

UNIVERSITY OF CALIFORNIA
Los Angeles

Essays in Labor Economics

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Nicholas Anthony Carollo

2021

© Copyright by
Nicholas Anthony Carollo
2021

ABSTRACT OF THE DISSERTATION

Essays in Labor Economics

by

Nicholas Anthony Carollo

Doctor of Philosophy in Economics

University of California, Los Angeles, 2021

Professor Moshe Buchinsky, Chair

This dissertation contains three essays in labor economics with a focus on economic and institutional differences in regional labor markets. It separately explores the causes and consequences of two major trends in the United States - the declining geographic concentration of immigrant location choices and the increasing prevalence of state-level occupational licensing requirements.

Chapter one shows that the geographic concentration of the foreign-born population in the United States fell sharply between 1980 and 2010 as immigrants were increasingly drawn to areas with historically low migrant inflows. This trend was driven primarily by the changing location choices of new immigrant cohorts, though secondary migration has played a minor role as well. An analysis of the determinants of location choice across four decades suggests that immigrants remain highly responsive to local labor market conditions, but the traditionally strong pull of ethnic enclaves has diminished over time.

Chapter two describes the construction of a novel dataset that compiles over one hundred years of occupational licensing, certification, and registration requirements in all fifty states and the District of Columbia. The data are assembled through a comprehensive analysis

of numerous primary and secondary sources and currently identify major state and federal policy changes for 250 unique occupation categories. It is the first occupational licensing database to link each policy to both current statutes or administrative regulations, as well as to historical legislation covering the entire twentieth century. A comprehensive analysis of state session laws, in particular, allows me to observe the exact text of all legislative acts enacting, amending, or replacing statutes that reference specific occupations. Using the content of these laws, I record the enactment and effective dates of regulatory changes and several variables that characterize the type of regulation that was adopted. Relative to existing sources, my data offer a significantly longer time series, the ability to observe superseded legislation, and a more complete coding of legal prohibitions that differentiates between practice and title restrictions.

Chapter three studies the short- and long-run impact of occupational licensing on labor market outcomes in the United States using the data described in chapter two. I implement an event study design that exploits within-occupation variation in the timing of licensing statutes across states to trace out the dynamic response of earnings and employment to policy changes. I find consistent evidence across several independent employer and household surveys that the typical licensing statute adopted during the past half-century increased worker earnings, but had null or weakly positive effects on employment. Twenty-five years after licensing statutes were adopted, cumulative wage growth in treated state-occupation cells exceeded that of untreated controls by 4 to 7%. Over the same time period, my results rule out an average disemployment effect greater than -5%. The data show much larger decreases in employment, however, among occupations that have little potential to cause serious harm. In cases where the consumer protection rationale for licensing is more plausible, I find simultaneous increases in both earnings and employment following the adoption of licensing requirements.

The dissertation of Nicholas Anthony Carollo is approved.

Jennie Brand

Rodrigo Antunes Pinto

Dora L. Costa

Moshe Buchinsky, Committee Chair

University of California, Los Angeles

2021

To my wife and parents.

TABLE OF CONTENTS

1 The Declining Pull of Ethnic Enclaves and the Diffusion of Immigrant Location Choices: 1980-2010	1
1.1 Introduction	1
1.2 Data and Measurement	4
1.2.1 Data	4
1.2.2 Measuring Immigrant Concentration	5
1.3 Trends in Immigrant Concentration	6
1.3.1 Changes in Immigrant Composition	6
1.3.2 Changes in Immigrant Dissimilarity	7
1.3.3 Heterogeneity	9
1.4 Concentration Among New and Internal Migrants	11
1.4.1 Cohorts and Assimilation	11
1.4.2 Return Migration	13
1.4.3 Native Migration and Robustness	15
1.5 Where are Immigrants Moving?	16
1.5.1 Empirical Strategy	17
1.5.2 New Immigrants	20
1.5.3 Internal Migrants	21
1.6 Conclusion	23
1.7 Data Appendix	25

2	Historical Data on Professional and Occupational Credentialing Requirements in the United States	44
2.1	Introduction	44
2.2	Background	46
2.2.1	Regulatory Terminology	46
2.2.2	Institutional Details	47
2.2.3	Implications for Data Collection	49
2.3	Data Description	50
2.3.1	Historical Credentialing Laws	50
2.3.2	Comparison to Existing Datasets	52
2.4	Dataset Construction	54
2.4.1	Identifying Regulated Occupations	54
2.4.2	Current Regulatory Requirements	57
2.4.3	Historical Policy Changes	59
2.5	Data Validation Exercises	60
2.5.1	Professional Directories	61
2.5.2	Regulation and Worker Credential Attainment	63
2.6	Conclusion	69
3	The Impact of Occupational Licensing on Earnings and Employment: Evidence from State-Level Policy Changes	79
3.1	Introduction	79
3.2	Institutional Background	85
3.2.1	Occupational Regulation	85

3.2.2	Potential Impacts on the Labor Market	88
3.3	Data	89
3.3.1	Occupational Licensing Legislation	90
3.3.2	Earnings and Employment	91
3.4	Research Design	94
3.4.1	Identification	95
3.4.2	Treatment Measures	97
3.4.3	Event Study Specifications	99
3.5	Main Results	102
3.5.1	Evidence from the Occupational Employment Statistics	102
3.5.2	Impact of Licensing on Earnings	104
3.5.3	Impact of Licensing on Employment	107
3.5.4	Robustness to Treatment Aggregation	109
3.6	Heterogeneity and Mechanisms	110
3.6.1	Mandated Investment and Worker Sorting	110
3.6.2	Heterogeneous Impacts of Licensing Requirements	112
3.6.3	Additional Heterogeneity Results	114
3.7	Conclusion	117
3.8	Data Appendix	119
3.8.1	Consistent Occupation Codes	119
3.8.2	Processing of Wage and Employment Data	120

LIST OF FIGURES

1.1	Share Foreign-Born in Working Age Population by State	28
1.2	Geographic Aggregations	29
1.3	Changes in Dissimilarity by Education and Citizenship Status	30
1.4	Changes in Dissimilarity Among Mexican Immigrants	31
1.5	Counterfactual Change in Dissimilarity	32
1.6	Change in Immigrant and Native Population Share by City	33
2.1	Share of Workers in Universally Licensed Occupations Reporting Credential At- tainment by State (CPS 2015-2018)	71
2.2	Semiparametric Estimates of the Relationship between Credential Attainment Rates and Regulation Coverage (CPS 2016-2018)	72
2.3	Effect of Alternative Measures of Regulation Coverage on Credential Attainment	73
3.1	Share of Workers Licensed by Consequence of Error	123
3.2	Licensing, Certification, and Registration Events by Detailed Occupation	124
3.3	Hypothetical Outcome Paths and Dynamic Treatment Effects With Multiple Pol- icy Changes	125
3.4	Dynamic Response of Hourly Wages and Employment to Licensing Events (Oc- cupational Employment Statistics State Panel 1999-2018)	126
3.5	Pooled Case Study Estimates of Wage and Employment Effects (Occupational Employment Statistics State Panel 1999-2018)	127
3.6	Dynamic Response of Hourly Wages to Licensing Events (Current Population Survey State Panel 1983-2018)	128

3.7	Robustness of Hourly Wage Estimates to Alternative Sample Definitions (State-Occupation Aggregated CPS Panel 1983-2018)	129
3.8	Effect of Licensing Events on the Self-Employment Rate (Current Population Survey State Panel 1983-2018)	130
3.9	Dynamic Response of Employment to Licensing Events (Current Population Survey State Panel 1983-2018)	131
3.10	Robustness of Employment Estimates to Alternative Sample Definitions (State-Occupation Aggregated CPS Panel 1983-2018)	132
3.11	Dynamic Response of Average Worker Age to Licensing Events (Current Population Survey State Panel 1983-2018)	133
3.12	Cross-Sectional Differences in Earnings and Employment (Occupational Employment Statistics State Sample 2015)	134
3.13	Dynamic Response of Industry Employment to Licensing Events (Quarterly Census of Employment and Wages 1990-2018)	135
3.14	Wage and Occupation Imputation Rates in the American Community Survey, Census, and Current Population Survey	136
3.15	Distribution of Employment Share Estimates for Regulated Suboccupations (Occupational Employment Statistics 2000)	137

LIST OF TABLES

1.1	Compositional Changes in the Foreign-Born Population 1980-2010	34
1.2	Changes in Immigrant Dissimilarity by Country of Origin	35
1.3	Changes in Immigrant Dissimilarity within Historical Destination States	36
1.4	Correlates of Immigrant Dissimilarity 1980-2010	37
1.5	Changes in Dissimilarity Within and Across Arrival Cohorts	38
1.6	Changes in Dissimilarity Within Arrival Cohorts	39
1.7	Effect of Employment Growth and Enclaves on New Immigrants' Location Choices	40
1.8	Effect of Employment Growth and Enclaves on New Immigrants' Location by Decade	41
1.9	Effect of Employment Growth and Enclaves on Internal Migration	42
1.10	Effect of Employment Growth and Enclaves on Internal Migration by Decade . .	43
2.1	Examples of Occupation Title Harmonization and Crosswalk Construction	74
2.2	Comparing Coverage of the Session Laws Data to Alternative Sources	75
2.3	Assessing the Consistency of Regulation Dates with Alternative Sources	76
2.4	Estimates of Licensed or Certified Workers Covered by Regulation Data (CPS 2015-2018)	77
2.5	Effect of Regulation Coverage on Credential Attainment (CPS 2016-2018)	78
3.1	Source and Characteristics of Worker Credentials (Survey of Income and Program Participation 2008 Panel, Wave 13)	138
3.2	Summary Statistics for a Subset of Licensed Occupations with Recent Policy Changes (State Legislation and Occupational Employment Statistics as of 2015)	139

3.3	Political and Economic Factors Affecting the Timing of State Licensing (Occupational Licensing Legislation Adopted 1975-2010)	140
3.4	Dynamic Response of Average Hourly Wages to Licensing Events (State-Occupation Aggregated CPS Panel 1983-2018)	141
3.5	Robustness of Earnings Estimates to Alternative Income Measures (State-Occupation Aggregated Census 1980-2000 and ACS 2001-2017)	142
3.6	Dynamic Response of Occupational Employment to Licensing Events (State-Occupation Aggregated CPS Panel 1983-2018)	143
3.7	Dynamic Response of Occupational Employment to Licensing Events (State-Occupation Aggregated Census 1980-2000 and ACS 2001-2017)	144
3.8	Robustness of Preferred Earnings Estimates to Alternative Treatment Aggregation Decisions (State-Occupation Aggregated CPS Panel 1983-2018)	145
3.9	Robustness of Preferred Employment Estimates to Alternative Treatment Aggregation Decisions (State-Occupation Aggregated CPS Panel 1983-2018)	146
3.10	Dynamic Response of Average Educational Attainment to Licensing Events (State-Occupation Aggregated CPS Panel 1983-2018)	147
3.11	Impact of Licensing on Hourly Wages with Demographic and Compositional Controls (State-Occupation Aggregated CPS Panel 1983-2018)	148
3.12	Heterogeneous Impact of Licensing on Average Hourly Wages and Employment by Occupation Characteristics (State-Occupation Aggregated CPS Panel 1983-2018)	149
3.13	Response of Hourly Wage Distribution to Licensing Events (Occupational Employment Statistics State Panel 1999-2018)	150
3.14	Construction of Consistent Standard Occupational Classification Codes	151
3.15	Aggregations Extending the Occ1990dd Balanced Occupation Panel to Cover 2010 Census Codes	152

3.16 Impact of Licensing Events on Wage and Occupation Imputation Rate (State- Occupation Aggregated CPS Panel 1983-2018)	153
3.17 Impact of Licensing Events on Earnings and Occupation Imputation Rate (State- Occupation Aggregated Census 1980-2000 and ACS 2001-2017)	154

ACKNOWLEDGMENTS

I am deeply grateful to my advisor Moshe Buchinsky, whose constant support and encouragement saw me through the Ph.D. and whose feedback greatly improved these papers. I am also thankful for the support of my committee members Dora Costa, Rodrigo Pinto, and Jennie Brand as well as the many people who provided helpful comments and suggestions including Felipe Goncalves, Adriana Lleras-Muney, Lee Ohanian, Daniel Rees, Evan Soltas, Till von Wachter and seminar participants at UCLA, the Young Economists Symposium, and the Western Economic Association International graduate student workshop. I am particularly grateful for many conversations with my classmates, especially Elicor Cohen, Brett McCully, Jacob Berman, and Jingyi Huang. I would also like to thank Simon Dong and Katie Xu for their excellent assistance in compiling the occupational licensing data and the Institute for Humane Studies for financial support. I thank the California Center for Population Research for providing an outstanding research environment. Finally, I thank my family members for their support and encouragement.

VITA

2019	C.Phil. Economics University of California, Los Angeles
2014	MSc. Economics (with Distinction) London School of Economics and Political Science
2012	B.S. Economics and B.A. International Studies (Magna Cum Laude) University of Wyoming
2020	Teaching Associate Department of Economics, University of California, Los Angeles
2019	Teaching Assistant Department of Economics, University of California, Los Angeles
2017-2018	Research Assistant California Policy Lab, University of California, Los Angeles
2016-2017	Teaching Assistant Department of Economics, University of California, Los Angeles
2014-2015	Research Assistant John F. Kennedy School of Government, Harvard University

CHAPTER 1

The Declining Pull of Ethnic Enclaves and the Diffusion of Immigrant Location Choices: 1980-2010

1.1 Introduction

The majority of immigrants living in the United States have historically chosen to locate in a small number of states and large metropolitan areas. Traditional gateway cities such as New York, Los Angeles, and Chicago attracted large migrant inflows, while the foreign-born population of the South Atlantic and Midwestern states remained relatively low. In recent decades, however, there has been a significant increase in the diffusion of immigrants' location choices. Although the foreign-born population remains highly concentrated relative to natives, over the past thirty years, new arrivals increasingly chose to settle in non-traditional destinations. This diffusion in location choices has been especially pronounced among immigrants from Latin America, and has broad implications for the economics of immigration and regional labor markets.

This chapter begins by providing descriptive evidence on the evolution of immigrants' location choices. I show, first, that more recent immigrant groups tend to be more concentrated in all years than immigrants from countries with a longer history of migration to the United States. Within groups defined by country of origin, the college-educated tend to be less geographically concentrated than individuals with a high-school education or less. Over time, concentration has fallen broadly, but the magnitude of these changes are generally larger for less-educated workers. Finally, I show that both immigrants and natives have

been moving away from (and toward) the same set of cities, which suggests that even groups whose concentration has remained stable over time have also shifted away from the cities where they were over-represented relative to natives in the 1980s.

Next, I examine the potential mechanisms driving the diffusion of immigrant location choices. I show that geographic concentration has been declining most rapidly among newer cohorts of immigrants at the time they arrive in the United States. Importantly, this implies that concentration among the stock of immigrants in the U.S. has declined because new arrivals are more likely to locate in a *larger* set of cities rather than just a *different* set of cities than earlier migrants. After a cohort has arrived in the United States, concentration also declines as secondary migrants leave cities where immigrants are over-represented. I show that these within-cohort changes are likely driven by the accumulation of human capital over time. Finally, I find that both initial and secondary migrants respond strongly to economic conditions when making their location choices, but the traditional pull of immigrant enclaves has diminished over time.

To highlight the significance of the trends discussed in this chapter, [Figure 1.1](#) displays the share of immigrants in each state's working-age population over time. In 1980 immigrants made up more than 10% of the labor force in only five states. By 2010 however, this was true of twenty states, with the foreign-born share exceeding 20% in nine of them. At the same time, changes in the geographic distribution of immigrants led to disproportionate increases in the share foreign-born across the South and Midwest. Consequently, many states that historically received few immigrants now have sizable foreign-born populations, especially in cities such as Atlanta, Raleigh-Durham and Minneapolis.

Despite the potential impact that these changes in immigrant concentration may have had on the labor markets and demographic composition of cities, little research has explored the diffusion of immigrant location choices in detail. Traditionally, the literature has studied location choice from a static perspective, consistently finding that new immigrants are attracted to cities with favorable labor market conditions and existing ethnic enclaves ([Bartel](#),

1989; Borjas, 2001; Zavodny, 1999). Less is known, however, about the dynamics of location choice among immigrant groups, or how the spatial distribution of arrival cohorts changes over time. Funkhouser (2000) examines changes in the distribution of immigrants between 1980 and 1990, though as I will argue, this misses the largest declines in concentration that took place after 1990. Jaeger (2008) presents a longer time series analysis, but does not explore the mechanisms driving trends in immigrant concentration. Others, such as Cutler et al. (1999) have focused on trends in segregation within rather than across cities.

A small literature has focused on the diffusion of Mexican immigrants in particular. Card and Lewis (2007) and Cadena and Kovak (2016) focus on the responsiveness of Mexicans to labor market changes between 1990 and 2000 and during the Great Recession, respectively. Others, such as Bohn and Pugatch (2015) and Watson (2013) have argued that immigration policies and border enforcement reduced the share of Mexicans moving to California and Texas. While this may have contributed to the rapid decline in concentration among Mexican migrants, my results suggest that immigration enforcement is not entirely responsible. First, I show that the concentration of Mexican immigrants has been declining *within* California and Texas as well. If border enforcement increases the fixed cost of migrating to these states, it is unclear why it would effect location choices within states. Second, concentration has been declining, though not as rapidly, among many immigrant groups who would be unaffected by enforcement on the southern border, including Puerto Ricans, who as American citizens, face no formal immigration restrictions.

The remainder of the chapter proceeds as follows. Section 1.2 describes the data used in the analysis and defines a measure of immigrant concentration. Section 1.3 presents descriptive statistics on changes in the concentration of immigrant groups between 1980 and 2010. Section 1.4 discusses whether these changes can be attributed to the location choices of new immigrants, internal migrants, return migrants, or the movement of natives. Section 1.5 examines location choices at the city and state level for new and internal migrants. Section 1.6 concludes.

1.2 Data and Measurement

1.2.1 Data

The primary data sources for this paper are the 5% samples of the 1980, 1990, and 2000 decennial censuses and the 2010 five year pooled American Community Survey sample obtained from IPUMS ([Ruggles et al., 2019](#)). Unless otherwise noted, the sample is limited to adults 18 to 64 years old, who are not living in group quarters or attending school. Immigrants are defined as individuals reporting a birthplace other than the United States who were not born abroad to American parents. For my main analysis I restrict the sample to the 99 immigrant groups with over 20,000 individuals living in the U.S. as of the 2000 census who are from a country that is consistently identifiable in all years between 1980 and 2010¹. Migrants from the American territories of Puerto Rico, American Samoa, and Guam are included in this definition for the purpose of this analysis. All other individuals who do not meet the definition of immigrants are considered natives. This necessarily includes the children of immigrants who were born in the United States since it is not possible to identify second-generation immigrants in the census microdata after 1970.

Measuring how immigrant location choices respond to changes in labor market conditions requires employment estimates for intercensal years. These come from the County Business Patterns (CBP) series collected by the Census Bureau. The CBP reports annual mid-March employment at the industry level for all counties in the United States. To ensure comparability with the industry classifications used in the IPUMS data, I base my industry definitions on 1987 Standard Industrial Classification (SIC) codes. Several industries in sectors such as construction and wholesale trade are combined to avoid issues created by the reclassification of more detailed industry codes over time. Unfortunately, the CBP does not report employment for agriculture, private households, public administration, rail transportation, or the

¹In some cases, countries must be combined for comparability over time. For example, the independent countries formed by the breakup of the Soviet Union and Yugoslavia have been combined.

U.S. Postal Service. I therefore supplement the CBP data with employment estimates for agriculture and government from the Bureau of Economic Analysis, which are also reported annually at the county level. The BEA data do not include employment estimates at the county level for the remaining industries missing from the CBP, and they are therefore omitted from my analysis. Additional details on the harmonization of industry codes for this paper are provided in [Section 1.7](#).

1.2.2 Measuring Immigrant Concentration

Immigrant concentration is measured using the following dissimilarity index,

$$D_{jt} = \frac{1}{2} \sum_{s=1}^S \left| \frac{F_{jst}}{F_{jt}} - \frac{N_{st}}{N_t} \right| \quad (1.1)$$

where F denotes foreign-born, N natives, j country of origin among the foreign-born, t year, and s a geographic area such as a state or city. The index is sometimes referred to as the Duncan index after [Duncan and Duncan \(1955\)](#) and measures the difference between two groups' spatial distribution. Each term in the summation is the difference between the share of immigrants and the share of natives living in location s . The sum has been divided by two to account for the fact that over-representation of a group in one location necessarily implies under-representation in another. This index has previously been applied to measure the concentration of immigrants in enclave cities by [Funkhouser \(2000\)](#) and residential segregation by [Cutler et al. \(1999, 2008\)](#), among others. Relative to alternative measures of concentration such as the Herfindahl index, D_{jt} has an intuitive interpretation as the share of immigrants who would have to move from an area where they are over-represented to an area where they are under-represented to match the spatial distribution of natives. The index attains its upper bound, $D_{jt} = 1$ when no immigrants and natives live in the same location, and its lower bound $D_{jt} = 0$ when the immigrant population is exactly proportional to the native population in every location.

Measuring immigrant concentration using the Duncan index requires geographic units that are consistent over time to avoid spurious changes in concentration driven by changes in geographic boundaries. Although states would clearly meet this requirement, a state-level analysis of location choice has two drawbacks in this context. First, immigrant locations have shifted both across and within states. Limiting the discussion to changes in concentration at the state-level ignores important changes in location choice *within* large immigrant states such as California, Florida, and Texas. Second, most states are too large to be considered homogeneous labor markets, so using states may mask important variation in local labor market conditions.

For these reasons, it is standard practice in the literature to use Metropolitan Statistical Areas (MSAs) as the unit of analysis. However, MSA boundaries are continuously updated to reflect population changes and commuting patterns, which means that they are not geographically consistent. Fortunately, IPUMS defines an alternative set of 543 geographic units called consistent public use microdata areas or conspumas, which are directly comparable in all microdata samples between 1980 and 2010. These areas are the smallest level of consistently identifiable geography in the public-use samples (Ruggles et al., 2019). In many cases, conspumas nest within MSAs. New York City, for example, contains eight. As described in [Section 1.7](#), I aggregate these sub-MSA conspumas to construct a geographic unit that is both constant over time and approximates an MSA. This results in 379 regions covering the entire land area of the United States, which for simplicity, I refer to as cities in the remainder of the paper. See [Figure 1.2](#) for a map of these areas.

1.3 Trends in Immigrant Concentration

1.3.1 Changes in Immigrant Composition

Before turning to this paper’s main focus on the changing spatial distribution of immigrants in the United States, it is useful to first consider how the composition of the foreign-born

population has changed over the sample period. Summary statistics for the 25 largest immigrant groups in the United States as of the 2000 Census are reported in [Table 1.1](#). The first two columns display the size of each group as a share of the total foreign-born population in 1980 and 2010. Observations are sorted by growth in this share (reported in Column 3), which shows that the composition of sending countries has shifted away from Western Europe and toward Latin America and Asia. The most striking change has been the growth in immigration from Mexico, with the share of Mexican immigrants in the foreign-born population nearly doubling from 16 percent in 1980 to 30 percent in 2010.

The remaining columns of [Table 1.1](#) characterize the distribution of years spent in the United States. Overall the average length of time since the typical immigrant arrived in the U.S. has increased. In 1980, 22% of immigrants arrived within the previous five years, but by 2010 this share had fallen to 9%, while the share of immigrants arriving more than 20 years ago rose by 12 percentage points. Although some of this effect can be attributed to a fall in immigration rates during the Great Recession, the share arriving in the previous five to ten years fell among the majority of immigrant groups as well. Although I do not report statistics on demographic characteristics by group, previous literature has shown that the changing composition of sending countries has also led to a shift in the immigrant population toward individuals who are, on average, less educated and less likely to speak English than immigrants arriving prior to 1980. See [Borjas \(2015\)](#) for a discussion of changes in immigrant composition and assimilation over time.

1.3.2 Changes in Immigrant Dissimilarity

The main patterns that this chapter explores are presented in [Table 1.2](#), which displays the dissimilarity index over time for large immigrant groups using cities as the unit of analysis. The first row shows that among the entire foreign-born population concentration fell by five percentage points, or just over 10%, a change which would have been even larger if not for an increase in concentration between 1980 and 1990. While the magnitude of this decline

is not trivial, it is somewhat misleading for two reasons. First, the index measures only changes in the distribution of the *stock* of immigrants. [Section 1.4](#) shows that concentration among the *flow* of new migrants has fallen much more. Second, since immigrants from different countries tend to settle in different cities, the dissimilarity index for all foreign-born cannot be represented as a weighted average of dissimilarity for individual nationalities. The aggregate index therefore masks significant heterogeneity in changes within specific groups. For example, the decline in concentration among Mexican immigrants was driven largely by a shift from immigrant dense cities in California and Texas to immigrant dense cities on the East Coast, and therefore contributes relatively little to changes in overall immigrant concentration.

For this reason, this chapter focuses on patterns of geographic concentration within nationalities over time. As shown in [Table 1.2](#), both the level and the change in dissimilarity varies substantially by country of origin. First, immigrants from Latin America and Asia tend to be more concentrated than immigrants from countries with a longer history of immigration to the U.S. such as the United Kingdom, Germany, and Canada. Dissimilarity has declined among the majority of large immigrant groups, though it has fallen most rapidly among groups that were highly concentrated at the beginning of the sample, and in particular for Latin Americans. Recalling that the dissimilarity index can be interpreted as the share of the group that would have to move to match the spatial distribution of natives, the 23 percentage point decline in dissimilarity among Mexican immigrants represents a substantial shift in their geographic distribution. Among the minority of nationalities which experienced an increase in concentration, the magnitudes of these changes tend to be small with the exception of the 8 percentage point increase among immigrants from Taiwan.

Are the changes in concentration documented in [Table 1.2](#) driven entirely by cross-state movements, or have location choices changed within states as well? [Table 1.3](#) addresses this question by computing dissimilarity for the largest immigrant groups within the six states that historically attracted the largest number of immigrants, California, Florida, Illinois,

New Jersey, New York, and Texas. Concentration among all foreign-born remained roughly constant in New York and New Jersey and fell in the remaining states. Although there have been large increases in concentration among Chinese and Vietnamese immigrants in California, Indian immigrants in New Jersey, and Russians in New York, within-state concentration has generally fallen. Importantly, the concentration of Mexican immigrants has declined in every state which suggests that the diffusion of Mexican immigrants is not entirely driven by the state-level immigration and border-enforcement policies studied by [Bohn and Pugatch \(2015\)](#) and [Watson \(2013\)](#).

1.3.3 Heterogeneity

Having shown the variation in dissimilarity across different immigrant groups, this section explores sources of heterogeneity in immigrant concentration by examining the correlates of dissimilarity across and within groups. Expanding the sample to include the 99 largest immigrant groups in the U.S. as of 2000, I estimate OLS regressions of the form,

$$D_{jkt} = \alpha_t + X'_{jkt}\beta + \gamma_j + \delta_k + \epsilon_{jkt}. \quad (1.2)$$

An observation in this regression is a birthplace-skill-year cell, where k denotes skill level. The dependent variable D_{jkt} is the dissimilarity index which has been computed separately for immigrants with a high-school education or less and some college or more, hereafter referred to as unskilled and skilled workers. When computing the education-specific index, immigrants are compared to the distribution of natives with the same level of education. The explanatory variables in the regression include fixed effects for time, birthplace, and skill, as well as a vector of demographic characteristics X . All regressions are weighted by the size of the birthplace-skill cell and report standard errors clustered by country of origin.

The results of this regression are reported in [Table 1.4](#). Column 1 includes only year dummies to display the average change in dissimilarity over time. Concentration rose from

1980 to 1990, and fell afterward, though these changes are not statistically significant. Column 2 adds birthplace fixed effects to the regression, sweeping out between-group variation to focus on changes within groups over time. Consistent with the tabulations in [Table 1.2](#), the decline in dissimilarity is large and statistically significant at the 5% level. Column 3 omits the birthplace fixed effects, but adds demographic characteristics. This shows that college-educated immigrants are, on average, 7 percentage points less concentrated than immigrants with a high-school education or less. Concentration is also decreasing in the size of the group and the share who speak English fluently, while groups with a larger share of unmarried individuals or American citizens tend to be more concentrated. Column 4 reports the full regression including birthplace fixed effects. The most noticeable change is that controlling for changes in composition reduces the coefficients on the year dummies substantially, indicating that much of the observed decline in dissimilarity within groups can be explained by compositional changes.

The final two columns estimate the regression separately for skilled and unskilled immigrants to allow all coefficients to vary by educational attainment. The coefficients on the year dummies show that concentration would have remained unchanged for the unskilled and would have risen for the skilled if not for changes in the demographic composition of immigrant groups over time. Changes in population size have a particularly large effect among unskilled immigrants. The implied elasticity of dissimilarity with respect to population evaluated at the mean is approximately 0.25, so a 10% increase in population implies a 2.5% decline in the dissimilarity index. Given the increasing rates of immigration over the sample period, this is a significant effect which may suggest the presence of crowding out effects in the location choices of unskilled immigrants. These columns also show that even within ethnic groups, correlations between demographic composition and dissimilarity vary substantially by skill level. It is natural then to ask whether immigrants of different skill levels from the same country experienced similar changes in concentration, or whether skill groups behave independently of each other. To answer this, [Figure 1.3](#) plots the change in

dissimilarity between 1980 and 2010 by country of origin, skill level, and citizenship status, which shows that changes are highly correlated within immigrant groups.

This section has shown that geographic concentration of immigrants relative to the native population has fallen among the majority of immigrant groups between 1980 and 2010, and that Latin Americans in particular experienced significant diffusion. Although the results reported so far suggest that these declines are related to changes in the demographic composition within different immigrant groups over time, the next section explores in more detail the mechanisms through which immigrant dissimilarity has changed by separating changes among the stock and the flow of immigrants.

1.4 Concentration Among New and Internal Migrants

Ignoring the potential effects of aging into or out of the sample between census years, there are four ways that the concentration index can decline. First, new immigrants may choose to locate in cities where the stock of immigrants already in the United States is under-represented. Second, after settling in the U.S., secondary migrants may move out of the immigrant-dense gateway cities. Third, out-migration can lower concentration if return migrants are disproportionately drawn from cities with high immigrant shares. And fourth, even if the spatial distribution of immigrants remains unchanged, concentration may fall if natives move into cities where immigrants are over-represented.

1.4.1 Cohorts and Assimilation

To separate the changes in dissimilarity driven by new cohorts from changes occurring within cohorts as time spent in the United States increases, I modify the baseline regression as follows,

$$D_{jkct} = Z'_{jkct}\theta + X'_{jkct}\beta + \gamma_j + \psi_c + \epsilon_{ijkct} \quad (1.3)$$

where c denotes arrival cohort and the dissimilarity index is now calculated separately for each birthplace-education-cohort cell. As before, the regression contains birthplace fixed effects, γ_j , and demographic controls X_{jkct} , but now includes cohort fixed effects ψ_c which capture changes in average dissimilarity across cohorts, and a set of indicators Z_{jkct} for time (in five-year bins) that the cohort has spent in the United States. All regressions are weighted by the population of the birthplace-education-cohort-year cell and have standard errors clustered by birthplace.

The results of this regression are presented in [Table 1.5](#). Column 1 includes only time spent in the United States and the cohort fixed effects, where the omitted category is immigrants arriving in 1970 to 1975. Concentration has declined significantly among newer cohorts at the time they enter the United States, and falls over time within cohorts. Column 2 adds fixed effects for country of origin without significantly affecting these trends. Next, Columns 3 and 5 repeat this specification separately for immigrants with and without a college education to allow all coefficients to vary by educational attainment. This shows that the decline in concentration across cohorts has been nearly twice as large among less educated immigrants, who also experience a more rapid decline in concentration within cohorts. Recall from [Section 1.3](#), however that the college educated are less concentrated to begin with. Columns 4 and 6 add additional demographic controls which reduces the size of the coefficients slightly, implying that some of these changes are related to the composition of immigrant cohorts, though overall the trends remain highly significant. Unreported regressions confirm that these results are robust to dropping Mexican and Central American immigrants from the sample, though this further reduces the magnitude of the coefficients by approximately one-third.

[Table 1.6](#) provides additional evidence on why concentration may decline within cohorts over time by including separate fixed effects for each cohort-birthplace pair. Variation therefore comes from changes over time within groups of immigrants from the same country of origin who arrived in the United States at approximately the same time. These fixed effects

absorb all time-invariant characteristics of the cohort such as demographic composition, provided that there is not significant selective attrition. Columns 1,3 and 5 show that the coefficients change little from the specification reported in [Table 1.5](#). Columns 2,4, and 6 add controls for the share of the cohort who report speaking English very well or who have attained U.S. citizenship, both of which tend to increase within cohorts over time. English language ability has a highly significant effect, which is much larger for less-educated immigrants. Interestingly, controlling for English language ability also reduces the coefficients on time spent in the United States to essentially zero, suggesting that human capital acquisition may be associated with lower concentration. This relationship, however, is not necessarily causal since individuals moving to cities with fewer immigrants may have a greater incentive to learn English.

1.4.2 Return Migration

[Table 1.5](#) and [Table 1.6](#) showed that within cohorts, concentration declines as time spent in the United States increases. However, this is not necessarily driven by secondary migration away from the cities where these cohorts were over-represented when they first arrived. Rather, the trends discussed above could also be explained by return migration if immigrants who return to their home countries are more concentrated when they arrive than those who plan to remain in the United States permanently. Unfortunately providing direct evidence on how return migration affects immigrant concentration is difficult since few contemporary datasets record return migration and fewer still provide information on where return migrants lived before they left the United States. I therefore turn to indirect evidence on the relationship between return migration and immigrant concentration using data from the Current Population Survey (CPS)².

After 1994 immigrants in the Annual Social and Economic Supplement to the CPS can

²The CPS samples are also obtained through IPUMS ([Flood et al., 2018](#)).

be identified by country of origin and approximate year of immigration. Since the ASEC is considerably smaller than the 5% samples of the Decennial Census, in this section I focus only on immigrants from Mexico, who in addition to being the largest immigrant group in the United States, have relatively high rates of return migration (Borjas and Bratsberg, 1996). Figure 1.4 plots the time series of dissimilarity, computed at the state level, for Mexican immigrants arriving in 1965-1984, 1985-1989, 1990-1995, and 1996-2001. Shaded areas indicate the 90% confidence intervals for these estimates³. Because the CPS reports year of immigration in intervals which vary in length by year, dissimilarity is only reported for each cohort once all of its members are consistently observable, typically between three and eight years after their arrival in the United States. The patterns in the figure are consistent with the regressions discussed in Section 1.2. Concentration declines over time within cohort, and has fallen at the time of entry among more recent arrivals.

If return migration is responsible for the decline in concentration within cohorts, it must be the case that the immigrants who remained in the U.S. were less concentrated than those who left. Although we cannot observe return migrants, using five-year migration status recorded by the CPS in 1995, 2005, and 2015, we can compute a separate dissimilarity index for immigrants observed during year t , who lived in the same state in $t - 5$ ⁴. If these non-movers are significantly less concentrated than the cohort was in $t - 5$, it suggests that concentration fell through selective attrition. If instead, non-movers are more concentrated than the cohort as a whole at time t , it implies that concentration fell through secondary migration. The estimates of concentration among non-movers are shown as markers in Figure 1.4 and support the latter case. With the possible exception of the two oldest cohorts

³Since the distribution of the dissimilarity index is unknown, the confidence intervals are constructed by bootstrapping. This involves re-sampling (with replacement) data from each year of the CPS for 1000 repetitions and calculating the dissimilarity index for each iteration. Confidence intervals are based on the distribution of these estimates. This method is not feasible in the Census datasets due to the much larger sample sizes.

⁴Although the CPS also contains one-year migration status, too few people move in any year to differentiate between the effect of internal migration and noise from the changing sample. This exercise is not repeated in the census data since the actual distribution of cohorts in intercensal years is not known.

in 2015, concentration is indeed consistently higher among non-movers and quite similar to concentration five years prior in the years when it is observable.

Of course, these results do not rule out that return migration may still have some role in reducing immigrant concentration, especially since we are only able to observe a small number of data points for a single immigrant group. [Dustmann and Görlach \(2016\)](#), for example, find that out-migration rates are highest during the first decade a cohort spends in the host country. Since we can only observe cohorts four to eight years after they arrive in the U.S., the estimates presented in this section may miss the effects of early return migration, though this concern is mitigated in part by earlier results which found that concentration declines little during a cohort's first ten years in the United States. It is unlikely however, based on the evidence in this section, that the decline in concentration within cohorts is entirely driven by return migration. Internal migration therefore is one channel through which immigrant concentration has declined, implying that immigrants continue to re-optimize their location decisions after settling in the U.S., and when they do, tend to move to cities where their cohort was previously under-represented.

1.4.3 Native Migration and Robustness

An implicit assumption of the analysis so far has been that changes in dissimilarity are driven by the location choices of immigrants. At least in principle, this need not be the case since even if the distribution of immigrants remained constant, dissimilarity could fall if natives move into the cities where immigrants are over-represented. While it is unlikely that changes in the spatial distribution of natives are significant enough to explain the trends discussed in this paper, it is still interesting to ask how immigrant concentration would have changed had the share of immigrants in each city remained constant. To answer this question, I compute a counterfactual dissimilarity index which holds the distribution of immigrants fixed in 1980. Changes in this index are therefore driven entirely by the movement of natives.

[Figure 1.5](#) plots the change in the counterfactual index against the change in the true

dissimilarity index separately for college and non-college immigrants. Two features of this graph are noteworthy. First, if immigrant location choices did not change, concentration would have actually risen for all but two groups, Mexicans and Vietnamese without a college education. Since in reality concentration has been declining among most groups, this implies that both immigrants and natives have been moving away from cities which had high immigrant shares in 1980, an observation I return to in the next section. Second, most points lie above the 45 degree line, indicating that the counter-factual change in dissimilarity is greater than the actual change in dissimilarity. Thus, even many of the groups that became more concentrated between 1980 and 2010 have been moving away from the cities where they were over-represented in 1980.

1.5 Where are Immigrants Moving?

The previous section alluded to the fact that in order to explain the observed patterns in the concentration of immigrants relative to natives, it must have been the case that immigrants and natives moved away from (and toward) the same set of cities. This is confirmed in [Figure 1.6](#), which plots the change in the share of natives living in each city between 1980 and 2010 against the same measure for immigrants. Due in part to the much larger size of the native population, the magnitude of these changes are smaller among natives, but the measures are indeed highly correlated. Sunbelt cities such as Atlanta, Phoenix, and Dallas became much more popular among both groups, while the traditional immigrant gateway cities of New York, Los Angeles, and Chicago saw large declines in their population share. The fall in the share of immigrants locating in New York, at 7 percentage points percentage points, is so large that it is omitted from the graph in order not to obscure the patterns among the remaining cities⁵.

The fact that the immigrant and native populations have shifted in the same direction is

⁵New York is still included in the fitted regression line.

consistent with existing literature such as [Borjas \(2001\)](#) and [Cadena and Kovak \(2016\)](#) who argue that although the location choices of both immigrants and natives respond to changing labor market conditions, immigrants tend to be more responsive. If labor market conditions in cities such as Atlanta became more favorable than those in traditional immigrant cities, we would indeed expect concentration to fall among the stock of immigrants as new migrants clustered in a different set of cities than those already in the country. However, a key finding of this paper is that dissimilarity has fallen most rapidly among the initial location choices of more recent cohorts at the time they enter the United States. This distinction is important because it implies that new immigrants are not simply concentrating in a *different* set of cities than earlier immigrants, but that they are increasingly choosing to settle in a larger *number* of cities as well. With that in mind, this section addresses the factors that attracted immigrants to different cities, and whether the importance of labor market conditions and migrant networks have changed over time.

1.5.1 Empirical Strategy

Focusing first on the determinants of location choice among new immigrants, the basic regression reported in this section has the form

$$y_{jkst} = \alpha_t + \beta_1 \Delta \ln(Emp)_{kst} + \beta_2 Share_{jks,t-1} + \beta_3 Dist_{js} + \gamma_j + \delta_k + \epsilon_{jkst}. \quad (1.4)$$

I follow [Borjas \(2001\)](#) in defining the dependent variable as

$$y_{jkst} = \frac{F_{jkst}^{new} / F_{jkt}^{new}}{N_{kst} / N_{kt}} \quad (1.5)$$

where F_{jkst}^{new} refers to immigrants from country j with education level k living in city s who arrived in the United States within the last five years. This index, which Borjas refers

to as an index of relative supply, is defined as the share of new immigrants moving to a city, deflated by the share of natives who live in that city. Expressing the dependent variable in this way addresses two potential scale effects that prevent us from measuring the immigrant inflow in levels. First, the number of immigrants moving to any city is a function of the size of the immigrant group, so expressing the inflow as a share ensures that the dependent variable has a comparable scale for all groups, regardless of population size. Second, the share of immigrants moving to a city is highly correlated with the size of the city. Dividing by the share of the native population in the city controls for this second scale effect.

Further, y_{jkst} is closely related to the dissimilarity index used throughout this paper. In particular, if $y_{jkst} > 1$ it implies that new immigrants are overrepresented in city s relative to natives, while if $y_{jkst} < 1$ they are underrepresented. In fact, the dissimilarity index calculated for new immigrants of skill level k can be equivalently expressed as

$$D_{jkst}^{new} = \frac{1}{2} \sum_{s=1}^S \frac{N_{kst}}{N_{kt}} |y_{jkst} - 1| \quad (1.6)$$

or the sum of absolute deviations of y_{jkst} from one, weighted by the native population share of each city.

The explanatory variables in regression (4) are $\Delta \ln(Emp_{kst})$, the change in log skill-specific employment, $Share_{jks,t-1}$ the share of the immigrant group in the city's population in the previous decade, $Dist_{js}$, the distance between country j and city s ⁶, and fixed effects for year, birthplace, and skill group. Employment growth is an endogenous variable, which may be driven by supply shifts, including the arrival of the immigrants themselves. To isolate changes in labor demand, I follow the standard approach in the literature and instrument

⁶Measured as the Vincenty distance between the population-weighted centroid of country j and city s .

for employment changes using the shift-share, or [Bartik \(1991\)](#) instrument,

$$Bartik_{st} = \sum_{l=1}^L \theta_{s,l,t-1} \ln \left(\frac{E_{-s,l,t}}{E_{-s,l,t-1}} \right). \quad (1.7)$$

where $E_{-s,l,t}$ denotes *national* employment in industry l excluding city s from the calculation, and the weights $\theta_{s,l,t-1}$ are the share of total employment in industry l in city s at time $t - 1$. The idea of the instrument is to measure changes in labor demand by instrumenting for employment changes in a city using a weighted average of national industry-level employment growth, with weights corresponding to the city's industry composition. By excluding the city itself from the calculation of employment growth, the instrument is mechanically purged of the effect of supply shifts unique to the city. In general, the exclusion restriction is satisfied if employment growth at the national industry level is orthogonal to local labor supply changes.

Since immigrants of different skill levels work in different industries, I also create education-specific Bartik instruments for college and non-college workers. I do this by first calculating the fraction of college and non-college workers in each industry using the entire population of both immigrants and natives from the 1980 to 2010 Censuses and linearly interpolating these shares for intercensal years. Using these weights, I construct separate employment series for skilled and unskilled employment growth, and recompute the instrument. Although it would be ideal to include employment growth in the year before immigrants arrive, exact year of arrival is only observed in census data from 2000 onwards. For this reason ΔEmp_{kst} will be measured as the average employment growth between $t-5$ and t , as will be the corresponding instrument. For ease of interpretation, throughout this section I follow [Wozniak \(2010\)](#) by first standardizing the Bartik measure to have a mean zero and variance of one, and report the reduced-form regression from including it as an exogenous regressor in (4) rather than reporting the two-stage least squares coefficients. The interpretation of the coefficient on the Bartik measure is therefore the effect of a one-standard deviation increase in employment

relative to the national average of employment growth.

1.5.2 New Immigrants

Table 1.7 reports the results of regression (4) for new immigrants. Since the flow of new immigrants is smaller than the stock already in the United States, to ensure that there are not too many birthplace-skill-city cells with no immigrants in them, I limit the sample in this section to the 50 largest immigrant groups and 100 largest cities as of the 2000 Census. Odd numbered columns use the original Bartik instrument for total employment changes, while even numbered columns use the education-specific variable. Columns 1 and 2 include immigrants of all skill levels and education fixed effects to account for the greater dispersion of location choices among the college-educated. As expected, both the original and education-specific Bartik measures have a positive effect on the relative share of immigrants moving to a city, indicating that immigrants are indeed responsive to labor market conditions. Although the difference is not large, the coefficients suggest that new immigrants may be slightly more responsive to employment growth in general rather than growth in the specific industries that others of their skill level tend to be employed in.

Consistent with the previous literature, the presence of ethnic enclaves has a strong effect on immigrants' location choices. Increasing the share of a particular ethnic group in a city by 1 percentage point is associated with an increase in y_{jkst} by 1.2 units, or approximately half a standard deviation. Likewise, immigrants are more likely to settle in cities that are closer to their country of origin. The remaining columns report regressions run separately for skilled and unskilled immigrants. Here too we see that immigrants of both skill levels may be slightly more responsive to total employment growth rather than education-specific employment growth. Overall, however, immigrants with a high-school education or less are more responsive to labor market conditions than their college-educated co-nationals, though they are also more likely to settle closer to their home countries and are more influenced by the size of existing immigrant communities than high-skilled migrants.

Are these patterns in [Table 1.7](#) stable over time, or have there been any changes in the responsiveness of immigrants to variables influencing location choices that might also be related to declining immigrant concentration? To answer this question, [Table 1.8](#) reports the regression separately for the years 1990-2010, using the education-specific Bartik⁷. The most striking result is that among both college and non-college immigrants, the importance of ethnic enclaves seems to have declined over time. Specifically, the coefficients on the lagged population share of an immigrant group fell in magnitude by approximately 60% between 1990 and 2010 for both skilled and unskilled immigrants. Some of the diffusion in immigrant location choices may therefore be explained by this declining association between the size of an existing ethnic enclave in a city and the number of new immigrants moving to it. The coefficients on the Bartik instruments are similar in 1990 and 2000, though they decline and even switch sign for the college educated in 2010.

1.5.3 Internal Migrants

I take a different approach to internal migrants by first examining individual internal migration decisions. The specification reported in this section is a linear probability model of the form

$$\mathbb{I}(\text{move})_{ijkst} = \alpha_t + \beta_1 \Delta \ln(\text{Emp})_{kst} + \beta_2 \text{Share}_{jks,t-1} + X'_{ijkst} \theta + \gamma_j + \delta_k + \epsilon_{ijkst} \quad (1.8)$$

where i denotes an individual, s now denotes state⁸, and $\mathbb{I}(\text{move})_{ijkst}$ is an indicator for whether the individual has moved between states in the five years prior to observation. The covariates have the same definition as before, but are measured in the individual's state

⁷The 1980 Census is omitted from this sample because it is not possible to calculate the lagged immigrant shares at the city level in 1970.

⁸I am not able to identify the city where internal migrants lived 5 years ago at the level of geography I use in this paper.

of residence in $t - 5$ to identify the factors that lead immigrants to *leave* certain states. The time fixed effects control for declining rates of internal migration among all groups, the skill fixed effects for higher cross-state migration probabilities among the skilled, and the birthplace fixed effects for varying migration rates across ethnic groups. As before, I report the reduced-form coefficients from including the Bartik instrument directly. Since five year migration status is only recorded in the Census for 1980-2000, the 2010 ACS is omitted from the sample.

The results of this regression are reported in [Table 1.9](#). As before, odd numbered columns use the Bartik instrument for total employment changes, while even numbered columns use the education-specific variable. As we would expect, the sign on the Bartik instrument is negative, indicating that greater employment growth decreases the probability that an individual moves out of a state. The estimates in Columns 1 and 2 imply that a 1 standard deviation negative employment shock increases the probability of leaving a state by 1 percentage point relative to an unconditional mean migration rate of 9%. Columns 1 and 2 also show that immigrants are less likely to move out of states where their ethnic group is highly concentrated, though recall that earlier results have shown that conditional on migration, immigrants tend to move to less concentrated locations. The remaining rows show that English-speakers are more likely to move, citizens less so, and that the probability of migration declines by 1.2 percentage points for every 10 years an immigrant has spent in the U.S.. The results by education level in columns 4 to 6 show that, in contrast to the results for new migrants, college-educated immigrants are much more responsive to labor market conditions than less-skilled immigrants when moving within the United States. While this result is surprising, [Wozniak \(2010\)](#) reaches similar conclusions about the migration behavior of college and non-college educated natives.

To determine whether secondary migrants became more or less responsive to these variables over time, [Table 1.10](#) repeats the preferred specification using the education-specific Bartik variable for each Census year separately. As for new immigrants, we see that among

both college and non-college immigrants, the importance of ethnic enclaves seems to have declined over time. While immigrants in states with a larger share of their ethnic group are still less likely to move, the magnitude of this effect fell significantly. In 1980, a 1 percentage point larger share of a particular ethnic group in a state's population was associated with a 3.7 percentage point lower probability that a low-skilled individual from that group left the state. By 2000, this effect fell to 0.5 percentage points. Among skilled immigrants, the same effect fell from 3.9 percentage points to 1 percentage point. Over the same time period, the responsiveness of both education levels to employment shocks increased, though these changes are much smaller and are not statistically significant.

This analysis therefore reveals a consistent decline in the importance of the share of an immigrant group already in a state on the location choices of new and internal migrants of all skill levels. While this has not necessarily had a causal effect on declining immigrant concentration since unobserved changes in immigrant-dense cities may have induced immigrants to seek new locations, the consistency of these patterns does provide interesting evidence suggesting that the traditionally strong pull of ethnic enclaves has diminished over time.

1.6 Conclusion

This chapter has shown that in contrast to the conventional wisdom that the spatial distribution of the foreign-born in the United States has remained stable, immigrant concentration has in fact been declining on average since 1980. Latin Americans in particular experienced rapid geographic diffusion as their location choices shifted away from California and Texas and toward other states in the West and South which historically received few immigrants. These changes have largely been driven by the locations of new immigrants, who are increasingly likely to choose from a larger set of cities when they make their initial location choices than immigrants did in the past. Internal migrants, too, have contributed to these trends

as individuals re-optimize their location decisions as they spend more time in the U.S., and, possibly, as they accumulate additional human capital such as language skills.

The cities favored by these new and internal migrants tend to be those that have also attracted natives, while the traditional gateway cities of New York, Los Angeles, and Chicago have had declining population shares among both groups. An analysis of the determinants of immigrants' location choices shows that while immigrants continue to be attracted by favorable labor market conditions, as well as the presence of other members of their own group, the importance of ethnic enclaves in these decisions seems to have been declining over time among both new and internal migrants of all skill levels. While I stress that the evidence is not strong enough to interpret these results as a causal mechanism for declining immigrant concentration, they represent interesting correlations which, to my knowledge, have not yet been addressed in the literature. The trends in immigrant location choices identified in this chapter have potentially significant implications for the economics and politics of immigration in the United States, and merit further research on the causal mechanisms underlying them.

1.7 Data Appendix

Geography

As discussed in the text, my analysis requires a unit of geography which is comparable over time and more closely approximates labor markets than states. Fortunately, IPUMS maintains a set of 543 geographic units called consistent public use microdata areas or conspumas, which are directly comparable in all microdata samples between 1980 and 2010. These areas were created on the basis of 1980 county groups and represent the smallest level of consistently identifiable geography for the years that they are available ([Ruggles et al., 2019](#)). Since conspumas were constructed from county boundaries, over 97% of U.S. counties map to a unique conspuma. The remaining 3% contain multiple conspumas which divide the county into smaller units. Los Angeles County, for example, is composed of two conspumas. Whenever this is the case, I first aggregate the conspumas to the county level before performing any additional aggregations.

Due to confidentiality requirements, no level of geography containing fewer than 100,000 individuals can be observed in public-use microdata, and consequently conspumas can vary greatly in size. For example, the entire state of Wyoming is a single conspuma, while New York City alone is composed of eight separate conspumas. In addition, conspumas do not cross state borders, so they partition cities which cross state lines such as Washington, D.C. and Kansas City. To construct a geographic unit that more closely approximates a local labor market, I aggregate conspumas to the MSA level whenever possible. If a single MSA such as Washington, D.C. contains multiple unique conspumas, I aggregate them to a single unit. Many conspumas, however, overlap with two or more MSAs. When this is the case, I assign a conspuma to an MSA if over 50% of the conspuma's population lived in the MSA as of 1980. Otherwise, I keep the original boundaries of conspuma. After consolidating conspumas in this way, I am left with 379 consistent geographic units as shown in [Figure 1.2](#).

Employment

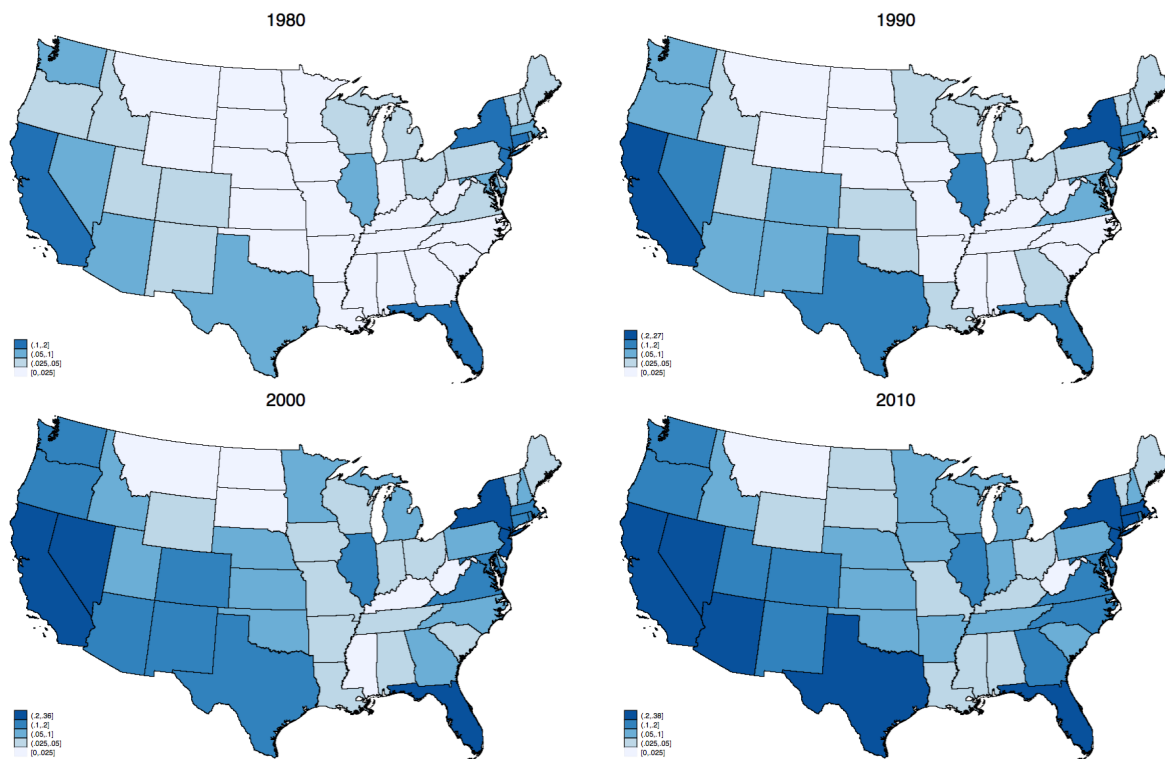
A significant issue that arises when considering employment growth at the industry level between 1980 and 2010 is that the industry definitions used by the census underwent several revisions during this time. The most significant of these changes was the transition from the Standard Industry Classification (SIC) system to the North American Industry Classification System (NAICS) between 1997 and 1998, although additional revisions to industry classifications took place in 1987, 2002, and 2007. The raw data in the County Business Patterns (CBP) therefore must be harmonized to be comparable over time. To address this issue, I choose two-digit 1987 SIC codes as the basis for my analysis and make use of a crosswalk constructed by David Dorn ([Dorn, 2009](#)) which provides weights indicating the share of employment in each six-digit 1997 NAICS industry that maps back to a four-digit 1987 SIC industry.

I begin by converting the 2007 and 2002 NAICS codes back to the 1997 NAICS basis. Several six-digit 2002 industries were combined in 2007. However, none of these changes affect employment estimates at the two-digit level, so I simply assign the 2007 codes to one of the 2002 codes they were aggregated from. More significant changes took place between 1997 and 2002. Five six-digit 2002 industries do not map to a unique 1997 code, and I was unable to find a weighted crosswalk with information on how to reallocate employment from these industries to the 1997 basis. Thus, I assign each of the five 2002 industries to the 1997 industry which contains the majority of the cell's employment. In four of the five cases this decision is inconsequential since the industries are in the construction and wholesale sectors, which will be further aggregated later. The only case where there may be concern that employment is incorrectly allocated is in the 2002 internet publishing and broadcasting industry, which I assign to the the 1997 publishing industry, although a fraction of its employment also maps to the 1997 information services industry.

Having standardized the 1997-2010 data to a 1997 NAICS basis, I can match to David

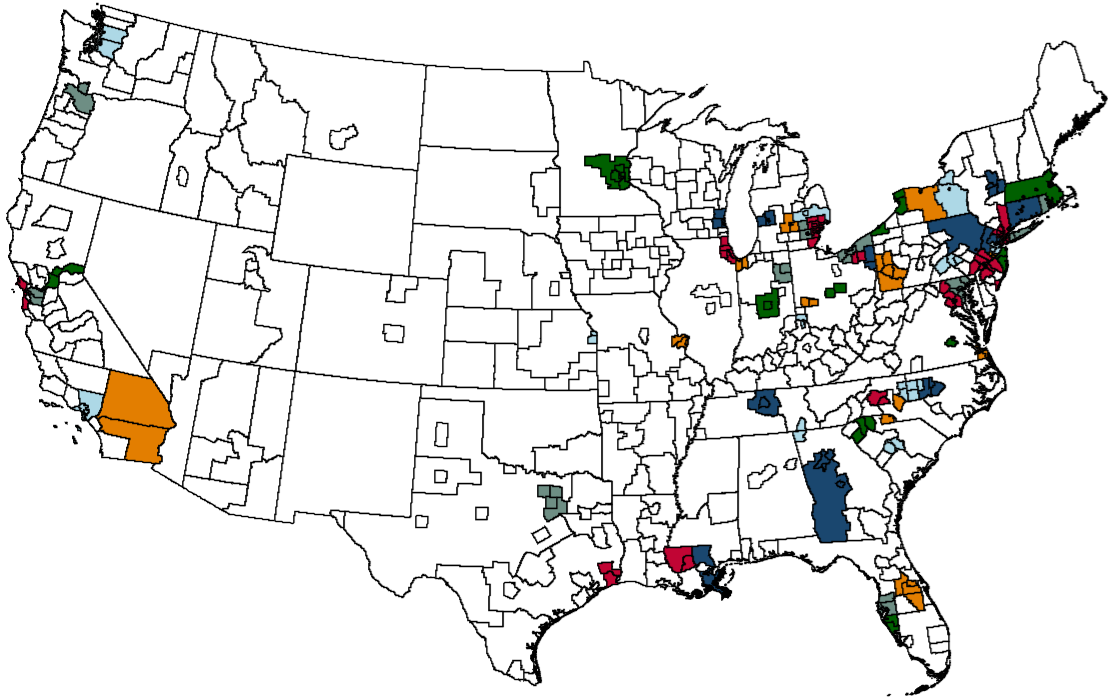
Dorn's crosswalk. Since my analysis uses two-digit industries rather than four-digit industries, I obtain the weights by merging the crosswalk with the 1998 CBP national employment estimates then calculating the share of employment in each three-digit NAICS industry that maps to each two-digit SIC industry. Finally, the SIC revisions in 1987 require additional aggregations of some two-digit SIC codes to ensure comparability with earlier employment figures and with the industry definitions used by IPUMS. For example, "Instruments and Related Products" is combined with "Electronic and Other Electric Equipment," and "Business Services", "Engineering and Management," and "Other Services" must be combined. After performing these aggregations, I am left with 60 industries which will be used in my analysis. As a final check, I plot the national time series of employment in each of these industries and verify that no discontinuities remain at the years when classification schemes changed.

Figure 1.1: Share Foreign-Born in Working Age Population by State



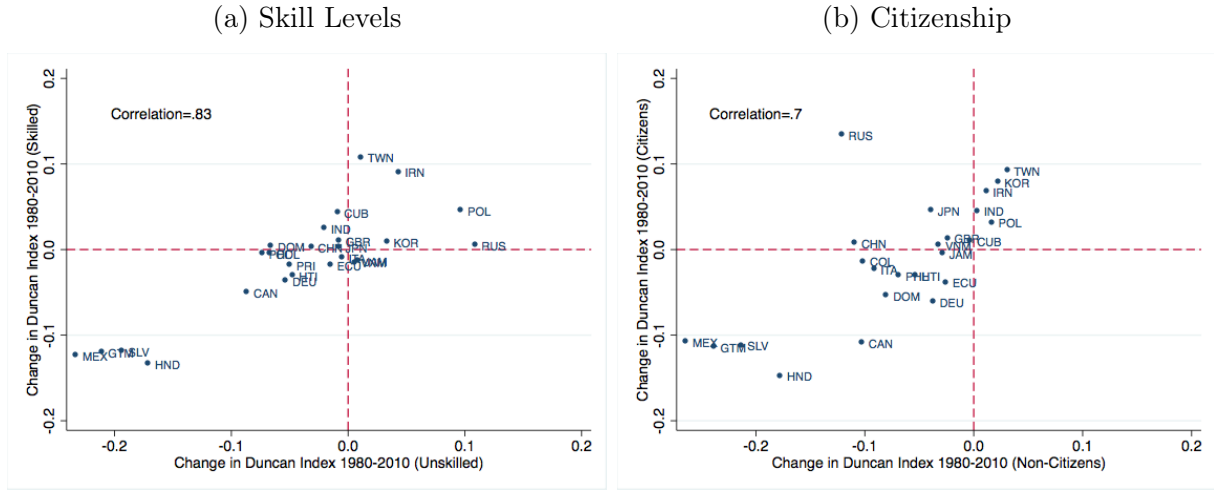
Notes: This figure displays the share of each state's working age population that is foreign-born. From lightest to darkest, bins are 0-2.5%, 2.6-5%, 5.1-10%, 10.1-20%, and 20% or more. The data are from the 1980-2000 5% decennial Census and 2010 ACS 5-year estimates from IPUMS and are limited to adults aged 18-64 not living in group quarters or attending school. Foreign-born is defined to be all individuals born outside the United States who did not have American parents.

Figure 1.2: Geographic Aggregations



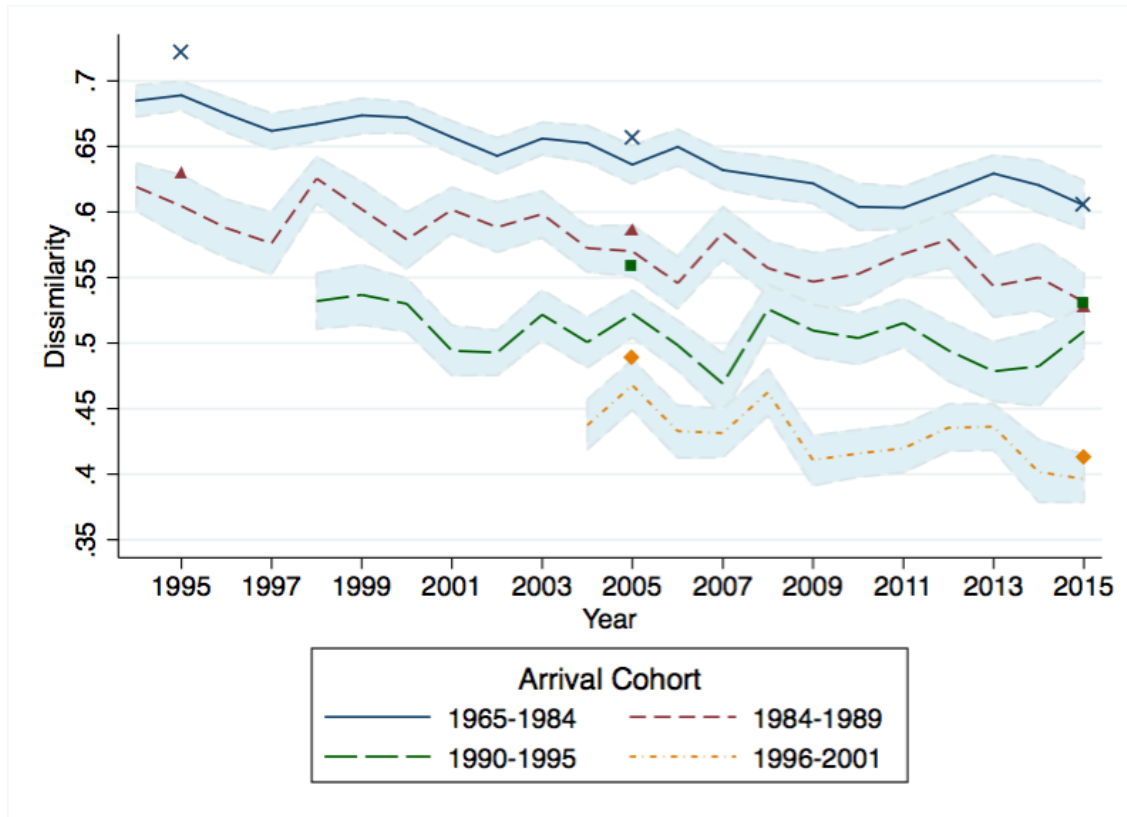
Notes: The map displays the boundaries of conspumas, the smallest consistently-identifiable geographic areas available in the IPUMS data between 1980 and 2010. Shaded areas have been aggregated to a single unit to more closely match the definition of a metropolitan area. Alaska and Hawaii do not appear in the map, but are included in the analysis.

Figure 1.3: Changes in Dissimilarity by Education and Citizenship Status



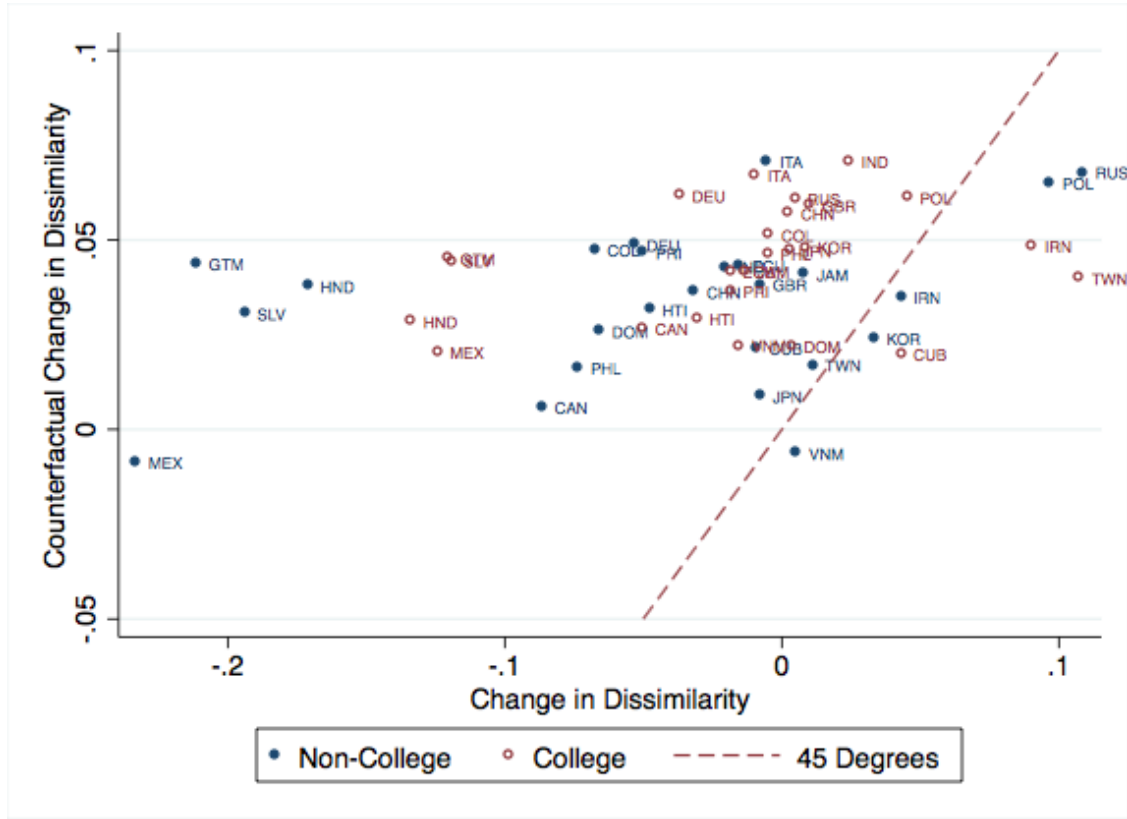
Notes: Figure displays changes in the dissimilarity index for the 25 largest immigrant groups in the United States as of 2000, calculated at the city level by education and citizenship status. Skilled refers to individuals with some college or more, while unskilled refers to individuals with a high-school education or less. Dissimilarity computed based on education compares immigrants to natives with the same level of education. Dissimilarity based on citizenship status compares immigrants to all natives since natives are all citizens by definition. Data are from the 1980-2000 5% decennial Census and 2010 ACS 5-year estimates from IPUMS and are limited to adults aged 18-64 not living in group quarters or attending school.

Figure 1.4: Changes in Dissimilarity Among Mexican Immigrants



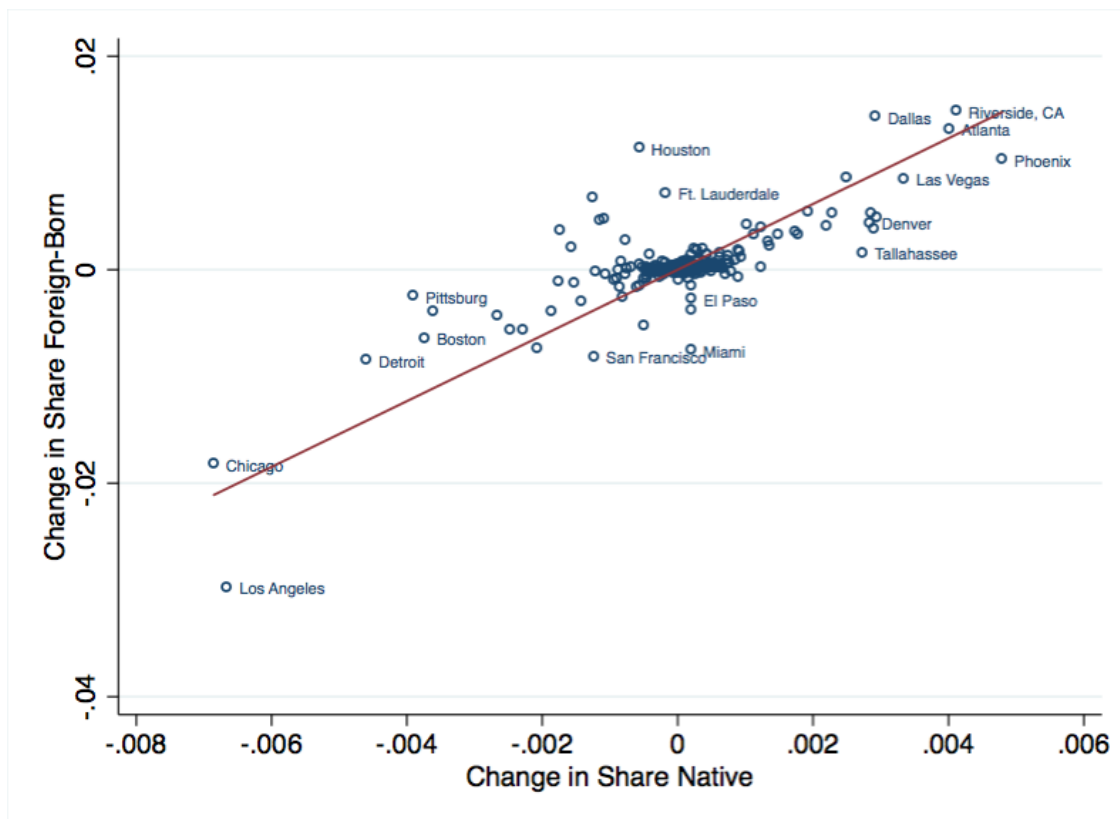
Notes: Figure displays trends in dissimilarity among different cohorts of Mexican immigrants, calculated at the state level. Shaded areas indicate 90% confidence intervals which have been computed by resampling yearly data for 1000 bootstrap repetitions, as described in the text. Markers indicate dissimilarity among immigrants who have not moved between states in the previous five years. The sample is Mexican immigrants in the 1994-2015 CPS between the ages of 18 and 64 who are not living in group quarters or attending school. "

Figure 1.5: Counterfactual Change in Dissimilarity



Notes: Table displays the actual change in dissimilarity and the counterfactual change in dissimilarity separately by education status for the 25 largest immigrant groups in the United States as of 2000. As discussed in the text, the counterfactual dissimilarity index is constructed by holding the distribution of immigrants constant in 1980 so changes in the index are driven by the movement of natives alone. Data are from the 1980-2000 5% decennial Census and 2010 ACS 5-year estimates from IPUMS and are limited to adults aged 18-64 not living in group quarters or attending school.

Figure 1.6: Change in Immigrant and Native Population Share by City



Notes: Figure displays the change in the share of natives and immigrants living in each city between 1980 and 2010. New York is omitted from the graph. In New York the share foreign-born decline by 0.07 and declined by 0.006 among natives. New York is included in the regression line. Data are from the 1980-2000 5% decennial Census and 2010 ACS 5-year estimates from IPUMS and are limited to adults aged 18-64 not living in group quarters or attending school.

Table 1.1: Compositional Changes in the Foreign-Born Population 1980-2010

	Share of Population by Time Since Arrival								
	Share of Foreign-Born			5 Years or Less		5-10 Years		10-20 Years	
	1980	2010	Change	1980	2010	1980	2010	1980	2010
All Foreign-Born	1	1	-	0.22	0.09	0.18	0.16	0.29	0.30
Mexico	0.16	0.30	0.14	0.29	0.08	0.26	0.18	0.26	0.32
India	0.02	0.05	0.03	0.38	0.16	0.35	0.21	0.22	0.33
El Salvador	0.01	0.03	0.02	0.45	0.07	0.29	0.16	0.19	0.30
Vietnam	0.01	0.03	0.02	0.87	0.05	0.09	0.08	0.03	0.34
Philippines	0.04	0.05	0.01	0.33	0.09	0.30	0.14	0.27	0.26
China	0.03	0.04	0.01	0.24	0.09	0.22	0.15	0.32	0.36
Korea	0.02	0.03	0.01	0.47	0.07	0.32	0.12	0.16	0.23
Russia/USSR	0.01	0.02	0.01	0.36	0.09	0.06	0.18	0.10	0.54
Guatemala	0.01	0.02	0.01	0.35	0.14	0.31	0.22	0.27	0.28
Colombia	0.01	0.02	0.01	0.24	0.08	0.26	0.20	0.41	0.28
Dominican Republic	0.01	0.02	0.01	0.26	0.08	0.24	0.12	0.43	0.34
Honduras	0.00	0.01	0.01	0.21	0.14	0.27	0.23	0.37	0.35
Taiwan	0.00	0.01	0.01	0.44	0.04	0.30	0.08	0.24	0.24
Iran	0.01	0.01	0.00	0.57	0.06	0.15	0.11	0.19	0.18
Haiti	0.01	0.01	0.00	0.31	0.06	0.31	0.13	0.33	0.28
Ecuador	0.01	0.01	0.00	0.21	0.08	0.26	0.21	0.45	0.32
Jamaica	0.01	0.01	0.00	0.25	0.06	0.27	0.11	0.37	0.25
Poland	0.02	0.01	-0.01	0.10	0.05	0.09	0.13	0.26	0.36
Japan	0.02	0.01	-0.01	0.27	0.18	0.14	0.16	0.26	0.24
United Kingdom	0.04	0.02	-0.02	0.13	0.09	0.08	0.13	0.28	0.23
Canada	0.05	0.02	-0.03	0.09	0.09	0.05	0.12	0.25	0.25
Cuba	0.05	0.02	-0.03	0.05	0.09	0.18	0.14	0.63	0.23
Italy	0.05	0.01	-0.04	0.05	0.04	0.12	0.05	0.30	0.10
Puerto Rico	0.08	0.03	-0.05	-	-	-	-	-	-
Germany	0.06	0.01	-0.05	0.05	0.08	0.04	0.08	0.25	0.16

Notes: This table displays summary statistics on the composition of the foreign-born population of the United States in 1980 and 2010. Year of arrival in the mainland United States is not recorded for the territory of Puerto Rico. Rows are sorted by change in share foreign-born.

Table 1.2: Changes in Immigrant Dissimilarity by Country of Origin

	1980	1990	2000	2010	Change 1980-2010
All Foreign-Born	0.45	0.48	0.44	0.40	-0.05
Mexico	0.73	0.70	0.57	0.50	-0.23
Philippines	0.59	0.60	0.58	0.55	-0.04
Puerto Rico	0.67	0.63	0.61	0.58	-0.09
Germany	0.26	0.24	0.21	0.20	-0.06
China	0.57	0.59	0.56	0.53	-0.04
India	0.40	0.43	0.43	0.41	0.01
Vietnam	0.46	0.53	0.48	0.46	0.00
Canada	0.42	0.38	0.33	0.31	-0.11
Korea	0.45	0.50	0.50	0.48	0.03
Cuba	0.75	0.75	0.75	0.75	0.00
Russia/USSR	0.47	0.58	0.57	0.51	0.04
United Kingdom	0.30	0.30	0.30	0.30	0.00
El Salvador	0.78	0.74	0.64	0.59	-0.19
Dominican Republic	0.84	0.83	0.79	0.76	-0.08
Jamaica	0.69	0.69	0.68	0.67	-0.02
Italy	0.58	0.57	0.55	0.52	-0.06
Colombia	0.67	0.66	0.63	0.59	-0.08
Japan	0.46	0.44	0.45	0.46	0.00
Guatemala	0.71	0.68	0.56	0.50	-0.21
Poland	0.57	0.57	0.61	0.60	0.03
Haiti	0.81	0.80	0.78	0.75	-0.06
Taiwan	0.47	0.52	0.55	0.55	0.08
Ecuador	0.73	0.74	0.73	0.70	-0.03
Iran	0.51	0.56	0.56	0.56	0.05
Honduras	0.66	0.65	0.56	0.50	-0.16

Notes: This table displays the dissimilarity index for the 25 largest immigrant groups in the United States as of 2000, calculated at the city level. Rows are sorted by the group's share of the foreign-born population. As described in the text, the dissimilarity index measures the concentration of immigrants relative to the population of natives and is interpreted as the share of the immigrant group that would have to move from an over-represented to an under-represented region to match the geographic distribution of natives.

Table 1.3: Changes in Immigrant Dissimilarity within Historical Destination States

		Share of Foreign-Born	1980	1990	2000	2010	Change 1980-2010
CA	All Foreign-Born	-	0.22	0.22	0.21	0.18	-0.04
	Mexico	0.44	0.29	0.26	0.22	0.19	-0.10
	Philippines	0.07	0.26	0.27	0.27	0.23	-0.03
	China	0.05	0.34	0.41	0.45	0.44	0.10
	Vietnam	0.05	0.26	0.31	0.35	0.35	0.09
FL	All Foreign-Born	-	0.45	0.44	0.39	0.35	-0.10
	Cuba	0.22	0.72	0.72	0.69	0.64	-0.08
	Puerto Rico	0.09	0.35	0.31	0.31	0.28	-0.07
	Mexico	0.06	0.44	0.35	0.31	0.26	-0.18
	Haiti	0.06	0.67	0.61	0.55	0.52	-0.15
IL	All Foreign-Born	-	0.35	0.35	0.35	0.32	-0.03
	Mexico	0.38	0.37	0.35	0.32	0.29	-0.08
	Poland	0.09	0.43	0.43	0.46	0.42	-0.01
	India	0.05	0.38	0.35	0.39	0.34	-0.04
	Philippines	0.04	0.37	0.38	0.36	0.32	-0.05
NJ	All Foreign-Born	-	0.23	0.24	0.25	0.24	0.01
	Puerto Rico	0.09	0.23	0.20	0.18	0.16	-0.07
	India	0.07	0.18	0.28	0.36	0.41	0.23
	Dominican Republic	0.05	0.51	0.51	0.46	0.42	-0.09
	Mexico	0.04	0.31	0.32	0.32	0.25	-0.06
NY	All Foreign-Born	-	0.41	0.43	0.44	0.41	0.00
	Dominican Republic	0.10	0.56	0.57	0.57	0.53	-0.03
	Puerto Rico	0.09	0.50	0.49	0.44	0.39	-0.11
	China	0.06	0.51	0.52	0.54	0.52	0.01
	Russia/USSR	0.05	0.40	0.49	0.52	0.49	0.09
TX	All Foreign-Born	-	0.31	0.26	0.25	0.23	-0.08
	Mexico	0.64	0.40	0.32	0.27	0.24	-0.16
	Vietnam	0.04	0.35	0.32	0.32	0.28	-0.07
	El Salvador	0.03	0.59	0.54	0.51	0.49	-0.10
	India	0.03	0.38	0.32	0.32	0.33	-0.05

Notes: This table displays the dissimilarity index for the largest foreign born groups within the six states which historically attracted the largest share of immigrants. The index is computed over aggregated conspumas within states which cover the entire geography of the state. See the text and appendix for additional details on the construction of these areas. There are 27 regions in CA, 15 in FL, 6 in IL, 9 in NJ, 13 in NY, and 26 in TX. Share of the state's foreign-born population is measured in 2000.

Table 1.4: Correlates of Immigrant Dissimilarity 1980-2010

	All				Non-College	College
	(1)	(2)	(3)	(4)	(5)	(6)
Year 1990	0.0227** (0.0104)	-0.0015 (0.0046)	0.0113 (0.0143)	0.0321** (0.0134)	0.0528*** (0.0136)	0.0395*** (0.0142)
Year 2000	-0.0140 (0.0289)	-0.0524** (0.0247)	-0.0292 (0.0276)	0.0000 (0.0204)	0.0282 (0.0212)	0.0474** (0.0210)
Year 2010	-0.0492 (0.0396)	-0.0924** (0.0353)	-0.0727* (0.0433)	-0.0218 (0.0303)	0.0092 (0.0258)	0.0502* (0.0289)
College	-	-	-0.0738** (0.0289)	-0.1074*** (0.0207)	-	-
log(Population)	-	-	-0.0242*** (0.0075)	-0.0350** (0.0137)	-0.1015*** (0.0175)	-0.0440** (0.0186)
Share Female	-	-	-0.1573 (0.1880)	-0.2010* (0.1018)	-0.3330** (0.1410)	0.0228 (0.1083)
Mean Age	-	-	0.0020 (0.0059)	-0.0031 (0.0040)	-0.0105*** (0.0024)	-0.0029 (0.0034)
Share Unmarried	-	-	0.6298*** (0.2294)	-0.1157 (0.1846)	0.1302 (0.1265)	0.0167 (0.1162)
Share Fluent English	-	-	-0.2399*** (0.0845)	-0.1751*** (0.0503)	-0.2945** (0.1200)	-0.0828 (0.0974)
Share Citizens	-	-	0.1177** (0.0470)	0.2150*** (0.0755)	0.5118*** (0.0778)	0.1319*** (0.0436)
Birthplace FE		✓		✓	✓	✓
Observations	792	792	792	792	396	396
R-squared	0.0401	0.7963	0.5023	0.9158	0.9696	0.9735

Notes: The dependent variable is the dissimilarity index, which has been computed separately for each birthplace-education cell. All regressions are weighted by cell size and have standard errors clustered by birthplace. The sample includes the 100 largest immigrant groups in the United States as of 2000.

Table 1.5: Changes in Dissimilarity Within and Across Arrival Cohorts

	All		Non-College		College	
	(1)	(2)	(3)	(4)	(5)	(6)
Arrived 1975-1979	-0.0429*** (0.0064)	-0.0335*** (0.0047)	-0.0350*** (0.0053)	-0.0337*** (0.0046)	-0.0223*** (0.0043)	-0.0217*** (0.0043)
Arrived 1980-1984	-0.0657*** (0.0122)	-0.0646*** (0.0125)	-0.0796*** (0.0129)	-0.0799*** (0.0140)	-0.0358*** (0.0054)	-0.0354*** (0.0045)
Arrived 1985-1989	-0.1040*** (0.0203)	-0.1021*** (0.0207)	-0.1214*** (0.0230)	-0.1161*** (0.0210)	-0.0560*** (0.0070)	-0.0553*** (0.0067)
Arrived 1990-1994	-0.1242*** (0.0251)	-0.1191*** (0.0263)	-0.1425*** (0.0293)	-0.1355*** (0.0281)	-0.0701*** (0.0086)	-0.0703*** (0.0082)
Arrived 1995-1999	-0.1711*** (0.0314)	-0.1683*** (0.0324)	-0.1932*** (0.0361)	-0.1840*** (0.0348)	-0.1045*** (0.0126)	-0.1041*** (0.0129)
Arrived 2000-2004	-0.1798*** (0.0344)	-0.1767*** (0.0351)	-0.2055*** (0.0378)	-0.1992*** (0.0417)	-0.1111*** (0.0151)	-0.1119*** (0.0149)
Arrived 2005-2010	-0.2025*** (0.0355)	-0.1989*** (0.0377)	-0.2285*** (0.0439)	-0.2194*** (0.0465)	-0.1225*** (0.0152)	-0.1221*** (0.0164)
In U.S. 5-10 Years	-0.0278*** (0.0056)	-0.0324*** (0.0037)	-0.0317*** (0.0047)	-0.0294*** (0.0043)	-0.0208*** (0.0040)	-0.0195*** (0.0067)
In U.S. 11-15 Years	-0.0430*** (0.0056)	-0.0469*** (0.0037)	-0.0494*** (0.0041)	-0.0444*** (0.0061)	-0.0313*** (0.0034)	-0.0297*** (0.0097)
In U.S. 16-20 Years	-0.0748*** (0.0073)	-0.0797*** (0.0053)	-0.0831*** (0.0061)	-0.0653*** (0.0133)	-0.0523*** (0.0056)	-0.0492*** (0.0140)
In U.S. 21-25 Years	-0.0810*** (0.0076)	-0.0842*** (0.0057)	-0.0886*** (0.0074)	-0.0686*** (0.0143)	-0.0581*** (0.0055)	-0.0544*** (0.0163)
In U.S. more than 25 Years	-0.1164*** (0.0110)	-0.1193*** (0.0093)	-0.1239*** (0.0121)	-0.0930*** (0.0199)	-0.0805*** (0.0077)	-0.0750*** (0.0221)
Birthplace Fixed Effects		✓	✓	✓	✓	✓
Demographic Controls				✓		✓
Observations	3,941	3,941	1,970	1,970	1,971	1,971
R-squared	0.2765	0.7365	0.8015	0.8200	0.8454	0.8458

Notes: The dependent variable is the dissimilarity index, which has been calculated separately for each birthplace-education-cohort cell. The omitted category for arrival cohort is immigrants arriving in 1970-1974. Demographic controls include mean age and share female, married, U.S. citizens and English-speaking. All regressions are weighted by cell population and have standard errors clustered by birthplace. The sample includes the 100 largest immigrant groups in the United States as of 2000. Results are robust to dropping extremely small cells, and to excluding Hispanic immigrants.

Table 1.6: Changes in Dissimilarity Within Arrival Cohorts

	All		Non-College		College	
	(1)	(2)	(3)	(4)	(5)	(6)
In U.S. 5-10 Years	-0.0317 (0.0246)	-0.0216 (0.0253)	-0.0333 (0.0390)	-0.0162 (0.0361)	-0.0202* (0.0111)	0.0028 (0.0106)
In U.S. 11-15 Years	-0.0473*** (0.0136)	-0.0347* (0.0185)	-0.0499** (0.0223)	-0.0206 (0.0167)	-0.0312*** (0.0055)	0.0043 (0.0071)
In U.S. 16-20 Years	-0.0792*** (0.0195)	-0.0594** (0.0286)	-0.0852*** (0.0307)	-0.0400 (0.0328)	-0.0516*** (0.0091)	-0.0002 (0.0116)
In U.S. 21-25 Years	-0.0851*** (0.0173)	-0.0640** (0.0275)	-0.0905*** (0.0282)	-0.0386 (0.0239)	-0.0586*** (0.0074)	-0.0018 (0.0104)
In U.S. more than 25 Years	-0.1197*** (0.0170)	-0.0923*** (0.0328)	-0.1283*** (0.0261)	-0.0667** (0.0322)	-0.0806*** (0.0080)	-0.0143 (0.0126)
Share Fluent English	-	-0.2049*** (0.0127)	-	-0.4154*** (0.1474)	-	-0.1265*** (0.0293)
Share U.S. Citizens	-	0.0397 (0.0347)	-	0.0094 (0.0510)	-	-0.0438** (0.0179)
Birthplace-Cohort FE	✓	✓	✓	✓	✓	✓
Observations	3,941	3,941	1,970	1,970	1,971	1,971
R-squared	0.8228	0.8797	0.8845	0.9075	0.9367	0.9440

Notes: The dependent variable is the dissimilarity index, which has been calculated separately for each birthplace-education-cohort cell. Share Fluent English refers to the share of the group who report speaking only English or English very well. All regressions are weighted by cell population and have standard errors clustered by birthplace-cohort cell. The sample includes the 100 largest immigrant groups in the United States as of 2000. Results are robust to dropping extremely small cells, and to excluding Hispanic immigrants.

Table 1.7: Effect of Employment Growth and Enclaves on New Immigrants' Location Choices

	All		Non-College		College	
	(1)	(2)	(3)	(4)	(5)	(6)
Employment Growth	0.5037*** (0.0846)	0.4345*** (0.0777)	0.5822*** (0.1010)	0.5428*** (0.0968)	0.4253*** (0.0738)	0.3577*** (0.0783)
Lag Group Share	1.2206*** (0.4138)	1.2324*** (0.4138)	1.3846*** (0.3933)	1.3881*** (0.3924)	1.0566** (0.4445)	1.0719** (0.4448)
Distance (1000 miles)	-0.4274*** (0.1094)	-0.4176*** (0.1110)	-0.4687*** (0.1094)	-0.4937*** (0.1142)	-0.3861*** (0.1142)	-0.3558*** (0.1126)
Birthplace FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Education FE	✓	✓				
Observations	30,000	30,000	15,000	15,000	15,000	15,000
R-squared	0.1935	0.1876	0.1999	0.1952	0.1923	0.1861

Notes: This table displays the results of regression of the immigrant inflow rate for a city on the indicated variables, where the inflow rate is defined as the share of immigrants choosing to locate in a city deflated by the share of the native population living in the city. All regressions have standard errors clustered by city. Even numbered columns use the education-specific Bartik shift-share variable. Note that 1980 is omitted since is not possible to calculate lag group share in 1970 at the city level. The sample is limited to the 50 largest immigrant groups and 100 largest cities as of 2000.

Table 1.8: Effect of Employment Growth and Enclaves on New Immigrants' Location by Decade

	Non-College			College		
	1990 (1)	2000 (2)	2010 (3)	1990 (4)	2000 (5)	2010 (6)
Employment Growth	0.5313*** (0.1124)	0.5914*** (0.0989)	0.3817*** (0.1245)	0.4692*** (0.1133)	0.4901*** (0.0580)	-0.1039 (0.1173)
Lag Group Share	2.6994*** (0.5252)	1.7626*** (0.4662)	1.0531*** (0.2828)	2.0901*** (0.6844)	1.3095** (0.5822)	0.7986** (0.3066)
Distance (1000 miles)	-0.4331*** (0.1198)	-0.5288*** (0.1136)	-0.4177*** (0.1231)	-0.3021** (0.1192)	-0.3894*** (0.1090)	-0.3411*** (0.1133)
Birthplace FE	✓	✓	✓	✓	✓	✓
Observations	5,000	5,000	5,000	5,000	5,000	5,000
R-squared	0.2740	0.2512	0.1654	0.2752	0.2503	0.1562

Notes: This table displays the results of regression of the immigrant inflow rate for a city on the indicated variables, where the inflow rate is defined as the share of immigrants choosing to locate in a city deflated by the share of the native population living in the city. All regressions have standard errors clustered by city. Even numbered columns use the education-specific Bartik shift-share. Note that 1980 is omitted since is not possible to calculate lag group share in 1970 at the city level. The sample is limited to the 50 largest immigrant groups and 100 largest cities as of 2000.

Table 1.9: Effect of Employment Growth and Enclaves on Internal Migration

	All		Non-College		College	
	(1)	(2)	(3)	(4)	(5)	(6)
Employment Growth	-0.0098* (0.0050)	-0.0131*** (0.0046)	-0.0034 (0.0034)	-0.0031 (0.0031)	-0.0172** (0.0074)	-0.0259*** (0.0092)
Lag Group Share	-0.0052** (0.0025)	-0.0060** (0.0023)	-0.0045** (0.0021)	-0.0049** (0.0021)	-0.0144*** (0.0040)	-0.0121** (0.0050)
Speaks English	0.0169*** (0.0026)	0.0169*** (0.0026)	0.0077*** (0.0014)	0.0078*** (0.0014)	0.0346*** (0.0049)	0.0340*** (0.0051)
Citizen	-0.0168*** (0.0033)	-0.0166*** (0.0034)	-0.0077*** (0.0017)	-0.0076*** (0.0017)	-0.0282*** (0.0050)	-0.0282*** (0.0050)
Years in U.S.	-0.0012*** (0.0001)	-0.0012*** (0.0001)	-0.0009*** (0.0002)	-0.0009*** (0.0002)	-0.0015*** (0.0002)	-0.0015*** (0.0002)
Birthplace FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Demographic Controls	✓	✓	✓	✓	✓	✓
Education FE	✓	✓				
Observations	1,488,082	1,488,082	970,486	970,486	517,596	517,596
R-squared	0.0350	0.0358	0.0199	0.0198	0.0464	0.0494

Notes: This table displays the results of linear regressions of an indicator for cross-state migration on the indicated variables. All regressions are weighted by Census sampling weights and have standard errors clustered by state. Demographic controls include gender, marital status, and a quadratic in age. Even numbered columns use the education-specific Bartik shift-share. The sample includes individuals in the 100 largest immigrant groups in the United States as of 2000. Note that 2010 is omitted from the sample because the ACS asks 1 year rather than 5 year migration status.

Table 1.10: Effect of Employment Growth and Enclaves on Internal Migration by Decade

	Non-College			College		
	1990 (1)	2000 (2)	2010 (3)	1990 (4)	2000 (5)	2010 (6)
Employment Growth	-0.0012 (0.0057)	-0.0019 (0.0056)	-0.0020 (0.0028)	-0.0188* (0.0102)	-0.0234*** (0.0085)	-0.0271** (0.0112)
Lag Group Share	-0.0365*** (0.0053)	-0.0169*** (0.0026)	-0.0041*** (0.0013)	-0.0387*** (0.0118)	-0.0205*** (0.0070)	-0.0113*** (0.0037)
Speaks English	0.0146*** (0.0021)	0.0082*** (0.0019)	0.0048*** (0.0014)	0.0321*** (0.0046)	0.0261*** (0.0032)	0.0372*** (0.0070)
Citizen	-0.0014 (0.0034)	-0.0043 (0.0028)	-0.0105*** (0.0011)	-0.0133** (0.0056)	-0.0135*** (0.0048)	-0.0371*** (0.0047)
Years in U.S.	-0.0011*** (0.0003)	-0.0010*** (0.0001)	-0.0007*** (0.0002)	-0.0027*** (0.0006)	-0.0019*** (0.0003)	-0.0012*** (0.0002)
Birthplace FE	✓	✓	✓	✓	✓	✓
Demographic Controls	✓	✓	✓	✓	✓	✓
Observations	104,657	306,496	559,333	44,002	180,046	293,548
R-Squared	0.0409	0.0309	0.0190	0.0536	0.0433	0.0572

Notes: This displays the results of linear regressions of an indicator for cross-state migration on the indicated variables. All regressions are weighted by Census sampling weights and have standard errors clustered by state. Demographic controls include gender, marital status, and a quadratic in age. Even numbered columns use the education-specific Bartik shift-share. The sample includes individuals in the 100 largest immigrant groups in the United States as of 2000. Note that 2010 is omitted from the sample because the ACS asks 1 year rather than 5 year migration status.

CHAPTER 2

Historical Data on Professional and Occupational Credentialing Requirements in the United States

2.1 Introduction

Professional and occupational credentialing is an important labor market institution through which federal, state, and municipal regulators establish the qualifications necessary to legally work in at least a quarter of U.S. jobs. This form of regulation – including occupational licensing, certification, and registration requirements – proliferated during the second half of the twentieth century, and now applies to hundreds of unique professions ([Kleiner and Krueger, 2013](#)). With the vast majority of these requirements set at the state level, the decentralized nature of occupational regulation has made it challenging for researchers to assemble data on the historical development of this institution. In addition to having to confront thousands of individual policies adopted independently across political jurisdictions, researching the history of each policy requires extensively searching multiple sources of legal documentation, some of which have only recently been digitized.

In this chapter, I describe the construction a novel dataset that provides the most comprehensive history of state and federal occupational credentialing laws assembled to date. It covers all fifty states and the District of Columbia, containing data for 250 unique occupation categories and over 7,000 individual regulatory events implemented as early as the late 1800s and extending to the present day. The database includes all major laws that enacted, amended, or repealed legal provisions governing occupational practice, titles, or advertis-

ing. It is the first dataset to link each credentialing requirement to both current statutes or administrative regulations, as well as to the original text of the underlying legislation that adopted or amended each policy. Every unique occupation is assigned an identifier based on the Standard Occupational Classification system, which can then be used to crosswalk the data to common statistical occupation codes.

Relative to the few existing datasets compiling similar information, my data offers a significantly longer uninterrupted timeseries, the ability to differentiate between various methods of regulation, and greater historical accuracy. These advantages are a direct result of my data collection methodology, which incorporates documentation from numerous primary and secondary sources. A comprehensive analysis of state session laws, in particular, allows me locate superseded legislation and accurately determine the original effective date of specific policies even when the statutes that initially implemented them have since been replaced by newer legislation. Another benefit to my approach is that, unlike other sources, I am able to identify transitions between different regulatory mechanisms. Many occupations that are currently licensed were previously subject to weaker certification or registration policies.

In addition to a comprehensive search of statutes and session laws, I also digitize and process dozens of secondary sources to validate the main legislative data. These include online databases, professional directories, policy reports, and law review articles. To the best of my knowledge, they cover essentially all published material compiling policy information for more than a handful of occupations or states. Although these secondary sources are insufficient to determine the exact year of policy changes, by cross-referencing the timing of the legislative events I find against the publication dates of these sources, I am able to identify and evaluate potential inconsistencies in the timeseries. I show that my data are highly consistent with these independent reports, providing strong evidence that the data are historically accurate. As a final quality test, I use self-reported licensing status from the Current Population Survey to show that my policy data strongly predicts within-occupation variation in the share of workers reporting a government-issued credential across states.

2.2 Background

Professional and occupational credentialing is a complex institution involving multiple levels of government and various approaches to regulation. Before introducing the data, this section explains important aspects of the legal context for this project. First, I formally define licensing, certification, and registration. Next, I provide a brief overview of relevant legal documentation and describe how this informs my approach to data collection.

2.2.1 Regulatory Terminology

In the United States, federal, state, and municipal laws regulate occupational practice through a variety of credentialing requirements that individuals must meet before they may legally work or advertise services to the public. The most common regulatory approaches typically fall into one of three distinct tiers known as licensing, certification, and registration ([Shimberg, 1980](#)). Licensing, or formally right-to-practice, is the most restrictive method, and also the most prevalent. Under licensing, regulators define the tasks that fall within an occupation’s scope of practice and make preforming this type of work without meeting predetermined entry standards a criminal offense. Requirements differ across occupations and jurisdictions, but almost always involve some combination of minimum education or experience, qualifying examinations, and fees.

State certification, or statutory title protection, is a less restrictive regulatory method that reserves the use of specific occupation titles for workers who have obtained a credential. Although training requirements may be similar to licensing, certification does not limit the tasks uncertified workers may preform, provided that they do not describe these services using legally-protected terms. In contrast to private certification, state certification is a direct method of labor market regulation with standards set and enforced by a government agency.¹ Finally, registration requirements are a weaker form of oversight under which work-

¹Many practice acts also include title protections, in particular for healthcare and social service jobs, where

ers must only file their name and qualifications with the government before beginning their practice. Unlike licensing and certification, registration requirements do not require workers to demonstrate a minimum level of competency, though proof of bonding and insurance or criminal background checks may still be required.²

Although the definitions of licensing, certification, and registration presented here are commonly accepted ([National Conference of State Legislatures, 2017](#)), research on these policies can be complicated by the fact that regulatory terminology is often used inconsistently in practice. Registered dietitians, for example, may be licensed or certified depending on the state where they work, but must always meet educational qualifications and pass an exam. When assembling the data, I therefore record practice and title protections directly, allowing me to classify regulations based on these conventions alone. Doing so maximizes comparability across states and occupations, ensuring that consistent definitions are applied regardless of what the regulating agency calls the credential.

2.2.2 Institutional Details

Assembling historical data on occupational licensing, certification, and registration requirements presents a number of unique challenges and requires examining legal documentation from multiple sources. To better understand the data collection methodology adopted in this project, it is therefore helpful to briefly overview how state and federal laws are organized.

a profession's scope of practice may be difficult to exhaustively define or overlap with those of other regulated occupations. See, for example, [Fischer \(1980\)](#) for a discussion of certification and licensing legislation in the field of psychology. An additional distinction is sometimes made between laws that restrict only the use of modifiers such as "state certified" and stronger statutes that protect unmodified titles such as "accountant," "interior designer," or "psychologist." While I refer to all these regulations as certification, I record the strength of title protections in the data.

²Governments may also indirectly regulate occupations through deceptive trade practice acts, limitations on eligibility for public insurance reimbursement, or minor training requirements such as food safety courses that do not require a formal credentialing process. These methods of regulation are widely considered to be less restrictive than licensing or certification ([Hemphill and Carpenter, 2016](#)) and are not included in the data.

Statutes and Regulations. The vast majority of occupational credentialing requirements occur at the state level and are adopted by statute, the general laws enacted by the state legislature. A typical occupational licensing statute defines the occupation’s scope of practice, specifies a list of tasks or titles reserved for licensed workers, and delegates implementation and enforcement of the law to a professional board, often dominated by members of the profession itself (Allensworth, 2017). The rules adopted by the board under the authority of the statute are known as administrative laws or regulations, and though subordinate to statutes, also have the force of law. Regulations written by the board usually cover specific provisions such as the content of entry examinations or procedures for disciplinary action. In rare circumstances, a board’s statutory authority is broad enough to adopt credentialing requirements for occupations that are not explicitly covered by the governing statute, but these cases are the exception rather than the norm.³

Session Laws. Statutes are published in multiple formats, the most readily accessible of which are called codes. Statutory code compiles laws that are currently in effect and organizes them topically, so that related laws are grouped together. Annotated codes provide additional historical references to the legislative acts that adopted or amended the statute, as well as to any relevant court cases.⁴ Session laws, by contrast, compile all legislative acts adopted during a particular session of the legislature. They are published annually or biennially depending on the length of the state’s legislative session and contain the original text of all newly-enacted statutes, as well as comprehensive documentation of acts amending, replacing, or repealing portions of existing code. Simply put, codes organize the stock of existing laws, while session laws record changes as they are implemented by the legislature.

³For example, in 1991 the Minnesota Department of Public Health adopted rules implementing a state certification policy for speech-language pathologists and audiologists under the authority of a statute regulating hearing aid specialists. The state legislature adopted a new statute codifying this regulation five years later.

⁴With the exception of a small number of rulings that overturned practice or title restrictions, court decisions rarely affect occupational regulation on the extensive margin. Case law is therefore less relevant for this project than statutory or administrative law.

Federal and Municipal Laws. Federal laws are also organized into statutes and regulations, but differ from state laws in that most occupational licensing requirements are adopted through regulations rather than statutes. While there is no comprehensive source of historical administrative law at the state level, changes to federal regulations are well documented in annual publications of the Code of Federal Regulations. Municipal codes, by contract, are much less accessible and therefore fall outside the scope of this project except when legislation adopted at the state level explicitly delegates the regulation of specific occupations to local jurisdictions.⁵

2.2.3 Implications for Data Collection

At the state level, annotated statutes are the most useful starting point for data collection since they reflect laws currently in effect. Without also examining session laws, however, codes are insufficient to construct a complete and accurate timeline of all relevant legislation affecting a particular occupation. This is because references to superseded laws are not well-documented in many statutes, potentially leading to incorrect conclusions about the date policies went into effect if relying on annotations alone.

More specifically, research based on statutory code suffers from three related limitations that can be addressed by examining the underlying legislation. First, many states periodically repeal and replace entire chapters of code, reenacting amended laws under a new statute. The earliest citations listed in the current version of the law may therefore reference the replacement act rather than the original one. This issue is pervasive and becomes increasingly severe the longer an occupation has been regulated.⁶ Second, many statutes

⁵Municipal laws cover an extremely small share of licensed workers. In the Survey of Income and Program Participation, the only official household survey reporting detailed information on the source of worker credentials, fewer than 3% of respondents holding a government-issued license or certificate reported obtaining their credential from a local (as opposed to state or federal) agency.

⁶For example, the earliest reference in the Arizona Nurse Practice Act to the licensing of registered nurses is to a law passed during the 1997 legislative session. In fact, Arizona adopted voluntary certification for registered nurses in 1921 and mandatory licensing in 1952. These early statutes are easily located in Arizona's

have been amended numerous times without detailed references to what specific changes were made in the annotations. This makes it difficult to determine, for example, whether an occupation that is currently licensed was previously registered or certified. Finally, laws repealed or overturned in court may not appear in the current code at all. When collecting the data, I therefore preform a comprehensive search of both statutory code and session laws, which considerably improves the reliability of my historical data.⁷

2.3 Data Description

2.3.1 Historical Credentialing Laws

The main dataset I produce in this project identifies all major state and federal regulation events affecting a broad set of 250 unique occupation categories. It records the enactment and effective dates of any laws or regulations adopting occupational credentialing requirements, as well as transitions between different approaches to regulation, i.e. from registration or certification to licensure. The data cover all fifty states and the District of Columbia, and capture legislation passed as early as the late 1800s, including laws adopted by territorial legislatures. Each observation in the data corresponds to a specific policy change, with implementation dates coded annually. Only occupation-state pairs that were ever regulated appear in the data.

The database includes three key variables that are used to differentiate between licensing, certification, and registration requirements. First, it records whether the law limits any elements of the occupation’s scope of practice to workers who have obtained a credential. This variable also notes whether the practice protections are complete, covering all work typically

session laws and are consistent with the timing of regulation documented by secondary sources ([Monheit, 1982](#); [White, 1987](#)).

⁷Unfortunately, historical documentation of state administrative regulations is less complete. Additional sources are therefore needed to confirm enactment dates for the minority of occupations that are not regulated by statute. See [Section 4](#) for details.

performed by members of the occupation, or partial, applying only to a subset of advanced tasks.⁸ Second, the data indicate whether the law limits the use of titles or advertising in any way. As with practice restrictions, title protections may be strong, applying to titles such as “accountant,” or weak, limiting only the use of modifiers such as “state-certified” to workers who have met the state’s requirements. Finally, I note whether or not the law requires a worker to meet predetermined competency standards to obtain the credential. Any legislative event affecting these three variables is considered a major policy change and is recorded in the dataset.

To facilitate linking the data to external sources, each occupation is assigned a unique identifier based on a refinement of the Standard Occupational Classification (SOC) system. The SOC defines over 800 occupation categories covering all civilian jobs and is the most detailed level at which the federal government collects wage and employment data. It is based on a six-digit classification system that groups occupations into major, minor, and detailed categories. At the detailed level, these codes are sufficiently precise that the statistical and regulatory definitions of many occupations are nearly equivalent. To completely accommodate all remaining regulations in this system, I create an eight-digit identifier for occupations, such as acupuncturists, that do not have their own six-digit code. My system therefore retains distinct regulated occupation categories as the unit of observation while preserving the hierarchical structure of the SOC.

In addition to the main event-based dataset, I also construct a balanced annual panel that assigns a regulatory status to each occupation, state, and year cell from 1935 to 2018. For most regulations, this simply involves populating the timeseries between policy changes. In some cases, however, state and federal laws regulate the same occupation, resulting in overlapping coverage. For example, farm labor contractors and investment advisers are

⁸Most medical and service professions are subject to the former type of licensing. Engineering and scientific regulations on the other hand usually only require individuals in responsible charge of the work to be licensed. Likewise, nearly all public accountancy statutes require a certified public accountant to attest financial statements, but do not otherwise restrict general accounting services if the individual does not represent that they are certified.

regulated at both the state and federal level. When this is the case, I assume the federal law is binding unless the state law imposes more stringent requirements, as is the case when workers must register with a federal agency, but also obtain a state license to practice.

2.3.2 Comparison to Existing Datasets

This project is not the first to collect historical data on occupational licensing requirements, though to the best of my knowledge it is the most complete. I am aware of only two other datasets that report the enactment dates of statutes regulating a broad set of occupations. It is therefore worth highlighting how my data differ in scope and methodology from these alternatives.

The Northwestern Licensing Database. My dataset is most similar to the NLD, an ambitious project that conducted a census of state laws, recording the enactment dates of occupational licensing statutes adopted between 1970 and 2012 ([Redbird, 2017](#)). The key difference relative to the data I assemble is methodological. While I link each policy to both modern codes *and* underlying legislation published in session laws, the NLD infers policy enactment dates from the current statute. As discussed in [Section 2](#), this can result in the left-censoring of enactment dates. Citations provide evidence that a regulation existed by a particular time, but not if it was the first applying to the occupation of interest. Because the probability that a law has been replaced increases over time, measurement error may become increasingly likely the earlier an occupation was initially regulated.

By consulting session laws in addition to statutes, my approach to data collection is not susceptible to this censoring issue and allows me to fully observe the text of all relevant legislation affecting the occupation. In addition to improving the historical accuracy of the data and allowing me to construct a significantly longer panel, my method also identifies when the stringency of regulations changed, while the NLD does not differentiate between licensing, certification, and registration requirements. In my data, I find that that in many

cases, legislation granting certified practitioners right-to-title proceeded practice acts by years, or even decades. Finally, because it collects data from current statutes only, the NLD does not include historical regulations that have been repealed, while my data does.

Council of State Governments. The second source of data compiling legislative enactment dates for a large number of occupational licensing statutes is a report published by the [Council of State Governments \(1952\)](#). This seminal study undertook a comprehensive analysis of licensing regulations for 38 occupations, the near universe of policies known to be in effect at the time. It remains extensively cited to highlight the diffusion of professional regulation in the second half of the twentieth century, and is used as the primary data source for several recent studies including [Law and Kim \(2005\)](#) and [Law and Marks \(2009\)](#). Like the Northwestern Licensing Database, this data does not differentiate between licensing, certification, and registration requirements, while my data does. This is a particularly important distinction in the early twentieth century, as many states initially regulated occupations including architects, accountants, engineers, and nurses through certification rather than licensure.

The Council of State Governments data were also collected through an examination of state statutes rather than primary legislation, which in some cases appears to have resulted in incorrect or unknown enactment dates. In the following section, I find that the data display a number of inconsistencies, not only with my database, but also with secondary sources published before 1950. By locating the exact text of early legislation that the enactment dates are based on and verifying that there are no superseded laws, my approach improves the historical accuracy of even these early credentialing laws. Relative to the Council of State Governments data, I also fill in a number of unknown enactment dates and add observations for requirements adopted by the federal government, Alaska and Hawaii (then territories), and the District of Columbia.

2.4 Dataset Construction

This section provides a detailed description of the data collection procedure for this project. I begin by assembling a list of occupations known to be regulated in at least one state from dozens of contemporary and historical sources. Next, I verify whether states currently regulate each occupation by searching statutory codes and administrative regulations. I then manually record several key variables necessary to differentiate between licensing, certification, and registration requirements. Each regulation is subsequently linked to the full text and effective date of the original legislation that enacted it, as well as to any earlier laws or major amendments that altered the occupation’s regulatory status.

2.4.1 Identifying Regulated Occupations

In the first step of the data collection process, I digitize and clean a variety of independent sources that compiled information on professional and occupational regulations at various points in time from the late 1930s to the present. These secondary sources are almost all cross-sectional, and therefore insufficient to construct a complete timeline of regulatory changes for any individual occupation. Taken together, however, they provide the most comprehensive enumeration of regulated occupations presently available and a useful foundation for organizing the data. This step results in a crosswalk file that lists unique regulated occupations and assigns each to a code in the Standard Occupational Classification. It also produces a set of cleaned datasets based on the supplemental sources I collect for use in data validation exercises.

Contemporary Documentation. I begin with the CareerOneStop License Finder, an online database that publishes occupational credentialing requirements collected by state labor market information units under a grant from the U.S. Department of Labor.⁹ The data

⁹Recent papers using policy data from this database include [Blair and Chung \(2019\)](#), [Bailey and Belfield \(2018\)](#), and [Gittleman and Kleiner \(2016\)](#).

list approximately 10,000 state and federal regulations and provides a crosswalk file mapping each occupation title to detailed statistical codes based on the Standard Occupational Classification system. However, because the database is collected from secondary sources, it is not possible to consistently differentiate between various methods of regulation, and few observations report underlying legislative citations. Moreover, not all states participate in the data collection initiative, so the information provided may not accurately reflect regulation requirements in all jurisdictions, a point that will become important when discussing data validation in the following section.

Using the crosswalk provided with the CareerOneStop database as a starting point, I harmonize similar occupation titles to form an initial list of regulated occupations. Since many regulations cover jobs even more specific than the detailed six-digit codes defined in the Standard Occupational Classification system, I construct a eight-digit code that allows me identify distinct suboccupations. To do this, I manually work through the original crosswalk, creating new codes as necessary whenever a regulated occupation is a proper subset of an existing job category. In the process, I review and correct apparent errors in the original mapping as well. I then assign each occupation a standardized title and group codes into three tiers. First tier occupations are equivalent to an existing six-digit code. Second tier occupations may be equivalent to an existing code if two or more are regulated simultaneously, for example, school and public librarians. Third tier occupations are nested entirely within an existing code and are often found in residual or broadly-defined categories.¹⁰

I supplement the CareerOneStop database with additional contemporary documentation from the Institute for Justice ([Carpenter et al., 2017](#)) and the [National Conference of State Legislatures \(2017\)](#). Both datasets are derived from primary sources and report extremely detailed information on the requirements workers must meet to obtain a license, through

¹⁰Note that this procedure preserves the original SOC mapping except in cases where errors are identified. Illustrative examples are shown in [Table 2.1](#). In [Section 5](#), I show that my edits to the CareerOneStop crosswalk significantly improve the ability of the data to predict variation in credential attainment at the individual level.

for a comparatively small set of occupations. To incorporate the occupation titles used in these sources into the master crosswalk, I first apply an autocoding application developed for O*Net, the successor of the Dictionary of Occupation Titles, to assign each descriptive title to a six-digit SOC code. I then compare this mapping to my crosswalk and follow the same procedure described above to create new eight-digit codes as necessary.¹¹ Finally, I review the set of titles assigned to the same code in order to maximize consistency in the mapping across datasets.

Historical Documentation. Next, I digitize additional documentation that compiled data on occupational licensing and certification requirements at various points in time between 1939 and 2005. For simplicity, I refer to these sources as directories, but note that they include official government publications, law review articles, and numerous other policy reports. A handful, including a professional directory published by the Gale Research Corporation (Bianco, 1996), several studies conducted by the Council of State Governments (1952) and its affiliate, the Council on Licensure, Enforcement and Regulation (1990, 1994), and a report issued by the U.S. Department of Labor (Greene, 1969) attempted to catalog all state-level regulations in effect at the time of publication. Others, such as the Book of the States series, periodically tabulate credentialing requirements for a smaller set of commonly-regulated occupations.¹²

As the scope and coverage of these directories vary, they are generally not sufficient to pin down specific regulation dates, but nonetheless provide a useful supplement for historical

¹¹I update the crosswalk as necessary when new information about the correct mapping for certain occupation titles becomes available. For example, some titles are not specific enough to reliably place them in an SOC, but once I locate the statute covering the occupation, the full regulatory definition can often be used to find the correct code. This only affects a small number of titles since the vast majority of occupations have a clear and direct mapping to a six-digit code.

¹²These regulation studies include the U.S. Department of Commerce (1942), the U.S. Department of Health, Education, and Welfare (1967, 1977), the Council of State Governments (1968), Stokes-Berry and Roederer (1982); Stokes-Berry (1986); Stokes-Berry and Brinegar (1988); Brinegar and Schmitt (1992, 1994); Tracy (1998); Wall (2000), and Brinegar (2004, 2005, 2006).

data collection.¹³ After converting the data to text, I follow the same procedure described above for the NCSL and IJ data to assign each descriptive occupation title to an eight-digit code, expanding the master crosswalk file to include any occupation categories not already identified by contemporary sources. The final crosswalk file produced in this step identifies hundreds of occupations that are, or were, subject to some form of regulation in the United States. A cleaned version of the CareerOneStop database, recoded using this crosswalk, provides preliminary data on where each occupation is currently regulated. Finally, the processed directory files identify roughly when many regulations may have initially been adopted in each state.

2.4.2 Current Regulatory Requirements

In the second step of the data collection process, I compile primary legal documentation for each policy currently in effect. From the text of the law, I hand-code several variables that allow me to differentiate between licensing, certification, and registration requirements. At this stage, however, I do not assign an implementation date to these policies.

Compiling Statutes and Regulations. For each occupation that secondary sources indicate is regulated in at least one state, I preform a search of state (or federal) statutes and regulations using the LexisNexis legal database to confirm that a regulation does in fact exist. Since my secondary sources may not have identified all states that regulate the occupation, I repeat this process for each of the fifty states and the District of Columbia to minimize the probability of false negatives. I also search the websites of professional organizations whenever they are available as an additional check on the data, since these often compile their own fifty-state regulatory information.

In the current version of the database, I prioritize collecting data for first-tier occupations,

¹³I am able to locate only three sources, [Graves \(1938\)](#), the [Council of State Governments \(1952\)](#) and the [U.S. Department of Health, Education, and Welfare \(1967\)](#) that provide any information about the enactment dates of the statutes they identify.

as these are the only regulated jobs for which direct measures of wages and employment are available, as well as second and third-tier occupations in the same minor (three-digit) groups as these. I further focus on credentialing requirements that include competency standards, so occupations that are subject to registration requirements alone, such as commercial fishermen, casino workers, and various dealers in agricultural and manufactured products, do not currently appear in the data unless they are also licensed or certified in at least one state. I also exclude business licenses unless the credential applies to an individual rather than an establishment and requires the proprietor to meet specific training and examination requirements, such as a general construction contractor. Finally, I am unable to collect municipal credentialing regulations unless a state law specifically delegates the authority to license a particular occupation.

Coding Legal Provisions. After compiling primary legal documentation for each policy, I begin constructing the main database by manually coding several key variables that characterize the scope of the regulation. First, I note whether or not a credential is mandatory to practice by examining how the occupation is defined, what specific actions are considered unlawful, and if there are any major exemptions from the licensing requirement. I differentiate between two types of practice acts. The first, which I call complete, require all or most workers to eventually obtain a license. The second, which I call partial, only apply to workers in supervisory roles or those performing a subset of tasks within the occupation’s usual scope of practice.¹⁴ The extent of practice restrictions are typically a fixed characteristic of the occupation itself, and rarely vary across states except for minor differences in exemptions that are too specific to code.

Next, I create a variable that records the strength of any title protections specified by the

¹⁴This type of partial licensing is a distinct concept from the the partial coverage of a statistical occupation code discussed above because even when practice restrictions do not cover all workers, title restrictions may still apply broadly. For example, I consider accountants a tier-one occupation since accounting practice acts occasionally restrict even advertising oneself an “accountant” without a license, even though many accounting tasks can be legally preformed by an unlicensed worker.

law. As with practice restrictions, workers who have not obtained a credential and advertise themselves using restricted terms are committing a criminal offense and face potential fines or even jail time. Weak title protections cover only the use of modifiers such as “licensed” or “certified,” while stronger protections limit full, unmodified titles. For example, Georgia protects only the title “registered interior designer” and is coded as having weak title protection, while Florida, which has a practice act, restricts the use of the title “interior designer” in any form and is coded as having strong title protection in addition to complete practice restrictions. For the vast majority of observations, there is a clear distinction between practice and title acts. In the small number of cases where the restriction is ambiguous, I code the strongest implied requirement and create a data quality flag to identify these observations.

To differentiate between licensing and registration requirements, I then record whether the law requires a demonstration of competency. This may include passing an exam, meeting a predetermined level of education beyond high school, fulfilling an experience requirement, or obtaining certification from a non-governmental organization. I do not require that these qualifications be met before obtaining the credential, only that compliance is eventually necessary to remain in good legal standing. Minor training requirements that take less than a day to complete are not considered a demonstration of competency for my purposes. Finally, I note the level of government issuing the credential and whether the relevant law appears in statutes or administrative code.

2.4.3 Historical Policy Changes

The third and final step of the data collection process constructs the history of each policy by assigning implementation and amendment dates. This involves locating the first law to regulate the occupation, repeating the coding procedure to identify practice and title restrictions, then searching for additional amendments as necessary if a jurisdiction’s initial regulatory approach differs from the one currently in effect.

Compiling State Session Laws. To link each policy to the underlying legislation implementing it, I turn to the HeinOnline database of state session laws, a fully-digitized collection of laws passed during nearly every legislative session in the history of the United States. As in the previous step, I preform a search for references to each occupation’s title by state to identify the first law (if any) that adopted a credentialing requirement. I then code the same set of variables described above and compare them to the current law. If there are any differences, I continue with a chronological search of legislation, recording any session laws that affect the key variables until I locate the amendment(s) that enacted the current requirement. I then note both the enactment and statutory effective dates of all major legislation to produce a list of events that completely characterize how the key variables have changed over time. At this point I also search for any secondary sources reporting the regulatory history of the occupation to aid in this process.

Additional Data Sources. Historical data on federal regulations also come from HeinOnline, which includes both annual versions of the Code of Federal Regulations from 1938 to the present, and the U.S. Statutes at Large dating back to the 1700s. I identify laws adopted in Washington, D.C. by searching the District of Columbia Register. Finally, I note that due to the absence of historical administrative regulations, it is much more difficult to determine when the small number of credentialing requirements appearing only in administrative code were adopted. Whenever possible, I rely on secondary sources to determine these enactment dates on a case-by-case basis. If any relevant policy change dates cannot be determined, I code the regulation as having an unknown adoption date. This affects less than one percent of the database.

2.5 Data Validation Exercises

This section discusses the coverage and accuracy of the regulation data I collect. I first cross-validate the session laws database with the professional directories described in the

previous section. Next, I use recent waves of the Current Population Survey to estimate the share of workers holding government-issued credentials who work in occupations included in the policy data. I then assess the relationship between regulation coverage and variation in reported credential attainment across states.

2.5.1 Professional Directories

Coverage of Regulated Occupations. For each secondary source of policy data, [Table 2.2](#) reports the number of unique occupation titles and regulations identified by the directory and the share of these observations that are covered by the session laws data. Column one lists the number of unique directory titles prior to standardization. Column three reports the total number of regulated occupation-state pairs, which adjusts for the fact that the majority of occupations are not regulated in every state or may be regulated under different titles. The session laws data cover approximately 70% of the regulations identified in the most comprehensive of these sources, with higher coverage of observations in the more limited reports.¹⁵ Occupations that are not covered are typically regulated in only a handful of states, and are more narrowly defined than those identifiable at the six-digit level, implying that they each affect only a small number of regulated workers.

Consistency of Regulation Dates. The directories also provide an opportunity to assess the accuracy of the policy enactment dates collected from state session laws. After standardizing occupation titles, I first define an occupation-state pair as regulated in directory d if the directory identifies any licensing, certification, or registration requirements for the pair. I then match these directory indicators to occupations with known regulation

¹⁵Coverage of the Institute for Justice’s License to Work report ([Carpenter et al., 2017](#)) is an outlier. This is because the License to Work report focuses largely on specialty contractor licenses and differentiates between contractors in commercial and residential work. Since these licenses apply only to the contractor, rather than the tradesmen they employ, contracting occupations do not cover complete occupation codes. Further, classifications of specialty contractor licenses are typically adopted through administrative regulations rather than statutory law, making it extremely difficult to date when the regulations were initially implemented.

dates in the main policy data, where I define regulation as any practice or title restrictions, including those that do not require a demonstration of competency.¹⁶ Table 2.3 summarizes the conflicts between these datasets. The first thing to note is that the directories themselves contain a number of potential inconsistencies. Column one reports the share of observations that are either regulated in d but unregulated in $d - 1$ and $d + 1$, or unregulated in d and regulated in $d - 1$ and $d + 1$. Fortunately, the share of conflicts is relatively low and exceeds 4% for only 3 of the 21 directories.¹⁷

Column two reports the share of occupation-state cells that are identified as regulated in the directory data, but unregulated in the main policy data as of the year the directory was published. These “false negatives” are the most concerning type of error, since they imply that these units may have been treated earlier than can be identified from the session laws alone. This is possible, for example, if another regulation was sufficiently broad to cover an occupation without explicitly specifying its title. If this is the case, the share of false negatives provides some information about the extent of measurement error in treatment timing. It is reassuring, therefore, that false negatives are rare, accounting for fewer than 2% of observations in most directories, including in the most comprehensive. Column four repeats this comparison, excluding directory observations identified as potential conflicts. As expected, this lowers estimates of the false negative rate, which uniformly falls below 3.5% in this comparison. Across all directories, approximately 5% of state-occupation cells in the session laws data have at least one potential false negative.

Columns three and five report another type of inconsistency, where an occupation-state cell is identified as regulated in the main policy data, but not in the directory data. These “false positives,” though much more common than false negatives, are less concerning since

¹⁶Due to inconsistencies in terminology, the directories cannot be used to accurately differentiate between various types of credentials. I therefore use a broad definition of regulation in this exercise.

¹⁷To avoid over-counting conflicts, I compare only the previous and subsequent directories that an occupation appears in. For example, if the sequence $\{0, 0, 1, 1, 0\}$ is observed, this method identifies the fifth observation as a conflict. This is intuitively more reasonable than treating the third and fourth as conflicts as well.

any observation in the policy data is based on a session law with a verified enactment date. False positives are therefore likely to reflect omissions from the directories rather than errors in the main data, a point I return to below. Overall, this table shows a high degree of consistency between the data used in this project and that collected from numerous other independent sources, though some measurement error in treatment timing is clearly possible. To ensure that analysis using the data are robust to these inconsistencies, I create a set of sample flags identifying cells associated with probable errors that can be use to drop these observations.

2.5.2 Regulation and Worker Credential Attainment

The second set of validation exercises I conduct assess the relationship between state regulations and self-reported credential attainment at the individual level. I use data from the Current Population Survey ([Flood et al., 2018](#)), which since 2015 has asked respondents whether they hold an active professional certification, state, or industry license, and if so, whether it was issued by a government agency or is required for their main job. Before introducing the validation tests, it is important to note that, as the CPS is a survey, credential attainment may be measured with error. For example, a pilot study conducted in 2010 to evaluate the proposed CPS questions found that 14 percent of respondents who were known to hold a license based on administrative data did not report having a credential when surveyed ([Allard, 2016](#)). Even in the most highly-regulated occupations, therefore, licensing rates in the CPS may not perfectly correlate with policy differences.

To illustrate this point, [Figure 2.1](#) plots the share of respondents who reported holding any professional credential by state-occupation cell for a subset of occupations that are regulated in all states. The data show significant variation in credential attainment, despite the sample being limited to occupations that are universally licensed. Focusing first on the variation between occupations, average credentialing rates range from approximately 30% for accountants and auditors to 95% for chiropractors. Lower credentialing rates are expected

for many of these jobs, which may only require a worker to obtain a license if they preform certain tasks. Indeed, occupations subject to this type of partial licensing are clustered at the bottom of the distribution. For others, such as physicians and licensed practical nurses, however, the lack of complete credential attainment is more likely the result of measurement issues.¹⁸

More importantly, the figure also displays significant differences in credential attainment within occupations, the relevant margin of variation for the validation tests presented below. Regardless of the profession’s average credentialing rate, the share of workers reporting a license or certification typically differs by 20 percentage points or more across states, with zeros occasionally observed in small cells.¹⁹ Nonetheless, assessing the relationship between regulation coverage and credential attainment is still a useful validation exercise given the relative scarcity of data on professional credentials. With the caveats discussed above in mind, I first use the CPS to estimate the share of the share of regulated employment covered by the session laws data, then show that regulation strongly, though imperfectly, predicts the probability that workers in non-universally licensed occupations hold a government-issued credential.

Coverage of Licensed Employment. Above, I found that the main data collected in this project covers approximately two-thirds of regulated occupation titles identified by independent sources. However, because employment is not uniformly distributed across occupations, this estimate does not necessarily reflect the share of licensed workers covered

¹⁸Not all unlicensed workers necessarily misreported their credentialing status. Some unlicensed physicians observed in the data, for example, are likely medical residents who have not yet received their license to practice independently (Cunningham, 2019). Errors in the assignment of occupation codes (Mellow and Sider, 1983) and observations of workers who are exempt from licensing requirements (Ingram, 2019) may also contribute to this pattern, though importantly, these channels are unlikely to reflect significant underlying policy differences.

¹⁹In many occupations, workers can also obtain private certifications that are not legally required. If the credential’s source is misreported, it may lead to implied licensing rates greater than zero in states that do not regulate the occupation. For example, 28 states do not certify or license dispensing opticians, but 18% of opticians report a government-issued credential in these states. See Furth (2016) for additional examples and a more detailed discussion of the CPS data.

by the data. For that reason, I turn to the CPS to produce employment-weighted coverage estimates. To do this, I first use the complete list of regulated occupations from the directory data to identify all known regulated occupation titles that map into each three-digit CPS occupation code. Because it is not possible to assign workers to sub-occupations below the three-digit level, I report upper and lower bounds for coverage by noting whether all, some, or none of the regulations affecting each occupation category are included in the session laws data.

Table 2.4 reports these coverage estimates, as well as the share of individuals in the CPS who hold a government-issued credential by major occupation group. Column four presents a lower bound on the coverage share. In this case, a three-digit occupation is considered covered by the main data only if all regulations that map into it are known, including those below the six-digit level. If the directories identify any regulation mapping into the code that is not included in the policy data, it is not considered covered. Column five reports an upper bound that treats a three-digit occupation as covered if any policy mapping into it is known.²⁰ These estimates show that 60-70% of certified or licensed employment is covered by this paper’s session laws data, which is comparable to the share of occupation titles covered. Coverage of the highly-regulated healthcare, legal, and education sectors is nearly complete, while it is poor for occupation groups where regulation is uncommon such as agricultural and food service workers, or where regulations fall below the six-digit level, as with many managerial jobs.

Does Regulation Coverage Predict Credential Attainment? Next, I test whether the policy data I collect in this project can predict within-occupation variation in credential attainment across states at the worker level. I preform this analysis using the most detailed

²⁰For example, the six-digit codes “security guards” and “gaming surveillance officers” are combined into a single three-digit code in the CPS. Since both these occupations are regulated in some states, but policy data is only collected for the former, credentialed workers in this CPS occupation are considered covered in the upper bound calculation only. The upper and lower bound estimates are equivalent to making the extreme assumptions that within a thee-digit code any worker reporting a credential either works in the sub-occupation(s) included in the policy data, or those that are not.

level of occupation codes available in the sample. The policy variables are therefore defined as the share of a CPS occupation code covered by each type of regulation, and are constructed using national six-digit occupational employment share estimates from the Occupational Employment Statistics program. Due to a technical issue during data collection, the variable indicating whether or not a worker’s credential is required for their job is not available in 2015, limiting the sample to 2016 to 2018 in this exercise.

Using the CPS data, I estimate variants of the following linear probability model,

$$Y_{ijst} = \alpha_j + \sum_k \beta_k \times D_{jst}^k + X'_{ijst} \Lambda + \delta_t + \Omega_{jst} + \epsilon_{ijst} \quad (2.1)$$

where Y_{ijst} is an indicator equal to one if individual i reports holding a credential and $D_{jst}^k \in [0, 1]$ measures the share of their three-digit occupation j covered by regulation policy k in state s . Occupation fixed effects α_j sweep out between-occupation variation in regulation and credential attainment, implying that $\hat{\beta}_k$ is identified from variation along the margin of occupations whose regulatory status differs across states. The vector X_{ijst} includes individual and state-level controls and the year fixed effects δ_t absorb common time trends across all states and occupations.²¹ Finally, Ω_{jst} controls for occupation-specific linear terms in the number of minor (below six-digit) regulations affecting the three-digit code. These do not contribute toward the identification of $\hat{\beta}_k$, but improve the precision of the main effect when included.

The results of these regressions are reported in [Table 2.5](#), where credential attainment is alternatively measured as holding any professional credential, holding a government-issued credential, and holding a credential that is required for the worker’s job (regardless of source).

²¹Note that because treatment occurs at the state and occupation level, this specification is not a difference-in-differences estimator. Rather, it uses cross-sectional variation in regulation status and credential attainment within three-digit occupation codes, after partialling out common time trends across all occupations. Although a difference-in-differences estimate would be informative about the change in the likelihood of credential attainment before and after a regulation is adopted, the CPS data has not been available long enough to estimate this first-stage effect with any precision for recent policy changes.

This allows for the possibility that workers misreport the type or source of their credential.²² In Panel A, I do not differentiate between certification, licensing, and registration policies, but define treatment as any policy being in effect. The point estimate in Column 1 indicates that fully regulating a state-occupation cell increases the probability of credential attainment by approximately 13 percentage points. Including state fixed effects and demographic, education, and experience controls has only a minor effect on this estimate. Point estimates are approximately 25% larger in columns 3-6, however, where the definition of a credential is limited to those issued by a government agency or required for the worker's job, as we would expect if these definitions effectively exclude some workers who hold a voluntary certification issued by a private organization.

Panel B of the table differentiates between different types of policies. All predict a statistically significant increase in the probability of worker credential attainment, and, as before, the effect is larger when the dependent variable is limited to government or required credentials. Across specifications, the point estimates are uniformly larger and more precisely estimated for licensing policies compared to state certification or registration, as expected given that the certification is legally voluntary and policies other than licensing are infrequently adopted by state legislators. The results show that being covered by a licensing regulation increases the average probability of reporting a government-issued credential by as much as 15 percentage points. Although this point estimate is somewhat small in absolute terms, the relative effect is reasonable given the likely measurement error in credential attainment and shows that differences in policy do indeed predict differences in credential attainment for workers in the same occupation.²³

²²In the Survey of Income and Program Participation, which has a much smaller sample, but more detail on the source of workers' credentials, I find that 15% of registered nurses report a credential issued by a professional association (as opposed to the state government), which they may have interpreted as the state's nursing board.

²³For example, the CPS data show that 70% of radiologic technologists and technicians report a government-issued credential in states that currently license this occupation, compared to 55% in states that have no standards for radiography, a difference approximately equal to the estimates found in this exercise.

The only comparable estimate in the literature that I am aware of is reported by [Kleiner and Soltas \(2020\)](#) who use a subset of policy data on 55 occupations from the Institute for Justice and the National Conference of State Legislatures. They find that regulation of these occupations predicts a 6.6 percentage point increase in credential attainment. I also follow Kleiner and Soltas by reporting semiparametric estimates of the relationship between credential attainment and regulation coverage in [Figure 2.2](#). This figure limits the sample to occupations with state-level policy variation and plots a local-linear regression of the regulation coverage variable on the share of workers reporting a government-issued credential by state-occupation cell, after partialing out state and occupation fixed effects. It shows that cells where a 10 percent higher share of workers report a credential are approximately 7 percent more likely to be covered by some form of occupational regulation.

Comparison to the CareerOneStop License Database. Above, I found that the policy data collected in this project identifies regulations in a large number of state-occupation cells that were not reported in other sources. If these cases are indeed false positives, it raises serious concerns about potential measurement error in policy variation, which could contribute to the relatively low point estimates found above. Although the lack of historical data on credential attainment makes it impossible to assess the data in earlier years, in this section, I replicate the CPS validation exercise using alternative regulation measures obtained directly from the CareerOneStop licensing database. Doing so serves as an additional validation check on the accuracy of the policy data, and provides an alternative benchmark for the relationship between regulation and reports of credential attainment in the CPS.

I create two policy variables using the CareerOneStop data. For the first, I use the database’s original crosswalk, which maps each regulation to a six-digit occupation code in the 2010 Standard Occupational Classification system. I then treat a state-occupation cell as regulated if the database contains at least one observation for the pair. This measure therefore uses only the original CareerOneStop data. For the second measure, I use the

cleaned version of the database that corrects apparent errors in the crosswalk, differentiates between regulations that cover a full or partial occupation code, and imposes the same sample selection criteria as the main policy data. I then aggregate both variables to CPS occupation codes using the same method as above. To obtain comparable definitions of regulation, I do not differentiate between licensing, certification and registration policies in this exercise, since it is not possible to make this distinction in the CareerOneStop data. I then estimate Equation 1 including the full set of controls, which is the specification used to produce the estimates in the even-numbered columns of [Table 2.5](#), Panel A.

[Figure 2.3](#) displays the results of this exercise. It is clear that the original CareerOneStop data preforms poorly in both absolute and relative terms. Although it predicts a statistically significant difference in credential attainment between regulated and unregulated cells, the point estimates are extremely small and fall below 1.5 percentage points for all credential definitions. Cleaning the data improves its accuracy substantially, raising the point estimates to approximately 5 percentage points. However, this is still less than half estimate obtained from this paper’s data, suggesting that the inconsistencies found in [Table 2.3](#) likely reflect measurement error in the CarrerOneStop data rather than the other way around. These results provide further evidence that this paper’s data substantially improves on existing data sources and accurately classifies state-level policy variation.

2.6 Conclusion

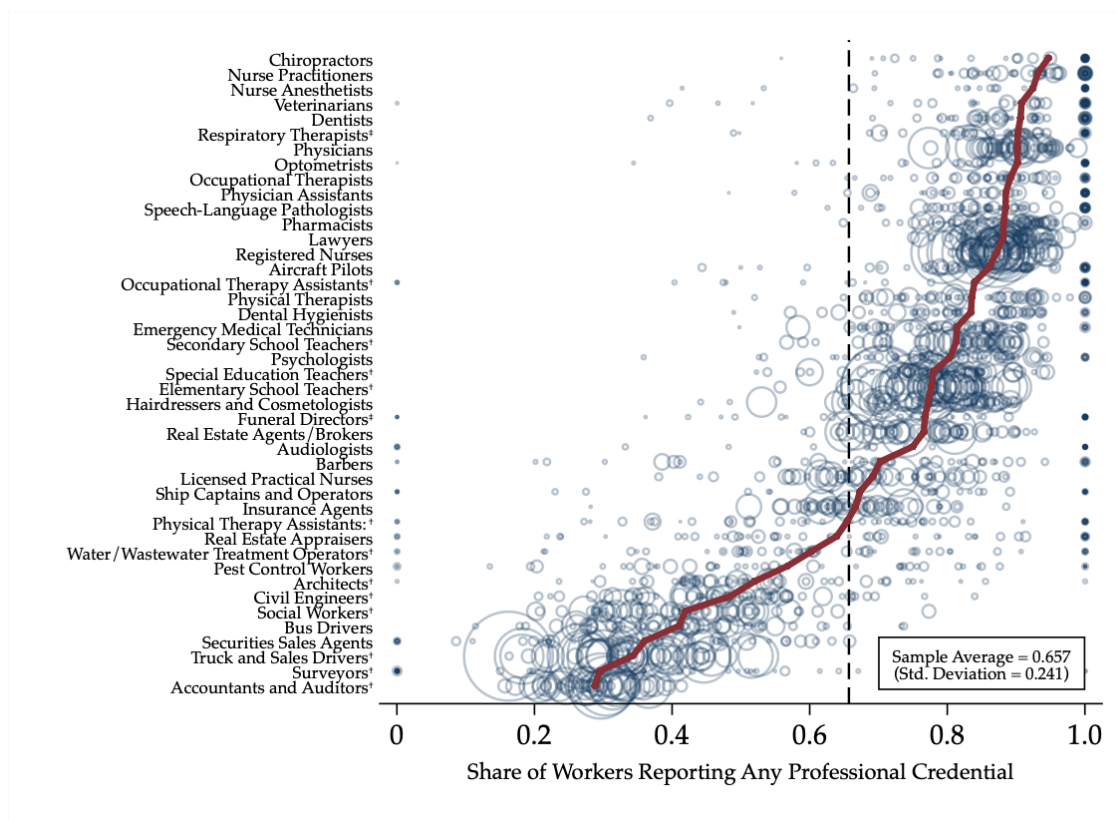
In this chapter I present a novel database that for the first time compiles a comprehensive history of professional and occupational credentialing regulations for a broad set of occupations. The data collection methodology I introduce overcomes the limitations of existing approaches by allowing me to observe the full text of all state legislation exactly as it appeared when enacted. As a result, I am able to locate superseded laws and trace the statutory development of each policy from its origin. My data are also the first to differentiate between

alternative approaches to regulation with differing legal restrictions and economic implications. I create a classification system that preserves unique regulated occupations as the unit of observation, while also mapping each job to statistical occupation codes that can be used to merge the policy data with external datasets. Finally, I digitize and clean numerous secondary data sources that provide independent snapshots of the regulations in effect at various points in time.

Although the dataset I create provides the most comprehensive history of occupational licensing and certification requirements assembled to date, it has several limitations. First, the data do not yet cover all regulated occupations. Further, I do not record the various levels of credentialing that may be present within an occupation. Workers will sometimes require one license at the entry level, but need additional endorsements to preform specialized tasks or practice independently. For example, there are numerous license classifications for emergency medical technicians, but it is beyond the scope of this project to distinguish between them. More importantly, I do not record the specific requirements necessary to obtain a credential, or how these have changed over time. It is well known that competency standards also differ across states conditional on regulation and would better characterize variation in the intensive margin of regulation.

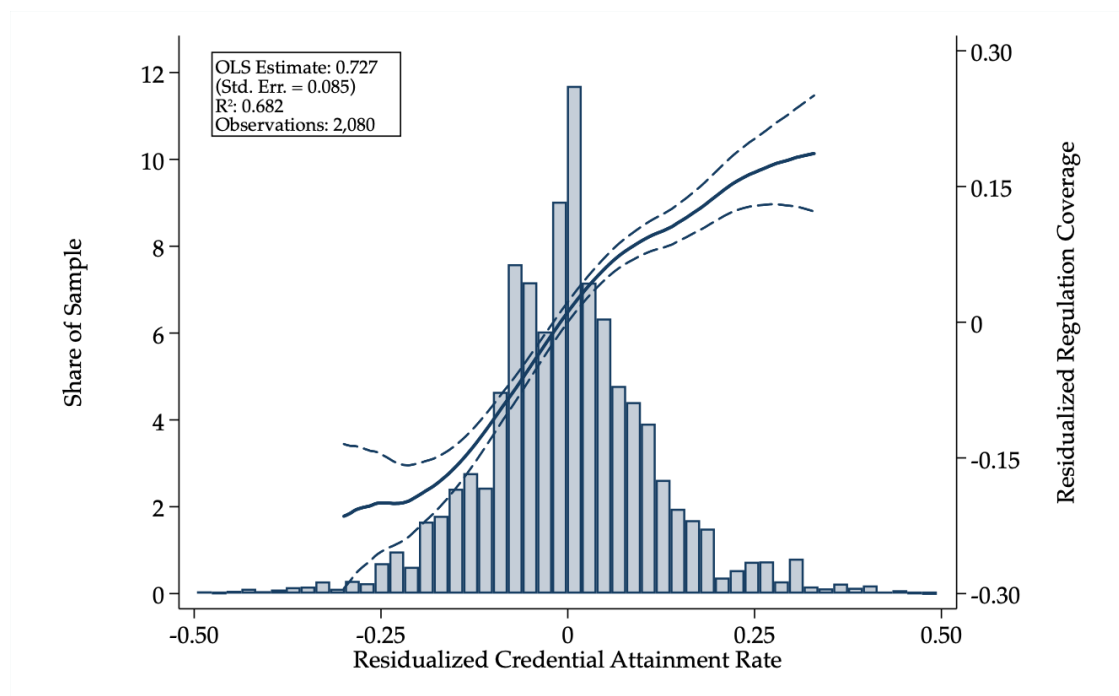
The data collection project described in this paper is ongoing, and will eventually be expanded to address these limitations. I am currently in the processes of extending the database to cover the remaining set of occupations identified by secondary sources. The data collection methodology I introduce is novel in the occupational licensing literature and in principle could be used to reconstruct each statute exactly as it would have appeared every year since it was initially adopted. The dataset could then be augmented to include a richer set of variables describing the intensive margin of regulation, as well as other provisions such as the composition of board members, reciprocity agreements between states, and grandfathering provisions for which there is currently little or no historical data.

Figure 2.1: Share of Workers in Universally Licensed Occupations Reporting Credential Attainment by State (CPS 2015-2018)



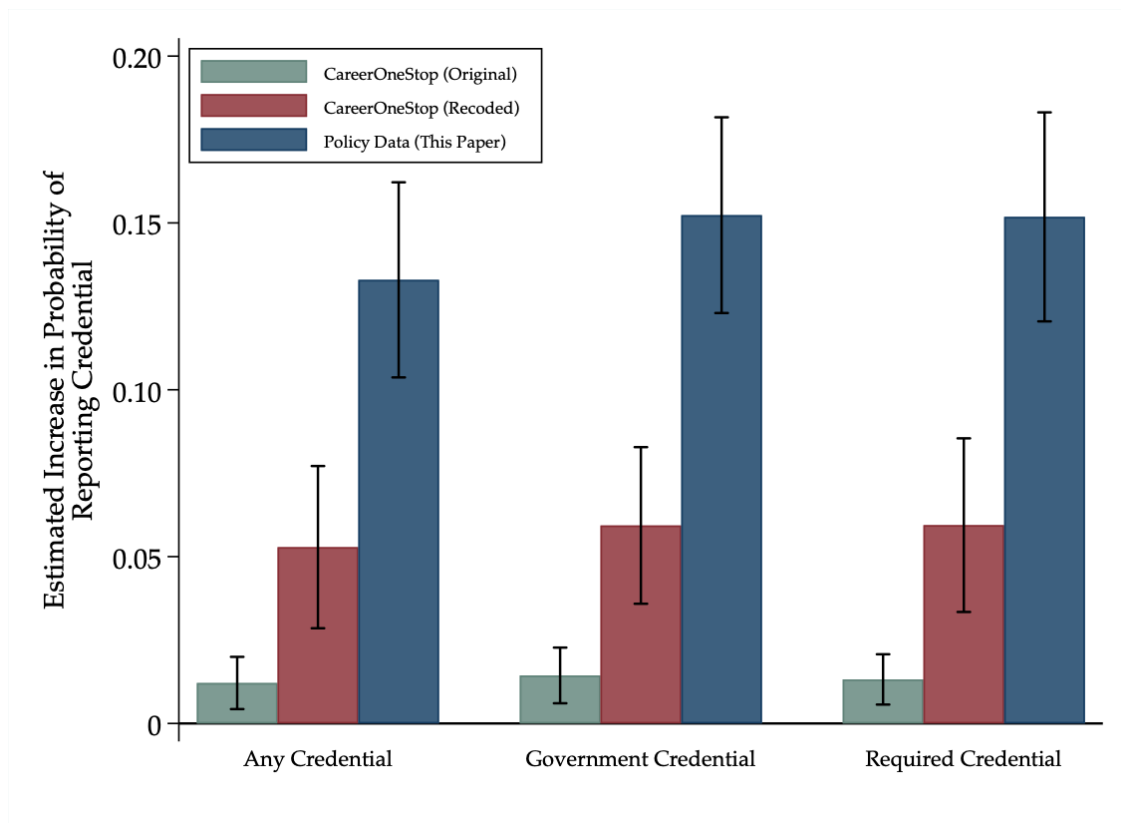
Notes: This figure displays the sample-weighted share of workers reporting any professional credential attainment by state and occupation for a subset of occupations that are regulated in every state. The sample is limited to employed respondents between the ages of 16 and 64 in the 2015 to 2018 Current Population Survey outgoing rotation groups. Observations with imputed occupation or credentialing status are excluded. The solid red line displays the mean level of credential attainment by occupation, and the dashed black line the average across all occupations in this sample. Circular markers show the share of workers reporting any professional credential by state-occupation cell, where marker sizes are proportional to the number of individual-level observations used to construct the state-occupation means. [†] Policy toward this occupation does not vary across states, but only workers performing certain tasks (or working in the public sector) must be licensed. [‡] Respiratory therapists are not regulated in Alaska, and funeral directors are not licensed in Colorado, both of which are excluded from the sample.

Figure 2.2: Semiparametric Estimates of the Relationship between Credential Attainment Rates and Regulation Coverage (CPS 2016-2018)



Notes: This figure presents graphical evidence of the relationship between credential attainment rates and regulation coverage using data from the 2016 to 2018 Current Population Survey. The figure is constructed by first selecting three-digit occupations with state-level policy variation during the sample window and computing the share of workers reporting a government-issued credential by state-occupation cell. The histogram then plots the cell-level credential attainment rate after residualizing with respect to state and occupation fixed effects. The solid line is a local linear regression of the residualized regulation coverage variable on the residualized credential attainment rate, weighted by cell size. Dashed lines show the 95% confidence interval for these estimates, with standard errors clustered by state. The figure also reports estimates from the corresponding two-way fixed effect OLS regression.

Figure 2.3: Effect of Alternative Measures of Regulation Coverage on Credential Attainment



Notes: This figure displays the estimated effect of regulation coverage on the probability that individuals report having attained a credential using data from the Current Population Survey matched to the main policy data used in this paper and the CareerOneStop license database. The original CareerOneStop regulation measure treats any state by occupation cell as regulated if it appears in the database. The recoded measure uses a cleaned version of the database as described in the text. All specifications and samples correspond to those used in the odd-numbered columns of [Table 2.5](#), Panel A. Bars denote 95% confidence intervals for the point estimates with standard errors clustered by state.

Table 2.1: Examples of Occupation Title Harmonization and Crosswalk Construction

Original CareerOneStop Title	New Project		2010 SOC		Standard Occupational Classification Title	Coverage Tier
	Code	New Occupation Title	Code			
Podiatrist	29-1081.00	Podiatrist	29-1081	Podiatrists	Tier One	
Podiatric Medicine, Doctor Of	29-1081.00	Podiatrist	29-1081	Podiatrists	Tier One	
Limited Podiatry Tmp	29-1081.00	Podiatrist	29-1081	Podiatrists	Tier One	
Speech Pathologist	29-1127.00	Speech-Language Pathologist	29-1127	Speech-Language Pathologists	Tier One	
Speech-Language Pathologist - Intern	29-1127.00	Speech-Language Pathologist	29-1127	Speech-Language Pathologists	Tier One	
Speech Pathologist & Audiologist	29-1127.00	Speech-Language Pathologist	29-1127	Speech-Language Pathologists	Tier One	
Speech Pathologist & Audiologist	29-1811.00	Audiologist	29-1181	Audiologists	Tier One	
Manicurists	39-5092.00	Manicurist	39-5092	Manicurists and Pedicurists	Tier One	
Nail Technician	39-5092.00	Manicurist	39-5092	Manicurists and Pedicurists	Tier One	
Cosmetology: Nail Technology	39-5092.00	Manicurist	39-5092	Manicurists and Pedicurists	Tier One	
Certified Public Librarian	25-4021.01	Public Librarian	25-4021	Librarians	Tier Two	
School Library/Media Specialist	25-4021.02	School Librarian	25-4021	Librarians	Tier Two	
Security And Fire Alarm Installers	49-2098.01	Fire Alarm Installer	49-2098	Security and Fire Alarm Systems Installers	Tier Two	
Security And Fire Alarm Installers	49-2098.02	Security Alarm Installer	49-2098	Security and Fire Alarm Systems Installers	Tier Two	
Fire Alarm Technician	49-2098.01	Fire Alarm Installer	49-2098	Security and Fire Alarm Systems Installers	Tier Two	
Security Alarm Installer	49-2098.02	Security Alarm Installer	49-2098	Security and Fire Alarm Systems Installers	Tier Two	
Art Therapist	29-1129.01	Art Therapist	29-1129	Therapists, All Other	Tier Three	
Art Therapist, School	29-1129.01	Art Therapist	29-1129	Therapists, All Other	Tier Three	
Dance Therapist	29-1129.02	Dance Therapist	29-1129	Therapists, All Other	Tier Three	
Dance Therapy	29-1129.02	Dance Therapist	29-1129	Therapists, All Other	Tier Three	
Music Therapist, Working In Schools	29-1129.03	Music Therapist	29-1129	Therapists, All Other	Tier Three	
Music Therapy	29-1129.03	Music Therapist	29-1129	Therapists, All Other	Tier Three	
Home Inspector, Certified (Ccb)	47-4011.01	Home Inspector	47-4011	Construction and Building Inspectors	Tier Three	
Building Inspector, Level I	47-4011.02	Building Inspector	47-4011	Construction and Building Inspectors	Tier Three	
Commercial Electrical Inspector	47-4011.05	Electrical Inspector	47-4011	Construction and Building Inspectors	Tier Three	
Local Plumbing Inspector, Certified	47-4011.09	Plumbing Inspector	47-4011	Construction and Building Inspectors	Tier Three	

Notes: This table provides illustrative examples of how the descriptive occupation titles appearing in the CareerOneStop database are assigned to eight-digit occupation codes based on the Standard Occupational Classification system. Note that the legal definition of manicurists includes pedicurists, which is why this occupation is assigned to tier one rather than tier two.

Table 2.2: Comparing Coverage of the Session Laws Data to Alternative Sources

	Unique Titles in Directory (1)	Share of Titles Covered by Data (%) (2)	Total Regulations in Directory (3)	Share of Regulations Covered by Data (%) (4)
<i>Panel A: Comprehensive Sources</i>				
CareerOneStop License Finder (2018)	5,377	64.92	8,989	66.35
Gale Research Inc. (Bianco 1996)	2,042	57.84	6,122	68.08
CLEAR (1994)	699	58.66	5,721	73.47
CLEAR (1990)	803	55.17	4,530	77.53
Department of Labor (Greene 1969)	723	39.97	2,969	67.33
Council of State Governments (1952)	69	78.26	1,209	96.28
<i>Panel B: Book of the States</i>				
Brinegar (2006)	63	95.24	2,469	97.29
Brinegar (2005)	64	95.31	2,443	97.46
Brinegar (2004)	58	98.28	2,336	98.63
Wall (2000)	58	98.28	2,134	98.64
Tracy (1998)	58	98.28	2,123	98.63
Brinegar & Schmitt (1994)	59	98.31	2,100	98.62
Brinegar & Schmitt (1992)	60	98.33	2,261	97.79
Stokes-Berry & Brinegar (1988)	42	100.00	1,322	100.00
Stokes-Berry (1986)	30	100.00	1,052	100.00
Stokes-Berry & Roederer (1982)	81	81.48	2,277	84.58
<i>Panel C: Other Sources</i>				
NCSL Licensing Database (2019)	31	96.77	1,381	97.39
Institute for Justice (Carpenter et al. 2017)	96	45.83	2,577	43.81
U.S. Public Health Service (1977)	31	87.10	892	97.09
Council of State Governments (1968)	130	63.08	1,782	86.42
U.S. Public Health Service (1967)	25	84.00	782	96.16
Department of Commerce (1942)	29	93.10	930	97.10
Graves (1939)	21	95.24	831	96.99

Notes: This table displays the number of occupational regulations appearing in various contemporary and historical sources and the share of these that are covered by the main policy data created in this project. Due to inconsistencies in reporting, federal licenses are not included in this table. Business, gambling, racing, and martial arts licenses are excluded from all samples. Unique titles refer to the original occupation name listed in the directory prior to standardization. Total regulations count the number of state-occupation pairs listed in the directory.

Table 2.3: Assessing the Consistency of Regulation Dates with Alternative Sources

	Share Directory Conflicts (%)	Share of Observations Conflicting with Main Policy Data			
		Baseline Comparison		Excluding Conflicting Data	
		Share False Negatives (%)	Share False Positives (%)	Share False Negatives (%)	Share False Positives (%)
<i>Panel A: Comprehensive Sources</i>	(1)	(2)	(3)	(4)	(5)
CareerOneStop License Finder (2018)	3.27	1.16	11.63	1.16	9.08
Gale Research Inc. (Bianco 1996)	3.15	1.65	11.10	1.06	9.85
CLEAR (1994)	3.45	1.68	9.03	0.98	7.96
CLEAR (1990)	5.01	2.80	12.19	1.92	9.50
Department of Labor (Greene 1969)	4.32	5.39	6.92	3.53	6.05
Council of State Governments (1952)	2.13	0.67	5.84	0.51	4.22
<i>Panel B: Book of the States</i>					
Brinegar (2006)	0.35	1.36	3.25	1.22	3.18
Brinegar (2005)	0.10	1.22	4.09	1.19	4.02
Brinegar (2004)	0.15	1.35	2.40	1.28	2.36
Wall (2000)	0.04	0.71	8.00	0.71	7.96
Tracy (1998)	0.08	0.86	7.32	0.86	7.24
Brinegar & Schmitt (1994)	0.68	0.86	5.18	0.86	4.58
Brinegar & Schmitt (1992)	2.80	4.47	2.36	3.31	1.71
Stokes-Berry & Brinegar (1988)	2.13	2.93	2.72	2.02	2.50
Stokes-Berry (1986)	1.38	2.63	2.82	2.32	2.63
Stokes-Berry & Roederer (1982)	4.89	2.76	9.36	2.31	6.87
<i>Panel C: Other Sources</i>					
U.S. Public Health Service (1977)	0.65	1.58	4.38	1.51	4.16
Council of State Governments (1968)	1.43	3.27	7.66	3.10	6.78
U.S. Public Health Service (1967)	1.24	1.34	0.67	1.05	0.48
Department of Commerce (1942)	0.23	0.93	3.40	0.77	3.40
Graves (1939)	0.00	1.14	7.49	1.14	7.49

Notes: This table compares the regulatory status of occupation-state cells appearing in various contemporary and historical sources to the main policy data used in this project. Due to inconsistencies in reporting, federal licenses are not included in this table. The sample is limited to occupation-state cells with complete data on policy changes. Column one reports the share of observations in each directory that are inconsistent with data from the others. An observation is considered conflicting if it either (i) appears as unregulated in directory d , but regulated in directory $d - 1$ and $d + 1$, or (ii) appears as regulated in directory d , but unregulated in directories $d - 1$ and $d + 1$. False negatives are defined as cases where a state-occupation cell is identified as regulated by the directory, but not the policy data. False positives are defined as cases where a state-occupation cell is identified as unregulated in the directory but regulated in the policy data. Columns two and three compare the policy data to the original directory data. Columns four and five exclude conflicting directory observations from the comparison.

Table 2.4: Estimates of Licensed or Certified Workers Covered by Regulation Data (CPS 2015-2018)

	Occupations in Regulation Data (1)	Credentialing and Employment in Current Population Survey Data		Share of Credentialed Workers Covered by Regulation Data	
		Share of Workers With Credential (%) (2)	Average Total Employment (1000's) (3)	Coverage Lower Bound (%) (4)	Coverage Upper Bound (%) (5)
All Occupation Groups	250	22.29	133,862	59.79	68.75
Healthcare Practitioners and Technical	71	73.78	8,143	98.39	99.28
Legal	9	62.16	1,526	99.57	99.57
Education, Training, and Library	13	53.51	8,123	69.34	83.70
Healthcare Support	25	47.32	3,182	41.88	99.48
Protective Service	4	37.23	2,757	39.86	57.40
Community and Social Service	17	33.17	2,306	87.05	87.05
Personal Care and Service	14	29.08	4,822	42.78	55.87
Life, Physical, and Social Science	18	22.60	1,274	63.07	71.53
Architecture and Engineering	11	21.36	2,801	88.55	88.55
Business and Financial Operations	16	19.57	6,552	53.07	53.07
Transportation and Material Moving	39	19.24	8,305	77.72	77.72
Management	19	19.08	15,195	17.58	41.88
Installation, Maintenance, and Repair	19	18.79	4,485	48.79	52.76
Construction and Extraction	32	16.87	7,161	60.24	60.50
Sales and Related	9	13.58	13,220	50.28	50.28
Office and Administrative Support	4	8.20	15,688	5.03	5.03
Production	6	8.03	7,674	16.41	16.41
Farming, Fishing, and Forestry	0	7.70	936	0.00	0.00
Arts, Design, Entertainment, Sports, and Media	7	7.54	2,773	41.01	41.01
Building and Grounds Cleaning and Maintenance	5	6.86	5,085	64.62	64.62
Computer and Mathematical	1	6.82	4,301	2.22	2.22
Food Preparation and Serving Related	0	6.65	7,553	0.00	0.00

Notes: This table reports estimates of the share of individuals who report holding a government-issued credential in the Current Population Survey who work in occupations covered by the main policy data used in this paper. The CPS sample is limited to employed workers, excluding those with imputed occupation or credentialing status. No policy data is available for the District of Columbia, which is dropped from the sample. Column one lists the number of unique occupational regulations included in the policy data by major occupation group. Column two reports the share of workers in the CPS who report a government-issued credential. Column 3 reports total annual average employment in these occupation groups. Column four gives a lower bound estimate of the share of credentialed workers covered by the policy data. A worker is considered covered in this case only if all regulations at the six-digit level and below that map into their three-digit occupation code are included in the policy data. Column five gives an upper bound estimate. Here, a three-digit occupation is considered covered if any regulations in the main policy data map into it.

Table 2.5: Effect of Regulation Coverage on Credential Attainment (CPS 2016-2018)

	Dependent Variable: Indicator for Worker Credential Attainment					
	Any Credential (1)	(2)	Government Credential (3)	(4)	Required Credential (5)	(6)
<i>Panel A: Pooled Estimates</i>						
Any Occupational Regulation	0.131*** (0.017)	0.133*** (0.015)	0.151*** (0.017)	0.152*** (0.015)	0.150*** (0.018)	0.152*** (0.016)
Observations	470,710	470,710	470,710	470,710	470,710	470,710
R-squared	0.294	0.319	0.306	0.329	0.354	0.373
<i>Panel B: Estimates by Regulation Type</i>						
Certification Policy	0.095*** (0.024)	0.103*** (0.022)	0.126*** (0.024)	0.133*** (0.023)	0.106*** (0.025)	0.112*** (0.024)
Licensing Policy	0.134*** (0.017)	0.135*** (0.015)	0.155*** (0.017)	0.156*** (0.015)	0.154*** (0.018)	0.156*** (0.016)
Other Policy	0.099*** (0.035)	0.103*** (0.028)	0.103*** (0.030)	0.107*** (0.024)	0.093** (0.038)	0.097*** (0.033)
Observations	470,710	470,710	470,710	470,710	470,710	470,710
R-squared	0.294	0.319	0.306	0.329	0.354	0.373
Occupation FE	✓	✓	✓	✓	✓	✓
Controls		✓		✓		✓

Notes: This table reports the effect of regulation coverage on the probability that individuals report having attained a credential using data from the outgoing rotation groups of the 2016 to 2018 Current Population Survey. The sample is limited to employed civilian adults between the ages of 16 and 64. Observations with imputed occupation or credentialing status are excluded. Data on licensing and certification policies occurs at the six-digit SOC level and is aggregated to three-digit CPS occupation codes using national employment share weights derived from the 2015 Occupational Employment Statistics survey. Controls include sex, fourth-order polynomials in age and potential experience, race, Hispanic origin, educational attainment, marital status, citizenship status, veteran status, indicators for local, state, federal, and self-employment, and state fixed effects. All regressions include occupation and year fixed effects and occupation-specific counts of minor (sub six-digit) policies. Regressions are weighted using CPS earnings supplement weights and standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

CHAPTER 3

The Impact of Occupational Licensing on Earnings and Employment: Evidence from State-Level Policy Changes

3.1 Introduction

Over 30 million workers in the United States – roughly a fifth of the total labor force – are required to obtain an occupational license before they may legally work or advertise services to the public. The number of jobs covered by state licensing statutes has risen substantially over the past half-century and now includes hundreds of distinct occupations, the majority of which require a license in some, but not all, states.¹ Growing awareness of these discrepancies, together with limited empirical evidence on the costs and benefits of licensing, have made this method of labor market regulation an increasingly prominent public policy issue ([Council of Economic Advisors, 2015](#)).²

Understanding how occupational licensing impacts earnings and employment in regulated occupations is essential to an evaluation of the trade-offs at the core of ongoing policy debates. On the one hand, licensing may raise barriers to entry, restricting the supply of labor and

¹These occupations include art therapists, dietitians, electricians, interior designers, locksmiths, massage therapists, plumbers, private investigators, radiologic technologists, and security guards, among many others. Occupations that require a license in every state include architects, attorneys, commercial drivers, emergency medical technicians, dentists, insurance agents, nursing home administrators, physicians, public school teachers, and registered nurses.

²In addition to the attention the topic has received from the White House, a number of states have recently taken steps toward regulatory reform by increasing recognition of out-of-state licenses, harmonizing credentialing requirements, or deregulating occupations entirely. See [Thornton et al. \(2017\)](#) and [Kilmer \(2018\)](#) for a discussion of these reforms.

increasing the market power of incumbent workers ([Friedman and Kuznets, 1945](#); [Stigler, 1971](#)). On the other hand, licensing may also increase demand for professional services by resolving informational asymmetries in markets where worker quality is otherwise difficult to observe and consumer health or safety is at risk ([Arrow, 1963](#); [Leland, 1979](#)). While both views predict that occupational licensing requirements will increase worker earnings, total employment in regulated occupations could rise or fall depending on the relative importance of these competing supply and demand channels.

Two important challenges, however, have hindered empirical research on occupational licensing. The first is the difficulty of assembling reliable data on the timing of state statutes. As a result, most researchers have either focused on a narrow set of occupations or adopted designs comparing licensed to unlicensed workers at a single point in time.³ Second, because any number of political or economic factors could plausibly affect both regulation and labor market outcomes, the causal effect of occupational licensing is difficult to discern. For instance, larger or more successful professional associations may be more influential when lobbying state legislators for favorable treatment. Moreover, the growth of licensing over time might correlate with other changes in state-level institutions or economic fundamentals that are difficult to net out without tracking both outcomes and policy changes for a large number of occupations over time.

In this paper I study the short- and long-run impact of occupational licensing on earnings and employment using newly-compiled data on the timing of major regulation events for over 200 unique occupations in all fifty states. This allows me to better exploit the rich cross-sectional and longitudinal variation in state laws covering the most economically significant

³The latter approach was introduced by [Kleiner and Krueger \(2010, 2013\)](#), who collected the first custom survey data on occupational licensing. Following their seminal work, questions related to licensing and certification have been added to both the Current Population Survey and the Survey of Income and Program Participation. Unfortunately, these questions were introduced too recently to capture meaningful changes in the prevalence of licensing within states and occupations over time, limiting designs using these datasets to cross-sectional comparisons.

occupations licensed during the past half-century.⁴ Importantly, my data are the first to link every policy to both a current statute or regulation, as well as to the full text of the underlying legislation that enacted, amended, or replaced any relevant laws. Relative to existing data on occupational licensing, my approach offers a complete panel of legal changes for many occupations, the ability to differentiate between licensing and weaker methods of occupational regulation, and greater historical accuracy. I match my policy data to several independent employer and household surveys and report findings based on labor market outcomes observed between 1980 and 2018.

My analysis exploits the fact that state legislatures typically pass licensing requirements for the same occupation at very different points in time, if ever. I adopt a difference-in-differences design that uses this staggered treatment timing to contrast changes in earnings and employment within detailed occupation categories across states.⁵ Since the impact of licensing may vary over time, I estimate an event study specification that traces out the full cumulative response of outcomes relative to the year of the policy change. As in other difference-in-differences designs, causal identification requires that earnings and employment in licensing states would have followed the trend observed for the same occupation in control states if the licensing requirement had not been enacted. To lend credibility to this assumption, my analysis flexibly controls for alternative methods of occupational regulation, heterogeneity in regional labor markets, and common state-level trends that may be correlated with the adoption of licensing statutes.

I first document that licensing has a clear positive effect on average hourly wages that increases with treatment duration when pooling across all occupations and policy changes.

⁴My estimation sample includes the near universe of licensed occupations with a precise definition in the Standard Occupational Classification system. These six-digit codes cover approximately two-thirds of all licensed workers.

⁵Fewer than a third of the occupations I study are currently licensed in all fifty states. Since there are no untreated units for these “universally-licensed” occupations, relative treatment timing is the only source of identifying variation. Variation for the remaining occupations comes from both relative treatment timing and comparisons between licensing and non-licensing states ([Goodman-Bacon, 2018](#)).

Twenty-five years after the adoption of a licensing requirement, cumulative wage growth in licensing states exceeded that of untreated controls by 5% on average. The magnitude of this intention-to-treat estimate is economically significant and shows that licensing does indeed have a meaningful impact on the labor market. In the long-run, the licensing wage premium is roughly equivalent in size to the wage differential that would be expected as a result of increasing the average educational attainment of all workers in the occupation by six months (Card, 1999). It is considerably smaller, however, than the 15% wage premium commonly attributed to union membership (Hirsch, 2004). In contrast to its notable long-run effects, licensing has little impact on average earnings in the short-run, most likely due to the grandfathering of incumbent workers (Han and Kleiner, 2017).

Despite experiencing a significant increase in long-run earnings, I find no evidence that employment in the typical licensed occupation fell relative to control states. My results show that adopting a licensing requirement had no effect on the number of workers employed in the short-run, and a weakly positive – but often statistically insignificant – impact in the long-run. After twenty-five years, point estimates range from 2 to 7% across various samples and rule out a long-run decrease in employment greater than -5% at conventional levels of statistical significance. I do find some evidence, however, suggesting that licensing may have reduced the share of self-employed workers over time. Nonetheless, the absence of large disemployment effects stands in sharp contrast to the view that the average occupational licensing statute increases earnings primarily through a reduction in the extensive margin of labor supply.

One potential explanation for these findings is that licensing simply induces higher-skilled workers to sort into the occupation without affecting net employment. Observable changes in the stock of human capital and demographic composition of workers, however, explains only a small fraction of the estimated wage premium. Although I find that licensing requirements have the expected effect of raising average educational attainment, this impact is modest and explains at most 15% of the long-run change in wages. Rather, exploring heterogeneity

in earnings and employment effects, I find evidence that is more consistent with licensing increasing demand for certain occupations. The data show that when the risk of consumer harm is serious, both earnings and employment increase following the adoption of licensing statutes. Where the consumer protection rationale for licensing is less plausible, the supply channel appears to dominate, increasing earnings, but significantly reducing the number of workers employed in the occupation.

This paper makes several contributions to a growing empirical literature on the labor market effects of occupational licensing. First, my analysis addresses an important trade-off between internal and external validity by analyzing a broad range of licensing policies. For the most part, other studies exploiting quasi-experimental variation in the timing of licensing statutes have focused on specific occupations. These include dental assistants ([Xia, 2020](#)), massage therapists ([Thornton and Timmons, 2013](#)), midwives ([Anderson et al., 2020](#)), nurses ([Law and Marks, 2017](#)), and opticians ([Timmons and Mills, 2018](#)), as well as a handful of professions licensed during the early twentieth-century ([Han and Kleiner, 2017](#); [Law and Kim, 2005](#); [Law and Marks, 2009](#)). While these studies have led to important insights, it is difficult to draw strong conclusions about the average impact of licensing requirements from these policy settings alone. My results, by contrast, provide more general evidence on the labor market effects of occupational licensing and an assessment of heterogeneity in earnings and employment responses across various occupation groups.

Second, I show that occupational licensing consistently raises earnings, despite no apparent decrease in employment for the typical occupation in my sample. Like this paper, a large number of studies have documented a substantial licensing wage premium, often using cross-sectional variation in self-reported license attainment from household survey data ([Blair and Chung, 2018](#); [Gittleman et al., 2018](#); [Ingram, 2019](#); [Kleiner and Krueger, 2013](#)). Comparing otherwise similar workers with and without a license, these studies have estimated wage effects ranging from 4 to 18%, which are larger on average than those I find. Often, this wage differential has been interpreted as indirect evidence that licensing restricts the supply

of occupational practitioners ([Council of Economic Advisors, 2015](#)). Direct evidence on employment effects, however, is more limited and the literature has yet to reach a consensus on the size or sign of these impacts.

In a recent paper, [Kleiner and Soltas \(2020\)](#) document that licensing increases wages by 15% while reducing employment by as much as 30% using cross-sectional variation in the share of workers reporting a license by state and occupation. In a related approach, [Blair and Chung \(2019\)](#) also find evidence of disemployment effects on the order of 20 to 30%. In contrast to these studies, I adopt a panel design that tracks earnings and employment over time and measures state-level policy changes directly. The difference in our empirical strategies offers one potential explanation for why my results – especially for employment – differ from those above. Further, my sample includes a number of occupations such as emergency medical technicians that became licensed in every state relatively recently. Estimates based on current data do not use variation from this type of occupation, which could also contribute to the differences in our findings.

The previous work most closely related to my own is [Redbird \(2017\)](#), who also uses state-level policy variation for many occupations to identify the impact of occupational licensing on labor market outcomes. Unlike this study, Redbird’s analysis found no effect of licensing on wages and a negative impact on educational attainment. By contrast, I find robust positive effects for both. Like Redbird, I find that the average occupational licensing statute may have had a weakly positive impact on employment, though my point estimates are considerably smaller. Taken together, Redbird interprets her results as evidence that licensing increases labor supply by drawing new workers into the occupation, whereas my findings are more consistent with increases in labor demand offsetting any negative impact of higher entry costs on overall participation. Although our research designs differ on a number of dimensions, the discrepancies in our findings are most likely explained by important differences in our policy data, which I outline in chapter two.

Finally, in addition to providing new evidence on the direct effects of occupational licens-

ing on earnings and employment, my results also contribute to a broader literature studying the role of labor market frictions in the United States. Given the prevalence of occupational licensing requirements across sectors and skill levels, several recent papers have cited licensing as a potential factor contributing to the declining U.S. employment-to-population ratio ([Abraham and Kearney, 2020](#); [Austin et al., 2018](#)) as well as falling occupational and geographic mobility ([Barrero et al., 2020](#); [Johnson and Kleiner, 2020](#); [Kleiner and Xu, 2020](#)). The macroeconomic implications of occupational licensing, however, remain largely unexplored. While a comprehensive general equilibrium analysis is beyond the scope of this paper, my results suggests that the impact of licensing on aggregate employment may be less pronounced than is commonly assumed.

The rest of this chapter is structured as follows. [Section 3.2](#) provides key details of the institutional background related to my research design and discusses the hypothesized effects of licensing on the labor market in greater detail. [Section 3.3](#) gives an overview of the data, including how my policy and outcome variables are constructed. [Section 3.4](#) discusses identification and introduces my empirical specification. [Section 3.5](#) presents my main findings on earnings and employment. [Section 3.6](#) assesses potential mechanisms and explores heterogeneous treatment effects. [Section 3.7](#) concludes.

3.2 Institutional Background

This section provides an overview of occupational regulation in the United States and discusses the main channels through which licensing may impact the labor market. Occupational regulation is a complex institution that involves multiple levels of government and at least three distinct regulatory approaches. Although this paper focuses on occupational licensing – the most prevalent and restrictive method of regulation – occupations licensed by some states at one point in time may be subject to weaker forms of regulation elsewhere. These policy alternatives generate additional variation in treatment status that I leverage in

my empirical analysis. Moreover, because licensing statutes usually contain broad exemptions for incumbent workers, policy changes may have little immediate effect, implying that a long-run analysis is necessary to fully characterize the impact of licensing on employment and wages.

3.2.1 Occupational Regulation

There are three general methods of occupational regulation that require workers to obtain government approval before they may legally practice or advertise their services to the public: licensing, certification, and registration ([Shimberg, 1980](#)).⁶ Most occupational licensing and regulation occurs at the state level, though a small number of occupations are regulated by federal agencies, individual municipalities, or multiple levels of government simultaneously ([Kleiner and Krueger, 2013](#)). This paper focuses on state and federal policies, which cover the vast majority of licensed jobs and workers.⁷ Because states often take different approaches to regulating the same occupation, understanding precisely how these policies differ is necessary for both the coding of my legal variables and ultimately the research design.

Methods of Regulation. Licensing, or formally “right-to-practice,” is the most restrictive form of occupational regulation. Policymakers adopting licensing statutes define a set of tasks that fall within the occupation’s scope of practice and make performing this type of work without meeting predetermined competency standards a criminal offense.⁸ The specific

⁶Occupational licensing is distinct from business licensing, such as a license to operate a dry cleaning shop or restaurant. Whereas a business license attaches to an establishment, an occupational license is a credential that individual workers must obtain to demonstrate that they are legally qualified for their job.

⁷[Table 3.1](#) displays summary statistics on the source and characteristics of worker credentials using data from the Survey of Income and Program Participation, the only large household survey that reports detailed information on how workers obtained their credentials. Less than one percent of the population holds a license or certification issued at the local level, and these account for only 2.3% of government-issued credentials overall. Omitting municipal licenses is therefore unlikely to have a significant impact on my results.

⁸Unlicensed practice is punishable by severe fines or even incarceration. While there is little data on enforcement actions, the majority of licensing boards are overseen by members of the regulated occupation itself, and have strong incentives to restrict unlicensed practice ([Allensworth, 2017](#)). Licensing boards

qualifications necessary to obtain a license vary by occupation and state, but applicants are usually required to demonstrate a minimum level of education or experience and pass an entry examination. Many licenses require periodic renewal, which may be accompanied by continuing education requirements.

Certification and registration are less restrictive alternatives to licensing. In certification statutes, the government establishes a “right-to-title” for workers who have obtained a credential. Only certified workers may legally use the reserved occupation titles in connection with their practice.⁹ Importantly, however, certification does not place limitations on the tasks that uncertified workers may perform, provided that they do not advertise these services using legally-protected terms. For that reason, state certification is sometimes referred to as “voluntary licensure” (Law and Marks, 2013). Finally, registration requirements are the weakest method of occupational regulation, under which workers must only file their name and qualifications with a government agency before beginning their practice. While registration is mandatory, workers are not required to demonstrate a minimum level of competency, though proof of bonding and insurance or criminal background checks may still be required.¹⁰

Implementation and Grandfathering. When occupational licensing policies are initially implemented, the minimum training and examination requirements they adopt may not apply to incumbent workers. Instead, licensing statutes usually include “grandfathering”

and professional organizations are also known to actively defend their scope of practice, including from encroachment by other occupations (Kleiner, 2016).

⁹Unlike the credentials and educational certificates issued by professional organizations (Deming et al., 2016), state certification is a direct method of labor market regulation with standards set and enforced by a government agency. Title protection statutes can also be quite expansive, limiting the use of unmodified occupation titles such as “accountant,” “interior designer,” or “psychologist” to workers with a credential approved by state regulators.

¹⁰Governments may also indirectly regulate occupations through deceptive trade practice acts, limitations on eligibility for public insurance reimbursement, or minor training requirements such as food safety courses that do not require formal government approval to work. Because these methods of regulation are not occupation-specific and are widely considered to be less restrictive than even registration requirements (Hemphill and Carpenter, 2016), I do not consider them in this paper.

provisions, allowing applicants who can demonstrate that they were working in the occupation before the law was approved to qualify for a license under relaxed standards ([Han and Kleiner, 2017](#)). Although complete exemptions from competency requirements are rare in modern statutes, licensing examinations are frequently waived if the incumbent otherwise qualifies by education or experience. Similarly, workers who do not meet the new educational standards may be allowed to substitute their work experience or continue practicing conditional on passing the examination. New workers, by contrast, must comply with all requirements of the law as soon as it becomes effective.

Grandfathering provisions imply that licensing requirements will have a greater impact on the flow of entrants than the stock of incumbents in the occupation. Because few workers are immediately forced to exit, changes in total labor supply depend on the entry decisions of new workers and may take several years to detect in aggregate. Likewise, average earnings will reflect the both the composition of workers through the stock of human capital, as well as any changes in the supply and demand for labor over time. As a result, the labor market may take many years to converge to a new steady-state, implying that the causal effect of licensing on earnings and employment may differ significantly in the short- and long-run. My empirical strategy therefore tracks outcomes up to twenty-five years after policy changes and places minimal structure on the evolution of the treatment effects to account for these stock-flow dynamics.

3.2.2 Potential Impacts on the Labor Market

As discussed by [Kleiner and Soltas \(2020\)](#), theory unambiguously predicts that earnings will rise following the adoption of licensing statutes, but the total number of workers employed in the occupation may rise or fall depending on the extent to which licensing shifts the relative supply and demand for labor. By mandating that workers demonstrate some minimum level of occupation-specific human capital, licensing requirements increase the cost of entry

for new workers, which is expected to decrease labor supply on the extensive margin.¹¹ In a frictionless labor market with complete information, supply effects dominate, average earnings rise, and incumbents covered by grandfathering provisions benefit from the decrease in competition at the expense of entrants and consumers (Friedman and Kuznets, 1945; Friedman, 1962; Stigler, 1971).

In markets with incomplete information, however, the negative effect of higher barriers to entry on labor supply may be offset by an increase in demand for the occupation's services, especially when consumer health and safety is at risk or employers are liable for harm caused by unqualified workers.¹² By establishing minimum qualifications for professional practice, licensing addresses the problem of asymmetric information by providing a credible signal of worker competency and increasing the average level of human capital in the occupation (Arrow, 1963; Akerlof, 1970; Leland, 1979). To the extent that consumers are willing to pay for these benefits, rising demand further increases earnings and dampens (or potentially reverses) any disemployment effects.

Even in markets where asymmetric information is not a concern, licensing requirements could impact earnings and employment through various institutional channels. For health care occupations in particular, the growth of licensing over time may reflect the specialization of labor and the emergence of new professions and technologies rather than rent-seeking by incumbent practitioners in established professions (Law and Kim, 2005; Lin, 2011). If the tasks preformed by emerging occupations overlap with the scope of practice reserved for existing occupations, licensing statutes may clarify potential regulatory issues and grant workers legal recognition as a distinct profession (Nunn and Scheffler, 2018). In some cases,

¹¹This view assumes that the standards set by regulators are at least as stringent as the minimum qualifications that would otherwise prevail in the market. Naturally, licensing will be less of a barrier to entry if it simply formalizes credentialing requirements that were already demanded by most employers (Redbird, 2017).

¹²As shown in Figure 3.1, the prevalence of occupational licensing is strongly correlated with the potential harm from worker mistakes. While individual licensing requirements often appear arbitrary, it is clear that licensed occupations differ significantly from unlicensed occupations on average.

licensure may even be a requirement for services to qualify for insurance reimbursement (Stokes-Berry and Roederer, 1982). If this is the case, licensing is expected to increase employment while earnings may rise or fall as new workers enter the occupation.

3.3 Data

In this section, I document the construction of my legal variables and describe the earnings and employment data used in the analysis. My policy data covers all fifty states and 250 unique occupations, of which 139 can be directly matched to a statistical code in the Standard Occupational Classification system. These occupations cover 180 detailed SOC codes (the near universe of occupations regulated at this level of aggregation), of which, 111 were licensed in fewer than fifty states by 1980 and potentially contribute identifying variation for my empirical analysis.¹³ The remaining occupations in the policy data are too small to exploit when estimating treatment effects, but are used to construct the control variables used in certain specifications.

This project is not the first attempting to assemble data on the enactment of occupational licensing policies (Redbird, 2017). My data are unique, however, on two important dimensions. First, I adopt a more detailed coding of specific legal provisions that allows me to differentiate between licensing requirements and less stringent methods of occupational regulation. Second, in addition to linking each policy to a current statute or regulation, I perform a comprehensive search of state session laws to gather the original text of all relevant historical legislation.¹⁴ Because statutes are frequently amended – or even replaced entirely

¹³The number of SOC codes covered is larger than the number of unique regulated occupations because some regulatory definitions, such as physicians, cover multiple statistical codes. I currently lack policy data for a small number of occupations that could potentially provide additional variation for my analysis. Together, these cover less than 15% of workers employed in licensed six-digit occupations.

¹⁴Session laws compile changes to statutes made during each legislative session. They include all new laws as well as amendments to existing code. On the extensive margin that I study, nearly all occupational licensing requirements are specified in statutory law. The small number that were adopted or overturned through administrative or case law are handled on a case-by-case basis with the aid of secondary sources.

– accurately determining when the regulatory status of a particular occupation changed is extremely difficult without this step. My approach allows me to observe superseded legislation, track exactly how each amendment modified the law, and construct a complete history of policy changes for each state and occupation.

3.3.1 Occupational Licensing Legislation

I compile data on occupational licensing and regulation from state and federal statutes, regulations, and sessions laws, then use this information to construct a balanced panel that assigns a treatment status to each state, occupation, and year cell between 1950 and 2018. An in-depth discussion of data collection and validation can be found in Chapter 2.

Sample Construction. To link the policy data to outcomes observed in employer and household surveys, I assign each regulated occupation a numerical identifier based on a refinement of the Standard Occupational Classification (SOC) system. The SOC defines over 800 occupation categories covering all civilian jobs and is the most detailed level at which the federal government collects wage and employment data. It is based on a six-digit classification system that groups occupations into major, minor, and detailed categories with similar work content. At the detailed level, these codes are sufficiently precise that the statistical and regulatory definitions of many occupations are nearly equivalent. For occupations that nest entirely within a six-digit code, I create a new eight-digit identifier to preserve the hierarchical structure of the SOC.

I use the underlying event data to create an annual panel recording whether each state-occupation cell is unregulated, licensed, certified, or registered in each year between 1950 and 2018. When an occupation is regulated at both the state and federal level, I assume that the federal law is binding unless the state law imposes more stringent requirements. The length

An entire state-occupation pair is excluded from the sample if the date of relevant policy changes cannot be determined.

of the regulation panel is chosen to reflect the possibility that licensing has long-run impacts on labor market outcomes. Because events occurring before the beginning of my outcome samples (typically in 1980) may still affect outcomes, assigning treatment and control groups in this setting requires observing each cell’s regulatory status at least twenty-five years before the beginning of the outcome window, which is the length of the event window I consider.

3.3.2 Earnings and Employment

Estimating the impact of occupational licensing on the labor market requires data on earnings and employment for detailed occupation categories over a relatively long time horizon. I therefore draw on three distinct datasets, the Current Population Survey, the Census and American Community Survey, and the Occupational Employment Statistics program to construct my estimation samples. Alone, each of these datasets have certain limitations that make it challenging to study the dynamic effects of licensing statutes. I show, however, that despite differences in sample coverage and aggregation, they produce remarkably similar treatment effect estimates. The consistency of my main findings across samples provides strong evidence that data quality concerns specific to particular surveys are not a major source of bias.

Current Population Survey. My primary source of data on earnings and employment are the Current Population Survey (CPS) outgoing rotation group extracts from 1983 to 2018 ([Flood et al., 2018](#)). These samples provide information on hourly wages for approximately 150,000 unique workers per year. The CPS is a relatively small survey, but it is the largest dataset where the occupational affiliation of workers is observed annually prior to 1999. Unless otherwise noted, the sample is limited to employed civilian adults between the ages of 16 and 64, excluding unpaid family workers. Wages are measured using the straight-time hourly wage for workers paid by the hour and usual weekly earnings divided by usual weekly hours otherwise. Observations with imputed earnings, hours, or occupation data are omitted

from any analysis using these variables.¹⁵ Although the CPS does not collect earnings information for the self-employed, I retain these observations when calculating employment counts.

Occupation codes in the CPS are based on Census classifications, which are revised every decade. To ensure that definitions and coverage remain comparable throughout the sample, I use a slightly aggregated version of the coding system developed by [Dorn \(2009\)](#) that extends the original balanced panel of occupations to incorporate additional revisions implemented after 2010. As described in [Section 3.8](#), these aggregations reduce the number of consistently-identifiable occupations from 330 to 310, but mainly affect unregulated production jobs. Data coded using Census classifications prior to 1980 cannot be fully accommodated in this system without substantial loss in detail, preventing the use of earlier data in balanced samples. I construct the main CPS estimation sample using earnings weights to collapse hourly wages to annual averages by state-occupation cell, the level at which treatment is assigned. Annual employment counts and worker characteristics are derived from the full monthly extracts.¹⁶

Census and American Community Survey. Due to its small sample size, a concern when using the CPS data is that some state-occupation cells may be extremely sparse. For that reason, I also use data from the decennial Census from 1980 to 2000 and the American Community Survey from 2001 to 2017 ([Ruggles et al., 2019](#)). These datasets are considerably larger, but present a trade-off between a reduction in cross-sectional sampling error on the one hand and potential data frequency issues related to the lack of annual data prior to 2000 on the other. As in the CPS, I limit the sample to employed civilian adults and aggregate the

¹⁵As shown in [Figure 3.14](#), imputation rates are significantly higher for wages than occupations, and have been increasing over time. A potential concern is that treatment may directly affect the imputation rate through changes in worker non-response. In [Table 3.16](#) and [Table 3.17](#) I replicate my main estimates with the imputation rate as the outcome variable and show that licensing is not related to changes in the share of imputed observations.

¹⁶Excluding imputed occupations results in an underestimate of total employment. I adjust for this bias using an approach similar to [Cengiz et al. \(2019\)](#) by first computing employment shares using non-imputed data, then scaling up the shares by total state employment to generate counts.

data to a balanced state-occupation panel using the same set of 310 consistent occupational codes discussed above. Although the CPS is my preferred source of data on hourly wages, the ACS-Census has the advantage of also recording total earned income, which is useful to verify that the main results are not driven by the omission of self-employment earnings.

Occupational Employment Statistics. Finally, I use data from the Occupational Employment Statistics (OES) program from 1999 to 2018. The OES is an employer-based survey that reports annual estimates of total employment and hourly wages for over 800 detailed Standard Occupational Classification codes.¹⁷ Surveyed establishments are drawn from the universe of firms covered by state unemployment insurance programs, which make up over 95% of public and private-sector employment, excluding certain agricultural industries and the unincorporated self-employed. Respondents are asked to provide employment counts for their establishment by occupation and wage bracket, which are then used to estimate the distribution of wages by occupation and state. The public-use data are released as three-year moving averages, and are derived from 1.1 million establishment-level surveys covering an average of 57% of all U.S. workers.

The OES data permits an analysis of the effect of regulation on labor market outcomes using much finer occupation categories and larger underlying sample sizes than are currently available in any household survey. Further, because the data are based on employer job classifications rather than an individual’s self-reported occupation, measurement error resulting from the misclassification of occupations is less of a concern ([Abraham and Spletzer, 2010](#)). However, because the data are not available prior to 1999, fewer policy changes provide identifying variation and the relatively short outcome window makes it more difficult to interpret long-run treatment effects, which I show are crucial for understanding the full impact of occupational licensing on the labor market.

¹⁷I harmonize revisions to the Standard Occupational Classification system adopted in 2010 by aggregating a small number of occupations that were split or combined using a crosswalk described in [Section 3.8](#). This results in a balanced set of 788 six-digit occupation codes with consistent longitudinal coverage.

3.4 Research Design

My research design exploits the rich policy variation resulting from differences in the timing of regulatory changes within occupations across states. To estimate the impact of occupational licensing on long-run labor market outcomes, I adopt an event study framework, which generalizes the canonical difference-in-differences estimator to a setting with time-varying treatment effects.¹⁸ This specification traces out the cumulative response of earnings and employment to licensing events within narrowly-defined state-by-occupation cells relative to changes for the same occupation in states that never licensed the occupation or adopted licensure at a different point in time.

Because licensing statutes do not necessarily apply to all workers within a statistical occupation code, my results should be interpreted as intention-to-treat effects that measure changes in outcomes at the cell level in response to an increase in licensing coverage. Further, since my design compares outcomes within occupation categories over time, occupations with no variation in treatment timing or policy differences during the sample window do not provide identifying variation. My results are therefore not applicable to occupations such as dental hygienists, physicians, or hairdressers that were already licensed in all states prior to 1980. [Appendix Table 3.2](#) displays summary statistics for a subset of the licensed occupations that provide identifying variation in my analysis.

¹⁸The canonical difference-in-differences estimator simply contrasts pre- and post-treatment periods for each unit, which requires the assumption that treatment effects are realized immediately and remain constant over time. As discussed in [Section 3.2.1](#), this assumption is unlikely to hold in my setting. In staggered treatment designs with time-varying treatment effects, difference-in-differences estimates confound long-treated observations for treated units with untreated controls, potentially resulting in biased estimates. In some cases, the bias can be severe enough that point estimates may be negative even when the underlying treatment effects are uniformly positive ([De Chaisemartin and d’Haultfoeuille, 2020](#); [Goodman-Bacon, 2018](#); [Borusyak and Jaravel, 2018](#)).

3.4.1 Identification

The underlying variation in my data is illustrated by [Figure 3.2](#), which plots the date and type of regulation events for a subset of fifty licensed occupations. Importantly, initial licensing dates for the same occupation often differ by decades across states. This implies that it is feasible to estimate relatively long-run treatment effects by comparing early to late adopters, even for the minority of occupations that are eventually licensed in all states. The figure also highlights the importance of controlling for alternative approaches to regulation. At any point in time, an occupation may be licensed by some states, but certified or registered in others and each state-occupation pair may experience multiple changes to its regulation status during the sample window. Many current licensing statutes were preceded by certification or registration requirements, and a small number of regulations were repealed or overturned in court.¹⁹

Identifying the causal effect of licensing on earnings and employment requires that within-occupation variation in the timing of policy changes is exogenous to potential outcomes. Because licensing is a policy choice made by state governments, however, the adoption of regulations is unlikely to be completely random ([Law and Kim, 2005](#); [Stigler, 1971](#)). To assess possible threats to identification, I analyze the political and economic factors associated with the timing of licensing policies in an event history analysis. [Table 3.3](#) displays the results of a conditional logistic regression with time-varying hazard rates for each occupation. Unsurprisingly, the partisan composition of the state government is strongly associated with the timing of policy changes, though there is also evidence that policies are adopted earlier in states with lower aggregate employment-to-population ratios and possibly less union coverage. The average wage and total employment in the three-digit occupation code do not

¹⁹Few units, however, experience more than two events, and the maximum observed in the data is four. Alabama adopted a state certification policy for interior designers in 1982, which was replaced with a licensing requirement in 2001. In 2004 the practice act was declared unconstitutional by a state court. After the ruling was upheld by the Alabama Supreme Court in 2007, the legislature adopted a revised certification act in 2010 ([Thornton and Timmons, 2015](#)).

appear to influence the timing of legislation.

The correlation between the adoption of licensing policies and time-varying state characteristics raises concerns that states with a greater propensity to regulate occupations may differ along other unobserved dimensions as well. Fortunately, because I study many regulated occupations (and not all occupations are regulated) my design is able to flexibly control for these potential confounds. In addition to the standard two-way unit and time fixed effects, my preferred empirical specification also introduces state-year fixed effects that vary across six major occupation groups. These terms absorb any aggregate labor market shocks or institutional differences that have a common effect on all occupations. Allowing the fixed effects to vary by major occupation group further nets out any state-specific trends affecting similar licensed and unlicensed occupations – such as the polarization of local labor markets (Autor and Dorn, 2013) – that may have occurred contemporaneously with the expansion of licensing requirements.

3.4.2 Treatment Measures

Earnings and employment are observed at different levels of occupational aggregation in my samples, which may or may not coincide with the coverage of licensing statutes. I therefore construct treatment measures first at the six-digit level, then use employment weights to map these variables to three-digit Census codes.

Treatment Events. The policy changes that provide identifying variation in this paper occur at the level of a state s and a detailed six-digit occupation code k . In each calendar year t , a state-occupation cell is either unregulated, certified, licensed, or potentially subject to a weaker form of regulation such as a registration requirement. Treatment events are identified by a set of indicators e_{kst}^l which are equal to one if unit ks experiences a policy change of type l in year t and zero otherwise. Although the adoption of licensing regulations are the main events of interest in this paper, I control for other policy changes throughout

the analysis.

To increase statistical power, I treat all licensing events as equivalent, regardless of whether or not weaker regulations were already in place at the time the law was passed. Repeals, while uncommon, are modeled as a separate type of event to avoid placing any symmetry restrictions on the effect of adopting and repealing a policy. My baseline specification therefore includes four types of events: certification, licensing, repeal, and a single residual category capturing all other occupation-specific regulations. The treatment variables are constructed from these event indicators by summing them over time,

$$d_{kst}^l = \sum_{r \leq t} e_{ksr}^l. \quad (3.1)$$

For the vast majority of state-occupation pairs that experience at most a single policy change of each type, d_{kst}^l is equivalent to the familiar post-treatment indicator of the canonical difference-in-differences specification. As in [Autor et al. \(2006\)](#), however, the indicators for multiple policies may be on at the same time if the unit has experienced more than one treatment event by time t . Since repeals are captured by a separate variable, all treatment sequences are non-decreasing, which implies that d_{kst}^l may be greater than one if a policy is adopted, repealed, and later re-adopted. In practice, these cases are extremely rare, and although I include them in the sample, dropping these cells has no effect on the results.

Aggregation. With the exception of the Occupational Employment Statistics sample, outcomes are observed only for occupation categories defined by the Census, which are typically less detailed than the six-digit level where treatment is assigned. Census occupational classifications may therefore contain both regulated and unregulated suboccupations, as well as units whose treatment timing differs. However, because the 310 consistent occupation codes used in this project are direct aggregations of the Standard Occupational Classification, it is possible to estimate the share of each three-digit code j covered by its six-digit components k using the OES data. These estimates can then be used to construct

an employment-weighted average of d_{kst}^l ,

$$D_{jst}^l = \sum_k \hat{\pi}_{jk} d_{kst}^l, \quad (3.2)$$

where $\hat{\pi}_{jk}$ is an estimate of the share of j 's total national employment in suboccupation k as of 2000, the approximate midpoint of the samples using these aggregated codes.²⁰

Equation 3.2 defines the treatment variables used in the main analysis, which map policy changes for detailed occupations to the balanced occupational classification system used in the CPS and ACS-Census samples. Although the identifying variation is limited to events that affect an entire six-digit occupation code, at this level of aggregation, these units may be grouped together with other detailed codes containing regulations that are even more narrowly defined (Kleiner and Soltas, 2020). Since it is not possible to assign employment share weights to regulations below the six-digit level, I control for them instead using a separate set of variables that count the number of these "minor" policies by three-digit occupation. I include in this set any six-digit occupations that cover less than 20% of the corresponding three-digit category to avoid identifying from events that have very little bite at the level outcomes are observed.²¹

²⁰The Occupational Employment Statistics program does not report employment data for all six-digit occupations prior to 2004, so employment for some occupations must be estimated. To do this, I use data from 2004 to 2016, where the panel is fully balanced, and log-linearly interpolate the missing employment data prior to calculating shares. I use national estimates due to concerns about the endogenous effect of regulation on employment at the state level. A time-invariant measure is also preferred so the weights do not vary with treatment timing. All within-occupation variation in Equation 3.2 therefore comes from policy changes rather than differences in employment shares across states or over time.

²¹A small number of occupations do not cover an entire six-digit code, but reasonably partition one, for example, water and wastewater treatment plant operators or school and public librarians. I assign these suboccupations a weight of 50% prior to aggregation rather than exclude them. Appendix Figure 3.15 plots the final distribution of aggregation weights $\hat{\pi}_{jk}$ for the regulated six-digit occupations included in the data. In Section 3.5.4 I show that the main results are similar when using alternative thresholds or methods of aggregation.

3.4.3 Event Study Specifications

My baseline empirical specification takes the standard two-way state and time fixed effects design that would be appropriate for a case study of a single occupation and stacks across all occupations in the data to estimate the average effect of occupational licensing statutes on labor market outcomes. Specifically, I estimate a distributed lag regression of the form,

$$Y_{jst} = \alpha_{js} + \sum_{l \in \mathcal{L}} \sum_{\tau \in \mathcal{T}} \gamma_{\tau}^l D_{js,t-\tau}^l + X'_{jst} \Lambda + \Omega_{jst} + \delta_{jt} + \epsilon_{jst} \quad (3.3)$$

where Y_{jst} denotes an outcome for occupation j in state s at time t . The unit fixed effects α_{js} absorb initial level differences in occupation-specific outcomes across states, and the time fixed effects δ_{jt} control for occupation-specific time trends common to all states. The vector X_{jst} includes additional controls that vary by state and occupation, and Ω_{jst} denotes an additional set of distributed lags that count the number of minor regulations in effect by three-digit occupation.

The key explanatory variables are a set of leads and lags of the treatment variables $D_{js,t-\tau}^l$ (the inner summation) which span an event window $[\underline{\tau}, \bar{\tau}]$. The coefficients on these variables identify the contemporaneous effect of a policy change occurring τ years relative to time t . I include a full set of these indicators for each type of policy change (the outer summation), which is a specification similar to the multiple event study design studied by [Sandler and Sandler \(2014\)](#). The coefficients of interest are event study estimates $\hat{\beta}_{\tau}^l$, which trace out the cumulative response of the outcome to the policy over time. These are recovered by taking the running sum of the contemporaneous policy responses,

$$\hat{\beta}_{\tau}^l = \sum_{\rho \leq \tau} \hat{\gamma}_{\rho}^l \quad (3.4)$$

which I normalize by the period prior to treatment. Recovering the event study coefficients from a distributed lag model in this way produces point estimates and standard errors that

are numerically identical to estimating the event study directly with a sequence of relative time indicators ([Schmidheiny and Siegloch, 2020](#)). The distributed lag model is preferred in my setting because it is less cumbersome to implement with non-binary treatment variables and recurring events.²² Given the small sample sizes in some surveys, I also report simple averages of the event study coefficients along the lines of the estimation strategy proposed by [Borusyak and Jaravel \(2018\)](#).

My baseline specification implicitly allows any states not regulating an occupation at time t to contribute toward the estimation of counterfactuals for that occupation. This is undesirable if labor market trends display spatial heterogeneity ([Dube et al., 2010](#)). Moreover, the results presented in [Section 3.4.1](#) suggest that certain time-varying state characteristics may be correlated with the adoption of licensing policies. To guard against potential bias from these sources, my preferred empirical specification allows the occupation-year fixed effects to vary by census division d and introduces state-year fixed effects for six broad occupation categories g ,

$$Y_{jst} = \alpha_{js} + \sum_{l \in \mathcal{L}} \sum_{\tau \in \mathcal{T}} \gamma_{\tau}^l D_{js,t-\tau}^l + X'_{jst} \Lambda + \Omega_{jst} + \delta_{djt} + \phi_{gst} + \epsilon_{jst}. \quad (3.5)$$

The fixed effects δ_{djt} address concerns related to spatial heterogeneity by limiting comparisons to the same occupation in geographically proximate states. The ϕ_{gst} terms further control for any state-level trends common to similar groups of occupations regardless of their regulation status, as in a triple-differences design.

Event Window Trimming. I estimate event study coefficients over an event window beginning 10 years prior to treatment and ending 25 years after treatment. Treated units are dropped after 25 years have elapsed without an event to prevent long-treated units from re-entering the sample as controls. This is because once licensed, regulation may increase on

²²[Figure 3.3](#) provides a graphical illustration of the relationship between contemporaneous and cumulative treatment effects in a setting where units experience multiple treatment events.

the intensive margin (i.e. more stringent entry requirements or changes to scope of practice), making long-treated units poor controls for recent adopters. I keep all observations prior to treatment but interact the unit fixed effects with an event window indicator. This allows units that are eventually treated to serve as controls more than 10 years before their treatment date, but excludes these distant periods from estimation of the unit fixed effects, which may bias estimates if the parallel trends assumption is violated in years long before the event occurs ([Autor et al., 2006](#)).

Weighting and Standard Errors. In regressions using state-occupation aggregated data, I weight observations by cell-level employment counts. Weights are derived from the person-level survey weights, and are computed to correspond to the universe of the outcome variable, i.e. wage and salary workers in the CPS wage regressions and all workers including the self-employed for employment regressions. Unless otherwise noted, standard errors are clustered at the state level to allow for arbitrary correlation in the error terms both within state-by-occupation cells over time and between different occupations in the same state.

3.5 Main Results

In this section, I study the average impact of occupational licensing statutes on earnings and employment using data from three independent labor force surveys. These estimates measure the change in outcomes within state-occupation cells in response to the adoption of licensing legislation for six-digit occupations with state-level policy differences during the sample. I find consistent evidence that licensing increases average hourly wages by 4-7% after being in effect for twenty-five years, but no evidence of a contemporaneous decline in employment. The data rule out long-run disemployment effects greater than 5% and reject a pure regulatory capture interpretation of licensing for the typical occupation in my sample.

3.5.1 Evidence from the Occupational Employment Statistics

I begin with a graphical preview of my main findings by estimating the baseline two-way fixed effect event study specification using data from the Occupational Employment Statistics sample. Unlike the other datasets used in this paper, the OES reports average hourly wages and employment counts by detailed six-digit occupation, so these initial estimates are free from any potential biases resulting from the aggregation of occupation codes. Moreover, because the data are constructed from large underlying samples and are based on employer occupational classifications rather than worker self-response, the outcome variables are less susceptible to measurement error than those obtained from household surveys. Although there are limitations to using the OES data alone – in particular the short panel and lack of data for the self-employed – the remaining analysis is largely consistent with the pattern shown in these figures.²³

Figure 3.4 displays estimates from the baseline two-way fixed effect specification, where all coefficients are expressed relative to the year prior to treatment. Panel A shows that licensing has a clear positive impact on average hourly wages. As expected given that licensing requirements are not immediately binding for all workers, the wage premium appears only gradually and is not statistically significant at the 5% level for more than a decade after the policy change. By contrast, there is no evidence of differential wage trends prior to policy adoption. Panel B displays estimates of the impact of licensing on employment. Unlike for wages, there is some evidence that employment in treated states was already declining prior to regulation. After treatment, however, there is no evidence that the number of workers employed in the occupation fell. If anything, the declining trend observed prior to

²³Long-run treatment effects in this sample are identified only through a comparison of units treated before the beginning of the sample to late adopters and never-treated controls. This raises concerns that the estimated cumulative response function may reflect the changing composition of events rather than the average dynamic treatment effect. Figure 3.5 reports the results of an alternative specification that matches treated units to untreated controls and balances the sample in event time (Cengiz et al., 2019). This design fixes the composition of events so the full cumulative response function is identified from the same set of units. While balancing the sample prevents the estimation of long-run effects, short and medium-run estimates are similar in the balanced and unbalanced samples.

the treatment date reverses, though the standard errors are too large to rule out a uniform null effect on employment during the outcome window.

Plotting the full set of event study coefficients is useful to visually assess treatment dynamics, but the point estimates are imprecise, particularly for longer leads and lags. To increase statistical precision and summarize the estimates, I also report simple averages of the coefficients and their corresponding standard errors. This washes out higher frequency variation in the cumulative response function to focus on longer-run dynamics. Taken at face value, the long-run estimate for the impact of licensing on wages is 4.4 log points (s.e. 1.5). This is an economically significant effect that is roughly equivalent in magnitude to increasing the average educational attainment of all workers in the cell by six months (Card, 1999). My intention-to-treat effects are smaller, however, than most estimates of the licensing wage premium reported in the literature, which typically fall between 4 and 10 percent when comparing licensed to unlicensed workers in cross-sectional survey data (Ingram, 2019; Gittleman et al., 2018).

3.5.2 Impact of Licensing on Earnings

My main estimates use data from the Current Population Survey from 1983 to 2018, which roughly doubles the length of the panel relative to the Occupational Employment Statistics sample. This enhances the credibility of my long-run estimates and allows for a richer set of controls, but comes at the cost of less detailed occupation codes and smaller underlying sample sizes. To ensure that a sufficient number of observations are used to estimate each coefficient, I replace the annual leads and lags of the treatment variables with three-year leads and lags, then report nine-year averages of the binned event study coefficients, as described above. Despite major differences in the samples, I find that the CPS estimates are remarkably similar to those obtained from the OES, providing strong evidence that licensing has a meaningful and plausibly causal impact on earnings.

Baseline Estimates. [Table 3.4](#) presents my main earnings results in specifications with progressive policy and spatial controls. Column one reports a specification that includes leads and lags of licensing policies only. In this specification, states adopting certification, registration, or minor regulations form part of the control group, along with states that do not regulate the occupation at all. Introducing policy controls in column two increases the magnitude of the medium and long-run estimates somewhat, implying that weaker methods of regulation also have some impact on worker outcomes. More importantly, the estimates in column two are extremely similar to those found in the equivalent regression reported in [Figure 3.4](#), with an estimated long-run wage premium of 4.2% (s.e. 2.2%). This is reassuring and shows that the aggregation of treatment to the three-digit level is unlikely to be a major source of bias. Column three adds state-year fixed effects specific to six broad occupational categories, with little impact on the estimates.

Columns four to six add separate occupation by year fixed effects for each of the nine census divisions. These sweep out variation between divisions and restrict the identifying variation to differential timing in policy adoption among a set of geographically proximate states. If the adoption of licensing policies are correlated with unobserved spatial heterogeneity, these comparisons form better control groups than allowing any untreated state to form part of the counterfactual ([Allegretto et al., 2017](#)). Comparing outcomes within census divisions increases the estimated impact of licensing on wages over time, but the underlying pattern of dynamics is qualitatively similar. As in the baseline specification, controlling for common time effects by state and occupation group has no effect on the estimates. Column six, which is my preferred specification, implies a long-run wage premium of 6.9% (s.e. 2.1%) due to licensing.

Across specifications, there is clear evidence that licensing increases the average hourly wages of workers in regulated occupations. Although they are never statistically significant, the pooled pre-treatment effects are uniformly positive in these regressions, and roughly comparable in magnitude to the short-run treatment effects. Since the pre-treatment effects

are also expressed relative to the leave-out period one to three years prior to treatment, these estimates imply that wages were *declining* slightly on average before licensing statutes were adopted. The consistency of this finding across increasingly saturated specifications could be evidence that policy changes are anticipated, which is plausible given that legislation may be introduced in several legislative sessions before it is ultimately passed. [Figure 3.6](#) plots the full set of event study estimates for each specification and shows that any potential decrease in wages before licensing occurs is quickly reversed in the post-treatment period.

Alternative Earnings Measures. One concern when estimating the impact of licensing on earnings using data from either the CPS or the OES is that neither include workers who are self-employed. If the impact of licensing on self-employment income differs from its effect on wage and salary income, or if licensing selectively shifts workers into or out of self-employment, the earnings estimates presented so far may be biased.²⁴ For that reason, I also report earnings estimates using data from Census and American Community Survey, which includes measures of wage and salary income, as well as total earned income from all sources. While this is an advantage of the Census data, measuring hourly compensation is less straightforward than in the CPS. Specifically, hourly wages (and earnings) must be imputed based on annual income, weeks worked at all jobs in the previous year, and usual weekly hours. Because this introduces greater potential for measurement error, I report estimates for the impact of licensing on both hourly and annual income, through the latter also captures changes in the intensive margin of labor supply.

[Table 3.5](#) reports the impact of licensing on the Census earnings variables. All regressions include separate occupation by year fixed effects for each census division, with my preferred triple-difference specification reported in the even-numbered columns. Overall, the Census results show the same pattern of treatment dynamics as found above, though the estimates

²⁴[Figure 3.8](#) plots event study estimates for the share of workers who are self-employed in the CPS. Although the estimates are not statistically significant in my preferred specification, there is some evidence licensing has a negative effect on self-employment as large as 10-20% relative to the mean self-employment rate in licensed occupations.

are more comparable in magnitude to those obtained from the baseline two-way fixed effect specification in the CPS and OES samples. Importantly, however, within the Census sample, the effect of licensing on earnings is similar regardless of whether or not business income is measured. I find a long-run impact of licensing on annual wage and salary income of 4.3% (s.e. 2.0%) compared to an impact on total earned income of 4.4% (s.e. 2.1%). This results alleviates concern that that failure to measure self-employment income biases the results of the previous section, and provides further evidence that licensing has a positive impact on earnings in another independent sample of workers.

Robustness to Sample Definitions. [Figure 3.7](#) displays the results of a number of robustness checks for my preferred CPS earnings estimates. Specification one drops public-sector jobs, which I define to be any occupation where 50% or more of workers are employed by the local, state, or federal government. In practice, the largest licensed occupations excluded from the sample under this definition are law enforcement officers, water and wastewater treatment plant operators, and speech-language pathologists (most of whom are employed in public schools). The second specification includes additional controls for the number of contiguous states licensing the same occupation. This is potentially a concern if licensing has geographic spillovers or the presence of nearby unregulated markets affects the return to licensing. The third specification drops 1994 and 1995 from the sample due to a lack of reliable wage imputation flags. Specification four drops policies that were repealed and later re-enacted. Specification five allows long-treated units to re-enter the sample as controls and specification six reports estimates that are not employment-weighted. The remaining rows drop each of the nine census divisions separately.

None of these tests raise serious concerns about the robustness of the wage estimates, though allowing units that have been licensed for twenty-five years or more to serve as controls lowers the estimated wage premium somewhat and results in a statistically significant pre-treatment effect. This finding is consistent with graphical evidence shown in [Figure 3.6](#), which implies that wage growth may not have returned to trend in licensing states by the end

of the event window. If wages continue growing in early-adopting states more than twenty-five years after treatment, including these observations in the set of potential controls biases estimates of the counterfactual trends for late-treated units upward, resulting in smaller estimated treatment effects overall. Even under the assumption that these comparisons are valid, the results show an economically significant increase in average hourly wages following the adoption of licensing statutes.

3.5.3 Impact of Licensing on Employment

Turning now to my main results on employment, I again begin with data from the Current Population Survey. I then replicate my findings in the Census and American Community Survey to address concerns that small sample sizes are driving the results. In both datasets I find that the impact of licensing on employment is weakly positive, implying that the licensing wage premium is not driven by a reduction in the extensive margin of labor supply.

Baseline Estimates. [Table 3.6](#) presents my main employment results and shows that licensing has a uniformly positive impact on the number of workers employed in the occupation as a share of the state’s total population. Comparing the estimates in column two to the equivalent Occupational Employment Statistics specification, I find a larger positive effect of licensing on employment in the CPS that is statistically significant at the 1% level in the medium to long-run. This specification implies that employment per capita increased by 7% (s.e. 2.5%) on average two decades after the occupation became licensed. Restricting the identifying variation to comparisons within census division has little impact on this finding. The magnitude of my two-way fixed effect estimates are comparable to those reported by [Redbird \(2017\)](#), who found that licensing increases an occupation’s share of total state labor hours between 5% and 10% after twenty years.

In contrast to my findings for earnings, however, netting out common state by occupation time trends reduces the magnitude of the employment effects substantially and renders them

statistically insignificant in both the baseline and census division specifications. A significant share of the increase in occupational employment observed following the adoption of licensing statutes therefore appears to have been driven by broad changes in the composition of the state's labor market rather than the causal effect of occupational licensing. Nonetheless, even in the most saturated specification, there is no evidence of large disemployment effects in either the short or long-run.²⁵ My preferred specification in column six implies a moderate increase in employment of 4.4% (s.e. 3.4%) after twenty years, ruling out a long-run decrease in employment of more than 2.5% at conventional levels of statistical significance.

ACS-Census Replication. The small sample size of the CPS is potentially a concern when estimating employment changes, as many state by occupation cells contain few observations. I therefore replicate my employment analysis using data from the Census and American Community Survey, where such sparsity is less of an issue. The results are shown in [Table 3.7](#) and are extremely similar to my preferred CPS estimates. As above, the point estimates are uniformly positive and statistically significant in the two-way fixed effect specifications, but substantially smaller and insignificant in the triple-difference specifications. In the Census sample, my preferred specification rules out a long-run decrease in employment greater than 5.4%. Interestingly, the Census results also display a decrease in employment *before* licensing takes effects of 2-4%, which is comparable to the pattern observed in the Occupational Employment Statistics.

Robustness to Sample Definitions. [Figure 3.10](#) displays sample definition robustness checks for my preferred CPS employment results. These are equivalent to the alternative specifications estimated for earnings, with the exception of specification three, which estimates employment effects in log levels rather than log shares. As with the wage results,

²⁵[Figure 3.9](#) displays the full set of underlying event study coefficients. In the census divisions specifications, there is a small dip in employment immediately following the enactment date of licensing. Given the imprecision of the estimates, it is difficult to interpret this effect, but it could reflect some incumbents who do not meet grandfathering criteria being forced out of the occupation in the short-run.

assumptions about the validity of long-treated units as potential controls appear to be the most consequential for the estimates. Allowing units licensed more than twenty-five years to serve as controls for late adopters results in a negative, but statistically insignificant employment effect of about -6%. As discussed above, this estimate is only valid under the relatively strong assumption that treatment effects have been fully realized after twenty-five years, which [Figure 3.9](#) suggests is not the case. By contrast, dropping either of two census divisions – East North Central or East South Central – results in positive effects that are marginally significant at the 95% level.

3.5.4 Robustness to Treatment Aggregation

[Table 3.8](#) and [Table 3.9](#) show the robustness of my preferred triple-difference results to alternative methods of aggregating the treatment variables from the six-digit to three-digit level. In addition to measuring sub-occupation employment shares in 2000, I also report measures that use 1980 and 2015 as the year of measurement. Because the Occupational Employment Statistics data is only available beginning in 1999, I log-linearly interpolate 1980 employment shares based on observed employment trends from 1999 to 2018. The 2015 shares use data directly from the OES without additional adjustments. For each of these three samples, I also report estimates that use all six-digit events as well as specifications excluding suboccupations covering less than 20% or 50% of their three-digit superset.²⁶ Throughout these tables, the underlying pattern and magnitude of the effects are similar to my baseline estimates.

²⁶As another check, I computed aggregation weights using a leave-out method. Because there are few cases where a single state has a large influence on sub-occupation employment share estimates, the results were nearly identical to those reported in these tables.

3.6 Heterogeneity and Mechanisms

The baseline estimates discussed in the previous section represent an average across all occupations in the data. This section assesses potential mechanisms for my findings and explores heterogeneous responses to occupational licensing requirements. I show that although licensing increases average educational attainment, observable changes in worker composition explain only a small share of the licensing wage premium. Next, I allow the effects of licensing to vary across different groups of occupations. As in the main analysis, I find that licensing is consistently associated with higher earnings, but has no discernible impact on employment for most subgroups. An important exception are occupations that pose a low risk of consumer harm, for which the evidence that licensing functions primarily as a barrier to entry is stronger. In cases where the consumer protection rationale for licensing is plausible, I find simultaneous increases in earnings and employment.

3.6.1 Mandated Investment and Worker Sorting

Occupational licensing statutes establish uniform minimum qualifications that workers must meet to legally enter the occupation. How much these mandatory investments in human capital raise the cost of entry, however, depends on how burdensome the statutory requirements are relative to the prevailing distribution of human capital in the market. In this section, I show that licensing increases average educational attainment by one to two months in the long-run, with larger effects for occupations that typically require at least an associate's degree. Controlling for this increase in the stock of human capital, as well as changes in the demographic composition of workers, explains at most 15% of the wage premium due to licensing. This finding rules out worker sorting as the primary mechanism through which licensing impacts the labor market.

Educational Attainment. To study the impact of occupational licensing requirements on human capital, I use data from the Current Population Survey and estimate my event

study design with average years of completed education by state-occupation cell as the dependent variable.²⁷ I present both average estimates using the full sample and separate event studies for jobs that typically require at least an associate’s degree and those that require a high school education or less based on information from O*Net Online, the successor of the Dictionary of Occupation Titles. [Table 3.10](#) shows that licensing has a uniformly positive impact on educational attainment, through these estimates are highly imprecise and only statistically significant in the long-run. My preferred specification implies that licensing increased average educational attainment across all occupation by 0.12 years (s.e. 0.07), or approximately a month and a half.

The table also shows that the increase in educational attainment is slightly larger for jobs that usually require workers to attend some college or more, with average educational attainment increasing by two months in the long-run. Because years of education are measured in discrete intervals, increasing educational attainment likely reflects shifting some workers from high school to associate’s degrees and from bachelor’s to master’s degrees. Assuming that each takes an additional two years to complete, an alternative interpretation of the magnitude of this coefficient is that licensing increased the share of workers with associate’s or master’s degrees by 10% relative to non-licensing states. For comparison, the average increase in educational attainment I find is approximately one-third as large as that reported by [Kleiner and Soltas \(2020\)](#), who find licensing raises mean education by 0.4 years when proxying for policy differences using the recent addition of self-reported license attainment to the CPS questionnaire.²⁸

²⁷The Current Population Survey changed its coding of educational attainment from a definition based on years of completed education to highest degree attained in 1992. I follow the replication materials from [Acemoglu and Autor \(2011\)](#) to impute years of education following the revision. In addition, I note that vocational training is not well measured in the Census definition of education, which may bias the estimated impact of licensing on educational attainment downward.

²⁸[Figure 3.11](#) displays the impact of licensing on the average age of workers. The results show that licensing potentially delays entry for jobs typically requiring some college education or more, but these estimates are extremely imprecise.

Does Sorting Explain the Wage Premium? One potential explanation for the finding that licensing increases earnings without reducing employment is that wage premium simply reflects changes in the composition of workers in the occupation. [Table 3.11](#) suggests that this is not the case. In the table, I present my preferred earnings specification with progressive controls for worker characteristics using data from the Current Population Survey. After controlling for increases in average educational attainment and the potential experience of workers, I still find a long-run wage premium of 6% (s.e. 1.9%) relative to my baseline estimate of 6.9%, a reduction consistent with the small increase in average educational attainment discussed above. Additional controls for the demographic composition of the occupation, union coverage, and the share of workers who are self-employed has little effect on the estimates. While this analysis does not account for selection on unobservables, it suggests that worker sorting alone does not explain the wage premium.

3.6.2 Heterogeneous Impacts of Licensing Requirements

Licensing policies apply to an incredibly diverse set of occupations. To assess the extent to which my main results mask important heterogeneity in earnings and employment responses, I allow the impact of licensing to vary based on the characteristics of the occupation.

By Difficulty of Entry. Panel A of [Table 3.12](#) reports the results of estimating my preferred event study specification separately by the level of preparation needed to enter the occupation based on data from O*Net. The first group includes jobs with low to moderate entry requirements. These usually require at most an associate's degree, apprenticeship, or vocational training. The second group includes higher-skilled jobs that usually require at least a bachelor's degree. I find that the licensing wage premium is larger from occupations in the former group, with a long-run wage premium of 7.5% (s.e. 2.8%), compared to a statistically insignificant 3.8% (s.e. 5.3%) for the latter. I find no evidence of a decline in employment for either group.

By Occupation Group. To assess the extent to which the results are driven by the licensing of high-skilled services, I next split the sample into the highly-licensed professional specialty occupations based on the (Dorn, 2009) occupational classification and all other occupations in the sample. The former contains many health care occupations, as well as social service jobs such as psychologists and social workers. In Panel B of Table 3.12, I find a substantial long-run wage premium of 17.7% (s.e. 6.5%) for professional specialty occupations, though this estimate is highly imprecise owing to the smaller sample share of these occupations. This estimate is large relative to my baseline estimates, but it is comparable in magnitude to the average wage premium found by (Kleiner and Krueger, 2013) using custom survey data. All other occupations experience a smaller average wage premium of 6.6% (s.e. 2.8%), closer to my main estimates. Again, however, neither group shows any evidence of disemployment effects.

By Consequence of Error. The main justification for licensing is that it protects consumers from potential harm caused by unqualified workers. To measure these risks, I use the O*Net Online variable “consequence of error,” which ranks how serious the result would be if a worker were to make a mistake that was not easily correctable. I then estimate my preferred event study specification separately for licensed occupations above and below the overall median on this scale. Although the prevalence of licensing correlates strongly with consequence of error, the possibility that it increases demand by providing a signal of quality is more plausible for some occupations than others. Licensed occupations that appear to pose little risk to consumers include barbers, cosmetologists, florists, dispensing opticians, interior designers, and massage therapists. However, even many occupations with a high potential for harm – including crane operators, electricians, medical assistants, pharmacy technicians, radiation therapists, and stationary engineers – are still licensed by fewer than fifty states.

The results shown in Panel C of Table 3.12 are consistent with licensing increasing the demand for services when the consequences of worker error are high, but functioning pri-

marily as a barrier to entry when these risks are low. For occupations below the median consequence of error, I find a statistically insignificant long-run increase in wages of 4.7% (s.e. 3.9%), but a 24.8% (s.e. 8.3%) decrease in employment. Employment is also increasing for these occupations prior to regulation, which provides some evidence in favor of the regulatory capture view of licensing for these low-risk occupations. By contrast, for occupations above the median consequence of error, I find simultaneous increases in earnings and employment of 7.3% (s.e. 2.3%) and 9.1% (s.e. 3.5%) respectively, which is more consistent with the view that licensing addresses a market failure. As most licensing policies in my data were adopted for occupations with relatively high potential for harm, these results provide some evidence reconciling my findings with negative labor supply estimates in the literature (Kleiner and Soltas, 2020; Blair and Chung, 2019).

3.6.3 Additional Heterogeneity Results

Given the significant heterogeneity in the characteristics of licensed occupations, an ideal analysis would estimate a separate set of dynamic treatment effects for each policy. Unfortunately, small sample sizes in the CPS and the short OES panel prevent this type of analysis for most occupations. To provide some occupation-specific evidence along these lines, this section first exploits the precise cross-sectional wage and employment estimates from the OES to assess the distribution of earnings and employment differences by state and occupation. Next, I present separate event studies for a small number of individual industries that are highly exposed to occupational licenses using data on precise employment counts derived from administrative records. Finally, I show that the increase in average hourly wages documented in the main analysis is driven by gains in the lower tail of the earnings distribution, consistent with theoretical predictions.

Cross-Sectional Estimates. The large underlying samples used to produce wage and employment estimates for the Occupational Employment Statistics survey make this data

uniquely suited to a cross-sectional analysis of differences in outcomes across states for individual occupations. For each licensed six-digit occupation in my sample with policy variation of as 2015, I use the OES data to compute the difference in log average hourly wages and employment per capita between states that license the occupation and those where it is entirely unregulated. Although these comparisons should not be interpreted as causal, they are useful to assess how earnings and employment differences in the raw data are related to the estimates from my main empirical design.

[Figure 3.12](#) shows that hourly wages are higher in licensing states for two-thirds of the occupations in the sample. This pattern is reversed for employment, however, with point estimates more likely to be negative in licensing states. Pooling across all occupations, a simple employment-weighted OLS regression with occupation fixed effects implies that licensing is associated with 5% (s.e. 2.3%) higher hourly wages and 3.4% (s.e. 2.9%) higher employment on average.²⁹ These estimates are comparable in sign and magnitude to the long-run impact of licensing identified in [Section 3.5.1](#). Subject to the caveat that the OES does not capture self-employed workers, there is little evidence that licensing decreases employment on average, though this may clearly be the case for certain occupations. The high likelihood of finding statistically significant negative (or positive) employment effects for a single occupation also cautions against drawing strong conclusions about the average impact of occupational licensing from studies limited to a single policy setting.

Industry Case Studies. In the United States, data derived from administrative records such as the Quarterly Census of Employment and Wages (QCEW) report exact employment counts, and are therefore not subject to the sampling error that could potentially affect my estimates. While the data unfortunately do not contain information on the occupational affiliation of workers, in some cases, industry definitions are sufficiently detailed that a six-

²⁹Unweighted estimates imply that licensing is associated with an increase in wages of 3.7% (s.e. 1.1%) and a decrease in employment of -3.8% (s.e. 2.9%), which is consistent with the distribution of occupation-specific differences shown in the figure.

digit NAICS code is highly likely to be affected by a specific occupational license. For example, the industry “interior design services” (NAICS 541510) is clearly associated with the occupation “interior designer.” Industry-level administrative data can therefore be used to estimate the impact of certain occupational licenses on precise state employment counts for wage and salary workers. Although these estimates reflect the total impact of the policy on the employment of both licensed and unlicensed workers in the industry, if licensing causes an economically meaningful reduction in employment, this analysis should detect those effects.

Using data from the QCEW for 1990 to 2018, I report case study estimates for six industries that meet two sample selection criteria. First, the industry must be highly exposed to a single occupational license.³⁰ Second, a sufficient number of policy changes must have occurred within the sample window to identify the cumulative response of employment to licensing over the full -10 to +25 year event window. For each of these industries, I estimate a separate regression with state and census division by year fixed effects using the timing of the relevant occupational licensing policies. The results are shown in [Figure 3.13](#). Analyzing each industry separately results in a substantial loss in power, but as in the main analysis, the employment estimates are concentrated around zero. The only exception is that registration of locksmiths appears to have reduced employment, though this effect is driven by California’s statute alone, which the Institute for Justice ranks as the third *least* burdensome locksmith regulation in the United States ([Carpenter et al., 2017](#)).

Distribution of Wages. Throughout this paper, my analysis of earnings has focused on the impact of occupational licensing requirements on average hourly wages. Licensing, however, may not have a uniform impact across the wage distribution. By raising entry costs, the model of [Shapiro \(1986\)](#) predicts that licensing raises the price of lower-quality services while simultaneously decreasing the price of higher-quality services, implying larger

³⁰This rules out, for example, “Offices of physical, occupational, and speech therapists, and audiologists” (NAICS 621340), which is affected by the regulation of at least nine distinct occupation categories.

wage effects at the bottom of the earnings distribution. I test this hypothesis by estimating my baseline event study specification using data from the Occupational Employment Statistics, where the dependent variable is now the log of hourly wages at the 10th, 25th, 50th, 75th and 90th percentiles within the state by occupation cell. In [Table 3.13](#), I find that increasing average hourly wages are indeed driven by gains for workers at or below median earnings, including a moderate short-run wage premium. Consistent with theoretical predictions, there is no effect on the 90th percentile of hourly wages.³¹

3.7 Conclusion

Economic theory predicts that occupational licensing requirements will increase worker earnings, but total employment could rise or fall depending on the extent to which regulation influences the demand for professional services. Consistent with these predictions, I find that licensing increased earnings by 4 to 7% across all occupations in my sample. This wage premium, however, did not coincide with a significant decline in employment relative to non-licensing states. Rather, employment effects appear to depend on the type of occupation being licensed, in particular how severe worker errors might be. When the risk to consumers is high, I find that licensing increases employment, which is consistent with the view that regulation addresses a market failure. When the risk is low, employment falls, which suggests that licensing primarily increases barriers to entry instead. The robustness of these findings across datasets, time periods, and levels of occupational aggregation supports a causal interpretation of these estimates

Although understanding the impact of occupational licensing on earnings and employment is central to ongoing policy debates, it is important to stress that this paper does not attempt to fully characterize the costs and benefits of licensing. Even without negatively

³¹Even at the 90th percentile, fewer than 1% of wages are top-coded for the occupations in my sample, so this finding is not explained by the truncation of the wage distribution in the data.

affecting total employment, licensing may impose costs on workers by reducing geographic mobility ([Johnson and Kleiner, 2020](#); [Kleiner and Xu, 2020](#)). Likewise, an assessment of consumer benefits would require either a clearly-defined measure of service quality ([Anderson et al., 2020](#); [Farronato et al., 2020](#)), or a structural model of the labor and product markets ([Kleiner and Soltas, 2020](#)). Further, this paper looks only at the extensive margin of regulation. Conditional on becoming licensed, variation in specific requirements – such as banning workers with a criminal history – may matter a great deal for labor market outcomes ([Blair and Chung, 2018](#)). At a minimum, my finding of both a wage premium and null or positive employment effects across a broad range of occupations shows that the potential for licensing to increase labor demand should be taken seriously.

My results, together with the new data I compile in this paper, suggest several topics for future research. First, my findings depart substantially from those obtained using cross-sectional survey data on occupational licenses. Understanding why these alternative approaches arrive at different conclusions will be crucial for reconciling my findings with the previous literature. Second, what explains the growth in occupational licensing over time and its variation across states? Traditionally, the public choice view has interpreted the diffusion of regulation as evidence of rent-seeking behavior. At least in the health care sector, however, new license categories might also have been created in response to the increasing specialization of workers and the evolution of their scope of practice over time. Finally, the data used here should also be useful for analyzing the general equilibrium effects of occupational licensing and its macroeconomic implications.

3.8 Data Appendix

This appendix provides additional details about the earnings and employment data used in this project. Section one describes how revisions to the various occupational coding systems used by government agencies are harmonized to create occupation groups that are

consistently identifiable over time. Section two documents the processing of my outcome variables.

3.8.1 Consistent Occupation Codes

Standard Occupational Classification. The Standard Occupational Classification (SOC) system defines over 800 occupation categories, and is the most detailed level at which the federal government reports wage and employment estimates. The data used in this project span two vintages of these codes, introduced first in 2000 and then revised in 2010. Although 90% of occupations are directly comparable across these systems, the rest must be aggregated to construct a balanced panel. In cases where existing codes were split or combined, the changes are harmonized by simply recombining the more detailed components. If instead, a new code was broken out from two or more existing ones, a single category is created out of all occupations with overlapping coverage.³² This results in a balanced set of 788 non-military Standard Occupational Classification codes that are comparable in data coded using either the 2000 or 2010 systems. Aggregated codes and their original components are tabulated in [Table 3.14](#).

Census Occupational Classifications. Census occupation codes are less detailed than the Standard Occupational Classification system, and are revised every decade. In addition, some microdata released through IPUMS further aggregates small occupations when disclosure requirements are not met, resulting in additional changes to occupation codes implemented in the 2005 and 2012 American Community Survey. To harmonize the Census classifications,

³²Changes resulting from the introduction of new occupation codes for transportation security screeners and solar photovoltaic installers are not completely harmonized. This is because doing so would require combining several occupations across major groups in a way that significantly affects the implied regulation coverage of jobs such as electricians and security guards. Instead, I assign these occupations to the main categories they were broken out from, which are "compliance officers" and "construction and related workers, all other," respectively. Since total employment in the new categories is extremely small relative to the original classifications they were split from, the impact on employment share estimates in these occupations is negligible.

I begin with the “occ1990dd” system developed by (Dorn, 2009), who provides a balanced panel of occupations valid in the 1980 to 2000 Census samples and the ACS from 2005 to 2009.³³ Extending this system to cover the period from 2010 to 2017 requires additional aggregations, which reduce the total number of non-military occupation codes from 330 to 310. As shown in Table 3.15, these aggregations mostly affect manufacturing jobs and therefore have little impact on the licensed occupations studied in this project.

3.8.2 Processing of Wage and Employment Data

Current Population Survey Outgoing Rotation Groups. I use data from the IPUMS Current Population Survey outgoing rotation group extracts from 1983 to 2018 (Flood et al., 2018). The sample is limited to adults ages 16 to 64 in the civilian labor force, excluding unpaid family workers. I process this sample following the replication materials provided by Acemoglu and Autor (2011), with the exception of minor differences in the construction of hourly wages.

Usual hourly wages for workers paid by the hour are measured as the reported straight-time hourly wage, excluding overtime, tips, and commissions. For non-hourly workers, I construct the implied hourly wage by dividing weekly earnings by usual weekly hours. If usual weekly hours are missing or variable, actual hours worked in the previous week are used. Top coded earnings are multiplied by a factor of 1.5. Earnings are windsorized at half the federal minimum wage on the lower tail and the implied hourly wage corresponding to 1.5 times the maximum weekly earnings for a full-time hourly worker on the upper tail. Wages and earnings are adjusted to 2018 dollars using the CPI-U-RS.

Observations with allocated earnings or occupations are excluded from all analysis using these variables. A flag for imputed occupations is consistently available throughout the CPS sample, but the method for identifying wage imputations varies by year. With the exception

³³The occupation codes used in the 2001 to 2004 ACS provide slightly more detail than the 2005 classification, but can be directly aggregated to match the latter.

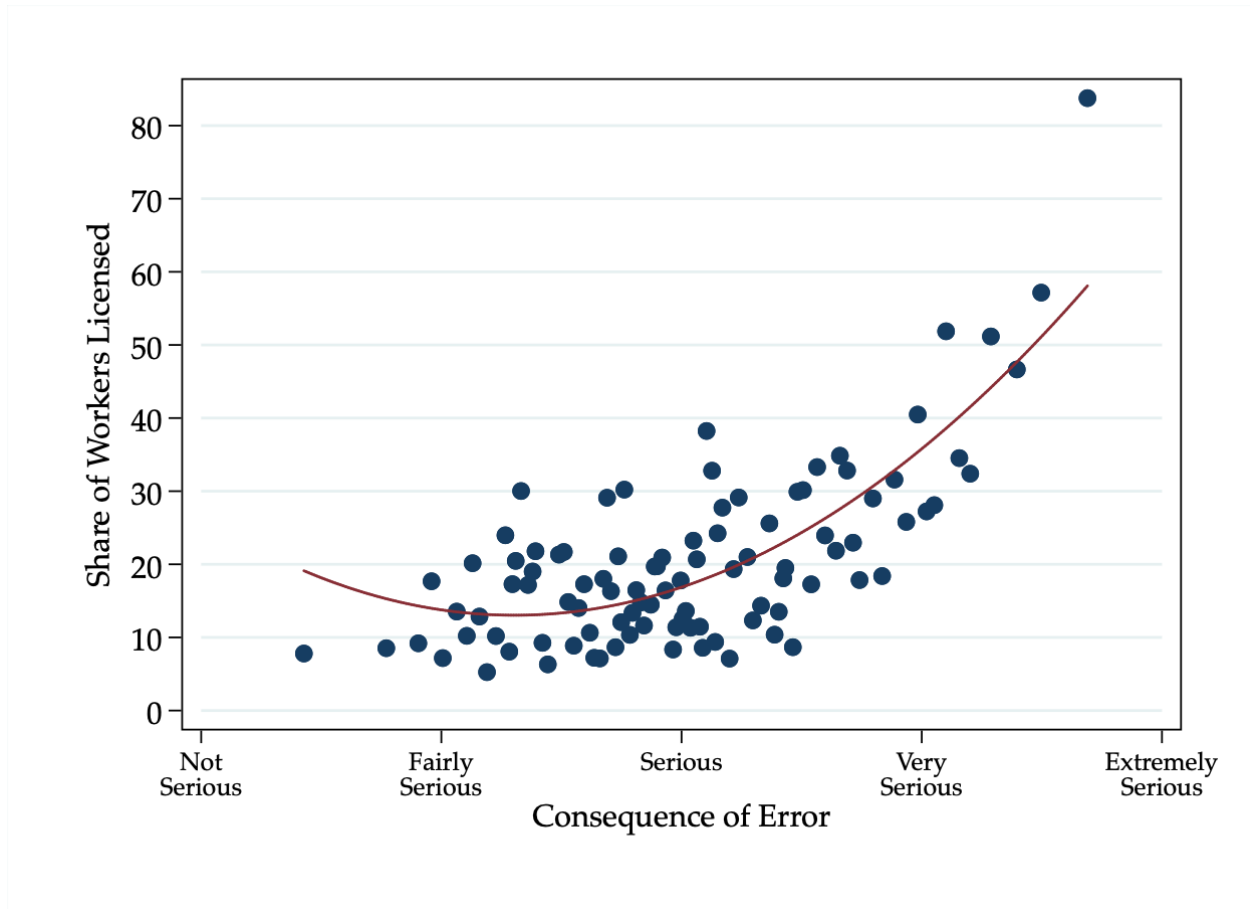
of 1989 to 1995, wage imputations are identified with positive allocation values for weekly earnings, hourly wages, or hours worked. Between 1989 and 1993, however, these flags are incomplete, and wage imputations are therefore identified as observations with positive edited earnings, wages, or hours, but missing unedited versions of these variables ([Hirsch and Schumacher, 2004](#)). As no method is known to identify wage imputations in 1994 and 1995, I retain all observations from these years in the main sample, but exclude these years in robustness checks.

Census and American Community Survey. I use data from the IPUMS Census extracts from 1980 to 2000 and the 2001-2017 American Community Survey ([Ruggles et al., 2019](#)). The sample is limited to civilian adults between the ages of 16 and 64 who are not living in group quarters. Demographic and education variables are processed following [Acemoglu and Autor \(2011\)](#). Hourly wages are constructed by dividing annual wage and salary income by the product of weeks worked and usual weekly hours. Top coded incomes are multiplied by a factor of 1.5, and implied hourly wages are truncated at 1.5 times the maximum feasible income for a full time worker given the top code. I then windsorize the bottom 1% of hourly wages in each year.

I also create a measure of hourly earnings that includes self-employment income by first applying the top coding rule to business income, then dividing the sum of wage and business income by hours of labor supply. Hourly earnings are windsorized below the first percentile of the hourly wage distribution for workers with positive wage and salary income, and below the first percentile of hourly earnings for those with business income only. The upper tail of hourly earnings is windsorized at the 99th percentile for workers with positive business income, and the adjusted top code of hourly wages otherwise. The hourly wage and earnings variables are therefore equivalent for workers with no business income. Observations with imputed occupation, wage, or earnings are excluded from any analysis that requires these variables.

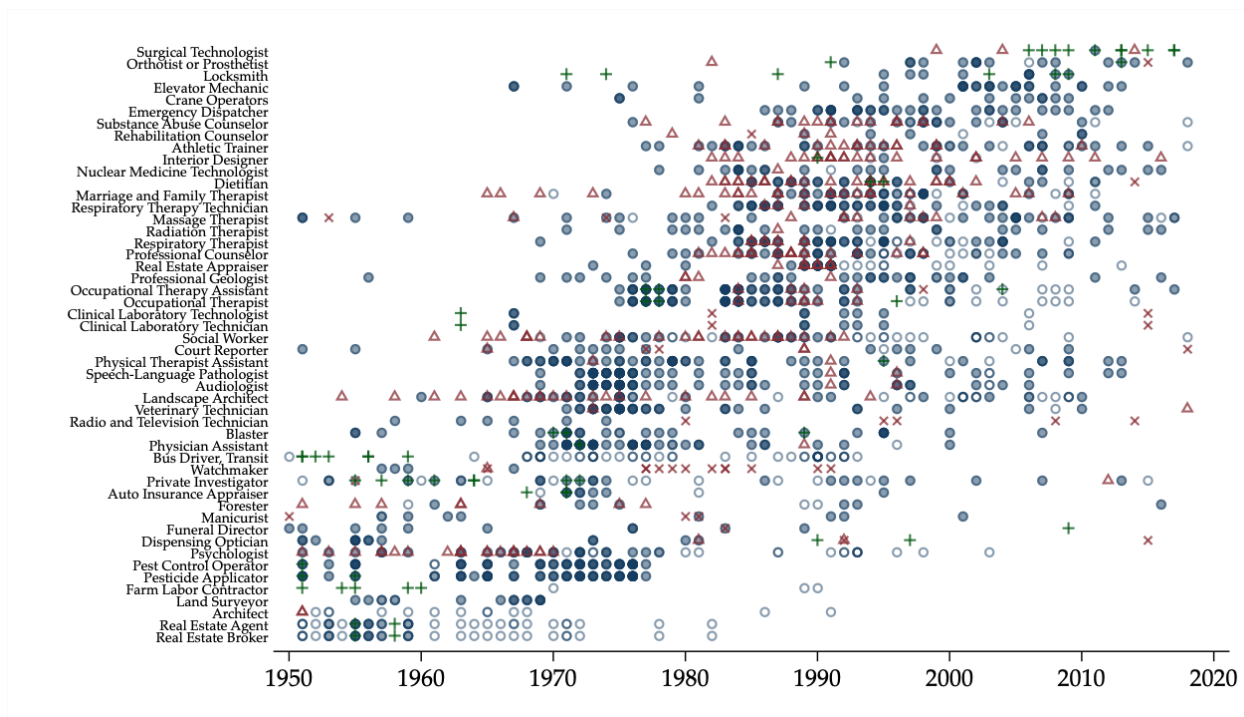
Occupational Employment Statistics. In addition to the individual-level surveys, I also use data from the Occupational Employment Statistics program, which provides wage and employment estimates by state and six-digit Standard Occupational Classification code for most occupations between 2000 and 2018. This data is based on three-year rolling averages of semi-annual business establishment surveys, and therefore excludes self-employed workers. In cases where only annual earnings are reported, hourly wages are constructed by dividing earnings by 2,080, which is the definition of full-time hours of labor supply used in the survey. Top coded wages are then multiplied by a factor of 1.5. Due to revisions to the Standard Occupational Classification system in 2010, I aggregate a small number of occupations to consistent definitions using the crosswalk described above in order to achieve a balanced panel. Hourly wages for aggregated occupations are constructed as employment-weighted averages across suboccupations.

Figure 3.1: Share of Workers Licensed by Consequence of Error



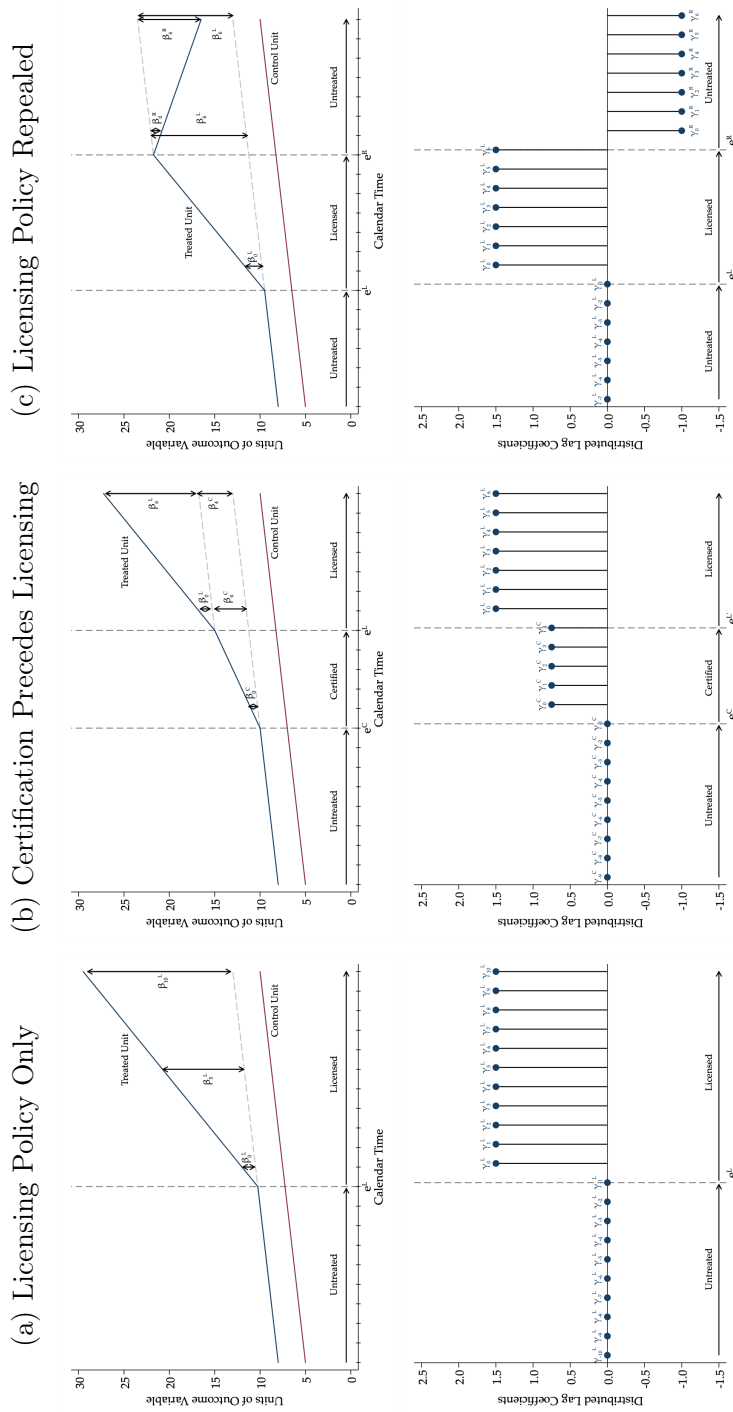
Notes: This figure uses data from the 2015-2018 Current Population Survey and O*Net Online. It displays a binned scatterplot of the relationship between the share of workers in the CPS reporting that they hold an occupational license against the potential severity of a worker error. The consequence of error measure is taken from O*Net, and measures on a one to five scale, “How serious would the result usually be if the worker made a mistake that was not readily correctable?”

Figure 3.2: Licensing, Certification, and Registration Events by Detailed Occupation



Notes: This figure plots the date and type policy changes for a subset of the larger six-digit occupations experiencing ten or more regulation events after 1950. The data are obtained from an analysis of state session laws. Solid blue circles mark the licensing of a previously unregulated occupation and open blue circles the licensing of an occupation that was previously registered or certified. Red triangles denote the certification of a previously unregulated occupation. Red crosses indicate that a regulation was repealed, and green plus signs show the adoption of a regulation other than state licensing or certification. Markers are darker when multiple states adopted the policy simultaneously.

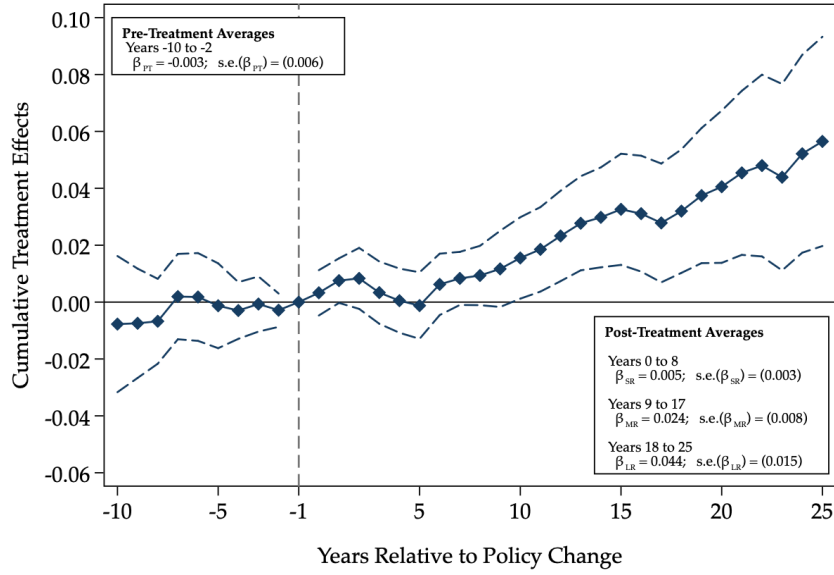
Figure 3.3: Hypothetical Outcome Paths and Dynamic Treatment Effects With Multiple Policy Changes



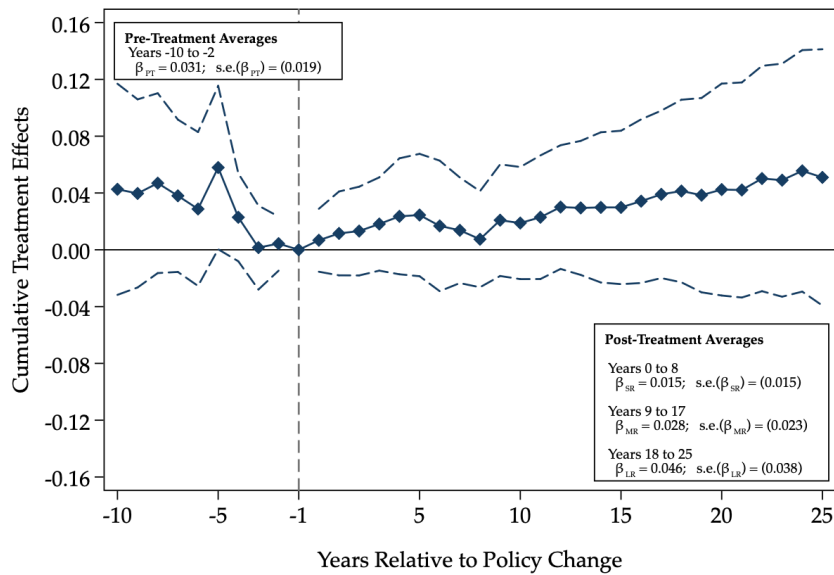
Notes: This figure displays three hypothetical outcome paths and the relationship between contemporaneous (distributed lag) and cumulative (event study) coefficients when units may experience multiple policy changes. For simplicity, the example assumes that the parallel trends assumption for treatment and control units is satisfied. Note that all past distributed lag coefficients (γ 's) must be accumulated to explain the difference between treated and control units in levels (β 's).

Figure 3.4: Dynamic Response of Hourly Wages and Employment to Licensing Events
(Occupational Employment Statistics State Panel 1999-2018)

(a) Log Average Hourly Wage



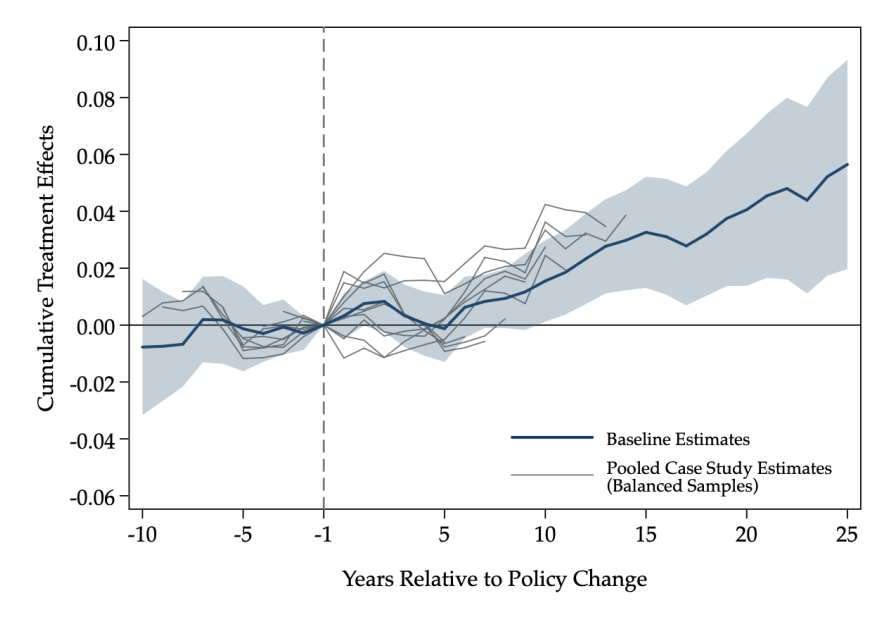
(b) Log Employment-to-Population Ratio



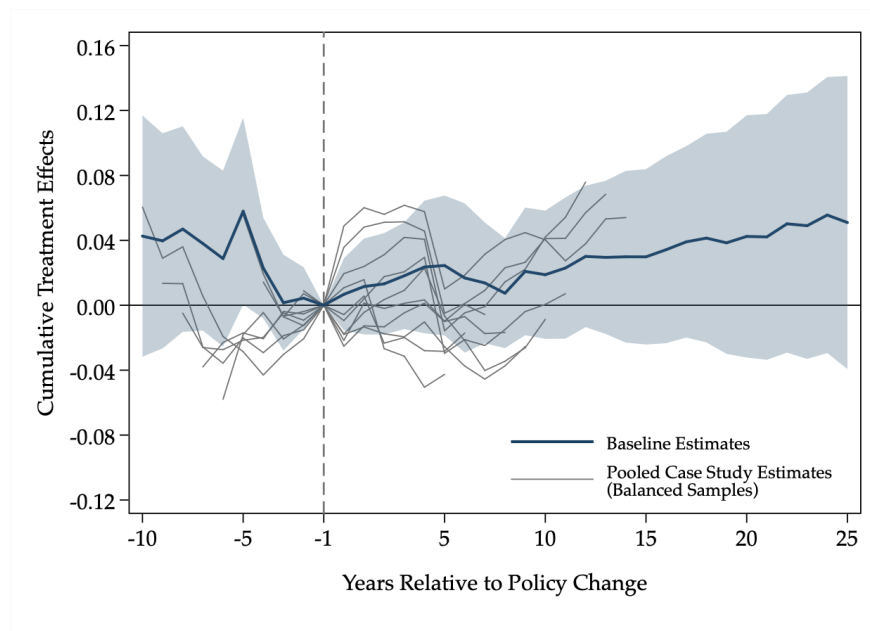
Notes: This figure displays event study estimates for the effect of licensing on wages and employment using data from the Occupational Employment Statistics program. Dashed lines denote 95% confidence intervals for the point estimates based on standard errors clustered by state.

Figure 3.5: Pooled Case Study Estimates of Wage and Employment Effects
(Occupational Employment Statistics State Panel 1999-2018)

(a) Log Average Hourly Wage



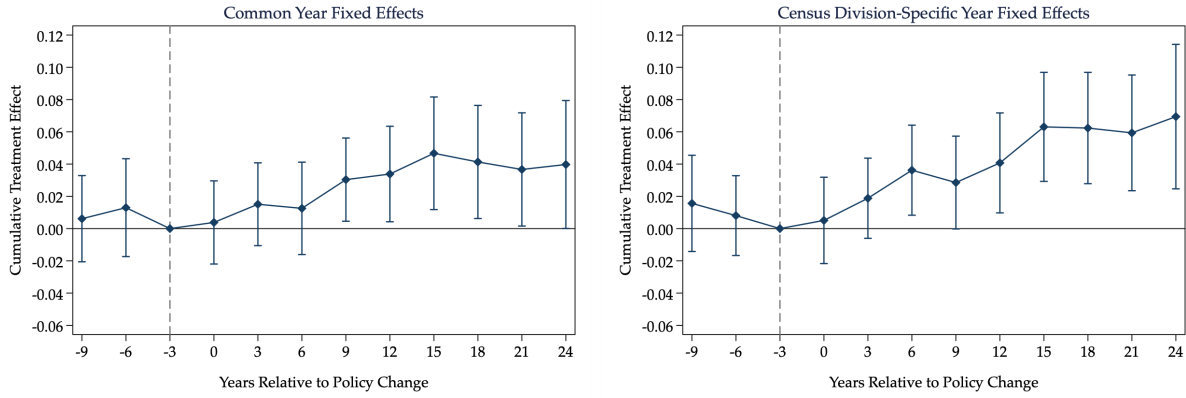
(b) Log Employment-to-Population Ratio



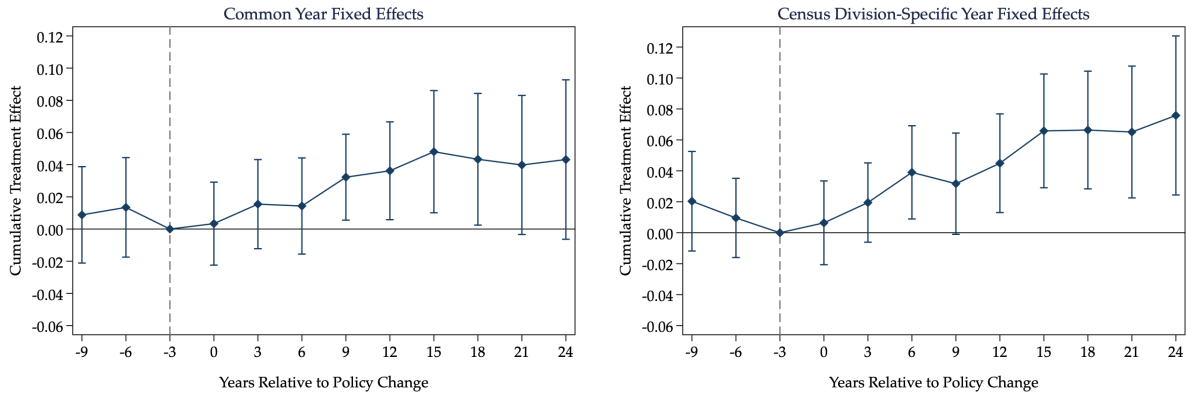
Notes: This figure displays pooled case study estimates for the effect of licensing on wages and employment using data from the Occupational Employment Statistics program. Grey lines denote estimates from samples that have also been balanced in event time, so that each coefficient is estimated using data from the same set of events.

Figure 3.6: Dynamic Response of Hourly Wages to Licensing Events
(Current Population Survey State Panel 1983-2018)

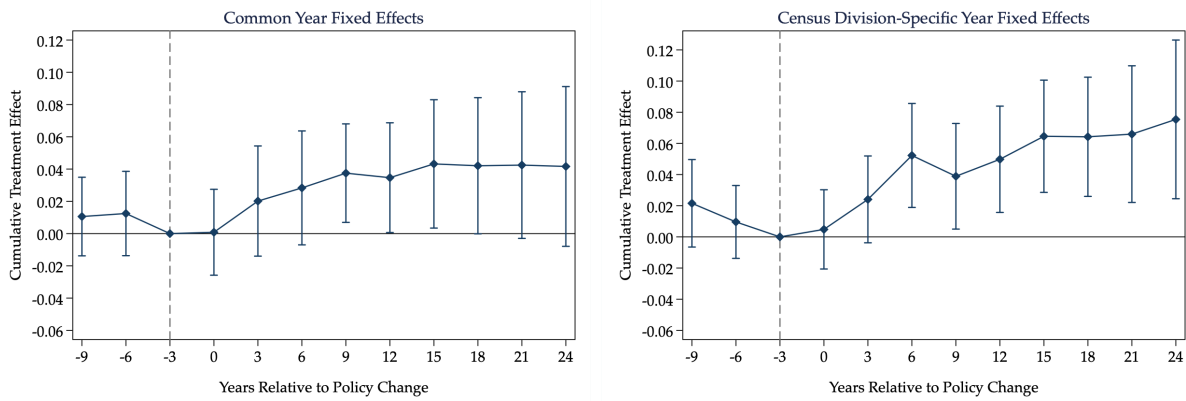
(a) Two-Way Fixed Effects, Licensing Only



(b) Two-Way Fixed Effects and Policy Controls

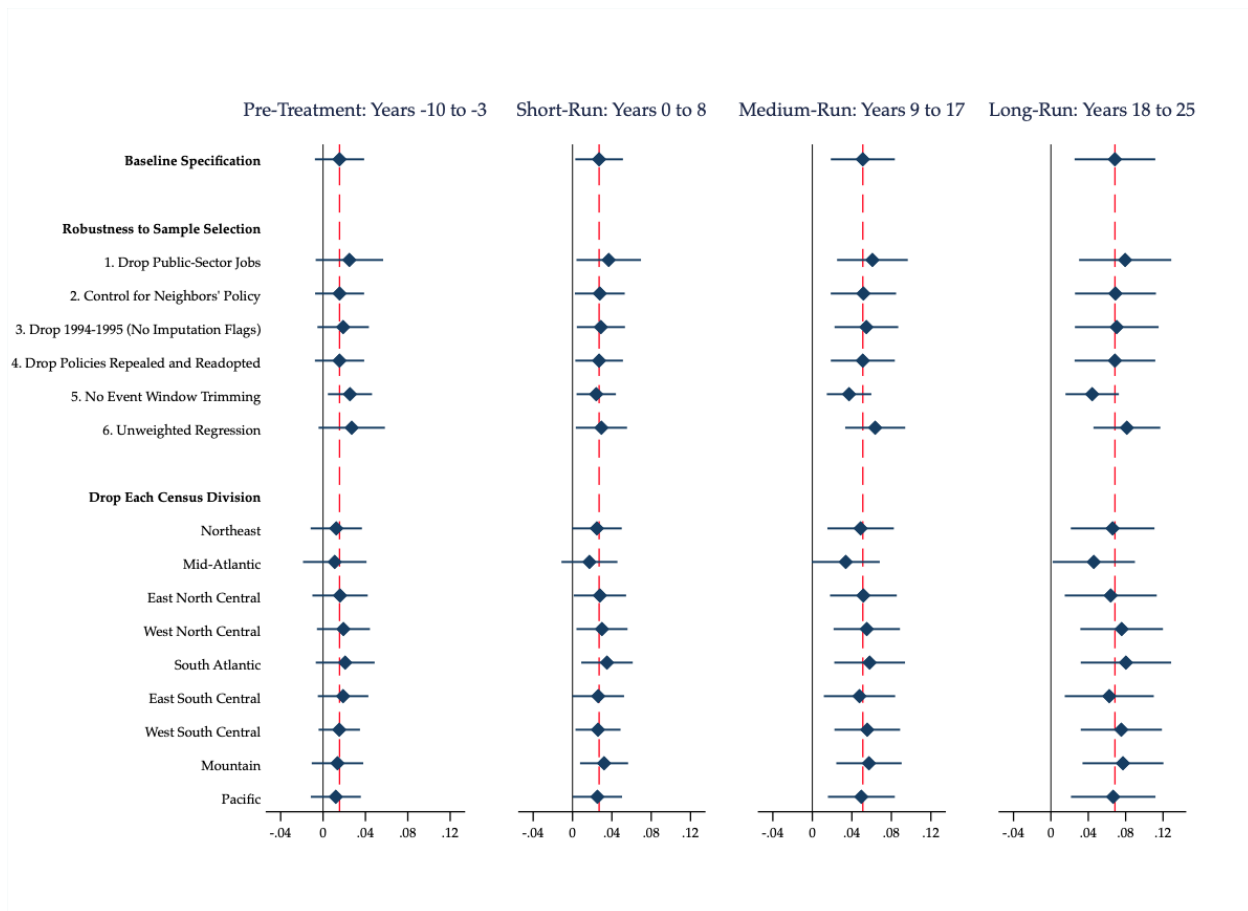


(c) Two-Way Fixed Effects, State by Occupation Group Trends, and Policy Controls



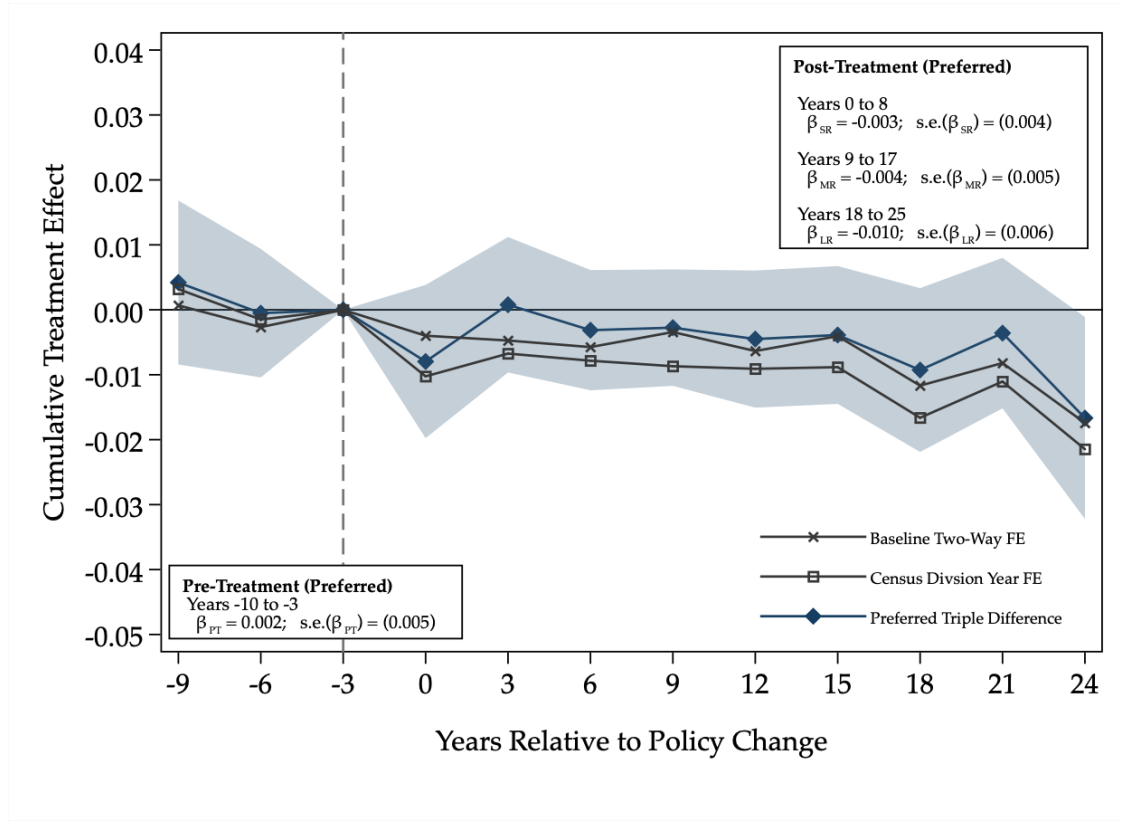
Notes: This figure displays the underlying event study coefficients used to construct the averages reported in Table 3.4. Vertical bands denote 95% confidence intervals based on standard errors clustered by state.

Figure 3.7: Robustness of Hourly Wage Estimates to Alternative Sample Definitions
(State-Occupation Aggregated CPS Panel 1983-2018)



Notes: This figure assesses the robustness of the main hourly wage event study estimates to alternative sample selection decisions. All regressions include occupation by state, occupation by census division and year, and occupation group by state and year fixed effects, which corresponds to the specification estimated in column six of Table 3.4. The top row of the figure replicates this estimate for reference. Horizontal bands denote 95% confidence intervals with standard errors clustered at the state level. Specification 1 drops any occupation where 50% or more of workers are employed in the public sector. Specification 2 controls for policies adopted by neighboring states. Specification 3 drops years 1994 and 1995, for which there are no reliable flags for earnings imputations. Specification 4 drops cases where regulations were repealed and later readopted. Specification 5 allows long-treated units to remain in the sample as controls. Specification 6 reports a regression that is not employment-weighted. The remaining rows drop each of the nine census divisions separately and estimate the regression with data from the remaining eight.

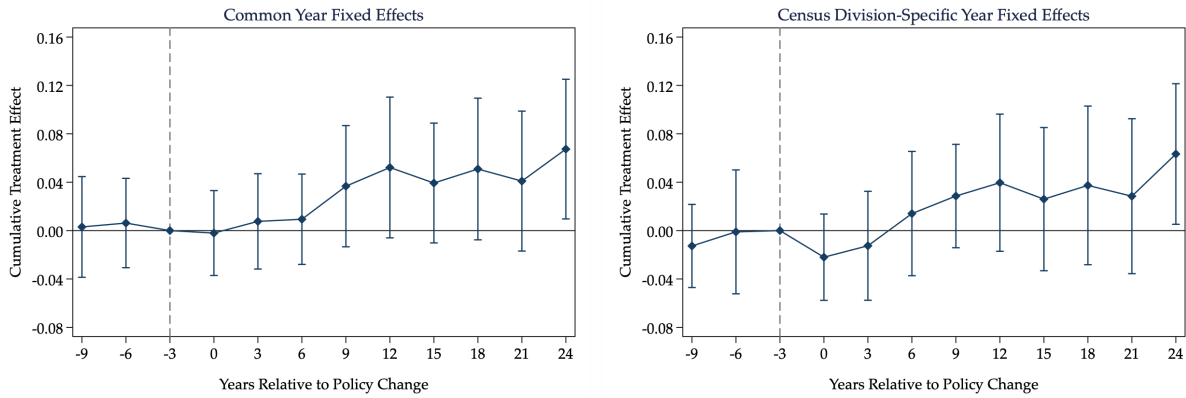
Figure 3.8: Effect of Licensing Events on the Self-Employment Rate (Current Population Survey State Panel 1983-2018)



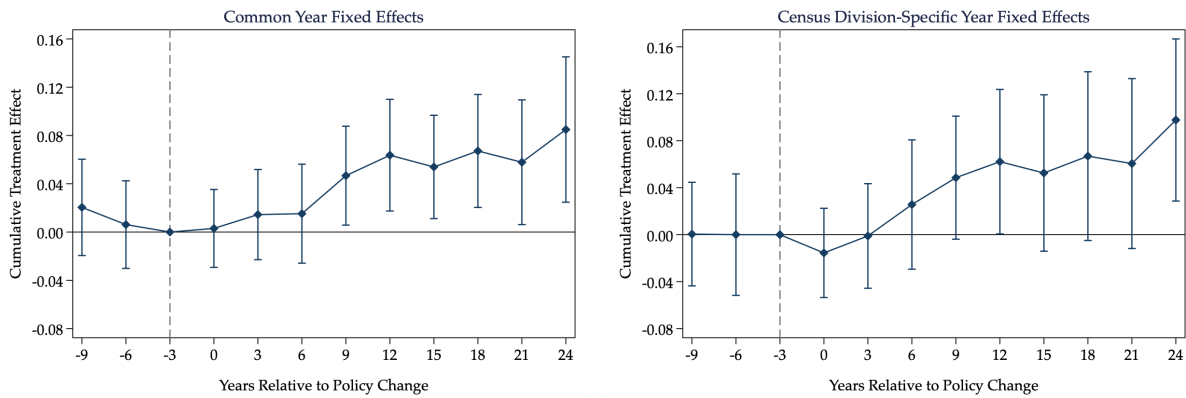
Notes: This figure plots event study estimates for the impact of licensing on the share of workers who are self-employed. The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. All coefficients are normalized relative to the period one to three years prior to treatment. The preferred specification includes state-year fixed effects specific to six broad occupation groups in addition to Census-division specific occupation-year effects. The shaded area denotes the 95% confidence interval based on standard errors clustered by state for the preferred specification.

Figure 3.9: Dynamic Response of Employment to Licensing Events
(Current Population Survey State Panel 1983-2018)

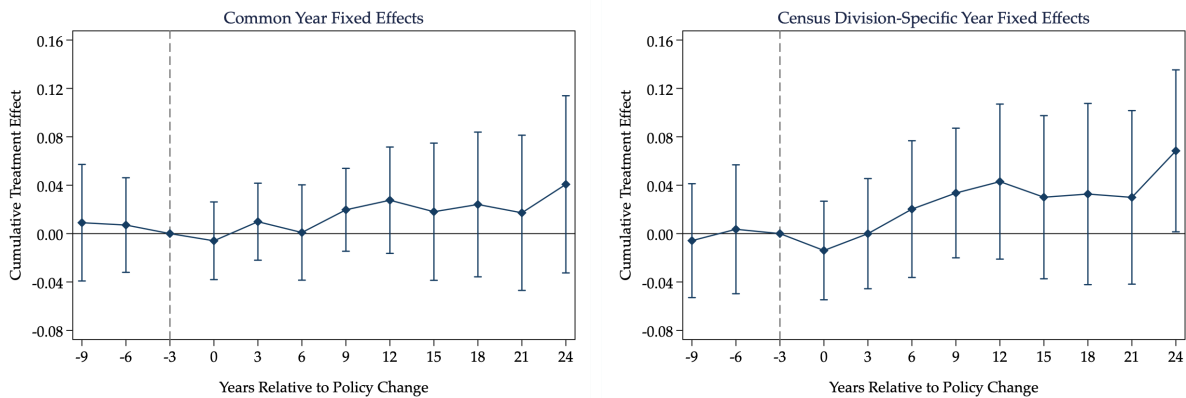
(a) Two-Way Fixed Effects, Licensing Only



(b) Two-Way Fixed Effects and Policy Controls

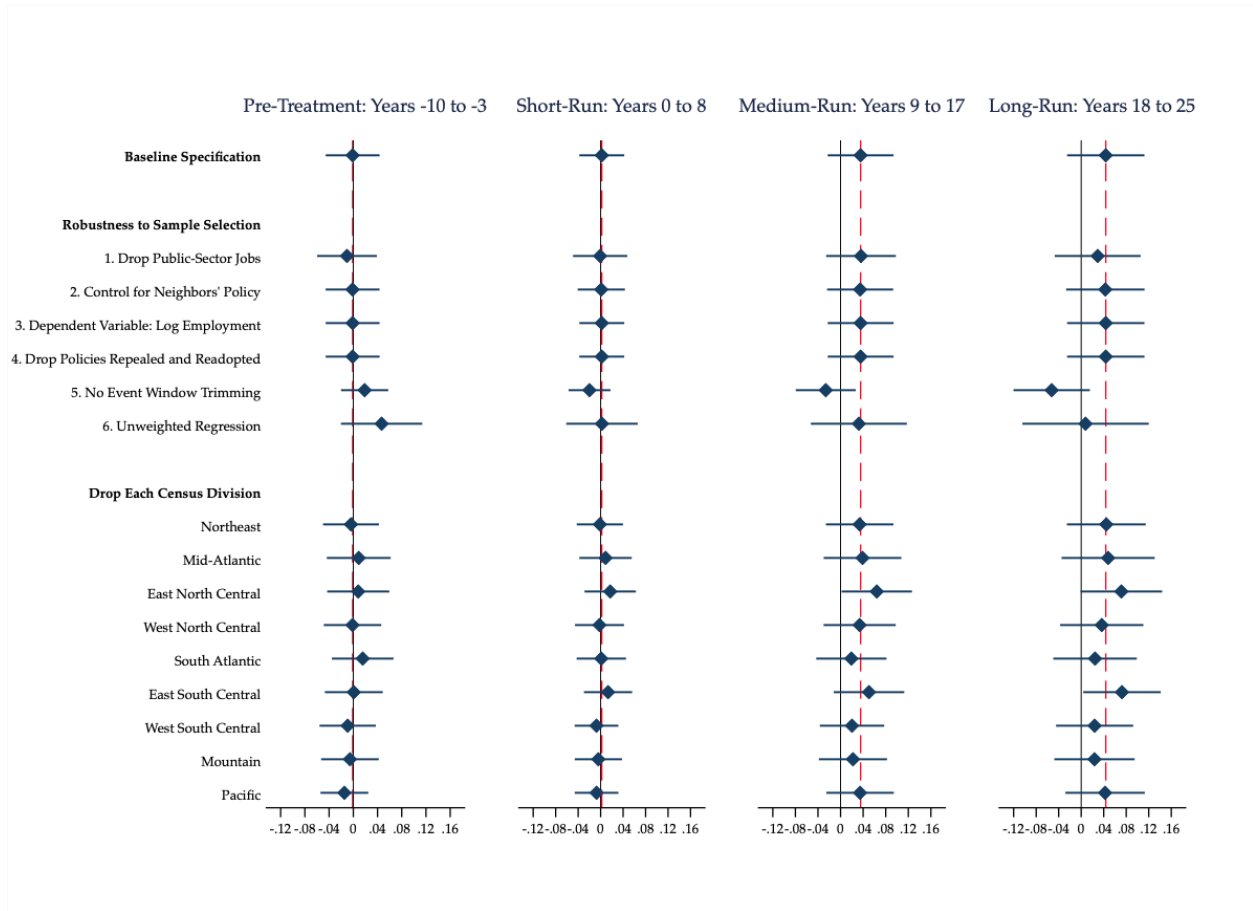


(c) Two-Way Fixed Effects, State by Occupation Group Trends, and Policy Controls



Notes: This figure displays the underlying event study coefficients used to construct the averages reported in Table 3.6. Vertical bands denote 95% confidence intervals based on standard errors clustered by state.

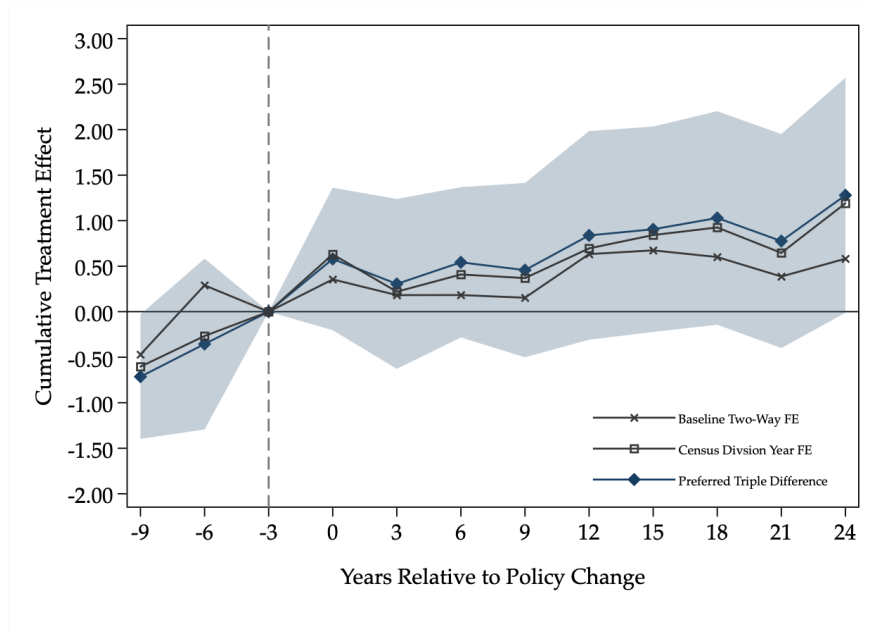
Figure 3.10: Robustness of Employment Estimates to Alternative Sample Definitions
(State-Occupation Aggregated CPS Panel 1983-2018)



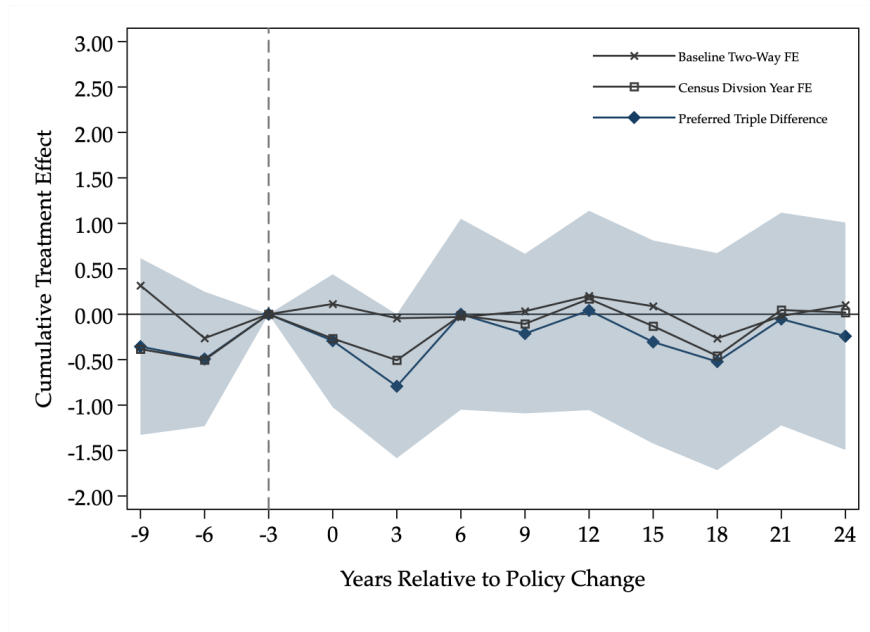
Notes: This figure assesses the robustness of the main employment event study estimates to alternative sample selection decisions. All regressions include occupation by state, occupation by census division and year, and occupation group by state and year fixed effects, which corresponds to the specification estimated in column six of [Table 3.6](#). The top row of the figure replicates this estimate for reference. Horizontal bands denote 95% confidence intervals with standard errors clustered at the state level. Specification 1 drops any occupation where 50% or more of workers are employed in the public sector. Specification 2 controls for policies adopted by neighboring states. Specification 3 uses log employment rather than log employment per-capita as the dependent variable. Specification 4 drops cases where regulations were repealed and later readopted. Specification 5 allows long-treated units to remain in the sample as controls. Specification 6 reports a regression that is not employment-weighted. The remaining rows drop each of the nine census divisions separately and estimate the regression with data from the remaining eight.

Figure 3.11: Dynamic Response of Average Worker Age to Licensing Events
(Current Population Survey State Panel 1983-2018)

(a) Some College or More



(b) High School or Less



Notes: This figure reports event study estimates for the effect of licensing on the average age of workers. The shaded area denotes 95% confidence intervals based on standard errors clustered by state for the preferred specification.

Figure 3.12: Cross-Sectional Differences in Earnings and Employment
(Occupational Employment Statistics State Sample 2015)

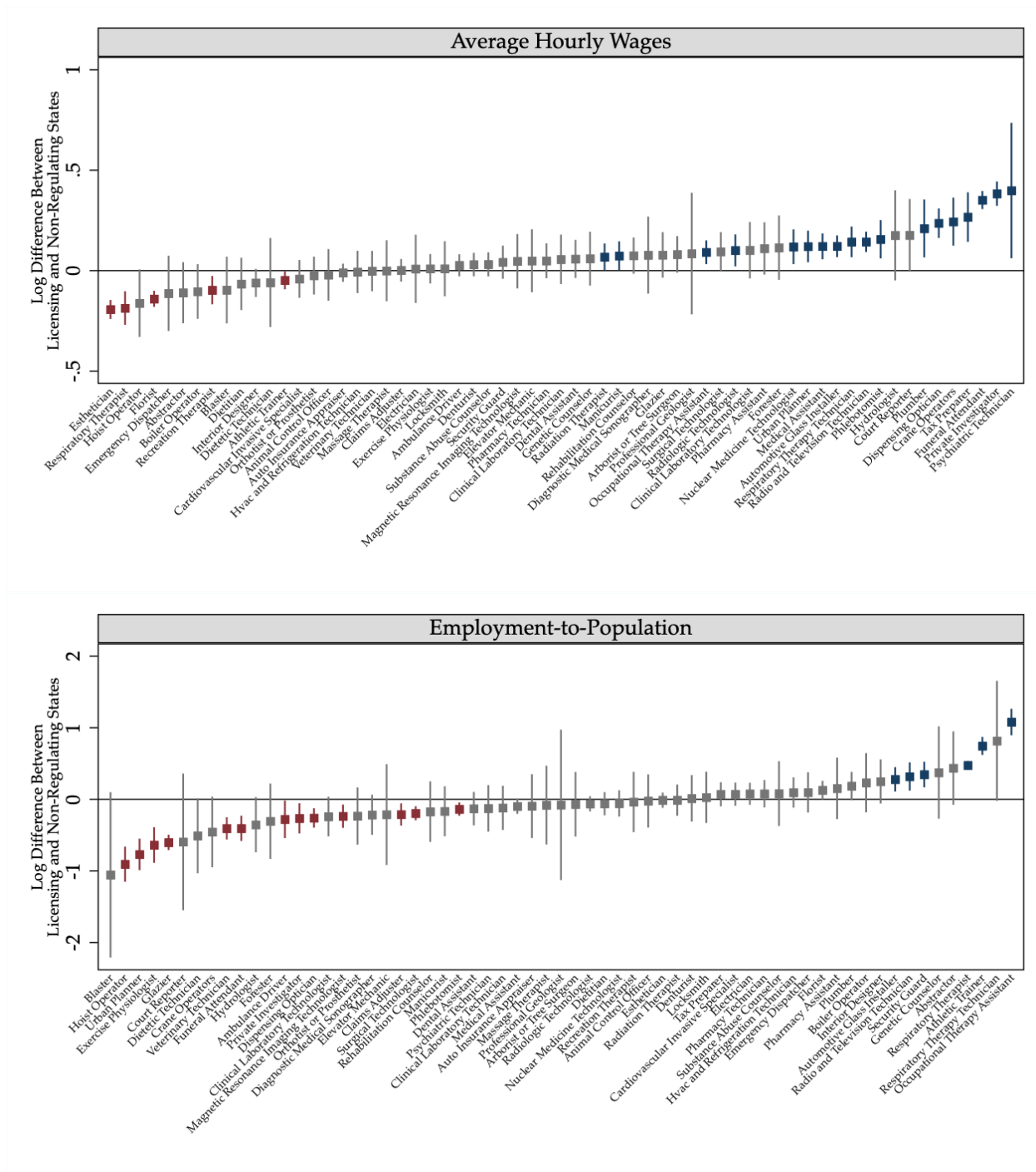
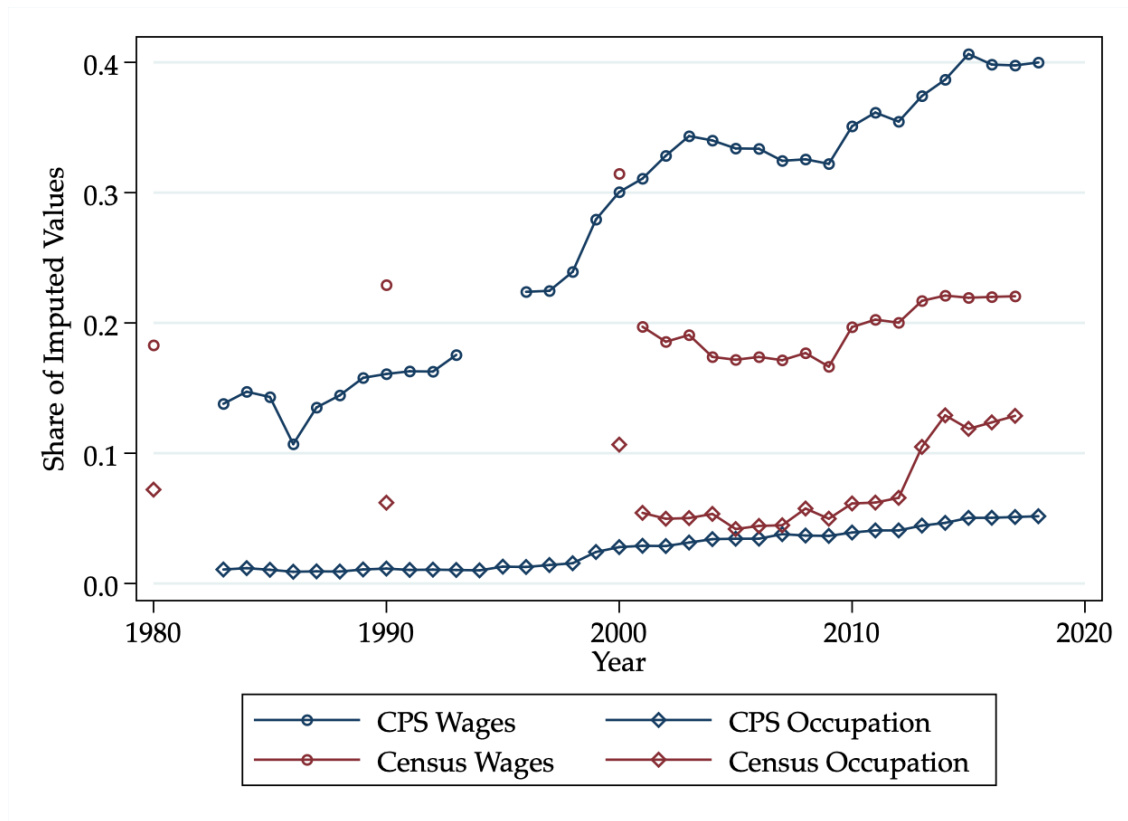


Figure 3.13: Dynamic Response of Industry Employment to Licensing Events (Quarterly Census of Employment and Wages 1990-2018)

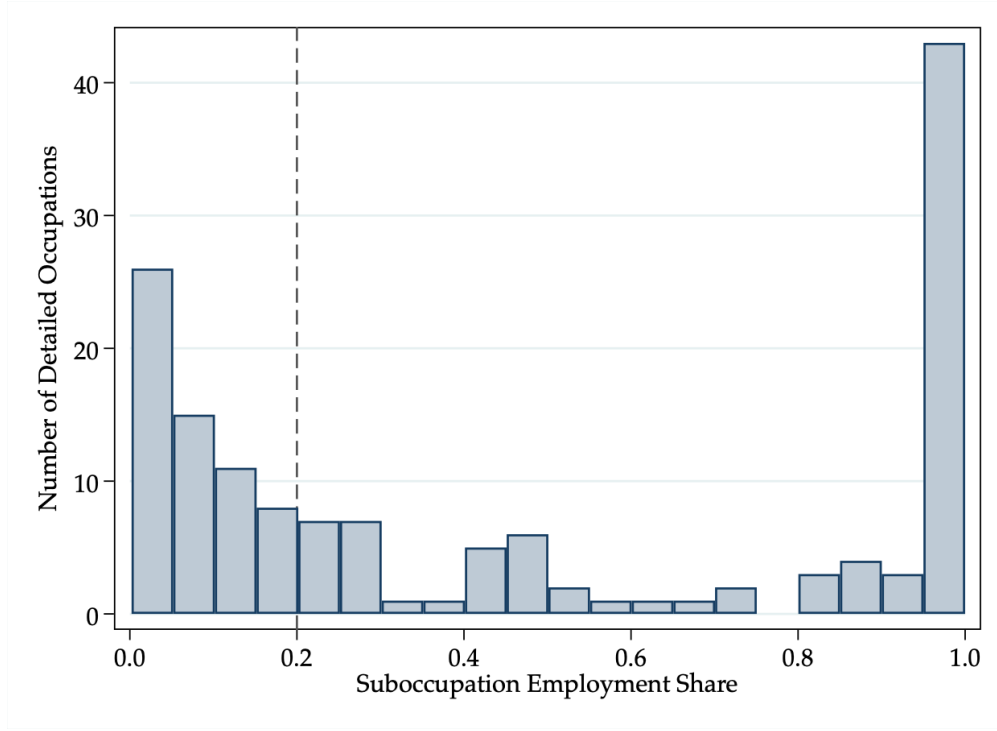


Figure 3.14: Wage and Occupation Imputation Rates in the American Community Survey, Census, and Current Population Survey



Notes: This figure displays the share of observations in each of the main survey data sets used in the project that have imputed wage or occupation variables. These observations are dropped prior to any analysis using these variables.

Figure 3.15: Distribution of Employment Share Estimates for Regulated Suboccupations
(Occupational Employment Statistics 2000)



Notes: This figure displays the distribution of weights, $\hat{\pi}_{jk}$ used to construct the aggregated policy variables described in [Section 3.4.2](#). These weights measure the share of each consistent three-digit Census occupation's total national employment covered by its disaggregated six-digit components.

Table 3.1: Source and Characteristics of Worker Credentials
(Survey of Income and Program Participation 2008 Panel, Wave 13)

	Organization Issuing Professional Certification or License				
	All Sources (1)	Federal Government (2)	State Government (3)	Local Government (4)	Private Organization (5)
<i>Panel A: Credential Attainment</i>					
Share of Respondents Holding Credential	25.92	1.38	16.57	0.60	7.22
Number of Respondents Holding Credential	6,196	329	3,954	144	1,722
Total Observations	23,906	23,906	23,906	23,906	23,906
<i>Panel B: Share of Workers Reporting Requirements to Obtain Credential</i>					
Coursework or Training	92.48	92.46	93.06	83.72	91.96
Examination or Skill Demonstration	91.69	91.67	92.41	78.71	91.14
Continuing Education or Testing	68.01	66.81	72.45	48.42	60.03
Any Demonstration of Competency	97.48	97.59	97.60	89.60	97.82
<i>Panel C: Share of Workers Reporting Reason for Holding Credential</i>					
Required for Current Job	75.52	69.42	81.32	76.70	63.53
Mainly Obtained for Work-Related Reason	95.90	93.96	96.65	92.46	94.83
Mainly Obtained for Personal Interest	4.10	6.04	3.35	7.54	5.17

Notes: This table uses data from the professional licensing, certification, and educational certificates topical module conducted during Wave 13 of the 2008 Survey of Income and Program Participation panel. Interviews were administered between August and November 2012. The sample is limited to adults between the ages of 16 and 64 who participated in the topical module and whose responses are self-reported. Private organizations include professional associations, businesses or non-profits, industry groups, and all other non-governmental sources. Any demonstration of competency indicates that at least one of the listed requirements is needed to obtain or maintain the worker's credential. Averages are computed using the sampling weights provided in the SIPP.

Table 3.2: Summary Statistics for a Subset of Licensed Occupations with Recent Policy Changes
(State Legislation and Occupational Employment Statistics as of 2015)

	States Regulating (1)	States Licensing (2)	Median Year of Licensing (3)	Average Education Level (4)	Average Hourly Wage (5)	Total Employment (6)
Tractor-Trailer Truck Driver	50	50	1977	High School	20.43	1,678,280
Nursing Assistant	50	50	1990	High School	12.89	1,420,570
Certified Public Accountant	50	46	1956	Bachelor's	36.19	1,226,910
Security Guard	33	27	1984	High School	13.68	1,097,660
Plumber	45	37	1950	High School	26.49	391,680
Claims Adjuster	34	34	1958	Some College	30.91	271,600
Emergency Medical Technician	50	50	1975	Some College	17.04	236,890
Physical Therapist	50	50	1959	Master's	41.25	209,690
Radiologic Technologist	41	41	1987	Some College	28.13	195,590
Bus Driver, Transit	50	50	1977	High School	19.31	168,620
Clinical Laboratory Technologist	10	10	1978	Bachelor's	29.74	162,950
Clinical Laboratory Technician	10	10	1978	Bachelor's	19.91	157,610
Speech-Language Pathologist	50	50	1976	Master's	36.97	131,450
Professional Counselor	50	45	1993	Bachelor's	21.67	128,200
Respiratory Therapist	49	49	1992	Some College	28.67	120,330
Phlebotomist	4	4	2001	High School	15.76	118,160
Occupational Therapist	50	50	1984	Bachelor's	39.27	114,660
Physician Assistant	50	50	1973	Bachelor's	47.73	98,470
Veterinary Technician	37	37	1976	Some College	16.00	95,790
Architect	50	50	1933	Bachelor's	39.83	93,720
Massage Therapist	45	42	1996	High School	20.76	92,090
Substance Abuse Counselor	40	33	1998	Bachelor's	20.64	87,090
Manicurist	49	49	1935	High School	11.36	83,840
Dispensing Optician	23	21	1955	Some College	17.70	73,520
Diagnostic Medical Sonographer	3	3	2012	Bachelor's	34.08	61,250
Dietitian	47	27	1992	Bachelor's	28.08	59,740
Psychiatric Technician	5	5	1970	Some College	17.44	58,450
Abstractor	10	7	1936	High School	23.96	54,620
Interior Designer	25	3	1997	Bachelor's	26.69	51,050
Glazier	1	1	2000	High School	21.84	44,230
Funeral Director	50	49	1936	Bachelor's	29.19	42,500
Occupational Therapy Assistant	49	49	1985	Some College	28.05	35,460
Magnetic Resonance Imaging Technologist	3	3	2013	Some College	32.86	33,460
Marriage and Family Therapist	50	44	1996	Master's	25.73	32,070
Professional Geologist	31	29	1994	Master's	50.83	31,800
Private Investigator	45	43	1970	Some College	25.41	30,460
Dietetic Technician	4	2	1989	High School	14.03	28,950
Landscape Architect	50	46	1987	Bachelor's	32.98	19,820
Nuclear Medicine Technologist	37	37	1994	Some College	36.06	19,740
Recreation Therapist	6	4	2006	Bachelor's	22.98	17,880
Locksmith	15	9	2006	High School	19.84	17,800
Automotive Glass Installer	1	1	2001	High School	16.93	17,160
Radiation Therapist	37	37	1990	Some College	40.61	16,930
Animal Control Officer	7	7	1992	High School	16.98	13,180
Audiologist	50	50	1976	Professional	37.22	12,070
Forester	15	11	1974	Bachelor's	29.16	8,590
Orthotist or Prosthetist	16	14	2003	Bachelor's	33.63	7,100
Exercise Physiologist	1	1	1996	Bachelor's	23.91	6,620
Hearing Aid Fitter	50	50	1972	Some College	25.41	5,920
Genetic Counselor	20	17	2013	Master's	35.85	2,520

Notes: This table provides summary statistics for a subset of fifty major occupations that experienced at least one policy change after 1980 (i.e. those that potentially contribute identifying variation in the main analysis). Columns one to three use data collected from state statutes, regulations, and historical session laws. Column four uses data on educational requirements from O*Net Online, the successor of the Dictionary of Occupation Titles. Columns five and six use national data from the Occupational Employment Statistics program.

Table 3.3: Political and Economic Factors Affecting the Timing of State Licensing
(Occupational Licensing Legislation Adopted 1975-2010)

	Dependent Variable: Indicator for Licensing Statute					
	(1)	(2)	(3)	(4)	(5)	(6)
A. Partisan Control of State Government: Legislature x Governor						
Republican × Democrat	1.043 (0.098)	1.044 (0.098)	1.031 (0.099)	1.032 (0.099)	1.027 (0.099)	1.045 (0.085)
Democrat × Republican	1.242** (0.115)	1.237** (0.115)	1.216** (0.113)	1.212** (0.113)	1.207** (0.112)	1.239*** (0.098)
Democrat × Democrat	1.501*** (0.132)	1.499*** (0.131)	1.446*** (0.129)	1.445*** (0.129)	1.445*** (0.129)	1.451*** (0.109)
B. State Economic Conditions						
Log Employment-to-Population Ratio			0.862** (0.051)	0.865** (0.052)	0.863** (0.052)	0.868*** (0.045)
Log Gross State Product			1.075 (0.182)	1.069 (0.181)	1.064 (0.181)	1.006 (0.146)
Urbanization Rate			1.029 (0.042)	1.030 (0.042)	1.031 (0.042)	1.039 (0.036)
Union Coverage			0.964 (0.036)	0.964 (0.036)	0.960 (0.036)	0.915*** (0.030)
C. Occupation-Specific Characteristics						
Log Average Hourly Wage		1.025 (0.031)		1.033 (0.044)	1.032 (0.044)	1.028 (0.036)
Log Employment per capita		0.855 (0.088)		0.875 (0.091)	0.867 (0.091)	0.909 (0.086)
Previously Certified or Registered					1.166 (0.153)	1.296** (0.143)
Neighboring States Licensing (%)					0.893 (0.143)	1.021 (0.138)
P-values for Joint Significance						
Partisan Control of State Government	0.000	0.000	0.000	0.000	0.000	0.000
State Economic Conditions			0.094	0.111	0.092	0.001
Occupation-Specific Characteristics		0.296		0.390	0.445	0.142
Number of Events	1,206	1,206	1,206	1,206	1,206	1,656
Sample Observations	25,777	25,777	25,777	25,777	25,777	39,268
Occupation Sample	6-digit	6-digit	6-digit	6-digit	6-digit	All

Notes: This table reports the results of a discrete-time hazard model estimating the factors correlated with the timing of state licensing statutes. Each column is produced from a separate conditional logistic regression that absorbs occupation-specific hazard rates. All estimates are expressed as the change in the log-odds of adopting a policy in a given year conditional on not having licensed the occupation in the past. State-level variables and occupation-specific wages and employment are reported as the impact of a one standard deviation increase in the variable. Alaska and Hawaii are omitted because they do not border any other states. Nebraska is omitted because it has a unicameral legislature. Occupational wages and employment per capita are measured at the three-digit level by linearly interpolating data between census years prior to 2000. Data on the timing of policy adoption is compiled through an analysis of state session laws. The remaining data sources are as follows: partisan composition of state governments (Klarner, 2013); total state employment (Quarterly Census of Employment and Wages); state population, urbanization, occupational wages and employment (Census Bureau); gross state product (Bureau of Economic Analysis); union membership (Hirsch & Macpherson, 2019). Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.4: Dynamic Response of Average Hourly Wages to Licensing Events
(State-Occupation Aggregated CPS Panel 1983-2018)

<i>Cumulative Treatment Effects</i>	Dependent Variable: Log Average Hourly Wage					
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Treatment Effect	0.010	0.011	0.012	0.012	0.015	0.016
$\hat{\beta}_{PT}$: Years -10 to -3	(0.014)	(0.015)	(0.012)	(0.013)	(0.013)	(0.012)
Short-Run Effect	0.010	0.011	0.016	0.020*	0.022*	0.027**
$\hat{\beta}_{SR}$: Years 0 to 8	(0.010)	(0.011)	(0.014)	(0.011)	(0.012)	(0.012)
Medium-Run Effect	0.037**	0.039**	0.038**	0.044***	0.047***	0.051***
$\hat{\beta}_{MR}$: Years 9 to 17	(0.014)	(0.015)	(0.017)	(0.014)	(0.016)	(0.016)
Long-Run Effect	0.039**	0.042*	0.042*	0.064***	0.069***	0.069***
$\hat{\beta}_{LR}$: Years 18 to 25	(0.018)	(0.022)	(0.022)	(0.018)	(0.021)	(0.021)
Number of Events	1,166	1,166	1,166	1,109	1,109	1,109
Sample Observations	348,249	348,249	348,249	339,660	339,660	339,660
Total Worker Observations	21,745,118	21,745,118	21,745,118	21,618,794	21,618,794	21,618,794
Occupation-Year FE	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓
Policy Controls		✓	✓		✓	✓
Occupation Group-State-Year FE			✓			✓
Occupation-Division-Year FE				✓	✓	✓

Notes: This table reports binned averages of event study estimates for the effect of licensing on hourly wages. The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.5: Robustness of Earnings Estimates to Alternative Income Measures
(State-Occupation Aggregated Census 1980-2000 and ACS 2001-2017)

<i>Cumulative Treatment Effects</i>	Log Hourly		Log Annual		Log Total		Log Total	
	Wage/	Hourly	Wage/	Annual	Hourly	Total	Annual	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-Treatment Effect								
$\hat{\beta}_{PT}$: Years -10 to -3	-0.011 (0.011)	-0.009 (0.011)	-0.007 (0.010)	0.000 (0.009)	-0.005 (0.013)	-0.001 (0.012)	-0.001 (0.009)	0.011 (0.009)
Short-Run Effect								
$\hat{\beta}_{SR}$: Years 0 to 8	0.000 (0.011)	0.000 (0.011)	0.003 (0.010)	0.008 (0.010)	-0.002 (0.012)	-0.003 (0.013)	0.003 (0.010)	0.009 (0.010)
Medium-Run Effect								
$\hat{\beta}_{MR}$: Years 9 to 17	0.023* (0.012)	0.024* (0.014)	0.025* (0.013)	0.030* (0.015)	0.019 (0.013)	0.020 (0.016)	0.027* (0.014)	0.033** (0.016)
Long-Run Effect								
$\hat{\beta}_{LR}$: Years 18 to 25	0.038** (0.018)	0.034* (0.018)	0.043** (0.020)	0.043** (0.020)	0.035* (0.019)	0.030 (0.020)	0.043** (0.021)	0.044** (0.021)
Number of Events	1,062	1,062	1,062	1,062	1,065	1,065	1,065	1,065
Sample Observations	235,737	235,737	235,737	235,737	238,426	238,426	238,439	238,439
Total Worker Observations	26,991,334	26,991,334	26,991,334	26,991,334	26,997,863	26,997,863	26,997,881	26,997,881
Occupation-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Occupation-Division-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Policy Controls	✓	✓	✓	✓	✓	✓	✓	✓
Occupation Group-State-Year FE		✓		✓		✓		✓

Notes: This table reports binned averages of event study estimates for the effect of licensing on annual earnings. The estimation sample is an aggregated state-occupation panel constructed from the 1980-2000 Census and 2001-2017 American Community Survey. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.6: Dynamic Response of Occupational Employment to Licensing Events
(State-Occupation Aggregated CPS Panel 1983-2018)

<i>Cumulative Treatment Effects</i>	Dependent Variable: Log Employment-to-Population Ratio					
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Treatment Effect	0.005	0.013	0.008	-0.007	0.000	-0.001
$\hat{\beta}_{PT}$: Years -10 to -3	(0.017)	(0.017)	(0.020)	(0.018)	(0.020)	(0.022)
Short-Run Effect	0.005	0.011	0.002	-0.007	0.003	0.002
$\hat{\beta}_{SR}$: Years 0 to 8	(0.015)	(0.014)	(0.013)	(0.018)	(0.019)	(0.020)
Medium-Run Effect	0.043*	0.055***	0.022	0.031	0.054*	0.036
$\hat{\beta}_{MR}$: Years 9 to 17	(0.025)	(0.020)	(0.021)	(0.024)	(0.028)	(0.029)
Long-Run Effect	0.053*	0.070***	0.027	0.043	0.075**	0.044
$\hat{\beta}_{LR}$: Years 18 to 25	(0.028)	(0.025)	(0.032)	(0.030)	(0.034)	(0.034)
Number of Events	1,189	1,189	1,189	1,145	1,145	1,145
Sample Observations	410,831	410,831	410,831	406,459	406,459	406,459
Total Worker Observations	22,060,862	22,060,862	22,060,862	21,982,018	21,982,018	21,982,018
Occupation-Year FE	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓
Policy Controls		✓	✓		✓	✓
Occupation Group-State-Year FE			✓			✓
Occupation-Division-Year FE				✓	✓	✓

Notes: This table reports binned averages of event study estimates for the effect of licensing on employment by occupation. The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.7: Dynamic Response of Occupational Employment to Licensing Events
(State-Occupation Aggregated Census 1980-2000 and ACS 2001-2017)

<i>Cumulative Treatment Effects</i>	Dependent Variable: Log Employment-to-Population Ratio					
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Treatment Effect	0.026	0.021	0.033*	0.040**	0.038**	0.041**
$\hat{\beta}_{PT}$: Years -10 to -3	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.019)
Short-Run Effect	0.011	0.011	0.005	0.016	0.016	0.011
$\hat{\beta}_{SR}$: Years 0 to 8	(0.013)	(0.015)	(0.014)	(0.021)	(0.022)	(0.025)
Medium-Run Effect	0.048***	0.045***	0.012	0.040*	0.042*	0.015
$\hat{\beta}_{MR}$: Years 9 to 17	(0.015)	(0.015)	(0.018)	(0.021)	(0.022)	(0.026)
Long-Run Effect	0.062***	0.061***	0.019	0.054*	0.053*	0.017
$\hat{\beta}_{LR}$: Years 18 to 25	(0.022)	(0.021)	(0.027)	(0.029)	(0.031)	(0.036)
Number of Events	1,103	1,103	1,103	1,066	1,066	1,066
Sample Observations	243,006	243,006	243,006	241,744	241,744	241,744
Total Worker Observations	27,096,327	27,096,327	27,096,327	27,002,162	27,002,162	27,002,162
Occupation-Year FE	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓
Policy Controls		✓	✓		✓	✓
Occupation Group-State-Year FE			✓			✓
Occupation-Division-Year FE				✓	✓	✓

Notes: This table reports binned averages of event study estimates for the effect of licensing on employment by occupation. The estimation sample is an aggregated state-occupation panel constructed from the 1980-2000 Census and 2001-2017 American Community Survey. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.8: Robustness of Preferred Earnings Estimates to Alternative Treatment Aggregation Decisions
(State-Occupation Aggregated CPS Panel 1983-2018)

	Dependent Variable: Log Average Hourly Wage								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pre-Treatment Effect									
$\hat{\beta}_{PT}$: Years -10 to -3	0.016 (0.012)	0.012 (0.012)	0.021** (0.010)	0.010 (0.012)	0.016 (0.012)	0.019* (0.010)	0.010 (0.012)	0.015 (0.012)	0.019* (0.010)
Short-Run Effect									
$\hat{\beta}_{SR}$: Years 0 to 8	0.026** (0.013)	0.029** (0.012)	0.033*** (0.011)	0.020 (0.012)	0.027** (0.012)	0.033*** (0.011)	0.021 (0.013)	0.027** (0.012)	0.031*** (0.012)
Medium-Run Effect									
$\hat{\beta}_{MR}$: Years 9 to 17	0.053*** (0.017)	0.060*** (0.016)	0.068*** (0.016)	0.040** (0.016)	0.051*** (0.016)	0.067*** (0.016)	0.041** (0.017)	0.051*** (0.016)	0.065*** (0.017)
Long-Run Effect									
$\hat{\beta}_{LR}$: Years 18 to 26	0.069*** (0.023)	0.081*** (0.021)	0.093*** (0.020)	0.050** (0.022)	0.069*** (0.021)	0.092*** (0.020)	0.052** (0.023)	0.069*** (0.022)	0.091*** (0.020)
Number of Events	1,360	1,056	867	1,428	1,109	807	1,427	1,149	817
Sample Observations	337,513	340,621	343,345	334,466	339,660	344,742	334,392	340,658	344,689
Total Worker Observations	19,480,896	21,618,555	21,943,168	19,396,662	21,618,794	21,958,198	19,380,075	21,672,842	21,957,894
Occupation-Division-Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Policy Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Occupation Group-State-Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year of Estimated Employment Share	1980	1980	1980	2000	2000	2000	2015	2015	2015
Coverage Threshold for Event Sample	None	20%	50%	None	20%	50%	None	20%	50%

Notes: This table replicates my preferred event study specification for hourly wages using alternative parameters for treatment aggregation (Equation 3.2). The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text under the assumptions shown in the bottom rows of the table. Column five repeats the specification presented in the main text for reference. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.9: Robustness of Preferred Employment Estimates to Alternative Treatment Aggregation Decisions
(State-Occupation Aggregated CPS Panel 1983-2018)

	Dependent Variable: Log Employment-to-Population Ratio								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pre-Treatment Effect									
$\hat{\beta}_{PT}$: Years -10 to -3	-0.012 (0.023)	-0.013 (0.024)	-0.013 (0.028)	0.001 (0.024)	-0.001 (0.022)	-0.012 (0.027)	0.003 (0.023)	0.002 (0.022)	-0.013 (0.026)
Short-Run Effect									
$\hat{\beta}_{SR}$: Years 0 to 8	-0.009 (0.019)	-0.005 (0.020)	0.004 (0.024)	-0.003 (0.021)	0.002 (0.020)	0.006 (0.024)	-0.003 (0.021)	0.002 (0.020)	0.003 (0.024)
Medium-Run Effect									
$\hat{\beta}_{MR}$: Years 9 to 17	0.023 (0.027)	0.026 (0.028)	0.037 (0.029)	0.026 (0.032)	0.036 (0.029)	0.043 (0.028)	0.025 (0.032)	0.035 (0.030)	0.038 (0.029)
Long-Run Effect									
$\hat{\beta}_{LR}$: Years 18 to 26	0.039 (0.036)	0.040 (0.034)	0.056 (0.040)	0.035 (0.042)	0.044 (0.034)	0.061 (0.038)	0.032 (0.040)	0.040 (0.035)	0.053 (0.039)
Number of Events	1,394	1,089	900	1,462	1,145	842	1,461	1,185	852
Sample Observations	404,599	408,839	411,731	400,691	406,459	413,565	400,617	407,468	413,476
Total Worker Observations	19,845,802	21,993,173	22,317,632	19,757,374	21,982,018	22,333,803	19,740,787	22,036,206	22,333,422
Occupation-Division-Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Policy Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Occupation Group-State-Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year of Estimated Employment Share	1980	1980	1980	2000	2000	2000	2015	2015	2015
Coverage Threshold for Event Sample	None	20%	50%	None	20%	50%	None	20%	50%

Notes: This table replicates my preferred event study specification for employment using alternative parameters for treatment aggregation (Equation 3.2). The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text under the assumptions shown in the bottom rows of the table. Column five repeats the specification presented in the main text for reference. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.10: Dynamic Response of Average Educational Attainment to Licensing Events
(State-Occupation Aggregated CPS Panel 1983-2018)

	Dependent Variable: Average Years of Education					
	Full Sample (1)	Some College or More (2)	High School or Less (3)	Full Sample (4)	Some College or More (5)	High School or Less (6)
<i>Cumulative Treatment Effects</i>						
Pre-Treatment Effect	0.032	0.021	0.049	0.044	0.012	0.078
$\hat{\beta}_{PT}$: Years -10 to -3	(0.038)	(0.053)	(0.050)	(0.037)	(0.055)	(0.048)
Short-Run Effect	0.017	0.035	-0.001	0.039	0.028	0.063
$\hat{\beta}_{SR}$: Years 0 to 8	(0.038)	(0.049)	(0.047)	(0.040)	(0.044)	(0.058)
Medium-Run Effect	0.082	0.112	0.060	0.079	0.109*	0.082
$\hat{\beta}_{MR}$: Years 9 to 17	(0.054)	(0.068)	(0.070)	(0.055)	(0.064)	(0.073)
Long-Run Effect	0.137**	0.169**	0.122	0.122*	0.181**	0.140
$\hat{\beta}_{LR}$: Years 18 to 25	(0.068)	(0.079)	(0.092)	(0.065)	(0.079)	(0.085)
Number of Events	1,145	678	467	1,145	678	467
Sample Observations	406,459	149,756	255,712	406,459	149,756	255,712
Total Worker Observations	21,982,018	8,658,441	13,323,577	21,982,018	8,658,441	13,323,577
Occupation-Division-Year FE	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓
Policy Controls	✓	✓	✓	✓	✓	✓
Occupation Group-State-Year FE				✓	✓	✓

Notes: This table reports binned averages of event study estimates for the effect of licensing on average educational attainment by occupation. Columns one and four use the full sample. The remaining columns split the sample into occupations typically requiring some college education or more versus a high school degree or less based on data from O*Net. The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.11: Impact of Licensing on Hourly Wages with Demographic and Compositional Controls
(State-Occupation Aggregated CPS Panel 1983-2018)

<i>Cumulative Treatment Effects</i>	Dependent Variable: Log Average Hourly Wage					
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Treatment Effect	0.017	0.016	0.015	0.017	0.017	0.016
$\hat{\beta}_{PT}$: Years -10 to -3	(0.014)	(0.013)	(0.013)	(0.012)	(0.012)	(0.011)
Short-Run Effect	0.020*	0.020*	0.021**	0.024**	0.024**	0.024**
$\hat{\beta}_{SR}$: Years 0 to 8	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)	(0.009)
Medium-Run Effect	0.041***	0.041***	0.044***	0.046***	0.044***	0.045***
$\hat{\beta}_{MR}$: Years 9 to 17	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)	(0.012)
Long-Run Effect	0.059***	0.059***	0.063***	0.060***	0.058***	0.058***
$\hat{\beta}_{LR}$: Years 18 to 25	(0.019)	(0.019)	(0.019)	(0.019)	(0.017)	(0.017)
Number of Events	1,125	1,125	1,123	1,109	1,109	1,107
Sample Observations	339,660	339,660	338,640	339,660	339,660	338,640
Total Worker Observations	21,618,794	21,618,794	21,611,000	21,618,794	21,618,794	21,611,000
Occupation-Division-Year FE	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓
Education and Potential Experience	✓	✓	✓	✓	✓	✓
Demographic Composition		✓	✓		✓	✓
Class of Worker and Union Coverage			✓			✓
Occupation Group-State-Year FE				✓	✓	✓

Notes: This table reports binned averages of event study estimates for the effect of licensing on hourly wages. The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.12: Heterogeneous Impact of Licensing on Average Hourly Wages and Employment by Occupation Characteristics
(State-Occupation Aggregated CPS Panel 1983-2018)

	Log Average Hourly Wage				Log Employment-to-Population Ratio			
	Pre-Treatment (1)	Short-Run (2)	Medium-Run (3)	Long-Run (4)	Pre-Treatment (5)	Short-Run (6)	Medium-Run (7)	Long-Run (8)
A. Level of Education, Experience, or Training Required								
Little to Medium Preparation	0.020 (0.018)	0.033* (0.020)	0.054*** (0.020)	0.075*** (0.028)	-0.004 (0.027)	0.004 (0.028)	0.052* (0.030)	0.049 (0.039)
Considerable or Extensive Preparation	0.013 (0.015)	0.015 (0.016)	0.035 (0.037)	0.038 (0.053)	0.004 (0.041)	-0.008 (0.033)	-0.010 (0.056)	0.005 (0.070)
B. Occupation Group								
Professional Specialty Occupations	-0.008 (0.010)	0.042** (0.019)	0.117*** (0.043)	0.177*** (0.065)	0.017 (0.039)	0.018 (0.037)	0.060 (0.076)	0.097 (0.112)
All Other Occupations	0.024 (0.020)	0.036* (0.020)	0.052** (0.020)	0.066** (0.028)	-0.017 (0.027)	-0.003 (0.029)	0.035 (0.034)	0.032 (0.039)
C. Consequence of Error								
Below Median Seriousness	0.015 (0.024)	0.004 (0.035)	0.044 (0.039)	0.047 (0.046)	-0.122** (0.050)	-0.143** (0.059)	-0.184** (0.071)	-0.248*** (0.083)
Above Median Seriousness	0.011 (0.012)	0.028** (0.012)	0.050*** (0.017)	0.073*** (0.023)	0.016 (0.023)	0.025 (0.022)	0.065** (0.030)	0.091** (0.035)

Notes: This table reports binned averages of event study estimates for the effect of licensing on hourly wages. The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. Treatment events occur at the six-digit level and are aggregated using the procedure explained in the text. These regressions use variation from all available licensing events in the data, including those with overlapping event windows and units treated before the outcome window. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.13: Response of Hourly Wage Distribution to Licensing Events (Occupational Employment Statistics State Panel 1999-2018)

	Dependent Variable: Log Hourly Wage at Percentile				
<i>Cumulative Treatment Effects</i>	(1)	(2)	(3)	(4)	(5)
Pre-Treatment Effect	0.002	0.006	-0.000	-0.003	-0.008
$\hat{\beta}_{PT}$: Years -10 to -2	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
Short-Run Effect	0.011**	0.016***	0.010*	0.001	-0.007
$\hat{\beta}_{SR}$: Years 0 to 8	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Medium-Run Effect	0.040***	0.044***	0.037***	0.016**	-0.005
$\hat{\beta}_{MR}$: Years 9 to 17	(0.012)	(0.012)	(0.011)	(0.008)	(0.011)
Long-Run Effect	0.054***	0.058***	0.062***	0.041***	0.012
$\hat{\beta}_{LR}$: Years 18 to 25	(0.019)	(0.018)	(0.022)	(0.015)	(0.016)
Wage Percentile	10th	25th	50th	75th	90th
Number of Events	1,429	1,429	1,429	1,429	1,429
Sample Observations	530,908	530,908	530,908	530,908	530,908
Occupation-Year FE	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓
Policy Controls	✓	✓	✓	✓	✓

Notes: This table displays event study estimates for the impact of licensing on the distribution of hourly wages using data from the Occupational Employment Statistics program. In this sample, treatment and outcome variables are observed at the six-digit level. All coefficients are normalized relative to the year prior to the statutory effective date of the law. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.14: Construction of Consistent Standard Occupational Classification Codes

Aggregated Occupation Title	Original Components	
	2000 SOC Codes	2010 SOC Codes
Farmers, Ranchers, and Other Agricultural Managers	11-9011, 11-9012	11-9013
Funeral Directors	11-9061	11-9061, 39-4031
Compliance Officers	13-1041	13-1041, 33-9093
Human Resources, Training, and Labor Relations Specialists, All Other	13-1071, 13-1079	13-1071, 13-1075
Business Operations, Marketing, and Public Relations Specialists	13-1121, 13-1199, 19-3021	13-1121, 13-1161, 13-1199
	27-3031	27-3031
Computer Systems Analysts and Computer Scientists	15-1051, 15-1071, 15-1081	15-1121, 15-1122, 15-1134, 15-1142, 15-1143, 15-1152
Community and Social Service Specialists, All Other	21-1091, 21-1099	21-1091, 21-1094, 21-1099
Paralegals, Legal Assistants and Law Clerks	23-2011, 23-2092	23-1012, 23-2011
Special Education Teachers, Preschool, Kindergarten, and Elementary School	25-2041	25-2051, 25-2052
Teachers and Instructors, All Other	25-3099	25-2059, 25-3099
Registered Nurses	29-1111	29-1141, 29-1151, 29-1161 29-1171
Therapists, All Other	29-1129	29-1128, 29-1129
Radiologic Technologists and Technicians	29-2034	29-2034, 29-2035
Health Technologists and Technicians, All Other	29-2099	29-2057, 29-2092, 29-2099
Healthcare Practitioners and Technical Workers, All Other	29-9099	29-9092, 29-9099
Nursing Aides, Orderlies, and Attendants	31-1012	31-1014, 31-1015
Healthcare Support Workers, All Other	31-9099	31-9097, 31-9099
Sales and Related Workers, All Other	41-9099	13-1131, 41-9099
Office and Administrative Support Workers, All Other	43-9199	43-3099, 43-9199
Construction and Related Workers, All Other	47-4099	47-2231, 47-4099
Installation, Maintenance, and Repair Workers, All Other	49-9099	49-9081, 49-9099
Printing and Binding Workers	51-5011, 51-5012, 51-5021 51-5023	51-5112, 51-5113
Photographic Process Workers and Processing Machine Operators	51-9131, 51-9132	51-9151
Production Workers, All Other	51-9199	51-3099, 51-9199

Notes: This table lists aggregations of the Standard Occupational Classification system that are made to bridge revisions to occupation definitions adopted in 2010 and create a balanced panel of codes with consistent longitudinal coverage. Codes that do not appear in this table are already comparable in the 2000 and 2010 SOC systems.

Table 3.15: Aggregations Extending the Occ1990dd Balanced Occupation Panel to Cover 2010 Census Codes

Aggregated Occupation Title	Occ1990dd Codes
Managers, administrators, and researchers, n.e.c.	13, 22, 37, 166
Health technologists and technicians, n.e.c.	208, 678
Farmers (managers, owners, and tenants)	473, 475
Masons, tilers, carpet, and structural metal workers	563, 597
Painters and paperhangers	579, 583
Model makers, patternmakers, and molding machine setters	645, 719
Other precision and craft workers	644, 684, 703, 706, 708, 709, 723, 724
Shoe and leather workers	669, 745
Printers, binders, and typesetters	679, 734, 736
Miscellaneous production workers	764, 779
Subway, streetcar, and other rail transportation workers	824, 825

Notes: This table lists aggregations of the occ1990dd classification system that are made to extend the balanced occupation panel to cover revisions to Census occupation definitions adopted after 2010. No changes are necessary for codes that do not appear in this table.

Table 3.16: Impact of Licensing Events on Wage and Occupation Imputation Rate
(State-Occupation Aggregated CPS Panel 1983-2018)

<i>Cumulative Treatment Effects</i>	Hourly Wage Imputation Rate			Occupation Imputation Rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Treatment Effect	-0.014	-0.010	-0.006	0.003	-0.001	-0.000
$\hat{\beta}_{PT}$: Years -10 to -3	(0.010)	(0.009)	(0.009)	(0.003)	(0.003)	(0.003)
Short-Run Effect	0.005	0.009	0.002	0.000	-0.002	-0.002
$\hat{\beta}_{SR}$: Years 0 to 8	(0.008)	(0.009)	(0.009)	(0.003)	(0.003)	(0.003)
Medium-Run Effect	0.009	0.007	-0.004	-0.001	-0.004	-0.005
$\hat{\beta}_{MR}$: Years 9 to 17	(0.012)	(0.011)	(0.010)	(0.004)	(0.003)	(0.005)
Long-Run Effect	0.005	0.003	-0.016	-0.005	-0.006	-0.006
$\hat{\beta}_{LR}$: Years 18 to 25	(0.015)	(0.012)	(0.010)	(0.005)	(0.005)	(0.007)
Number of Events	1,172	1,123	1,123	1,172	1,123	1,123
Sample Observations	369,929	362,919	362,919	370,024	363,028	363,028
Total Worker Observations	21,943,216	21,838,698	21,838,698	21,944,842	21,840,577	21,840,577
Occupation-Year FE	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓
Policy Controls	✓	✓	✓	✓	✓	✓
Occupation-Division-Year FE		✓	✓		✓	✓
Occupation Group-State-Year FE			✓			✓

Notes: This table replicates the main event study specifications with the imputation rate for hourly wages and occupation as the dependent variables. The estimation sample is an aggregated state-occupation panel constructed from the 1983-2018 Current Population Survey and a balanced panel of three-digit occupation codes. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell, which is constructed using sample-weighted annual occupational employment share estimates from the full monthly CPS scaled by total employment. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Table 3.17: Impact of Licensing Events on Earnings and Occupation Imputation Rate
(State-Occupation Aggregated Census 1980-2000 and ACS 2001-2017)

<i>Cumulative Treatment Effects</i>	Earned Income Imputation Rate			Occupation Imputation Rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Treatment Effect	-0.004	-0.007	-0.005	-0.006	-0.002	-0.003
$\hat{\beta}_{PT}$: Years -10 to -3	(0.006)	(0.007)	(0.005)	(0.004)	(0.004)	(0.003)
Short-Run Effect	0.001	-0.002	-0.005	-0.002	0.003	0.002
$\hat{\beta}_{SR}$: Years 0 to 8	(0.004)	(0.006)	(0.005)	(0.003)	(0.003)	(0.002)
Medium-Run Effect	0.006	0.002	-0.005	-0.003	-0.001	-0.002
$\hat{\beta}_{MR}$: Years 9 to 17	(0.005)	(0.008)	(0.006)	(0.004)	(0.004)	(0.003)
Long-Run Effect	0.008	0.008	-0.002	-0.004	-0.001	-0.002
$\hat{\beta}_{LR}$: Years 18 to 25	(0.007)	(0.010)	(0.007)	(0.005)	(0.005)	(0.004)
Number of Events	1,103	1,072	1,072	1,103	1,072	1,072
Sample Observations	242,875	241,593	241,593	243,126	241,858	241,858
Total Worker Observations	27,140,075	27,046,126	27,046,126	27,140,337	27,046,426	27,046,426
Occupation-Year FE	✓	✓	✓	✓	✓	✓
Occupation-State FE	✓	✓	✓	✓	✓	✓
Policy Controls	✓	✓	✓	✓	✓	✓
Occupation-Division-Year FE		✓	✓		✓	✓
Occupation Group-State-Year FE			✓			✓

Notes: This table replicates the main event study specifications with the imputation rate for hourly wages and occupation as the dependent variables. The estimation sample is an aggregated state-occupation panel constructed from the 1980-2000 Census, and the 2001-2017 ACS and a balanced panel of three-digit occupation codes. Policy controls include additional distributed lags for six-digit registration and certification requirements, as well as occupation-specific controls for the number of minor (sub six-digit) regulations. The regressions are weighted by the number of workers in each state-occupation-year cell, which is constructed using sample-weighted annual occupational employment share estimates from the full monthly CPS scaled by total employment. All estimates are expressed relative to the leave-out category of 1 to 3 years prior to treatment. Standard errors are clustered by state. Significance levels are indicated by *** 1%; ** 5%; and * 10%.

Bibliography

- Abraham, Katharine G., and Melissa S. Kearney.** 2020. “Explaining the Decline in the U.S. Employment-to-Population Ratio: A Review of the Evidence.” *Journal of Economic Literature* 58 (3): 585–643.
- Abraham, Katharine G., and James R. Spletzer.** 2010. “Are the New Jobs Good Jobs?” In *Labor in the New Economy*, 101–143, Chicago, IL: University of Chicago Press.
- Acemoglu, Daron, and David Autor.** 2011. “Skills, Tasks and Technologies: Implications for Employment and Earnings.” In *Handbook of Labor Economics*, Volume 4. 1043–1171, Elsevier.
- Akerlof, George A.** 1970. “The Market for Lemons: Quality Uncertainty and the Market Mechanism.” *The Quarterly Journal of Economics* 84 (3): 488–500.
- Allard, Mary D.** 2016. “Adding Questions on Certifications and Licenses to the Current Population Survey.” *Monthly Labor Review*.
- Allegretto, Sylvia, Arindrajit Dube, Michael Reich, and Ben Zipperer.** 2017. “Credible Research Designs for Minimum Wage Studies: A Response to Neumark, Salas, and Wascher.” *International Labor Relations Review* 70 (3): 559–592.
- Allensworth, Rebecca H.** 2017. “Foxes at the Henhouse: Occupational Licensing Boards Up Close.” *California Law Review* 105 (6): 1567–1610.
- Anderson, Mark D., Ryan Brown, Kerwin Kofi Charles, and Daniel I. Rees.** 2020. “Occupational Licensing and Maternal Health: Evidence from Early Midwifery Laws.” *Journal of Political Economy* 128 (11): 4337–4383.
- Arrow, Kenneth J.** 1963. “Uncertainty and the Welfare Economics of Medical Care.” *American Economic Review* 53 (5): 941–973.

- Austin, Benjamin, Edward Glaeser, and Lawrence H. Summers.** 2018. "Saving the Heartland: Place-Based Policies in 21st Century America." *Brookings Papers on Economic Activity*.
- Autor, David H., John J. Donohue III, and Stewart J. Schwab.** 2006. "The Costs of Wrongful-Discharge Laws." *The Review of Economics and Statistics* 88 (2): 211–231.
- Autor, David H., and David Dorn.** 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review* 103 (5): 1553–97.
- Bailey, Thomas R., and Clive Belfield.** 2018. "The Impact of Occupational Licensing on Labor Market Outcomes of College-Educated Workers." *CCRS Working Paper No. 104*.
- Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis.** 2020. "COVID-19 is Also a Reallocation Shock." *National Bureau of Economic Research Working Paper No. 27137*.
- Bartel, Ann P.** 1989. "Where do the New U.S. Immigrants Live?" *Journal of Labor Economics* 7 (4): 371–391.
- Bartik, Timothy J.** 1991. "Who Benefits from State and Local Economic Development Policies?" *Upjohn Institute for Employment Research*.
- Bianco, David P.** 1996. *Professional and Occupational Licensing Directory: A Descriptive Guide to State and Federal Licensing, Registration, and Certification Requirements*. Detroit, MI: Gale Research Inc., 2nd edition.
- Blair, Peter Q., and Bobby W. Chung.** 2018. "Job Market Signaling Through Occupational Licensing." *National Bureau of Economic Research Working Paper No. 24791*.
- Blair, Peter Q., and Bobby W. Chung.** 2019. "How Much of Barrier to Entry is Occupational Licensing?" *British Journal of Industrial Relations* 57 (4): 919–943.

- Bohn, Sarah, and Todd Pugatch.** 2015. "U.S. Border Enforcement and Mexican Immigrant Location Choice." *Demography* 52 (5): 1543–1570.
- Borjas, George J.** 2001. "Does Immigration Grease the Wheels of the Labor Market?" *Brookings Papers on Economic Activity* 2001 (1): 69–133.
- Borjas, George J.** 2015. "The Slowdown in the Economic Assimilation of Immigrants: Aging and Cohort Effects Revisited Again." *Journal of Human Capital* 9 (4): 483–517.
- Borjas, George J., and Bernt Bratsberg.** 1996. "Who Leaves? The Outmigration of the Foreign-Born." *The Review of Economics and Statistics* 78 (1): 165–176.
- Borusyak, Kirill, and Xavier Jaravel.** 2018. "Revisiting Event Study Designs, With an Application to the Estimation of the Marginal Propensity to Consume." *Working Paper, Harvard University*.
- Brinegar, Pamela.** 2004. "Trends and Issues in State Professional Licensing." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Brinegar, Pamela.** 2005. "Professional Licensing." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Brinegar, Pamela.** 2006. "Professional and Occupational Regulation." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Brinegar, Pamela, and Kara Schmitt.** 1992. "State Occupational and Professional Licensure." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Brinegar, Pamela, and Kara Schmitt.** 1994. "Licensure/Regulation." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Cadena, Brian C., and Brian K. Kovak.** 2016. "Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession." *American Economic Journal: Applied Economics* 8 (1): 257–90.

- Card, David.** 1999. “The Causal Effect of Education on Earnings.” *Handbook of Labor Economics* 3 1801–1863.
- Card, David, and Ethan G. Lewis.** 2007. “The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts.” In *Mexican Immigration to the United States*, 193–228, University of Chicago Press.
- Carpenter, Dick, Lisa Knepper, Kyle Sweetland, and Jennifer McDonald.** 2017. “License to Work: A National Study of the Burdens from Occupational Licensing.” Arlington, VA: Institute for Justice.
- Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer.** 2019. “The Effect of Minimum Wages on Low-Wage Jobs.” *The Quarterly Journal of Economics* 134 (3): 1405–1454.
- Council of Economic Advisors.** 2015. *Occupational Licensing: A Framework for Policymakers*. Washington, DC: Council of Economic Advisors.
- Council of State Governments.** 1952. *Occupational Licensing in the States*. Chicago, IL: Council of State Governments.
- Council of State Governments.** 1968. *Occupations and Professions Licensed by the States, Puerto Rico and the Virgin Islands*. Chicago, IL: Council of State Governments.
- Council on Licensure, Enforcement and Regulation.** 1990. *Occupational and Professional Regulation in the States: A Comprehensive Compilation*. Lexington, KY: Council on Licensure, Enforcement and Regulation.
- Council on Licensure, Enforcement and Regulation.** 1994. *The Directory of Professional and Occupational Regulation in the United States and Canada*. Lexington, KY: Council on Licensure, Enforcement and Regulation.

- Cunningham, Evan.** 2019. “Professional Certifications and Occupational Licenses.” *Monthly Labor Review*.
- 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.
- Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor.** 2008. “Is the Melting Pot Still Hot? Explaining the Resurgence of Immigrant Segregation.” *The Review of Economics and Statistics* 90 (3): 478–497.
- De Chaisemartin, Clément, and Xavier d’Haultfoeuille.** 2020. “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects.” *American Economic Review* 110 (9): 2964–2996.
- Deming, David J., Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F. Katz.** 2016. “The Value of Postsecondary Credentials in the Labor Market: An Experimental Study.” *American Economic Review* 106 (3): 778–806.
- Dorn, David.** 2009. “Essays on Inequality, Spatial Interaction, and the Demand for Skills.” *Dissertation University of St. Gallen No. 3613*.
- Dube, Arindrajit, T. William Lester, and Michael Reich.** 2010. “Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties.” *The Review of Economics and Statistics* 92 (4): 945–964.
- Duncan, Otis D., and Beverly Duncan.** 1955. “A Methodological Analysis of Segregation Indexes.” *American Sociological Review* 20 (2): 210–217.
- Dustmann, Christian, and Joseph-Simon Görlach.** 2016. “The Economics of Temporary Migrations.” *Journal of Economic Literature* 54 (1): 98–136.

- Farronato, Chiara, Andrey Fradkin, Bradley Larsen, and Erik Brynjolfsson.** 2020. “Consumer Protection in an Online World: An Analysis of Occupational Licensing.” *National Bureau of Economic Research Working Paper No. 26601*.
- Fischer, Jeanne.** 1980. “State Regulation of Psychologists.” *Washington University Law Review* 58 (3): 639–676.
- Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren.** 2018. “Integrated Public Use Microdata Series, Current Population Survey: Verion 6.0.” [dataset]. Minneapolis: University of Minnesota. <https://doi.org/10.18128/D030.V6.0>.
- Friedman, Milton.** 1962. *Capitalism and Freedom*. Chicago, IL: University of Chicago Press.
- Friedman, Milton, and Simon Kuznets.** 1945. “Income from Independent Professional Practice.” *National Bureau of Economic Research*.
- Funkhouser, Edward.** 2000. “Changes in the Geographic Concentration and Location of Residence of Immigrants.” *International Migration Review* 34 (2): 489–510.
- Furth, Salim.** 2016. “Understanding the Data on Occupational Licensing.” *The Heritage Foundation Background No. 3159*.
- Gittleman, Maury, Mark A. Klee, and Morris M. Kleiner.** 2018. “Analyzing the Labor Market Outcomes of Occupational Licensing.” *Industrial Relations* 57 (1): 57–100.
- Gittleman, Maury, and Morris M. Kleiner.** 2016. “Wage Effects of Unionization and Occupational Licensing Coverage in the United States.” *ILR Review* 69 (1): 142–172.
- Goodman-Bacon, Andrew.** 2018. “Difference-in-Differences with Variation in Treatment Timing.” *National Bureau of Economic Research Working Paper No. 25018*.

- Graves, Brooke W.** 1938. "Professional and Occupational Restrictions." *Temple Law Quarterly* 13 (3): 334–365.
- Greene, Karen.** 1969. *Occupational Licensing and the Supply of Nonprofessional Manpower*. Washington, DC: U.S. Department of Labor and Educational Testing Service.
- Han, Suyoun, and Morris M. Kleiner.** 2017. "Analyzing the Influence of Occupational Licensing Duration and Grandfathering on Labor Market Outcomes." *Federal Reserve Bank of Minneapolis Staff Report No. 556*.
- Hemphill, Thomas A., and Dick M. Carpenter.** 2016. "Occupations: A Hierarchy of Regulatory Options." *Regulation* 39 (3): 20–24.
- Hirsch, Barry T.** 2004. "Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic." *Journal of Labor Research* 25 (2): 233–266.
- Hirsch, Barry T., and Edward J. Schumacher.** 2004. "Match Bias in Wage Gap Estimates Due to Earnings Imputation." *Journal of Labor Economics* 22 (3): 689–722.
- Ingram, Samuel J.** 2019. "Occupational Licensing and the Earnings Premium in the United States: Updated Evidence from the Current Population Survey." *British Journal of Industrial Relations* 57 (4): 732–763.
- Jaeger, David A.** 2008. "Green Cards and the Location Choices of Immigrants in the United States." *Research in Labor Economics* 27 (1): 31–183.
- Johnson, Janna E., and Morris M. Kleiner.** 2020. "Is Occupational Licensing a Barrier to Interstate Migration?" *American Economic Journal: Economic Policy* 12 (3): 347–73.
- Kilmer, Marc.** 2018. "A Look at Occupational Licensing Reform across the United States." *Arkansas Center for Research in Economics*.
- Kleiner, Morris M.** 2016. "Battling Over Jobs: Occupational Licensing in Health Care." *American Economic Review: Papers & Proceedings* 106 (5): 165–70.

- Kleiner, Morris M., and Alan B. Krueger.** 2010. "The Prevalence and Effects of Occupational Licensing." *British Journal of Industrial Relations* 48 (4): 676–687.
- Kleiner, Morris M., and Alan B. Krueger.** 2013. "Analyzing the Extent and Influence of Occupational Licensing on the Labor Market." *Journal of Labor Economics* 31 (2): 173–202.
- Kleiner, Morris M., and Evan Soltas.** 2020. "A Welfare Analysis of Occupational Licensing in the U.S. States." *NBER Working Paper No. 26383*.
- Kleiner, Morris M., and Ming Xu.** 2020. "Occupational Licensing and Labor Market Fluidity." *National Bureau of Economic Research Working Paper No. 27568*.
- Law, Marc T., and Sukkoo Kim.** 2005. "Specialization and Regulation: The Rise of Professionals and the Emergence of Occupational Licensing Regulation." *The Journal of Economic History* 65 (3): 723–756.
- Law, Marc T., and Mindy S. Marks.** 2009. "Effects of Occupational Licensing Laws on Minorities: Evidence from the Progressive Era." *The Journal of Law and Economics* 52 (2): 351–366.
- Law, Marc T., and Mindy S. Marks.** 2013. "From Certification to Licensure: Evidence from Registered and Practical Nurses in the United States, 1950-1970." *The European Journal of Comparative Economics* 10 (2): 177–198.
- Law, Marc T., and Mindy S. Marks.** 2017. "The Labor-Market Effects of Occupational Licensing Laws in Nursing." *Industrial Relations: A Journal of Economy and Society* 56 (4): 640–661.
- Leland, Hayne E.** 1979. "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards." *Journal of Political Economy* 87 (6): 1328–1346.

- Lin, Jeffrey.** 2011. “Technological Adaptation, Cities, and New Work.” *Review of Economics and Statistics* 93 (2): 554–574.
- Mellow, Wesley, and Hal Sider.** 1983. “Accuracy of Response in Labor Market Surveys: Evidence and Implications.” *Journal of Labor Economics* 1 (4): 331–344.
- Monheit, Alan C.** 1982. “Occupational Licensure and the Utilization of Nursing Labor: An Economic Analysis.” In *Advances in Health Economics and Health Services Research*, Volume 3. 117–142, Greenwich, CT: JAI Press.
- National Conference of State Legislatures.** 2017. *The State of Occupational Licensing: Research, State Policies, and Trends*. Washington, DC: National Conference of State Legislatures.
- Nunn, Ryan, and Gabriel Scheffler.** 2018. “Occupational Licensing and the Limits of Public Choice Theory.” *Administrative Law Review Accord* 4 (2): 25–41.
- Redbird, Beth.** 2017. “The New Closed Shop? The Economic and Structural Effects of Occupational Licensure.” *American Sociological Review* 82 (3): 600–624.
- Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, and Matthew Sobek.** 2019. “Integrated Public Use Microdata Series: Version 9.0.” [dataset]. Minneapolis: University of Minnesota. <https://doi.org/10.18128/D010.V9.0>.
- Sandler, Danielle H., and Ryan Sandler.** 2014. “Multiple Event Studies in Public Finance and Labor Economics: A Simulation Study with Applications.” *Journal of Economic and Social Measurement* 39 (1): 31–57.
- Schmidheiny, Kurt, and Sebastian Siegloch.** 2020. “On Event Studies and Distributed-Lags in Two-Way Fixed Effects Models: Identification, Equivalence, and Generalization.” *IZA Institute of Labor Economics Discussion Paper Series No. 12-079*.

- Shapiro, Carl.** 1986. "Investment, Moral Hazard, and Occupational Licensing." *The Review of Economic Studies* 53 (5): 843–862.
- Shimberg, Benjamin.** 1980. *Occupational Licensing: A Public Perspective*. Princeton, NJ: Educational Testing Service.
- Stigler, George J.** 1971. "The Theory of Economic Regulation." *The Bell Journal of Economics and Management Science* 2 (1): 3–21.
- Stokes-Berry, Frances.** 1986. "State Regulation of Occupations and Professions." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Stokes-Berry, Frances, and Pamela Brinegar.** 1988. "State Regulation of Occupations and Professions." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Stokes-Berry, Frances, and Doug Roederer.** 1982. "State Regulation of Occupations and Professions." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Thornton, Robert J., Edward Timmons, and Dante DeAntonio.** 2017. "Licensure or License?: Prospects for Occupational Deregulation." *Labor Law Journal* 68 (1): 46–578.
- Thornton, Robert J., and Edward J. Timmons.** 2013. "Licensing One of the World's Oldest Professions: Massage." *The Journal of Law and Economics* 56 (2): 371–388.
- Thornton, Robert J., and Edward J. Timmons.** 2015. "The De-Licensing of Occupations in the United States." *Monthly Labor Review*.
- Timmons, Edward J., and Anna Mills.** 2018. "Bringing the Effects of Occupational Licensing Into Focus: Optician Licensing in the United States." *Eastern Economic Journal* 44 (1): 69–83.
- Tracy, LeeAnn.** 1998. "Management, Regulation and Personnel." In *The Book of the States*, Lexington, KY: Council of State Governments.

- U.S. Department of Commerce.** 1942. *State Occupational Legislation*. Washington, DC: U.S. Department of Commerce.
- U.S. Department of Health, Education, and Welfare.** 1967. *State Licensing of Health Occupations*. Washington, DC: U.S. Department of Health, Education, and Welfare.
- U.S. Department of Health, Education, and Welfare.** 1977. *State Regulation of Health Manpower*. Washington, DC: U.S. Department of Health, Education, and Welfare.
- Wall, Audrey.** 2000. "Licensure and Regulation." In *The Book of the States*, Lexington, KY: Council of State Governments.
- Watson, Tara.** 2013. "Enforcement and Immigrant Location Choice." *NBER Working Paper No. 19626*.
- White, William D.** 1987. "The Introduction of Professional Regulation and Labor Market Conditions: Occupational Licensure of Registered Nurses." *Policy Sciences* 20 (1): 27–51.
- Wozniak, Abigail.** 2010. "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources* 45 (4): 944–970.
- Xia, Xing.** 2020. "Barrier to Entry or Signal of Quality? The Effects of Occupational Licensing on Minority Dental Assistants." *Working Paper*.
- Zavodny, Madeline.** 1999. "Determinants of Recent Immigrants' Locational Choices." *International Migration Review* 33 (4): 1014–1030.