state of residence. If they have a child who was eight years or older and born in the same state, then it is likely they did not move to their current state after WWII started. We validate this approach using the 1940 Census, which asked for the place of residence five years prior and find it is highly accurate in identifying non-migrants.[5]

We use our standard difference-in-differences approach, except at the individual level with additional controls for age (cubic polynomial), marital staus, and region of birth:

[5]The 1950 Census asks only for the place of residence one year prior. Validation results available upon request.

$$Y_{irt} = \beta_1 WarExp_r \times Post_t + \beta_2 Draft_r \times Post_t + Post_t + \gamma_r + X_{irt}\rho + \varepsilon_{irt} \tag{B.2}$$

The results are presented in Table 2.3. The results excluding potential migrants are very similar to our main results. Therefore, it does not seem as though our results can be solely explained by selective migration.

Impact on younger cohorts. One potential explanation for the persistence of our results is that Black men gained valuable work experience during World War II, leading to persistent productivity improvements. If this explains the persistence, then workers who move to metropolitan areas with higher war expenditures or future generations would not benefit from the accumulated experience. A way to test this explanation is to see occupational gains for cohorts who were too young to have gained significant experience during the war.

Men who are ages 18-24 in 1950 would have been 18 or younger in 1944[6] and therefore would have not been able to accumulate significant experience or would have done so at the cost of reduced education. We compare 18-24 year olds in 1950 versus 1940 and then repeat the exercise using 18-34 year olds in 1960 versus 1940. We use our standard difference-in-differences approach, except at the individual level with additional controls for age, marital staus, and region of birth:

$$Y_{irt} = \beta_1 WarExp_r \times Post_t + \beta_2 Draft_r \times Post_t + Post_t + \gamma_r + X_{irt}\rho + \varepsilon_{irt} \tag{B.3}$$

The results are presented in Table 2.4. The coefficients for the full sample and the restricted age samples are very similar. Therefore, the results can be explained purely by the experience gained during the war. One note of caution when interpreting the results is that younger cohorts could have benefitted from increased education; however, results are similar if

[6] War expenditures were ramping down in 1945.

education is included as a control.

Who upgraded? A natural question is who upgraded? Our robustness Figure 2.6 provides some preliminary evidence that upgrading occurred across demographic groups. At an aggregate level, we can examine how occupational upgrading varied with age or educational status. Appendix Figure B.6 shows the share of men who are in skilled occupations by age for 1930, 1940, and 1950. The shares are roughly constant between 1930 and 1940, but then there is a major shift between 1940 and 1950. The most interesting finding is that the upgrading occurred in all age groups. This suggests that changes in discrimination, rather than compositional changes, might be important. Similar results can be seen when looking at upgrading by educational groups. The upgrading occurred for all educational groups, with the smallest changes for the highest education group. Again, large changes in occupational skill level even for the lowest levels of education (0-5 years) is most consistent with declines in discrimination rather than compositional changes.

A related question is what occupations within the skilled category did Black men enter? Table B.9 gives the relationship between war expenditures and changes in employment shares for 11 aggregate occupation categories. Black men left domestic service, service worker, farm labor, and common laborer occupations and primarily entered operative and craftsman occupations (semi-skilled / skilled blue-collar). These occupations are both very common in manufacturing, which was the key defense industry.

B.3.3 Instrumental variable analysis

One major potential concern is that war expenditures be endogenous with respect to the labor market outcomes for Black men. While there does not appear to be significant pre-existing trends (see Figure 2.4), there might be other potential issues. For example, there could be reverse causality; areas where many Black men upgraded might have had the capacity to receive more war contracts. Therefore, we check if the results are similar when using a Bartik instrument. The basic idea is to predict war expenditures using the baseline industry employment by location interacted with the national (leave-out) expenditures by industry.

We use firm-level data on the total value of war contracts from Li and Koustas (2019). We then allocate these contracts to 1950 Census industry codes using supplemental data from Bianchi and Giorcelli (2020). 1940 Census industry employment totals are used to convert the expenditures for industry i into expenditures per worker ($WarExpPerWorker_i$). Next, we create the baseline number of workers in each industry for each location ($Workers_{ir}$). The predicted shock for each region is:

$$IV_r = \frac{1}{Pop_r} \sum_i WarExpPerWorker_i * Workers_{ir}$$

We then predict war expenditures per capita in a first stage using IV_r. In practice, we use the leave-out version of $WarExpPerWorker_i$, i.e., for each region r we construct the measure excluding contracts and workers in region r.

The key identification assumption is that the "shocks," $WarExpPerWorker_i$, are as good as randomly assigned, conditional on covariates. Note that this does not require the exogeneity of exposure shares. The key potential threat to identification is if the industries that are more likely to receive war expenditure shocks were also industries that were more likely to receive some unobservable shock that caused skill upgrading for Blacks.

It is likely that the shocks are not as good as randomly assigned since manufacturing industries were more likely to receive contracts. Therefore, following the advice in Borusyak, Hull and Jaravel (2019), we control for the initial share in manufacturing, the initial share in durable goods manufacturing, and the share of workers in the labor force.

The results are presented in Appendix Table B.7. The OLS with standard controls are presented in column 1 and the OLS with the additional IV controls are presented in column 2. The IV results are in column 3. The first stage is very strong with an F-stat of over 30. The estimates for the effect on the share skilled are very similar to the OLS estimates. The estimates for the effect on wages are much noisier since the IV is less efficient.

The key result is the fact that we cannot reject the exogeneity of the war expenditures for either the share skilled or wages (see the endogeneity test p-values). We can reject exogeneity for the change in population, but in this case the IV estimates are significantly larger. Therefore, we do not believe there are significant endogeneity issues for our main OLS estimates.

B.3.4 Input-output analysis

War contracts represent the value of the final demand for industry output. The production of the final goods requires significant intermediate inputs. For example, the "direct demand" for a B-17 generates significant "indirect demand" for aluminum. We assign war contracts from Li and Koustas (2019) to 1958 SIC industry codes using supplemental data from Bianchi and Giorcelli (2020). The industry codes are assigned based on the pre-war industry of the firm receiving the contract. This gives the "direct demand" by industry.

We use historical benchmark BLS input-output tables to calculate the indirect demand.[7] We use the 1958 table, but results are similar using the 1947 table instead.[8] These tables give direct purchases from each industry i required to produce one dollar of output in industry j. Let A be the input-output table and d be the vector of direct demand. Then the direct demand industries will need to purchase Ad inputs to produce their output. But these input producers need to purchase their own inputs to produce the output, which adds the additional demand $A(Ad)$. This process can be continued iteratively, and it can be shown that the total gross output, g, required from all industries to produce direct demand d is:

$$g = (I - A)^{-1}d$$

[7] See https://www.bea.gov/industry/historical-benchmark-input-output-tables.

[8] We use the 1958 table because the 1947 table requires additional assumptions and imputations to convert to standardized industries.

Where I is an identify matrix. Therefore, the direct demand is d and the indirect demand is given by $(g - f)$. Finally, we convert the direct and indirect demand to value added by multiplying by the value added share for each industry.

The final step is assigning the industry-level shocks to metro areas. We divide the direct and indirect industry demand shocks by the number of workers in the industry in 1940. The shocks are then allocated to metro areas by multiplying the number of workers in each metro in each industry by the industry shocks to get the total shock. It is then converted to a per capita figure by dividing on the population in the metro.

B.3.5 Effect of war expenditures on migration

First, the migration of Black families cannot be understood without discussing the Great Migration. This overview paragraph draws heavily from Collins (2020), an excellent review of economic research on the Great Migration. Prior to WWI, around 90% of Black individuals lived in the South. Over the next six decades, millions migrated out of the South until less than half of Black individuals lived in the South in 1970. This migration took place in two waves. The first started due to labor shortages during WWI[9] and ended with the Great Depression. The second was precipitated by WWII and ended in the 1960s. Collins and Wanamaker (2014) show that migrants had large earning gains. Migrants also had major impacts on the receiving Northern cities. Boustan (2009) shows how migrants impacted the labor market outcomes of Black and white workers in the North. Large influxes of new migrants also reduced intergenerational mobility for Black individuals (Derenoncourt, 2019). Finally, Black migrants also caused "white flight" to the suburbs (Boustan, 2010).

World War II expenditures were an important influence on the decision to migrate. Boustan (2010) and Derenoncourt (2019) instrument for migrant flows to Northern cities by using pre-existing migration networks interacted with "push" shocks in Southern counties. One of these shocks they use is war expenditures per capita. They find that war

[9]Labor shortages were caused due to war demands combined with the sudden halt to European immigrant flows – see Collins (1997).

expenditures do predict migrant outflows, with higher expenditure areas associated with less out-migration. For this project, our concern is how war expenditures worked as a "pull" factor – i.e., were migrants more likely to go to areas with higher war expenditures. War expenditures are strongly associated with population increases for Black men between 1940 and 1950 but not between 1930 and 1940. For white men they are not strongly associated in 1940 to 1950 but there is a negative association in the pre-period that might indicate a positive impact relative to the existing trend. Appendix Figure B.4 provides the effect of war expenditures on migration for a variety of specifications.

B.3.6 Education of the next generation

School attendance 1930-40 and 1940-60. We repeat our analysis of the effect of war expenditures on school enrollment for the periods 1930-40 and 1940-60 instead of 1940-50. We follow our main difference-in-differences approach.

The results for 1930-40 are presented in Appendix Table B.13. We do not see any significant relationship between war expenditures and changes in school enrollment. Therefore, there does not seem to be significant positive pre-trends or a slight negative trend.

The results for 1940-60 are presented in Appendix Table B.12. We see a positive relationship between war expenditures and schooling for Black boys but no significant effect for Black girls. This result is consistent with our 1940-50 results that find stronger impacts on Black boys. The effect size is smaller, which does indicate the effect could fade with time. An alternative explanation could be that attendance of 16-18 year olds increased with time, reducing our ability to measure the treatment effect because there are fewer potential "switchers."

High school graduation rates. One concern with the effect of war expenditures on high school graduation rates (Figure 2.8) is the potential presence of pre-trends. The 5% sample for 1960 might not have sufficient power to rule out pre-trends. One alternative is to conduct the same analysis using the 1940 Census to see if there were pre-existing trends. Appendix Figure B.8 shows these results. For boys, the coefficients on "War exp * Black" seem to be consistently close to zero. For girls, there is little evidence of a positive trend in the years immediately leading up to WWII. There is potentially some trend in the early 1930s at the onset of the Great Depression that is driven by differential changes in the education of white girls. These results are consistent with our estimates for girls being noisier and less likely to be statistically significant than our estimates for boys.

B.4 Quantitative Appendix

B.4.1 Equilibrium

Consumption. First, consider the consumption side. Because trade is costless and preferences are identical across regions, consumption prices are equalized across space.

Consumption of industry i in region r is given by

$$C_{rit}P_{it}^{C} = \mu_i P_t^C C_{rt} \tag{B.4}$$

where

$$P_t^C = \prod_i \left(\frac{P_{it}^C}{\mu_i}\right)^{\mu_i} \tag{B.5}$$

denotes the final good price and where P_{it}^C denotes the consumption price of industry i. Consumption of industry i from origin j in destination r is given by

$$C_{jrit} = \mu_{jit}\left(\frac{P_{jit}^Y}{P_{it}^C}\right)^{-\rho} C_{rit} \tag{B.6}$$

where P_{jit}^Y is the production price in region j of industry i and where the consumption price of industry i in all regions is

$$P_{it}^{C} = \left(\sum_{j} \mu_{jit} \left(P_{jit}^{Y} \right)^{1-\rho} \right)^{\frac{1}{1-\rho}} \tag{B.7}$$

Production. Next, consider the production side. Industry i profit maximization implies that in region r the output of industry-occupation io pair is given by

$$Y_{riot} = \mu_{riot} \left(\frac{P_{riot}^{Y}}{P_{rit}^{Y}} \right)^{-\eta} Y_{rit} \tag{B.8}$$

where Y_{rit} is the region r output of industry i, where the output price in region r of industry i is

$$P_{rit}^{Y} = \left(\sum_{o} \mu_{riot} \left(P_{riot}^{Y} \right)^{1-\eta} \right)^{\frac{1}{1-\eta}} \tag{B.9}$$

and where P_{riot}^{Y} denotes the region r output price of industry-occupation pair io.

The share of workers in group g and region r who choose to work in industry-occupation io, denoted by $\pi_{riogt}^{L} \equiv N_{riogt}/N_{rgt}$ (where N_{rgt} denotes the measure of group g workers who choose to live in region r and N_{riogt} the measure who additionally choose to work in io), is given by

$$\pi_{riogt}^{L} = \left(A_{riogt} T_{riogt} P_{riot}^{Y} \right)^{\theta} \Big/ \Phi_{rgt} \tag{B.10}$$

and where

$$\Phi_{rgt} \equiv \sum_{io} \left(A_{riogt} T_{riogt} P_{riot}^{Y} \right)^{\theta} \tag{B.11}$$

The total efficiency units supplied by group g in industry-occupation io in region r is

$$L_{riogt} = \gamma T_{riogt} \left(\pi_{riogt}^{L} \right)^{\frac{\theta-1}{\theta}} \pi_{rgt}^{N} N_{gt} \tag{B.12}$$

In equation (B.12), $\gamma \equiv \Gamma\left(1 - \frac{1}{\theta}\right)$ where Γ is the gamma function, and $\pi_{rgt}^{N} \equiv N_{rgt}/N_{gt}$ is the share of workers in group g who choose to live in region r and is given by

$$\pi^N_{rgt} = \left(U_{rgt}\Phi^{\frac{1}{\theta}}_{rgt}\right)^{\nu} \Big/ \left[\sum_{r'} \left(U_{r'gt}\Phi^{\frac{1}{\theta}}_{r'gt}\right)^{\nu}\right] \tag{B.13}$$

Finally, the average wage of group g in region r and job io is given by

$$Wage_{riogt} = \gamma\Phi^{\frac{1}{\theta}}_{rgt}\Big/A_{riogt} \tag{B.14}$$

Market clearing. Region r's output of industry i must equal the sum of consumption across all regions for each ri pair

$$Y_{rit} = \sum_j C_{rjit} \tag{B.15}$$

Locally, markets must clear in each rio triplet

$$Y_{riot} = \sum_g L_{riogt} \tag{B.16}$$

Market clearing and balanced trade link production and consumption

$$P^C_t C_{rt} = \sum_{gio} Wage_{riogt}\pi^L_{riogt}N_{rgt} \tag{B.17}$$

Equilibrium. An equilibrium is a vector of consumption prices $\{P^C_t, P^C_{it}\}$, production prices $\{P^Y_{rit}, P^Y_{riot}\}$, aggregator $\{\Phi_{rgt}\}$ and wages $\{Wage_{riogt}\}$, quantities produced $\{Y_{rt}, Y_{rit}, Y_{riot}\}$, consumption levels $\{C_{rt}, C_{rit}, C_{jrit}\}$, and labor allocations $\{\pi^N_{rgt}, \pi^L_{riogt}, L_{riogt}\}$ for all region pairs jr, industries i, occupations o, and worker groups g that satisfy (B.4)-(B.17).

B.4.2 Decomposition

In this section, we provide the system of equations with which to solve for the implications of shocks and show how to measure these shocks. We define $\widehat{x} = x_{t+1}/x_t$ for any variable x;

it is the relative value of a variable in a "new equilibrium" $(t+1)$ relative to in the initial equilibrium (t). The point of writing the system in changes is that it dramatically reduces the set of parameters we need to estimate to conduct our decomposition and counterfactuals.

In practice, the shocks that we feed into the system are changes across time in productivity, T_{riogt}, in amenities, U_{rgt}, and in national populations, N_{gt}. Here, however, we allow for a more general set of shocks, additionally including shocks to demand across origin and industry pairs, μ_{jit}, changes in demand across occupations within industries, μ_{riot}, and changes in amenities for working in industry-occupation io within each region r, A_{riogt}. We show here that for given values of $\rho \neq 1$ and $\eta \neq 1$, it is without loss of generality to normalize μ_{jit} and μ_{riot} to be fixed over time, since any changes in these parameters can be absorbed by changes in T_{riogt} without affecting any results.

B.4.3 System in changes

We express our system of equations in changes as follows:

$$\widehat{P}_{it}^{C}\widehat{C}_{rit} = \widehat{P}_{t}^{C}\widehat{C}_{rt} \tag{B.18}$$

$$\widehat{P}_{t}^{C} = \prod_{i}\left(\widehat{P}_{it}^{C}\right)^{\mu_i} \tag{B.19}$$

$$\widehat{C}_{jrit} = \widehat{\mu}_{jit}\left(\frac{\widehat{P}_{jit}^{Y}}{\widehat{P}_{it}^{C}}\right)^{-\rho}\widehat{C}_{rit} \tag{B.20}$$

$$\widehat{Y}_{riot} = \widehat{\mu}_{riot}\left(\frac{\widehat{P}_{riot}^{Y}}{\widehat{P}_{rit}^{Y}}\right)^{-\eta}\widehat{Y}_{rit} \tag{B.21}$$

$$\widehat{P}_{it}^{C} = \left(\sum_{j} S_{jit}^{C}\widehat{\mu}_{jit}\left(\widehat{P}_{jit}^{Y}\right)^{1-\rho}\right)^{\frac{1}{1-\rho}} \tag{B.22}$$

where $S^C_{jit} \equiv \frac{\mu_{jit}(P^Y_{jit})^{1-\rho}}{\sum_r \mu_{rit}(P^Y_{rit})^{1-\rho}}$ denotes the share of each region's expenditure on industry i that is produced in region j.

$$\widehat{P}^Y_{rit} = \left(\sum_o S^Y_{riot} \widehat{\mu}_{riot} \left(\widehat{P}^Y_{riot} \right)^{1-\eta} \right)^{\frac{1}{1-\eta}} \tag{B.23}$$

where $S^Y_{riot} \equiv \frac{\mu_{riot}\left(P^Y_{riot}\right)^{1-\eta}}{\sum_{o'} \mu_{rio't}\left(P^Y_{rio't}\right)^{1-\eta}}$ denotes region r's share of expenditure within industry i on occupation o.

$$\widehat{\pi}^N_{rgt} = \frac{\left(\widehat{U}_{rgt} \widehat{\Phi}^{\frac{1}{\theta}}_{rgt} \right)^{\nu}}{\sum_{r'} \pi^N_{r'gt} \left(\widehat{U}_{r'gt} \widehat{\Phi}^{\frac{1}{\theta}}_{r'gt} \right)^{\nu}} \tag{B.24}$$

$$\widehat{\Phi}_{rgt} = \sum_{io} \pi^L_{riogt} \left(\widehat{A}_{riogt} \widehat{T}_{riogt} \widehat{P}^Y_{riot} \right)^{\theta} \tag{B.25}$$

$$\widehat{\pi}^L_{riogt} = \frac{(\widehat{A}_{riogt} \widehat{T}_{riogt} \widehat{P}^Y_{riot})^{\theta}}{\widehat{\Phi}_{rgt}} \tag{B.26}$$

$$\widehat{L}_{riogt} = \widehat{T}_{riogt} \left(\widehat{\pi}^L_{riogt} \right)^{\frac{\theta-1}{\theta}} \widehat{\pi}^N_{rgt} \widehat{N}_{gt} \tag{B.27}$$

$$\widehat{Wage}_{riogt} = \frac{\widehat{\Phi}^{\frac{1}{\theta}}_{rgt}}{\widehat{A}_{riogt}} \tag{B.28}$$

$$\widehat{Y}_{rit} = \sum_j s_{jt} \widehat{C}_{rjit} \tag{B.29}$$

where $s_{jt} \equiv \frac{C_{rjit}}{\sum_{j'} C_{rj'it}}$ is the share of region r's industry i output shipped to j.

$$\widehat{Y}_{riot} = \sum_g s_{riogt} \widehat{L}_{riogt} \tag{B.30}$$

where $s_{riogt} \equiv \frac{L_{riogt}}{\sum_{g'} L_{riog't}}$ is the share of output in r of io produced by group g.

$$\widehat{P}^C_t \widehat{C}_{rt} = \sum_{iog} v_{riogt} \widehat{Wage}_{riogt} \widehat{\pi}^L_{riogt} \widehat{\pi}^N_{rgt} \widehat{N}_{gt} \tag{B.31}$$

where $v_{riogt} \equiv \frac{Wage_{riogt}\pi^{L}_{riogt}N_{rgt}}{\sum_{i'o'g'} Wage_{ri'o'g't}\pi^{L}_{ri'o'g't}N_{rg't}}$ is the share of total labor income in r that accrues to g within io.

The system has 14 equations (B.18)-(B.31) and unknowns:

$$\left\{\widehat{P}^{Y}_{riot}, \widehat{P}^{Y}_{rit}, \widehat{\Phi}_{rgt}, \widehat{Wage}_{riogt}, \widehat{\pi}^{L}_{riogt}, \widehat{\pi}^{N}_{rgt}, \widehat{L}_{riogt}, \widehat{P}^{C}_{it}, \widehat{P}^{C}_{t}, \widehat{C}_{rt}, \widehat{C}_{rit}, \widehat{C}_{jrit}, \widehat{Y}_{riot}, \widehat{Y}_{rit}\right\}$$

Given shocks $\{\widehat{T}_{riogt}, \widehat{A}_{riogt}, \widehat{\mu}_{jit}, \widehat{\mu}_{riot}, \widehat{U}_{rgt}, \widehat{N}_{gt}\}$, elasticities $\{\rho, \theta, \eta, \nu\}$, and initial equilibrium shares $\{\mu_i, S^{C}_{jit}, S^{Y}_{riot}, \pi^{L}_{riogt}, \pi^{N}_{rgt}, s_{jt}, s_{riogt}, v_{riogt}\}$, we can solve for all changes using the previous system and a normalization. This algorithm requires that we have values for the elasticities, the shocks, and the initial equilibrium shares. We next describe how we choose these.

B.4.3.1 Calibrating the model to 1940 data

In constructing each share, we are using data only from the regions that we are considering. As an example, if we require total labor income earned in industry i across all regions, then we sum total labor income earned in industry i across all regions in our sample (rather than across all regions in the U.S.). All data used in constructing shares is from the 1940 census.

μ_i: μ_i is the share of expenditure on industry i, which we assume is constant across time and regions. The numerator of μ_i is the sum of labor income (since labor is the only factor of production) across regions in industry i and the denominator is the sum of labor income across regions and industries.

S^{C}_{jit}: S^{C}_{jit} is the share of national industry i expenditure that is produced in region j. Since labor is the only factor of production and there is no trade with the outside world, the numerator is labor income earned in industry i in region j and the denominator is labor income earned in industry i summed across all regions.

S^{Y}_{riot}: S^{Y}_{riot} is the share of region r's labor income in industry i that is earned in occupation o. To measure S^{Y}_{riot}, the numerator is labor income earned within io in region r and the denominator is labor income earned within i in region r.

π^L_{riogt}: π^L_{riogt} is the share of employed men (part- and full-time) within region r for group g that is worked within io. The numerator of π^L_{riogt} is employment of group g in region r in io and the denominator is the sum of employment of group g in region r across all io pairs.

π^N_{rgt}: π^N_{rgt} is the share of the employed male population within group g that lives in region r. The numerator is the employed male population of g in r and the denominator is the employed male population of g across all regions.

s_{jt}: s_{jt} is the share of total labor income earned in region j. The numerator is labor income in j and the denominator is labor income across all regions.

s_{riogt}: s_{riogt} is the share of labor income in region r and industry-occupation io that is paid to group g. The numerator is labor income in region r and industry-occupation io that is paid to group g and the denominator is the sum of labor income in region r and industry-occupation io across all g.

v_{riogt}: v_{riogt} is the share of total labor income in r (across all g and across all io) that accrues to g within io. The numerator is the payment to g in io within r and the dominator is total labor payments in r across all g and all io.

B.4.3.2 Calibrating η

To choose the value of η, we calibrate the model to 1940 data, matching the initial shares described in B.4.3.1. Then we take the following steps.

1. We pick a value of η.

2. We measure shocks, as described in B.4.3.3, where the value of two of these six shocks, β^T_1 and β^T_2, depends on the choice of η.

3. We feed into the model the government spending shocks (associated with all β^T and β^U parameters) and solve for the new equilibrium of the model.

4. We estimate regressions of the form in (2.1) in Section 2.3.3 using actual data and, separately, using model-generated data, where the dependent variables, Y_{rt} are the share of employment in skilled occupations and the ln average wage of Black and white workers in each metro area and year (for 1940 and 1950).

5. For each of the four coefficients of interest in the data, $\beta^{data}_{1,race,Y}$, and in the model, $\beta^{model}_{1,race,Y}$, where Y indicates the dependent variable and $race$ indicates the sample, we construct the sum of squared differences

$$\mathcal{L}(\eta) = \sum_{Y,race} \omega_{race}(\beta^{data}_{1,race,Y} - \beta^{model}_{1,race,Y})^2$$

   where ω_{race} is a weight that we set to 1 if $race =$ white and we set to 2 if $race =$ Black, given our focus on explaining Black labor market outcomes.

Finally, we iterate over values of η to minimize $\mathcal{L}(\eta)$. This procedure yields $\eta = 11$. Panels A and B of Appendix Table B.16 display the resulting values of $\beta^{data}_{1,race,Y}$ and $\beta^{model}_{1,race,Y}$, respectively, that result from our baseline calibration.

B.4.3.3 Measuring shocks

We focus on shocks between the years 1940 and 1950. We measure changes in national (across the regions in our analysis) employed male populations, $\widehat{N}_{gt}$, directly from the data. We express (without loss of generality) the structural productivity, net of the discriminatory wedge, of group g in region r and industry-occupation io at time t as

$$\ln T_{riogt} = \gamma^T_{riog} + \gamma^T_{iogt} + G_r \mathbb{I}_t \mathbb{I}_i \left[\beta^T_1 + \beta^T_2 \mathbb{I}_o + \beta^T_3 \mathbb{I}_g + \beta^T_4 \mathbb{I}_o \mathbb{I}_g\right] + \iota^T_{riogt} \tag{B.32}$$

and the amenity value for group g of living in region r at time t as

$$\ln U_{rgt} = \gamma^U_{rg} + G_r \mathbb{I}_t [\beta^U_1 + \beta^U_2 \mathbb{I}_g] + \iota^U_{rgt} \tag{B.33}$$

In addition to the shocks incorporated in the body of the paper, we include three additional shocks:

$$\ln \mu_{rit} = \gamma^i_{ri} + \gamma^i_{it} + G_r \mathbb{I}_t \mathbb{I}_i \beta^i_1 + \iota^i_{rit} \tag{B.34}$$

$$\ln \mu_{riot} = \gamma^o_{rio} + \gamma^o_{iot} + G_r \mathbb{I}_t \mathbb{I}_i \left[\beta^o_1 + \beta^o_2 \mathbb{I}_o\right] + \iota^o_{riot} \tag{B.35}$$

$$\ln A_{riogt} = \gamma^A_{riog} + \gamma^A_{iogt} + G_r \mathbb{I}_t \mathbb{I}_i \left[\beta^A_1 + \beta^A_2 \mathbb{I}_o + \beta^A_3 \mathbb{I}_g + \beta^A_4 \mathbb{I}_o \mathbb{I}_g\right] + \iota^A_{riogt} \tag{B.36}$$

Here, we will show that it is without loss of generality to assume that β^i_1, β^o_1, and β^o_2 are all set to zero for given values of $\rho \neq 1$ and $\eta \neq 1$. We will also show that the data requires that the β^A parameters also be set to zero.

Industry-occupation amenity shocks. From equation (B.14) we have

$$\ln Wage_{riogt} = \ln \gamma - \ln A_{riogt} + (1/\theta) \ln \Phi_{rgt}$$

The previous expression and (B.36) yield

$$\ln Wage_{riogt} = \gamma_{rgt} - \gamma^A_{riog} - \gamma^A_{iogt} - G_r \mathbb{I}_t \mathbb{I}_i \left[\beta^A_1 + \beta^A_2 \mathbb{I}_o + \beta^A_3 \mathbb{I}_g + \beta^A_4 \mathbb{I}_o \mathbb{I}_g\right] - \iota^A_{riogt} \tag{B.37}$$

where $\gamma_{rgt} \equiv \ln \gamma + (1/\theta) \ln \Phi_{rgt}$. Taking changes across time in equation (B.37), *changes* in average wages within *riog* cells conditional on *rg* and *iog* fixed effects identify the impact of government spending on *changes* in industry-occupation amenities. Intuitively, in the absence of any changes in these amenities, we would find zero values of all β^A parameters in model-generated data given our assumption of Fréchet distributed idiosyncratic productivites. Estimating (B.37), we find that all four β^A coefficients are economically small (with absolute values ranging from 0.000 to 0.008) and statistically insignificant (the highest t-statistic is 0.5); see column 1 of Table B.15. In summary, in the absence of any such changes in amenities, the assumption of Fréchet distributed idiosyncratic productivites for average wage changes matches our data well. Given this result, we impose that $A_{riogt} = A_{riog}$ throughout the remainder of the analysis.

Anti-discriminatory shocks. From (B.10), we have

$$\ln \pi^L_{riogt} = \theta \ln T_{riogt} + \theta \ln P^Y_{riot} - \ln\left(\sum_{i'o'} (T_{ri'o'gt} P^Y_{ri'o't})^\theta\right)$$

The previous expression and (B.32) yield

$$\begin{aligned}\ln \pi^L_{riogt} =& \theta\gamma^T_{riog} + \theta\gamma^T_{iogt} + G_r \mathbb{I}_t \mathbb{I}_i \left[\theta\beta^T_1 + \theta\beta^T_2 \mathbb{I}_o + \theta\beta^T_3 \mathbb{I}_g + \theta\beta^T_4 \mathbb{I}_o \mathbb{I}_g\right] \\ &+ \theta\iota^T_{riogt} + \theta \ln P^Y_{riot} - \ln\left(\sum_{i'o'} (T_{ri'o'gt} P^Y_{ri'o't})^\theta\right)\end{aligned}$$

which can be re-expressed as (2.9), where $\gamma_{rgt} \equiv -\ln\left(\sum_{i'o'} (T_{ri'o'gt} P^Y_{ri'o't})^\theta\right)$, $\gamma_{riot} \equiv \theta \ln P^Y_{riot} + G_r \mathbb{I}_t \mathbb{I}_i \left[\theta\beta^T_1 + \theta\beta^T_2 \mathbb{I}_o\right]$, $\gamma_{riog} \equiv \theta\gamma^T_{riog}$, $\gamma_{iogt} \equiv \theta\gamma^T_{iogt}$, $\iota_{riogt} \equiv \theta\iota^T_{riogt}$, $\beta_3 \equiv \theta\beta^T_3$, and $\beta_4 \equiv \theta\beta^T_4$. See column 1 of Table 2.5 for estimation results.

Equation (B.13) implies

$$\frac{1}{\nu} \ln \pi^N_{rgt} = \ln U_{rgt} + \frac{1}{\theta}\Phi_{rgt} + \gamma_{gt}$$

where $\gamma_{gt} \equiv -\frac{1}{\nu} \ln\left[\sum_{r'} \left(U_{r'gt} \Phi^{1/\theta}_{r'gt}\right)^\nu\right]$. Equation (B.14) implies

$$\frac{1}{\theta} \ln \Phi_{rgt} = \ln Wage_{riogt} - \ln \gamma + \ln A_{riog}$$

Combining the previous two expressions yields

$$\frac{1}{\nu} \ln \pi^N_{rgt} - \ln Wage_{riogt} = \ln U_{rgt} - \ln \gamma + \ln A_{riog} + \gamma_{gt}$$

The previous expression and (B.33) yield

$$\frac{1}{\nu} \ln \pi^N_{rgt} - \ln Wage_{riogt} = \tilde{\gamma}^U_{riog} + \gamma_{gt} + G_r \mathbb{I}_t [\beta^U_1 + \beta^U_2 \mathbb{I}_g] + \iota^U_{rgt} \tag{B.38}$$

where $\tilde{\gamma}^U_{riog} \equiv \gamma^U_{rg} + \ln A_{riog} - \ln \gamma$. The previous expression simplifies to equation (2.11) where $\gamma_{rt} \equiv \beta_1^U G_r \mathbb{I}_t$. See column 4 of Table 2.5 for estimation results. This concludes the baseline identification of the anti-discriminatory effects of wartime spending.

Compositional shock I: amenity parameter β_1^U. Having estimated (2.11) to identify β_2^U, we subtract $\widehat{\beta}_2^U G_r \mathbb{I}_t \mathbb{I}_g$ from both the left- and right-hand side of (B.38), and estimate

$$\frac{1}{\nu} \ln \pi_{rgt}^N - \ln Wage_{rgt} - \widehat{\beta}_2^U G_r \mathbb{I}_t \mathbb{I}_g = \gamma_{rg} + \gamma_{gt} + \beta_1^U G_r \mathbb{I}_t + \iota_{rgt}^U \tag{B.39}$$

See column 5 of Table 2.5 for the resulting estimate of β_1^U. In robustness, we also estimate β_U^1 and β_U^2 together in a single step by estimating (B.38) directly. Results of this robustness exercise are shown in column 8 of Table 2.5; these results are very similar quantitatively to our baseline results diplayed in columns 4 and 5.

Compositional shock II: productivity parameter β_2. From (B.8), we obtain

$$\ln\left(P_{riot}^Y Y_{riot}\right) = \ln \mu_{riot} + (1-\eta)\ln P_{riot}^Y + \eta \ln P_{rit}^Y + \ln Y_{rit}$$

From (B.10) and (B.14), we have

$$\ln P_{riot}^Y = \ln Wage_{riogt} + \frac{1}{\theta}\ln \pi_{riogt} - \ln T_{riogt} - \ln \gamma$$

Combining the previous two expressions yields

$$y_{riogt} = \frac{\theta}{1-\eta}\left[-\ln \mu_{riot} + (1-\eta)\ln T_{riogt} - \eta \ln P_{rit}^Y - \ln Y_{rit} + (1-\eta)\gamma\right] \tag{B.40}$$

where we have defined

$$y_{riogt} \equiv \frac{-\theta}{1-\eta}\ln\left(P_{riot}^Y Y_{riot}\right) + \theta \ln Wage_{riogt} + \ln \pi_{riogt}^L$$

Combining equation (B.40) with (B.32) and (B.35) yields

$$y_{riogt} = \gamma_{riog} + \gamma_{iogt} + \gamma_{rit} + \beta_2 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_o + \beta_3 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g + \beta_4 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g \mathbb{I}_o + \iota_{riogt} \tag{B.41}$$

where $\gamma_{rit} \equiv -\frac{\theta}{1-\eta}\eta \ln P^Y_{rit} - \frac{\theta}{1-\eta}\ln Y_{rit} + \theta\gamma - \frac{\theta}{1-\eta}\beta^o_1 + \theta\beta^T_1$, $\iota_{riogt} \equiv -\frac{\theta}{1-\eta}\iota^o_{riot} + \theta\iota^T_{riogt}$, $\gamma_{riog} \equiv -\frac{\theta}{1-\eta}\gamma^o_{rio} + \theta\gamma^T_{riog}$, $\gamma_{iogt} \equiv -\frac{\theta}{1-\eta}\gamma^o_{iot} + \theta\gamma^T_{iogt}$, $\beta_2 \equiv -\frac{\theta}{1-\eta}\beta^o_2 + \theta\beta^T_2$, $\beta_3 \equiv \theta\beta^T_3$, and $\beta_4 \equiv \theta\beta^T_4$. We subtract our estimates of $\widehat{\beta}_3 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g$ and $\widehat{\beta}_4 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g \mathbb{I}_o$ from the left- and right-hand sides of the previous expression to obtain

$$\tilde{y}_{riogt} = \gamma_{riog} + \gamma_{iogt} + \gamma_{rit} + \beta_2 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_o + \iota_{riogt} \tag{B.42}$$

where $\tilde{y}_{riogt} \equiv y_{riogt} - \widehat{\beta}_3 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g - \widehat{\beta}_4 G_r \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g \mathbb{I}_o$. We estimate β_2 using (B.42), and report results in column 2 of Table 2.5. In robustness, we estimate β_2, β_3, and β_4 using (B.41), and report results in column 6 of Table 2.5. Results in column 6 are very similar to those reported in columns 1 and 2.

Compositional shock III: productivity parameter β_1. Equations (B.4), (B.6), and (B.15) yield

$$P^Y_{jit} Y_{jit} = \mu_{jit} \left(P^Y_{jit}\right)^{1-\rho} \times \left(P^C_{it}\right)^{\rho-1} \mu_i \sum_r P^C_t C_{rt}$$

The previous expression and (B.9) yield

$$P^Y_{jit} Y_{jit} = \mu_{jit} \left(\sum_o \mu_{jiot} \left(P^Y_{jiot}\right)^{1-\eta}\right)^{\frac{1-\rho}{1-\eta}} \times \left(P^C_{it}\right)^{\rho-1} \mu_i \sum_r P^C_t C_{rt}$$

The previous expression and the definition of S^Y_{jiot}, which implies

$$\sum_o \mu_{jiot} \left(P^Y_{jiot}\right)^{1-\eta} = \frac{\mu_{jio't} \left(P^Y_{jio't}\right)^{1-\eta}}{S^Y_{jio't}} \quad \text{for any } o'$$

yield

$$P^Y_{jit} Y_{jit} \left(S^Y_{jiot}\right)^{\frac{1-\rho}{1-\eta}} = \mu_{jit} \mu_{jiot}^{\frac{1-\rho}{1-\eta}} \left(P^Y_{jiot}\right)^{1-\rho} \times \left(P^C_{it}\right)^{\rho-1} \mu_i \sum_r P^C_t C_{rt}$$

Combining the previous expression with (B.10) and (B.14), which imply

$$\ln P^Y_{riot} = \ln Wage_{riogt} + \frac{1}{\theta}\ln \pi_{riogt} - \ln T_{riogt} - \ln \gamma$$

yields

$$\begin{aligned}\ln(P^Y_{jit}Y_{jit}) + \frac{1-\rho}{1-\eta}\ln S^Y_{jiot} = \ln \mu_{jit} + \frac{1-\rho}{1-\eta}\ln \mu_{jiot} + \ln\left[\left(P^C_{it}\right)^{\rho-1}\mu_i \sum_r P^C_t C_{rt}\right] \\ + (1-\rho)\left[\ln Wage_{jiogt} + \frac{1}{\theta}\ln \pi_{jiogt} - \ln T_{jiogt} - \ln\gamma\right]\end{aligned}$$

which is equivalent to

$$B_{jiogt} = -\frac{\theta}{1-\rho}\ln \mu_{jit} - \frac{\theta}{1-\eta}\ln \mu_{jiot} + \gamma_{it} + \theta \ln T_{jiogt}$$

where $\gamma_{it} \equiv \theta \ln \gamma - \frac{\theta}{1-\rho}\ln\left[\left(P^C_{it}\right)^{\rho-1}\mu_i \sum_r P^C_t C_{rt}\right]$ and where

$$B_{jiogt} \equiv -\frac{\theta}{1-\rho}\ln(P^Y_{jit}Y_{jit}) - \frac{\theta}{1-\eta}\ln S^Y_{jiot} + \theta \ln Wage_{jiogt} + \ln \pi_{jiogt} \tag{B.43}$$

We substitute out $\ln \mu_{jit}$, $\ln \mu_{jiot}$ and $\ln T_{jiogt}$ using (B.32), (B.34), and (B.35) to obtain

$$B_{jiogt} = \gamma_{jiog} + \gamma_{iogt} + \beta_1 G_j \mathbb{I}_t \mathbb{I}_i + \beta_2 G_j \mathbb{I}_t \mathbb{I}_i \mathbb{I}_o + \beta_3 G_j \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g + \beta_4 G_j \mathbb{I}_t \mathbb{I}_i \mathbb{I}_o \mathbb{I}_g + \iota_{jiogt} \tag{B.44}$$

where $\gamma_{iogt} \equiv \gamma_{it} - \frac{\theta}{1-\rho}\gamma^i_{it} - \frac{\theta}{1-\eta}\gamma^o_{iot} + \theta\gamma^T_{iogt}$, $\gamma_{jiog} \equiv -\frac{\theta}{1-\rho}\gamma^i_{ji} - \frac{\theta}{1-\eta}\gamma^o_{jio} + \theta\gamma^T_{jiog}$, $\iota_{jiogt} \equiv -\frac{\theta}{1-\rho}\iota^i_{jit} - \frac{\theta}{1-\eta}\iota^o_{jiot} + \theta\iota^T_{jiogt}$, $\beta_1 \equiv -\frac{\theta}{1-\rho}\beta^i_1 - \frac{\theta}{1-\eta}\beta^o_1 + \theta\beta^T_1$, $\beta_2 \equiv -\frac{\theta}{1-\eta}\beta^o_2 + \theta\beta^T_2$, $\beta_3 \equiv \theta\beta^T_3$, and $\beta_4 \equiv \theta\beta^T_4$.

Subtracting from both sides of the previous expression the terms associated with β_2, β_3, and β_4 which we have previously estimated, we estimate

$$\tilde{B}_{jiogt} = \gamma_{jiog} + \gamma_{iogt} + \beta_1 G_j \mathbb{I}_t \mathbb{I}_i + \iota_{jiogt} \tag{B.45}$$

where

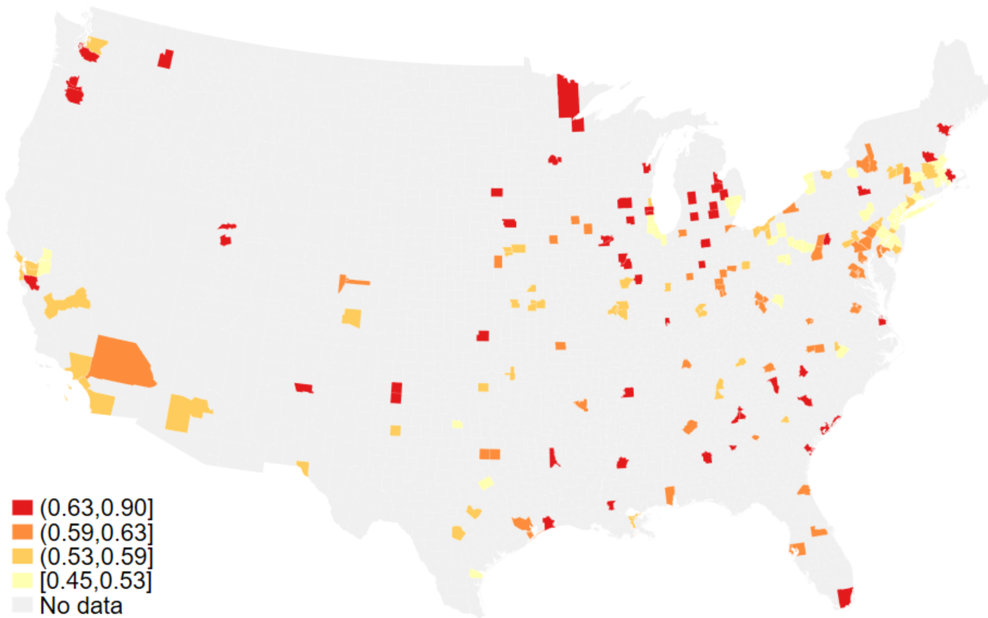
$$\tilde{B}_{jiogt} \equiv B_{jiogt} - \hat{\beta}_2 G_j \mathbb{I}_t \mathbb{I}_i \mathbb{I}_o - \hat{\beta}_3 G_j \mathbb{I}_t \mathbb{I}_i \mathbb{I}_g - \hat{\beta}_4 G_j \mathbb{I}_t \mathbb{I}_i \mathbb{I}_o \mathbb{I}_g$$

We report results of estimating regression (B.45)in column 3 of Table 2.5. In robustness, we estimate β_1, β_2, β_3, and β_4 all together using (B.44) and report results in column 7 of Table 2.5. Results in column 7 are very similar to those reported in columns 1, 2, and 3 as well as to those reported in column 6.

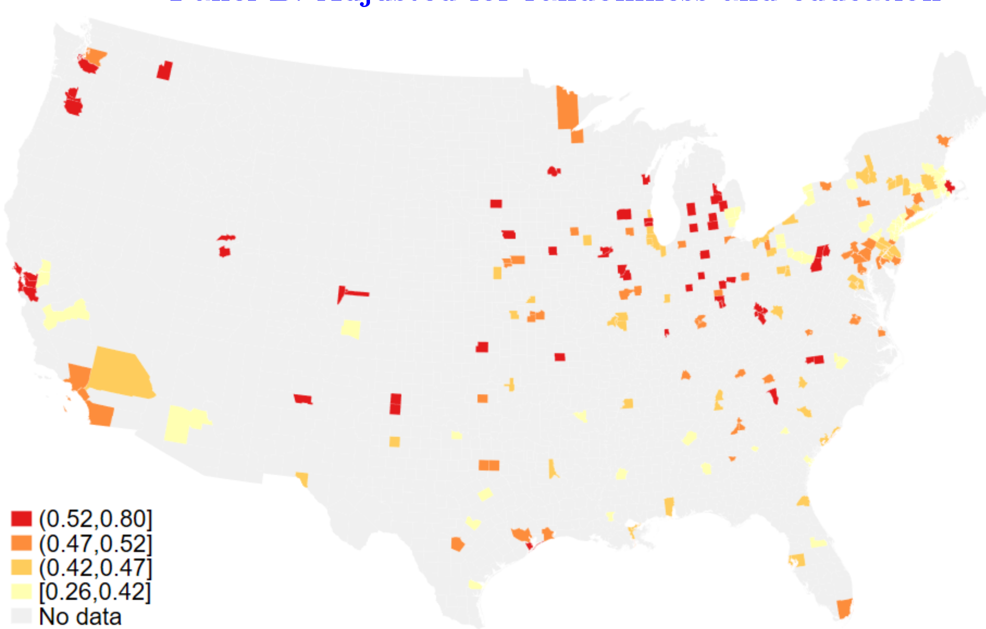
B.5 Appendix tables and figures

Figure B.1: Black vs. white men occupational dissimilarity index by metro area (1940)

Panel A: Unadjusted

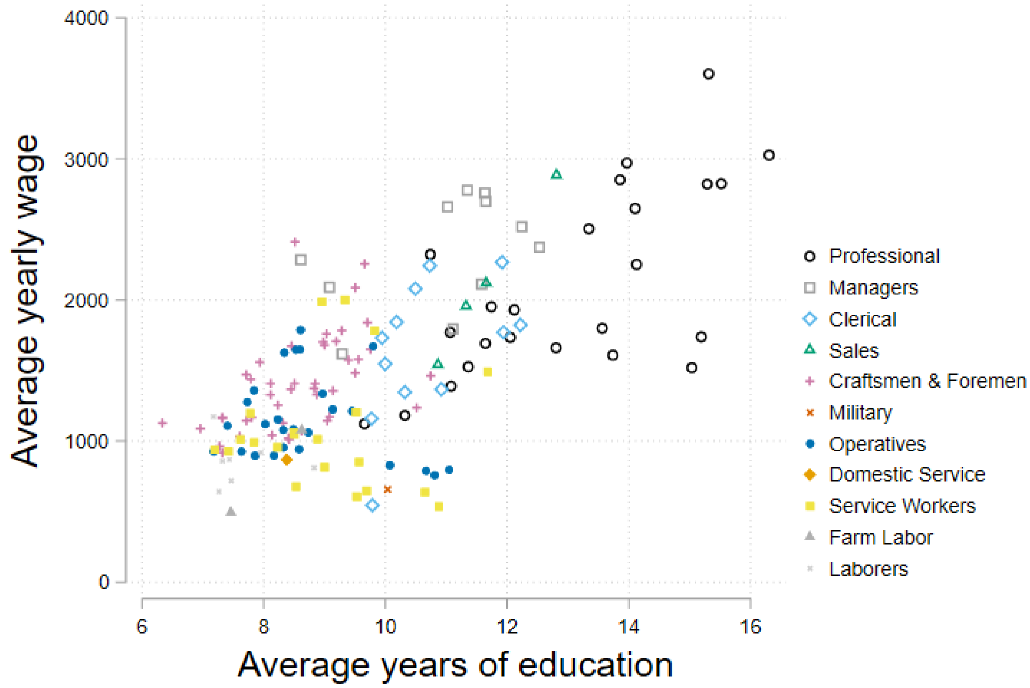


Panel B: Adjusted for randomness and education



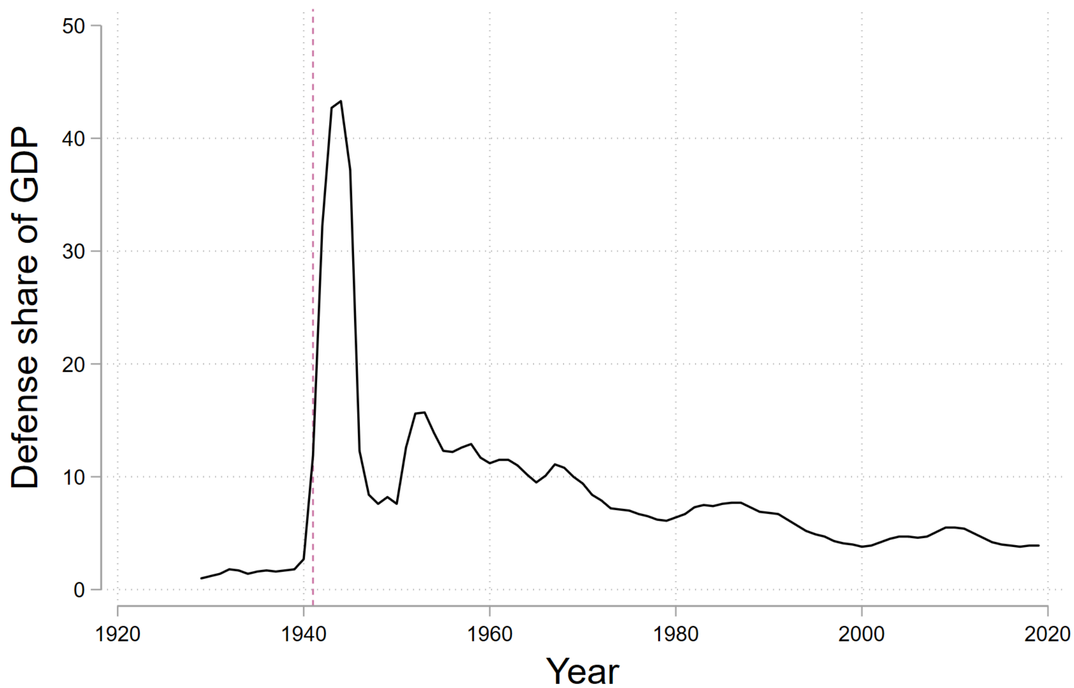
Note: For 146 metro areas, as defined by the Census Bureau in 1940 and 1950. Panel A presents unadjusted occupational dissimilarity indices, while Panel B adjusts for education (5 groups) and randomness. For more details on the creation of these measures, please see Appendix Section B.2.2.

Figure B.2: Average yearly wage and years of education by occupation for white men (1940)



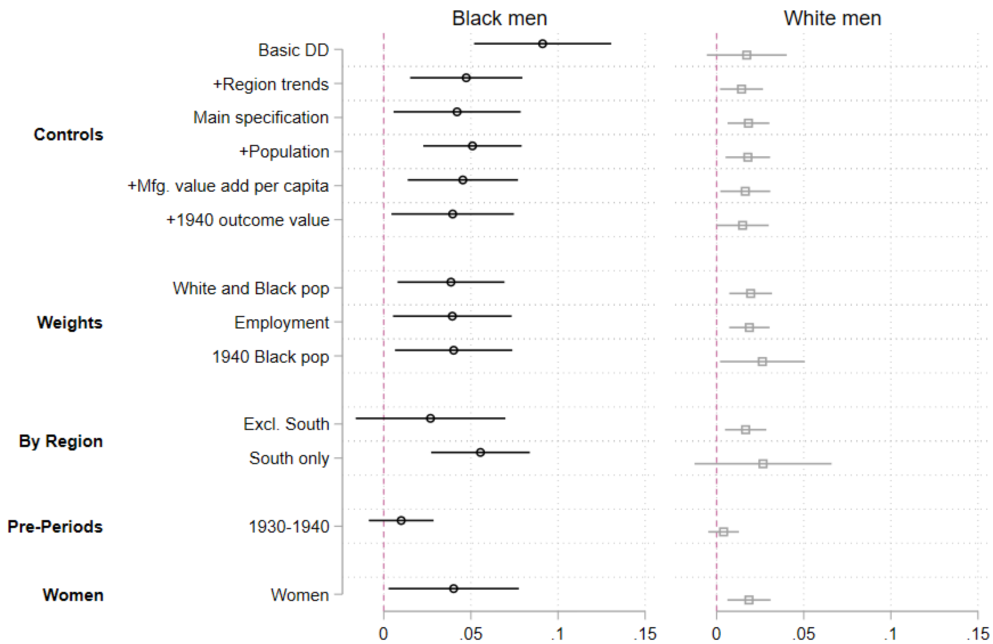
Note: Each point is an occupation from the 1950 Census occupational coding scheme with at least 10,000 employed white men. Average yearly wage is average total wage earnings within the occupational group (1940 dollars) in the previous year for men who are currently employees.

Figure B.3: Defense expenditures as a share of GDP



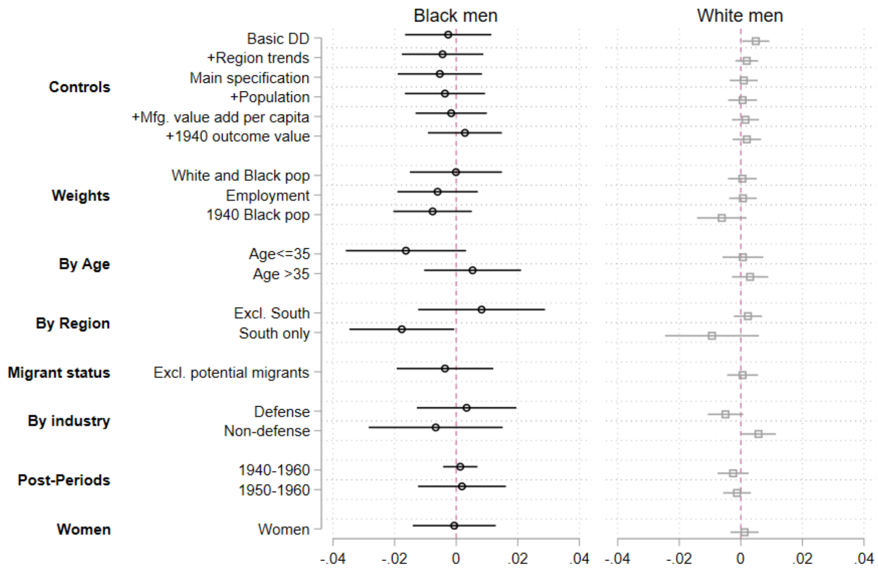
Note: Data is from U.S. Bureau of Economic Analysis series "Shares of gross domestic product: Government consumption expenditures and gross investment: Federal: National defense [A824RE1A156NBEA]," retrieved from FRED.

Figure B.4: Robustness of effects of war expenditures on ln(male population)



Note: See equation 2.1 for the basic specification. Intervals are 95% confidence intervals. All controls are interacted with an indicator for post. "Main specification" is our standard specification with controls for region, average years of education, share in manufacturing, share in agriculture, share Black, and predicted draft rate. "+Population" adds controls for the (log of) total population and Black popuation in 1940. "+1940 outcome value" adds controls for 1940 share employed, share skilled, and (log of) average yearly wage. "Excl. potential migrants" means excluding individuals in 1950 who were not born in their current state of residence and are not living with a child eight years or older born in the current state of residence. There are 146 metropolitan areas, and data comes from the 1920-1960 Census samples.

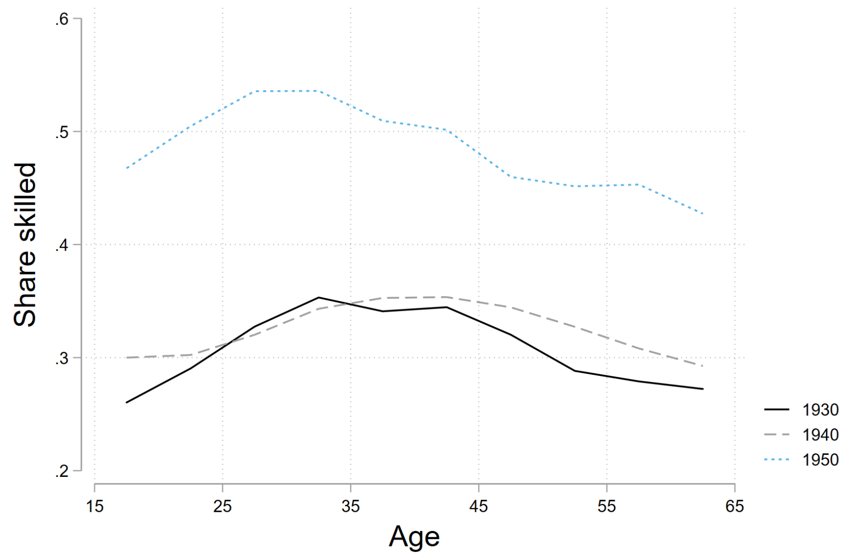
Figure B.5: Effect of war expenditures on share of prime-age men who completed high school



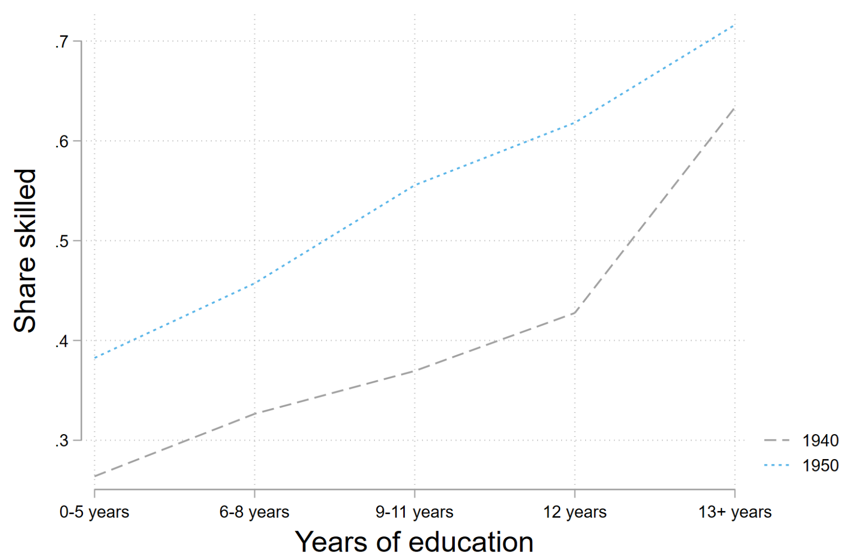
Note: See equation 2.1 for the basic specification. Intervals are 95% confidence intervals. All controls are interacted with an indicator for post. "+Base controls" is our standard specification with controls for region, average years of education, share in manufacturing, share in agriculture, share Black, and predicted draft rate. "+Population" adds controls for the (log of) total population and Black popuation in 1940. "+Baseline outcomes" adds controls for 1940 share employed, share skilled, and (log of) average yearly wage. "Excl. potential migrants" means excluding individuals in 1950 who were not born in their current state of residence and are not living with a child eight years or older born in the current state of residence. There are 146 metro areas, and data comes from the 1940-1960 Census samples.

Figure B.6: Black occupational upgrading by age and education

Panel A: By age

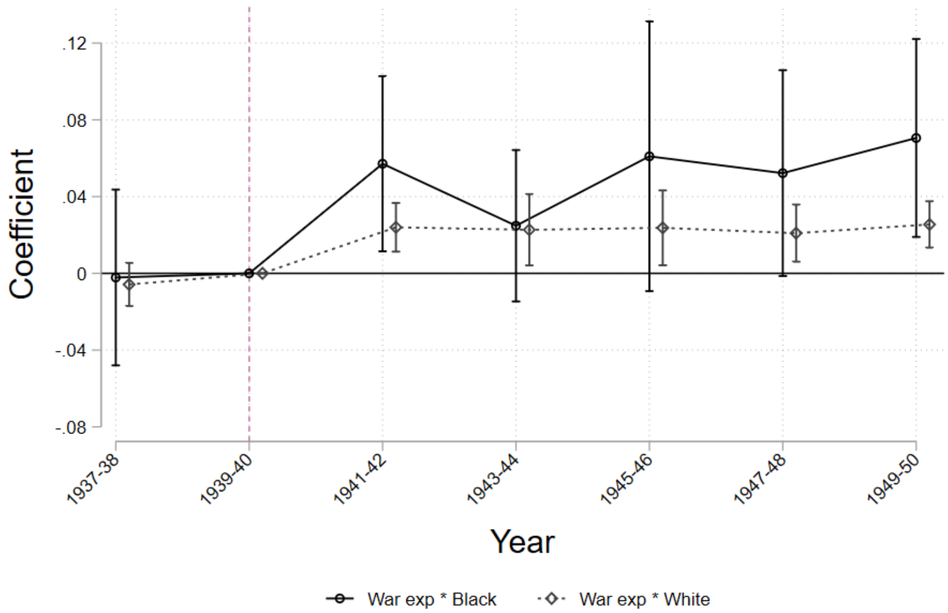


Panel B: By education



Note: Sample is employed men living in metro areas. The figure shows the share of employed Black men in skilled occupations for each education and age grouping by Census year. Data comes from the 1930 (5%), 1940 Census (100%), and 1950 (1%) samples.

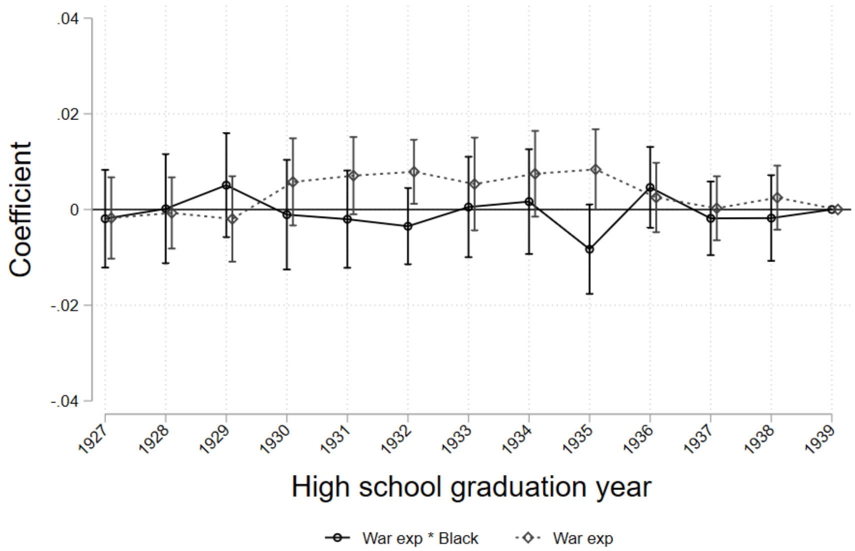
Figure B.7: Effect of war expenditures on unionization rates by race



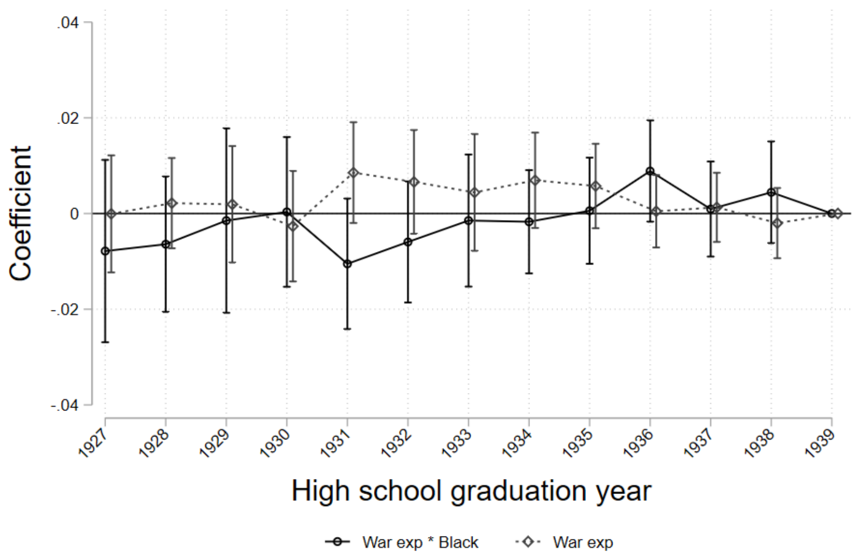
Note: The estimating equation is (for individual i, in year t, in state $r(i)$): $Union_{it} = \sum_{r\neq 1939}(\beta_r^{Bl} WarExp_r \times \mathbb{I}_{t=j} \times Black_{it} + \beta_r^{Wh} WarExp_r \times \mathbb{I}_{t=j} \times White_{it}) + \gamma_{r(i)}^{Bl} + \gamma_{r(i)}^{Wh} + \alpha_t^{Bl} + \alpha_t^{Wh} + X_{rt}\rho + \varepsilon_{it}$. The graph shows the estimates for β_r^{Wh} and β_r^{Bl} with 95% confidence intervals (SEs are clustered at the state level). In practice, time fixed effects are allowed to vary separately for the South. The sample is restricted to non-farmers and male respondents. Data is from the Gallup polls from Farber et al. (2021). If re-estimated in a simple DD framework, then coefficient on $WarExp_r \times Post_t \times Black_{it}$ is 0.029 higher than white men and the difference is statistically significant at the 5% level.

Figure B.8: Pre-trends in high school graduation rates (1940 Census)

Panel A: Boys



Panel B: Girls



Note: See equation 2.3 for the estimating equation. Intervals are 95% confidence intervals. Cohorts grouped by expected graduation year and the sample excludes the South. Graduating high school is defined as having completed 12 years of schooling in 1960. Fixed effects include metro-race FE and cohort-race FE. Other controls interacted with race include include whether born in the South interacted with cohort indicators. Results are similar if controls for veteran status are included. Data comes from the 1940 Census (100% sample; 5% sub-sample for whites).

Table B.1: Occupational distribution for Black and white men in 1940 and 1950

	Black men		White men	
	1940	1950	1940	1950
Professional	1.9	1.9	5.8	7.4
Farmers	19.6	13.0	13.5	10.2
Managers	1.4	2.0	9.9	11.1
Clerical	1.1	2.7	6.9	6.5
Sales	0.6	1.0	6.2	6.5
Craftsmen & Foremen	4.3	7.6	15.6	19.2
Military	0.2	2.6	0.7	2.4
Operatives	12.0	20.8	17.9	19.7
Domestic Service	2.9	0.8	0.3	0.1
Service Workers	11.7	13.1	5.8	5.8
Farm Laborer	18.5	10.8	7.0	4.4
Laborers	25.8	23.6	10.6	6.8

Note: Occupational distribution is for employed men in 1940 and is not limited to men in metro areas. Data is from the 1940 Census (100%).

Table B.2: Example unions policies toward Black workers during the Great Depression

Union
Example unions explicitly or effectively barring Blacks
Blacksmiths, Drop Forgers and Helpers', Brotherhood of
Boilermakers, Iron Shipbuilders and Helpers of America, International Brotherhood of
Carmen of America, Brotherhood of Railway
Clerks, Freight Handlers, Express and Station Employees, Brotherhood of Railway and Steamship
Conductors, Brotherwhood of Dining Car
Conductors Order of Sleeping Car
Conductors of America, Order of Railway
Electrical Workers, International Brotherhood of
Engineers, Grand International Brotherhood of Locomotive
Fireman and Enginemen, Brotherhood of Locomotive
Flint Glass Workers
Granite Cutters, International Association of
Journeyman Tailors
Machinists, International Association of
Mail Association. Railway
Maintence of Way Employees, Brotherhood of
Masters, Mates and Pilots, National Organization
National Rural Letter Carriers' Association
Neptune Association
Plasterers Union
Plumbers and Steam Fitters, United Association of Journeyman
Railroad Workers, American Federation of
Sheet Metal Workers
Switchmen's Union of North America
Telegraphers, Order of Railroad
Telegraphers, Union of America, Commercial
Train Dispatchers Association, American
Wire Weavers' Protective Association, American
Yardmasters of North America, Railroad
Example unions with segregated Locals
Carpenters and Joiners Unions
Painters, Decorators and Paperhangers
Hotal and Restaurant Workers

Note: Union policies are taken from "The Negro Year Book: An annual Encyclopedia of the Negro, 1937-1938" by Monroe Work and Jessie Guzman. The list is not complete.

Table B.3: Most over and under represented occupations for Black men, conditional on education and location (living in metro, 1940)

	Top 15 over-represented Occupation	$\frac{Actual}{Expected}$		Top 15 under-represented Occupation	$\frac{Actual}{Expected}$
1	Janitors and Porters	3.35	1	Tool Makers	0.05
2	Clergymen	2.93	2	Motormen	0.06
3	Private Household Workers	2.77	3	Mechanical Engineers	0.06
4	Elevator Operators	2.64	4	Civil Engineers	0.09
5	Musicians	2.22	5	Electrical Engineers	0.09
6	Service Workers, Except Private Household	2.14	6	Other Technical Engineers	0.10
7	Cooks	2.07	7	Bookkeepers	0.10
8	Recreation Workers	1.97	8	Salesmen, Wholesale	0.10
9	Teachers	1.91	9	Salesmen, Manufacturing	0.10
10	Laborer - Construction	1.86	10	Tinsmiths, Coppersmiths, and Sheet Metal Workers	0.11
11	Laundry Workers	1.86	11	Locomotive Engineers	0.11
12	Waiters and Bartenders	1.82	12	Printing Craftsmen	0.12
13	Mail Carriers	1.8	13	Foremen, Durable Goods	0.15
14	Laborer - Other	1.79	14	Foremen, Non-Durable Goods	0.15
15	Laborer - Primary Metal	1.77	15	Designers and Draftsmen	0.15

Note: For employed men living in metro areas in 1940. Expected employment is based on random assignment within educational group (5 groups) and location. For more details, see Appendix Section B.2.2. Occupation groupings are based on aggregations used in 1950 Census publications. Data is from the 1940 Census (100%).

Table B.4: Relationship between actual and predicted draft rate

	(1) Predicted draft rate	(2) Draft rate	(3) War exp per capita	(4) War exp per capita
Predicted draft rate		0.258**	-0.174***	
		(0.107)	(0.064)	
Draft %				-0.150*
				(0.081)
ln(Avg yearly wage)	0.006	-0.026	0.012	0.007
	(0.153)	(0.136)	(0.140)	(0.138)
% Agriculture	0.390**	0.158	0.021	-0.008
	(0.176)	(0.188)	(0.141)	(0.139)
% Government	0.164**	-0.113	0.078	0.039
	(0.074)	(0.104)	(0.153)	(0.137)
% Manufacturing	0.240**	0.085	0.397***	0.377**
	(0.115)	(0.140)	(0.139)	(0.145)
ln(Mfg. value added per capita)	-0.174*	-0.091	0.176**	0.186**
	(0.097)	(0.090)	(0.081)	(0.073)
% Skilled	-0.034	0.213	0.151	0.188
	(0.196)	(0.187)	(0.176)	(0.172)
% Unemployed	0.251**	0.168	-0.042	-0.051
	(0.100)	(0.102)	(0.089)	(0.096)
% Black	0.077	0.094	0.039	0.043
	(0.165)	(0.152)	(0.075)	(0.075)
ln(Population)	-0.006	0.131	-0.027	-0.007
	(0.074)	(0.090)	(0.077)	(0.078)
Northeast	-0.085	0.245*	0.021	0.069
	(0.144)	(0.142)	(0.095)	(0.094)
Midwest	-0.520***	0.111	0.034	0.121
	(0.129)	(0.159)	(0.146)	(0.146)
West	-0.224*	-0.171	0.111	0.116
	(0.114)	(0.112)	(0.097)	(0.097)
R2	0.43	0.36	0.33	0.33
N	146	146	146	146

Note: An observation is a metro area, and all variables have been standardized to have $\mu = 0$ and $\sigma^2 = 1$ and are based on 1940 values. The denominator for percentage variables is the number of employed men except for the % unemployed for which it is the number of men in the labor force. Omitted regional category is the South. War expenditure per capita in 1940 dollars. For a discussion of the draft measures see Appendix Section B.1.2. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table B.5: Effect of war expenditures on defense industry employment (1940-1944)

	Black	White	Men	Women
War exp per capita (1940)	0.034***	0.056***	0.052***	0.065***
	(0.011)	(0.009)	(0.009)	(0.012)
Mean change	.06	-.05	-.09	.07
Mean War Exp PC (1000s)	1.83	1.83	1.83	1.83
Region FE	X	X	X	X
Baseline controls	X	X	X	X
Draft control	X	X	X	X
Metro areas	146	146	146	146

Note: Sample is 146 metro areas. The outcome is the change in the share employed in defense industries. See Appendix Section B.3.1 for more details. War expenditure is $1000s per capita. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) Census sample and ES-270 reports. Metro area definitions based on 1940 and 1950 Census Bureau definitions. All values are in 1940 dollars. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table B.6: Effect of war expenditures on occupational segregation (1940-1950)

	(1)	(2)	(3)	(4)
	Occ dissimilarity index		ln(Occ dissimilarity index)	
	Basic	Adjusted	Basic	Adjusted
Panel A: All metros				
War exp per capita * Post	-0.007**	-0.008**	-0.013**	-0.017**
	(0.003)	(0.003)	(0.005)	(0.007)
N	270	270	270	270
Panel B: Excluding South				
War exp per capita * Post	-0.008**	-0.010**	-0.015**	-0.020**
	(0.003)	(0.004)	(0.006)	(0.008)
N	172	172	172	172
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro FE	X	X	X	X
Division-Year FE	X	X	X	X
Baseline controls	X	X	X	X
Draft control	X	X	X	X

Note: Full sample is 146 metro areas; occupational segregation indices are only available for 135 metro areas. See equation 2.1 for the basic specification. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table B.7: Effect of war expenditures on Black men with Bartik IV approach (1940-1950)

	(1) OLS - Controls	(2) OLS - IV Controls	(3) IV - IV Controls
Panel A: Share skilled (1940-50)			
War exp per capita * Post	0.013***	0.016***	0.015**
	(0.005)	(0.004)	(0.007)
Endogeneity test P-value			0.90
Panel B: Share skilled (1940-60)			
War exp per capita * Post	0.018***	0.021***	0.027*
	(0.006)	(0.006)	(0.016)
Endogeneity test P-value			0.45
Panel C: ln(Average yearly wage) (1940-50)			
War exp per capita * Post	0.025**	0.022**	0.004
	(0.012)	(0.011)	(0.019)
Endogeneity test P-value			0.13
Panel D: ln(Average yearly wage) (1940-60)			
War exp per capita * Post	0.018**	0.015**	0.024
	(0.008)	(0.007)	(0.020)
Endogeneity test P-value			0.42
Panel E: ln(Male population) (1940-50)			
War exp per capita * Post	0.042**	0.050***	0.089***
	(0.018)	(0.017)	(0.029)
Endogeneity test P-value			0.03
Panel F: ln(Male population) (1940-60)			
War exp per capita * Post	0.044	0.060***	0.113**
	(0.028)	(0.023)	(0.045)
Endogeneity test P-value			0.03
Metro areas	146	146	146
Mean war exp per capita	1.83	1.83	1.83
Metro FE	X	X	X
Division-Year FE	X	X	X
1st stage F-stat	-	-	32.26

Note: See equation 2.1 for the basic specification. For a discussion of the Bartik instrument, please see Appendix Section B.3.3. War expenditure is $1000s per capita. Occupational shares are shares of employed men. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Metro area definitions based on 1940 and 1950 Census Bureau definitions. All values are in 1940 dollars. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table B.8: Effect of war expenditures on Black men by geography (1940-1950)

	(1) Metro Controls	(2) CZ - Metros Controls	(3) CZ - All Controls	(4) State Controls
Panel A: Share skilled (1940-50)				
War Exp PC * Post	0.013***	0.023***	0.017***	0.020*
	(0.005)	(0.006)	(0.004)	(0.010)
Mean Y - 1940	0.33	0.28	0.22	0.22
Mean Y - 1950	0.48	0.44	0.38	0.38
Panel B: Share skilled (1940-60)				
War Exp PC * Post	0.018***	0.020**	0.011*	0.025
	(0.006)	(0.009)	(0.006)	(0.024)
Mean Y - 1940	0.33	0.28	0.22	0.22
Mean Y - 1950	0.51	0.48	0.45	0.45
Panel C: ln(Average yearly wage) (1940-60)				
War Exp PC * Post	0.018**	0.022*	0.024**	0.070**
	(0.008)	(0.012)	(0.010)	(0.034)
Mean Y - 1940	6.59	6.50	6.35	6.38
Mean Y - 1950	7.42	7.36	7.25	7.26
Geo areas	146	134	722	049
Mean war exp per capita	1.83	1.39	0.49	1.01
Geo FE	X	X	X	X
Division-Year FE	X	X		X

Note: See equation 2.1 for the basic specification. War expenditure is $1000s per capita. Occupational shares are shares of employed men. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Metro area definitions based on 1940 and 1950 Census Bureau definitions; commuting zones are 1990 definitions; SEAs are based on 1950 Census Bureau definitions. All values are in 1940 dollars. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table B.9: Effect of war expenditures on Black occupational composition

	(1) Professional	(2) Farmers	(3) Managers	(4) Clerical	(5) Sales	(6) Craftsmen	(7) Operatives	(8) Domestic service	(9) Service worker	(10) Farm labor	(11) Laborer
War exp per capita * Post	-0.001	-0.001	-0.000	0.001	-0.000	0.002	0.015***	-0.002	-0.005	-0.003*	-0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.005)	(0.001)	(0.003)	(0.002)	(0.006)
Mean Y - 1940	0.03	0.02	0.02	0.02	0.01	0.07	0.18	0.04	0.22	0.04	0.34
Mean Y - 1950	0.02	0.01	0.03	0.05	0.01	0.10	0.26	0.01	0.20	0.02	0.27
Observations	284	284	284	284	284	284	284	284	284	284	284
Mean war exp per capita	1.83	1.83	1.83	1.83	1.83	1.83	1.83	1.83	1.83	1.83	1.83
Metro FE	X	X	X	X	X	X	X	X	X	X	X
Division-Year FE	X	X	X	X	X	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X	X	X	X	X	X
Draft control	X	X	X	X	X	X	X	X	X	X	X

Note: Sample is 146 metro areas. See equation 2.1 for the basic specification. War expenditure is $1000s per capita. Occupational shares are shares of employed men. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Metro area definitions based on 1940 and 1950 Census Bureau definitions. All values are in 1940 dollars. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table B.10: Oaxaca-Blinder ln(yearly wage) decomposition

	ln(Wage) gap		Δ1940-50		
	1940	1950	Pure change	Price change	Total change
Overall	0.63	0.38			-0.25
Explained	0.37	0.22	-0.03	-0.11	-0.14
Education	0.14	0.08	-0.01	-0.06	-0.06
Occupation	0.21	0.13	-0.03	-0.04	-0.07
Industry	0.05	0.02	-0.01	-0.02	-0.03
Region	0.02	0.01	-0.01	0.00	-0.01
Age	-0.04	-0.01	0.02	0.01	0.03
Unexplained	0.26	0.16			-0.10

Note: Sample includes Black and native born white men who are wage earners. See Appendix Section B.2.3 for more detail. Education includes years of education interacted with division of birth and dummies for high school and college completion. Occupation includes dummies for ten aggregate occupational groupings. Industry includes dummies for twelve aggregated industry groupings. Region includes dummies for nine Census divisions. Age includes a cubic polynomial in age. "Explained" gaps evaluated at coefficients for white men. "Pure" change captures compositional changes, while "price' captures changing coefficients. Data comes from 1940 (100%; 5% sub-sample for white men) and 1950 (1%) Census samples.

Table B.11: Effect of war expenditures on school attendance (1940-1950)

	(1)	(2)	(3)	(4)
	Black Children		White Children	
	Boys, 16-18	Girls, 16-18	Boys, 16-18	Girls, 16-18
Panel A: All metros				
War exp per capita * Post	0.029**	0.014	-0.001	-0.001
	(0.014)	(0.013)	(0.003)	(0.003)
N - Pre	127,085	144,710	81,192	81,192
N - Post	502	531	4,341	4,341
Panel B: Excluding South				
War exp per capita * Post	0.045***	0.017	-0.000	-0.000
	(0.017)	(0.011)	(0.004)	(0.004)
N - Pre	56,707	62,974	68,712	68,712
N - Post	264	276	3,506	3,506
Mean Y - Pre	0.50	0.50	0.65	0.61
Mean Y - Post	0.54	0.57	0.69	0.64
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro-Age FE	X	X	X	X
Division-Year-Age FE	X	X	X	X
Draft control	X	X	X	X

Note: Regression at the individual level and only includes children living in one of 146 metro areas. See equation 2.1 for the basic specification. School attendance is an indicator for whether the child attended any school in the past month (1940) or two months (1950). Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Regressions weighted by sample line weights. Standard errors clustered at the metro-year level. *p<.1; **p<.05; ***p<.01

Table B.12: Effect of war expenditures on school attendance (1940-1960)

	(1)	(2)	(3)	(4)
	Black Children		White Children	
	Boys, 16-18	Girls, 16-18	Boys, 16-18	Girls, 16-18
Panel A: All metros				
War exp per capita * Post	0.012*	0.003	0.001	0.001
	(0.007)	(0.005)	(0.004)	(0.004)
N - Pre	127,085	144,710	81,192	81,192
N - Post	11,578	12,713	40,630	40,630
Panel B: Excluding South				
War exp per capita * Post	0.018**	0.004	-0.002	-0.002
	(0.008)	(0.007)	(0.005)	(0.005)
N - Pre	56,707	62,974	68,712	68,712
N - Post	6,340	7,169	32,181	32,181
Mean Y - Pre	0.50	0.50	0.65	0.61
Mean Y - Post	0.66	0.62	0.76	0.70
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro-Age FE	X	X	X	X
Division-Year-Age FE	X	X	X	X
Draft control	X	X	X	X

Note: Regression at the individual level and only includes children living in one of 146 metro areas. See equation 2.1 for the basic specification. School attendance is an indicator for whether the child attended any school in the past month (1940) or two months (1960). Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1960 (5%; 40% sub-sample for whites) Census samples. Regressions weighted by sample line weights. Standard errors clustered at the metro-year level. *p<.1; **p<.05; ***p<.01

Table B.13: Placebo effect of war expenditures on school attendance (1930-1940)

	(1)	(2)	(3)	(4)
	Black Children		White Children	
	Boys, 16-18	Girls, 16-18	Boys, 16-18	Girls, 16-18
Panel A: All metros				
War exp per capita * Post	-0.004	0.000	0.000	0.000
	(0.006)	(0.005)	(0.002)	(0.002)
N - Pre	4,983	6,119	73,194	73,194
N - Post	127,085	144,710	81,192	81,192
Panel B: Excluding South				
War exp per capita * Post	-0.013	0.004	-0.002	-0.002
	(0.008)	(0.005)	(0.003)	(0.003)
N - Pre	2,034	2,414	63,215	63,215
N - Post	56,707	62,974	68,712	68,712
Mean Y - Pre	0.38	0.38	0.50	0.46
Mean Y - Post	0.50	0.50	0.65	0.61
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro-Age FE	X	X	X	X
Division-Year-Age FE	X	X	X	X
Draft control	X	X	X	X

Note: Regression at the individual level and only includes children living in one of 146 metro areas. See equation 2.1 for the basic specification. School attendance is an indicator for whether the child attended any school in the past six months (1930) or month (1940). Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1930 (5%) and 1940 (100%; 5% sub-sample for whites) Census samples. Regressions weighted by sample line weights. Standard errors clustered at the metro-year level. *p<.1; **p<.05; ***p<.01

Table B.14: Effect of war expenditures on returns to education - Actual vs. model predicted (1940-1950)

	(1)	(2)	(3)	(4)
	Actual		Model	
	ln(Yearly wage)	Skilled	ln(Yearly wage)	Skilled
War exp per capita * Education $\geq$ 9 years * Post	-0.006	-0.001	0.002	-0.001
	(0.014)	(0.016)	(0.010)	(0.007)
War exp per capita * Post	0.021*	0.017**	0.028***	0.018***
	(0.012)	(0.008)	(0.005)	(0.002)
Education $\geq$ 9 years * Post	-0.022	-0.026	-0.002	0.001
	(0.036)	(0.031)	(0.029)	(0.020)
War exp per capita * Education $\geq$ 9 years	-0.009*	-0.012***	-0.027***	-0.009***
	(0.005)	(0.004)	(0.007)	(0.005)
Education $\geq$ 9 years	-0.044***	0.170***	0.192***	0.143***
	(0.015)	(0.010)	(0.021)	(0.014)
Observation Level	Individual	Individual	rgt	rgt
N - Pre	978,057	1,243,358	258	258
N - Post	5,163	5,925	258	258
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro FE	X	X	X	X
Division-Year FE	X	X	X	X
Baseline controls	X	X	X	X
Draft control	X	X	X	X
Individual controls	X	X	-	-

Note: Regression at the individual level for columns (1) and (2) but at metro-education group level for columns (3)-(6). Only includes men living in one of 146 metro areas. Wages are total wage earnings (1940 dollars) in the previous year for men who are currently employees. Individual controls include a cubic in age, whether born in the South, and whether married. Primary data sources are 1940 (100%; 5% sub-sample for whites) and 1950 (1%) Census samples. Regressions weighted by sample line weights ((1) and (2)) or employed population ((3)-(6)). Standard errors clustered at the metro-year level. *p<.1; **p<.05; ***p<.01

Table B.15: Effect of war expenditures wages, education, and age within group, industry and occupation

	(1) ln(Avg. wage)	(2) Avg. education	(3) Avg. age
β_1	-0.002	-0.003	0.069
	(0.008)	(0.025)	(0.211)
β_2	-0.000	-0.014	0.005
	(0.009)	(0.022)	(0.184)
β_3	-0.001	-0.005	0.101
	(0.014)	(0.096)	(0.370)
β_4	0.008	0.043	0.252
	(0.016)	(0.066)	(0.507)
Observations	3,530	3,530	3,530
R-squared	0.980	0.997	0.960
γ_{riog}	X	X	X
γ_{iogt}	X	X	X
γ_{rgt}	X	X	X

Note: Sample is 146 metro areas. Metro area definitions based on 1940 and 1950 Census Bureau definitions. Regressions are weighted by relevant population. Standard errors are clustered at the metro-year level. *p<.1; **p<.05; ***p<.01

Table B.16: Evaluating actual estimated changes versus model data

	(1)	(2)	(3)	(4)
	Share skilled		ln(Avg yearly wage)	
	Black	White	Black	White
Panel A: Actual data				
War exp per capita * Post	0.013***	0.000	0.025**	0.006*
	(0.005)	(0.001)	(0.012)	(0.004)
Panel B: Model-generated data				
War Exp PC * Post	0.017***	0.001***	0.028***	0.006***
	(0.002)	(0.000)	(0.004)	(0.000)
Metro areas	146	146	146	146
Mean war exp per capita	1.83	1.83	1.83	1.83
Metro FE	X	X	X	X
Division-Year FE	X	X	X	X
Baseline controls	X	X	X	X
Draft control	X	X	X	X

Note: Sample is 146 metro areas. Wages are total wage earnings in the previous year for men who are currently employees. Baseline controls are 1940 variables interacted with a post indicator: average years of education, share employed in manufacturing, share employed in agriculture, and share Black. Draft control is predicted draft rate based on 1940 demographics. Primary data sources are 1940 (100%) and 1950 (1%) Census samples. Metro area definitions based on 1940 and 1950 Census Bureau definitions. Regressions are weighted by relevant population. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

APPENDIX C

Appendix Materials for Chapter 3

C.1 Valuing working conditions

C.1.1 Theoretical framework

Gronberg and Reed (1994) propose a basic model that allows the estimation of the marginal willingness to pay for job attributes. Take a basic search model where jobs are characterized by some vector of attributes, X, such as wages and benefits. Let $s\lambda$ be the arrival rate of job offers where λ is a firm determined offer rate and s is search effort on the part of the worker. There is some cost of search effort, $c(s)$ where $c(0) = 0$, $c'(s) > 0$, and $c''(s) > 0$ – i.e. the marginal cost of search is increasing in search effort. Finally, let b be the value of the outside option and $v(X)$ be the utility of job with characteristics X.

Firms offer some exogenous distribution of utility offers, $F(w)$ and δ is an exogenous rate of job destruction. Given this setup, a worker has some optimal search effort, $s^*(v(X))$, that depends on their current job. The hazard rate of a job spell is given by:

$$h(v(X)) = \delta + \lambda s^*(v(X))(1 - F(v(X)))$$

The derivative of the hazard rate with respect to attribute i is:

$$\frac{\partial h}{\partial X_i} = \frac{\partial v}{\partial X_i}(\frac{\partial s^*}{\partial v}\lambda(1 - F(v(X))) + \frac{\partial(1 - F(v))}{\partial v}\lambda s^*)$$

If we take the ratio of this expression for any attribute relative to the same expression for wages we get:

$$\frac{\frac{\partial h}{\partial X_i}}{\frac{\partial h}{\partial w}} = \frac{\frac{\partial v}{\partial X_i}}{\frac{\partial v}{\partial w}}$$
$$= MWP_i(X)$$

The hazard rate over a year is the number of leavers divided by the initial employment. In the steady state, the number of leavers equals the number of recruits in order to maintain constant employment so the hazard rate is approximately equal to the turnover rate. Therefore, comparing the relative coefficients in a regression with turnover gives the marginal willingness to pay for attributes. I parameterize the hazard rate as $h(v(X)) = \lambda \exp X\beta$.

C.1.2 Estimation

Specification: The regression specification for facility i at time t is:

$$\ln Turnover_{it} = \beta^w \ln W_{it} + \beta^e \ln E_{it} + \beta^b \ln B_{it} + \theta_t + \varepsilon_{it}$$

where W_{it} is the average hourly wage, E_{it} is effort (the workload), B_{it} are benefits, and θ_t are county-year fixed effects. I use the wages and turnover rates for nursing assistants. ε_{it} is the error term and it is clustered at the facility level. I also try a specification that includes facility fixed effects, μ_i. For the within-facility specification I omit benefits since it appears there is significant noise in the time series variation. One potential issue is that these variables are likely endogenous. Therefore I instrument for E_{it} using variation in patient severity since facilities do not seem to adjust staffing based on temporary variation in patient severity.

Results: The regression results can be seen in Appendix Table C.6. All three measures are strongly correlated with turnover. The coefficient for wages is roughly twice the magnitude of workload, which implies that a 2% decrease in workload is equivalent to a 1% increase in wages. The coefficient on benefits is about one fifth of the coefficient on wages. Total expenditure on nursing benefits is about one fourth of the total expenditures on nursing salaries, so I cannot reject the hypothesis that a dollar increase in benefits is equivalent to a dollar increase in wages.[1]

The instrument for workload does not substantially change the estimates. This suggests that there are not significant endogeneity concerns. However, I have not instrumented wages, so there might be omitted variables that affect these estimates. Therefore, I also compare my estimate with others in the literature.

The best identified estimates involve the labor supply elasticity rather than the elasticity of turnover with respect to wages. Following arguments in Manning (2003), the elasticity of labor supply is approximately twice the separation elasticity. Therefore, these estimates imply a labor supply elasticity of ≈ 2. In a recent meta-analysis, Sokolova and Sorensen (2020) find a median estimate of 1.69. Therefore, these estimates seem reasonable.

To place a value on amenities, I use the following values based:

- Marginal willingness to pay for workload: a 2% decrease in workload is equivalent to a 1% increase in wages
- Marginal willingness to pay for benefits: a \$1 increase in benefits is equivalent to a \$1 increase in wages

[1]Some benefits are taxed, such as vacation or sick days, while others enjoy significant tax breaks.

C.2 Returns to quality

One potential issue with interpreting the results is if unobserved worker quality plays a large role in determining wages and working conditions. In this section I provide evidence that there are low returns to observable quality measures for nursing assistants.

Returns to education: First, I look at the returns to education for nursing assistants in nursing homes. I use the 2005-2018 ACS and regress the log of the hourly wage on education. I control for a variety of demographic characteristics, including a cubic polynomial in age, gender, race, language, and citizenship status. The results are in Appendix Table C.2. Column (1) shows high returns to education for all workers. Column (2) shows that there are no observable returns to education for nursing assistants. However, selection on unobserved quality could result in this finding if the workers with higher education who choose to be nursing assistants are negatively selected. For comparison, I construct a set of comparable occupations: cleaning workers, childcare workers, beauticians, secretaries, grocery cashiers, and food preparation workers. Any selection argument would likely affect these occupations as well. Column (3) shows that these occupations still have significantly higher returns to education, suggesting that the results for nursing assistants are not solely due to selection.

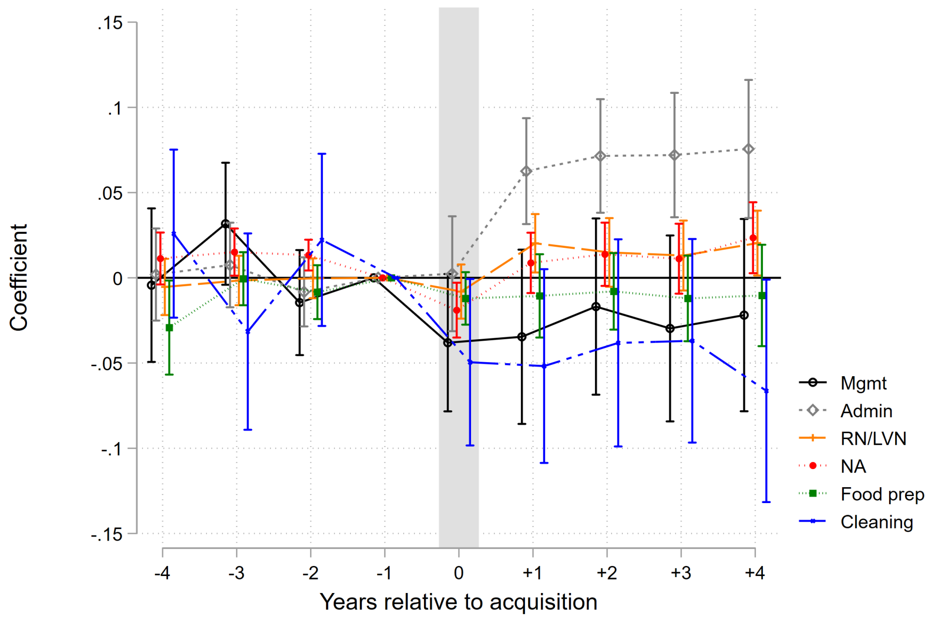
Returns to experience: An alternative approach is to look at the returns to experience. I use the panel structure of the CPS to construct a measure of occupational experience. Experience is an indicator if the worker was in the same occupation one year previously. The results are in Appendix Table C.3. Column (1) shows the results for all workers and again shows high returns to experience. Column (2) shows that there are limited returns to experience for nursing assistants. Column (3) shows that similar occupations have higher returns to experience.

Unobserved worker quality: Another measure of unobserved quality is if a worker has been previously fired from their job. In 2004, the U.S. Department of Health and Human Services conducted the National Nursing Assistant Survey. The survey was a nationally representative sample of nursing assistants in nursing homes; as part of the survey workers were asked about employment history. I regress current wages on whether a worker had been fired from a job in the previous two years. I also include some basic demographic and facility controls. The results are in Appendix Table C.4. The results show that there is no relationship between current wages and whether the worker was previously fired.

A second measure of the returns to unobserved quality is to see how wages in previous jobs predict current wages. If wages are differentiated among workers based on ability, then there should be a correlation between previous wages (at a different job) and current wages. If instead, wages are primarily determined at the firm level then wages of co-workers should be the strongest predictor. Using the NNAS survey, I regress worker's wages on previous wages and on the wages of coworkers. Note that these wages are self-reported and there are only a few observations from each sampled facility so there is noise in the co-workers' average wage. The results are in Appendix Table C.5.The wages of co-workers are a much stronger predictor of a worker's wages than the worker's previous wages. This result suggests that wages are set at the facility level.

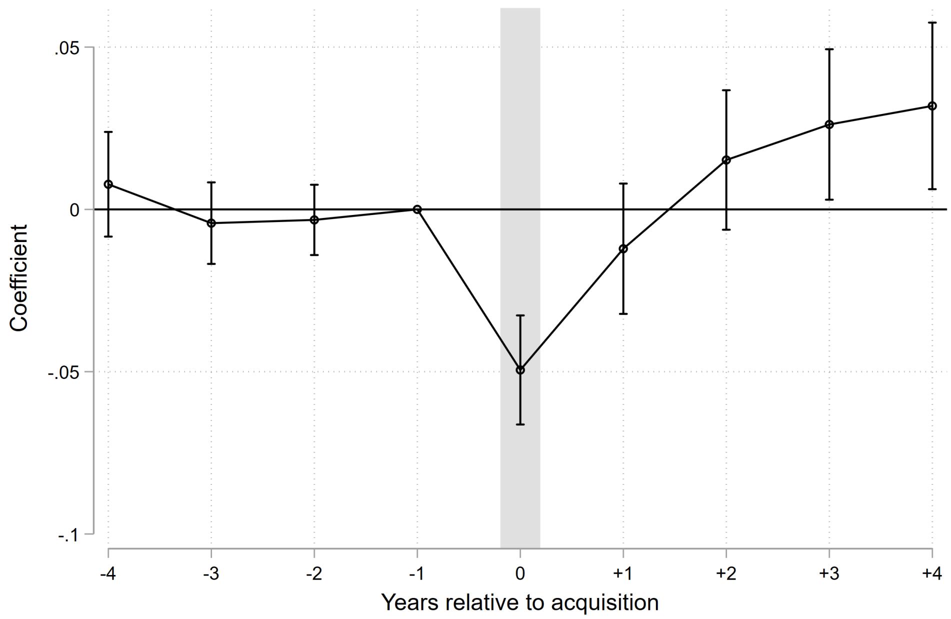
C.3 Appendix tables and figures

Figure C.1: Effect of acquisitions on log of total hours worked by occupation



Note: For California for-profit nursing home facilities (1997-2019). Includes county-year and facility FE.

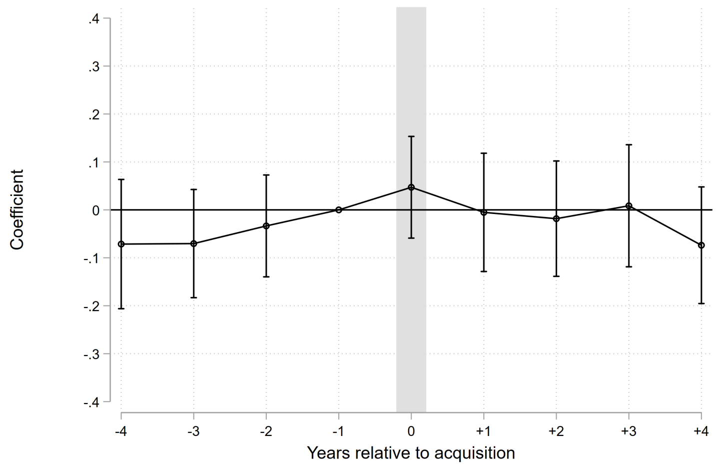
Figure C.2: Effect of acquisitions on log of employment



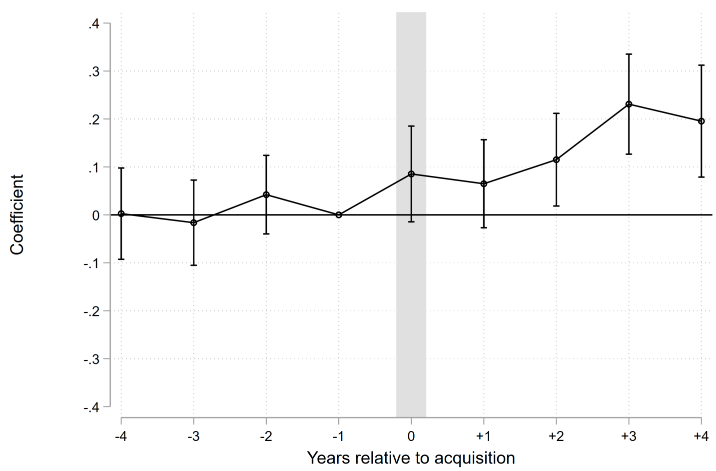
Note: For California for-profit nursing home facilities (1997-2019). Includes county-year and facility FE.

Figure C.3: Effect of acquisitions on standardized patient outcomes

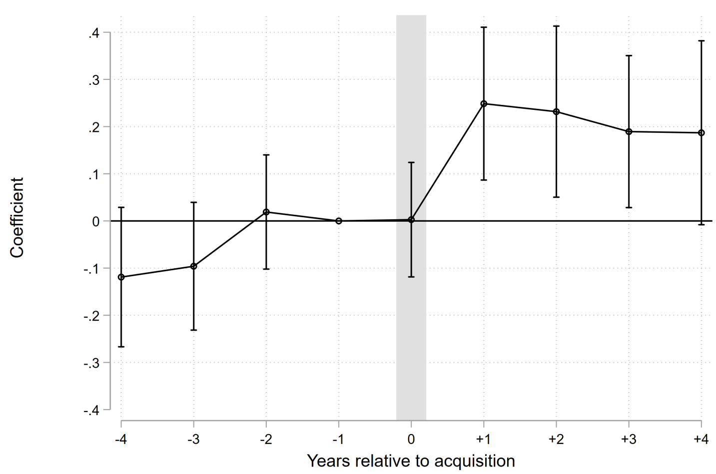
Panel A: Total Deficiency Score



Panel B: Discharges to Death or Hospital per patient day

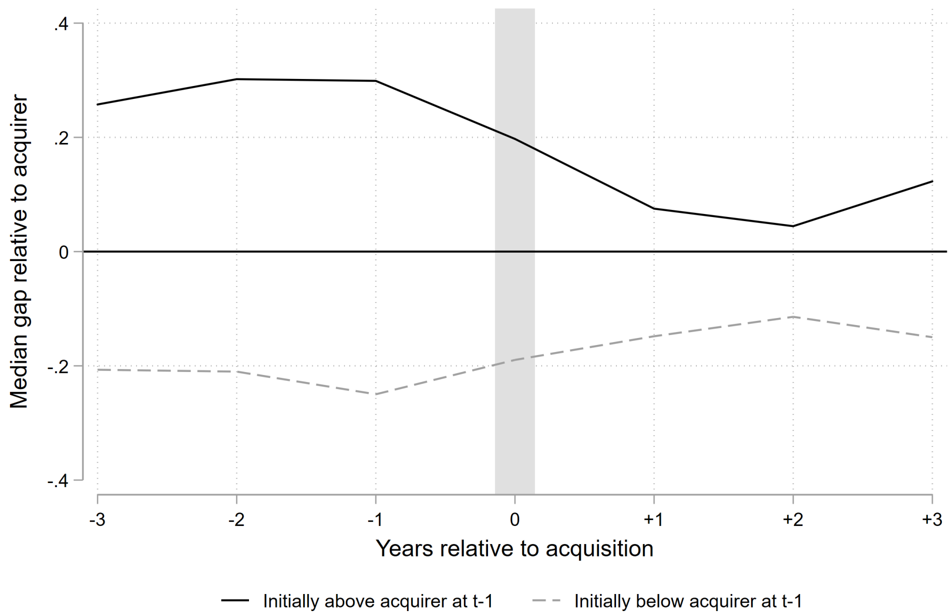


Panel C: ADL Decline



Note: For California for-profit nursing home facilities (2000-2019). Includes county-year and facility FE. Ranges are 95% confidence intervals.

Figure C.4: Gap in log of benefits between acquired facility and acquiring firm by year relative to acquisition



Note: For California for-profit nursing home facilities (1997-2019) that were acquired by an identified chain with at least three establishments. Gaps are relative to acquiring chain's median value at time t-1 and have been adjusted for general time trends

Table C.1: Nursing home worker characteristics by occupations (CA 2005-2018)

	NA	LVN	RN	Other med	Admin	Other	Overall
Share of total empl	0.39	0.14	0.13	0.09	0.12	0.13	1.00
Compensation							
Hourly wage	13.51	22.45	32.24	28.67	30.15	12.66	20.62
Employer insurance	0.54	0.69	0.76	0.75	0.76	0.52	0.63
Demographics							
Female	0.81	0.82	0.84	0.75	0.76	0.63	0.78
Age	40.54	38.03	43.50	43.35	45.01	44.20	41.84
White, non-Hispanic	0.15	0.17	0.26	0.40	0.36	0.15	0.21
White, Hispanic	0.39	0.29	0.11	0.22	0.27	0.60	0.33
Black	0.12	0.10	0.09	0.08	0.04	0.05	0.09
Asian	0.32	0.41	0.52	0.28	0.30	0.19	0.34
Immigrant	0.52	0.48	0.57	0.34	0.36	0.57	0.49
Language at home - Spanish	0.33	0.21	0.09	0.16	0.19	0.52	0.27
Language at home - Filipino	0.21	0.24	0.30	0.15	0.18	0.11	0.20
Education							
Less than high school	0.15	0.01	0.00	0.03	0.05	0.32	0.11
High school	0.36	0.24	0.03	0.13	0.21	0.43	0.27
Some college	0.34	0.57	0.35	0.23	0.29	0.19	0.34
College	0.15	0.18	0.61	0.62	0.45	0.06	0.28

Note: Data is from the 2005-2018 ACS samples and is for all workers who are employed by a nursing home and live in California at the time of the survey.

Table C.2: Returns to education for Nursing Assistants in Nursing Homes (ACS 2005-2018)

	(1)	(2)	(3)
ln(Hourly wage)	All	NA in NH	Similar jobs
Years of education	0.0611***	0.00813	0.0307***
	(0.000505)	(0.00780)	(0.00209)
Observations	753,822	1,639	32,114
R-squared	0.464	0.117	0.149
County FE	X	X	X
Year FE	X	X	X
Occ FE	X	-	X
Ind FE	X	-	X
Controls	X	X	X

Note: Data is from the 2005-2018 ACS samples for California. Other controls include a cubic polynomial in age, gender, race, language, and citizenship status. Comparison jobs include cleaning, childcare, beauticians, secretaries, grocery cashiers, and food prepartion workers. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table C.3: Returns to experience for Nursing Assistants in Nursing Homes (CPS 1994-2019)

ln(Hourly wage)	(1) All	(2) NAs in NHs	(3) Similar jobs
Occ. Experience	0.0461***	-0.00627	0.00976*
	(0.00170)	(0.0161)	(0.00523)
Year	0.00172***	0.00340***	0.00300***
	(0.000110)	(0.00108)	(0.000340)
Age	0.0690***	0.0655***	0.0639***
	(0.00176)	(0.0178)	(0.00489)
Age sq.	-0.00119***	-0.00137***	-0.00120***
	(4.70e-05)	(0.000472)	(0.000132)
Age cu.	6.60e-06***	9.49e-06**	7.41e-06***
	(3.97e-07)	(3.96e-06)	(1.12e-06)
Observations	302,159	2,078	23,694
R-squared	0.437	0.078	0.179
Controls	X	X	X

Note: Data is from the 1994-2019 CPS samples who were part of the Earner Study in their 8th interview. Experience is based on whether individuals were employed in the same occupation one year prior. Other controls include gender and race. Comparison jobs include cleaning, childcare, beauticians, secretaries, grocery cashiers, and food prepartion workers. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table C.4: Relationship between cross-sectional wages and whether a nursing assistant has been laid off in the previous two years (NNAS 2004)

ln(Wage)	(1)	(2)
Ever laid off	-0.00557	-0.00239
	(0.0145)	(0.0171)
Ever laid off from NH		-0.00844
		(0.0279)
Previous job NA in NH	0.0186	0.0205
	(0.0144)	(0.0158)
Observations	1,355	1,355
R-squared	0.178	0.179
Controls	X	X

Note: Data is from the 2004 NNAS. Ever laid off means whether the wokrer had been laid off from any job in the previous two years (including in industries other than nursing homes). Controls include type of ownership, facility size, and whether located in a metro. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table C.5: Relationship between cross-sectional wages with co-workers' wages and wages in previous job for job-switchers (NNAS 2004)

ln(Hourly wage)	(1)
ln(Leave-out facility avg. wage)	0.718***
	(0.0430)
ln(Wage) in previous job	0.0806***
	(0.0277)
Previous job NA in NH	0.0186
	(0.0186)
Observations	1,230
R-squared	0.627
Controls	X

Note: Data is from the 2004 NNAS and includes only individuals who have been employed in a new job in the previous two years. Controls include type of ownership, facility size, and whether located in a metro. Robust standard errors in parentheses. *p<.1; **p<.05; ***p<.01

Table C.6: Relationship between workplace conditions and log of NA turnover)

	(1)	(2)	(3)	(4)
	Within county		Within facility	
	OLS	IV - Severity	OLS	IV - Severity
ln(NA Workload)	0.476***	0.609***	0.418***	0.423*
	(0.0929)	(0.145)	(0.120)	(0.240)
ln(Avg. NA wages)	-1.037***	-1.063***	-1.076***	-1.077***
	(0.148)	(0.152)	(0.167)	(0.174)
ln(Benefits per empl.)	-0.207***	-0.208***		
	(0.0500)	(0.0498)		
Observations	12,213	12,213	12,206	12,206
R-squared	0.228	0.046	0.453	0.012
County-Year FE	X	X	X	X
Facility FE	-	-	X	X

Note: Data is from OSHPD and includes for-profit California nursing homes from 2002-2017. Standard errors are clustered at the facility level *p<.1; **p<.05; ***p<.01

Table C.7: Effect of acquisitions on nursing working conditions - first difference approach

	(1)	(2)	(3)	(4)
	Δ ln(Wage [Adj.])	Δ ln(Benefits)	Δ ln(Workload)	Δ ln(Turnover)
Panel A: Basic				
Acquired	-0.00254	-0.0578***	0.0241***	0.131***
	(0.00300)	(0.00983)	(0.00720)	(0.0312)
Panel B: Controlling for trend in Medicare days				
Acquired	-0.00517	-0.0662***	0.0260***	0.120***
	(0.00335)	(0.0112)	(0.00714)	(0.0346)
Δ ln(Medicare days)	0.00185	-0.00186	0.0111***	-0.0387***
	(0.00166)	(0.00603)	(0.00287)	(0.0143)
Observations	12,004	6,363	11,902	11,896
Year FE	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table C.8: Effect of acquisitions on patient outcomes - first difference approach

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Total Deficiency Score		Δ Discharge to Hosp/Death per day		Δ ADL Decline	
Acquired	-0.0293	-0.0295	0.196***	0.217***	0.275***	0.242***
	(0.0520)	(0.0603)	(0.0501)	(0.0460)	(0.0711)	(0.0793)
Trend $t-4$ to $t-1$		0.0192**		0.146***		0.0734***
		(0.00950)		(0.0288)		(0.0228)
Observations	8,862	7,710	6,277	5,163	3,798	2,766
R-squared	0.002	0.003	0.005	0.056	0.007	0.020
Year FE	X	X	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Table C.9: Effect of acquisitions on working conditions by whether above or below acquiring firm

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ ln(Wage [Adj.])		Δ ln(Benefits)		Δ ln(Staffing)	
Acquired	0.0127**	0.00944	-0.0475**	-0.0616*	0.0139*	-0.0125
	(0.00585)	(0.00880)	(0.0228)	(0.0315)	(0.00740)	(0.0109)
Δ Above Acquirer - Wage	-0.356***	-0.407***		0.409**		0.173**
	(0.0585)	(0.0602)		(0.204)		(0.0775)
Δ Below Acquirer - Wage	0.168**	0.257***		-0.0909		0.193
	(0.0766)	(0.0755)		(0.308)		(0.143)
Δ Above Acquirer - Benefits		0.0298	-0.461***	-0.543***		0.0307
		(0.0221)	(0.0792)	(0.0815)		(0.0287)
Δ Below Acquirer - Benefits		-0.0526**	0.654***	0.669***		0.0266
		(0.0204)	(0.0717)	(0.0775)		(0.0329)
Δ Above Acquirer - Staffing		0.110**		0.0937	-0.551***	-0.565***
		(0.0477)		(0.214)	(0.0992)	(0.0969)
Δ Below Acquirer - Staffing		-0.0334		-0.0139	0.380***	0.387***
		(0.0370)		(0.138)	(0.0807)	(0.0772)
Δ Above Median	0.0114	0.0114	5.56e-05	-0.00110	-0.105***	-0.106***
	(0.00937)	(0.00936)	(0.0123)	(0.0123)	(0.0174)	(0.0173)
Δ Below Median	0.0648***	0.0648***	0.0503***	0.0510***	0.113***	0.112***
	(0.0111)	(0.0110)	(0.0159)	(0.0158)	(0.0392)	(0.0392)
Observations	12,104	12,104	11,243	11,243	12,050	12,050
R-squared	0.560	0.560	0.071	0.071	0.138	0.139
Year FE	X	X	X	X	X	X

Note: Sample is California non-specialized nursing homes from 1997 to 2019. Standard errors are clustered at the facility level. *p<.1; **p<.05; ***p<.01

Bibliography

Aaronson, Daniel, and Bhashkar Mazumder. 2011. "The Effect of Rosenwald Schools on Black Achievement." *Journal of Political Economy*, 119(5): 821–888.

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering." Working Paper 24003.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251–333.

Acemoglu, Daron, David Autor, and David Lyle. 2004. "Women, War and Wages: The Effect of Female Labor Supply on the Wage Structure at Mid-Century." *Journal of Political Economy*, 112(3): 497–551.

Antwi, Yaa Akosa, and John R. Bowblis. 2018. "The Impact of Nurse Turnover on Quality of Care and Mortality in Nursing Homes: Evidence from the Great Recession." *American Journal of Health Economics*, 4(2): 131–163.

Arcidiacono, Peter, Paul B. Ellickson, Carl F. Mela, and John D. Singleton. 2020. "The Competitive Effects of Entry: Evidence from Supercenter Expansion." *American Economic Journal: Applied Economics*, 12(3): 175–206.

Arnold, David. 2020. "Mergers and Acquisitions, Local Labor Market Concentration, and Worker Outcomes." Working Paper.

Ashenfelter, Orley, and Alan B. Krueger. 2018. "Theory and Evidence on Employer Collusion in the Franchise Sector." Working Paper 24831.

Asker, John. 2010. "A Study of the Internal Organization of a Bidding Cartel." *American Economic Review*, 100(3): 724–762.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2019. "(Mis)Allocation, Market Power and Global Oil Extraction." *American Economic Review*, 109(4): 1568–1615.

Asker, John, and Volker Nocke. 2021. "Collusion, Mergers, and Related Antitrust Issues." Working Paper.

Baye, Michael R., John Mogan, and Patrick Scholten. 2006. "Information, Search, and Price Dispersion." In *Economics and Information Systems.* , ed. Terrence Hendershott, 322–377. Elsevier Science Ltd.

Bayer, Patrick, and Kerwin Kofi Charles. 2018. "Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940." *The Quarterly Journal of Economics*, 133(3): 1459–1501.

Bertrand, Marianne, and Sendhil Mullainathan. 2003. "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences." *Journal of Political Economy*, 111(5): 1043–75.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Bianchi, Nicola, and Michela Giorcelli. 2020. "Not All Management Training Is Created Equal: Evidence from the Training Within Industry Program." Working Paper.

Billips, Robert S. 1936. "Hourly Entrance Rates of Common Laborers in 20 Industries, July 1936." *Monthly Labor Review*, 44(4): 938–952.

Bishop, John. 1989. "Why the Apathy in American High Schools." *Educational Researcher*, 18(1): 6–10.

Blonigen, Bruce A., and Justin R. Pierce. 2016. "Evidence for the Effects of Mergers on Market Power and Efficiency." Working Paper 22750.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2019. "Quasi-experimental Shift-share Research Designs." Working Paper.

Bos, Iwan, and Joseph E. Harrington. 2010. "Endogenous Cartel Formation with Heterogeneous Firms." *The RAND Journal of Economics*, 41(1): 92–117.

Bourveau, Thomas, Guoman She, and Alminas Zaldokas. 2020. "Corporate Disclosure as a Tacit Coordination Mechanism: Evidence from Cartel Enforcement Regulations." *Journal of Accounting Research*, 58(2): 295–332.

Boustan, Leah Platt. 2009. "Competition in the Promised Land: Black Migration and Racial Wage Convergence in the North, 1940-1970." *Journal of Economic History*, 69(3): 755–782.

Boustan, Leah Platt. 2010. "Was Postwar Suburbanization 'White Flight'? Evidence from the Black Migration." *Quarterly Journal of Economics*, 125(1): 417–443.

Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson. 2015. "Acquisitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry." *The American Economic Review*, 105(7): 2086–2119.

Bresnahan, Timothy F. 1987. "Competition and Collusion in the American Automobile Industry: The 1955 Price War." *Journal of Industrial Economics*, 35(4): 457–482.

Brouillette, Jean-Felix, Charles I. Jones, and Peter J. Klenow. 2021. "Race and Economic Well-being in the United States." Working Paper.

Brown, Charles, and James L. Medoff. 1988. "The Impact of Firm Acquisitions on Labor." In *Corporate Takeovers: Causes and Consequences.* , ed. Alan J. Auerbach, 9–32. NBER Books.

Brunet, Gillian. 2018. "Stimulus on the Home Front: The State-level Effects of WWII Spending." Working Paper.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel. 2019. "Changes in Between-group Inequality: Computers, Occupations, and International Trade." *American Economic Journal: Macroeconomics*, 11(2): 348–400.

Byrne, David P., and Nicolas de Roos. 2019. "Learning to Coordinate: A Study in Retail Gasoline." *American Economic Review*, 109(2): 591–619.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*, 36(S1): S13–S69.

Card, David, and Alan Krueger. 1992. "School Quality and Black-White Relative Earnings: A Direct Assessment." *The Quarterly Journal of Economics*, 107(1): 151–200.

Carrington, William J., and Kenneth R. Troske. 1997. "On Measuring Segregation in Samples with Small Units." *Journal of Business and Economic Statistics*, 15(4): 402–409.

Carruthers, Celeste K., and Marianne H. Wanamaker. 2017. "Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap." *Journal of Labor Economics*, 35(3): 655–692.

Castle, Nicholas G., John Engberg, Ruth Anderson, and Aiju Men. 2007. "Job Satisfaction of Nurse Aides in Nursing Homes: Intent to Leave and Turnover." *The Gerontologist*, 47(2): 193–204.

Christie, William G., and Paul H. Schultz. 1994. "Why Do NASDAQ Market Makers Avoid Odd-Eighth Quotes?" *Journal of Finance*, 49: 1813–1840.

Clay, Karen, Jeff Lingwall, and Melvin Stephens. 2012. "Do Schooling Laws Matter? Evidence from the Introduction of Compulsory Attendance Laws in the United States." Working Paper 18477.

Collins, William. 2000. "African-American Economic Mobility in the 1940s: A Portrait from the Palmer Survey." *Journal of Economic History*, 60(3): 756–781.

Collins, William. 2001. "Race, Roosevelt, and Wartime Production: Fair Employment in World War II Labor Markets." *The American Economic Review*, 91(1): 272–286.

Collins, William J. 2020. "The Great Migration of Black Americans from the US South: A Guide and Interpretation." Working Paper 27268.

Collins, William J., and Marianne H. Wanamaker. 2014. "Selection and Income Gains in the Great Migration of African Americans." *American Economic Journal: Applied*, 6(1): 220–252.

Collins, William J., and Robert A. Margo. 2003. "Historical Perspectives on Racial Differences in Schooling in the United States." Working Paper 9770.

Connor, John M., and Yuliya Bolotova. 2006. "Cartel Overcharges: Survey and Meta-Analysis." *International Journal of Industrial Organization*, 24.

Cooper, David J. 1997. "Barometric Price Leadership." *International Journal of Industrial Organization*, 15(3): 301–325.

Cooper, David J., and Kai-Uwe Kuhn. 2014. "Communication, Renegotiation, and the Scope for Collusion." *American Economic Journal: Microeconomics*, 6(2): 247–278.

Currie, Janet, Mehdi Farsi, and W. Bentley Macleod. 2005. "Cut to the Bone? Hospital Takeovers and Nurse Employment Contracts." *ILR Review*, 58(3): 471–493.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. 1999. "The Rise and Decline of the American Ghetto." *Journal of Political Economy*, 107(3): 455–506.

David, Joel M. 2020. "The Aggregate Implications of Mergers and Acquisitions." Working Paper.

Davis, Steven J., John Haltiwanger, Kyle Handley, Ron Jarmin, Josh Lerner, and Javier Miranda. 2014. "Private Equity, Jobs, and Productivity." *The American Economic Review*, 104(12): 3956–3990.

DellaVigna, Stefano, and Matthew Gentzkow. 2019. "Uniform Pricing in U.S. Retail Chains." *The Quarterly Journal of Economics*, 134(4): 2011–2084.

Derenoncourt, Ellora. 2019. "Can You Move to Opportunity? Evidence from the Great Migration." Working Paper.

Derenoncourt, Ellora, and Claire Montialoux. 2019. "Minimum Wages and Racial Inequality." Working Paper.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of Diverging Location Choices by Skill: 1980-2000." *American Economic Review*, 106(3): 479–524.

DOJ, and FTC. 2010. *Horizontal Merger Guidelines.* Washington, D.C.:U.S. Federal Government.

Donohue, John J., and James Heckman. 1991. "Continuous Versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks." *Journal of Economic Literature*, 29(4): 1603–1643.

Duncan, Greg J., and Katherine Magnuson. 2012. "Socioeconomic Status and Cognitive Functioning: Moving from Correlation to Causation." *Wiley Interdisciplinary Review: Cognitive Science*, 3(3): 377–386.

Duncan, Otis, and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indexes." *American Sociological Review*, 20(2): 210–217.

Eckbo, B. Espen. 2014. "Corporate Takeovers and Economic Efficiency." *Annual Review of Financial Economics*, 61: 51–74.

Eckert, Fabian, Adres Gvirtz, and Michael Peters. 2018. "A Consistent County-level Crosswalk for US Spatial Data Since 1790." Working Paper.

Eliason, Paul J., Benjamin Heebsh, Ryan C. McDevitt, and James W. Roberts. 2020. "How Acquisitions Affect Firm Behavior and Performance: Evidence from the Dialysis Industry." *The Quarterly Journal of Economics*, 135(1): 221–267.

Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suarez Serrato, and Owen Zidar. 2019. "State Taxes and Spatial Misallocation." *The Review of Economic Studies*, 86(1): 333–376.

Farber, Henry S., Daniel Herbst, Ilyana Kuziemko, and Suresh Naidu. 2021. "Unions and Inequality Over the Twentieth Century: New Evidence from Survey Data." *The Quarterly Journal of Economics*, 136(3): 1325–1385.

Ferrara, Andreas. 2020. "World War II and Black Economic Progress." Working Paper.

Fishback, Price, and Joseph A. Cullen. 2013. "Second World War Spending and Local Economic Activity in U.S. Counties, 1939-58." *The Economic History Review*, 66(4): 975–992.

Fishback, Price V., and Valentina Kachanovskaya. 2015. "The Multiplier for Federal Spending in the States During the Great Depression." *The Journal of Economic History*, 75(1): 125–162.

Flood, John. 1989. "Megalaw in the U.K.: Professionalism or Corporatism? A Preliminary Report." *Indiana Law Journal*, 64(3): 569–592.

Frazier, Edward K., and Jacob Perlman. 1939. "Entrance Rates of Common Laborers, July 1939." *Monthly Labor Review*, 46(12): 1450–1465.

Friedrich, Benjamin U., and Martin B. Hackmann. 2021. "The Returns to Nursing: Evidence from a Parental-Leave Program." *Review of Economic Studies.*

Galanter, Marc, and Thomas Palay. 1991. *Tournament of Lawyers.* Chicago:University of Chicago Press.

Galanter, Mark, and William Henderson. 2010. "The Elastic Tournament: A Second Transformation of the Big Law Firm." *The Stanford Law Review*, 60: 1867–1930.

Galle, Simon, Andres Rodriguez-Clare, and Moises Yi. 2018. "Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade." Working Paper.

Gandhi, Ashvin. 2019. "Picking Your Patients: Selective Admissions in the Nursing Home Industry." Working Paper.

Gandhi, Ashvin, Huizi Yu, and David C. Grabowski. 2021. "High Nursing Staff Turnover in Nursing Homes Offers Important Quality Information." *Health Affairs*, 40(3).

Gandhi, Ashvin, YoungJun Song, and Prabhava Upadrashta. 2021. "Private Equiity, Consumers, and Competition." Working Paper.

Garin, Andy. 2019. "Public Investment and the Spread of 'Good-paying' Manufacturing Jobs: Evidence from World War II's Big Plants." Working Paper.

Gaynor, Martin, Kate Ho, and Robert J. Town. 2015. "The Industrial Organization of Health-Care Markets." *Journal of Economic Literature*, 53: 235–284.

Genesove, David, and Wallace Mullin. 2001. "Rules, Communication, and Collusion: Narrative Evidence from the Sugar Institute Case." *American Economic Review*, 91(3): 379–398.

Goldin, Claudia. 1991. "The Role of World War II in the Rise of Women's Employment." *American Economic Review*, 81(4): 741–756.

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): 782–818.

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

Graham, John R., Hyunseob Kim, Si Li, and Jiaping Qiu. 2019. "Employee costs of corporate bankruptcy." Working Paper 25922.

Gronberg, Timothy J., and W. Robert Reed. 1994. "Estimating Workers' Marginal Willingness to Pay for Job Attributes Using Duration Data." *The Journal of Human Resources*, 29(3): 911–931.

Gupta, Atul, Sabrina T. Howell, Constantine Yannelis, and Abhinav Gupta. 2021. "Does Private Equity Investment in Healthcare Benefit Patients? Evidence from Nursing Homes." Working Paper.

Hackmann, Martin B. 2019. "Incentivizing Better Quality of Care: The Role of Medicaid and Competition in the Nursing Home Industry." *The American Economic Review*, 109(5): 1684–1716.

Hackmann, Martin B., and R. Vincent Pohl. 2018. "Patient vs. Provider Incentives in Long Term Care." Working Paper 25178.

Haines, Michael R., and ICPSR. 2010. "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002."

Halm, Margo. 2019. "The Influence of Appropriate Staffing and Healthy Work Environments on Patient and Nurse Outcomes." *American Journal of Critical Care*, 28(2): 152–156.

Harrington, Charlene, Brian Olney, Helen Carrillo, and Taewoon Kang. 2012. "Nurse Staffing and Deficiencies in the Largest For Profit Nursing Home Chains and Chains Owned by Private Equity Companies." *Health Services Research*, 47(1): 106–128.

Harrington, Charlene, Clarilee Hauser, Brian Olney, and Pauline Rosenau. 2011. "Ownership, Financing, and Management Strategies of the Ten Largest For-Proft Nursing Home Chains in the United States." *International Journal of Health Services*, 41(4): 725–746.

Harrington, Charlene, John F. Schnelle, Margaret McGregor, and Sandra F. Simmons. 2016. "The Need for Higher Minimum Staffing Standards in U.S. Nursing Homes." *Health Services Insights*, 9: 13–19.

Harrington, Joseph E. 2005. "Detecting Cartels." 526.

Harrington, Joseph E. 2011. "Posted Pricing as a Plus Factor." *Journal of Competition Law & Economics*, 7: 1–35.

Harrington, Joseph E. 2016. "Heterogeneous Firms can Always Collude on a Minimum Price." *Economics Letters*, 138: 46–49.

Harrington, Joseph E. 2017. "A Theory of Collusion with Partial Mutual Understanding." *Research in Economics*, 71: 140–158.

Harrington, Joseph E. 2018. "Lectures on Collusive Practices."

Head, Keith, and Thierry Mayer. 2014. "Gravity Equations: Workhorse, Toolkit, and Cookbook." In *Handbook of International Economics*. Vol. 4, , ed. G. Gopinath, E. Helpman and K. Rogoff, Chapter 3, 131–195. Elsevier.

He, Alex Xi, and Daniel le Maire. 2020. "Mergers and Managers: Manager-Specific Wage Premiums and Rent Extraction in M&As." Working Paper.

Hjort, Jonas, Xuan Li, and Heather Sarsons. 2020. "Across-Country Wage Compression in Multinationals." Working Paper.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica*, 87(5): 1439–1474.

Huang, Jingyi. 2020. "Monopsony, Cartels, and Market Manipulation: Evidence from the U.S. Meatpacking Industry, 1903-1918."

Jaworski, Taylor. 2017. "World War II and the Industrialization of the American South." *The Journal of Economic History*, 77(4): 1048–1082.

Kawai, Kei, and Jun Nakabayashi. 2020. "Detecting Large-Scale Collusion in Procurement Auctions." Working Paper.

Kesselman, Amy. 1990. *Fleeting Opportunities: Women Shipyard Workers in Portland and Vancouver During World War II and Reconversion.* New York:State University of New York Press.

Knittel, Christopher R., and Victor Stango. 2003. "Price Ceilings as Focal Points for Tacit Collusion: Evidence from Credit Cards." *The American Economic Review*, 93: 1703–1729.

Krueger, Alan. 2017. "The Rigged Labor Market."

Lafotune, Julien, Jesse Rothstein, and Diane W. Schanzenbach. 2018. "School Finance Reform and the Distribution of Student Achievement." *American Economic Journal: Applied Economics*, 10(2): 1–26.

Lang, Kevin, and Ariella Kahn-Lang Spitzer. 2020. "Race Discrimination: An Economic Perspective." *Journal of Economic Perspectives*, 34(2): 68–89.

Lewis, Robert. 2007. "World War II Manufacturing and the Postwar Southern Economy." *Journal of Southern History*, 73(4): 837–866.

Lichtenberg, Frank R., and Donald Siegel. 1990. "The Effect of Ownership Changes on the Employment and Wages of Central Office and Other Personnel." *The Journal of Law & Economics*, 33(2): 383–408.

Li, Xiaoyang. 2012. "Workers, Unions, and Takeovers." *J Labor Res*, 33: 443–460.

Li, Zhimin, and Dmitri Koustas. 2019. "The Long-run Effects of Government Spending on Structural Change: Evidence from Second World War Defense Contracts." *Economics Letters*, 178: 66–69.

Loomer, Lacey, David C. Grabowski, Huizi Yu, and Ashvin Gandhi. 2021. "Association between Nursing Home Staff Turnover and Infection Control Citations." Working Paper.

Loos, Shoshana. 2005. "Women and Unions During World War Two: How Social Climate Affected Women's Labor Participation in World War Two." Working Paper.

Maksimovic, Vojislav, and Gordon Phillips. 2001. "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and are there Efficiency Gains?" *Journal of Finance*, 56(6): 2019–2065.

Manning, Alan. 2003. *Monopsony in Motion.* New Jersey:Princeton University Press.

Margo, Robert. 1995. "Explaining Black-White Wage Convergence, 1940-1950." *ILR Review*, 48(3): 470–481.

Margo, Robert A. 1993. "Explaining Black-White Wage Convergence 1940-1950: The Role of the Great Compression." Working Paper 44.

McGuckin, Robert H., Sang V. Nguyen, and Arnold P. Reznek. 1998. "On Measuring the Impact of Ownership Change on Labor: Evidence from U.S. Food-Manufacturing Plant-Level Data." In *Labor Statistics Measurement Issues.* , ed. John Haltiwanger, Marilyn E. Manser and Robert Topel, 207–248. University of Chicago Press.

Miller, Amalia R., and Carmit Segal. 2012. "Does Temporary Affirmative Action Produce Persistent Effects? A Study of Black and Female Employment in Law Enforcement." *The Review of Economics and Statistics*, 94(4): 1107–1125.

Miller, Conrad. 2017. "The Persistent Effect of Temporary Affirmative Action." *American Economic Journal: Applie Economics*, 9(3): 152–190.

Miller, Nathan H., Gloria Sheu, and Matthew C. Weinberg. 2021. "Oligopolistic Price Leadership and Mergers: The United States Beer Industry." Working Paper.

Myers, Charles A., and Rupert MacLaurin. 1943. *The Movement of Factory Workers.* New York:Wiley and Sons.

Porter, Robert H., and J. Douglas Zona. 1999. "Ohio School Milk Markets: An Analysis of Bidding." *The RAND Journal of Economics*, 30: 263–288.

Posner, Richard A. 1968. "Oligopoly and the Antitrust Laws: A Suggested Approach." *Stanford Law Review*, 21: 1562–1606.

Prager, Elena, and Matt Schmitt. 2021. "Employer Consolidation and Wages: Evidence from Hospitals." *The American Economic Review*, 111(2): 397–427.

Reardon, Sean F. 2011. "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." In *Whither opportunity? Rising Inequality, Schools and Children's Life Chances.* , ed. Greg J. Duncan and Richard J. Murnane, Chapter 5, 91–116. Russell Sage Foundation.

Reardon, Sean F., Ericka S. Weathers, Erin M. Fahle, Heewon Jang, and Demetra Kalogrides. 2019. "Is Separate Still Unequal? New Evidence on School Segregation and Racial Academic Achievement Gaps." Working Paper 19-06.

Reynolds, Lloyd G. 1951. *The Structure of Labor Markets.* New York:Harper & Brothers.

Rhode, Paul W., James M. Snyder Jr., and Koleman Stumpf. 2018. "The Arsenal of Democracy: Production and Politics During WWII." *Journal of Public Economics*, 166: 145–161.

Rose, Evan. 2018. "The Rise and Fall of Female Labor Force Participation During World War II in the United States." *The Journal of Economic History*, 78(3): 1–39.

Rosen, Sherwin. 1986. "The Theory of Equalizing Differences." In *The Handbook of Labor Economics.* , ed. Orley Ashenfelter, 641–692. Elsevier.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. "IPUMS USA: Version 10.0 [dataset]."

Saez, Emmanuel, Benjamin Schoefer, and David Seim. 2019. "Hysteresis from Employer Subsidies." Working Paper 26391.

Schmidheiny, Kurt, and Sebastian Siegloch. 2019. "On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications." Discussion Paper 12079.

Schmieder, Johannes F., Till von Wachter, and Joerg Heining. 2019. "The Costs of Job Displacement over the Business Cycle and Its Sources: Evidence from Germany." Working Paper.

Schmitt, Matt. 2017. "Do Hospital Mergers Reduce Costs?" *Journal of Health Economics*, 52: 74–94.

Selective Service System, U.S. 1956. *Special Monographs of the Selective Service System.* Washington, D.C.:Government Printing Office.

Shapiro, Carl, and Joseph Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review*, 74(3): 433–444.

Shleifer, Andrei, and Lawrence H. Summers. 1988. "Breach of Trust in Hostile Takeoevers." In *Corporate Takeovers: Causes and Consequences.* , ed. Alan J. Auerbach, 33–57. NBER Books.

Siegel, Donald S., and Kenneth L. Simons. 2010. "Assessing the Effects of Mergers and Acquisitions on Firm Performance, Plant Productitivy, and Workers: New Evidence from Matched Employer-Employee Data." *Strategic Management Journal*, 31(8): 903–916.

Sitkoff, Harvard. 1971. "Racial Militancy and Interracial Violence in the Second World War." *The Journal of American History*, 58(3): 661–681.

Smigel, Erwin O. 1969. *The Wall Street Lawyer, Professional Organization Man?* Bloomington:Indiana University Press.

Smith, James, and Finis Welch. 1989. "Black Economic Progress After Myrdal." *Journal of Economic Literature*, 27(2): 519–564.

Sokolova, Anna, and Todd Sorensen. 2020. "Monopsony in Labor Markets:A Meta-Analysis."

Squillace, Marie R., Anita Bercovitz, Emily Rosenoff, and Robin Remsburg. 2008. "An Exploratory Study of Certified Nursing Assistants' Intent to Leave."

Stigler, George J. 1947. "The Kinky Oligopoly Demand Curve and Rigid Prices." *Journal of Political Economy*, 55(5): 432–449.

Stigler, George J. 1961. "The Economics of Information." *Journal of Political Economy*, 69(3): 213–225.

Sundstrom, William A. 1994. "The Color Line: Racial Norms and Discrimination in Urban Labor Markets, 1910-1950." *The Journal of Economic History*, 54(2): 382–396.

Taras, Daphne, and A. Gesser. 2003. "How New Lawyers Use E-Voice to Drive Firm Compensation: The "Greedy Associates" Phenomenon." *Journal of Labor Research*, 24(1): 9–29.

Taubman, Paul. 1989. "Role of Parental Income in Educational Attainment." *The American Economic Review*, 79(2): 57–61.

Troy, Leo. 1965. *Trade Union Membership, 1897-1962.* NBER.

Turner, Donald F. 1962. "The Definition of Agreement Under the Sherman Act: Conscious Parallelism and Refusals to Deal." *Harvard Law Review*, 75: 655–706.

Turner, Sarah E., and John Bound. 2003. "Closing the Gap or Widening the Divide: The Effects of the G.I. Bill and World War II on the Educational Outcomes of Black Americans." *Journal of Economic History*, 63(1): 145–177.

van Damme, Eric, and Sjaak Hurkens. 2004. "Endogenous Price Leadership." *Games and Economic Behavior*, 47: 404–420.

War Manpower Commission, U.S. Government. 1945. *The Labor Market.* Reports and Analysis Service.

Whatley, Warren. 1990. “Getting a Foot in the Door: Learning, State Dependence, and the Racial Integration of Firms.” *The Journal of Economic History*, 50(1): 43–66.

Whatley, Warren. 1991. “African American Strikebreaking from the Civil War to the New Deal.” *Social Science History*, 13.

Wright, Gavin. 1986. *Old South New South: Revolutions in the Southern Economy Since the Civil War.* Basic Books.

Wynn, Neil A. 1976. *The Afro American and the Second World War.* Holmes and Meier Publishers.