NBER WORKING PAPER SERIES

THE ORIGINS AND EVOLUTION OF OCCUPATIONAL LICENSING IN THE UNITED STATES

Nicholas A. Carollo
Jason F. Hicks
Andrew Karch
Morris M. Kleiner

Working Paper 33580
http://www.nber.org/papers/w33580

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2025

The views expressed in this paper are those of the authors and do not represent the views or policies of the Federal Reserve Board of Governors, the Federal Reserve Bank of Minneapolis, or any other members of the Federal Reserve System or its staff. We thank Peter Blair, Darwyyn Deyo, Casey Mulligan, Elena Pastorino and seminar participants at the ASSA-AEA Labor and Employment Relations Association annual meeting, the American Political Science Association, Federal Reserve Bank of Minneapolis, Midwest Political Science Association, NBER Summer Institute Development of the American Economy, Queen Mary University of London, Society of Labor Economics, Stanford University, University of Minnesota, West Virginia University, and the Western Economic Association International for helpful feedback and suggestions. Brandon Dunk provided outstanding research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Origins and Evolution of Occupational Licensing in the United States
Nicholas A. Carollo, Jason F. Hicks, Andrew Karch, and Morris M. Kleiner
NBER Working Paper No. 33580
March 2025
JEL No. J01, J29, J4, J44, J48

ABSTRACT

The analysis of occupational licensing has concentrated largely on its labor market and consumer welfare effects. By contrast, relatively little is known about how occupational licensing laws originated or the key factors in their evolution. In this paper, we study the determinants of U.S. licensing requirements from 1870 to 2020. We begin by developing a model where licensing arises as an endogenous political outcome and use this framework to study how market characteristics and political incentives influence regulators' choices. Our empirical analysis draws on a novel database tracking the initial enactment of licensing legislation for hundreds of unique occupations, as well as changes to the specific qualifications required to obtain a subset of licenses over time. We first show that, consistent with the predictions of our model, licensing requirements are more common and were adopted earlier for occupations whose tasks plausibly pose some risk to consumers. Second, large, urbanized states are significantly more likely to produce new policies. Third, among occupations regulated before 1940, licensing requirements appeared earlier in states with more practitioners and where incumbent workers likely experienced greater labor market competition. After 1980, state-level factors are more strongly associated with the timing of policy adoption. Finally, political organization, as measured by the establishment of a state professional association, significantly increases the probability of regulation. Together, our findings suggest that both public and private interests have contributed to the diffusion of licensing requirements across states and occupations.

Nicholas A. Carollo
Federal Reserve Board of Governors
Constitution Ave NW &, 20th St NW
Washington, DC 20551
carollo.nicholas@gmail.com

Jason F. Hicks
University of Victoria
3800 Finnerty Road
Victoria BC V8P 5C2
Canada
jasonhicks@uvic.ca

Andrew Karch
University of Minnesota
Social Sciences Building
Minneapolis, MN 55455
ajkarch@umn.edu

Morris M. Kleiner
University of Minnesota
Humphrey School of Public Affairs
260 Humphrey Center
301 19th Street South
Minneapolis, MN 55455
and NBER
kleiner@umn.edu

1 Introduction

Occupational licensing has become one of the most prevalent forms of labor market regulation in the United States.[1] Recent estimates from the Bureau of Labor Statistics, for instance, show that about 20% of workers currently hold an active license, twice the share who belong to a labor union (Cunningham, 2019). While a growing body of research has shown that licensing requirements have economically significant effects on workers and consumers, relatively little is known about the development of the institution itself.[2] That is, how did licensing laws originate, what were the key factors in their evolution, and why has this form of regulation spread to such a broad swath of the labor market?

In this paper, we address these questions by analyzing the economic and political determinants of occupational licensing policy. Leveraging novel regulatory data spanning 1870 to 2020, we first document a set of stylized facts regarding the historical composition and timing of licensing laws. Building on these insights, we offer new evidence on the places where licensing requirements were initially enacted, the ways they diffused across states, and the forces driving the evolution of specific qualifications for licensure. In doing so, our analysis is the first to systematically study changes to both the extensive and intensive margins of regulation over time, using information on thousands of individual policies. We are also the first to relate both occupational task content and the organization of professional associations to the implementation of licensing statutes and regulatory procedures, which provides new insights into the political economy of U.S. labor regulation.

We structure our analysis around a model that integrates a competitive market for professional services with the problem of a representative politician who may choose to enact licensing requirements in response to influence from professional associations. We follow Kleiner and Soltas (2023) in that licensing raises entry costs, reducing labor supply, but may also benefit consumers, increasing labor demand. When licensing costs vary across workers (owing, for instance, to heterogeneous ability), professional association members may prefer a higher level of regulation than is socially-optimal, even if licensing is beneficial on net. Our framework therefore embeds the main private and social motives for licensing emphasized by classic theories of economic regulation (Shleifer, 2005). Further, our theoretical framework highlights what we view as the key determinants of policy adoption and diffusion in our setting, including occupational task content, market size, and the political organization of workers who stand to benefit from licensure.

Guided by the predictions of our model, our empirical analysis draws on two unique regulatory

[1] An occupational license is a credential awarded by a government agency that constitutes legal authority to do a specific job. Unlike business licenses, which attach to an establishment or firm, occupational licenses attach to individual workers and usually require a formal demonstration of competency, such as passing a qualifying examination. In the United States, the vast majority of licensing requirements are implemented at the state level.

[2] Recent studies analyzing the effect of occupational licensing on the labor market and worker mobility include Kleiner and Krueger (2010, 2013); Gittleman et al. (2018); Blair and Chung (2019); Johnson and Kleiner (2020); Kleiner and Xu (Forthcoming); Han and Kleiner (2021); Deyo and Plemmons (2022) and Blair and Chung (Forthcoming). Others, including Kleiner and Kudrle (2000); Kleiner et al. (2016); Anderson et al. (2020); Larsen et al. (2020); Blair and Fisher (2022) and Farronato et al. (2024) study the effect of licensing on the product market, service quality, and prices. Kleiner and Soltas (2023) propose a sufficient statistic approach to estimate the welfare effects of occupational licensing in a general equilibrium model.

datasets. First, we use data from Carollo (2024), which records the enactment of state and federal licensing laws from 1870 to 2020. In this dataset, we observe the timing of policy changes for over 250 detailed occupations, which covers the vast majority of licensed jobs in the United States. Second, we leverage new data from the Occupational Licensing Law Research Project (OLLRP), which records the evolution of specific qualifications workers must demonstrate to attain a license. Here, we observe a panel of training standards, continuing education, and other requirements for a subset of licensed occupations.[3] Consistent with previous estimates, our data show that the share of the workforce covered by licensing requirements rose from about 5% in 1950 to at least 20% today. Most of this growth reflects the enactment of new laws.[4]

To test our predictions on the determinants of occupational licensing policy, we link our regulatory data to a broad range of state- and occupation-level characteristics. In addition to drawing on familiar datasets such as the U.S. Census, we use information from the Occupational Information Network (O*NET) and the Dictionary of Occupational Titles, together with natural language processing methods, to construct a set of indices that plausibly reflect public health and safety concerns based on the tasks workers perform. We also assemble original data on the establishment of state and national professional associations as measures of an occupation's political organization. Further, we compile and harmonize dozens of additional variables describing the evolution of state demographic, economic, and political characteristics that we expect to be related to the diffusion of occupational licensing over the past 150 years.

Consistent with the predictions of our model, we find, first, that occupations are more likely to be regulated when public health, safety, or welfare concerns related to their tasks are plausible. Specifically, occupations with greater task complexity, more interpersonal interaction, and greater professional autonomy are all significantly more likely to be licensed than occupations that rank lower on these attributes. Further, we find that the higher an occupation scores on these measures, the earlier it was regulated, on average, relative to other occupations that eventually became licensed in the same state. In other words, we find that licensing tends to diffuse from more to less risky and complex jobs over time. While these results underscore the importance of occupational task content for understanding regulatory choices, we also note that many occupations have become widely-licensed despite little apparent risk to consumers.

Next, we explore patterns of policy innovation and diffusion at the state-level. We find that certain states are much more likely to create new license categories or quickly adopt these policies after other jurisdictions. Roughly a quarter of the occupations we study were initially regulated in just three states: California, New York, and Texas. Moreover, California was one of the first

[3] The OLLRP data collection project is ongoing. In this draft, we present case-study results for a handful of occupations with preliminary data, though we plan to update our analysis as additional data is finalized.

[4] We obtain these numbers by aggregating policy indicators for detailed state-by-occupation cells weighted by estimates of their national employment shares. While our approach provides methodologically-consistent estimates of licensing coverage over time, we emphasize that these figures cannot be interpreted as the share of workers who actually held a license. Survey data on licensing at the individual level have become available only in the past decade, which means that it is not possible to estimate historical aggregate credential attainment rates. Although our estimate of policy coverage in 2020 is similar to the share of workers who reported having attained a license in the Current Population Survey, these numbers are based on different concepts and are not directly comparable.

10 states to regulate nearly half of all licensed occupations. Reflecting persistent demographic, ideological, and institutional differences, early policy adopters tend to be larger, more urbanized, and have accumulated a larger volume of regulation across all legal domains. Once enacted, we find that licensing laws tend to diffuse geographically, consistent with long-standing research on state policymaking (Walker, 1969). In contrast to other policy diffusion studies, however, we find little evidence that occupational licensing laws have become more politically polarized in recent decades (DellaVigna and Kim, 2022; Mallinson, 2021).

Third, we leverage the panel structure of our data to study the relative timing of policy adoption within occupations. Among those initially regulated before 1940, licensing laws appeared earlier in larger markets (as measured by the number of practitioners in a state), a finding that has been documented for other types of regulation as well (Mulligan and Shleifer, 2005; Law and McLaughlin, 2022). Like these papers, we see this as consistent with the view that the fixed cost of establishing new legal institutions limits the supply of regulation in small markets. However, we also find evidence that, for a given occupation, greater exposure to labor market competition from recent immigrants further increases the probability of regulation, consistent with theories (including our own) that emphasize the demand for regulation among incumbent workers. For policies enacted since 1980, we find that common state-level factors are more strongly associated with the timing of initial regulation than variation in occupation-specific employment and earnings.

Fourth, we assess the role of professional associations in the diffusion of occupational licensing laws. Historically, organizations such as the American Medical Association and its local affiliates supported licensing requirements and actively lobbied state legislatures to enact statutes for their professions. We show empirically that these efforts were often successful. Using data on state-level associations representing nine major occupations in an event-study design, we find that the probability of an occupation-state cell becoming regulated increases by 20 percentage points within five years of an association's founding. This provides clear and plausibly causal evidence that local political organization led to favorable regulatory outcomes for these groups.

The final set of results in this paper shifts our focus from the origin of licensing policies to the evolution of qualifications for licensure. As our data collection process is ongoing, this draft presents a case study of training hours for barbers and cosmetologists. Around the turn of the 20th century, barbers, who were regulated before cosmetologists, initially required as much as three years of training to obtain a license. By the 1930s, however, most states had cut these requirements dramatically. When states began licensing cosmetologists in the 1920s, training hours were initially lower than those for barbers, but the two had largely converged by the 1950s. For these occupations, we find evidence that subsequent changes to licensing qualifications within states reflected movement toward national standards rather than a response to state-specific trends.

Taken together, our findings provide support for elements of both public-interest (Pigou, 1938) and special-interest (Stigler, 1971; Peltzman, 1976) theories of economic regulation. Indeed, we emphasize that for many of the occupations we study, both social welfare and rent-seeking motivations are empirically plausible, and need not be mutually exclusive. For example, our theoretical framework shows how political and economic incentives can lead to overregulation, even in cases where licensing is welfare improving on net. These private incentives could also be one reason why licensing has become the dominant form of occupational regulation when less restrictive policies such as state certification (voluntary licensure) might be sufficient to address potential market failures. Similarly, how licensing is implemented matters, and researchers have long noted that licensing boards may adopt excessive or arbitrary standards to attain a license (Maurizi, 1974). Thus, even if some level of regulation is in the public interest, less may be better on the margin.

Related literature. Our analysis contributes to several branches of literature on occupational licensing, the political economy of regulation, and labor economics. First, our findings complement other work on the historical development of labor market policies and institutions, including collective bargaining (Schmick, 2018; Medici, 2024), fair employment laws (Collins, 2003), minimum wages (Sobel, 1999; Neumark and Wascher, 2008), and workers' compensation (Fishback and Kantor, 1998). We offer a range of new evidence on the factors associated with the timing and diffusion of licensing requirements, and our results speak to long-standing debates on the role of private incentives and social welfare concerns in explaining the proliferation of state regulation.

We contribute most directly to the literature on occupational licensing. While most research has concentrated on understanding the effects of licensing policies on consumers and the labor market, comparatively little has been written about the political economy of these laws. Notable exceptions include Smith (1982); Graddy (1991a,b); Law and Kim (2005) and Pagliero (2011). Relative to these papers, ours leverages vast new data, extending many of their key themes beyond the analysis of specific time periods or occupations. Further, while existing work has focused exclusively on the origins of licensing legislation, we are the first to study the evolution of qualifications for licensure over time. We also contribute theoretically by providing a unified framework that ties the distributional effects of occupational licensing in the labor market to the incentives of professional associations and politicians to support the enactment of licensing legislation.

Our paper is also related to research on long-run trends in the occupational structure of the U.S. labor market. One branch of this literature examines the relationship between technological change, skills, and occupational employment (Goldin and Katz, 2008; Autor et al., 2006; Autor and Dorn, 2013; Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018; Michaels et al., 2019). Others consider the emergence and diffusion of "new work" (Lin, 2011; Autor et al., 2024; Kim et al., 2024) or the evolution of tasks within occupations (Autor et al., 2003; Atalay et al., 2020; Nedelkoska et al., 2023). Relative to these studies, ours highlights the economic significance of regulatory responses to major structural changes, including the growth of service-sector employment, increasing task complexity, and occupational professionalization. Secondarily, we propose a tractable method for measuring changes in task content within occupations over time, which researchers could readily extend to other occupational attributes or source data.

Beyond our focus on labor market regulation, our paper draws from rich bodies of literature in political science and economics on policy innovation and diffusion across U.S. states (Karch, 2007; Mooney, 2020; Mallinson, 2020, 2021; DellaVigna and Kim, 2022). In keeping with long-standing themes in this literature, our findings underscore the importance of geographic proximity and professional associations, and also highlight the interaction between state-level characteristics and the attributes of specific policies – in our context, occupations (Nicholson-Crotty, 2009; Makse and Volden, 2011). Further, while most studies have focused on policies enacted relatively recently, ours spans several distinct political eras, which date back to the emergence of the modern regulatory state.

Road map. Section 2 introduces a theoretical framework for analyzing the political economy of occupational licensing. Section 3 describes our primary datasets. Section 4 documents trends in the coverage, timing, and composition of U.S. occupational licensing statutes. Section 5 presents our main results on the origins of licensing statutes, and Section 6 our main results on the evolution of training standards and qualifications. In Section 7, we conclude by discussing the implications of our findings for policymakers and researchers.

2 Theoretical framework

Our study of the origins and evolution of occupational licensing begins by considering the incentives of state policymakers and professional associations. In the spirit of classic theories of economic regulation such as Stigler (1971) and Peltzman (1976), we develop a model in which licensing requirements are chosen by a representative politician in response to influence from special interest groups. This theoretical framework motivates our empirical analysis by highlighting what we view as the key determinants of policy adoption and diffusion in our setting.

2.1 Licensing as a political outcome

We model occupational licensing as a mandatory investment in occupation-specific training and a fee that is required to enter the market.[5] Licensing requirements are chosen by the government, and may affect wages, occupation choice, and consumer welfare in a way that we specify below. Our setup closely follows that of Grossman and Helpman (1994), who assume that the government consists of an incumbent politician with preferences over both direct contributions from interest groups and the general welfare of unorganized voters.

Government's problem. Consider an economy populated by $i = 1, ..., N$ individuals working in many occupations $j = 1, ..., J$. Workers are partitioned into groups g, of which a subset G' have organized themselves as professional associations (which may or may not represent all workers in an occupation). A representative politician chooses a training requirement z_j and a license fee τ_j for each occupation. Let the vector $\mathbf{p} \in P$ denote the set of all feasible policies, which we assume is bounded. Professional associations, though their lobbying arms, announce a contribution schedule

[5]In our setting, this means that licensing increases the cost of entry but, conditional on meeting the training requirements, does not directly limit the number of entrants (unlike certain business licenses that are in fixed supply). While we think that this is a realistic assumption for our exposition, we note that there is some historical evidence that licensing boards may have achieved de facto quotas by controlling examination pass rates (Maurizi, 1974).

$b_g(\mathbf{p}) \geq 0$, which maps each feasible policy choice to a donation that is made to the politician conditional on its enactment. These contributions do not necessarily reflect explicit vote-buying or persuasion; rather they reflect a mutual understanding that associations will tend to support legislators whose actions align with their own private interests.[6]

The politician then chooses a policy vector to maximize a weighted sum of political contributions and social welfare subject to administration and enforcement costs,

$$\max_{\mathbf{p}\in P} \quad \Omega(\mathbf{p}) = \sum_{g\in G'} b_g(\mathbf{p}) + \psi\mathcal{W}(\mathbf{p}) \quad \text{s.t.} \quad \kappa_j(z_j) \leq R_j(\tau_j) \;\; \forall \;\; j \in J \tag{1}$$

where $\psi \geq 0$ is the weight placed on social welfare. The presence of social welfare in the politician's objective function can be understood to reflect either ideological concerns (if there is some regard for the public interest) or electoral concerns (if enacting policies that are socially costly eventually reduces political support).[7] As $\psi \to \infty$, the politician acts as a social planner, and as $\psi \to 0$, they set policy only to collect the maximum possible contributions from lobbyists.

Our model departs from Grossman and Helpman (1994) in two respects. First, regulation is not necessarily distortionary but may raise or lower aggregate welfare, depending on the underlying characteristics of the market. Second, regulation is costly to implement and enforce, which is captured by the cost function $\kappa_j(z_j) > 0$. As in Mulligan and Shleifer (2005), we assume that setting up a new regulatory program entails a fixed cost, so that $\lim_{z_j\to 0} \kappa(z_j) = \kappa_0$. For simplicity, we require each policy to be self-financing and assume that excess revenue $R_j(\tau_j)$ cannot be redistributed. This assumption keeps the focus of our model on occupational licensing rather than expanding the set of policy instruments to include general taxes and transfers, as in Becker (1983).[8]

Association's problem. Before policy is made, associations announce their contribution schedules $b_g(\mathbf{p})$, taking as given the actions of other groups. After observing these schedules, the politician chooses a policy vector, and lobbyists make their contributions (we assume that contribution schedules are binding). Associations choose these functions to maximize the total welfare of their group net of lobbying costs, or

$$\Omega_g(\mathbf{p}) = \mathcal{W}_g(\mathbf{p}) - b_g(\mathbf{p}), \tag{2}$$

where $\mathcal{W}_g(\mathbf{p})$ denotes group-specific welfare. An equilibrium of this lobbying subgame is a feasible policy vector $\mathbf{p}^*$ and a set of contribution schedules $b^*_g(\mathbf{p})$ that maximize the politician's objective

[6]Since contributions are made only after policies have been enacted, we interpret lobbying in our setting as a mechanism for associations to reveal their policy preferences and reward friendly policymakers ex-post, rather than expenditures directed at persuading undecided or hostile ones.

[7]We choose to abstract from elections in our model for two reasons. First, occupational licensing is rarely, if ever, a prominent campaign issue, so we view lobbying efforts as more relevant to our setting than electoral competition is. Second, electoral models such as Grossman and Helpman (1996) yield similar predictions but introduce additional complexity that we abstract from here.

[8]The Census Annual Survey of State Government Tax Collections shows that since 1951, only Delaware has consistently received more than 5% of total tax revenue from business and occupational licenses. The average across all states is 1.5%. Further, if revenue generation alone were the primary motive for licensing occupations, it is not clear why states would choose to bundle fees with human capital requirements rather than simply taxing proprietors.

function and that of each professional association.

Assuming that contribution schedules are differentiable, Grossman and Helpman (1994) prove two key results. First, $b_g^*(\mathbf{p})$ reveals the association's true policy preferences, as the slope of the contribution schedule equals the association's marginal benefit of a small change in licensing policy. Second, given equilibrium contribution schedules, the politician's problem is equivalent to the maximization of a weighted social welfare function where unorganized groups receive a weight ψ and organized groups reveive a weight of $1+\psi$. Thus, the extent to which the politically-optimal level of regulation differs from the socially-optimal level reflects only (i) which groups are organized as professional associations, (ii) the weight politicians place on social welfare, and (iii) how much the preferences of association members diverge from those of the public.

2.2 Consumer's problem and the labor market

To say more about the implications of the lobbying game in the context of occupational licensing, we must specify how these laws affect the welfare of different groups. Following Kleiner and Soltas (2023), we consider an economy where individuals maximize a utility function of the form:

$$\max_{\{c_{ij}\},y_i,J_i} \quad \ln U_i = \ln\left[\left(\sum_{j=1}^{J}\gamma_j(z_j)c_{ij}^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}\right] - \rho_k(y_i + z_{J_i}) + \mu_{iJ_i} \tag{3}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} w_j c_{ij} \leq (1-\tau_{J_i})A_{J_i}(y_i)w_{J_i}.$$

Here, c_{ij} denotes individual i's consumption of occupation j's services, y_i is years of schooling, z_{J_i} is the training required to obtain a license, and J_i is their occupation. Workers have idiosyncratic preferences over occupations μ_{ij} that are drawn from a multivariate Fréchet distribution with dispersion parameter σ. Both education and licensing entail a cost ρ_k to obtain, which varies across worker types $k = 1, ..., K$. We interpret these differences in training costs as reflecting workers' heterogeneous ability.

Each worker inelastically supplies one unit of labor; $A_j(y_i)$ is an occupation-specific function mapping educational attainment to labor productivity; and w_j is the common wage per efficiency unit. The training required to obtain a license may or may not be valued by consumers, which is captured in the consumption weights of the CES composite good $\gamma_j(z_j)$. These potential benefits can vary by occupation, but we assume that training is subject to diminishing returns. This implies that there always exists a point at which the marginal cost of additional training exceeds consumers' marginal willingness to pay for quality. In addition, licensing fees τ_j operate as a tax on labor earnings that is paid by workers in the occupation.

In Appendix C, we solve for equilibrium in the labor market and show that aggregate demand for occupation j's services is given by

$$C_j = \frac{(w_j/\gamma_j(z_j))^{-\epsilon}}{P^{1-\epsilon}} \sum_{j=1}^{J} \sum_{k=1}^{K} n_k s_{jk} A_j(y_{jk})(1-\tau_j) w_j, \tag{4}$$

where P is the CES price index, n_k is the measure of type-k workers, and s_{jk} is the fraction of type-k workers who choose occupation j. Aggregate labor supplied to the occupation is

$$H_j = \sum_{k=1}^{K} n_k s_{jk} A_j(y_{jk}), \tag{5}$$

where w_j and s_{jk} in these two equations are determined in equilibrium and depend on the government's policy choices $\mathbf{p} = \{(z_j, \tau_j)\}_{j=1,..,J}$.

2.3 Equilibrium and welfare

Starting from an initial policy vector, workers solve equation (3) to maximize their utility. After their choices have been made, a subset of workers in an occupation organize themselves as a professional association and engage in the lobbying game described above.[9] If a policy change is enacted, workers may reoptimize, and the labor market adjusts to the new equilibrium.

Definition 1. *Given occupation characteristics $\{\gamma_j(\cdot), A_j(\cdot)\}$, regulatory costs $\{\kappa_j(\cdot)\}$, and parameters $\{\psi, \epsilon, \sigma, \rho_k\}$, an equilibrium is defined by an endogenous policy choice $\boldsymbol{p}^*$, political contribution schedules $\{b_g(\cdot)\}_{g\in G'}$, prices $\{w_j\}$, and quantities $\{\{c_{ij}\}, J_i\}$ such that*

i. Workers optimize: Consumption $\{c_{ij}\}$ and occupation choice J_i solve the worker's problem;

ii. Markets clear: For each occupation, the wage w_j is such that $C_j = H_j$;

iii. The politician and professional associations optimize: $\{\{b_g(\cdot)\}, \boldsymbol{p}^*\}$ is a subgame-perfect Nash equilibrium of the lobbying game.

Let z_j^* denote the *politically-optimal* training standard for occupation j. A central question for our analysis is how z_j^* might differ from the socially-optimal level of regulation, denoted z_j^s.

Proposition 1. *The social welfare effect of licensing a small occupation j is*

$$\frac{\partial \ln \mathcal{W}}{\partial z_j} = s_j^E \left[\frac{\partial \ln w_j}{\partial z_j} - \frac{\tau_j}{1-\tau_j} \frac{\partial \ln \tau_j}{\partial z_j} - \sum_{k=1}^{K} s_{jk}^E \rho_k \right], \tag{6}$$

where s_j^E is the employment share of occupation j.

To build intuition, consider the case where licensing fees are small. Then, the effect of a change in regulation on the equilibrium wage is

[9] We do not take a stand on why associations choose to organize, though our analysis does speak to the conditions under which rent-seeking may be more profitable. For a model of endogenous lobby formation, see Mitra (1999).

$$\frac{\partial \ln w_j}{\partial z_j} = \frac{1}{\epsilon + \sigma}\left(\epsilon \frac{\partial \ln \gamma_j(z_j)}{\partial z_j} + \sigma \sum_{k=1}^{K} s_{jk}^{L} p_k\right). \tag{7}$$

This expression is strictly positive, reflecting both consumers' marginal willingness to pay for higher licensing standards on the demand side and the labor-share weighted average of workers' training costs on the supply side. Substituting (7) into (6), we have

$$\frac{\partial \ln \mathcal{W}}{\partial z_j} \approx \frac{\epsilon s_j^E}{\sigma + \epsilon}\left(\frac{\partial \ln \gamma_j(z_j)}{\partial z_j} - \sum_{k=1}^{K} s_{jk}^{E} \rho_k\right), \tag{8}$$

which shows that from the social planner's perspective, the optimal level of regulation equates consumers' marginal WTP for licensing with the marginal social cost of training workers.[10] Further, $\gamma_j'(z_j) > 0$ is a necessary, but not sufficient, condition for $z_j^s > 0$. It cannot be socially optimal to impose regulations that have no benefit to consumers.

Proposition 2: *The effect of licensing on the welfare of a professional association g is*

$$\frac{\partial \ln \mathcal{W}_g}{\partial z_j} = \frac{\partial \ln w_j}{\partial z_j} - \frac{\tau_j}{1 - \tau_j}\frac{\partial \ln \tau_j}{\partial z_j} - \rho_g. \tag{9}$$

This proposition states that the optimal level of regulation from the association's perspective equates the marginal effect of licensing on the post-tax wage with its members' marginal cost of training. Comparing equations (6) and (9) shows how lobbying can distort regulatory outcomes in our setting. When marginal private costs differ from marginal social costs, there are opportunities for rent-seeking, and the association's preferred level of regulation differs from the social optimum. For any $\psi < \infty$ (if politicians care about campaign contributions), there is scope for the association to nudge policy decisions in the direction of its own preferred outcome by signaling these preferences to policymakers and rewarding those who enact favorable legislation.

2.4 Discussion and empirical implications

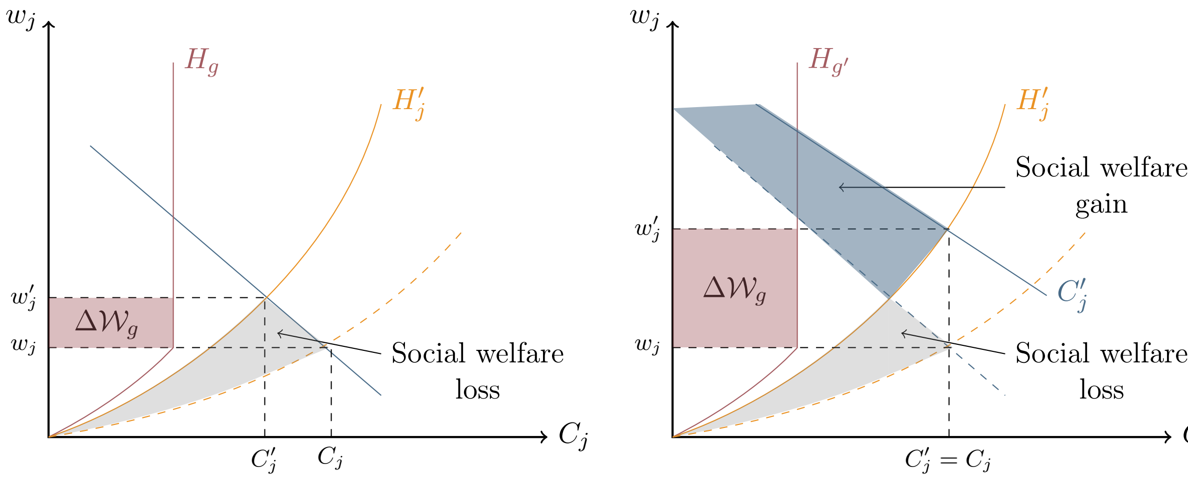
We now summarize the key implications of our model for the political economy of occupational licensing and discuss how these motivate our empirical analysis.

Proposition 3. *If $\kappa_0 > 0$, regulation is limited by the size of the market. Conversely, a decrease in $\kappa_j(z_j)$ weakly increases the politically-optimal level of regulation z_j^*.*

We predict that, as in the model of Mulligan and Shleifer (2005), smaller markets will be less regulated if enacting new legislation entails a fixed cost. In our model, this is because higher per-capita licensing fees are required to finance the regulatory program when the number of workers

[10]Equation (8) holds with equality if $\tau_j = 0$ and labor shares are equal to employment shares, $s_{jk}^L = s_{jk}^E$. The second condition is required because a change in the wage reflects the labor-share weighted average of training costs ρ_k, while welfare costs are weighted by employment, not total labor supply. This discrepancy will be larger in occupations where education has a larger effect on labor productivity.

Figure 1: Labor Market Effects of Occupational Licensing



(a) Supply shift only, $\gamma'_j(z_j) = 0$

(b) Supply and demand shift, $\gamma'_j(z_j) > 0$

Note: This figure shows two stylized examples of the labor market effects of occupational licensing in our model. For simplicity, we assume $\tau_j = 0$. Here, H_g represents the professional association's labor supply and $\Delta\mathcal{W}_g$ its change in welfare (before licensing costs). Total labor supply to the occupation, which includes workers who are not members of the association, is represented by H_j.

in the occupation is small, offsetting any public or private benefits from regulation. As fixed and variable costs fall, market size matters less, and the supply of regulation rises.

Although we abstract from dynamics and strategic interaction between states, we note that to the extent that regulatory costs decrease if similar policies have already been enacted by other jurisdictions, we would also expect regulation to diffuse from larger to smaller markets over time. Once a policy has been enacted in one jurisdiction, policymakers in others can use this legislation as a model, and may also observe additional information, such as ex-post outcomes and how key constituencies responded to the law. For this reason, we expect regulatory costs to be a decreasing function of prior enactments in other states. We also note that, although we have restricted our attention to monetary costs in our model, the same logic applies to the opportunity cost of legislators' time and political capital. Lobbying by professional associations may also relax these constraints on policymakers, as emphasized by Hall and Deardorff (2006).

Proposition 4: *For all $\psi \geq 0$, the politically-optimal level of regulation z_j^* is higher if $\gamma'_j(z_j) > 0$ (licensing raises consumer willingness to pay) than if $\gamma'_j(z_j) = 0$.*

We expect occupations with a plausible public interest case for licensing to be more extensively regulated. This is because, if all other parameters are held fixed, both social and private welfare are increasing functions of consumer WTP. While the professional association benefits from policies that restrict competition, they *also* benefit from policies that increase demand, as Figure 1 highlights. Importantly, this is true even if the politician places zero weight on social welfare, since public and private interests are better aligned when $\gamma'_j(z_j) > 0$. As shown in the following proposition, however, quality effects are not a sufficient condition for political outcomes to be socially optimal, and we expect that, in practice, many occupations will become overregulated.

Proposition 5: *For all $\psi < \infty$, the difference between the politically-optimal level of regulation z_j^* and the socially-optimal level of regulation z_j^s depends on the marginal private cost of licensing for the association relative to licensing's marginal social cost. We have,*

i. $z_j^* > z_j^s$ *if* $\rho_g < \sum_k s_{jk}^E \rho_k$,

ii. $z_j^* = z_j^s$ *if* $\rho_g = \sum_k s_{jk}^E \rho_k$,

iii. $z_j^* < z_j^s$ *if* $\rho_g > \sum_k s_{jk}^E \rho_k$.

This proposition states that when licensing is less costly for association members than it is for the average worker, opportunities for rent-seeking exist and the market will be overregulated in equilibrium, even if licensing is welfare-improving on net. Conversely, if marginal private costs are higher than marginal social costs, regulation will fall short of the social optimum. We expect the first case to be more common empirically, both because we expect professional association members to be positively selected, and because most licensing statutes offer at least some grandfathering exemptions, lowering compliance costs for incumbents relative to entrants. Such exemptions may also provide incentives for licensing boards to periodically raise their qualifications over time if these additional training costs are not internalized.

Proposition 5 also highlights a special case where political outcomes in our model *are* socially optimal: when marginal private costs are equal to marginal social costs. In this case, there are no opportunities for rent-seeking, so private incentives and the public interest are perfectly aligned. This could either occur when there is no heterogeneity in workers' ability meet licensing requirements or, equivalently, when all workers in an occupation have equal political representation.[11] Holding potential consumer benefits fixed, our model therefore predicts more extensive regulation for occupations with greater dispersion in training costs or more selective association membership.

3 Data

Our empirical analysis leverages two new datasets that track the history of occupational licensing requirements in the United States, obtained from Carollo (2024) and the Minnesota Occupational Licensing Law Research Project. To test the implications of our model for the political and economic determinants of licensing policy, we link this unique regulatory data to measures of occupational task content, the establishment of professional associations, and a broad range of other state- and occupation-level characteristics. Here, we highlight only the most important features of our data. Additional details on the construction of specific variables are provided in Appendix B.

[11] In the context of trade policy, Grossman and Helpman (1994) also show that when all voters are represented by special interests, political competition results in socially-optimal policy (i.e., free trade).

3.1 Occupational licensing and regulation

At the core of our project are two new datasets that we use to construct detailed event history panels tracking changes to the extensive and intensive margin of regulation over time. Because of limited data in the earlier part of our sample on key covariates for Alaska, Hawaii, and the District of Columbia, we restrict our attention to the contiguous 48 states throughout our analysis.

Timing of state policy changes. Our first source of policy data comes from Carollo (2024), which records the enactment date of major regulatory events for over 250 unique occupations that are estimated to cover more than 90% of licensed jobs in the United States. The dataset includes both state and federal laws and differentiates between licensure and less restrictive policies, such as state certification and registration.[12] This distinction is meaningful for our study, as regulatory stringency is a policy choice that frequently varies within an occupation.

The legal variables in the dataset were hand-coded according to the text of legislation, obtained from a comprehensive library of state session laws, as well as current statutes and administrative regulations. We observe changes to the extensive margin of policy over time, beginning with the year of initial regulation or licensure for each state-occupation pair. We can therefore study the origin of policies adopted as early as the late 1800s, when modern occupational licensing legislation began to emerge.[13] We structure the dataset as a balanced annual panel, focusing on the period from 1870 to 2020. Each occupation is classified according to the 2010 Standard Occupational Classification system, which facilities linking policies to information on task content and employment.

Licensing qualifications. Our second source of policy data comes from the Minnesota Occupational Licensing Law Research Project (OLLRP), an ongoing project that documents the evolution of specific licensing qualifications such as training hours and continuing education. We are in the process of collecting data on these legal requirements in each state from 1991 to 2021 for roughly 50 occupations.[14] This includes both "universally-licensed" occupations (those regulated by all states) and "partially-licensed" occupations (those where licensing is required in only some states). In Section 6 of this draft, we present case study results for a subset of occupations with preliminary data, though we plan to update our analysis as more data is compiled.

The key variables we observe in the OLLRP include secondary education requirements (high school completion or minimum grade level) and requirements associated with training, apprentice-

[12] Licensing is legally mandatory to work in the occupation and requires a specific demonstration of competency. State certification gives individuals who have met the government's competency standards the right to use specific job titles, but unlike licensure, does not prevent certified and uncertified workers from performing the same tasks. Registration simply requires that workers file their name and qualifications with a government agency before practicing, and potentially that they provide evidence that they are bonded and insured.

[13] The HeinOnline database used to compile the data includes laws passed by territorial legislatures, so initial licensing statutes are observable even when their enactment predates statehood. However, a small share of the policies we consider were adopted through administrative regulations rather than legislative acts. Since historical documentation of administrative law is less comprehensive than statutory law, enactment dates in these cases are typically pinned down using additional information gleaned from secondary sources.

[14] For a subset of these, we are extending our data back to the first licensing statute in each state. This is not possible for all occupations as many specify specific qualifications only in administrative regulations, which, unlike state session laws, have little historical documentation.

ships, on-the-job experience, and continuing education needed for licensees to obtain and maintain their license. We also observe whether a credential issued by a nongovernmental organization,such as a national professional association, is required for licensure and the legal authority (state agency or occupational licensing board) that directly regulates the occupation. Lastly, the database documents when different qualification requirements can be substituted for each other to fulfill the initial licensure requirement, which provides workers with multiple pathways to fulfill initial state licensure requirements.

3.2 Occupational attributes and task content

To summarize differences in the work performed by licensed and unlicensed occupations, we use data on task content and job attributes from the Occupational Information Network (O*NET) and its predecessor, the Dictionary of Occupational Titles (U.S. Department of Labor, Employment and Training Administration, 1939, 1977). As the latter does not provide numerical rankings for the characteristics we are interested in, we use natural language processing to predict these features from textual descriptions of job tasks (Gentzkow et al., 2019).

Occupational attributes. We begin by summarizing occupational attributes, using data from O*NET, a workforce survey sponsored by the U.S. Department of Labor. O*NET compiles information on job tasks, skill requirements, and hundreds of other features for over 900 occupations. From these descriptors, we select a subset that, in our view, plausibly reflects the public interest rationale for occupational licensing. Using a principal component analysis, we collapse these attributes into three summary indices. Although we do not claim that our measures are the *only* occupational attributes that might affect economic welfare, we do think that they capture several important characteristics that policymakers and interest groups have in mind when they highlight the risks of unlicensed practice (or lack thereof).

Our first indicator is constructed using three elements from the O*NET work context category "criticality of position": (i) consequence of error, (ii) freedom to make decisions, (iii) and frequency of decision-making. We see these features – in particular, "consequence of error" – as a measure of the potential risk that an unqualified worker might pose to the public. Our second indicator measures the degree to which a worker is likely to interact directly with consumers. Here, we use the elements: (i) assisting and caring for others, (ii) performing for or working directly with the public, (iii) contact with others, and (iv) physical proximity. Lastly, we measure the complexity of an occupation's task content using a set of 35 descriptors identified by Caines et al. (2017). Complex tasks require higher-order skills, which are difficult for workers to learn, and thus may also be challenging for non-experts to evaluate. We therefore think of task complexity as a proxy for asymmetric information between workers and consumers.[15]

[15] We are not the first to identify related factors as relevant for the adoption of licensing laws. For example, in a study of six licensed healthcare occupations, Graddy (1991a) uses educational requirements to proxy for service complexity and liability insurance rates to capture potential risks to consumers. Our approach extends this idea by offering measures of complexity and risk that can be defined for any occupation. Further, unlike educational

Predicting attributes from task descriptions. Although it provides a rich set of variables to draw from, O*NET has two key limitations in our context. First, it is measured near the end our sample, but occupations and tasks have changed materially over time (Atalay et al., 2020). Second, licensing itself might induce workers to view their job as risky or complex, which would bias our measures. To address these issues, we predict each occupation's attributes, using only the semantic meaning embedded in descriptions of its tasks, and use these predicted scores throughout our analysis. This approach allows us to extract attributes from historical editions of the DOT circa 1940 and 1980 that are directly comparable with those we compute from O*NET. Moreover, because our approach looks for associations between tasks and job attributes across the universe of O*NET occupations, it mitigates our concerns about potential reporting bias.

Briefly, we first collect the text of all core tasks listed under each O*NET occupation. We then follow the method described by Kogan et al. (2023) to convert each occupation's tasks into a single document embedding, a geometric representation of the text's meaning as a high-dimensional vector. Next, we train a regression model to estimate weights corresponding to each element of these document embeddings, with the goal of predicting each numerical attribute.[16] Finally, we take the weights we estimate from the O*NET data and apply them to document embeddings constructed from the DOT to predict our criticality, personal interaction, and complexity indices using historical task descriptions.

3.3 Establishment of professional associations

Lastly, we leverage the federated structure of professional associations to assess the influence of special interest groups on licensing policy. Professional associations are formed to represent the interests of incumbent practitioners in an occupation, which includes lobbying legislatures for policies that potentially benefit their members. Many associations have both a national headquarters and state chapters. These are usually organized at different times in different states, a fact that we exploit to capture variation in an occupation's local political organization.

We collected information on the founding dates of state-level professional associations for nine occupations: accountants, architects, civil engineers, dentists, lawyers, pharmacists, physicians, nurses, and realtors. This was obtained from a variety of sources, including association websites, occupational histories, and contacting the associations themselves. We researched associations representing many other occupations, but were unable to find reliable data on the history of their state-level associations. However, information about the founding of national organizations was often readily available, which we observe for 61 professional occupations.

attainment and insurance rates – which may be endogenous outcomes of the regulatory environment – our measures are directly related to the underlying task content of each occupation.

[16] Our approach is similar in spirit to that of Nedelkoska et al. (2023), who use O*NET task descriptions to train a neural network-based model, which they use to measure routine and non-routine task content in the DOT. It is also related to Mühlbach (2022), who uses O*NET word embeddings predict target characteristics without the use of training data, such as the "age friendliness" of an occupation (Acemoglu et al., 2022).

4 Long-run trends in occupational licensing requirements

We begin our empirical analysis by documenting a set of novel facts related to the growth and composition of licensed jobs over the past 150 years. We first leverage our rich data to introduce new annual estimates of historical licensing coverage and discuss how regulatory activity has tracked the evolution of work in the United States. Next, we show that both policy diffusion and sectoral employment trends are qualitatively important for explaining the rising share of U.S. workers who are subject to occupational licensing requirements.

4.1 Descriptive statistics

Modern occupational licensing statutes in the United States – characterized by minimum entry qualifications enforced by a government agency or regulatory board – began to emerge in the late 1800s as state governments took a more active role in regulating the medical and legal professions.[17] As shown in the left panel of Figure 2, licensing coverage rose significantly over the 20th century and has become a pervasive feature of the labor market. Our data show that the share of workers employed in licensed occupations grew from about 6% in 1950 to more than 20% today, broadly consistent with previous estimates of historical policy coverage (Kleiner and Krueger, 2010).[18] State certification and registration, by contrast, have declined in both absolute and relative importance. This finding suggests that policymakers have increasingly opted to regulate occupations through licensure rather than less restrictive policy alternatives.

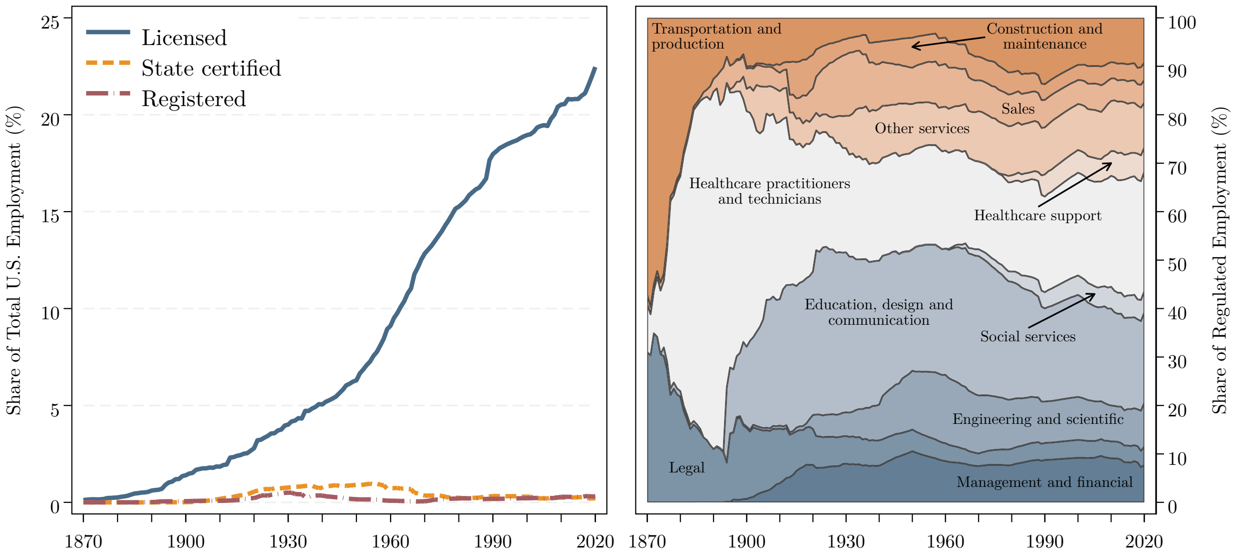
Composition of regulated employment. The right panel of Figure 2 shows that the composition of employment covered by licensing, certification, and registration laws has changed markedly over time, especially during the first half of our sample. Before 1900, licensing requirements applied mostly to healthcare and legal professionals, as well as some transportation workers (reflecting the federal licensing of ship captains, officers, and steam engineers). Over the first half of the 20th century, regulation spread to virtually all other major occupation groups, from professional and personal services to sales occupations and construction trades. While compositional changes are less pronounced after 1950, we do find a notable increase in the share of regulated employment attributable to healthcare and social service occupations starting around 1970.

Another feature worth highlighting in Figure 2 is that throughout our sample period, the com-

[17] Other forms of licensing existed much earlier, but these were significantly weaker and less uniform than the statutes we consider. Private medical societies, for instance, "licensed" physicians before the establishment of state medical boards. Although practicing medicine without a license was not illegal, unlicensed physicians had no legal recourse to recover unpaid fees (Kett, 1968). Attorneys were admitted to practice law by individual courts, and similarly, responsibility for teacher certification was placed with municipal or county authorities. We control for these policies in our main analysis below, but exclude them from our definition of licensure.

[18] Appendix B.1 provides additional details on how we define licensing coverage and Appendix B.2 on how we estimate occupational employment. While our policy data are the most comprehensive assembled to date, they are not completely exhaustive, and therefore our estimates of licensing coverage are likely understated by a few percentage points toward the end of our sample. Further, while the data are broadly representative across occupation groups, licenses for managerial occupations are underrepresented. That said, we think that the data are sufficiently comprehensive and representative to accurately characterize the aggregate trends that we discuss in this section.

Figure 2: Coverage and Composition of Occupational Licensing Policies



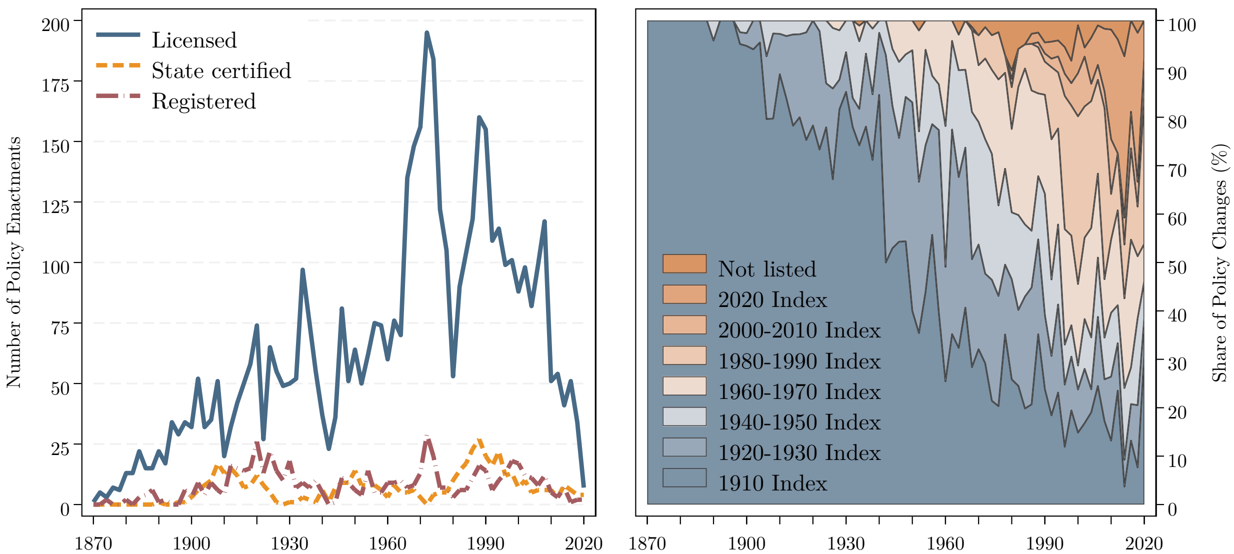
Notes: The estimates shown in this figure are constructed by aggregating policy indicators for state-by-occupation cells weighted by cell-level employment shares. Appendix B.2 describes our procedure for estimating employment in regulated occupation categories, which are typically more granular than Census classification codes. The left panel shows policy coverage as a share of total U.S. employment. The right panel shows the sectoral composition of regulated employment across major occupation groups. Both state and federal policies are included in this figure.

position of regulated employment differs significantly from that of the economy as a whole. At the end of our sample, for instance, healthcare and social services jobs employed just over 10% of U.S. workers, but represented around a third of licensed employment. More generally, professional and personal services are overrepresented among licensed occupations, while manufacturing occupations are significantly underrepresented. This fact already suggests that licensed occupations are highly selected with respect to their task content, which is consistent our model's predictions and a key point that we return to in the following section.

Timing of policy changes. Turning our attention from the stock of covered employment to the flow of policy changes, the left panel of Figure 3 plots the number of new licensing, certification, and registration laws enacted in each two-year interval since 1870. We find that licensing activity peaked in the early 1970s, but has declined sharply in recent years. While certification and registration account for a larger share of unweighted laws than their employment-weighted coverage suggests, licensing requirements still make up the vast majority of policy changes in our data.

The right panel of Figure 3 decomposes enactments into bins based on the age of the occupation being regulated. Following Lin (2011) and Autor et al. (2024), we identify when each occupation first appeared in the Census Classified Index of Occupations and Industries, allowing us to differentiate between older occupations, like barbers (listed in the first edition), and newer ones, such as polysomnographic technologists (added in 2020). We find that regulatory activity has tracked the emergence of new work over time, though usually with a considerable lag. In the past 30 years, for example, roughly 60% of policy changes applied to occupations that did not exist in 1950. While the

Figure 3: Timing of Policy Enactments and Composition of Occupations by Year of Addition to the Census Classified Index



Notes: The left panel of this figure shows the number of new licensing, certification, and registration policies enacted in each two-year interval since 1870. We aggregate across years in this way to smooth though annual variation driven by the timing of state legislative sessions. The right panel shows the unweighted composition of policy changes by the year each occupation was first included in the Census Classified Index of Occupations and Industries, which reflects the "age" or "vintage" of regulated occupations. Federal polices are excluded from this figure.

figure suggests that some licensing requirements began to diffuse very early in an occupation's history, the majority tend be enacted after the job has become established enough to warrant statistical classification, consistent with a positive correlation between market size and regulation.

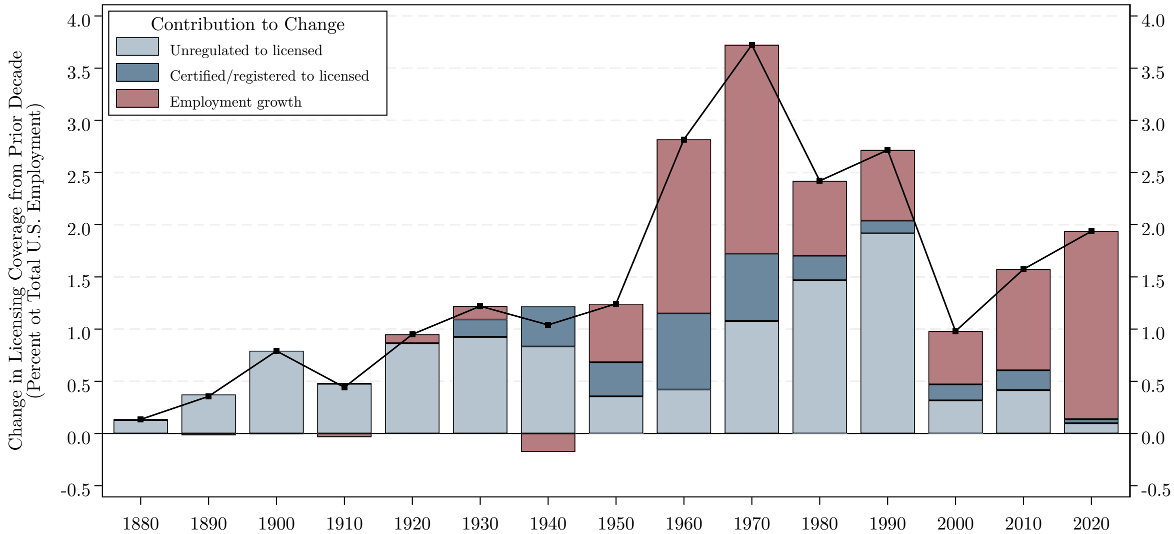
4.2 Sources of growth in licensing coverage

To what extent does the growth of licensing coverage reflect the diffusion of regulation to new states and occupations versus changes in the occupational composition of the U.S. economy? To answer this question, we decompose the decade-over-decade change in national licensing coverage shown in Figure 2 into three components: the contribution of new state-occupation cells becoming licensed, the replacement of existing certification or registration laws with licensing statutes, and changes in employment within previously-licensed cells. Noting that licensing coverage is an employment-weighted aggregation of policy indicators, we can write:

$$\Delta Coverage = \underbrace{\sum_i \sum_s \sum_k \omega_{ist} \left(\Delta L_{ist} \mid D_{is,t-1} = k\right)}_{\substack{\text{Contribution of} \\ \text{policy changes}}} + \underbrace{\sum_i \sum_s \Delta\omega_{ist} L_{is,t-1}}_{\substack{\text{Contribution of} \\ \text{employment growth}}}, \tag{10}$$

where $L_{ist} = 1$ if occupation i is licensed in state s at time t, ω_{ist} is the share of national employment in state-occupation cell i, s, and $D_{is,t-1} \in \{U, C, R\}$ denotes whether cell i, s was unregulated, certified, or registered in the previous decade.

Figure 4: Decomposition of Changes in Licensing Coverage



Notes: This figure decomposes decade-over-decade changes in licensing coverage (black line) into contributions from policy changes (blue bars) and employment growth in state-by-occupation cells that were previously regulated (red bars). Both state and federal policies are included in this figure.

Figure 4 shows that before 1950, the growth of licensing coverage was almost entirely driven by the adoption and diffusion of new policies. Between 1950 and 1990, the period when coverage was growing the fastest on average, policy changes and employment growth contributed about equally. The contribution of new policy changes peaked in 1990, owing largely to federal legislation requiring states that had not already done so to license commercial drivers and nursing assistants in long-term care facilities. Consistent with the decline in policy enactments shown in Figure 3, most of the increase in licensing coverage over the past 20 years reflects employment growth in previously-licensed occupations rather than the adoption of new policies.[19]

Lastly, we note that a sizeable share of the contribution of policy growth comes from the replacement of certification or registration policies by licensing requirements. In fact, over half the certification and registration policies in our data were eventually strengthened in this way, whereas we find that it is extremely rare for licensing policies to be weakened or repealed. We see this, together with the fact that certification and registration policies are relatively uncommon to begin with, as potentially consistent with the incentives for overregulation implied by our model.

[19] Since this decomposition depends on our estimates of occupational employment, we also conducted a counterfactual exercise where we estimated the share of employment that would be regulated today (given current employment) had policy diffusion stopped entirely in 1950. We find that licensing coverage would have risen from 6% to 12% rather than 23%. Thus, about two-thirds of the observed increase in licensing coverage after 1950 occurred in state-occupation cells that were initially licensed after that year, though as noted above, many occupations regulated after 1950 did not exist yet.

5 Origins of occupational licensing policies

In this section, we present our main findings on the origins of occupational licensing policies.[20] Following the implications of our model, we first assess whether the likelihood of regulation increases as public health, safety, or welfare concerns become more plausible. We then turn to the state-level determinants of regulation and show that, conditional on an occupation becoming licensed, policy innovation and diffusion follow predictable patterns across states. Lastly, we examine within-occupation factors driving the timing of initial regulation, with a focus on the role of market size, labor market competition, and political organization.

5.1 Occupational attributes and task content

Which occupations become licensed, and how does the likelihood of regulation depend on the tasks workers perform? Our model implies that although occupational regulation is not necessarily optimal, or even welfare improving in aggregate, policies that confer at least some benefit to consumers should receive greater political support than those with large social costs. Consistent with this prediction, we find that occupations whose tasks plausibly pose a greater risk to the public's health, safety, or welfare are indeed more widely regulated. That said, we also document significant heterogeneity in the task content of licensed occupations, and find that regulation has tended to diffuse from more to less risky and complex work over time.

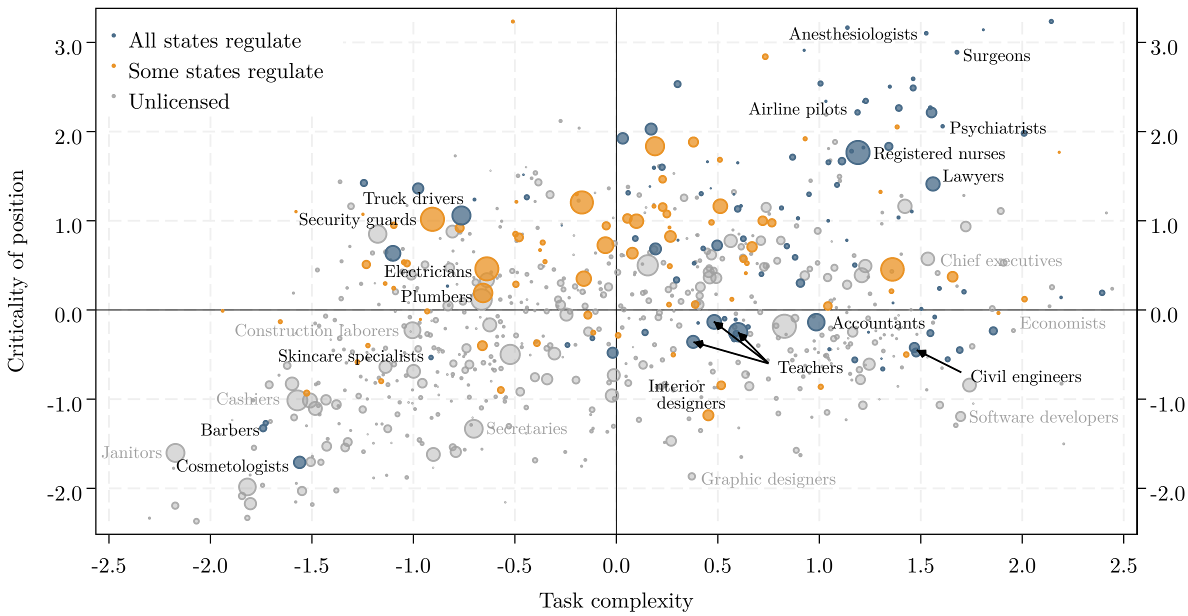
Cross-sectional evidence. In Figure 5, we plot our predicted criticality and task complexity measures for all six-digit occupations with task descriptions in O*NET.[21] Occupations regulated in all states as of 2020 are shaded in blue, while occupations regulated in only some states are shaded in orange. Most (though not all) licensed occupations appear in the top-right of the figure. That is, our text classification approach suggests that they are riskier and more complex than average. Conversely, occupations that are both less critical and complex than average (bottom left) are rarely regulated, with the notable exception of barbers and cosmetologists, occupations that all states currently license.

Table 1 presents a set of cross-sectional regressions summarizing the relationship between occupational attributes and regulation. Here, an observation is a state-occupation cell, and the outcome is an indicator for whether the cell is regulated. All specifications include state fixed effects, which control for potential differences in states' general propensity to regulate (see Section 5.2). Column one shows that each standard deviation increase in an occupation's criticality or complexity raises the probability of regulation by 7%. An increase in interpersonal interaction has a similar effect. The estimates are qualitatively similar in column two, which further compares occupations in the

[20] As state certification and registration policies are relatively uncommon and were generally adopted before licensing, in this section, we do not differentiate between regulatory methods, unless otherwise noted. Further, because policies (state-occupation pairs) are the unit of observation, our analysis is generally unweighted since we see no clear reason to discount information from small states and occupations when studying the determinants of policy adoption. However, we explore the robustness of our findings to these choices in Appendix A.

[21] In this subsection, we focus on policies with six-digit equivalents only. Although our data contain policy information for more granular occupations, these cannot be assigned to task measures.

Figure 5: Occupational Attributes and Number of Regulating States



Notes: This figure plots our predicted criticality and task complexity scores for all 6-digit occupations with O*NET task descriptions. Each attribute is standardized to have mean zero and standard deviation one. Markers shaded in blue indicate occupations that are regulated by all 48 states (or federally) in our sample as of 2020, while markers shaded in orange indicate occupations regulated by some, but not all, states. Marker sizes are proportional to estimated employment in 2019 based on ACS and OEWS data.

same two-digit group such as healthcare practitioners or construction trades. Columns three to six show that the positive correlation between these attributes and regulation also holds historically when using concurrent measures of occupational task content, though within occupation groups, we find a weaker relationship to task complexity.

In Appendix A.1, we present several additional findings and robustness checks. First, we show that using survey-based measures of license attainment from the Current Population Survey yields estimates extremely close to those reported in Table 1. This confirms that our main results are robust to an alternative measure of licensing that is entirely independent of our policy data. Second, we look at the type of regulation states choose to enact and find that mandatory licensing tends to be adopted for occupations that are both risky and complex, while the smaller number of voluntary state certification policies apply to occupations that are complex, but not risky. Third, we show that our estimates are robust to excluding healthcare, weighting policies by covered employment, and considering potential risk to coworkers as another possible rationale for licensing.

Regulatory timing and diffusion. So far, we have focused on differences between licensed and unlicensed occupations. We now ask whether these same attributes are related to the relative timing of policy adoption, conditional on an occupation becoming licensed in at least one state.

Figure 6 plots the mean of our three main occupational attributes over the flow of newly-enacted policies in each year. To smooth through year-to-year volatility resulting from variation in the total

Table 1: Occupational Attributes and Probability of Regulation

	Regulated in 2020		Regulated in 1980		Regulated in 1940	
	(1)	(2)	(3)	(4)	(5)	(6)
Criticality of position	0.071***	0.055**	0.058***	0.043**	0.041**	0.038*
	(0.015)	(0.020)	(0.018)	(0.019)	(0.018)	(0.020)
Task complexity	0.070***	0.040*	0.049***	0.004	0.030**	-0.000
	(0.013)	(0.022)	(0.012)	(0.022)	(0.012)	(0.022)
Interpersonal interaction	0.091***	0.102***	0.051***	0.052**	0.053***	0.045**
	(0.016)	(0.024)	(0.017)	(0.023)	(0.018)	(0.020)
Total observations	36,960	36,960	35,472	35,472	28,128	28,128
Occupations	770	770	739	739	586	586
R-squared	0.219	0.351	0.123	0.283	0.116	0.304
State FE	✓	✓	✓	✓	✓	✓
Major group FE		✓		✓		✓
Task measures	O*NET 2017		DOT 1977		DOT 1939	

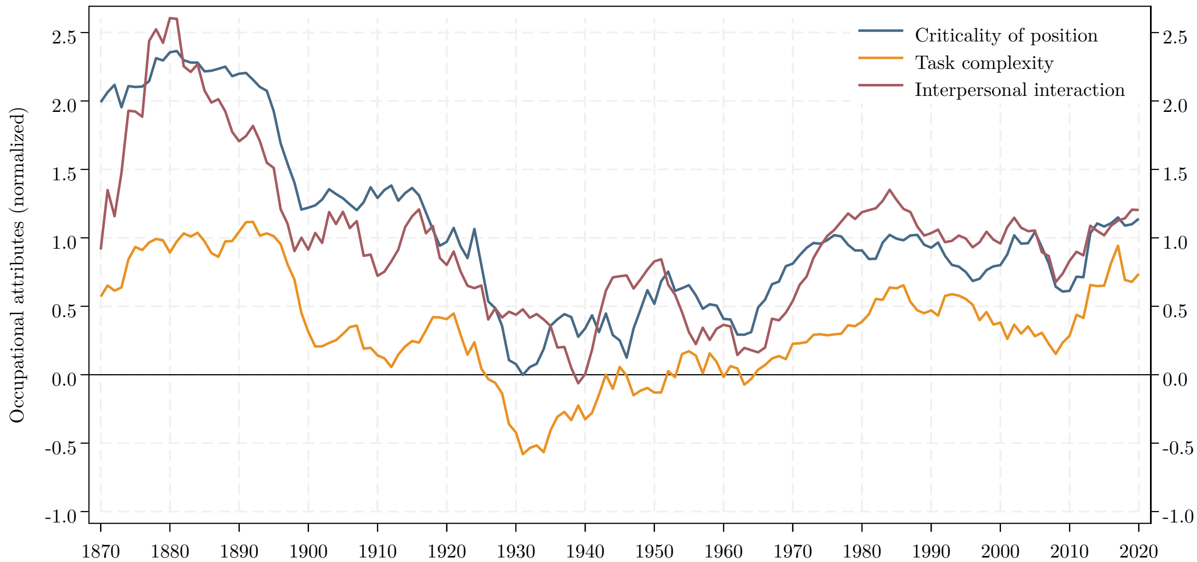
Notes: This table reports the results of regressing an indicator for whether a state-occupation cell is subject to any occupational regulation on our predicted occupational attributes. Major group fixed effects are defined at the level of 2-digit occupations. Significance levels based on standard errors clustered by state and occupation are indicated by *** 1%; ** 5%; and * 10%.

number of polices enacted, we focus on five-year moving averages. The figure implies that the first occupations to be regulated (which included attorneys, pharmacists, and physicians) were among the highest ranking on our task measures. From 1900 to 1950, however, licensing requirements diffused to less risky, complex, and interactive occupations. In fact, toward the end of this period, our measures show that newly-licensed occupations were actually somewhat *less* complex than the typical occupation, and no more risky (partially reflecting the licensing of hairdressers, manicurists, and other service occupations in these years). After 1950, these declining trends reversed, and the indices' averages stabilized around their current values by the 1980s.

One potential drawback of looking at aggregate trends is that set of occupations itself is changing as technological innovation and specialization drive the emergence of new jobs (Autor et al., 2024). The average occupational attributes we show may therefore mask patterns of regulatory diffusion within a given cohort or vintage of occupations. To address this issue, we estimate a discrete-time logistic hazard regression that models the order of policy adoption within states as a function of an occupation's attributes and cohort, which we define to be the year it was added to the Census Classified Index of Occupations and Industries.[22] Specifically, we estimate

[22] We estimate discrete-time models for two reasons. First, given variation in the timing of state legislative sessions, not every state is continuously "at risk" of policy adoption in our panel. Second, we study the regulation of hundreds of distinct occupations, essentially pooling many individual policy diffusion studies. Accounting for differences in baseline hazard rates across groups is straightforward in the conditional logistic regressions we estimate, as it requires only specifying the appropriate set of fixed effects to be interacted with time.

Figure 6: Average Occupational Attributes for Newly-Enacted Policies



Notes: This figure plots 5-year moving averages of our occupational attributes for newly-enacted policies. Here, we measure attributes using the vintage of task data that is closest to the median year of first regulation for each occupation. Each attribute is standardized to have mean zero and standard deviation one. Sample includes state and federal policies with six-digit SOC equivalents.

$$\ln\left(\frac{p(t; \mathbf{X}_{icst})}{1 - p(t; \mathbf{X}_{icst})}\right) = \alpha_{sct} + \sum_{d \in \mathcal{D}} \beta_d T_{id} + \epsilon_{icst} \tag{11}$$

where the left-hand side is the log odds of policy adoption for occupation i in state s at time t, T_{id} are our task measures, and α_{sct} is an indicator for state-cohort-years. These allow the baseline probability of policy adoption to vary non-parametrically over time within states and occupational cohorts. Observations in the panel are removed once policy adoption occurs, as they are no longer "at risk" of an event, but may reenter the estimation sample if laws are repealed or overturned.

The estimates in Table 2 confirm that licensing has generally diffused from more to less risky, complex, and interactive occupations, even within the set of occupations that some states ultimately chose to regulate. For ease of interpretation, we report the marginal effects implied by our logistic regressions, computed for an occupation with average attributes. The first column shows that a one standard deviation increase in each index (from zero) is associated with a 5–8% increase in the probability of regulation. This implies that as new occupations emerge, those ranking higher on these attributes were licensed earlier. Column two shows that within major occupation groups, criticality is the main attribute associated with early adoption. Columns three to six repeat this analysis, splitting the sample in 1950, the approximate inflection point of the trends shown in Figure 6. In both periods, we see similar patterns of policy diffusion after accounting for the emergence of new jobs, though these are slightly less pronounced after 1950 and, within major groups, less complex

Table 2: Event History Estimates of Occupational Attributes and Regulatory Diffusion

	Full Sample		Enacted 1870–1950		Enacted 1951–2020	
Marginal effect of attribute	(1)	(2)	(3)	(4)	(5)	(6)
Criticality of position	0.051***	0.086***	0.055***	0.049*	0.037***	0.059**
	(0.008)	(0.017)	(0.012)	(0.025)	(0.010)	(0.025)
Task complexity	0.080***	0.030	0.115***	0.153***	0.052***	-0.144***
	(0.006)	(0.019)	(0.009)	(0.030)	(0.008)	(0.030)
Interpersonal interaction	0.057***	0.009	0.084***	0.037**	0.034***	-0.011
	(0.006)	(0.015)	(0.008)	(0.018)	(0.008)	(0.029)
Total observations	83,648	11,056	54,993	6,382	28,655	4,674
Number of events	4,442	2,997	1,934	1,633	2,508	1,364
Occupations	146	112	78	65	118	80
State × cohort × year FE	✓	✓	✓	✓	✓	✓
Major group interactions		✓		✓		✓

Notes: This table shows the effect of occupational attributes on the probability of regulation in a discrete-time logistic hazard model. We report marginal effects evaluated at a vector of zeros (i.e., for an occupation with average attributes). Major group fixed effects are defined at the level of 2-digit occupations. Federal policies are excluded from this analysis as there is no variation in regulatory timing. Significance levels based on standard errors clustered by state-occupation cell are indicated by *** 1%; ** 5%; and * 10%.

occupations were licensed earlier on average.

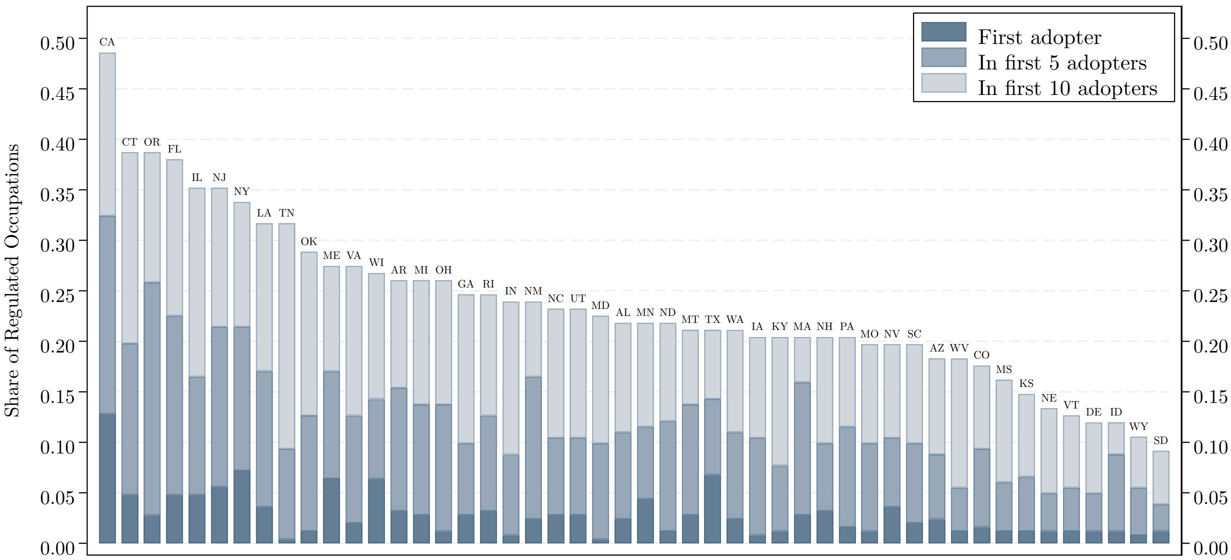
Discussion and interpretation. Our results in this section underscore the empirical significance of task content for explaining variation in regulation across occupations. As noted above, we think of our task measures as reasonable, if imperfect, proxies for the potential benefits of licensing that are defined for all occupations, regardless of actual regulatory outcomes.[23] Although these indices suggest that licensing is, on average, more common for occupations that plausibly pose some risk to consumers or the public, we emphasize (as our theoretical framework shows) that this is not a sufficient condition for the net benefits of licensing to be positive. Nor do we make any statement about the effectiveness of licensing requirements in practice, which is beyond the scope of this paper. Consistent with our analysis of occupations' task content, however, Kleiner and Soltas (2023) find that both consumer willingness to pay for licensing and its net welfare effects are more favorable in occupations that are widely regulated. Like our results, this suggests that socially beneficial policies are more likely to diffuse to all states than those with large social costs.

5.2 Policy innovation and diffusion

Are certain states more likely than others to originate and spread occupational licensing requirements? In this section, we abstract from occupation-specific characteristics and instead focus our

[23]The literature has long recognized that the earnings and employment effects of occupational licensing may be informative about its potential social benefits. Estimating these causal labor market effects, however, requires making comparisons between appropriate treatment and control units. When treatment effects are heterogeneous across occupations, we cannot obtain credible counterfactual wage and employment (or quality) effects for jobs that were never licensed, as there is no variation in treatment status. Addressing this measurement issue is one of the reasons we choose to focus on tasks when assessing the selective adoption of licensing policies.

Figure 7: Frequency of Early Policy Adoption by State, 1870–2020



Notes: This figure displays how often each state was within the first 1, 5, or 10 states to adopt a new licensing, certification, or registration policy. Since our policy data is coded at the annual level, ties occur when multiple states initially regulate an occupation in the same year. Rather than breaking these ties arbitrarily, a state enacting a policy at time t is assigned a rank equal to 1 + the number of states that adopted the policy at $t-1$ or earlier. Further, states that have not adopted a specific policy by the end of our sample cannot be ranked, so our results should be interpreted conditional on the total number of regulating states exceeding the specified adoption threshold.

attention on the state-level correlates of policy innovation and diffusion. We find that new licensing policies tend to be enacted earlier in large urbanized states with more regulation overall, then gradually diffuse along geographic, demographic, and political lines. This is consistent with our theoretical framework, where policymaker preferences, fixed costs, and market size influence regulatory decisions in a way that we would expect to be correlated within states.

Characteristics of early adopters. We first identify early policy adopters by ranking states according to the order in which they first licensed, certified, or registered the occupations in our data. We then compute the share of policies that originated in each state as well as how often they were among the first five or ten initial regulators. Figure 7 plots these rankings and shows that some states are consistently early adopters of licensing requirements while others rarely produce or take up new policies. We find, for instance, that roughly 25% of policies originated in three states: California, New York, and Texas. California stands out, not only because it is the leading origin of licensing legislation but also because it was within the first 10 states to regulate nearly half the occupations in our data. At the other end of the distribution, South Dakota and Wyoming were early adopters of only one in 10 policies.

Figure 7 suggests that early adopters tend to be large states (or their neighbors), while late adopters are often small and rural. The descriptive statistics in Table 3 confirm this observation. Since the relationship between these characteristics may have changed over time, and because addi-

Table 3: **State Descriptive Statistics by Frequency of Early Policy Adoption**

	1870–1950			1951–2020		
	Early Adopters (1)	Late Adopters (2)	Difference: Early-Late (3)	Early Adopters (4)	Late Adopters (5)	Difference: Early-Late (6)
Log total resident population	14.76	13.21	1.55	15.55	14.24	1.31
	(0.86)	(1.34)	[0.01]	(0.77)	(1.09)	[0.01]
Urban population share (%)	53.72	25.91	27.81	74.03	62.90	11.13
	(13.69)	(9.50)	[0.00]	(14.15)	(13.17)	[0.09]
Share of population that is white (%)	92.42	84.01	8.41	80.47	90.33	–9.86
	(12.66)	(18.27)	[0.25]	(6.50)	(5.25)	[0.00]
Share of population that is foreign-born (%)	23.23	14.40	8.83	11.17	5.42	5.75
	(9.70)	(11.02)	[0.07]	(7.63)	(1.69)	[0.03]
Log income per capita	9.48	9.14	0.34	10.39	10.37	0.01
	(0.25)	(0.38)	[0.03]	(0.19)	(0.09)	[0.83]
Republican presidential vote share (%)	50.50	45.00	5.50	51.17	55.97	–4.80
	(9.89)	(17.40)	[0.40]	(3.80)	(6.42)	[0.06]
Unified Democratic government	0.20	0.42	–0.22	0.42	0.15	0.26
	(0.29)	(0.42)	[0.18]	(0.19)	(0.16)	[0.00]
Unified Republican government	0.48	0.39	0.09	0.15	0.40	–0.26
	(0.27)	(0.35)	[0.53]	(0.07)	(0.23)	[0.00]
Southern state	0.10	0.50	–0.40	0.50	0.10	0.40
	(0.32)	(0.53)	[0.05]	(0.53)	(0.32)	[0.05]
Unemployment rate (%)				6.26	5.02	1.24
				(0.70)	(0.96)	[0.00]
Union representation (%)				16.01	14.30	1.70
				(6.64)	(3.82)	[0.49]
Change in union representation (1970–2020)				–13.74	–13.53	–0.21
				(4.96)	(5.20)	[0.93]
Right-to-work state				0.50	0.50	0.00
				(0.53)	(0.53)	[1.00]
Days in legislative session				129.07	132.81	–3.74
				(61.11)	(43.62)	[0.88]
Log expenditure per legislator				13.44	12.47	0.97
				(0.92)	(0.90)	[0.03]
Log legislator salary				11.23	10.48	0.75
				(0.58)	(0.95)	[0.05]

Notes: This table reports descriptive statistics for states ranked by their average order of policy adoption. We first compute the fraction of occupations each state was among the first 10 to regulate. Early adopters are then defined as the top 10 states sorted by this measure, and late adopters as the bottom 10. Columns 1, 2, 4, and 5 report variable means and standard deviations. Columns 3 and 6 report the difference in means between early and late adopters with the p-values in brackets. We take averages over the full sub-period but note that not all variables are available in every year. Variable definitions and sources are provided in Section B.5

tional variables become available in later years, we report our results separately for policies enacted in the first (1870–1950) and second (1951–2020) half of our sample. Across both time periods, early adopters are larger, are more urbanized, and have higher foreign-born population shares. Second, since 1950, unified Democratic and Republican governments have tended to be early and late adopters, respectively. Before 1950, the opposite was true. Third, the locus of early adoption has shifted from non-Southern to Southern states. Finally, we find little difference in average legislative session length between groups, but early adopters have higher expenditures per legislator, which potentially reflect greater access to professional support staff (Bowen and Greene, 2014).

Policy diffusion. A vast political science literature has examined the determinants of policy diffusion across states, often focusing on the roles of geographic proximity and ideological similarity (Mallinson, 2020). While we abstract from strategic interaction between states in our theoretical framework, these factors may be important in practice. For instance, policymakers might borrow regulatory models from nearby jurisdictions because their neighbors are culturally, demographically, or economically similar to their own. Geographic proximity could also generate economic competition or externalities between states (Baybeck et al., 2011). Further, states governed by the same political party likely have similar policy preferences and hence may be quicker to embrace laws that were adopted by other Democrat- or Republican-controlled states.[24]

To understand which of these factors are relevant for the diffusion of occupational licensing, we follow DellaVigna and Kim (2022) and summarize the clustering of early adopters using a version of Geary's C-statistic. Let L_{is} be an indicator for whether occupation i is licensed, certified, or registered in state s when the total number of states regulating the occupation first reaches 10. Further, define weights $\omega_{sr} \in \{0, 1\}$ indicating whether state r is among the one-third of states closest to s based on some distance metric. Then, the clustering statistic for occupation i is defined as

$$C_i = 1 - \frac{\frac{1}{48 \cdot 16} \sum_{s=1}^{48} \sum_{s \neq r} \omega_{sr} (L_{is} - L_{ir})^2}{\frac{1}{48 \cdot 47} \sum_{s=1}^{48} \sum_{s \neq r} (L_{is} - L_{ir})^2}. \tag{12}$$

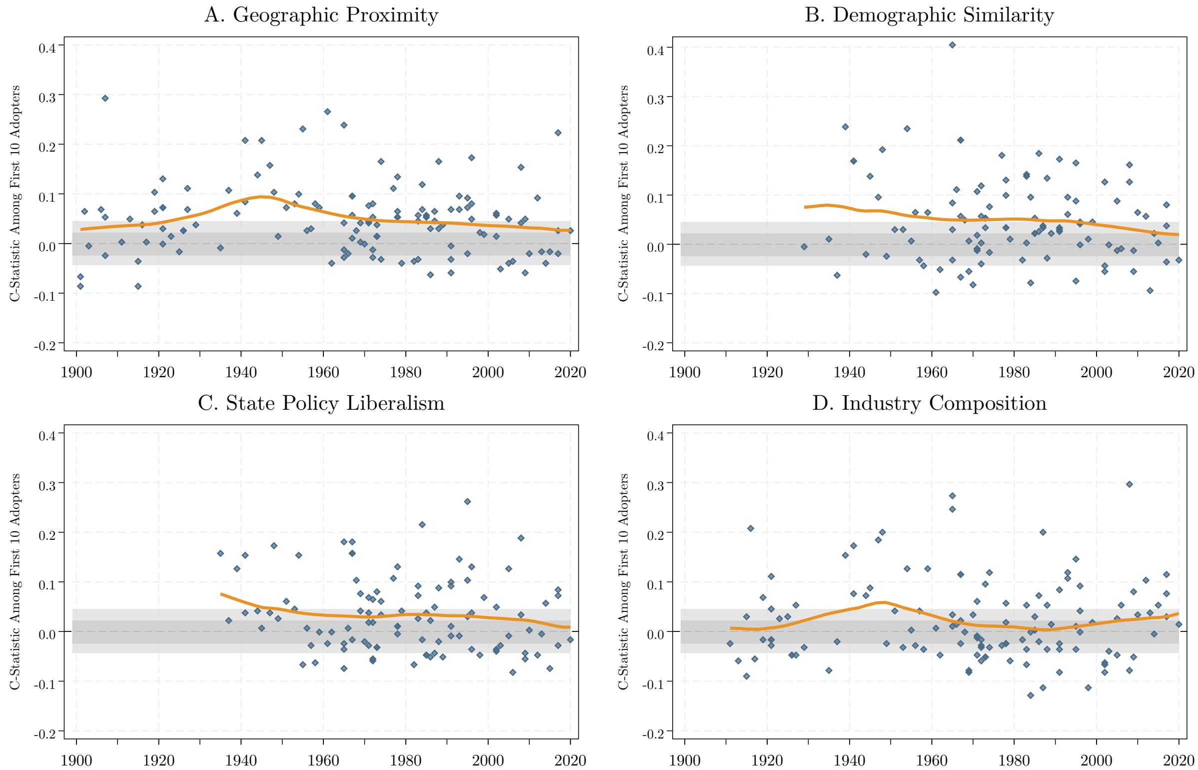
This measure compares the average of proximity-weighted policy differences across states with the corresponding unweighted average. If policy enactments display no clustering, the weighted and unweighted averages are equal, and $C_i = 0$. If instead similar states tend to adopt (or not adopt) the same laws, the weighted average will be small relative to the unweighted average, and $C_i > 0$. Conversely, $C_i < 0$ implies that polices initially diffuse across dissimilar states.

Figure 8 plots the distribution of C_i over time based on four distance metrics: geographic proximity, demographic similarity, political ideology, and industry composition. Panels A and B show that early adopters of licensing requirements tend to be highly clustered geographically and with respect to their demographic characteristics. For most of our sample, the average licensing policy displays levels of geographic and demographic clustering at or above the 75th percentile of what we would expect if policies were enacted randomly. In Panel C, we find evidence that licensing has historically diffused along ideological lines as well, though all three relationships appear to have weakened somewhat in the last 30 years. Finally, Panel D shows a weaker relationship, on average, between policy diffusion and variation in state industry composition.

Discussion and interpretation. The patterns we identify in this section for the adoption of

[24] The ideological position of previous adopters can also serve as an informational shortcut, enabling decision-makers to "minimize the uncertainty about how issues fit in the liberal-conservative policy space" (Grossback et al., 2004). That is, a policymaker might infer latent partisan support from the observed decisions of other states.

Figure 8: State Characteristics and Correlated Policy Diffusion



Notes: This figure shows the distribution of Geary's C-statistic for early policy enactments, using alternative distance metrics. Each point represents a distinct occupation that was licensed, certified, or registered in at least 10 states. Values of C greater than zero indicate that the first 10 states to enact a policy were clustered geographically or were similar with respect to their demographic characteristics, legislative ideology, or industry composition. To measure the ideology of state legislators, we use the state policy liberalism index constructed by Caughey and Warshaw (2014), which captures the enactment of legislation generally associated with the political left. Solid lines depict a local polynomial trend with a bandwidth of 10 years. The darker shaded area represents the 25th to 75th percentiles of the C-statistic under the null of random policy adoption, and the lighter shaded region represents the 10th to 90th percentiles of this distribution.

occupational licensing laws are broadly consistent with those documented in other policy diffusion studies. Like Mulligan and Shleifer (2005) and Law and McLaughlin (2022), we find that regulation increases with population size. Similarly, the impact of geographic proximity we identify echoes a long-standing theme of state policymaking research (Walker, 1969; Mooney, 2020). Unlike several other recent papers, however, our results show little evidence that the enactment of licensing laws has become increasingly politically polarized over time (Caughey et al., 2017; Mallinson, 2021; DellaVigna and Kim, 2022). While Democrat-led states were more likely than Republican-led states to enact new policies after 1950, licensing has, if anything, become less ideologically clustered since 2000.

Given that the diffusion of occupational licensing has historically followed patterns similar to other policies, we might expect early adopters of licensing legislation to be prolific regulators in general. Indeed, early adopters also license a larger number of occupations, as shown in Appendix A.2.

Table 4: Correlation Between Early Policy Adoption and Total Volume of State Law

	Log kilobytes of statutory code			Log words in administrative code		
	First Adopter (1)	In First 5 Adopters (2)	In First 10 Adopters (3)	First Adopter (4)	In First 5 Adopters (5)	In First 10 Adopters (6)
Volume of state law	0.018***	0.044***	0.060***	0.011***	0.033***	0.045***
	(0.003)	(0.008)	(0.012)	(0.003)	(0.007)	(0.010)
Observations	35	35	35	45	45	45
R-squared	0.527	0.470	0.429	0.208	0.340	0.315
Dependent variable mean	0.028	0.124	0.236	0.028	0.124	0.236

Notes: This table displays the results of regressing the share of occupations that each state was the first, or among the first five or 10 states, to regulate on measures of the state's total volume of statutes and administrative code. Kilobytes of statutory law come from Mulligan and Shleifer (2005). This variable is measured in 2001–2003 and is missing for the following states: CO, GA, KS, KY, LA, MD, NH, NM, ND, OH, OK, VA, and WI. Total words in state administrative code from 2017 to 2021 are obtained from QuantGov. Data are missing for AR, NJ, and VT. Robust standard errors are in parentheses and significance levels are indicated by *** 1%, ** 5%, and * 10%.

Moreover, Table 4 documents a strong positive correlation between licensing and the total volume of legal text that states have accumulated across all subjects. Column three, for example, shows that a one standard deviation increase in statutory law is associated with a 6 percentage point increase in a state's share of "first 10" licensing policy adoptions. Further, this relationship is highly predictive, with R^2 values ranging from 0.4 to 0.5. Looking across occupations, therefore, we see that state-level factors appear to explain a significant share of the variation in licensing requirements and their relative timing across jurisdictions.

5.3 Within-occupation evidence on the timing of licensing statutes

We now present our main findings on the determinants of policy timing within individual occupations. Leveraging the full panel structure of our data, we estimate a discrete-time logistic hazard model of the form

$$\ln\left(\frac{p(t;\mathbf{X}_{ist})}{1-p(t;\mathbf{X}_{ist})}\right) = \alpha_{it} + \beta'\mathbf{X}_{ist} + \epsilon_{ist}, \tag{13}$$

where $p(t;\mathbf{X}_{ist})$ is the conditional probability of policy adoption given a vector of time-varying state and occupation covariates $\mathbf{X}_{ist}$. Whereas the hazard regressions estimated in Section 5.1 captured the diffusion of regulation across occupations within states, in this section, we study policy diffusion across states for a given occupation, holding fixed common occupational characteristics such as task content.

Event history estimates, 1870–1940. We initially estimate Equation 13 using data from 1870 to 1940, which cover the first major wave of occupational regulation in the United States. It is also the period during which we observe our most precise measure of occupational employment, which we construct using write-in job titles from full-count Census microdata (Ruggles et al., 2024a). Table 5 presents estimates from our hazard model as well as marginal effects computed at sample means.

Column one reports estimates from a model that includes state-level variables only. Consistent with our findings in the previous section, we find that larger states enacted occupational regulation earlier. We also find some evidence that states enacting more legislation associated with the progressive movement were early adopters of occupational regulation, as were states that granted women the right to vote before ratification of the 19th Amendment.[25] The marginal effects of these variables, however, are not statistically significant when evaluated at sample means.

In columns three and four, we add occupation-level characteristics to the regression. Both market size, as measured by the total number of workers in the occupation, and the share of neighboring states that have already regulated the occupation are positively associated with the hazard of regulation: a one hundred log point increase in employment increases the probability of regulation by about 9% on the margin. Interestingly, once we control for occupation size, the coefficient on state population becomes negative. That is, holding the number of workers in an occupation fixed, regulation was more likely to be adopted in smaller states. In other words, large states tend to be early regulators because they usually have more workers in regulated occupations. Finally, we find that hazard rates are state dependent. Specifically, private credentialing and locally enforced regulation tend to slow the enactment of state-level regulation.[26]

Columns five and six repeat this analysis, restricting attention to licensing laws only (i.e., excluding about 300 state certification and registration policies). Unsurprisingly, the results are broadly similar to those reported in columns three and four, though the point estimates are a bit less precise. In contrast to private credentialing requirements and local regulation, we find that state certification and registration increase the hazard of licensing. This may be because jurisdictions that previously enacted some form of state regulation have already realized the fixed costs associated with drafting statutes and setting up administrative boards, lowering the marginal cost of future amendments. It could also reflect the incentives of incumbents to raise barriers to entry over time, especially once some form of regulation has already been established.

Labor market competition. Our finding that larger markets are regulated earlier (as measured by the number of workers in an occupation and state) is consistent with the view that fixed costs limit the supply of regulation (Mulligan and Shleifer, 2005). However, it is also possible that greater competition in the labor market increases demand for regulation among incumbents. We test this hypothesis by leveraging historical immigration flows to isolate plausibly exogenous variation in labor market competition experienced by native-born workers.[27]

[25]Several female-majority occupations, such as teachers and nurses, were initially regulated before 1920.

[26]Private credential requirements are not administered by a government licensing board but still place some legal restrictions on professional practice or advertising. An 1881 law in Wisconsin, for example, made it unlawful for physicians without a diploma or membership in a private medical society to "assume the title of doctor, physician, or surgeon" (1881 Wisconsin Acts, Chapter 256). Locally enforced regulation, in which states delegated responsibility for licensing and regulation to county or municipal governments, was more common. For example, responsibility for teacher examination and certification was initially placed with county superintendents in most states.

[27]In addition to raising training costs, many licensing statutes enacted during the early 20th Century included citizenship or state residency requirements, which may have been especially effective at limiting competition from immigrants. With respect to Wisconsin's Certified Public Accountant law, Cross (1998) notes that "Requiring U.S. citizenship was important in protecting the fledgling American profession from overwhelming competition from its

Table 5: Event History Estimates of Factors Influencing the Timing of Occupational Regulation (Licensing, Certification and Registration Laws Enacted 1870–1940)

	Initial Regulation		Initial Regulation		Initial Licensure	
	(1)	(2)	(3)	(4)	(5)	(6)
	Coef.	dy/dx	Coef.	dy/dx	Coef.	dy/dx
A. State demographic and political characteristics						
Log total state population	0.164***	0.008***	-0.142**	-0.035*	-0.113*	-0.025
	(0.042)	(0.003)	(0.061)	(0.018)	(0.068)	(0.021)
Urbanization rate (%)	0.001	0.000	-0.004*	-0.001*	-0.005*	-0.001*
	(0.003)	(0.000)	(0.002)	(0.001)	(0.003)	(0.001)
Share foreign-born (%)	0.005	0.000	0.010**	0.002**	0.012**	0.003*
	(0.005)	(0.000)	(0.005)	(0.001)	(0.005)	(0.001)
Unified Democratic government	-0.103	-0.005	-0.067	-0.017	-0.116	-0.026
	(0.117)	(0.007)	(0.117)	(0.029)	(0.140)	(0.033)
Unified Republican government	0.069	0.003	0.079	0.019	0.071	0.016
	(0.081)	(0.004)	(0.076)	(0.018)	(0.092)	(0.020)
Women's suffrage	0.257**	0.013	0.249**	0.061**	0.215	0.048
	(0.129)	(0.009)	(0.119)	(0.029)	(0.143)	(0.034)
Progressive legislation index	0.091***	0.005	0.090**	0.022**	0.042	0.009
	(0.034)	(0.003)	(0.035)	(0.009)	(0.041)	(0.009)
B. Occupational characteristics						
Log occupational employment			0.352***	0.086***	0.373***	0.082***
			(0.061)	(0.020)	(0.065)	(0.031)
Neighboring states regulating (%)			0.864***	0.212***	0.891***	0.197***
			(0.141)	(0.037)	(0.144)	(0.057)
C. State dependence: Lagged regulation variables						
Private credential requirement			-0.227	-0.056	-0.304	-0.067
			(0.322)	(0.079)	(0.322)	(0.074)
Local registration/license			-0.506*	-0.124*	-0.657**	-0.145*
			(0.289)	(0.073)	(0.320)	(0.084)
State registration					1.174***	0.259***
					(0.310)	(0.091)
State certification					0.652	0.144
					(0.436)	(0.108)
Total sample observations	22,264	22,264	21,648	21,648	18,520	18,520
Number of events	1,427	1,427	1,412	1,412	1,140	1,140
Number of occupations	80	80	76	76	71	71

Notes: This table reports the results of a discrete-time hazard model estimating the impact of various political and economic characteristics on the probability of adopting an occupational regulation or licensing law. Each pair of columns is obtained from a separate conditional logistic regression that absorbs occupation-by-year fixed effects. Marginal effects are computed at the sample mean. Sample excludes federally regulated occupations and state-occupation cells with no workers. Standard errors are clustered at the occupation level, and significance levels are indicated by *** 1%, ** 5%, and * 10%.

Assessing the relationship between immigration and regulation is natural over this time period for three reasons. First, the late 19th and early 20th centuries were marked by large inflows of foreign-born workers, and there is evidence that other labor market institutions such as unions developed, at least in part, to shield native workers from increased competition (Medici, 2024).

British counterpart. Citizenship and state residency played a part in the political process that preceded passage of the 1913 Certified Public Accountant Act."

Second, the results reported above show that states with a larger stock of foreign-born workers regulated occupations earlier, on average. Finally, our model implies that regulation is most effective at raising the wages of incumbents when licensing costs are heterogeneous, and obtaining a license may be especially costly for immigrants because of language barriers or differences in educational credentials (Federman et al., 2006). Thus, we expect stronger incentives for native-born workers to seek licensure in states and occupations where they experience more direct competition with recent immigrants.

To measure labor market competition between immigrants and natives, we construct an occupation specific shift-share instrument (Card, 2001). Let $\hat{\alpha}^j_{s,t-5}$ be the share of all European migrants between the ages of 14 and 65 from origin county j living in state s, lagged five years.[28] This term captures the well-documented fact that new immigrants tend to settle in locations where previous migrants from the same origin country have already established themselves. We modify the standard immigration shift-share with an additional term that captures variation in ethinic-occupational composition, which reflects factors such as differences in skill background (Locke, 2024). Let $\hat{\gamma}^j_{-si,t-5}$ be the share of all migrants from county j who are employed in occupation i in all states *except* s, lagged five years. We use a leave-out measure of occupational employment to purge the instrument of state-specific demand shocks that might be correlated with both the timing of regulation and the share of immigrants who choose to enter the occupation in a given state.

Using data on the number of arrivals from county j in the previous five years (Barde et al., 2006), we construct the following measure of predicted labor market competition:

$$\hat{Z}_{ist} = \frac{1}{Natives_{is,t-5}} \sum_j \hat{\alpha}^j_{s,t-5} \times \hat{\gamma}^j_{-si,t-5} \times Immigrants^j_{t,t-5}. \tag{14}$$

This shift-share apportions the total number of new immigrants from each country of origin across states and occupations according to the geographic distribution and occupational employment of previous migrants. We normalize these counts by the number of native-born workers in the occupation in state s at time $t-5$. Finally, we windsorize $\hat{Z}_{ist}$ at the 99th percentile to reduce the influence of extreme outliers. Table A5 shows that, before initial regulation, an increase in $\hat{Z}_{ist}$ is associated with stronger growth in the share of foreign-born workers in an occupation, as expected.

Table 6 reports reduced-form estimates from our logistic hazard model, including predicted labor market competition as an additional covariate. As before, our specification includes occupation-by-year fixed effects, implying that the estimates are identified from variation in geographic distribution of immigrants. We include, but do not report, state demographic and political variables, including log population and the aggregate foreign-born share of each state. Consistent with theories of regulation that emphasize the private incentives of incumbent workers, we find that an increase in labor market competition raises the hazard of regulation, relative to the hazard for same occupation in other states (holding market size fixed). The size of this effect is economically significant: the

[28] As with the rest of our event history analysis, we observe stock variables in census years only, which we linearly interpolate to construct a balanced panel. We observe immigration flows annually.

Table 6: Event History Estimates of Immigrant Competition and Regulation Timing (Licensing, Certification and Registration Laws Enacted 1870–1940)

	Initial Regulation		Initial Licensure	
	(1)	(2)	(3)	(4)
Occupational characteristics	Coef.	dy/dx	Coef.	dy/dx
Predicted competition from new immigrants	0.509**	0.127**	0.652***	0.161**
	(0.223)	(0.056)	(0.247)	(0.067)
Log total occupational employment	0.356***	0.089***	0.394***	0.097***
	(0.074)	(0.018)	(0.079)	(0.025)
Neighboring states regulating (%)	0.792***	0.198***	0.796***	0.197***
	(0.157)	(0.040)	(0.164)	(0.040)
Total sample observations	20,297	20,297	17,354	17,354
Number of events	1,362	1,362	1,092	1,092
Number of occupations	76	76	71	71
Demographic and political controls	✓	✓	✓	✓
Lagged regulation variables	✓	✓	✓	✓

Notes: This table reports the results of a discrete-time hazard model estimating the impact of occupational characteristics on the probability of adopting an occupational regulation or licensing law. Each pair of columns is obtained from a separate conditional logistic regression that absorbs occupation-by-year fixed effects. Marginal effects are computed at the sample mean. Sample excludes federally regulated occupations and state-occupation cells with no workers. Standard errors are clustered at the occupation level, and significance levels are indicated by *** 1%, ** 5%, and * 10%.

marginal effect reported in column two implies that one additional predicted immigrant per native worker over a five-year window increases the probability of regulation by 12% (given other state and occupation characteristics). This effect is slightly larger when restricting the sample to licensing policies, which present greater barriers to entry than certification or registration laws.

Event history estimates, 1980–2020. Turning now to more recent policies, we estimate Equation 13 using data from the past 40 years, limiting our sample to occupations defined at the six-digit level or higher. As shown in Table 7, during this period state demographic and political characteristics are more strongly associated with the hazard of regulation than occupational employment and earnings are. Specifically, the hazard rates for regulation and licensing are higher in urban states and those with a smaller share of workers represented by unions, though not in larger states or those with more immigrants conditional on other factors. Unified Democratic governments are more likely than divided governments to regulate occupations, while unified Republican governments regulate less. Further, the hazard of regulation is higher for legislatures identified as more politically liberal, even after controlling for partisan balance (Caughey and Warshaw, 2014).

Unlike our results before 1940, our results after 1980 show no relationship between occupation size and regulation. There is some evidence that the hazard of regulation may be higher in states with lower occupation-specific earnings, but this relationship is not statistically significant. While employment and earnings are measured somewhat imprecisely in our post-1980 data, these results

Table 7: Event History Estimates of Factors Influencing the Timing of Occupational Regulation (Licensing, Certification and Registration Laws Enacted 1980–2020)

	Initial Regulation		Initial Regulation		Initial Licensure	
	(1)	(2)	(3)	(4)	(5)	(6)
	Coef.	dy/dx	Coef.	dy/dx	Coef.	dy/dx
A. State demographic and political characteristics						
Log total state population	0.023	0.005	-0.016	-0.004	-0.009	-0.002
	(0.045)	(0.009)	(0.111)	(0.026)	(0.112)	(0.026)
Urbanization rate (%)	0.010***	0.002***	0.011***	0.003**	0.012***	0.003**
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)
Share foreign-born (%)	-0.014	-0.003	-0.011	-0.003	-0.005	-0.001
	(0.009)	(0.002)	(0.008)	(0.002)	(0.009)	(0.002)
Union coverage (%)	-0.026***	-0.006**	-0.027***	-0.007***	-0.019**	-0.005**
	(0.009)	(0.002)	(0.009)	(0.002)	(0.008)	(0.002)
Unified Democratic government	0.265***	0.058***	0.261***	0.064***	0.231***	0.055**
	(0.073)	(0.020)	(0.075)	(0.022)	(0.081)	(0.023)
Unified Republican government	-0.134	-0.030	-0.156*	-0.038*	-0.285***	-0.068**
	(0.091)	(0.021)	(0.092)	(0.022)	(0.101)	(0.027)
State policy liberalism	0.196***	0.043**	0.196***	0.048***	0.119**	0.029**
	(0.058)	(0.018)	(0.057)	(0.015)	(0.059)	(0.014)
B. Occupational characteristics						
Log occupational employment			0.037	0.009	-0.039	-0.009
			(0.098)	(0.023)	(0.101)	(0.026)
Log average hourly earnings			-0.235	-0.057	-0.213	-0.051
			(0.201)	(0.049)	(0.231)	(0.054)
Neighboring states regulating (%)			0.090	0.022	0.017	0.004
			(0.210)	(0.052)	(0.206)	(0.049)
C. State dependence: Lagged regulation variables						
Private credential requirement			-0.466*	-0.114	-0.274	-0.066
			(0.266)	(0.071)	(0.274)	(0.068)
Local registration/license			1.074***	0.262**	1.135***	0.272**
			(0.374)	(0.105)	(0.383)	(0.115)
State registration					-0.111	-0.027
					(0.410)	(0.098)
State certification					-0.229	-0.055
					(0.288)	(0.072)
Total sample observations	17,727	17,727	17,362	17,362	16,947	16,947
Number of events	1,150	1,150	1,133	1,133	1,026	1,026
Number of occupations	86	86	85	85	79	79

Notes: This table reports the results of a discrete-time hazard model estimating the impact of various political and economic characteristics on the probability of adopting an occupational regulation or licensing law. Each pair of columns is obtained from a separate conditional logistic regression that absorbs occupation-by-year fixed effects. Marginal effects are computed at the sample mean. Sample is limited to occupations with six-digit equivalents and excludes federally regulated occupations and state-occupation cells with no workers. Standard errors are clustered at the occupation level, and significance levels are indicated by *** 1%, ** 5%, and * 10%.

suggest that market-specific factors may have become relatively less important for the timing of initial regulation within occupations.[29] Brinegar (2006), for instance, notes that while regulatory

[29] Unlike our pre-1940 data, which are based on full population counts, our post-1980 data is derived from 5% extracts of the decennial Census and the American Community Survey, which covers 1% of the population after

efforts were historically initiated by the profession itself, "State legislatures show signs of departing from their customary professional licensing approach as new professions gain state licensure without initiation by a profession or the public" (pg. 457). Further, the licensing of several occupations after 1980 – including commercial drivers, nursing assistants, real estate appraisers, and mortgage loan originators – was federally mandated, though our findings are little changed when dropping these occupations.

5.4 Professional associations and political influence

Finally, we assess the role of professional associations in the enactment of occupational licensing and regulation. As noted above, many associations historically supported licensing requirements and actively lobbied state policymakers for their introduction. Three years after its founding in 1903, for instance, the Maryland State Association of Graduate Nurses wrote that

> State Registration has been the principal feature of the preliminary work at least of all the Associations, and this has been successfully carried through by twelve of the twenty-one Societies. Those Societies not as yet successful are nothing daunted and declare their intention of continuing their efforts until bills meeting their requirements are finally passed. — Carr (1906, pg. 22)

Here, we quantify the relationship between local political organization and the enactment of occupational regulation, using data on the year professional associations representing nine occupations were initially organized in each state. We have information on associations representing accountants, architects, civil engineers, dentists, lawyers, pharmacists, physicians, professional nurses, and realtors. Together, these nine occupations represent approximately 20% of all licensed workers in 2019. Because certification or registration statutes were often enacted before licensing for several of these occupations, we treat any method of regulation administered by a state agency as the outcome of interest.

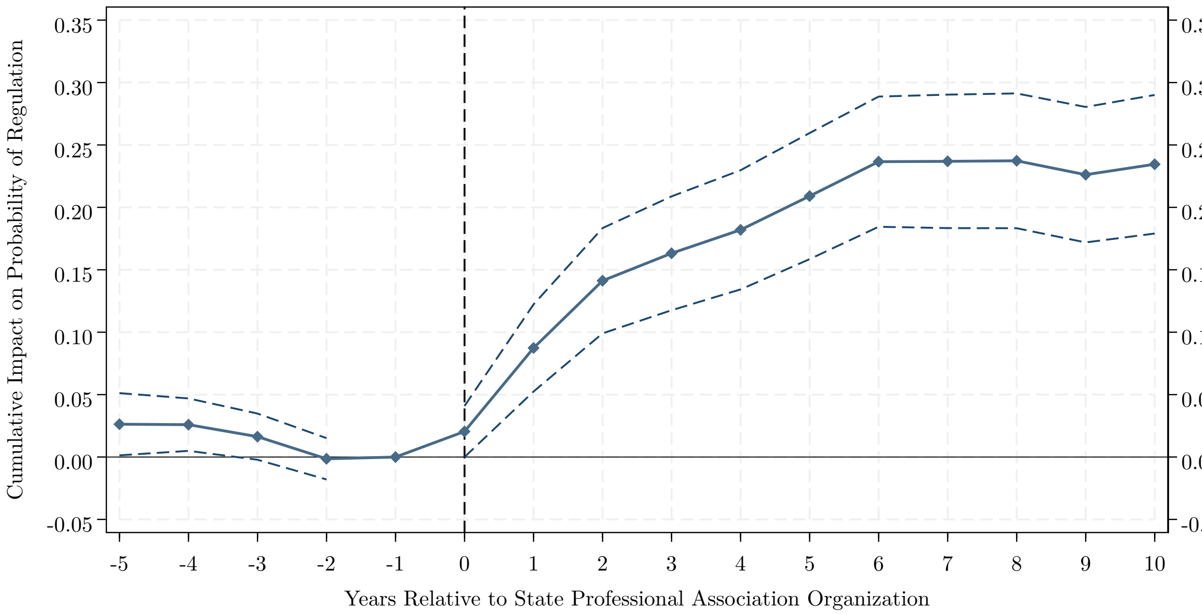
We estimate the effect of establishing an association on the cumulative probability of regulation using the following event study specification:

$$Regulated_{ist} = \alpha_{is} + \sum_{\tau=-10}^{20} \beta_\tau \times ProfAssociation_{ist}^{(\tau)} + \gamma' \mathbf{X}_{ist} + \delta_{it} + \epsilon_{ist}. \tag{15}$$

Here, $Regulated_{ist}$ is an indicator, equal to one if occupation i is licensed, certified, or registered in state s in calendar year t. We include occupation-by-state fixed effects α_{is}, occupation-by-year fixed effects δ_{it}, and a vector $\mathbf{X}_{ist}$, which includes other state and occupation-level correlates of regulation from Section 5.3. The indicator $ProfAssociation_{ist}^{(\tau)}$ denotes years in event time τ relative to the establishment of a professional association (with endpoints binned and normalizing $\beta_{-1} = 0$). The event study coefficients β_τ measure the cumulative effect in percentage points on the probability of a regulation being in force τ years before or after an association is organized.

2005. Further, we must measure hourly earnings at the 3-digit level, which we assign to all 6-digit sub-occupations in our data.

Figure 9: Event Study Estimates of the Relationship Between Professional Association Organization and Enactment of Initial Regulation

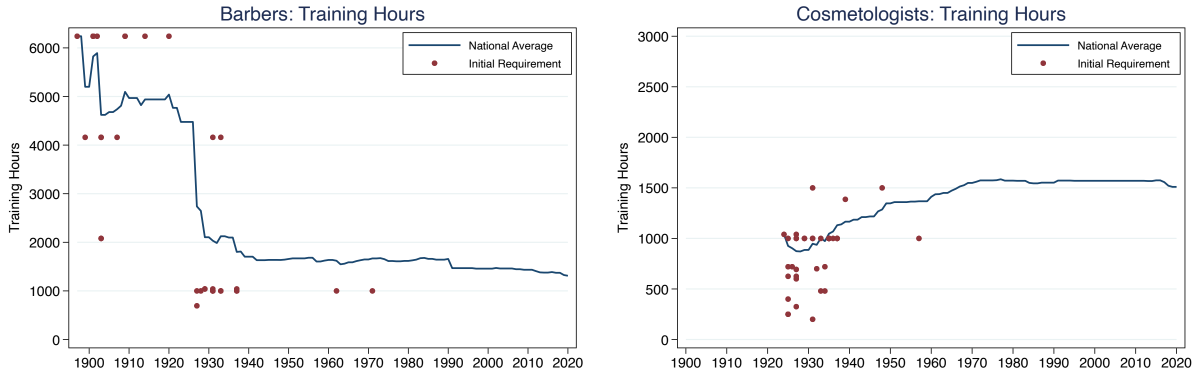


Notes: This figure plots event study estimates of the relationship between the organization of state professional associations and the enactment of initial licensing, certification, or registration legislation. We observe data on state-level professional associations for 9 occupations (a total of 411 associations, excluding those where we could not determine the timing of organization). The regression controls for the log of occupational employment, log state population, state urbanization rate, and a set of indicators for partisan control of the legislature and governorship. All coefficients are normalized relative to the year prior to professional association organization. Dashed lines denote 95% confidence interval estimates based on standard errors clustered by state-occupation cell.

Figure 9 plots our event study estimates, revealing a striking increase in the probability regulation immediately following the organization of a state professional association. We find that within five years, an occupation is about 20 percentage points more likely to be regulated in states with an association than in those without one. Beyond this, the contour of the estimates flattens out, which suggests that the associations we study were most effective at influencing legislation within a few years of their establishment. Importantly, we find no evidence that the probability of regulation was trending differently across states just before the formation of professional associations. This finding supports a causal interpretation of our estimates.

In Section A.4 we present an alternative design that simply counts the number of laws passed τ years relative to association organization. Like our event study estimates, Figure A4 shows a significant jump in the number of laws passed just after state professional associations were founded. In contrast to our results for state-level associations, however, we find little change in the probability of regulation around the formation of *national* associations, implying that local political organization is driving our findings. Finally, Figure A5 plots trends separately for each occupation. For at least six of the nine we study, we find a clear increase in the probability of regulation just after a state association is organized. Professional associations appear to have had the largest effect

Figure 10: Evolution of Minimum Training Hours for Barbers and Cosmetologists



Notes: This figure displays the national average of training hours required to obtain a license as a barber or cosmetologist over time. Solid markers indicate the qualifications specified in states' initial laws for these occupations.

on the timing of initial regulation for accountants, registered nurses, and pharmacists. By contrast, we find no relationship between regulation and the formation of civil engineering associations.[30]

6 The evolution of licensing qualifications

Thus far, our analysis has focused on the extensive margin of regulation. Even if they enact licensing policies, however, states often adopt very different training standards for the same occupations. In this section, we turn to the question of how these qualifications are determined and what drives their evolution over time. We focus on a case study of personal care services for which we currently have the most reliable data. Going forward, we plan to extend this analysis to around 50 occupations once we have validated our remaining legal data.

6.1 Case study of personal care services

Figure 10 plots a time series of average training hours (apprenticeship and/or trade school) required to obtain a license as a barber or cosmetologist, the two largest licensed occupations in the personal care services industry. They are different from many of the other occupations we consider in that qualifications for these licenses are usually specified in state statutes rather than administrative code. This feature allows us to trace the full history of minimum training hours back to the initial licensing laws for these occupations. Solid dots show the level of training specified in each state's first law.

Barbers were regulated earlier than cosmetologists, beginning in 1897 with Minnesota. Strikingly, early laws required over 6,000 hours (3 full-time years) of apprenticeship or attendance at a barber college to obtain a license. The national average fell dramatically in the 1920s, however, as

[30]Unlike most other associations we study, the American Society of Civil Engineers initially opposed licensing and actively worked to defeat proposed legislation in several states, which may explain this null effect (Spinden, 2015).

Table 8: Event History Estimates of Factors Influencing Amendments to Barber and Cosmetology Statutes

	Raise Requirement		Lower Requirement	
	Coef.	(Std. Err)	Coef.	(Std. Err)
Log Total State Population	0.300	(0.514)	-0.469	(0.794)
Unified Democratic	-0.299	(0.224)	0.153	(0.391)
Unified Republican	-0.160	(0.466)	0.391	(0.708)
Political Liberalism	-0.566***	(0.172)	0.008	(0.240)
Log Occupational Employment	-0.457	(0.428)	0.263	(0.761)
Log Hourly Earnings	0.389	(0.598)	0.561	(0.957)
Deviation from national average	-0.307***	(0.066)	0.044***	(0.014)
Log duration since last change	0.489***	(0.142)	0.142	(0.220)

Notes: This table reports the results of a discrete-time hazard model estimating the impact of various political and economic characteristics on the probability of raising or lowering training requirements for personal care services occupations. Significance levels based on robust standard errors are indicated by *** 1%, ** 5%, and * 10%.

later adopters of licensing set their training requirements at around 1,000 hours. During this period, states also began to regulate cosmetologists (at the time called beauticians). Notably, many states set significantly lower requirements for cosmetologists than those for barbers despite the similar nature of work these occupations perform. Over the next 30 years, training requirements for cosmetologists rose while those for barbers edged down; licensing qualifications for both occupations stabilized around 1,500 hours by the 1980s.

Event history estimates. Although it is not immediately apparent from the figure, we observe hundreds of individual amendments to barbering and cosmetology statutes in our data. To understand the evolution of training requirements, we therefore define two binary events – increases and decreases in training hours – and estimate a discrete-time event history model similar to Equation 13. Whereas in our previouss analyis, observations are removed from the risk set once the event has occurred, in thi scase states may in principle amend their laws at any time. Since in practice they are unlikely to do so, we further control for the log duration since the last amendment within state-occupation cells to capture the potential duration dependence of amendments. Further, we restrict our sample to 1960-2019, which allows us to include average hourly earnings in the regression.

The first two columns of Table 8 display our estimates for the determinants of choosing to raise training hours for barbers and cosmetologists. We find first that since 1960, more politically liberal states have been less likely to raise requirements, but we find no effect of the occupation's employment or earnings on the likelihood of passing an amendment. Second, the negative coefficient on the deviation of a state's requirement from the national average implies that states with training hours below the national average are more likely to raise their own requirements and vice versa. This confirms the intuition suggested by Figure 10 that qualifications tend to converge across states over time. Finally, we find evidence of duration dependence – states that recently changed their requirements are less likely to do so again over a short time horizon.

The last two columns repeat this analysis for amendments that lower training requirements. Here, we also find evidence of mean reversion. States with training hours above the national average were more likely to reduce their own. However, we do not see evidence that either state- or occupation-level variables predict these changes. Overall, our takeaway from this case study of the personal services industry is that the evolution of licensing qualification was driven primarily by convergence in regulatory requirements, both within and across these similar occupations.

7 Conclusion

Occupational licensing has become a pervasive feature of the labor market in the United States and other advanced economies. In this paper, we provide a theoretical framework to study the endogenous adoption of these laws and leverage vast new data to study the political and economic factors associated with their diffusion across U.S. states and occupations. We summarize our main empirical findings as follows:

1. Modern occupational licensing has its origins in the late 1800s. By 1950, about 6% of the workforce was employed in licensed occupations. This has risen to at least 20% today. About two-thirds of this growth reflects the enactment of new policies, although many recently regulated occupations did not exist (or had negligible employment shares) in 1950.

2. Occupations whose tasks plausibly pose some risk to consumers or the public are more likely to be regulated than less risky, complex, or interactive occupations. As new occupations emerge, those with a stronger public interest case were regulated earlier, on average. While licensing is not necessarily an optimal approach to regulation (as entry requirements may be too stringent or changed arbitrarily by licensing boards), our results suggest that socially beneficial policies are more likely to spread to all states than those with large social costs.

3. Large urban states, like California and New York, are much more likely to produce new licensing policies than small rural ones. They also tend to adopt policies originating in other states earlier and have accumulated a larger corpus of statutory and administrative law across all subjects. These findings suggest that these states have institutional environments that are more favorable to regulation in general.

4. Like other policies, licensing requirements tend to diffuse along geographic, demographic, and political lines across states. Since 1950, Democrat-led states were more likely than Republic-led states to license occupations, though licensing does not appear to be an especially partisan issue and has become somewhat less politically clustered in recent decades.

5. Before 1940, occupation-specific factors, including market size and the share of a state's neighbors that previously adopted a policy for the occupation, are strongly associated with the likelihood of regulation. Within occupations, greater exposure to labor market competition

from recent immigrants during the Age of Mass Migration further increases the hazard rate of licensing, which suggests that competition increases the demand for regulation among incumbent workers. After 1980, we find that state-level factors are more predictive of regulatory timing than occupation-specific factors are.

6. The probability of enacting a regulatory statute rises sharply after the establishment of a state professional association, which suggests that local political organization led to favorable regulatory outcomes for these groups and their members.

In 1952, a widely cited study of occupational licensing by the Council of State Governments opened with the following (possibly apocryphal) anecdote:

> Not long ago the Governor of a midwestern state was approached by representatives of a particular trade anxious to enlist the Governor's support in securing passage of legislation to license their trade.
>
> "Governor," the men said, "passage of this licensing act will ensure that only qualified people will practice this occupation; it will eliminate charlatans, incompetents or frauds; and it will thereby protect the safety and welfare of the people of this state."
>
> The Governor, from long experience, was somewhat skeptical. "Gentlemen," he asked, "are you concerned with advancing the health, safety and welfare of the people under the police powers of this state, or are you primarily interested in creating a monopoly situation to eliminate competition and raise prices?"
>
> The spokesman for the occupational group smiled and said, "Governor, we're interested in a little of each." — Council of State Governments (1952, pg. 1)

As this story illustrates, the potential conflict between public and private interests is central to the political economy of occupational licensing. Our results suggest the "little of each" characterization is an appropriate one. On the one hand, licensed occupations are highly selected with respect to their task content, which provides some support for the public interest view. On the other hand, local political organization and potentially greater exposure to labor market competition also influenced the adoption of regulatory requirements, consistent with theories that emphasize private incentives. All told, we see both motivations (in addition to broader economic trends) as important for understanding which occupations have become licensed, which states originated and adopted these policies, and why this institution has spread to such a broad swath of the labor market.

Our findings offer several potentially policy-relevant lessons. First, our theoretical framework highlights why groups seeking licensure may prefer a higher level of regulation (or more stringent qualifications) than is socially optimal, even when there is a plausible social-welfare case for some regulation. One way to mitigate these incentives may be to ensure that incumbent workers better internalize the costs of licensing – for instance, by limiting grandfathering exemptions. Another is by considering less restrictive policy alternatives such as state certification and registration, which, for many occupations, may address plausible market failures without overly burdensome restrictions on entry. Finally, like other forms of regulation, licensing policies are rarely repealed once established.

Yet, technology and tasks evolve over time, and so some occupations may have weaker cases for licensing now than when these statutes were initially adopted. Periodically reconsidering the need for licensure may identify areas for socially-beneficial policy reforms.

References

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, edited by Card, David, and Orley Ashenfelter Volume 4. 1043–1171, Elsevier.

Acemoglu, Daron, Nicolaj Søndergaard Mühlbach, and Andrew J. Scott. 2022. "The Rise of Age-Friendly Jobs." *Journal of the Economics of Ageing* 23, article 100416.

Acemoglu, Daron, and Pascual Restrepo. 2018. "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares and Employment." *American Economic Review* 108 (6): 1488–1542.

Althobaiti, Shahad, Ahmad Alabdulkareem, Judy Hanwen Shen, Iyad Rahwan, Morgan Frank, Esteban Moro, and Alex Rutherford. 2022. "Dictionary of Titles." Dataset, Harvard Dataverse, https://doi.org/10.7910/DVN/DQW8IP.

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.

Atalay, Enghin, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum. 2020. "The Evolution of Work in the United States." *American Economic Journal: Applied Economics* 12 (2): 1–34.

Autor, David, Caroline Chin, Anna Salomons, and Bryan Seegmiller. 2024. "New Frontiers: The Origins and Content of New Work, 1940–2018." *Quarterly Journal of Economics* 139 (3): 1399–1465.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–1597.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *American Economic Review* 96 (2): 189–194.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118 (4): 1279–1333.

Barde, Robert, Susan B. Carter, and Richard Sutch. 2006. "Immigrants, by Continent of Last Residence: 1820–1997." In *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*, edited by Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, Cambridge University Press.

Baybeck, Brady, William D. Berry, and David A. Siegel. 2011. "A Strategic Theory of Policy Diffusion via Intergovernmental Competition." *Journal of Politics* 73 (1): 232–247.

Becker, Gary S. 1983. "A Theory of Competition Among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98 (3): 371–400.

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.

Blair, Peter Q., and Bobby W. Chung. Forthcoming. "Job Market Signaling Through Occupational Licensing." *Review of Economics and Statistics.*

Blair, Peter Q., and Mischa Fisher. 2022. "Does Occupational Licensing Reduce Value Creation on Digital Platforms." *NBER Working Paper No. 30388.*

Bowen, Daniel C., and Zachary Greene. 2014. "Should We Measure Professionalism with an Index? A Note on Theory and Practice in State Legislative Professionalism Research." *State Politics & Policy Quarterly* 14 (3): 277–296.

Brinegar, Pam. 2006. "Professional and Occupational Regulation." In *The Book of the States*, The Council of State Governments.

Burnham, W. Dean. 1992. "Partisan Division of American State Governments, 1834-1985." Dataset, Inter-University Consortium for Political and Social Research. http://doi.org/10.3886/ICPSR00016.v1.

Caines, Colin, Florian Hoffmann, and Gueorgui Kambourov. 2017. "Complex-Task Biased Technological Change and the Labor Market." *Review of Economic Dynamics* 25: 298–319.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19 (1): 22–64.

Carollo, Nicholas A. 2024. "Historical Data on Occupational Regulation in the United States." *Working Paper.*

Carr, Ada M. 1906. "State Societies: Their Work and Opportunities." In *Proceedings of the Third Annual Convention of the Maryland State Association of Graduate Nurses*, 20–27, J.H. Furst Company.

Caughey, Devin, and Christopher Warshaw. 2014. "The Dynamics of State Policy Liberalism, 1936-2014." *American Journal of Political Science* 60 (4): 899–913.

Caughey, Devin, Christopher Warshaw, and Yiqing Xu. 2017. "Incremental Democracy: The Policy Effects of Partisan Control of State Government." *The Journal of Politics* 79 (4): 1342–1358.

CISER Data & Reproduction Archive. 2002. "Dictionary of Occupational Titles (DOT)." Dataset, Cornell University, https://doi.org/10.6077/j5/3iwlyl.

Collins, William J. 2003. "The Political Economy of State-Level Fair Employment Laws, 1940–64." *Explorations in Economic History* 40 (1): 24–51.

Council of State Governments. 1952. *Occupational Licensing Legislation in the United States.* Council of State Governments.

Cross, J. N. 1998. "The First Wisconsin Accountancy Bill: An Historical Perspective." *Accounting Historians Journal* 25 (2): 113–128.

Cunningham, Evan. 2019. "Professional Certifications and Occupational Licenses: Evidence from the Current Population Survey." *Monthly Labor Review*, U.S. Bureau of Labor Statistics.

DellaVigna, Stefano, and Woojin Kim. 2022. "Policy Diffusion and Polarization Across U.S. States." *NBER Working Paper No. 30142.*

Deming, David J. 2017. "The Growing Importance of Social Skills in the Labor Market." *Quarterly Journal of Economics* 132 (4): 1593–1640.

Deyo, Darwyyn, and Alicia Plemmons. 2022. "Have License, Will Travel: Measuring the Effects of Universal Licensing Recognition on Mobility." *Economic Letters* 219.

Farronato, Chiara, Andrey Fradkin, Bradley J. Larsen, and Erik Brynjolfsson. 2024. "Consumer Protection in an Online World: An Analysis of Occupational Licensing." *American Economic Journal: Applied Economics* 16 (3): 549–579.

Federman, Maya N., David E. Harrington, and Kathy J. Krynski. 2006. "The Impact of State Licensing Regulations on Low-Skilled Immigrants: The Case of Vietnamese Manicurists." *American Economic Review* 96 (2): 237–241.

Fishback, Price V. 1992. *Soft Coal, Hard Choices: The Economic Welfare of Bituminous Coal Miners, 1890–1930*. New York: Oxford University Press.

Fishback, Price V., and Shawn Everett Kantor. 1998. "The Adoption of Workers' Compensation in the United States, 1900–1930." *Journal of Law & Economics* 41 (2): 305–342.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as Data." *Journal of Economic Literature* 57 (3): 535–574.

Gittleman, Maury, Mark A. Klee, and Morris M. Kleiner. 2018. "Analyzing the Labor Market Outcomes of Occupational Licensing." *Industrial Relations* 57 (1): 57–100.

Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race Between Education and Technology.* Harvard University Press.

Graddy, Elizabeth. 1991a. "Interest Groups or the Public Interest – Why Do We Regulate Health Occupations?" *Journal of Health Politics, Policy and Law* 16 (1): 25–49.

Graddy, Elizabeth. 1991b. "Toward a General Theory of Occupational Regulation." *Social Science Quarterly* 72 (4): 676–695.

Grossback, Lawrence J., Sean Nicholson-Crotty, and David A. M. Peterson. 2004. "Ideology and Learning in Policy Diffusion." *American Politics Research* 32 (5): 521–545.

Grossman, Gene M., and Elhanan Helpman. 1994. "Protection for Sale." *American Economic Review* 84 (4): 833–850.

Grossman, Gene M., and Elhanan Helpman. 1996. "Electoral Competition and Special Interest Politics." *Review of Economic Studies* 63 (2): 265–286.

Hall, Richard L., and Alan V. Deardorff. 2006. "Lobbying as Legislative Subsidy." *American Political Science Review* 100 (1): 69–84.

Han, Suyoun, and Morris M. Kleiner. 2021. "Analyzing the Influence of Occupational Licensing Duration and Grandfathering on Wage Determination." *Industrial Relations* 60 (2): 147–187.

Hirsch, Barry T., David A. Macpherson, and Wayne G. Vroman. 2001. "Estimates of Union Density by State." *Monthly Labor Review*, U.S. Bureau of Labor Statistics.

Johnson, Janna E., and Morris M. Kleiner. 2020. "Is Occupational Licensing a Barrier to Interstate Migration?" *American Economic Journal: Economic Policy* 12 (3): 347–373.

Karch, Andrew. 2007. *Democratic Laboratories: Policy Diffusion Among the American States.* University of Michigan Press.

Kett, Joseph F. 1968. *The Formation of the American Medical Profession: The Role of Institutions, 1780–1860.* Yale University Press.

Kim, Gueyon, Cassandra Merritt, and Giovanni Peri. 2024. "Measuring and Predicting New Work in the United States: The Role of Local Factors and Global Shocks." *NBER Working Paper No. 32526.*

Klarner, Carl. 2003. "The Measurement of the Partisan Balance of State Government." *State Politics & Policy Quarterly* 3 (3): 309–319.

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 Robert T. Kudrle. 2000. "Does Regulation Affect Economic Outcomes? The Case of Dentistry." *Journal of Law and Economics* 43 (2): 547–582.

Kleiner, Morris M., Allison Marier, Kyoung Won Park, and Coady Wing. 2016. "Relaxing Occupational Licensing Requirements: Analyzing Wages and Prices for a Medical Service." *Journal of Law & Economics* 59 (2): 261–291.

Kleiner, Morris M., and Evan J. Soltas. 2023. "A Welfare Analysis of Occupational Licensing in U.S. States." *Review of Economic Studies* 90 (5): 2481–2516.

Kleiner, Morris M., and Ming Xu. Forthcoming. "Occupational Licensing and Labor Market Fluidity." *Journal of Labor Economics.*

Kogan, Leonid, Dimitris Papanikolaou, Lawrence D.W. Schmidt, and Bryan Seegmiller. 2023. "Technology and Labor Displacement: Evidence from Linking Patents with Worker-Level Data." *NBER Working Paper No. 31846.*

Larsen, Bradley, Ziao Ju, Adam Kapor, and Chuan Yu. 2020. "The Effect of Occupational Licensing Stringency on the Teacher Quality Distribution." *NBER Working Paper No. 28158.*

Law, Marc T., and Sukkoo Kim. 2005. "Specialization and Regulation: The Rise of Professionals and the Emergence of Occupational Licensing Regulation." *Journal of Economic History* 65 (3): 723–756.

Law, Marc T., and Patrick A. McLaughlin. 2022. "Industry Size and Regulation: Evidence from US States." *Public Choice* 192 (1-2): 1–27.

Lin, Jeffrey. 2011. "Technological Adaptation, Cities, and New Work." *Review of Economics and Statistics* 93 (2): 554–574.

Locke, Elijah. 2024. "Ethnic-Occupational Niches: Evidence from the Age of Mass Migration." *Working Paper.*

Makse, Todd, and Craig Volden. 2011. "The Role of Policy Attributes in the Diffusion of Innovations." *The Journal of Politics* 73 (1): 108–124.

Mallinson, Daniel J. 2020. "The Spread of Policy Diffusion Studies: A Systematic Review and Meta-Analysis, 1990–2018." *Working Paper.*

Mallinson, Daniel J. 2021. "Who Are Your Neighbors? The Role of Ideology and Decline of Geographic Proximity in the Diffusion of Policy Innovations." *Policy Studies Journal* 49 (1): 67–88.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. 2021. "IPUMS National Historical Geographic Information System: Version 16.0." Dataset, Minneapolis, MN: IPUMS.

Maurizi, Alex. 1974. "Occupational Licensing and the Public Interest." *Journal of Political Economy* 82 (2): 399–413.

Medici, Carlo. 2024. "Closing Ranks: Organized Labor and Immigration." *Working Paper.*

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding. 2019. "Task Specialization in U.S. Cities from 1880 to 2000." *Journal of the European Economic Association* 17 (3): 754–798.

Mitra, Devashish. 1999. "Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination." *American Economic Review* 89 (5): 1116–1134.

Mooney, Christopher Z. 2020. *The Study of US State Policy Diffusion: What Hath Walker Wrought?.* Cambridge University Press.

Mulligan, Casey B., and Andrei Shleifer. 2005. "The Extent of the Market and the Supply of Regulation." *Quarterly Journal of Economics* 120 (4): 1445–1473.

Mühlbach, Nicolaj Søndergaard. 2022. "occ2vec: A Principal Approach to Representing Occupations using Natural Language Processing." *arXiv:2111.02528.*

Nedelkoska, Ljubica, Gadgin Matha Shreyas, James McNerney, Andre Assumpcao, Dario Diodato, and Frank Neffke. 2023. "Eight Decades of Changes in Occupational Tasks, Computerization and the Gender Pay Gap." *Harvard University Center for International Development Working Paper.*

Neumark, David, and William L. Wascher. 2008. "The Political Economy of Minimum Wages." In *Minimum Wages*, The MIT Press.

Nicholson-Crotty, Sean. 2009. "The Politics of Diffusion: Public Policy in the American States." *The Journal of Politics* 71 (1): 192–205.

Pagliero, Mario. 2011. "What Is the Objective of Professional Licensing? Evidence from the US Market for Lawyers." *International Journal of Industrial Organization* 29 (4): 473–483.

Peltzman, Sam. 1976. "Toward a More General Theory of Regulation." *Journal of Law & Economics* 19 (2): 211–240.

Pennington, Jeffrey, Richard Socher, and Christopher Manning. 2014. "GloVe: Global Vectors for Word Representation." In *The 2014 Conference on Empirical Methods in Natural Language Processing*, 1532–1543, Proceedings of the Conference, Association for Computational Linguistics.

Pigou, Arthur. 1938. *The Economics of Welfare*. London: Macmillan.

Ruggles, Steven, Sarah Flood, Matthew Sobek et al. 2024a. "IPUMS USA: Version 15.0." Dataset, Minneapolis, MN: IPUMS, https://doi.org/10.18128/D010.V15.0.

Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren. 2024b. "IPUMS Ancestry Full Count Restricted Data: Version 4.0R." Dataset, Minneapolis, MN: IPUMS, https://doi.org/10.18128/D014.V4.0R.

Schmick, Ethan. 2018. "Collective Action and the Origins of the American Labor Movement." *Journal of Economic History* 78 (3): 744–784.

Shapiro, Carl. 1986. "Investment, Moral Hazard, and Occupational Licensing." *Review of Economic Studies* 53 (5): 843–862.

Shleifer, Andrei. 2005. "Understanding Regulation." *European Financial Management* 11 (4): 439–451.

Smith, Janet Kilholm. 1982. "Production of Licensing Legislation: An Economic Analysis of Interstate Differences." *Journal of Legal Studies* 11 (1): 117–137.

Sobel, Russell S. 1999. "Theory and Evidence on the Political Economy of the Minimum Wage." *Journal of Political Economy* 107 (4): 761–785.

Spinden, Paul M. 2015. "The Enigma of Engineering's Industrial Exemption to Licensure: The Exception That Swallowed a Profession." *UMKC Law Review* 83 (3): 637–686.

Stigler, George J. 1971. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2 (1): 3–21.

U.S. Department of Labor, Employment and Training Administration. 1939. *Dictionary of Occupational Titles: First Edition.* U.S. Government Printing Office.

U.S. Department of Labor, Employment and Training Administration. 1977. *Dictionary of Occupational Titles: Fourth Edition.* U.S. Government Printing Office.

Walker, Jack L. 1969. "The Diffusion of Innovations Among the American States." *American Political Science Review* 63 (3): 880–899.

Yamaguchi, Shintaro. 2012. "Tasks and Heterogeneous Human Capital." *Journal of Labor Economics* 30 (1): 1–53.

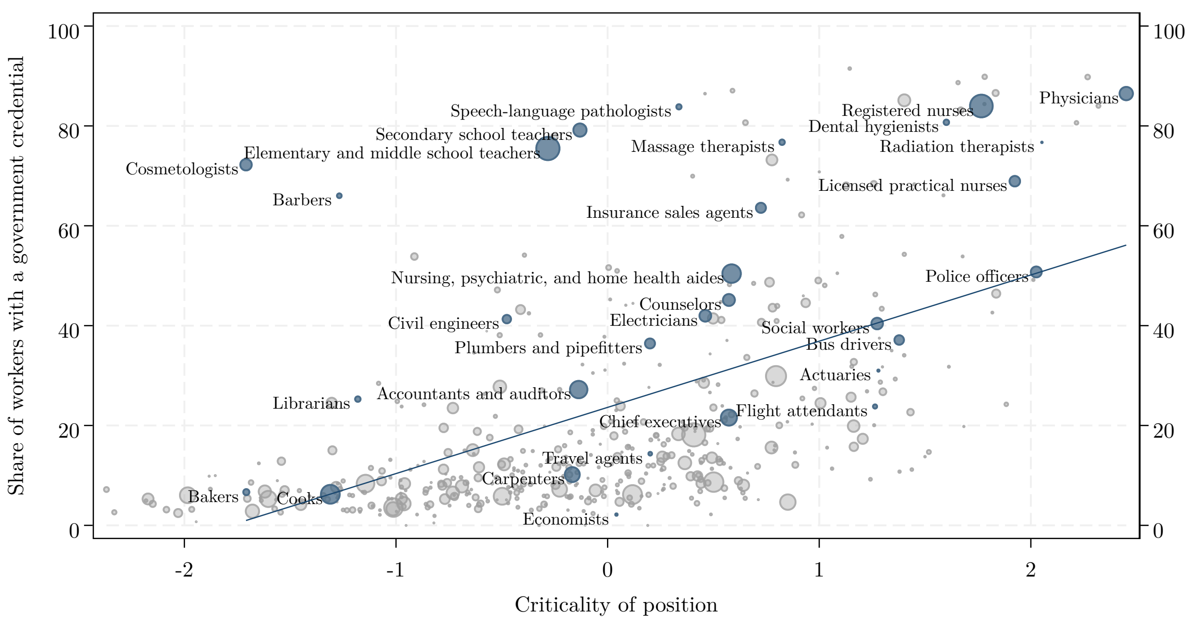
A Additional results and robustness checks

A.1 Occupational regulation and task content

This section presents additional results and robustness checks related to our analysis of task content and the diffusion of licensing across occupations.

Survey-based measures. We first show that our results are robust to an alternative measure of occupational licensing. We measure license attainment at the individual level using the 2015 to 2019 Current Population Survey (CPS). Following Kleiner and Soltas (2023), we define licenses using the "profcert" and "statecert" variables. The "profcert" variable indicates an affirmative response to the following question: "Do you have a currently active professional certification or state or industry license?" The "statecert" variable indicates an affirmative response to the follow-up question: "Were any of your certifications or licenses issued by the federal, state, or local government?" Workers are considered licensed if they answer yes to both of these questions.

Figure A1: Criticality of Position and License Attainment (CPS 2015–2019)



Notes: This figure plots the share of workers who report holding an occupational license in the CPS against our criticality of position index. Marker sizes are proportional to total employment.

Pooling across all years of data, we first compute the share of licensed workers nationally in each of the 483 occupation categories used by the CPS. Our sample is limited to employed civilian adults, excluding unpaid family workers and individuals whose occupation or licensing status is imputed. Because the CPS uses a less detailed occupational classification than O*NET, we aggregate our predicted occupational attributes to CPS codes using employment weights for six-digit occupations from the Occupational Employment and Wage Statistics (OEWS) survey.

Table A1: Occupational Attributes and License Attainment

	Dependent Vaiable: Share of Licensed Workers				
	(1)	(2)	(3)	(4)	(5)
Criticality of position	13.259***			5.040**	4.078*
	(2.024)			(2.156)	(2.445)
Task complexity		9.074***		6.482***	6.391***
		(2.043)		(1.735)	(1.791)
Interpersonal interaction			13.949***	11.536***	4.568*
			(2.281)	(2.174)	(2.485)
Observations	483	483	483	483	483
R-squared	0.313	0.167	0.294	0.481	0.711
Major group FE					✓

Notes: This table reports the results of regressing self-reported licensing rates on measures summarizing occupational task content. Each observation is a CPS occupation category. Here, we use predicted attributes based on O*NET task descriptions, which we aggregate to CPS occupations using national employment share weights from the 2015 OEWS survey. Occupation group fixed effects are defined at the 2-digit level of aggregation. Significance levels based on robust standard errors are indicated by *** 1%; ** 5%; and * 10%.

Figure A1 plots the relationship between our "criticality of position" index and licensing rates across CPS occupations. As was the case with our policy-based measures, we find that licensing is more prevalent among occupations that rank highly on this attribute. Importantly, this is true even after accounting for the fact that – unlike our policy measures – actual licensing rates in the CPS generally fall somewhere between zero and one. Column one of Table A1 reports the regression corresponding to the solid blue line in the figure.[1] We find that each standard deviation increase in criticality is associated with a 13% increase in the share of licensed workers in an occupation. Columns two and three show that licensing rises at about the same rate with increases in interpersonal interaction and task complexity.

In column four, we include all three attributes in the same regression and find that they remain individually and jointly statistically significant. As with our main results, this suggests that each index captures a separate margin of variation that is likely to be important for the prevalence of licensing. Notably, the R^2 of this regression is 0.48, which implies that these three measures alone explain nearly half of the cross-sectional variation in license attainment across occupations in the CPS. Moreover, the magnitude of the estimates in columns four and five are generally similar to those reported in Table 1, confirming that our main results are largely unaffected by the choices we make when assigning policies to statistical occupation codes, or by measurement error resulting from policies that we do not observe in our data.

Choice of regulatory methods. As certification and registration policies are fairly uncommon, we group them together with licensing requirements in our main analysis. Here, we ask whether

[1] Here, we depart from our main policy-based regressions in that we weight observations by occupational employment, since licensing rates for smaller occupations may be subject to greater sampling error.

Table A2: Occupational Attributes and Choice of Regulatory Method

	Regulated in 2020			Regulated in 1980			Regulated in 1940		
	Licensed	Certified	Registered	Licensed	Certified	Registered	Licensed	Certified	Registered
Criticality of position	0.63***	-0.19	0.30***	0.56***	0.09	0.46***	0.35***	1.54***	2.38***
	(0.02)	(0.14)	(0.07)	(0.03)	(0.17)	(0.07)	(0.04)	(0.16)	(0.23)
Task complexity	0.76***	0.80***	-0.04	0.50***	0.71***	-0.18***	0.42***	0.25***	0.15***
	(0.02)	(0.09)	(0.07)	(0.02)	(0.09)	(0.05)	(0.02)	(0.06)	(0.04)
Interpersonal interaction	0.62***	0.93***	0.58***	0.30***	0.90***	-0.34**	0.51***	-0.07	-1.25***
	(0.02)	(0.13)	(0.07)	(0.02)	(0.17)	(0.16)	(0.03)	(0.09)	(0.19)
Occupations	180	27	20	159	21	17	109	30	7
Policies	6,173	92	185	5,004	54	103	3,061	255	96
Total observations		36,960			35,472			28,128	
Task measures		O*NET 2017			DOT 1977			DOT 1939	

Notes: This table reports the results of a set of multinomial logistic regressions for states' choice of regulatory method relative to choosing not to regulate an occupation. All specifications include state fixed effects. Significance levels based on bootstrapped standard errors indicated by *** 1%; ** 5%; and * 10%.

occupational attributes also affect the *type* of policy states choose to enact. Table A2 presents a set of multinomial logit estimates that model the probability that states have enacted a licensing, certification, or registration law for each occupation, relative to a baseline of no regulation. As in our main analysis, we organize our data into three state-by-occupation cross-sections and include a set of state fixed effects to control for potential differences in each state's general propensity to regulate.

We find that as of 2020, criticality is associated with a higher probability of licensing and registration but has no effect on the adoption voluntary certification. Task complexity, on the other hand, increases the probability of licensing and certification, but has no effect on registration. Put differently, licensing tends to be adopted for occupations that are both risky and complex; voluntary certification for occupations that are complex, but not risky; and mandatory registration for occupations that are risky but not complex. We see these results as consistent with what one might expect given the signaling and screening functions of these credentials (Shapiro, 1986). Our estimates as of 1980 are qualitatively similar to those for 2020, though we find that the 1940 estimates differ more. Most notably, criticality has a larger impact on the probability of certification and registration than it does on licensing, possibly because our approach assigns high criticality rankings to the most commonly certified and registered occupations in 1940 (registered nurses and insurance sales agents), which were eventually licensed in all states.

Robustness. Table A3 presents three additional robustness checks on our cross-sectional results. Columns one and two exclude healthcare occupations from the regression sample (SOC groups 29 and 31). Estimates are generally similar to those reported in Table 1, confirming that our results are not driven by the overrepresentation of healthcare practitioners and support personnel among licensed occupations, but hold more generally. Second, our main results do not weight state-occupation cells by their size, as policies that cover more workers are not necessarily more informative for understanding regulators' choices in our setting. It is reassuring, however, that our conclusions are not especially sensitive to this choice. Columns three and four weight cells by their

Table A3: Robustness: Occupational Attributes and Probability of Regulation

	Regulated in 2020 with O*NET 2017 Task Measures					
	(1)	(2)	(3)	(4)	(5)	(6)
Criticality of position	0.056***	0.057***	0.091**	0.107**	0.043**	0.059**
	(0.016)	(0.020)	(0.044)	(0.043)	(0.020)	(0.023)
Task complexity	0.059***	0.029	0.063*	-0.001	0.089***	0.038*
	(0.013)	(0.023)	(0.036)	(0.037)	(0.017)	(0.022)
Interpersonal interaction	0.052***	0.074***	0.117***	0.012	0.097***	0.102***
	(0.017)	(0.026)	(0.041)	(0.048)	(0.016)	(0.024)
Co-workers health/safety					0.038**	-0.007
					(0.016)	(0.020)
Total observations	33,312	33,312	34,582	34,582	36,960	36,960
Occupations	694	694	770	770	770	770
R-squared	0.100	0.248	0.225	0.442	0.225	0.351
State FE	✓	✓	✓	✓	✓	✓
Major group FE		✓		✓		✓
Specification	Exc. Healthcare		Coverage-weighted		Main	

Notes: This table reports the results of regressing an indicator for whether a state-occupation cell is subject to any occupational regulation on our predicted occupational attributes. Major group fixed effects are defined at the level of 2-digit occupations. Significance levels based on standard errors clustered by state and occupation are indicated by *** 1%; ** 5%; and * 10%.

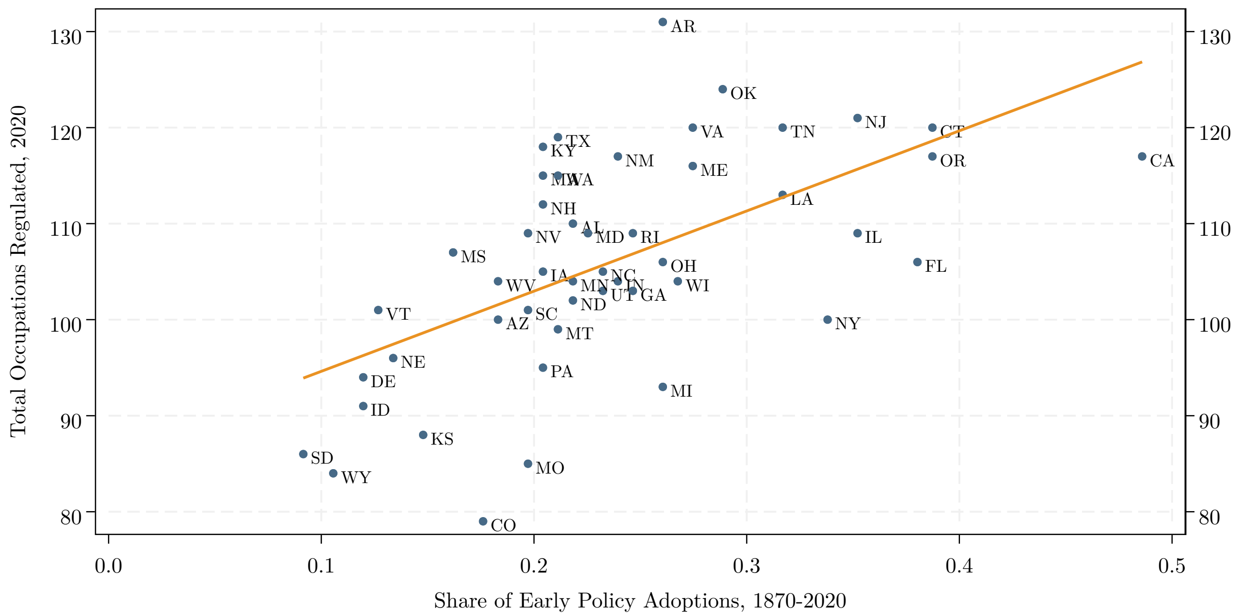
total employment as of 2019. The most significant difference compared with our preferred estimates is that within major occupation groups, criticality matters more, and interpersonal interaction less, when placing more weight on larger occupations and states.

Finally, another potential rationale for licensing could be the welfare of workers themselves. In the early 1900s, for instance, several states adopted licensing requirements for bituminous coal mine foremen and examiners, with worker safety as the stated justification (Fishback, 1992). Columns five and six include an additional attribute constructed from an O*NET question that measures workers' responsibility for the health and safety of their coworkers. While highly correlated with our criticality index, column five shows that both are associated with a higher likelihood of licensing. However, column six suggests that this primarily reflects differences across major occupation groups. Within similar groups of occupations, criticality has a point estimate close to our main results, while the coefficient on co-worker's health and safety falls to zero. Thus, while we do not rule out worker safety as another factor that may be related to licensing, it appears secondary to consumer welfare, the standard justification provided for the licensing of most occupations.

A.2 Innovation and diffusion

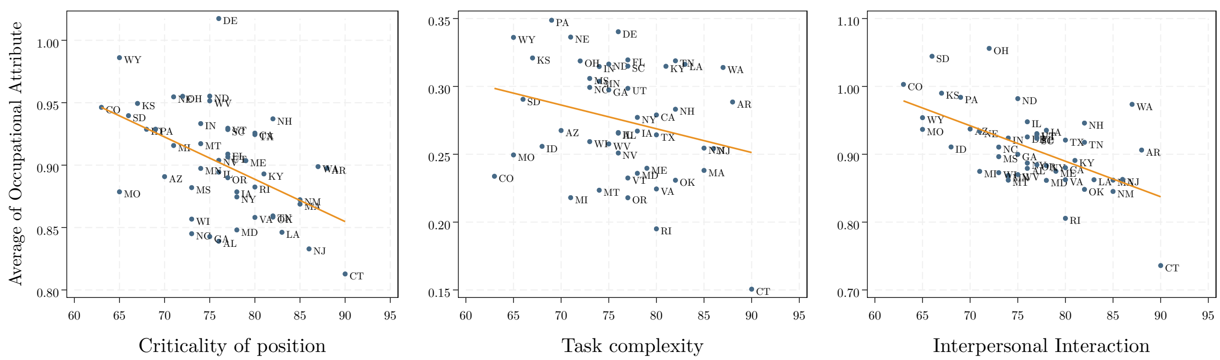
In the main text, we focus on the average timing of policy adoption across states, identifying jurisdictions that are consistently early adopters of occupational regulation. Figure A2 shows that early adopters also accumulate a larger stock of regulation, as measured by the total number of

Figure A2: Early Policy Adopters Accumulate More Regulation



Notes: This figure displays the relationship between the total number of occupations in our sample that each state licenses, certifies, or registers in 2020 against the fraction of policies enacted since 1870 for which they were one of the first 10 early adopters.

Figure A3: Average of Occupational Attributes by Number of Regulated Occupations



Notes: This figure plots an unweighted average of occupational attributes by state (y-axis) against the total number of six-digit occupations regulated by each state as of 2020 (x-axis).

occupations in our data that they regulate as of 2020.

Consistent with our analysis of task content, Figure A3 shows that as the number of regulated six-digit occupations in a state increases, the average criticality, complexity, and interpersonal interaction scores of those occupations declines. That is, the marginal occupation regulated by states like Connecticut, which regulates many occupations, appears to have a weaker public interest rationale than the marginal occupation regulated by states like Colorado, which regulates few. Together with the finding we report in the main text that occupational regulation is correlated with a state's total

Table A4: **Robustness of Clustering Statistics to Alternative Definitions**

	Initial Regulation			Initial Licensure		
A. Geographic proximity	In First 5 Adopters	In First 10 Adopters	In First 16 Adopters	In First 5 Adopter	In First 10 Adopters	In First 16 Adopters
1870-1949	0.05	0.05	0.06	0.04	0.03	0.05
1950-1989	0.08	0.05	0.05	0.05	0.05	0.05
1990-2020	0.04	0.03	0.04	0.04	0.05	0.04
B. Demographic similarity						
1870-1949	0.12	0.08	0.07	0.13	0.08	0.06
1950-1989	0.08	0.05	0.06	0.08	0.06	0.05
1990-2020	0.03	0.03	0.02	0.03	0.03	0.03
C. Industry composition						
1870-1949	0.05	0.03	0.02	0.06	0.03	0.03
1950-1989	0.00	0.01	0.03	0.02	0.02	0.03
1990-2020	-0.01	0.02	0.01	0.00	0.04	0.03
D. State policy liberalism						
1870-1949	0.08	0.08	0.08	0.08	0.09	0.07
1950-1989	0.06	0.03	0.04	0.04	0.04	0.04
1990-2020	0.01	0.03	0.02	0.01	0.04	0.02
E. Republican presidential vote share						
1870-1949	0.02	0.01	0.00	0.01	0.00	0.01
1950-1989	0.02	-0.01	0.02	0.02	-0.00	0.00
1990-2020	0.00	-0.00	-0.01	0.01	-0.00	-0.01
Null 25th/75th percentile	[-0.03, 0.03]	[-0.02, 0.03]	[-0.02, 0.02]	[-0.03, 0.03]	[-0.02, 0.03]	[-0.02, 0.02]

Notes: This table shows the average of Geary's C-statistic for early policy enactments using alternative distance metrics and adoption thresholds. Values of C greater than zero indicate that the first n states to enact a policy were clustered geographically, or were similar with respect to their demographic characteristics, legislative ideology, or industry composition. To measure the ideology of state legislators, we use the state policy liberalism index constructed by Caughey and Warshaw (2014), which captures the enactment of legislation generally associated with the political left. The 25th to 75th percentiles of the C-statistic under the null of random policy adoption are reported at the bottom of the table.

volume of legal text, this suggests that some states have institutional environments that are more favorable to regulation in general. It also suggests that the marginal regulations adopted by these states may be less socially beneficial than inframarginal policies that are common to all states.

Robustness. In our main analysis, we rank states according to the order in which they initially regulated each occupation. Results on the characteristics of early adopters are qualitatively similar if we restrict our attention to licensing policies or if we limit the sample to a consistent set of occupations that were eventually regulated by 10 or more states. For brevity, we omit these versions of our exhibits. Table A4 reports robustness exercises for our policy clustering analysis using different regulatory definitions and adoption thresholds. Like our main results, the results for these exercises show that early adopters of licensing or regulation are highly clustered geographically, demographically, and politically, though by most measures, policy clustering has declined over time. In Panel E, we report an alternative measure of state ideology – the vote share for the Republican candidate in the most recent presidential election (DellaVigna and Kim, 2022). We find less evidence

Table A5: **First Stage Estimates for Immigrant Labor Market Competition (1870–1940)**

	Immigration flows since t-5			Immigration flows since t-10		
Panel A: Estimates prior to regulation	(1)	(2)	(3)	(4)	(5)	(6)
Predicted competition from new immigrants	0.370***	0.246***	0.250***	0.481***	0.359***	0.363***
	(0.072)	(0.063)	(0.065)	(0.081)	(0.069)	(0.071)
Total sample observations	173,191	173,191	173,191	158,764	158,764	158,764
F-statistic	26.70	28.86	29.47	34.85	37.14	31.52
Panel B: Estimates after regulation						
Predicted competition from new immigrants	0.132***	0.094***	0.088***	0.262***	0.212***	0.208***
	(0.035)	(0.034)	(0.032)	(0.060)	(0.047)	(0.046)
Total sample observations	35,575	35,575	35,575	35,043	35,043	35,043
F-statistic	13.91	8.95	9.02	18.82	18.10	15.54
Occupation × year FE	✓	✓	✓	✓	✓	✓
Demographic and political controls		✓	✓		✓	✓
State FE			✓			✓

Notes: This table reports first-stage estimates for our measure of immigrant labor market competition. The dependent variable is the change in the number of immigrants in a state-occupation cell since $t-5$ or $t-10$, normalized by the lagged number of native workers in the state-occupation cell. The independent variable is our shift-share instrument, which is the predicted change in immigrants per native worker. Sample excludes federally regulated occupations and state occupations with no worker. Additionally, we exclude state-occupation cells after they have become regulated, and licensing may affect the entry of immigrants. Standard errors are clustered at the occupation level, and significance levels are indicated by *** 1%, ** 5%, and * 10%.

of policy clustering along this dimension than our measure of state policy liberalism, which we think better captures the revealed preferences of state policymakers.

A.3 Immigration and competition

In the main text, we construct a shift-share instrument to measure labor market competition between native workers and recent immigrants. Table A5 confirms the relevance of our competition measure. Here, we estimate a series of linear regressions of observed changes in the number of immigrants per native worker in a state-occupation cell on predicted immigration flows.

Note that observed changes in the number of immigrants per native worker in a state-occupation cell are constructed by linearly interpolating data between census years, as we observe policy changes annually. Since licensing might affect both total employment and the share of immigrants who enter an occupation, in Panel A of the table we restrict the sample for our first-stage estimates to years before regulation was adopted (the same data structure we use in our hazard model). However, both the linear interpolation and policy changes occurring between time t and $t+10$ could weaken the correlation between predicted and actual migration flows. Indeed, Panel B shows a significantly weaker relationship between predicted and actual flows in years *after* a regulatory policy has been enacted, suggesting that licensing was effective at deterring the entry of immigrants. This measurement issue is less concerning for our predicted competition measure than actual flows, since we observe flows annually and the share component of the instrument is defined excluding state s. For this reason, we focus on reduced form rather than two-stage estimates in our main analysis.

Table A6: Robustness Checks for Event History Estimates of Immigration Competition (Licensing, Certification, and Registration Laws Enacted 1870–1940)

	Immigration flows since t-5				Immigration flows since t-10			
	Regulated		Licensed		Regulated		Licensed	
Occupation characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted competition from new immigrants	0.509**	0.404	0.652***	0.643**	0.099	0.054	0.153**	0.149**
	(0.223)	(0.251)	(0.247)	(0.251)	(0.063)	(0.084)	(0.061)	(0.062)
Log total occupational employment	0.356***	0.361***	0.394***	0.433***	0.373***	0.368***	0.410***	0.447***
	(0.074)	(0.081)	(0.079)	(0.089)	(0.079)	(0.087)	(0.084)	(0.092)
Neighboring states regulating (%)	0.792***	0.835***	0.796***	0.807***	0.827***	0.882***	0.846***	0.862***
	(0.157)	(0.186)	(0.164)	(0.195)	(0.153)	(0.187)	(0.157)	(0.196)
Total sample observations	20,297	20,297	17,354	17,354	19,116	19,116	16,351	16,351
Number of events	1,362	1,362	1,092	1,092	1,301	1,301	1,035	1,035
Number of occupations	76	76	71	71	76	76	71	71
Demographic and political controls	✓	✓	✓	✓	✓	✓	✓	✓
Lagged regulation variables	✓	✓	✓	✓	✓	✓	✓	✓
State fixed effects		✓		✓		✓		✓

Notes: This table reports the results of a discrete-time hazard model estimating the impact of occupational characteristics on the probability of adopting an occupational regulation or licensing law. Here, we show untransformed logistic estimates. Sample excludes federally regulated occupations and state-occupation cells with no workers. Standard errors are clustered at the occupation level, and significance levels are indicated by *** 1%, ** 5%, and * 10%.

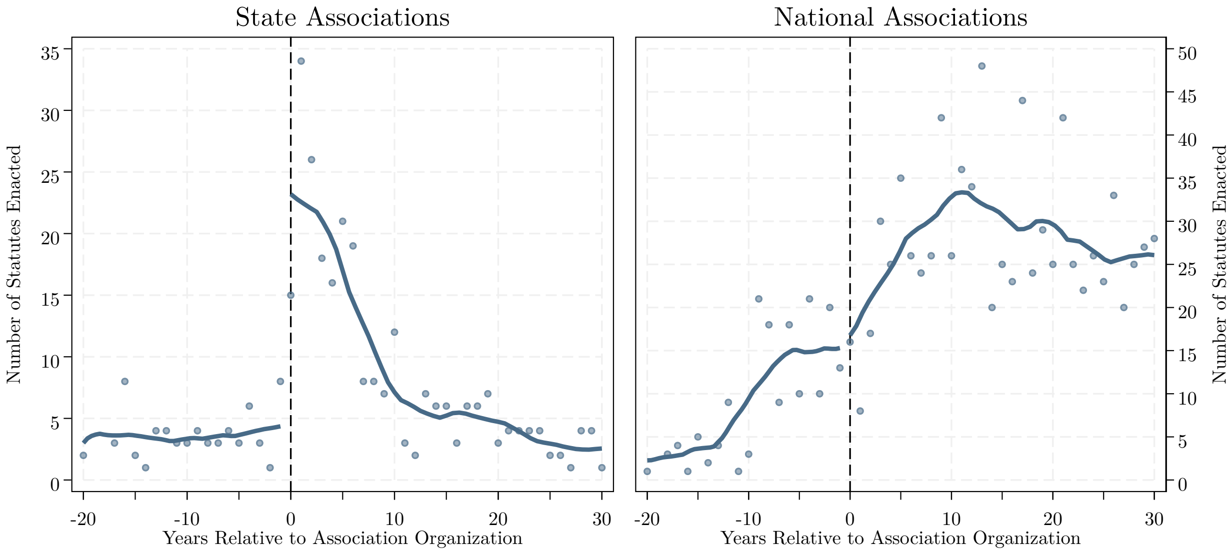
Table A6 reports alternative specifications for our analysis of the effect of labor market competition on occupational licensing and regulation. Here, we show untransformed logistic estimates rather than marginal effects. Columns one and three replicate our preferred specification. In even-numbered columns, we add state fixed effects to the regression, which control for time-invariant state characteristics. States like New York, for instance, are consistently more attractive to immigrants, while the magnitude, timing, and composition of migration flows varies within state over time. Columns five to eight use predicted migration flows over the previous decade rather than over the previous five years. Across specifications, we generally find stronger results for licensing than our broader definition of regulation, which we might expect if certification and registration are less effective at shielding native workers from labor market competition.

A.4 Professional associations

In the main text, we adopt an event study design to quantify the relationship between local political organization and initial regulation, leveraging the staggered timing of association formation. Here, we present an alternative approach that is closer to the raw data. First, we count the number of policies enacted τ years relative to the formation of a professional association in each state. We then fit a locally-weighted regression to this frequency distribution, allowing for a discontinuity at zero (essentially treating association formation as a regression discontinuity in time).

The left panel of Figure A4 uses data for state associations and shows a clear spike in the number of policies enacted just after an association was organized, while there is no trend in the decade before association organization. An advantage of this approach is that unlike our event-study design,

Figure A4: Timing of Initial Regulation Relative to State and National Professional Association Organization

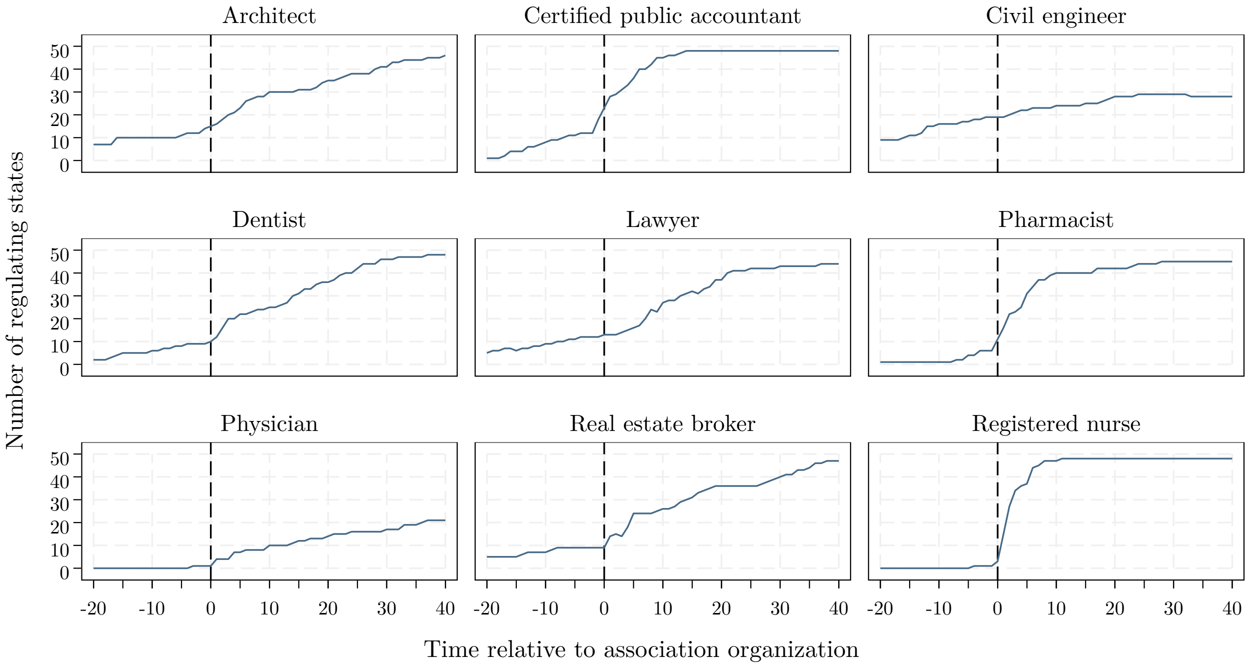


Notes: This figure plots the number of state licensing, certification, and registration laws passed relative to the year in which state and national professional associations were established. Blue lines show fitted locally weighted regression estimates with two year bandwidths, allowing for a discontinuity at zero. We observe data on state-level professional associations for 9 occupations (a total of 411 associations) and data on national-level associations for 61 occupations.

it allows us to consider the role of national associations, shown in the right panel. We find that most laws were passed after the formation of national professional associations, though without an appropriate control group, there is little we can infer from this fact. However, in contrast to the pattern we observe for state associations, we find no noticeable discontinuity in the density of laws passed around the establishment of national associations.

Finally, Figure A5 plots the total number of regulating states relative to the timing of state professional association organization, separately by occupation. This is analogous to our main event study specification, which captures the average change in the slope of these lines around zero. For at least six occupations (architects, accountants, dentists, pharmacists, real estate brokers, and registered nurses) we see a clear increase in the total number of laws in effect just after zero – implying that many states passed laws just after professional organizations were organized. This effect is especially pronounced for nurses and pharmacists, while we see no clear pattern for civil engineers. We also see less of an effect for physicians and lawyers, though in these cases most associations are quite old, and in many cases long pre-dated the diffusion of modern occupational licensing legislation.

Figure A5: Number of Regulating States Relative to Time of State Professional Association Organization



Notes: This figure plots the number of regulatory statutes in effect for each year relative to state professional association organization. State-occupation cells where the timing of professional association organization is unknown are excluded from the sample.

B Data appendix

In this appendix, we provide additional details of our data sources, variable construction, and validation exercises.

B.1 Occupational licensing and regulation

Our data on the timing of occupational licensing policies comes from Carollo (2024). This paper provides a detailed description of the data, but for convenience of the reader, we summarize the occupational classification and policy definitions used in this dataset below.

Occupational classification and coverage. Licensed occupations are classified following the Standard Occupational Classification system and harmonized so that policies regulating the same occupation under different job titles are comparable (e.g., licensed practical and vocational nurses). The data contain many policies with exact six-digit equivalents, but also some that are more or less granular. For example, the SOC system lists several physician subspecialties, while we observe policy data for the five-digit code "physicians" only. Conversely, art therapists are a subset of the SOC code "therapists, all other" and are assigned a custom eight-digit classification.

We use these codes to link our policy data to other datasets with information on occupations. Importantly, these links are based on an assessment of *policy coverage* for statistical occupation categories. We emphasize that not all workers in a covered occupation must necessarily be licensed themselves. Less than half of accountants and civil engineers, for example, actually *attain* a license, as only a subset of these occupations' tasks must legally be performed by a licensed worker. Nonetheless, we think that the detailed occupational categories we observe are the relevant market definition for our analysis. Even if some workers are, in practice, exempt from licensing requirements, they are still likely to be close substitutes for, and thus competing with, other workers in the same occupation who would become licensed if the option were available.

Regulatory taxonomy. Regulatory terminology is often used inconsistently or interchangeably in state legislation. To maximize comparability across states and occupations, Carollo (2024) adopts the following regulatory taxonomy, which we also use throughout our analysis:

1. We classify laws as licensing requirements if they (i) make it unlawful to preform certain tasks without a state credential and (ii) require workers to demonstrate their competency through any combination of experience, training, or examination requirements. We include in this definition laws that protect the use of specific, unmodified, occupational titles, such as "architect" or "physical therapist." These laws, which we term "effective licensing," account for about 5% of the policies we classify as licensing requirements.
2. We classify laws as state certification requirements if they (i) make it unlawful to use title modifiers such as "licensed," "certified," or "registered" without a credential, (ii) require workers to demonstrate their competency through any combination of experience, training, or examination requirements, and (iii) do *not* prevent uncertified workers from preforming specific tasks. In contrast to private certification, state certification explicitly codifies protected titles and enables a state agency to administer and enforce the certification program.
3. We classify laws as registration requirements if they (i) make it unlawful to preform certain tasks without first registering with the state government but (ii) do *not* require any specific qualifications to register. This lack of competency standards distinguishes registration from licensing, through both may impose other requirement, such as criminal background checks, posting a surety bond, or providing proof of insurance.

We observe only state and federal laws. Data limitations prevent us from extending our analysis to requirements adopted by municipal ordinance, which cover a relatively small share of licensed workers and appear mostly, but not exclusively, in the construction industry (Gittleman et al., 2018). That said, there are instances recorded in our data – which we term local regulation – where the *state* legislature enacted (i) a local act establishing a licensing requirement in a specific jurisdiction or (ii) a statewide licensing requirement with enforcement delegated to local authorities. Similarly, some state laws recognize non-governmental credentials for title protection or as minimum competency standards without direct state oversight. Unless otherwise noted, we control for these policies in our analysis but do not include them in our definition of licensure.

B.2 Occupational employment estimates

One of the main challenges we face in our empirical analysis is estimating the number of workers in licensed occupations. First, the Census Bureau classifies jobs using statistical definitions that are too coarse to identify many of the categories we are interested in. Second, occupational classifications are periodically revised, making longitudinal comparisons difficult without further aggregation and loss of detail (Autor and Dorn, 2013). We therefore supplement the Census microdata with additional information on occupations and job titles to construct our preferred employment estimates. Our specific approach varies with data availability across three time periods.

Employment estimates, 1870–1940. To measure occupational employment from 1870 to 1940, we use the original write-in job titles from the restricted full-count Census microdata maintained by IPUMS (Ruggles et al., 2024b). The "occstr" variable records each Census respondent's unedited occupation, exactly as it appeared on their enumeration form. Although IPUMS classifies this free-form text into categories based on the 1950 Census classification, by leveraging the source data directly, we are able to compute employment for occupations that are significantly more granular than those available in the public-use datasets. For instance, the 1950 Census code "therapists and healers" contains several distinct licensed occupations including massage practitioners, naturopaths, occupational therapists, physical therapists, and podiatrists.

We first clean the raw "occstr" text by removing non-alphabetical characters, correcting spelling errors, replacing common abbreviations, and consolidating word order and spacing duplicates. While the processed text still contains millions of unique strings, the top 0.5% of titles (those reported by approximately 100 individuals or more across all census years) account for over 95% of total employment. We therefore sort titles by their 1950 Census classification and manually review the most frequently occurring within each category to build a crosswalk from Census-enumerated job titles to licensed occupations. Finally, we aggregate the microdata across workers to produce a balanced panel of employment counts by state and occupation.

Employment estimates, 1950–1990. As we do not have access to complete Census data after 1940, we must rely on public-use sample extracts from IPUMS (Ruggles et al., 2024a).[1] The most granular information on workers' occupational affiliation in these datasets is the Census Bureau's original classification, which is revised every decade. To produce estimates of employment in licensed occupations, we disaggregate Census codes into six-digit components, using allocation weights that we derive from the distribution of occupations in a database of help-wanted advertisements.

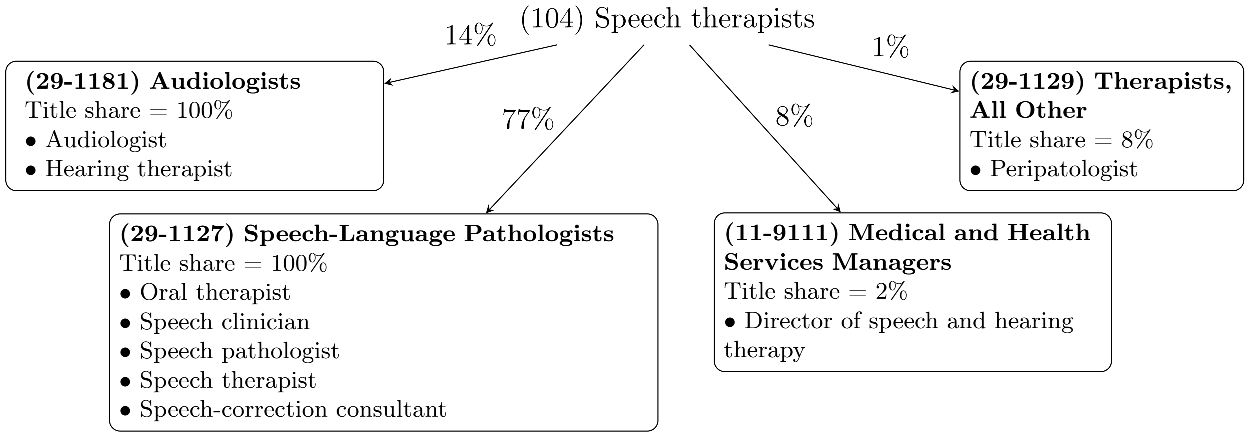
We begin by creating a mapping from each vintage of Census codes to six-digit Standard Occupational Classification (SOC) codes using the Census Classified Index of Occupations and Industries. The Census Index is a set of classification manuals that document how bureau staff were instructed to assign write-in occupation and industry responses to statistical codes. Importantly, the 2010 index classifies over 30,000 job titles using both the Census's four-digit classification and the six-digit SOC system it is derived from. Using a combination of automated matching, replication material provided by Autor et al. (2024), and manual review, we trace individual titles back across previous additions of the Census Index (manually classifying a small share of "dead" titles that do not appear in the 2010 index). For each job title and year, we therefore observe both the Census code it was assigned to and its corresponding six-digit SOC code.

Next, for each Census code, we need to estimate the share of employment that should be allocated to each of its six-digit components. To do this, we use a database of job advertisements from Atalay et al. (2020), who report annual counts of postings for SOC-coded occupations from 1940 to 2000, which they extract from the text of major newspapers.[2] Using this data as a proxy for the relative size of occupations, we compute allocation weights from three-digit Census codes to 6-digit SOC codes as follows:

[1] Although the 1950 Census was recently declassified, IPUMS has not yet processed string variables, meaning that we cannot extend our text classification approach beyond 1940 at this time.

[2] To the best of our knowledge, this is the only large dataset with information on the relative frequency of six-digit occupations spanning the second half of the 20th century. One concern with using job postings to proxy for employment is that the type of occupations employers recruit through help-wanted ads may be unrepresentative (for example, overrepresenting professional services). However, because our approach uses only relative frequencies for occupations that map to the same Census classification, we require only the weaker assumption that vacancy behavior is similar within narrowly defined sets of occupations. For the same reason, we are not concerned about aggregate trends in the level vacancies, provided that these do not differ across very similar jobs.

Figure B1: Example of Estimating Six-Digit Employment from the 1980 Census



Note: This figure illustrates our approach to estimating employment for six-digit occupations between 1950 and 1990. For all titles classified under the three-digit occupation "speech therapists," the boxes show the six-digit codes corresponding to these titles. Percentages reported in the boxes are the share of each six-digit code's titles that map to speech therapists. Percentages above the arrows shows the share of employment we allocate to each six-digit code based on the count of job postings for these categories.

1. Since advertisements are a flow rather than a stock, we count the total number of postings for six-digit occupation j between year t and $t-9$, assuming that the stock of employment in j is roughly proportional to total postings over the previous decade.
2. Some SOC codes include titles mapped to different Census codes, so we distribute postings across three-digit codes i in proportion to the share of *titles* that are classified in i. We note, however, that in practice, most titles in an SOC code typically map to a single Census occupation, as the latter are generally less detailed.
3. For each three-digit code i, we calculate the share of employment to attribute to each of its six-digit components j as the number of postings in the intersection of i and j divided by the total number of postings for i.

Finally, we multiply state-level Census employment by these national allocation factors to produce our estimates of six-digit employment. For licensed occupations with exact six-digit equivalents, we make no further adjustments. For licensed occupations more detailed than SOC codes, we assign employment in proportion to the number of job titles in their parent classification.

Figure B1 provides a graphical illustration of our approach for the 1980 Census code "speech therapists." The Census Index lists nine titles for this occupation, which map to four distinct SOC codes. Two of these, "audiologists" and "speech-language pathologists," are entirely contained within this Census code, while most titles for "medical and health services managers" and "therapists, all other" are classified elsewhere. Using the number of job postings for these SOC codes observed between 1971 and 1980, we assign 77% workers coded as speech therapists by the Census to speech-language pathologists and 14% to audiologists, both of which are licensed occupations. Our method therefore produces employment estimates for these narrow occupation categories that are fully consistent with published Census data when re-aggregated.

Employment estimates, 2000–2020. In 1999 the BLS began releasing employment estimates for six-digit occupations from the Occupational Employment and Wage Statistics (OEWS) survey.

However, state-level data are often suppressed for small occupations, and because they are derived from an establishment survey, they cover a different employment universe than household surveys like the Census. Most importantly, unincorporated self-employed workers are not covered by the survey. Rather than using the OEWS data directly, we therefore estimate employment following the disaggregation approach described above, using OEWS employment in place of job postings and annual ACS data in place of the decennial Census from 2005 onward. Further, after 2000, Census occupations are direct aggregations of six-digit SOC codes, so we do not have to make any further adjustments using job titles from the Census Index in this step.[3]

B.3 Occupational attributes and tasks

To summarize current occupational attributes and train our text classification model, we use O*NET version 22.1, which was released in October 2017. This vintage of the database classifies occupations according to the 2010 Standard Occupational Classification (SOC) system, the same as we use for our main regulation measures. Task descriptions and most numerical variables are available for 963 8-digit occupation categories. Data are not provided for most residual SOC codes, which we exclude from our task analysis.

We use the Dictionary of Occupational Titles (DOT) to summarize the historical task content of occupations. The first edition of the DOT was published in 1939, and until 1998, when it was replaced by O*NET, periodic updates were published. In our analysis, we use digitized versions of the first edition from Althobaiti et al. (2022) and the fourth edition from CISER (2002). The DOT defines over 10,000 job titles in each edition we use. To map these to O*NET occupation codes, we start with a crosswalk provided by O*NET that assigns every DOT occupation in the revised fourth edition to 2017-vintage O*NET codes. Next, we use the DOT's Conversion Table of Code and Title Changes to extend this crosswalk back to the fourth edition. We use a combination of hand-matching and additional conversion tables to classify first edition job titles. Although we are able to assign O*NET codes to only half of the job titles appearing in the first edition, the vast majority of unmatched titles appear in manufacturing industries. We match nearly all service occupations, which are those most critical to our analysis of occupational licensing.

Using NLP to summarize task content

We use natural language processing (NLP) to predict occupational attributes from textual descriptions of job tasks. Our approach proceeds in three steps: (i) define a set of job attributes, (ii) use word embeddings to represent task descriptions numerically, and (iii) estimate a text classification model to predict numerical attributes from these embeddings.

Defining occupation attributes. Starting with the set of numerical variables in O*NET, we group indicators into three summary measures that we believe reflect the public interest view of occupational licensing. For the first measure, we take three descriptors from the O*NET work context category "criticality of position," which captures potential damages from worker error and the extent to which their job requires unsupervised decision-making (all measured on a continuous 1–5 scale):

1. **Consequence of error**: How serious would the result be if the worker made a mistake that was not readily correctable?

[3]The 2000 and 2010 SOC systems are mostly comparable, though we make some minor adjustments for codes that split, as well as codes specific to the OEWS in 2009–2010 and after 2016. We also impute employment for residual categories, which the OEWS does not report before 2004, based on major group totals.

2. **Freedom to make decisions**: How much decision-making freedom, without supervision, does the job offer?
3. **Frequency of decision-making**: How frequently is the worker required to make decisions that affect other people, the financial resources, and/or the image and reputation of the organization?

For the second, we focus on interpersonal interaction (with all elements measured on a 0–7 scale):

1. **Assisting and caring for others**: Providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients.
2. **Performing for or working directly with the public**: Performing for people or dealing directly with the public. This includes serving customers in restaurants and stores, and receiving clients or guests.
3. **Contact with others**: How much does this job require the worker to be in contact with others (face-to-face, by telephone, or otherwise) in order to perform it?
4. **Physical proximity**: To what extent does this job require the worker to perform job tasks in close physical proximity to other people?

Lastly, we measure task complexity using the features selected by Caines et al. (2017). We think of this index as capturing occupations whose tasks are difficult for humans to learn (and presumably those which are difficult for non-experts to evaluate).

After selecting and grouping the O*NET variables above, we follow Yamaguchi (2012) and Caines et al. (2017) in collapsing each attribute to a single composite index using principal component analysis. Given a set of descriptors, $\mathcal{D}$ we compute a principal component score for each O*NET occupation i,

$$v_i = \sum_{d \in \mathcal{D}} \theta_d x_{id}, \tag{B1}$$

where x_{id} is the value of descriptor d for occupation i and θ_d is the principal component factor loading associated with this variable. We standardize v_i to have mean zero and standard deviation one across O*NET occupations.

In addition to our three primary indices, which focus on risk to the public, we also use responses to the O*NET question "How much responsibility is there for the health and safety of other [workers] in this job?" to capture risk to coworkers as another possible rationale for licensing. Finally, we reproduce existing measures capturing routine, cognitive, manual, and social tasks from Acemoglu and Autor (2011) and Deming (2017).

Representing task descriptions numerically. In the second step of our procedure, we represent textual descriptions of tasks from O*NET numerically. For each occupation, we concatenate its "core" task list into a document that describes, in words, what workers do precisely. Following Kogan et al. (2023), we then map each document to a weighted average of word embeddings: mathematical representations that encode the text's meaning in a high-dimensional vector.

First, we convert the task descriptions to lower-case; remove all numbers and punctuation; retain only nouns and verbs; convert these to their singular form and present tense; then drop any remaining words with little semantic content (such as "is," "works," or "does"). Next, we map each word to a 300-dimensional vector using pre-trained word embeddings from the Global Vector

Table B1: Summary of Task Descriptions

	Occupations	Words per Occupation			
		Mean	Std. Dev.	Min	Max
O*NET (2017)	963	106	47	4	303
DOT (1977)	876	634	2,044	10	36,221
DOT (1939)	623	282	791	2	13,696

Notes: This table shows the distribution of words across occupations after pre-processing the three sources we use to construct our document embeddings. The first column reports the number of O*NET occupations with any task description.

for Word Representation (GloVe) model of Pennington et al. (2014).[4] In addition to reducing the dimensionality of the data for our text classification step, word embeddings also help account for synonyms, which is especially important given that we aim to compare tasks across sources that may use different vocabulary to refer to similar concepts. Finally, we compute the weighted average of word embeddings within a document to construct a vector representation of each occupation. The weights are defined as:

$$w_{ik} = \underbrace{\frac{\text{count of words } k \text{ in document } i}{\text{total words in document } i}}_{\text{Term frequency}} \times \ln \underbrace{\left(\frac{\text{\# of documents in corpus}}{\text{\# of documents that include word } k} \right)}_{\text{Inverse document frequency}}. \tag{B2}$$

This weighting choice is common in natural language processing, as it places less weight on terms that occur frequently across many documents and hence are less informative about differences in work across occupations.[5]

Predicting attributes using task descriptions. In the final step, we use the O*NET data to train a text classification model and derive a mapping from document vectors to the numerical attributes defined in step one. Let v_i denote one of our principal component indices for occupation i and $\mathbf{c}_i$ its document embedding. We estimate a ridge regression to minimize an objective function of the form:

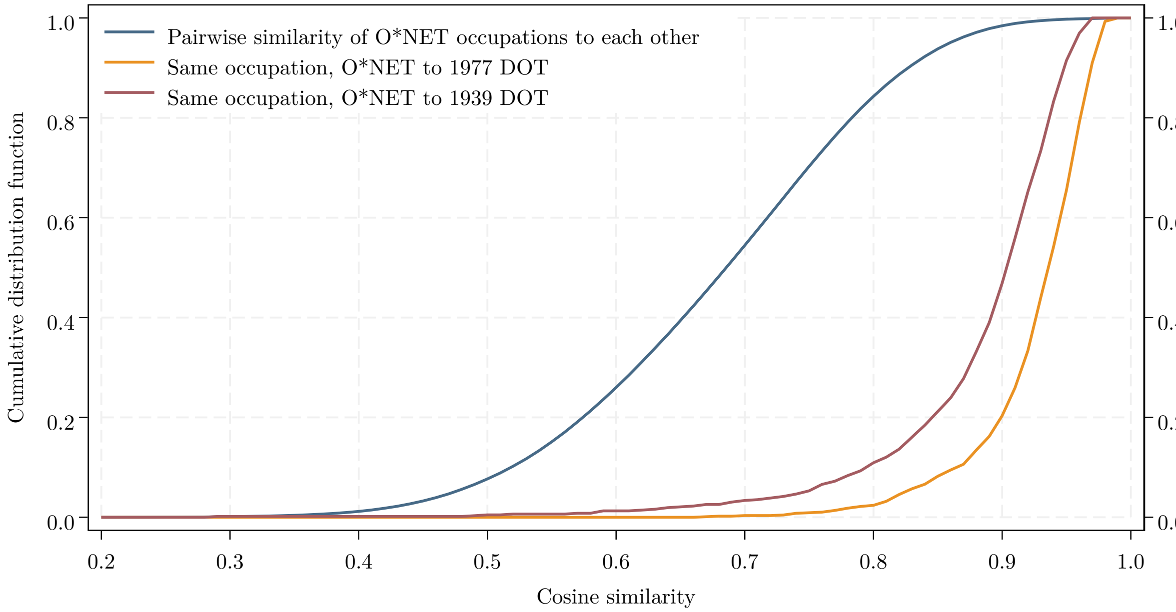
$$\min_{\omega}(\omega; \alpha) = \left[\sum_i \left(v_i - \sum_{k=1}^{300} \omega_k c_{ik} \right)^2 + \alpha \sum_{k=1}^{300} \omega_k^2 \right], \tag{B3}$$

where the first term is the sum of squared residuals and the second is a penalization term that is included to account for the high dimensionality of $\mathbf{c}_i$ (Gentzkow et al., 2019). To reduce the risk of overfitting, we use repeated five-fold cross-validation to select the penalization parameter

[4] Word embeddings represent text geometrically, so that words or segments of text with a similar meaning are mapped to locations close to each other in vector space. The GloVe embeddings we use are based on the co-occurance of millions words in a large corpus of online text containing billions of individual tokens.

[5] Table B1 reports summary statistics on the number of words we retain per occupation after implementing our text pre-processing for each dictionary. Figure B2 plots the cosine similarity of our document vectors within and across dictionaries. It shows that occupation descriptions are generally far more similar in meaning to each other when measured at different points in time than two different occupation descriptions are to each other in O*NET. Further, occupations are more similar, on average, when comparing definitions closer in time. We see this as strong evidence that the document vectors we construct are encoding reliable information about occupations.

Figure B2: Cosine Similarity of Occupational Task Descriptions



Notes: This figure plots three cumulative distribution functions for the cosine similarities between the document vectors we compute. The first (blue) compares O*NET occupations with each other using all pairwise comparisons. The second (red) and third (green) compare the same occupations in O*NET and the DOT.

α separately for each attribute we predict. Finally, we re-estimate the model on the full O*NET data with the optimal α to obtain $\hat{\omega}$, which we then apply to the DOT document embeddings to construct our predicted attributes from those sources,

$$\hat{v}_i = \sum_{k=1}^{300} \hat{\omega}_k c_{ik}. \tag{B4}$$

Once we have fit our text classification model, its output depends on only the text we feed into it. Thus, predictions from O*NET and the DOT are directly comparable to each other, which allows us to measure changes over time. For ease of interpretation, we standardize each predicted attribute using its mean and standard deviation in O*NET. We emphasize, however, that this normalization is rank preserving; it is the text alone that determines whether we classify one occupation as riskier or more complex than another. Figure B3 shows that our text-based approach is effective at predicting the numerical attributes we use, with R^2 values ranging from 0.48 to 0.86.

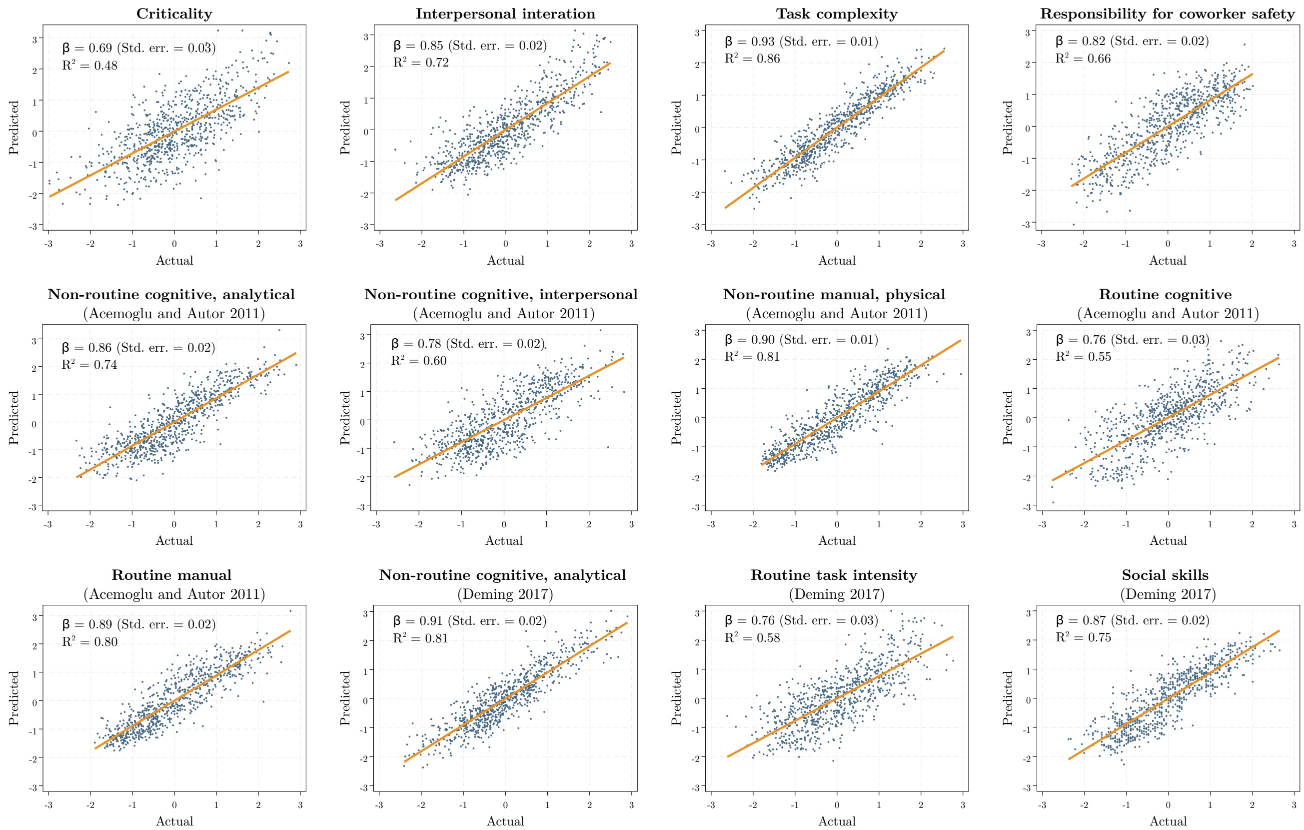
Examples and discussion

The following example highlights the intuition underlying our approach. In 1939, the Dictionary of Occupational Titles defined an *occupational therapist* as a worker who:

> Teaches handicrafts, such as basketry and rug making, to disabled patients in hospitals or institutions for the physically or mentally handicapped.

O*NET, on the other hand, enumerates tasks for an occupational therapist including the following:

Figure B3: Relationship Between Predicted and Actual Occupation Attributes



Notes: This figure plots the correlation between occupational attributes computed from numerical O*NET data and our predictions of these features from textual task descriptions. All variables are standardized to have mean zero and standard deviation one. Graphs are truncated at ± 3 standard deviations for viewability.

> Plan, organize, and conduct occupational therapy programs in hospital, institutional, or community settings to help rehabilitate those impaired because of illness, injury or psychological or developmental problems. Test and evaluate patients' physical and mental abilities and analyze medical data to determine realistic rehabilitation goals for patients. Select activities that will help individuals learn work and life-management skills within limits of their mental or physical capabilities.

Comparing the document vectors from these two sources yields a cosine similarity of 0.57, one of the lowest in our sample. Further, our approach captures the increasing criticality and complexity of this occupation suggested by the text. Our criticality measure for occupational therapists rises from -0.6 standard deviations in 1939 to +0.6 standard deviations in 2017. Similarly, the occupation's complexity rises from -0.5 standard deviations in 1939 to +0.9 standard deviations in 2017. Most of this change occurred before 1977, as the cosine similarity of the O*NET definition of occupational therapy rises to 0.98 when compared with the 4th edition of the DOT.[6]

Table B2 plots changes in each of our predicted occupational attributes since 1939 and 1977 for

[6]Interestingly, the first laws licensing occupational therapy were passed by Florida and New York in 1975. This fact suggests that by the time licensing of this occupation began to diffuse, its task content was fairly similar to what it is today (Fla. Laws, Ch. 75-179; 1975 Laws of New York, Chapter 779).

Table B2: **Changes in Occupational Attributes**

	Δ 1939–2017	Δ 1977–2017
Criticality of position	-0.14	-0.21
Personal interaction	-0.10	-0.08
Task complexity	0.26	-0.00
Responsiblity for coworkers	-0.15	-0.09
Acemoglu & Autor (2011)		
Non-routine cognitive, analytical	0.30	0.05
Non-routine cognitive, interpersonal	0.23	0.02
Non-routine manual, physical	-0.30	-0.15
Routine manual	-0.14	-0.06
Routine cognitive	0.09	0.00
Deming (2017)		
Non-routine cognitive, analytical	0.23	-0.00
Routine task intensity	0.28	0.12
Social skills	0.23	0.07

Notes: This table reports changes in the average of our predicted occupational attributes over time. All values are normalized to express units in terms of standard deviations in the set of O*NET occupations.

the set of 6-digit SOC occupations we study. These reflect a combination of (unweighted) changes in task content within occupations, as well as the growing set of occupations with task descriptions over time. We find that since 1939, average criticality, interpersonal interaction, and responsibility for coworkers has fallen somewhat, while tasks became more complex between 1939 and 1977.

While most studies of occupational task content use measures collected at a single point in time (and none include the main attributes we are interested in), Atalay et al. (2020) examine trends in the Acemoglu and Autor (2011) set of tasks between 1950 and 2000, using the text of help-wanted advertisements from major newspapers. It is reassuring, therefore, that like us, they find an increase in non-routine analytic and interpersonal tasks for 6-digit occupations, and a decline in routine and non-routine manual tasks. By contrast, they find a decrease in routine cognitive tasks, while we find a small increase since 1939.[7] Similarly, Nedelkoska et al. (2023) document the rising prevalence of analytical and interpersonal tasks and the decline of manual tasks in the DOT using a neural-network-based model. Our methodology therefore results in measures of occupational attributes that are broadly consistent with trends documented by the existing literature.

B.4 State similarity and policy diffusion

To study the diffusion of occupational licensing requirements across states, we first define the following distance metrics. We then use these measures to compute the clustering statistics described in Section 5.2.

1. **Geographic proximity.** We use the distance between states' geographic centroids, which we obtain from the Census Bureau's Master Address File/Topologically Integrated Geographic Encoding and Referencing database.

[7] Atalay et al. (2020) use a continuous bag of words approach, which counts the occurrence of pre-defined sets of words in their text.

2. **Demographic similarity.** We use data on states' log total population, urbanization rate, share of the population that is white, share of the population that is foreign-born, and log income per capita from 1929 to 2020. Following DellaVigna and Kim (2022), we first standardize each variable and then calculate demographic similarity as the mean absolute difference across these five characteristics in a given year.

3. **Political ideology.** Our preferred measure of political ideology uses the state policy liberalism index constructed by Caughey and Warshaw (2014), which is available annually from 1936 to 2021. We define ideological distance to be the absolute difference between two states' values of this measure in a given year. Alternatively, we use the absolute difference in Republican two-party presidential vote shares in the election year closest to time t.

4. **Industry composition.** We use data from the 1910 to 2000 decennial census and the 2001 to 2020 American Community Survey to compute the share of each state's workers who are employed in 12 broad industry groups, using the "ind1950" classification system. We then define the distance between states s and r using the dissimilarity index

$$\mathcal{D}_{srt} = \frac{1}{2} \sum_{j=1}^{12} \mid e_{jst} - e_{jrt} \mid , \tag{B5}$$

where e_{jst} is the employment share of industry j in state s at time t. This metric can be interpreted as the share of a state's workers that would have to switch industries in order to achieve the same industrial composition as another.

B.5 State demographic, economic, and political characteristics

In this section, we describe the sources of the remaining state-level variables we use in our analysis. Unless otherwise noted, we linearly impute decennial data to construct an annual panel. All monetary variables are adjusted to 2019 dollars.

- **Total resident population** (annual 1900–2020; decennial 1870–1900): We use intercensal population estimates from 1900 to 2020 from the Census Bureau, retrieved from the FRED database at the Federal Reserve Bank of St. Louis. From 1870 to 1900 we interpolate decennial estimates from the Census Bureau.

- **Urbanization** (decennial 1870–2020): We use the share of the state's population living in urban areas from the Census Bureau. Note that the definition of urban areas was revised in 1880, 1890, 1900, 1950, 2000, and 2020.

- **Demographic characteristics** (decennial 1870–2020) and **foreign-born population share** (decennial 1870–2000 exc. 1890; annual 2001–2020): We use data on the share of the state's population by sex and race (white, black, and other non-white) from the Census Bureau and NHGIS (Manson et al., 2021). Foreign born population shares are computed directly from Census and American Community Survey (ACS) microdata as the share of a state's non-institutionalized adult population who were born outside the United States, including in U.S. outlying areas (Ruggles et al., 2024a).

- **Personal income** (annual 1929–2020): We use total personal income per capita from the Bureau of Economic Analysis. This variable includes all income from wages, proprietors' earnings, dividends, interest, rents, and government benefits.

- **Tax revenue** (annual 1942–2020): We use total state tax revenue from the Census Bureau's Annual Survey of State Government Tax Collections, retrieved from the FRED database at the Federal Reserve Bank of St. Louis.

- **Employment and unemployment**: We use data on total non-farm payroll employment from the Bureau of Labor Statistics' Current Employment Statistics survey (annual 1939–2020) and state unemployment and labor force participation rates from the Current Population Survey (annual 1976–2020).

- **Unionization** (annual 1964–2020): We use the union density estimates compiled by Hirsch et al. (2001) and updated though 2020 at `unionstats.com`. This variable is defined as the percentage of non-agricultural wage and salary employees, including public sector employees, who are union members.

- **Right-to-work state**: Indicates that the state has adopted a statute or constitutional provision that prohibits unionized workplaces from negotiating contracts requiring non-union employees to contribute to the costs of union representation. Source: National Right to Work Committee. `https://nrtwc.org/facts/state-right-to-work-timeline-2016/`.

- **State government partisanship** (annual 1870–2020): We combine data from several sources to create an annual panel tracking the political affiliation of state governors and legislators. We use data on the partisan division of legislatures from 1870 to 1936 from Burnham (1992). From 1937 to 2011, we obtain this information from Klarner (2003) as well as updates available on the author's State Politics and Policy website: `http://academic.udayton.edu/SPPQ-TPR/klarner_datapage.html`. To complete the panel, we use data from the National Conference of State Legislatures for 2012 to 2020. Finally, we appended data on the political affiliation of state and territorial governors from various state-specific sources. We define unified Democratic and Republican governments to be those where the governor and majorities in both houses of the legislature belong to the same party. For Nebraska (which has had a non-partisan legislature since 1937) and for territorial governments (where we do not observe political affiliations) we assign unified party control based on the governor's party.

- **Presidential elections** (election years 1868–2020): We use data on state election returns from Congressional Quarterly, Inc.

- **Policy liberalism.** (annual 1936–2020): We use the state policy liberalism measure originally constructed by Caughey and Warshaw (2014) and later extended by the authors through 2021. This measure captures the ideological evolution of state governments over time by applying a dynamic latent-variable model to data on 148 state policies. Higher values of this index are associated with legislation generally associated with the U.S. political left.

- **Progressive legislation index.** (annual 1870–1950): This variable is taken from Fishback and Kantor (1998). It counts the number of the following laws a state has adopted by year t: compulsory school attendance, establishment of a state tax commission, establishment of a state welfare agency, establishment of a merit system, initiative and referendum, direct primary, minimum age for child labor, mothers' pension, and the establishment of a state commission to regulate electricity rates.

C Theory appendix

C.1 Model derivation

Our solution to the household's problem follows Kleiner and Soltas (2023), which we repeat below for the case of heterogeneous workers. For simplicity, we abstract from the intensive margin of labor supply and assume that occupational licensing has no direct effect on productivity.

Consumption decision and labor demand. Given income $I_i = (1 - \tau_{J_i}) A_{J_i}(y_i) w_{J_i}$, individual i's demand for occupation j's services solves the problem,

$$\max_{\{c_{ij}\}} \ C_i = \left(\sum_{j=1}^{J} \gamma_j(z_j) c_{ij}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} \quad \text{s.t.} \quad \sum_{j=1}^{J} w_j c_{ij} \leq I_i,$$

which implies that

$$c_{ij} = \frac{(w_j/\gamma_j(z_j))^{-\epsilon}}{P^{1-\epsilon}} I_i.$$

If we aggregate across occupations j, individual i's total consumption is $C_i = I_i/P$.

Educational attainment. Given wages and initial licensing requirements, individuals select their level of education to solve

$$\max_{y_i} \ \ln C_i - \rho_k y_i,$$

which implies that

$$\frac{\partial \ln A_j(y_i^*)}{\partial y_i} = \rho_k.$$

With diminishing marginal returns to education, this optimality condition implies that workers with a low marginal cost of training ρ_k (i.e., those with high ability) obtain more education conditional on their occupation choice and select into occupations where the marginal product of education is high. Moreover, conditional on entering occupation j, workers of type k choose the same level of education, which we assume serves as the basis for organizing a professional association.

Occupation choice and labor supply. Since education choices vary only by type and occupation, the common component of indirect utility for a worker of type k who enters occupation j is

$$v_{jk} = e^{-\rho_k(y_{jk}^* + z_j)} (1 - \tau_j) A_j(y_{jk}^*)(w_j/P). \tag{C1}$$

Occupation choice is then pinned down by the solution to the discrete choice problem,

$$J_i^* = \arg\max_j \mu_{ij} v_{jk},$$

which, given our assumptions on the distribution of idiosyncratic preferences, implies that the share of type k workers who enter occupation j is

$$s_{jk} = \frac{v_{jk}^{\sigma}}{\sum_j v_{jk}^{\sigma}}.$$

Expected utility. In equilibrium, expected utility must be equal across all occupations conditional on type. For type k workers, we have

$$\mathcal{W}_k = \Gamma(1 - \frac{1}{\sigma}) \left(\sum_{j=1}^{J} v_{jk}^{\sigma} \right)^{\frac{1}{\sigma}}. \tag{C2}$$

Equilibrium conditions. Aggregating across individuals, we can write demand for occupation j's services as

$$C_j = \frac{(w_j/\gamma_j(z_j))^{-\epsilon}}{P^{1-\epsilon}} \sum_{j=1}^{J} \sum_{k=1}^{K} n_k s_{jk} A_j(y_{jk})(1-\tau_j) w_j,$$

which is a function of the quality adjusted price and aggregate income. Aggregating across worker types, total labor supply to occupation j is

$$H_j = \sum_{k=1}^{K} n_k s_{jk} A_j(y_{jk}).$$

Throughout our analysis, we assume that occupations are "small," meaning that a change in licensing requirements has no first-order effect on aggregate income or the price level. We further assume that licensing occupation j' has no effect on the equilibrium wage of $j \neq j'$. Given these assumptions, in equilibrium, we have:

Consumption demand:

$$\frac{\partial \ln C_j}{\partial z_j} = \epsilon \left(\frac{\partial \ln \gamma_j(z_j)}{\partial z_j} - \frac{\partial \ln w_j}{\partial z_j} \right);$$

Labor supply:

$$\frac{\partial \ln H_j}{\partial z_j} = \sum_{k=1}^{K} s_{jk}^{L} \frac{\partial \ln s_{jk}}{\partial z_j};$$

Occupation choice:

$$\frac{\partial \ln s_{jk}}{\partial z_j} = \sigma \left(\frac{\partial \ln w_j}{\partial z_j} + \frac{\partial \ln(1-\tau_j)}{\partial z_j} - \rho_k \right);$$

Licensing fees:

$$\frac{\partial \ln(1-\tau_j)}{\partial z_j} = -\frac{\tau_j}{1-\tau_j} \left(\frac{\partial \ln \kappa_j(z_j)}{\partial z_j} - \frac{\partial \ln w_j}{\partial z_j} - \frac{\partial \ln H_j}{\partial z_j} \right);$$

Labor market clearing:

$$\frac{\partial \ln C_j}{\partial z_j} = \frac{\partial \ln H_j}{\partial z_j};$$

where

$$s_{jk}^{L} = \frac{n_k s_{jk} A_j(y_{jk}^*)}{\sum_k n_k s_{jk} A_j(y_{jk}^*)}$$

is the labor share of type k workers in occupation j. Simplifying this system of equations, we can write the change in the equilibrium wage as

$$\frac{\partial \ln w_j}{\partial z_j} = \frac{\epsilon}{\epsilon+\sigma}\frac{\partial \ln \gamma_j(z_j)}{\partial z_j} + \frac{\sigma}{\epsilon+\sigma}\left(\sum_{k=1}^{K} s_{jk}^{L}\rho_k + \frac{\tau_j}{1-\tau_j}\frac{\partial \ln \tau_j}{\partial z_j}\right),$$

which reflects (i) consumers' willingness to pay for licensing, (ii) average training costs, and (iii) the pass-through of licensing fees to service prices. Setting $\tau_j = 0$ gives equation (7) in the text.

Finally, the equilibrium change in consumption is

$$\frac{\partial \ln C_j}{\partial z_j} = \frac{\epsilon\sigma}{\epsilon(1-\sigma\tau_j)+\sigma}\left(\frac{\partial \gamma_j(z_j)}{\partial z_j} - \tau_j\frac{\partial \ln \kappa_j(z_j)}{\partial z_j} - (1-\tau_j)\sum_{k=1}^{K} s_{jk}^{L}\rho_k\right),$$

which may rise or fall depending on whether consumer WTP for licensing is large enough to offset any additional training and regulatory costs.

C.2 Welfare effects

Social welfare. We define the change in social welfare for a marginal change in licensing requirements as

$$\begin{aligned}
\frac{\partial \ln \mathcal{W}}{\partial z_j} &= \sum_{k=1}^{K} n_k \frac{\partial \ln \mathcal{W}_k}{\partial z_j} \\
&= \sum_{k=1}^{K} n_k s_{jk}\frac{\partial \ln v_{jk}}{\partial z_j} \\
&= \sum_{k=1}^{K} n_k s_{jk}\left(\frac{\partial \ln w_j}{\partial z_j} + \frac{\partial \ln(1-\tau_j)}{\partial z_j} - \rho_k\right) \\
&= s_j^{E}\left(\frac{\partial \ln w_j}{\partial z_j} + \frac{\partial \ln(1-\tau_j)}{\partial z_j} - \sum_{k=1}^{K} s_{jk}^{E}\rho_k\right) \\
&= s_j^{E}\left(\frac{\partial \ln w_j}{\partial z_j} - \frac{\tau_j}{1-\tau_j}\frac{\partial \ln \tau_j}{\partial z_j} - \sum_{k=1}^{K} s_{jk}^{E}\rho_k\right),
\end{aligned}$$

which establishes Proposition 1. Here, we define employment shares as

$$s_j^{E} = \sum_{k=1}^{K} n_k s_{jk} \quad \text{and} \quad s_{jk}^{E} = \frac{n_k s_{jk}}{\sum_k n_k s_{jk}}.$$

Setting $\tau_j = 0$ and substituting in $\frac{\partial \ln w_j}{\partial z_j}$ yields

$$\frac{\partial \ln \mathcal{W}}{\partial z_j} = \frac{s_{jk}^E}{\epsilon + \sigma}\left[\epsilon\left(\frac{\partial \ln \gamma_j(z_j)}{\partial z_j} - \sum_{k=1}^{K} s_{jk}^E \rho_k\right) + \sigma \sum_{k=1}^{K}(s_{jk}^L - s_{jk}^E)\rho_k\right].$$

Thus, when $s_{jk}^L \approx s_{jk}^E$,

$$\frac{\partial \ln \mathcal{W}}{\partial z_j} \approx \frac{\epsilon s_j^E}{\sigma + \epsilon}\left(\frac{\partial \ln \gamma_j(z_j)}{\partial z_j} - \sum_{k=1}^{K} s_{jk}^E \rho_k\right),$$

which is (6) in the text.

Association's welfare. Next, Proposition 2 follows by differentiating the common component of indirect utility for association members of type $k = g$ who have *already chosen* occupation j in the initial equilibrium,

$$\frac{\partial \ln \mathcal{W}_g}{\partial z_j} = \frac{\partial \ln w_j}{\partial z_j} - \frac{\tau_j}{1 - \tau_j}\frac{\partial \ln \tau_j}{\partial z_j} - \rho_g.$$

C.3 Proof of propositions

Proposition 3 states that with fixed costs, the supply of regulation is limited by the extent of the market. To see this, note that because both social welfare and the association's welfare are strictly decreasing in τ_j it is never optimal for the politician to choose a fee larger than necessary to cover regulatory costs. Thus, we can write the tax rate as a function of the licensing standard:

$$\tau_j = \frac{\kappa_j(z_j)}{w_j(z_j)H_j(z_j)}.$$

As market size $H_j(0) \to 0$, the tax base shrinks and $\tau_j \to 1$, which shuts down the market. This cannot be optimal for an occupation that provides positive social surplus, implying that holding all other parameters constant, there exists some market size below which $z_j^* = 0$. Conversely, τ_j^* decreases as $\kappa_j(z_j)$ falls, increasing the politically optimal level of regulation (if positive).

Proposition 4 states that the politically optimal level of regulation rises with consumer willingness to pay. To see this, note that both social welfare and the association's welfare depend on the change in the wage, which is increasing in $\gamma_j'(z_j)$. If we hold costs fixed, a higher level of regulation is (weakly) preferred by both the association and the social planner than when $\gamma_j'(z_j) = 0$.

Proposition 5 gives the conditions under which the politically optimal level of regulation coincides with the social optimum. From the planner's perspective, the optimal level of regulation satisfies

$$\epsilon\frac{\partial \ln \gamma_j(z_j)}{\partial z_j} = (\epsilon + \sigma)\sum_{k=1}^{K} s_{jk}^E \rho_k - \sigma \sum_{k=1}^{K} s_{jk}^L \rho_k + \epsilon\frac{\tau_j}{1 - \tau_j}\frac{\partial \ln \tau_j}{\partial z_j},$$

and from the association's perspective, it satisfies

$$\epsilon\frac{\partial \ln \gamma_j(z_j)}{\partial z_j} = (\epsilon + \sigma)\rho_g - \sigma \sum_{k=1}^{K} s_{jk}^L \rho_k + \epsilon\frac{\tau_j}{1 - \tau_j}\frac{\partial \ln \tau_j}{\partial z_j}.$$

These two expressions are identical save for training costs. Intuitively, the professional association and the planner agree on the marginal benefit of licensing (to consumers), but their marginal costs differ. With diminishing marginal returns to training, the association therefore prefers a higher level of regulation than the social planner when its members' training costs are lower than the occupational average. Since the lobbying game implies that the politically-optimal level of regulation solves a convex combination of the social planner's and association's problems, the result follows.