A Welfare Analysis of Occupational Licensing in U.S. States*

Morris M. Kleiner    Evan J. Soltas

July 1, 2019

Abstract

We assess the welfare consequences of occupational licensing for workers and consumers. We estimate a model of labor market equilibrium in which licensing restricts labor supply but also affects labor demand via worker quality and selection. On the margin of occupations licensed differently between U.S. states, we find licensing raises wages and hours but reduces employment. We estimate an average welfare loss of 15 percent of occupational surplus. Consumers and workers each bear about half of the incidence. Higher willingness to pay offsets only 60 percent of higher prices for consumers, and higher wages compensate workers for only 70 percent of the cost of mandated investment in occupation-specific human capital.

Keywords: Occupational Licensing, Labor Supply, Human Capital, Welfare Analysis

JEL Codes: D61, J24, J38, J44, K31

---

*Kleiner: Humphrey School of Public Affairs, University of Minnesota (kleiner@umn.edu) and NBER. Soltas: MIT Department of Economics (esoltas@mit.edu). For helpful feedback, we thank Abi Adams, David Autor, Alex Bryson, Ashley Craig, Amy Finkelstein, Jon Gruber, Alan Krueger, Jim Poterba, Ferdinand Rauch, Alex Tabarrok, Owen Zidar, and conference and seminar participants at APPAM, ASSA, LERA, MIT, YES-NYU, and the Upjohn Institute for Employment Research. This paper previously circulated under the title “Occupational Licensing, Labor Supply, and Human Capital.” Evan gratefully acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. 1122374.
1 Introduction

Occupational licensing policies, a major category of labor market regulation in the United States and other countries, have potential costs and benefits. Chief among the costs is that licensing may reduce the supply of labor in licensed occupations. Among the potential benefits are gains in product quality due to the resolution of inefficiencies from asymmetric information. Though there is often heated debate over the trade-offs posed by licensing, economists have thus far offered little guidance on how to conduct a welfare analysis of such policies. This paper develops a theoretical framework for evaluating the welfare effects of licensing and implements it empirically for occupations that some U.S. states license and others do not.

We introduce a model of licensing as a required upfront investment of time in training, to which workers respond by adjusting hours, occupation choice, and consumption. We allow this investment to affect labor quality, both directly and indirectly via the selection of workers who choose to enter an occupation. We prove that, within a set of assumptions that define a class of models, the changes in employment and the wage bill are sufficient statistics for welfare analysis (Chetty, 2009). The change in the share of workers in the occupation reveals the change in worker welfare, and the change in the wage bill—equivalent to consumption expenditure in our labor trading economy—reveals the change in consumer welfare. Our model captures the fundamental welfare trade-off in setting licensing policies between cost and quality and characterizes who, among workers and consumers, bears the welfare costs and benefits from such policies in equilibrium.

We estimate the model using variation among U.S. states and occupations in the share of workers who hold an active government-issued professional license as a proxy for licensing policy. Comparing similar workers across states and occupations in a two-way fixed effect design, we estimate causal effects of licensing on wages, hours, and employment that correspond to reduced-form moments of our model. We further develop a method to estimate the opportunity costs of licensing from its effects on the age structure of workers in occupations, and we substantiate these estimated costs with evidence that licensing increases and reallocates human capital investment. We use these reduced-form effects to estimate the welfare consequences of licensing and the structural parameters of our model.

We conclude that, for marginal occupations licensed by U.S. states, the welfare costs of licensing appear to exceed the benefits. In particular, we estimate that licensing an occupation reduces total surplus from the occupation by about 15 percent relative to no licensing. Consumers and workers each bear about half of these welfare costs. For consumers, licensing on net increases prices adjusted for willingness to pay (WTP), thereby reducing real incomes, with gains in WTP offsetting only about 60 percent of the increase in price. For workers, wage increases compensate for only about 70 percent of the opportunity cost of investments that licensing regulations mandate.

To reach these welfare conclusions, we first develop a model in which licensing is an entry barrier, which imposes welfare costs, but also generates gains in worker quality and selection that imply its net welfare consequences are ambiguous. Suppose a state government licenses an occupation. The licensing costs cause labor supply in the occupation to contract on the extensive margin,
raising occupational wages and consequently intensive-margin labor supply. Consumers respond to the wage increase by reducing the consumption of labor services from the occupation. To the extent that licensing raises consumer willingness to pay, however, the employment response can be reversed. Since licensing affects both occupational labor supply and demand, its effects on the wage bill and on total labor hours in the occupation are also ambiguous. Our model characterizes the unlicensed and licensed equilibria in terms of wages, hours, and shares of employment by occupation. It relates the division of the welfare costs and benefits of licensing among workers and consumers to these reduced-form responses. As in the Summers (1989) model of mandated benefits, whether a licensing policy raises welfare depends upon whether WTP for training mandated by the license exceeds its social cost of provision. In our model, these quantities are determined by the discount rate and responses to licensing via worker quality, worker selection, and consumer substitution.

Our empirical strategy is to use variation in the licensed share of workers by state and occupation to identify the effects of licensing. In particular, we implement a two-way fixed effect design that compares an outcome of interest, such as employment, in state–occupation cells where a relatively large or small share of workers are licensed relative to both the occupation and state. Our identification assumption is that, relative to the occupation and the state, highly-licensed state–occupation cells are otherwise comparable to cells with lower licensed rates. This approach addresses two fundamental challenges in recent empirical research on licensing. First, the policies are hard to measure in the data. In the U.S., occupations are mostly licensed at the state level and by a myriad of institutions (Kleiner, 2000), few if any of which design policies with statistical definitions of occupations in mind. Second, much of the literature has used research designs that compare individual outcomes between licensed and unlicensed workers. Such comparisons are vulnerable to selection into licensing, a significant concern given the imperfect correspondence between regulatory and statistical definitions of occupations and by analogy to selection into unions (Lewis, 1986) or education (Card, 1999). Using the licensed share of workers in a state–occupation cell as a proxy for policy naturally resolves the former problem and does much to address the latter. Our estimates thus reflect local average treatment effects of licensing occupations with interstate differences in policy, an estimate that approximates the margin intuitively relevant for policymaking. The data come from the U.S. Current Population Survey, which since January 2015 has included questions on license attainment.

We find that licensing increases wages and hours per worker but reduces employment. In our preferred specification, shifting an occupation in a state from entirely unlicensed to entirely licensed increases state average wages in the licensed occupation by 15 percent, increases hours per worker by 3 percent, and reduces employment by 29 percent.\(^1\) To assess the opportunity costs of licensing, we estimate its effects on the distribution of educational attainment and worker age. Most licensing regulations require workers to obtain specific credentials to be legally employed in an occupation (Gittleman et al., 2018). We estimate that licensing an occupation increases average education by

\(^1\)Per a recent U.S. government review (White House, 2015): “While there is compelling evidence that licensing raises prices for consumers, there is less evidence on whether licensing restricts supply of occupational practitioners, which would be one way in which it might contribute to higher prices.”
about 0.4 years. This masks a dramatic reallocation in the types of human capital workers acquire: We find large increases in the shares of workers whose highest degrees are vocational associate’s degrees or graduate degrees and decreases in high school degrees and bachelor’s degrees. We also find licensing delays the entry of younger workers into occupations. This delay is much greater than the increase in average years of education, suggesting additional opportunity costs beyond measured formal schooling. Our results are consistent with actual educational demands of licenses as well as substantial opportunity costs of licensing that could plausibly account for the reduction in labor supply.

We present a variety of robustness checks for our empirical results. First, we validate our findings with data on worker outcomes from other surveys. Second, we show our results hold when comparing workers across state–occupation cells who are similar on predetermined demographic characteristics. Our results are also robust to controlling for the unionized and certified shares in each state–occupation cell, and they remain unchanged in saturated specifications that seek to purge variation most likely related to endogenous political determinants, suggesting that other labor market regulations and institutions do not confound our identification strategy (Besley and Case, 2000). They are also not driven by predictable variation due to state demography and labor demand, and the results are largely robust to restricting comparisons to within closely-related occupational groups or groups of adjacent states. We show our worker-level regressions yield similar results with household fixed effects, making explanations of selection on unobservables implausible. To focus on the likely influence of policy, we show using only the most extreme variation in the state-occupation licensed share also yields similar results.

Our findings have considerable policy implications. They suggest that a general shift of policy towards delicensing marginal occupations would raise welfare, and particularly that of workers. Indeed, in the U.S. and elsewhere, policymakers appear increasingly favorable to deregulatory reforms of occupational licensing. For any such policy decision, the correct statistics for welfare analysis are the occupation-specific responses of employment and the wage bill, as the WTP effect and opportunity cost of licensing may vary among occupations. However, our research design is not adequately powered to evaluate licensing in individual occupations or groups of occupations. In lieu of such estimates, we formulate a potentially useful net-benefits test that compares the WTP effects and opportunity costs of licensing. We interpret our results in this paper as signing the difference between U.S. policy and the social optimum: More U.S. workers appear to be in licensed occupations that are optimally unlicensed than vice versa.

Our paper contributes both theoretically and empirically to the literature on labor market institutions in labor and public economics. Our model of licensing takes as inspiration a classic tradition of models (Akerlof, 1970; Leland, 1979; Shapiro, 1986) that portray how licensing may correct market imperfections. We build more directly, however, upon recent structural models of

---

2See, e.g., current U.S. Labor Secretary Alexander Acosta (The Wall Street Journal, 8 January 2018): “[O]verly burdensome licensure requirements weaken competition without benefiting the public.” Similarly, in a report (White House, 2015), the Obama administration wrote: “Too often, policymakers do not carefully weigh costs and benefits when making decisions about whether or how to regulate a profession through licensing.”
labor markets (Hsieh et al., 2013; Kline and Moretti, 2014; Suárez Serrato and Zidar, 2016; Redding and Rossi-Hansberg, 2017), yielding a framework with testable comparative statics about the effects of licensing on labor market outcomes and that maps directly to welfare and incidence. Economists have recently focused on estimating effects of licensing on wages (Kleiner and Krueger, 2010, 2013), labor supply (DePasquale and Stange, 2016; Redbird, 2017; Blair and Chung, 2018), migration (Kugler and Sauer, 2005; Johnson and Kleiner, 2017), and product quality (Kleiner and Kudrle, 2000; Kleiner, 2006; Angrist and Guryan, 2008; Larsen, 2013; Anderson et al., 2016; Kleiner et al., 2016; Barrios, 2018). Our model organizes such empirical evidence and explains its implications for welfare and incidence. In this focus, our paper is the first of the modern literature to revisit the welfare questions that originally inspired the seminal works by Kuznets and Friedman (1945) and Stigler (1971) on the economics of licensing.

To guide our empirical analysis, we present the theoretical model of licensing in Section 2. We introduce our data and empirical strategy in Section 3. Section 4 reviews the results and finds evidence for the model’s main testable predictions. Section 5 addresses the robustness of these results to confounding factors. Section 6 structurally estimates the model. Section 7 concludes.

2 A Model of Occupational Licensing

We model licensing as a mandatory upfront investment of time for individuals to enter an occupation and characterize the equilibrium responses of labor market outcomes to licensing. Our model is of a labor trading economy: Individuals supply labor for others’ consumption. They choose their occupations, schooling investments, hours of work, and consumption expenditures on labor by occupation, all given licensing requirements, wages, and their preferences for occupational employment, leisure, and consumption. We capture potential benefits of licensing by allowing for changes in labor quality and changes in worker selection into occupations, both of which may change consumers’ willingness to pay for licensed labor.\(^3\)

In equilibrium, licensing raises wages and hours per worker, but its effect on employment depends upon the balance of consumer substitution in response to wage increases and the extent to which these wage increases are offset by increases in WTP. Within a broad class of models, the effects of licensing on employment and the wage bill are sufficient statistics for welfare analysis. This class of models is defined by three conditions (Adao et al., 2017): The model admits a normative representative consumer, production is constant returns to scale, and all markets are perfectly competitive.\(^4\) Our model shows that whether licensing raises employment and the wage bill, and thus the welfare of workers and consumers respectively, depends upon relative magnitudes of occupational labor supply and demand responses, and in turn upon structural primitives. In Appendix

\(^3\)We set aside modeling of specific product- and labor-market imperfections that can generate efficiency gains from licensing. This is because the effect on consumer revealed WTP exhausts the welfare benefits of licensing if there are no externalities or behavioral frictions. Although one can imagine cases in which these assumptions may be incorrect, they are integral to any revealed preference analysis.

\(^4\)As in Adao et al. (2017), these conditions render it sufficient to consider a reduced factor demand system—here, for occupational labor services—rather than a demand system in product space.
B, we present a detailed solution of the model.

2.1 Preliminaries

Individuals are indexed by $i = 1, \ldots, N$ and occupations by $j = 1, \ldots, M$. The government chooses a training requirement $\tau_j$ for each occupation. Entering an occupation with a requirement $\tau_j$ delays individuals’ payoffs by a time interval $\tau_j$. Individuals also invest time $y_i$ in schooling, which similarly delays their payoff. Schooling and licensing, however, differ in two respects. First, the former is an individual choice, whereas the latter is mandatory conditional on occupational choice. Second, schooling raises individual productivity, whereas WTP effects of licensing depend upon the average behavior of all workers in the occupation. After observing $\{\tau_j\}$, individuals solve their respective problems. Individuals invest in schooling, enter one occupation, supply labor for other individuals’ consumption, and consume their entire labor income. We treat their payoffs from these consumption and labor supply decisions as if occurring in a single period. For conceptual clarity, we distinguish between individuals’ roles as workers and consumers, especially in our welfare analysis.

2.2 Problem

Statement. Individuals maximize a utility function with preferences over consumption and labor hours as well as the timing of this payoff, with consumption modeled as a constant elasticity of substitution (CES) composite good and an idiosyncratic occupation-specific preference term $a_{iJ_i}$:

$$
\max \left\{ \log \left( \sum_j q_j c_{ijj}^{\varepsilon / \varepsilon - 1} \right)^{\varepsilon - 1 / \varepsilon} - \frac{\psi}{1 + \eta} h_i^{1 + \eta} - \rho (y_i + \tau_{J_i}) + a_{iJ_i} \right\}
$$

s.t. $\sum_j w_j c_{ij} \leq A_{J_i}(y_i) w_{J_i} h_i.$

Individual $i$ chooses consumption $c_{ij}$ of labor services from each occupation $j$, labor hours $h_i$, years of schooling $y_i$, and an occupation $J_i$ in which to work. The consumption weights $q_j = q(\tau_j, E[a_{iJ_i} | J_i = j])$ are endogenous to training requirements, accommodating potential labor quality and selection effects that affect WTP for goods produced by licensed labor. The elasticity of substitution is $\varepsilon$, the Frisch intensive-margin elasticity of labor supply is $1 / \eta$, the annual discount rate is $\rho$, and $\psi$ scales labor supply. The occupation preference term $a_{ij}$ is distributed i.i.d. Type I Extreme Value with dispersion parameter $\sigma$, with a larger $\sigma$ implying less dispersion in occupational preferences. The wage in occupation is $w_j$ and is common across workers, and $A_j(\cdot)$ is an effective labor supply function, with $A_j'(\cdot) > 0$ for all $j$, so that individual investments in schooling raise effective labor supply. By contrast, training requirements $\tau_j$ affect consumption weights $q_j$ and $\vdots$

\footnote{A sufficient condition for uniqueness is $1 + \sigma (1 + \eta) + \eta \varepsilon \neq 0$. The economic content of this restriction on the model parameters is to ensure that the occupational labor supply and demand elasticities are not equal.}
the equilibrium wage $w_j$ but do not enter $A_j(\cdot)$.

The quality-adjusted price index of the CES composite good is $P = \left( \sum q_j^\epsilon w_j^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}}$. We normalize the wage $w_0 = 1$ of a reference occupation such that $\tau_0 = 0$.

**Willingness to Pay.** Licensing may raise welfare in our model insofar as it either directly raises labor quality or induces selection into licensed occupations that raises WTP for labor provided by licensed workers. For example, consumers may be willing to pay for barbers with more training, just as they may gain from screening out bad barbers who would otherwise pool with good barbers and thus whose services they might otherwise unwittingly purchase. We therefore capture the preceding literature’s two main proposed channels for welfare benefits of licensing—gains in quality and the restoration of efficiency in markets with asymmetric information—in a model that is nevertheless tractable for estimation and welfare analysis. We model the willingness to pay for licensed labor as a log-linear function of training time and the average value of the idiosyncratic preference term of workers in the occupation, capturing respectively quality and selection effects:

$$\log q_j = \kappa_0 + \kappa_1 \tau_j + \kappa_2 \log E[a_i; J_i = j],$$

where $\kappa_1$ and $\kappa_2$ are parameters governing the response of WTP to training time and to selection with respect to occupation preferences. Licensing, of course, may affect the selection of workers along many dimensions, but selection affects WTP only insofar as the attribute on which selection occurs, $a_i$, is itself valued by consumers or correlated with other valued attributes.

**Consumption.** Individual $i$’s consumption of occupation $j$’s labor is

$$c_{ij} = \frac{A_j(y_i)w_i h_i(w_j/q_j)^{-\epsilon}}{P^{1-\epsilon}},$$

and so aggregate consumption of occupation $j$’s labor is

$$C_j = \sum_i c_{ij} = \frac{N(w_j/q_j)^{-\epsilon}}{P^{1-\epsilon}} \sum_j s_j A_j(y_i^*; J_i = j)w_j h_i; J_i = j,$$

where $s_j$ denotes the share of workers in occupation $j$.

---

6The distinction between licensing and schooling is a strict generalization of a model with one human capital stock in which licensing is a occupation-specific mandated minimum for schooling. It is without loss of generality that we do not allow private investment in training.

7This functional form implies that the WTP “returns” to training and selection are constant and is best viewed as a local linear approximation, averaged over heterogeneous treatment effects (e.g., diminishing returns to training). One of many possible micro-foundations for a relationship between WTP and $E[a_i; J_i = j]$ is that workers may exert some costly unobserved effort beyond labor hours which consumers value and which reduces their probability of being banned from working in the occupation. High $a_i$ workers will exert more of this effort in equilibrium. For example, individuals who have a high preference term for being an accountant may be more unobservably diligent because they face an idiosyncratically high cost of a ban.
Schooling. Individual $i$’s schooling choice $y_i^*$ satisfies

$$\rho = \frac{1 + \eta}{\eta} \cdot \frac{A'_{J_i}(y_i^*)}{A_{J_i}(y_i^*)},$$

reflecting that, in equilibrium, individuals equate the marginal delay cost and the marginal individual productivity benefit of schooling (Mincer, 1974). Furthermore, $y_i^*$ is constant among individuals grouped by occupation choice $J_i$, and $y_i^*_{J_i=j}$ is independent of $\tau_j$. Most importantly, the outside option to invest in schooling at equilibrium return $\rho$ enforces, in a sense we make precise below, a required return on licensing requirements. We can also express the present value opportunity cost of licensing requirements as $\ell_j = \rho \tau_j$, where $\ell_j$ is the share of the present value of lifetime labor income dissipated by the requirement.

Labor Supply. The individual’s indirect utility conditional on entering occupation $j$ and the distributional assumption for occupational preferences imply that occupation shares are

$$s_j = \frac{\left( e^{-\rho(y_j^* + \tau_j)} A_j(y_j^*) w_j \right)^{\frac{1 + \eta}{\eta}}}{\sum_{j'} \left( e^{-\rho(y_{j'}^* + \tau_{j'})} A_{j'}(y_{j'}^*) w_{j'} \right)^{\frac{1 + \eta}{\eta}}},$$

Next, individual labor supply is

$$h_{i:J_i=j} = \psi \frac{1}{n} w_j^{\frac{1}{\eta}},$$

and we define total labor supply in occupation $j$ as $H_j = \sum_{i:J_i=j} h_i$.

2.3 Equilibrium

**Definition 1.** Given occupation characteristics $\{\kappa_{0j}\}$, parameters $\{\sigma, \rho, \psi, \eta, \varepsilon, \kappa_1, \kappa_2\}$, and a policy choice $\{L_j\}$, an equilibrium is defined by endogenous quantities $\{\{J_i, h_i, \{c_{ij}\}\}, \{w_j, q_j\}\}$ such that

1. Individuals optimize: For all $i$, occupation $J_i$, hours $h_i$, and consumption $\{c_{ij}\}$ solve Equation 1.
2. Markets clear: For all $j$, the wage $w_j$ is such that labor markets clear:

$$C_j = A_j(y_j^*_{J_i=j}) H_j \quad \forall j.$$

3. Beliefs are confirmed: For all $j$, willingnesses to pay $q_j$ are such that Equation 2 holds.

We now present four equilibrium relationships in the model which, together, compose the system of equations that we solve to obtain comparative statics. Equations 3 and 7 imply that

$$\frac{\partial \log C_j}{\partial \tau_j} = \frac{\partial \log H_j}{\partial \tau_j} = \varepsilon \left( \frac{\partial \log q_j}{\partial \tau_j} - \frac{\partial \log w_j}{\partial \tau_j} \right),$$

7
which states that consumption falls if licensing raises wages by more than it raises WTP. The partial derivative of Equation 5 with respect to $\tau_j$ is

$$\frac{\partial \log s_j}{\partial \tau_j} = \frac{\sigma(1 + \eta)}{\eta} \left( \frac{\partial \log w_j}{\partial \tau_j} - \rho \right),$$  

and, differentiating Equation 6, we obtain

$$\frac{\partial \log h_{i;J_i=j}}{\partial \tau_j} = \frac{\partial \log w_j}{\partial \tau_j}.$$  

The preceding equations show that the response of employment to licensing depends on whether the response of wages to licensing is greater or less than the return on education—that is, on the sign of the response of the present value of income to licensing. This is the sense in which worker responses to the earnings profile reflect a required return $\rho$. The effect on hours per worker depends only on the wage effect, showing licensing distorts the intensive margin of labor supply only indirectly.

Next, we differentiate Equation 2 and apply a result, which we prove in Appendix B, that

$$\frac{\partial \log \mathbb{E}[a_{i,J_i=j}|J_i=j]}{\partial \tau_j} = -\frac{1}{\sigma} \frac{\partial \log s_j}{\partial \tau_j},$$

yielding

$$\frac{\partial \log q_j}{\partial \tau_j} = \kappa_1 - \frac{\kappa_2}{\sigma} \frac{\partial \log s_j}{\partial \tau_j} \equiv \alpha.$$  

The first result relates the change in the conditional expectation of the idiosyncratic occupational preference term $a_{i,J_i}$ to the change in the share of workers in the occupation. Intuitively, if licensing drives out many workers from an occupation, only the “dedicated” types (i.e., high $a_{i,J_i}$) remain, which may raise WTP. We lack an empirical method to distinguish quality from selection—that is, to identify $\kappa_1$ and $\kappa_2$—and henceforth we use the constant $\alpha$ to summarize WTP effects.

### 2.4 Implications of Model

We summarize the model in four propositions. Proofs are in Appendix B.

**Proposition 1.** Consider the case $\alpha = \kappa_1 = \kappa_2 = 0$ (licensing has no effect on WTP). An increase in $\tau_j$ has the following effects in equilibrium:

1. Workers exit the occupation: \( \frac{\partial \log s_j}{\partial \tau_j} = -\frac{\rho \sigma (1 + \eta) (\varepsilon + 1/\eta)}{1 + \sigma (1 + \eta) + \eta \varepsilon} < 0. \)

2. The occupation’s gross wage rises, but its net wage falls: \( \frac{\partial \log w_j}{\partial \tau_j} = \frac{\rho \sigma (1 + \eta)}{1 + \sigma (1 + \eta) + \eta \varepsilon} \in (0, \rho). \)

3. Hours per worker in the occupation rise: \( \frac{\partial \log h_{i;J_i=j}}{\partial \tau_j} = \frac{\rho \sigma (1 + \eta) / \eta}{1 + \sigma (1 + \eta) + \eta \varepsilon} > 0. \)

This proposition demonstrates that, when licensing purely restricts entry, the model yields sensible predictions for outcomes in labor markets which follow from $\sigma$ and $\eta$, which determine the
intensive and extensive margin labor supply elasticities, and \( \varepsilon \), the labor demand elasticity. Licensing raises wages, but absent increases in WTP, these increases are insufficient to fully compensate workers for licensing costs. In response to these changes in gross and net wages, workers increase labor supply in the occupation on the intensive margin but reduce it on the extensive margin. Appendix B contains comparative static formulae for the general case (\( \alpha \) unrestricted), as summarized in the next proposition.

**Proposition 2.** The following inequalities hold for all \( \tau_j \) and \( \alpha \):

\[
\frac{\partial^2 \log w_j}{\partial \tau_j \partial \alpha} > 0, \quad \frac{\partial^2 \log h_j}{\partial \tau_j \partial \alpha} > 0, \quad \frac{\partial^2 \log s_j}{\partial \tau_j \partial \alpha} > 0,
\]

and there exists an \( \alpha < \infty \) such that

\[
\frac{\partial \log w_j}{\partial \tau_j} > \rho, \quad \frac{\partial \log s_j}{\partial \tau_j} > 0.
\]

This proposition states that, if licensing raises WTP, wages and hours per worker rise more, and employment declines less, in response to licensing than under \( \alpha = 0 \), as licensing now raises labor demand in addition to reducing labor supply. If the WTP effect is sufficiently large, employment and the present value of income rise. With Proposition 3, the employment result confirms that the sign of the social welfare impact of licensing is ambiguous and depends on model parameters.

**Proposition 3.** Define social welfare as \( W = \sum_i u_i \), and \( W_j \) as surplus from occupation \( j \). Then

\[
\frac{\partial \log W_j}{\partial \tau_j} = \frac{1}{\sigma} \frac{\partial \log s_j}{\partial \tau_j} + \frac{1 + \eta}{\eta(\varepsilon - 1)} \frac{\partial \log w_j H_j}{\partial \tau_j},
\]

which reflects a change in consumer welfare of

\[
\frac{\partial \log W^C_j}{\partial \tau_j} = \frac{s_j(1 + \eta)}{\eta(\varepsilon - 1)} \frac{\partial \log w_j H_j}{\partial \tau_j},
\]

and a change in worker welfare of

\[
\frac{\partial \log W^L_j}{\partial \tau_j} = s_j \frac{\partial \log s_j}{\partial \tau_j}.
\]

This proposition states that the change in occupational surplus from licensing reflects two considerations: the net change in employment in the licensed occupation, which captures the change in worker welfare, and the change in the quality-adjusted price level, which captures the change in consumer welfare. This result emerges from two revealed-preference arguments based on the responses to licensing of consumers and workers in the licensed occupation. Licensing raises the welfare of consumers insofar as increases in WTP offset increases in the occupation’s wage, which

---

\(^8\)We define the total surplus from occupation \( j \) as \( W_j = W(0, \{\tau_j\}) - \lim_{\tau_j \to \infty} W(\tau_j, \{\tau_j\}) \). Intuitively, this is the difference in social welfare between no licensing for \( j \) and banning entry into \( j \).
reduce consumers’ real income via a higher price level. Lacking data on quality-adjusted prices, we look to changes in the occupational wage bill to reveal changes in consumer welfare: Holding all other prices fixed, the change in the quality-adjusted price level equals \( s_j/(1 - \varepsilon) \) times the change in \( j \)'s wage bill. For workers, the welfare effect follows from a property of all discrete choice models satisfying the “connected substitutes” condition of Berry et al. (2013): One can invert a choice-share function to recover a value function.\(^9\) Consequently, employment shares are related to the common component of conditional utility in an occupation.

**Proposition 4.** Licensing reduces social welfare if

\[\rho > \frac{\alpha \varepsilon}{\varepsilon - 1},\]

This proposition provides a net-benefits test for licensing. It shows that whether the welfare effect of licensing is positive or negative depends upon the relative magnitudes of the WTP effect and consumption substitution elasticity, which together determine the social benefit of greater WTP, and the return on education, which determines the social cost of reduced occupational labor supply. In particular, the WTP effect cannot be too far below the equilibrium return on education if licensing is to raise social welfare. Increases in WTP are the sole motive for licensing in the model: If \( \alpha = 0 \), there are no values of the other parameters for which licensing raises welfare. Moreover, this proposition illustrates the close connection of our model to Summers (1989): Whether for employer-side benefits or worker training, the welfare cost of a mandate reflects the difference between willingness to pay and the social cost of provision.

**Proposition 5.** Workers and consumers respectively bear shares \( \gamma^L \) and \( \gamma^C \) of the incidence of a change in licensing, where

\[\gamma^L = \frac{\Delta W^L}{\Delta W} = \frac{(\alpha - \rho)\varepsilon - \rho}{(\alpha - \rho)\eta\varepsilon + \rho} \cdot \frac{\varepsilon - 1}{1 + \sigma(1 + \eta) + \eta\varepsilon}; \quad \gamma^C = 1 - \gamma^L.\]

A change in licensing raises consumer welfare but reduces worker welfare if

\[\Delta s_j < 0 < \Delta w_j H_j \iff \alpha \in \left( \frac{\rho\sigma(1 - 1/\varepsilon)}{1 + \sigma}, \rho \left( 1 - \frac{1}{\eta\varepsilon} \right) \right).\]

The first part of the proposition shows the incidence of licensing. Workers bear a smaller share of the incidence when \( \sigma \) is high (occupational choice is more elastic to net income), \( \rho \) is high (delay is more costly), or \( \varepsilon \) is low (consumers are inelastic). The second part of the proposition shows that, in the model, it is possible for changes in consumer and worker welfare from licensing to be differently signed, and that this case coincides with licensing reducing employment but raising the wage bill in the licensed occupation. In fact, this case appears to be of practical and not purely

\(^9\)Our results are also related to the “gains from trade” formula of Arkolakis et al. (2012). For large occupations, Appendix B proves the employment effect on the licensed occupation remains the sufficient statistic. Due to between-occupation spillovers, the Herfindahl index of employment shares is also required to estimate the magnitude, but not the sign, of the worker welfare effect. In our application the Herfindahl term is negligibly small.
theoretical interest: As long as $\sigma > 0$, this interval exists, and the behavioral responses to licensing estimated in Section 6 suggest that licensing in marginal U.S. occupations is near this trade-off.

3 Data and Empirical Strategy

3.1 Data

We use new survey questions in public microdata from the basic monthly U.S. Current Population Survey (CPS) from January 2015 to December 2018.\(^\text{10}\) The CPS asks adults in survey households three questions about certification and licensing. The questions are:

Q1. “Do you have a currently active professional certification or a state or industry license?”
Q2. “Were any of your certifications or licenses issued by the federal, state, or local government?”
Q3. “Is your certification or license required for your job?”

To match the U.S. government definition of an occupational license,\(^\text{11}\) we say a worker is licensed if he or she answers yes to both Q1 and Q2—that is, if the worker holds an active government-issued professional certification or license—and say the worker is not licensed otherwise. We say a worker is certified if he or she answers yes to Q1 but no to Q2—that is, if he or she holds an active professional certification or license but it is not government-issued—and use certification as a control in robustness checks. Our decision to use the CPS is informed by sample size, as precise estimates of state–occupation licensed shares are an essential component of our research design. The sample covers 624,697 unique workers, and Appendix Table A1 tabulates workers by their answers to these survey questions: 27.5 percent are licensed or certified, and 22.6 percent are licensed.\(^\text{12}\) These shares are consistent with those in other survey data (e.g., Kleiner and Krueger, 2013; Blair and Chung, 2018).

Our analysis defines occupations according to 2010 Census categories. The sample contains workers in 483 occupations.\(^\text{13}\) We measure licensing by the licensed share of workers in a state–occupation cell as a proxy for policy. Informing our approach, state and local governments define

\(^{10}\)Appendix Table A7 replicates our main results using microdata from the 2010–15 American Community Survey. We do so by merging our CPS-based estimates of state–occupation licensed shares into ACS microdata.

\(^{11}\)According to the Interagency Working Group on Expanded Measures of Enrollment and Attainment, an occupational license is “[a] credential awarded by a government agency that constitutes legal authority to do a specific job.” See https://nces.ed.gov/surveys/GEMEnA/definitions.asp. We do not restrict the sample to yes on Q3 because we found this leads to counterfactually low licensed shares of workers, both overall and in universally-licensed occupations.

\(^{12}\)All data are drawn from the Integrated Public Use Microdata Series (Flood et al., 2018). We limit the sample to employed adults age 16 to 64, except for age regressions, and follow Autor et al. (2008) to address topcoding and allocation of earnings by estimating hourly earnings for non-hourly workers and by Winsorizing for earnings below half the federal minimum wage. We also Winsorize usual weekly hours above 100 and map educational attainment to years of education using data from their replication materials.

\(^{13}\)The reciprocal of the Herfindahl index of occupation shares of employment—which measures the “effective” number of occupations—is 108.7, indicating workers are not concentrated in a few occupations.
licensed occupations at their discretion and obey no occupational classification scheme. For example, some states license occupations as specific as eyebrow threading (Carpenter et al., 2012). The many regulatory bodies that license occupations across states, as well as the challenge of harmonizing definitions of occupations, have made licensing particularly difficult to study.

Our proxy naturally resolves this mapping of regulations to Census categories. Workers in licensed occupations must by law themselves be licensed, and thus variation in licensed shares is unlikely to reflect selection into licensing conditional upon occupation. Misalignment between regulatory and statistical definitions of occupations, however, would result in Census occupational categories pooling some unlicensed occupations with licensed ones as defined by state regulations. Other factors, such as survey misresponse and individuals who hold licenses for occupations other than those in which they work, may also contribute to this phenomenon. Appendix Figure A1 shows that, due to these considerations, there is considerable mass of the cell licensed share distribution at values between 0 and 1. The mass suggests significant scope for selection into licensing—that is, into “suboccupations” unobservable to the researcher that differ in both licensing and outcomes of interest—that we resolve by using cell licensed shares as a measure of policy.

Does self-reported license status reflect the truth? Given data limitations, one feasible test is to evaluate how the probability of self-reported licensing differs between occupations that are and are not “universally licensed” by U.S. states. In the 32 occupations listed as universally licensed in Gittleman et al. (2018), we find 66.2 percent of workers are licensed, as compared to 13.2 percent of workers in the other 451 occupations. Conversely, 42.4 percent of workers who self-report that they are licensed work in universally licensed occupations, as compared to 5.4 percent of workers who self-report that they are not licensed. Both differences are highly significant and in the desired direction. Self-reported license status, and thereby the licensed share, does appear to be correlated with this measure of truth. A considerable fraction of workers do, however, self-report as unlicensed in universally licensed occupations.\(^{14}\) Due to considerations discussed above, it is hard to determine whether or not such self-reports are misresponses. In our main sample, we exclude workers in universally licensed occupations, but in the Appendix, we show our results are robust to their inclusion.

To address finite-sample bias (Goldsmith-Pinkham et al., 2018) and reduce sampling error in cells with few observations, we estimate licensed shares using the leave-out mean with an empirical Bayes adjustment:

\[
\%\text{License}_i = \frac{\hat{\alpha}_o + \sum_{j \in W_{os} : j \neq i} \text{License}_j}{\hat{\alpha}_o + \hat{\beta}_o + N_{os} - 1},
\]

where worker \(j\) is in the set \(W_{os}\) if and only if \(j\) is in occupation \(o\) in state \(s\). \(N_{os}\) is the number of such workers. The terms \(\hat{\alpha}_o\) and \(\hat{\beta}_o\) are occupation-specific constants that are derived from a beta-binomial model that we explain in Appendix D; they reduce measurement error by using prior

\(^{14}\)This is true even in occupations that are narrowly defined in the Census. For example, only 65.9 percent of workers who are “licensed practical and licensed vocational nurses” (occupation code 3500) self-report as licensed.
knowledge of each occupation’s distribution of cell licensed shares to efficiently shrink the raw cell licensed shares towards the national share. To estimate attenuation bias from sampling variation, we calculate for each cell the standard error of the licensed share using the standard deviation of the posterior distribution:

$$\sigma_{u_i} = \sqrt{\frac{\left(\hat{\alpha}_0 + \sum_{i' \in W_{os}, i' \neq i} \text{License}_{i'}\right)\left(\beta_0 + N_{os} - 1 - \sum_{i' \in W_{os}, i' \neq i} \text{License}_{i'}\right)}{\left(\hat{\alpha}_0 + \hat{\beta}_0 + N_{os} - 1\right)^2}}.$$

Bolstered by our empirical Bayes approach, we have sufficient data to offer precise estimates of licensed shares: The median worker is in a state–occupation cell whose licensed share has a standard error of 1.7 p.p., and the standard error for the 95th-percentile worker’s cell is 4.7 p.p. Appendix Table A2 shows that about 90 percent of variation in the licensed share is between occupations. By comparison, between-state variation is negligible (<1 percent). The remaining 10 percent is our identifying variation, and the standard deviation of these residuals is about 7.1 percentage points. Taken together, these results imply an attenuation bias on the order of 7 percent, which we henceforth ignore in our analysis.

We also use other CPS data on worker characteristics, some as outcomes and others in our standard set of controls. These are the hourly wage (from the Merged Outgoing Rotation Group sample), hours worked last week, age, educational attainment, sex, race (white, black, Asian, other), ethnicity (Hispanic and non-Hispanic), and indicators for certification status, union status (covered and non-covered), veteran status, marital status, disability status (any physical or cognitive), and metropolitan status (MSA resident or non-resident), and the presence of children at home. Throughout our analysis, we treat worker age, sex, race, ethnicity, veteran status, marital status, disability status, metropolitan status, and the presence of children as demographic characteristics which are predetermined with respect to licensing and thus use them in our controls. For our analysis of the opportunity costs of licensing, we restrict controls to worker sex, race, and ethnicity.

Splitting the sample on individual license status, we report summary statistics for these demographic variables in Appendix Table A3. Licensed and unlicensed workers differ along nearly every observable characteristic: The licensed are older, more educated, more likely to be female, married, non-Hispanic white, union members, U.S. citizens, non-disabled, veterans, and earn about 30 percent more than the unlicensed on average. The pervasive differences between licensed and unlicensed workers represent a key challenge in identifying causal effects of licensing. Using cell licensed shares largely frees us from problem of selection into licensing which threaten causal interpretations of previous research.

The adjustment is only of consequence for estimating licensed shares in cells with very few observations. See Appendix D. Results are similar without the correction and particularly so if we simply drop such small cells.
3.2 Empirical Strategy

We use variation in the state–occupation cell licensed share to estimate effects of licensing which correspond to reduced-form moments of our model. We estimate specifications of the form

\[ y_i = \alpha_o + \alpha_s + \beta \cdot %\text{Licensed}_{i(o,s)} + X_i' \theta + \varepsilon_i, \quad (10) \]

where \( \alpha_o \) and \( \alpha_s \) are occupation and state fixed effects and \( \beta = \gamma \tau \) is the average effect of licensing for some outcome \( y_i \) for worker \( i \). The independent variable \( %\text{Licensed}_{i(o,s)} \) is the estimated licensed share of workers in the same occupation and state as worker \( i \). The state and occupation fixed effects mean we identify the effect of licensing from occupations for which licensed shares of workers differ among states. In controls \( X_i \), we include fixed effects for the demographic strata as well as industry and survey month–year fixed effects. We cluster standard errors by cell, which we define to be a state–occupation pair.

This specification identifies effects of licensing by a two-way comparison of a state–occupation cell to the same occupation in other states and other occupations in same state. Abstracting from covariates, the formal identification assumption for \( \beta \) is that two-way differences in untreated potential outcomes are independent of two-way differences in licensed shares for any two occupations \( o_1, o_2 \) and any two states \( s_1, s_2 \):

\[ [y_{o_1, s_1}(0) - y_{o_2, s_1}(0) - y_{o_1, s_2}(0) + y_{o_2, s_2}(0)] \perp \%L_{o_1, s_1} - %L_{o_2, s_1} - %L_{o_1, s_2} + %L_{o_2, s_2}, \]

where \( y_{o_1}(0) = E[y_i(0) | i \in W_{o,s}] \) and \( y_{i}(0) \) is the outcome of interest for individual \( i \) if their cell was totally unlicensed. Relative to all occupations in a state and the occupation in all states, cell licensed shares must therefore be uncorrelated with unobservable determinants of the outcome of interest. Following de Chaisemartin and D'Haultfoeuille (2019), the estimator can be written as a weighted average of heterogenous treatment effects \( \Delta_{o,s} \) of licensing occupation \( o \) in state \( s \), weighted by the \( \omega_{o,s} \) terms:

\[ \beta = \sum_{o,s} \omega_{o,s} \Delta_{o,s} \]

where

\[ \Delta_{o,s} = E[y_i(1) - y_i(0) | i \in W_{o,s} : L_i = 1] \]

\[ \omega_{o,s} = \frac{\lambda_{o,s} \%L_{o,s}(\%L_{o,s} - \%L_{o} - \%L_{s} + \%L)}{\sum_{o,s} \lambda_{o,s} \%L_{o,s}(\%L_{o,s} - \%L_{o} - \%L_{s} + \%L)}, \]

with \( \lambda_{o,s} \) representing cell worker shares. For concision, we write \( L_i = 1 \) if worker \( i \) is licensed and \( %L_{(\cdot)} \) for a licensed share. Importantly, our approach requires variation in licensing shares within an occupation between U.S. states, and so our results do not pertain to occupations that are licensed by essentially all or no U.S. states. We identify instead an average treatment effect that approximates...
Table 1: For Which Occupations Does Licensing Vary Among U.S. States?

<table>
<thead>
<tr>
<th>Occupation</th>
<th>% Licensed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td><strong>Panel A: High Interstate Variance</strong></td>
<td></td>
</tr>
<tr>
<td>Brokerage Clerks</td>
<td>40.0</td>
</tr>
<tr>
<td>Dispensing Opticians</td>
<td>30.8</td>
</tr>
<tr>
<td>Elevator Installers</td>
<td>41.4</td>
</tr>
<tr>
<td>Electricians</td>
<td>43.9</td>
</tr>
<tr>
<td><strong>Panel B: Low Interstate Variance</strong></td>
<td></td>
</tr>
<tr>
<td>Lawyers</td>
<td>82.8</td>
</tr>
<tr>
<td>Registered Nurses</td>
<td>83.2</td>
</tr>
<tr>
<td>Economists</td>
<td>1.6</td>
</tr>
<tr>
<td>Cashiers</td>
<td>2.1</td>
</tr>
</tbody>
</table>

Notes: This table presents statistics on selected occupations with high or low variance in state-occupation licensed shares. In particular, we report their Census occupation code, their estimated average annual employment in our sample, the estimated national licensed share, and the sample-weighted standard deviation of the state-occupation licensed shares. See Appendix Table A4 for occupations ranked by their treatment-effect weights as in de Chaisemartin and D’Haultfoeuille (2019) and the most overweighted occupations relative to their population share.

the quantity relevant for policy analysis, insofar as weights $\omega_{os}$ are large for occupations with much between-state “disagreements” in licensing that may reflect areas of interest.\(^\text{16}\)

Which occupations have interstate variation in licensing and thus contribute most to empirical identification? Table 1 provides guidance. Panels A and B respectively list four occupations with high and low interstate variance in their licensed share. Many salient licensed occupations are universally licensed (and thus explicitly excluded from our sample, but included here) or have low interstate variance in the licensed share (and thus receive little weight). A characteristic marginally licensed occupation is the dispensing optician (Timmons and Mills, 2018): It is licensed by 21 U.S. states but unlicensed by 29. Though related to two health professions with universal licensing, ophthalmology and optometry, opticians’ scope of practice is narrower: They cannot diagnose eye diseases or perform eye examinations but can dispense eyeglasses and contact lenses. In such occupations, it is unclear whether the social gains from licensing compensate for its social costs.

4 Reduced-Form Effects of Licensing

Our reduced-form empirical analysis proceeds in several steps. First, we present evidence that suggests that licensing regulations have substantial bite: that is, their costs appear on average

\(^{16}\)In any two-way fixed effect design (de Chaisemartin and D’Haultfoeuille, 2019), weights $\omega_{os}$ may be positive or negative. In our application, all occupations receive on average a positive weight, and so $\beta$ is a convex combination of occupation-specific treatment effects. However, an assumption that within-occupation treatment effects of licensing are homogeneous between states is comparatively necessary, as 30.3 percent of treatment-effect weights $\omega_{os}$ are negative, where we calculate this fraction weighting by $|\omega_{os}|$. See Appendix C for a list of occupations by their regression weight and fully-interacted results that eliminate negative weights.
economically significant as a share of workers’ present value lifetime incomes. Second, we show that licensing raises average wages, compensating in part for licensing costs. Third, we show labor supply increases on the intensive margin, but contracts on the extensive margin, consistent with the combination of licensing costs and higher wages.

4.1 Education and the Opportunity Cost of Licensing

We present several pieces of evidence consistent with economically significant licensing costs, motivating our subsequent analysis of wage and labor supply responses. First, we show that licensing’s education requirements appear to bind, raising average investment in education. Second, we show that licensing reallocates human capital investment towards occupation-specific credentials. Third, we show that licensing appears to delay the entry into employment of young workers.

In Panel A of Table 2, we estimate effects of licensing on mean years of education and find that workers in highly-licensed cells, relative to that occupation in other states and other occupations in that state, have substantially more education than workers in less-licensed cells. Our estimate in Column 3, in particular, implies that fully licensing a cell raises mean education by 0.4 years.

Second, licensing reallocates human capital investment. Figure 1 displays the effects of licensing on shares of educational attainment by degree level, using our two-way fixed effect specification to compare distributions of educational attainment in cells with high and low licensed shares. We see a striking pattern: Licensing increases the shares of workers with more occupation-specific forms of educational credentials, such as occupational or vocational associates degrees or masters’ degrees, and decreases the shares of workers with educational credentials that are not specific to occupations, such as high school degrees or bachelor’s degrees. These results are consistent with actual licensing policies, a majority of which impose specialized educational requirements (Gittleman et al., 2018). Our estimates are noteworthy in magnitude, comparable to the G.I. Bill (Bound and Turner, 2002) or modern grant-aid programs (Dynarski, 2003). We summarize the extent of reallocation by estimating the total variation distance from fully licensing an unlicensed occupation, which represents the minimum share of workers in an occupation whose education level changes as a result of licensing policy: 10.7 percent.\footnote{For discrete random variables $X, Y$ over event space $\Omega$, variation distance equals $\frac{1}{2} \sum_{x \in \Omega} |P(X = x) - P(Y = x)|$. Our estimate reflects a bias correction for the effect of sampling error on estimated variation distance that we explain in Appendix D. This correction is inconsequential in magnitude for our application.}

Licensing thus substantially increases the occupational specificity of human capital.

The CPS definition of education, however, excludes much training required by licensing. For instance, legal entrance into the occupation of cosmetology requires, in a majority of U.S. states, instructional or apprenticeship programs requiring at least 1,500 work hours (Reddy, 2017). To assess the full opportunity cost of licensing—which, in our model, is the delay in entry to employment due to mandated training—we also consider worker age as an outcome. In particular, we estimate
Notes: This figure presents estimates from Equation 10 of the effects of licensing on the shares of workers in a cell by their highest level of educational attainment. Standard errors are clustered at the state–occupation cell level. Bars reflect 95-percent robust confidence intervals.

the horizontal shift in the age profile of employment with a specification

$$\text{Age}_{os,a} = \alpha_{o,a} + \alpha_{s,a} + \beta \cdot \%\text{Licensed}_{os} + \delta \log \text{Emp}_{os,a} + \varepsilon_{os,a}$$,

where $a$ is the worker age (so $\text{Age}_{os,a} = a$), $o$ is the occupation, and $s$ is the state. Therefore, $\alpha_{s,a}$ and $\alpha_{o,a}$ are respectively state–age and occupation–age fixed effects, and $\text{Emp}_{os,a}$ is the employment count in occupation $o$ and state $s$ for workers of age $a$. To focus on entry to employment, we restrict the sample to workers below age 35.

Panel B of Table 2 reports that licensing delays the entry to employment by about 1.1 years. This suggests that time in formal education indeed understates the opportunity cost of licensing. We also directly examine the effect of licensing on the age profile of employment, using a Poisson regression specification of Equation 10 that splits cell employment counts by worker age in years:

$$E[\text{Emp}_{os,a}] = \exp(\alpha_{o,a} + \alpha_{s,a} + \beta \cdot \%\text{Licensed}_{os})$$.

Figure 2 shows that there are fewer young workers in highly licensed state–occupation cells relative
Figure 2: Effect of Licensing on Employment Age Profile

Notes: This figure shows the estimated effects of licensing on the number of workers by age in a state–occupation cell as estimated by Equation 11. Gray dashed lines indicate 95-percent confidence interval with standard errors clustered at the level of the state–occupation cell.

To the same occupation in other states where the licensed share is lower, consistent with delayed worker entry into occupations. Employment of workers who are 25 years old or younger, for example, falls by 46 percent on average. Consequently, the opportunity costs of licensing appear substantial and reflective of time spent in formal education as well as unmeasured investments.

4.2 Wages

To what extent are workers compensated in equilibrium for licensing costs via higher wages? Panel B of Table 2 reports estimated wage effects of licensing. Column 1 reports the specification with demographic-strata controls and with individual license status as the treatment variable. Comparing the average hourly wages of observably similar licensed and unlicensed workers after state and occupation fixed effects, we find licensed workers earn about 15 percent more per hour than unlicensed workers.

This comparison is vulnerable to selection on unobservables of workers into licensing according to correlates of the wage. Column 2 replaces individual license status with the licensed share. We thus identify the wage effect of licensing using state–occupation variation in licensing rates, purging the comparison of within-cell selection. As occupations that are highly licensed in a state relative to the state’s overall licensing rate and the occupation’s overall licensing rate also pay relatively high
Table 2: Reduced-Form Worker Effects of Occupational Licensing

<table>
<thead>
<tr>
<th></th>
<th>Licensed = 1</th>
<th>% Licensed in Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Panel A: Years of Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.383***</td>
<td>0.418***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,865,209</td>
<td>1,865,209</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,321</td>
<td>20,321</td>
</tr>
<tr>
<td><strong>Panel B: Years of Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.282***</td>
<td>1.135***</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.243)</td>
</tr>
<tr>
<td>Observations</td>
<td>722,168</td>
<td>722,168</td>
</tr>
<tr>
<td>Clusters</td>
<td>17,842</td>
<td>17,842</td>
</tr>
<tr>
<td><strong>Panel C: Log Hourly Wage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.159***</td>
<td>0.201***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Observations</td>
<td>317,142</td>
<td>317,142</td>
</tr>
<tr>
<td>Clusters</td>
<td>18,753</td>
<td>18,753</td>
</tr>
<tr>
<td><strong>Panel D: Log Weekly Hours Per Worker</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.039***</td>
<td>0.044***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,865,209</td>
<td>1,865,209</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,321</td>
<td>20,321</td>
</tr>
<tr>
<td><strong>Panel E: Log Employment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.294***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.065)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from Equation 10 of effects of licensing on outcomes of interest which correspond to reduced-form moments of the model. The estimate in Column 1 refers to individual-worker licensing status, whereas those in Columns 2 and 3 refer to the state–occupation cell licensed share of workers. In Columns 1 and 3, we include strata fixed effects for predetermined demographic observables. All specifications include fixed effects for occupation, state, industry, and month, except in Panel E, which has only state and occupation fixed effects. Standard errors are clustered at the level of the state–occupation cell. Appendix Table A6 includes universally licensed occupations. ∗∗∗ = p < 0.01.

wages, the comparison finds positive wage effects of licensing. In Column 3, our baseline estimate of the causal effect of licensing on wages, we reintroduce the demographic strata controls and thus hold constant a list of predetermined covariates potentially related to wages. We find licensing raises wages by 15 percent in this specification.
4.3 Hours and Employment

If licensing raises the hourly wage but reduces the present value of net income, as Proposition 1 explains will occur when licensing has little effect on WTP, licensing should raise hours per worker but reduce employment. Panel B of Table 2 reports effects of licensing on log weekly hours per worker. Columns 1 to 3 find that licensing increases average hours in the state–occupation cell by about 3 to 4 percent. Reassuringly, the ratio of our estimated hours and wage responses to licensing are near benchmark estimates of the intensive-margin labor supply elasticity (Chetty, 2012). Panel A of Appendix Table A5 repeats these specifications using the level of hours and finds increases of about 1.4 to 1.8 hours per week attributable to licensing.

To evaluate employment effects of licensing, we calculate sample-weighted employment counts by state–occupation cell and regress the log cell count on the state–occupation licensed share:

$$\log \text{Emp}_{os} = \beta \cdot \%\text{Licensed}_{os} + \alpha_o + \alpha_s + \varepsilon_{os}.$$  

We report these results in Panel C. Across specifications, we estimate a significant disemployment effect of licensing of around 29 percent. Relative to the same occupation in other states and to other occupations in the same state, highly-licensed cells also have considerably lower employment than less-licensed cells. As our employment regressions cannot be meaningfully estimated at the worker level or with worker-level controls, we present only one estimate in Column 2 of Table 2.

These results, however, survive several checks. First, we estimate a Poisson regression specification on the employment counts and, as reported in Panel B of Appendix Table A5, we find similar employment declines. Second, in Appendix Table A7, we repeat this exercise with American Community Survey (ACS) microdata to calculate employment shares while using our CPS-based measure of licensing. We find a disemployment effect of about 25 percent.

5 Robustness Checks

Our research design identifies effects of licensing on outcomes using differences in licensed shares across states and occupations. A confounding variable must therefore correlate with the outcome of interest and the licensed share in a state–occupation cell relative to other cells in the same state or in the same occupation after controls.

One such variable is the existence of other labor market policies whose presence is correlated with licensing (Besley and Case, 2000), which could bias our results in either direction. We are unaware of comprehensive measures of such policies at the state–occupation level and thus cannot decisively evaluate the claim. Certification and unionization, however, could plausibly substitute or complement licensing in such a fashion. We add controls for the state–occupation certification and unionization rates to our baseline wage specification and report results in Column 1 of Table 3. These cell rates are produced using the same beta-binomial method described in Section 3. Certification and unionization controls do not much alter our estimates.
## Table 3: Robustness Checks

<table>
<thead>
<tr>
<th></th>
<th>(1) Unions &amp; Cert.</th>
<th>(2) Occ. &amp; Demo. Mix</th>
<th>(3) State–Occ. Group FE</th>
<th>(4) Div–Occ. FE</th>
<th>(5) Drop ± 1 SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Years of Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>0.410***</td>
<td>0.366***</td>
<td>0.308***</td>
<td>0.255***</td>
<td>0.372***</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.056)</td>
<td>(0.053)</td>
<td>(0.053)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,865,209</td>
<td>1,859,356</td>
<td>1,865,209</td>
<td>1,865,206</td>
<td>364,166</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,321</td>
<td>19,470</td>
<td>20,321</td>
<td>20,318</td>
<td>8,283</td>
</tr>
<tr>
<td><strong>Panel B: Years of Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>1.137***</td>
<td>1.152***</td>
<td>0.941***</td>
<td>0.615**</td>
<td>0.901***</td>
</tr>
<tr>
<td></td>
<td>(0.244)</td>
<td>(0.243)</td>
<td>(0.237)</td>
<td>(0.260)</td>
<td>(0.233)</td>
</tr>
<tr>
<td>Observations</td>
<td>722,168</td>
<td>720,225</td>
<td>722,168</td>
<td>722,143</td>
<td>130,029</td>
</tr>
<tr>
<td>Clusters</td>
<td>17,842</td>
<td>17,397</td>
<td>17,842</td>
<td>17,817</td>
<td>6,555</td>
</tr>
<tr>
<td><strong>Panel C: Log Hourly Wage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>0.124***</td>
<td>0.158***</td>
<td>0.138***</td>
<td>0.146***</td>
<td>0.161***</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.024)</td>
<td>(0.024)</td>
<td>(0.024)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Observations</td>
<td>317,142</td>
<td>316,123</td>
<td>317,141</td>
<td>317,045</td>
<td>62,336</td>
</tr>
<tr>
<td>Clusters</td>
<td>18,753</td>
<td>18,164</td>
<td>18,752</td>
<td>18,657</td>
<td>7,203</td>
</tr>
<tr>
<td><strong>Panel D: Log Weekly Hours Per Worker</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>0.031***</td>
<td>0.034***</td>
<td>0.028***</td>
<td>0.024**</td>
<td>0.026**</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,865,209</td>
<td>1,859,356</td>
<td>1,865,209</td>
<td>1,865,206</td>
<td>364,166</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,321</td>
<td>19,470</td>
<td>20,321</td>
<td>20,318</td>
<td>8,283</td>
</tr>
<tr>
<td><strong>Panel E: Log Employment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>-0.320***</td>
<td>-0.176***</td>
<td>-0.084</td>
<td>-0.193***</td>
<td>-0.172***</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.062)</td>
<td>(0.052)</td>
<td>(0.052)</td>
<td>(0.059)</td>
</tr>
<tr>
<td>Observations</td>
<td>20,524</td>
<td>19,481</td>
<td>20,524</td>
<td>20,435</td>
<td>7,596</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from variations on Equation 10 as explained in the text. All estimates refer to the coefficient on the licensed share of workers in the state–occupation cell. All specifications include fixed effects for occupation, state, industry, and month, except in Panel E, which has only state and occupation fixed effects. Standard errors are clustered at the level of the state–occupation cell. Appendix Table A8 includes universally licensed occupations. ∗ = p < 0.10, ∗∗ = p < 0.05, ∗∗∗ = p < 0.01.

In Column 2 of Table 3, we add two controls for predicted employment to Equation 10. The first control is a low-dimensional representation of the state occupational mix. In summary, we use principal component analysis to extract a vector of state labor market characteristics that explain variation across states in occupational employment shares that, a priori, we do not expect to be explained by licensing. The second is a Bartik-like control that removes the predictive power of the state demographic mix for occupational employment shares. Appendix D develops both controls in detail. Motivating these controls, it would be concerning if, for instance, general patterns such
as whether a state had high or low relative employment shares in occupations related to the rural economy or in occupations predominantly held by nonwhites were driving our identification. We find, reassuringly, that our estimates are essentially unchanged by these controls, even though they explain fully one quarter of the residual variation in occupational employment shares.

Columns 3 and 4 of Table 3 present the next set of robustness checks, in which we restrict identifying variation to related groups of states and occupations. In Column 4 we add fixed effects for the interaction of the state and Census detailed occupational group to our specification.\(^18\) We now identify the effect of licensing only from variation in licensing rates and wages within cells defined by the state and a group of similar occupations. Our results are mostly unchanged, though our estimated employment effect falls and becomes insignificant. In Column 5, we restrict the comparison to occupations within groups of states in the same Census geographic division by adding division–occupation fixed effects.\(^19\) Our estimates are essentially unchanged, suggesting that division-specific occupational differences do not confound our estimates of the effects of licensing. These comparisons bolster our results insofar as states in the same Census division, or occupations in the same Census occupational group, serve as more credible counterfactuals than pooling all U.S. states or all occupations.\(^20\)

Compared to a direct measure of policy, the licensed share of workers among cells as a measure of treatment comes at some cost to clarity in the underlying sources of variation in treatment. To build confidence that true variation in policy drives our results, we drop from our sample all state–occupation cells whose residual licensed share (i.e., after removing state and occupation fixed effects) is within one standard deviation of zero. That is, here we use an approximately one-third subsample with large differences in the licensed share of workers, as these are arguably most suggestive of the policy variation we wish to exploit for identification, rather than sampling variation or other considerations. Our results, reported in Column 6 of Table 3, are unchanged. Using only the large differences in cell licensed shares most suggestive of policy variation does not change the estimated wage effects of licensing.

Appendix C reports more robustness checks. We examine the sensitivity of our results to fixed effects by state and Census major occupational category, occupation- and state-specific effects of demographic characteristics, a more flexible specification of the two-way fixed effect structure, and

\(^18\)The regression equation is \(y_i = \alpha_o + \alpha_s + \gamma_{gs} + \beta \cdot \%\text{Licensed}_i + X_i' \theta + \varepsilon_i\), noting the subscripts on \(\gamma_{gs}\) indicate the coefficients are specific to a group–state pair, and so \(\beta\) is identified from “within” variation alone. Occupations assigned one of 10 major groups (e.g., “professional and related occupations”) and one of 23 detailed groups (e.g., “legal occupations”). For further detail, see Appendix B of the CPS March Supplement documentation.

\(^19\)The regression equation is \(y_i = \alpha_o + \alpha_s + \gamma_{od} + \beta \cdot \%\text{Licensed}_i + X_i' \theta + \varepsilon_i\), noting the subscripts on \(\gamma_{od}\) indicate the coefficients are specific to a division–occupation pair. The U.S. Census divides states into 9 divisions: New England, South Atlantic, Middle Atlantic, East North Central, West North Central, East South Central, West North Central, West South Central, Mountain, and Pacific. Divisions contain between 3 and 8 states.

\(^20\)These robustness checks can equally be interpreted respectively as checks against between-occupation, within-state and within-state, between-occupation spillovers from licensing, which would represent a violation of the the stable unit treatment value assumption (Rubin, 1978). We think spillovers of detectable magnitude are a priori unlikely here, as workers are not at all concentrated in a few occupations in our sample. To the extent spillovers matter, these estimates will be larger in absolute magnitude than our baseline results, and yet we find none increase notably.
a direct test for political endogeneity of licensing.

6 Welfare Effects and Incidence of Licensing

We translate the reduced-form estimates into welfare impacts in two ways using our model. First, we proceed by the sufficient statistics for worker and consumer welfare. Second, we structurally estimate the model. We view these approaches as complements, insofar as they reveal precisely when we require further assumptions to map from reduced-form responses to licensing to welfare and structural parameters.

In our sufficient-statistics approach, welfare effects rescale reduced-form responses of occupational employment and the wage bill to licensing, letting us move transparently from data to welfare. A structural approach, however, lets us say more about the welfare impacts of licensing at the price of two calibration assumptions, which we examine carefully. Structural estimation allows us, in particular, to decompose the reduced-form responses into effects of licensing occupational labor supply and labor demand; we also can examine the plausibility of the vector of estimated structural parameters implied by our reduced-form results. The gap between these two approaches reflects the fact that licensing shifts both occupational labor supply and demand, and our approach does not separately identify supply and demand elasticities—and thus not the gross shifts of supply and demand—but neither are these gross shifts necessary for welfare analysis. The welfare effects of licensing depend on whether increases in labor demand are large enough to offset reductions in labor supply on net, as in Summers (1989) and as we formalize in Proposition 4.

6.1 Welfare Analysis from Reduced-Form Estimates

Proposition 3 shows that, in our model, the reduced-form effect of licensing on occupational employment and the wage bill reveal the effects of licensing on worker and consumer welfare respectively. In Section 4, we estimate licensing reduces occupational employment. This implies licensing reduces worker welfare, with the implied worker welfare losses increasing in \( \sigma \), which moves inversely with occupational preference dispersion. We also find in Section 4 that licensing raises the average wage and weekly hours, but by amounts less than the employment decline. This implies that licensing reduces the occupational wage bill, though this estimate is imprecise, and therefore that licensing reduces consumer welfare. These consumer welfare losses are declining in the occupational labor demand elasticity \( \varepsilon \). Taken together, it is immediate from our reduced-form findings that licensing reduces both occupational employment and the wage bill that licensing in marginal occupations reduces social welfare.

6.2 Structural Estimation

We use the classical minimum distance estimator (Newey and McFadden, 1994) to estimate a vector structural parameters \( \theta \) that, by the mapping \( m(\cdot) \) implied by our model, best matches a vector
of reduced-form empirical moments $\hat{\beta}$ as weighted by the inverse of variance matrix $\hat{V}$. These estimated structural parameters are given by

$$\hat{\theta} = \arg \min_{\theta} \left\{ (\hat{\beta} - m(\theta))' \hat{V}^{-1} (\hat{\beta} - m(\theta)) \right\}. \quad (12)$$

The vector of reduced-form empirical moments $\hat{\beta}$ contains the four main results of Section 4, which are the effects of licensing on wages, hours per worker, employment, and the worker age profile. These moments just-identify four structural parameters: the return on education $\rho$, the intensive-margin labor supply elasticity $\eta$, the average required training time $\tau$, and the WTP effect $\alpha$.

We calibrate the two remaining structural parameters, which are the dispersion of occupational preferences $\sigma$ and the elasticity of occupational labor demand $\varepsilon$, from the literature. Following estimates in Hsieh et al. (2013) and Cortes and Gallipoli (2017) of occupational preference dispersion of U.S. workers, we consider values of $\sigma \in \{2, 3, 4\}$.\footnote{In our model, this range of calibrations implies a range of elasticities of the occupational employment share with respect to the present value of net income between 2.5 and 5. Direct evidence of this elasticity is limited, but see Powell and Shan (2012).} For estimates of the occupational labor demand elasticity $\varepsilon$, we look to the survey of Hamermesh (1996) and consider values of $\varepsilon \in \{1.5, 2, 2.5\}$, with the view that such an elasticity should be above the skilled–unskilled labor substitution or local labor demand elasticities in Autor et al. (1998) and Kline and Moretti (2013). We provide a constructive proof of identification in Appendix B which shows that our estimates of $\eta$ and $\tau$ are independent of our calibrated $\sigma$ and $\varepsilon$, but the calibration does matter for $\alpha$ and $\rho$.

After partialling out fixed effects and controls from our four outcomes and the licensed share, our model yields four linear moment conditions:

\[
\begin{bmatrix}
\hat{w}_j \\
\hat{h}_i \\
\hat{s}_j \\
\hat{a}_i \\
\end{bmatrix} = \begin{bmatrix}
\hat{\beta} \\
\end{bmatrix} = \begin{bmatrix}
\frac{\hat{w}_j \cdot \%\text{Licensed}_j}{1 + \sigma(1 + \eta) + \eta \varepsilon} \\
\frac{\rho \sigma (1 + \eta) + \alpha \eta \varepsilon}{\rho \sigma (1 + \eta) / \eta + \alpha \varepsilon} \\
\frac{\rho \sigma (1 + \eta) / \eta + \alpha \varepsilon}{\sigma (1 + \eta) (\alpha \varepsilon - \rho (\varepsilon + 1 / \eta))} \\
\frac{1 + \sigma(1 + \eta) + \eta \varepsilon}{1 + \sigma(1 + \eta) + \eta \varepsilon} \\
\end{bmatrix} \cdot m(\theta).
\]

Our approach to structural estimation is to consider the vector of structural parameters that would rationalize our reduced-form results in Section 4 while imposing minimal additional assumptions. In discussing our parameter estimates, therefore, we provide relevant benchmarks to assess whether they are reasonable. Furthermore, the sufficient-statistics result of Proposition 3 implies that our structural estimates only affect welfare and incidence through their implications for the responses of employment and the wage bill to licensing, which are pinned down by the reduced-form estimates. Our aim is to make as transparent as possible these steps from the reduced-form estimates to the welfare analysis of licensing.

Table 4 displays the results of the structural estimation for the various calibrations of $\sigma$ and
Table 4: Structural Estimates

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occ. Pref. Dispersion (σ)</strong></td>
<td>3.00</td>
<td>2.00</td>
<td>4.00</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td><strong>Demand Elasticity (ε)</strong></td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>1.50</td>
<td>2.50</td>
</tr>
</tbody>
</table>

**Panel A: Estimated Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensive Margin Elasticity (1/η)</td>
<td>0.199**</td>
<td>0.199**</td>
<td>0.199**</td>
<td>0.199**</td>
<td>0.199**</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.081)</td>
<td>(0.081)</td>
<td>(0.081)</td>
<td>(0.081)</td>
</tr>
<tr>
<td>Return to Schooling (ρ)</td>
<td>0.164***</td>
<td>0.189***</td>
<td>0.151***</td>
<td>0.164***</td>
<td>0.164***</td>
</tr>
<tr>
<td></td>
<td>(0.062)</td>
<td>(0.071)</td>
<td>(0.057)</td>
<td>(0.062)</td>
<td>(0.062)</td>
</tr>
<tr>
<td>WTP Effect (α)</td>
<td>0.035</td>
<td>0.035</td>
<td>0.035</td>
<td>0.009</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.031)</td>
<td>(0.031)</td>
<td>(0.034)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Licensing Cost in Years (τ)</td>
<td>1.350***</td>
<td>1.350***</td>
<td>1.350***</td>
<td>1.350***</td>
<td>1.350***</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.478)</td>
<td>(0.478)</td>
<td>(0.478)</td>
<td>(0.478)</td>
</tr>
</tbody>
</table>

**Panel B: Welfare Effects**

| Worker                   | -0.081***| -0.121***| -0.061***| -0.081***| -0.081***|
|                         | (0.018)  | (0.028)  | (0.014)  | (0.018)  | (0.018)  |
| Consumer                | -0.070   | -0.070   | -0.070   | -0.141   | -0.047   |
|                         | (0.076)  | (0.076)  | (0.076)  | (0.152)  | (0.051)  |
| Social                  | -0.151   | -0.192*  | -0.131   | -0.222   | -0.128*  |
|                         | (0.093)  | (0.102)  | (0.089)  | (0.169)  | (0.068)  |

**Panel C: Incidence Analysis**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost as Share of Income (\tilde{t})</td>
<td>0.221***</td>
<td>0.255***</td>
<td>0.204***</td>
<td>0.221***</td>
<td>0.221***</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.034)</td>
<td>(0.027)</td>
<td>(0.029)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Share of Cost Offset</td>
<td>0.695***</td>
<td>0.603***</td>
<td>0.752***</td>
<td>0.695***</td>
<td>0.695***</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.067)</td>
<td>(0.052)</td>
<td>(0.059)</td>
<td>(0.059)</td>
</tr>
<tr>
<td>WTP-Adj. Price Change</td>
<td>0.059</td>
<td>0.027</td>
<td>0.059</td>
<td>0.059</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.066)</td>
<td>(0.063)</td>
<td>(0.063)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Share of Price Change Offset</td>
<td>0.599</td>
<td>0.599</td>
<td>0.599</td>
<td>0.153</td>
<td>0.867</td>
</tr>
<tr>
<td></td>
<td>(1.151)</td>
<td>(1.151)</td>
<td>(1.151)</td>
<td>(0.751)</td>
<td>(1.394)</td>
</tr>
</tbody>
</table>

*Notes: This table reports structural parameters \( \hat{\theta} \) as estimated by Equation 12 in Panel A, welfare effects on workers and consumers in Panel B, and incidence analysis in Panel C. The sample pools the MORG and full CPS sample, using the earnings weights on the MORG sample and final person-level weights for the non-MORG sample. Standard errors are clustered at the level of the state-occupation cell. * = p < 0.10, ** = p < 0.05, *** = p < 0.01.
Panel A results report the structural parameter estimates. We estimate an intensive-margin labor supply elasticity $1/\eta = 0.20$, not far from the survey of Chetty (2012), which reaches a point estimate for $1/\eta = 0.33$. For the return to a year of schooling, we estimate a $\rho$ of about 16 percent, which is above estimates in Card (1999) but not implausible given the low level of education of most compliers (Heckman et al., 2018; Mountjoy, 2019). We estimate a mean training time $\tau$ of about 1.3 years, which is near the mean reported in the survey of licensing in low-wage occupations in Carpenter et al. (2012). Our estimates of the WTP effect $\alpha$ imply that, on average, one year of required training raises WTP by about 4 percent. To the best of our knowledge, there are no comparable estimates of WTP for licensing, but these estimates fit qualitatively with the small positive estimated effects of licensing on labor quality reviewed in Section 1.

In Panel B, we report the estimated welfare effects of licensing on workers and consumers. We find licensing makes workers significantly worse off and that consumer welfare declines insignificantly. Consumer welfare estimates are notably less precise than the worker welfare estimates. These welfare results follow, as above, from the reduced-form results that licensing reduces employment and that, combining our wage, hours, and employment estimates, the effect on the wage bill is negative but statistically insignificant. Taking these results together, the social welfare effects of licensing appear negative, although our estimates are imprecise due to the consumer side.

The incidence analysis in Panel C helps to interpret why licensing makes workers and consumers worse off. We estimate large opportunity costs of licensing—on the order of a fifth of the present value of lifetime income—and that workers are less than fully compensated for these opportunity costs by higher wages. In particular, wages offset about 60 to 75 percent of these costs, leaving workers worse off by about 25 to 40 percent of the opportunity cost. Consumer welfare also appears to decline, albeit insignificantly. Increases in WTP offset only about 60 percent of the increase in the price of labor services, leaving the quality-adjusted price level higher but not significantly so.

Motivated by Proposition 4, which provides a breakeven level of WTP gains at which the social costs and benefits of licensing are exactly equal, we can use our structural estimates to ask how far short $\hat{\alpha}$ falls of this breakeven. This empirical-to-breakeven ratio is

$$\frac{\hat{\alpha} \hat{\varepsilon}}{\hat{\rho}(\hat{\varepsilon} - 1)}$$

and in our baseline structural estimates, we find it to be about 0.4. This implies the true WTP effect would need to be more than twice as large as our empirical estimate of it for society to break even. Phrased differently, if we distrust welfare analysis from revealed preference, we would need a positive externality from licensing whose value in WTP terms exceeds its effect on private WTP.

## 7 Conclusion

We develop a new theoretical model of occupational licensing and empirical evidence on the effects of licensing to conduct a welfare analysis of licensing policies in U.S. states. We find that, on the
margin of occupations where policies differ across states, the average net social value of licensing appears negative: The social cost of reduced occupational labor supply appears to exceed the social benefit from higher WTP for labor from licensed occupations. We use these results to structurally estimate the model and to assess the incidence of licensing. Consumers and workers each appear to bear about half of the incidence: Increases in WTP do not fully offset the higher price of labor for consumers, nor do these wage increases do not fully compensate workers for licensing costs.

Within a broad class of models, the effects of licensing on employment and the wage bill are summary statistics for the welfare effect of licensing and respectively identify its incidence on workers in the occupation and on consumers. Our theoretical model also generates testable comparative statics for several labor market outcomes, which in principle only require data on a representative sample of workers, rather than data on product prices or quality, to be empirically evaluated. In our empirical analysis, we use variation in licensing policies across states and occupations as proxied by variation in the licensed share of workers. We find licensing raises average wages and hours per worker but reduces employment. Further results match prior expectations from the policy context and key comparative static predictions of the model: In particular, workers accumulate more occupation-specific human capital than they would absent licensing, delaying their entry to employment.

Two theoretical arguments exist for licensing. The first is about a missing technology: Absent licensing, workers may lack a credible signal of quality, leading to worker underinvestment in quality and excess entry. This argument is at the core of classic models of licensing, and it is the one we evaluate empirically in this paper. We find that, in marginal occupations, consumers appear to value the signal insufficiently to justify its social cost. However, we note this argument remains plausible for inframarginal occupations, such as those licensed by all U.S. states. Our theory offers a net-benefits test to assess licensing in such cases. The second argument is about externalities: There may be positive marginal social WTP for quality in some occupations, causing the private return on human capital to be inefficiently low and underinvestment even when workers’ quality is perfectly observable. As social WTP is not revealed by individual choices, we do not evaluate this argument here. A promising direction for future research, however, would be to translate existing evidence on quality effects of licensing to social WTP.

References


—, *Licensing Occupations: Ensuring Quality or Restricting Competition?*, WE Upjohn Institute, 2006.


Appendices for Online Publication

A  Additional Tables and Figures  
B  Model Appendix  
C  Further Results  
D  Econometric Extensions
### Additional Tables and Figures

**Table A1: Employed Population by License Status and Type**

<table>
<thead>
<tr>
<th>Has licensing or certification?</th>
<th>State issued?</th>
<th>Required for job?</th>
<th>Number</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
<td>No</td>
<td>452,667</td>
<td>0.725</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>23,713</td>
<td>0.038</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>37,026</td>
<td>0.059</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>7,052</td>
<td>0.011</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>104,239</td>
<td>0.167</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Notes:* This figure reports counts of unique employed workers according to their answers to Questions 1–3 as described in Section 3. Workers are here counted as answering affirmatively if they ever answer affirmatively while in the sample.
Table A2: Variance Components of License Status and State–Occupation Licensing Rate

<table>
<thead>
<tr>
<th>Component</th>
<th>Individual License Status</th>
<th>Licensing Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>0.002</td>
<td>0.005</td>
</tr>
<tr>
<td>Occupation</td>
<td>0.321</td>
<td>0.905</td>
</tr>
<tr>
<td>Residual</td>
<td>0.677</td>
<td>0.089</td>
</tr>
</tbody>
</table>

Notes: This table reports the results of a variance decomposition of individual license status and the state–occupation licensed rate in the CPS sample. For both variables, state fixed effects explain negligible shares of total variance, whereas occupation fixed effects explain considerable shares of variance, particularly after collapsing to state–occupation means.
### Table A3: Summary Statistics of Licensed and Unlicensed Workers

<table>
<thead>
<tr>
<th>Variable</th>
<th>Has state-issued occupational license</th>
<th>p-val. $(1) - (2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Variable</strong></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Age</td>
<td>40.33</td>
<td>35.90</td>
</tr>
<tr>
<td>Female</td>
<td>0.52</td>
<td>0.47</td>
</tr>
<tr>
<td>Married</td>
<td>0.52</td>
<td>0.39</td>
</tr>
<tr>
<td>Children at Home</td>
<td>0.47</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than HS</td>
<td>0.03</td>
<td>0.15</td>
</tr>
<tr>
<td>HS Graduate</td>
<td>0.21</td>
<td>0.30</td>
</tr>
<tr>
<td>Some College</td>
<td>0.32</td>
<td>0.29</td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td>0.24</td>
<td>0.19</td>
</tr>
<tr>
<td>More than Bachelor’s</td>
<td>0.19</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>0.77</td>
<td>0.74</td>
</tr>
<tr>
<td>Black</td>
<td>0.14</td>
<td>0.16</td>
</tr>
<tr>
<td>Asian</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Other</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.87</td>
<td>0.79</td>
</tr>
<tr>
<td>Citizen</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Lives in MSA</td>
<td>0.74</td>
<td>0.75</td>
</tr>
<tr>
<td>Paid by Hour</td>
<td>0.38</td>
<td>0.56</td>
</tr>
<tr>
<td>Hourly Wage</td>
<td>41.80</td>
<td>31.01</td>
</tr>
<tr>
<td>Weekly Labor Income</td>
<td>2,606.59</td>
<td>1,845.21</td>
</tr>
<tr>
<td>Union</td>
<td>0.14</td>
<td>0.07</td>
</tr>
<tr>
<td>Usually Full-Time</td>
<td>0.75</td>
<td>0.65</td>
</tr>
<tr>
<td>Any Disability</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Veteran</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Number of Workers</strong></td>
<td>74,086</td>
<td>470,905</td>
</tr>
</tbody>
</table>

**Notes:** This table reports summary statistics on the characteristics of unique workers by their licensing status according to the first survey month in the CPS. To be consistent across rows, only workers in the MORG are included in the sample.
### Table A4: Which Occupations Contribute Most to Empirical Identification?

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Code</th>
<th>Treat. Eff. Weight</th>
<th>Workers Per 10,000</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Most Influential Occupations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electricians</td>
<td>6355</td>
<td>0.0414</td>
<td>61.3</td>
<td>6.74</td>
</tr>
<tr>
<td>Nursing, psychiatric, and home health aides</td>
<td>3600</td>
<td>0.0282</td>
<td>146.2</td>
<td>1.93</td>
</tr>
<tr>
<td>Patrol officers</td>
<td>3850</td>
<td>0.0243</td>
<td>53.4</td>
<td>4.55</td>
</tr>
<tr>
<td>Pipe layers, plumbers, etc.</td>
<td>6440</td>
<td>0.0214</td>
<td>44.4</td>
<td>4.82</td>
</tr>
<tr>
<td>Teacher assistants</td>
<td>2540</td>
<td>0.0179</td>
<td>70.9</td>
<td>2.52</td>
</tr>
<tr>
<td>Construction managers</td>
<td>0220</td>
<td>0.0169</td>
<td>65.4</td>
<td>2.59</td>
</tr>
<tr>
<td>Social workers</td>
<td>2010</td>
<td>0.0151</td>
<td>58.1</td>
<td>2.60</td>
</tr>
<tr>
<td>Personal and home care aides</td>
<td>4610</td>
<td>0.0150</td>
<td>93.2</td>
<td>1.61</td>
</tr>
<tr>
<td>Dental assistants</td>
<td>3640</td>
<td>0.0143</td>
<td>22.1</td>
<td>6.48</td>
</tr>
<tr>
<td>Automotive service technicians and mechanics</td>
<td>7200</td>
<td>0.0137</td>
<td>67.1</td>
<td>2.04</td>
</tr>
<tr>
<td><strong>Panel B: Most Overweighted Occupations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brokerage clerks</td>
<td>5200</td>
<td>0.0014</td>
<td>0.3</td>
<td>42.63</td>
</tr>
<tr>
<td>Emergency management directors</td>
<td>0425</td>
<td>0.0030</td>
<td>0.7</td>
<td>40.66</td>
</tr>
<tr>
<td>Aircraft assemblers</td>
<td>7710</td>
<td>0.0013</td>
<td>0.5</td>
<td>27.16</td>
</tr>
<tr>
<td>Fire inspectors</td>
<td>3750</td>
<td>0.0046</td>
<td>1.7</td>
<td>26.94</td>
</tr>
<tr>
<td>Opticians, dispensing</td>
<td>3520</td>
<td>0.0098</td>
<td>3.7</td>
<td>26.10</td>
</tr>
<tr>
<td>Explosives workers</td>
<td>6830</td>
<td>0.0018</td>
<td>0.7</td>
<td>25.74</td>
</tr>
<tr>
<td>Manufactured building and home installers</td>
<td>7550</td>
<td>0.0013</td>
<td>0.5</td>
<td>24.91</td>
</tr>
<tr>
<td>Funeral service workers</td>
<td>4460</td>
<td>0.0017</td>
<td>0.7</td>
<td>24.85</td>
</tr>
<tr>
<td>Ambulance drivers and attendants, ex. EMTs</td>
<td>9110</td>
<td>0.0025</td>
<td>1.0</td>
<td>24.50</td>
</tr>
<tr>
<td>Septic tank servicers and sewer pipe cleaners</td>
<td>6750</td>
<td>0.0019</td>
<td>0.8</td>
<td>24.32</td>
</tr>
</tbody>
</table>

**Notes:** This table reports the top 10 most influential occupations according to two criteria. Panel A reports influential occupations according to the implicit weights on potentially heterogeneous treatment effects by occupation in the two-way fixed effect estimator, as derived by de Chaisemartin and D’Haultfoeuille (2019). Panel B reports overweighted occupations, as defined by the ratio of the implicit weight and the occupation’s sample share of workers.
Table A5: Additional Reduced-Form Effects of Occupational Licensing

<table>
<thead>
<tr>
<th></th>
<th>Licensed = 1</th>
<th>% Licensed in Cell</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Panel A: Weekly Hours Per Worker</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.690***</td>
<td>1.856***</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.313)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,149,992</td>
<td>2,149,992</td>
</tr>
<tr>
<td>Clusters</td>
<td>21,890</td>
<td>21,890</td>
</tr>
<tr>
<td><strong>Panel B: Employment Count (Poisson)</strong></td>
<td>-0.268***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>22,098</td>
<td></td>
</tr>
</tbody>
</table>

*Notes:* This table reports estimates from Equation 10 of effects of licensing on outcomes of interest which correspond to reduced-form moments of the model. The estimate in Column 1 refers to individual-worker licensing status, whereas those in Columns 2 and 3 refer to the state–occupation cell licensed share of workers. In Column 3, we include strata fixed effects for predetermined demographic observables. In Panel A, the dependent variable is the level of weekly hours per worker, and we include fixed effects for occupation, state, industry, and month. In Panel B, the dependent variable is the state-occupation employment count in a Poisson regression, and we include fixed effects for occupation and state. Standard errors are clustered at the level of the state-occupation cell. *** = p < 0.01.
Figure A1: Distribution of State–Occupation Licensed Shares

Notes: This figure shows the distribution of estimated shares of workers with a mandatory state-issued occupational license in each state–occupation cell, weighted by each cell’s total employment count. Licensed shares are estimated by the empirical Bayes procedure described in Section 3 and Appendix D.
B Model Appendix

This Appendix provides a detailed solution to the theoretical model of occupational licensing presented in Section 2. See the text for the structure of the main model. We restate here only the full optimization problem of worker \( i \):

\[
\max_{\{c_{ij}\}, h_i, y_i, a_i, J_i} \left\{ \log \left( \left( \sum_j q_j c_{ij}^{\frac{1}{\varepsilon}} \right)^{\frac{1}{1-\varepsilon}} - \frac{1}{1-\eta} h_i^{1+\eta} \right) - \rho (\tau J_i + y_i) + a_i J_i \right\}
\]

s.t. \( \sum_j w_j c_{ij} \leq A_j (y_i) w_i h_i. \)

The worker’s problem can be solved in three stages:

1. Given an income \( I_i = A_j (y_i) w_i h_i \), choose the consumption allocation \( \{c_{ij}\} \) that maximizes the value of the CES composite good.

2. Given an effective hourly wage \( A_j (y_i) w_j \), choose the hours \( h_i : J_i = j \) that maximize indirect utility in each occupation.

3. Given consumption–labor schedules \( \{\{c_{ij}\}, h_i : J_i = j\} \) for each occupation, choose the years of schooling \( y_i : J_i = j \) that maximize indirect utility in each occupation.

4. Given indirect utilities \( \tilde{V}_{ij} \) conditional upon entering each occupation \( j \), choose \( J_i = \text{arg max}_j \tilde{V}_{ij} \).

B.1 Consumption Decision

Begin with the CES utility maximization problem:

\[
\max_{\{c_{ij}\}} \sum_j q_j c_{ij}^{\frac{1}{\varepsilon}} \text{ s.t. } \sum_j w_j c_{ij} \leq I_i,
\]

where we hold \( I_i \) fixed. Given a large number of industries, the first order conditions with respect to \( c_{ij} \) are

\[ q_j c_{ij}^{-1/\varepsilon} + \lambda w_j = 0 \quad \forall n, \]

where \( \lambda \) is a Lagrange multiplier on the budget constraint. We omit the familiar CES derivations and proceed to results. Individual consumptions are:

\[ c_{ij} = \frac{A_j (y_i) w_i (w_j/q_j)^{-\varepsilon}}{\rho^{1-\varepsilon}}, \]
where the ideal price index is

\[ P = \left( \sum_j q_j^* w_j^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}, \]

such that the value of the optimal CES composite good available to the worker who has years of education \( y_i \) and works \( h_i \) hours in industry \( J_i \) has a consumption level

\[ C_i^*(h_i) = \frac{I(y_i, h_i)}{P} = \frac{A_{J_i} (y_i) w_{J_i} h_i}{P}. \]

We normalize the wage of a reference occupation \( w_0 = 1 \) such that \( \tau_0 = 1 \).

### B.2 Labor Supply Decision

Let \( V_j \) indicate the payoff-period utility apart from idiosyncratic occupation preferences and that is thus common across workers in occupation \( j \). We can rewrite the optimization problem at this stage as

\[
\max_{h_i} \left\{ C_i^*(h_i) - \frac{\psi}{1 + \eta} h_i^{1+\eta} \right\} \quad \text{s.t.} \quad C_i^*(h_i) \leq \frac{A_{J_i} (y_i) w_{J_i} h_i}{P}.
\]

This yields the first-order condition with respect to \( h_i \)

\[
\frac{A_{J_i} (y_i) w_{J_i}}{P} - \psi h_i^{\eta} = 0,
\]

and thereby the constant-elasticity labor supply function

\[ h_i^*(w_{J_i}) = \left( \frac{A_{J_i} (y_i) w_{J_i}}{\psi P} \right)^{\frac{1}{\eta}}. \]

We can now express \( V_j \) as a function of the wage \( w_{J_i} \), which the worker takes as given, and the schooling choice \( y_i \), which we endogenize in the next subsection:

\[
V_j(y_j) = \frac{A_{J_i} (y_i) w_{J_i}}{P} \left( \frac{A_{J_i} (y_i) w_{J_i}}{\psi P} \right)^{\frac{1}{\eta}} - \frac{\psi}{1 + \eta} \left( \frac{A_{J_i} (y_i) w_{J_i}}{\psi P} \right)^{\frac{1+\eta}{\eta}}.
\]

\[
= \frac{\eta}{1 + \eta} \left( \frac{A_{J_i} (y_i) w_{J_i}}{\psi P} \right)^{\frac{1+\eta}{\eta}}.
\]
B.3 Schooling

After observing \( \{\tau_j\} \), workers set their level of schooling to maximize their present-value utility. Taking logs, we have the solution to the schooling decision problem

\[
y^*_i = \arg \max_{y_i} \{ \log V_j(y_i) - \rho y_i \},
\]

which yields the first order condition

\[
1 + \frac{\eta}{\eta} \cdot \frac{A'_J(y^*_i)}{A_J(y^*_i)} = \frac{\rho}{\rho} = 0.
\]

We can therefore define \( v_j \) as the common indirect utility of the worker in occupation \( j \), which is

\[
v_j = e^{-\rho(y^*_i + \tau_j)} V_j(y^*_i) = \frac{\eta e^{-\rho(y^*_i + \tau_j)}}{1 + \eta} \left( \frac{A_J(y^*_i) w_j}{\psi P} \right)^{\frac{1+\eta}{\eta}}.
\]

B.4 Occupation Decision and Utility

The conditional indirect utility of a worker in occupation \( j \) is the product of common conditional indirect utility \( v_j \) and his or her idiosyncratic occupation preference term \( a_{ij} \):

\[
v_{ij} = a_{ij} v_j.
\]

As \( v_{ij} \) is increasing in the i.i.d. Fréchet random variable \( a_{ij} \), \( v_{ij} \) is itself distributed i.i.d. Fréchet. The worker’s problem at this stage is to pick the occupation \( j \) that maximizes \( V_{ij} \):

\[
J^*_i = \arg \max_j v_{ij}.
\]

By max-stability of \( v_{ij} \), \( v_{ij}^* \) is distributed i.i.d. Fréchet:

\[
v_{ij}^* = a_{ij} \left( \sum_j \left( \frac{\eta e^{-\rho(y^*_i + \tau_j)}}{1 + \eta} \right)^{\sigma} \left( \frac{A_J(y^*_i) w_j}{\psi P} \right)^{-\frac{1+\eta}{\eta}} \right)^{\frac{1}{\sigma}}.
\]

where \( a_{ij} \) is i.i.d. Fréchet with dispersion parameter \( \sigma \). Notice that the second term is independent of the choice \( J_i \). The choice probability of occupation \( j \) is

\[
s_j = P \left( J^*_i = \arg \max_j v_{ij} \right) = P \left( (a_{ij} - a_{ij'}) \geq \log \left( \frac{v_{ij}}{v_{ij'}} \right) \forall j \right) = \frac{v_{ij}^\sigma}{\sum_j v_{ij'}^\sigma}.
\]
The expected utility of workers in occupation \( j \) is

\[
\bar{u}_j = \mathbb{E}[V_{i,J_i}^* | J_i = j] = \mathbb{E} \left[ a_{i,J_i} \left( \sum_j \left( \frac{\eta e^{-\rho(y_{i,J_i}^*-\tau_j)}}{1 + \eta} \right)^\sigma \left( \frac{A_j(y_{i,J_i}^*)w_j}{\psi P} \right)^{\frac{\sigma(1+\eta)}{\eta}} \right) \left| J_i = j \right] \right] = \Gamma \left( 1 - \frac{1}{\sigma} \right) \left( \sum_j \left( \frac{\eta e^{-\rho(y_{i,J_i}^*-\tau_j)}}{1 + \eta} \right)^\sigma \left( \frac{A_j(y_{i,J_i}^*)w_j}{\psi P} \right)^{\frac{\sigma(1+\eta)}{\eta}} \right)^{\frac{1}{\sigma}}.
\]

Expected utility in occupation \( j \) is the same in all occupations and therefore equal to expected utility of all workers (\( \bar{u}_j = \bar{u} \) for all \( j \)).

### B.5 Willingness to Pay

We assume that willingness to pay is a function of the licensing cost and the expectation of workers’ idiosyncratic occupation preference term conditional upon entering the occupation:

\[
\log q_j = \kappa_0 j + \kappa_1 \log(1 - \ell_j) + \kappa_2 \log \mathbb{E}[a_{i,J_i} | J_i = j].
\]

For an occupation \( j \) that is sufficiently small, changes in \( \tau_j \) have a negligible effect on expected utility \( \bar{u} \). Also recall that

\[
\frac{\partial \log \bar{u}}{\partial \tau_j} \approx \frac{\partial \log v_j}{\partial \tau_j} + \frac{\partial \log \mathbb{E}[a_{i,J_i} | J_i = j]}{\partial \tau_j}.
\]

By the choice probability equation above, we also have

\[
\frac{\partial \log s_j}{\partial \tau_j} = \sigma \frac{\partial \log v_j}{\partial \tau_j}.
\]

Then combining these statements, we have

\[
\frac{\partial \log \mathbb{E}[a_{i,J_i} | J_i = j]}{\partial \tau_j} = \frac{\partial \log v_j}{\partial \tau_j} = -\frac{1}{\sigma} \frac{\partial \log s_j}{\partial \tau_j},
\]

and so

\[
\frac{d \log q_j}{\partial \tau_j} = \kappa_1 - \kappa_2 \frac{\partial \log \mathbb{E}[a_{i,J_i} | J_i = j']}{\partial \tau_j} = \kappa_1 - \frac{\kappa_2}{\sigma} \frac{\partial \log s_j}{\partial \tau_j} = \kappa_1 - \frac{\kappa_2}{\sigma} \frac{\partial \log s_j}{\partial \tau_j} \equiv \alpha,
\]

as employment shares are isoelastic in the net-of-license share.
B.6 Equilibrium Conditions

Consumption demand:
\[
\frac{\partial \log C_j}{\partial \tau_j} = \varepsilon \left( \frac{\partial \log q_j}{\partial \tau_j} - \frac{\partial \log w_j}{\partial \tau_j} \right)
\]

Willingness to pay:
\[
\frac{\partial \log q_j}{\partial \tau_j} = \alpha
\]

Intensive margin labor supply:
\[
\frac{\partial \log h_{i;J_i = j}}{\partial \tau_j} = \frac{1}{\eta} \frac{\partial \log w_j}{\partial \tau_j}
\]

Schooling:
\[
\frac{\partial \log y_{i;J_i = j}}{\partial \tau_j} = 0
\]

Extensive margin labor supply:
\[
\frac{\partial \log s_j}{\partial \tau_j} = \sigma (1 + \eta) \left( \frac{\partial \log w_j}{\partial \tau_j} - \rho \right)
\]

Labor market clearing:
\[
\frac{\partial \log C_j}{\partial \tau_j} = \frac{\partial \log H_j}{\partial \tau_j} = \frac{\partial \log s_j}{\partial \tau_j} + \frac{\partial \log h_{i;J_i = j}}{\partial \tau_j}
\]

B.7 Model Solution

The model can be solved by using the four labor market equilibrium conditions and the WTP equation. Let
\[
x' = \left[ \frac{\partial \log s_j}{\partial \tau_j}, \frac{\partial \log h_{i;J_i = j}}{\partial \tau_j}, \frac{\partial \log w_j}{\partial \tau_j}, \frac{\partial \log H_j}{\partial \tau_j}, \frac{\partial \log q_j}{\partial \tau_j} \right].
\]

The above results form a system of linear equations of the form \( Ax = Cx + b \), where \( A \) and \( C \) are 5-by-5 matrices and \( x' \) is a vector of length 5. If \( A \) and \( C \) are both of full rank and \( b \neq 0 \), the
system admits a unique solution \( x = (A - C)^{-1}b \). We confirm first that \( b \neq 0 \):

\[
b = \begin{bmatrix}
\frac{\rho \sigma(1 + \eta)}{\eta} \\
0 \\
\alpha \\
0 \\
\alpha
\end{bmatrix}.
\]

Thus, for \( b \neq 0 \), we require that either \( \sigma \neq 0 \) and \( |\eta| < \infty \), or \( \alpha \neq 0 \). The former two conditions will hold in all cases of interest. Since \( A = I \), we also have

\[
A - C = \begin{bmatrix}
1 & 0 & -\sigma(1 + \eta)/\eta & 0 & 0 \\
0 & 1 & -1/\eta & 0 & 0 \\
0 & 0 & 1 & 1/\varepsilon & 0 \\
-1 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}.
\]

The determinant of this matrix is

\[
|A - C| = -\frac{1 + \sigma(1 + \eta) + \eta \varepsilon}{\eta \varepsilon}.
\]

\( A - C \) is of full rank if and only if \( |A - C| \neq 0 \), thus if \( 1 + \sigma(1 + \eta) + \eta \varepsilon \neq 0 \) and \( |\eta \varepsilon| < \infty \). The economic content of this parameter restriction is to establish that, if a market-clearing wage exists, it is unique: It rules out the case in which the total labor supply elasticity—that is, the sum of the extensive and intensive margins—is exactly equal to the labor demand elasticity. This holds in any case of interest, as we assume \( \sigma > 0 \), \( \eta > 0 \), and \( \varepsilon < 0 \). With these restrictions, we have a unique solution to the model:

\[
\begin{bmatrix}
\frac{\partial \log s_j}{\partial \tau_j} \\
\frac{\partial \log h_{i,j,\omega j}}{\partial \tau_j} \\
\frac{\partial \log w_j}{\partial \tau_j} \\
\frac{\partial \log H_j}{\partial \tau_j} \\
\frac{\partial \log q_j}{\partial \tau_j}
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & -\sigma(1 + \eta)/\eta & 0 & 0 \\
0 & 1 & -1/\eta & 0 & 0 \\
0 & 0 & 1 & 1/\varepsilon & 0 \\
-1 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}^{-1}
\begin{bmatrix}
\frac{\rho \sigma(1 + \eta)}{\eta} \\
0 \\
\alpha \\
0 \\
\alpha
\end{bmatrix}
\times
\begin{bmatrix}
\sigma(1 + \eta)(\alpha \varepsilon - \rho(\varepsilon + 1/\eta)) \\
\rho \sigma(1 + \eta)/\eta + \alpha \varepsilon \\
\rho \sigma(1 + \eta) + \alpha \varepsilon \\
(1 + \eta)(\rho \sigma(1 + \eta) + \alpha \varepsilon)/\eta \\
\alpha(1 + \sigma(1 + \eta) + \eta \varepsilon)
\end{bmatrix}.
\]
B.8 Social Welfare

Expected utility is

\[ \bar{u} \propto \left[ \sum_j \left( \frac{A_j(y_j^*) w_j e^{-\rho(y_j^* + \tau_j)}}{P} \right)^{\sigma(1+\eta)/\eta} \right]^{1/\sigma}. \]

Taking the logarithm and collecting constants with respect to \( \tau_j \) in \( c \), we obtain

\[ \log \bar{u} = c + \frac{1}{\sigma} \log \left[ \sum_j \left( \frac{A_j(y_j^*) w_j e^{-\rho(y_j^* + \tau_j)}}{P} \right)^{\sigma(1+\eta)/\eta} \right], \]

and then we can use a first-order approximation for the partial derivative with respect to \( \tau_j \)

\[ \frac{\partial \log \bar{u}}{\partial \tau_j} = \frac{1 + \eta}{\eta} \sum_j s_j \left( \frac{\partial \log w_j}{\partial \tau_j} - \rho \frac{\partial \tau_j}{\partial \tau_j'} - \frac{\partial \log P}{\partial \tau_j'} \right). \]

By the envelope theorem,

\[ \frac{\partial}{\partial \tau_j'} \left[ \log A_j(y_j^*) - \rho y_j^* \right] = 0 \quad \forall j, \]

and thus

\[ \frac{\partial \log \bar{u}}{\partial \tau_j'} = \frac{1 + \eta}{\eta} \sum_j s_j \left( \frac{\partial \log w_j}{\partial \tau_j} - \rho \frac{\partial \tau_j}{\partial \tau_j'} - \frac{\partial \log P}{\partial \tau_j'} \right). \]

Splitting the sum into occupation \( j' \) whose \( \tau_{j'} \) changes and all others, we have that

\[ \frac{\partial \tau_{j'}}{\partial \tau_{j'}} = 1 \quad \text{and} \quad \frac{\partial \tau_j}{\partial \tau_{j'}} = 0 \quad \forall j' \neq j, \]

and so, simplifying further, we obtain which can be rearranged to

\[ \frac{\partial \log \bar{u}}{\partial \tau_{j'}} = \frac{1 + \eta}{\eta} \left[ s_{j'} \left( \frac{\partial \log w_{j'}}{\partial \tau_{j'}} - \rho \right) + \sum_{j:j \neq j'} s_j \frac{\partial \log w_j}{\partial \tau_{j'}} \right] - \frac{1 + \eta}{\eta} \frac{\partial \log P}{\partial \tau_{j'}}. \]

Inverting Equation 9, and doing this for both \( j' \) and \( j : j \neq j' \), we obtain

\[ \frac{\partial \log w_{j'}}{\partial \tau_{j'}} - \rho = \frac{\eta}{\sigma(1+\eta)} \frac{\partial \log s_{j'}}{\partial \tau_{j'}} \]

\[ \frac{\partial \log w_j}{\partial \tau_{j'}} = \frac{\eta}{\sigma(1+\eta)} \frac{\partial \log s_j}{\partial \tau_{j'}} \quad \forall j : j \neq j' \]
and substitutions yield
\[
\frac{\partial \log \bar{u}}{\partial \tau_{j'}} = s_{j'} \frac{\partial \log s_{j'}}{\partial \tau_{j'}} + \frac{1}{\sigma} \sum_{j:j \neq j'} s_j \frac{\partial \log s_j}{\partial \tau_{j'}} - \frac{1 + \eta}{\eta} \frac{\partial \log P}{\partial \tau_{j'}}.
\]

Independence of preferences across occupations gives us that displaced workers from occupation \( j' \) are apportioned to occupations \( j \neq j' \) according to the shares of \( j \) in total employment:
\[
\frac{\partial \log s_j}{\partial \tau_{j'}} = -\frac{s_j}{1 - s_{j'}} \frac{\partial \log s_{j'}}{\partial \tau_{j'}}.
\]

Under our assumption of a utilitarian social welfare function, \( W = \sum_i u_i = \sum N\bar{u} \). By these substitutions, we obtain
\[
\frac{\partial \log W}{\partial \tau_{j'}} = s_{j'} \frac{\partial \log s_{j'}}{\partial \tau_{j'}} - \frac{1}{\sigma} \sum_{j:j \neq j'} \frac{s_j^2}{1 - s_{j'}} \frac{\partial \log s_{j'}}{\partial \tau_{j'}} - \frac{1 + \eta}{\eta} \frac{\partial \log P}{\partial \tau_{j'}}.
\]

which rewrites to
\[
\frac{\partial \log W}{\partial \tau_{j'}} = \frac{1}{\sigma} \frac{\partial \log s_{j'}}{\partial \tau_{j'}} \left( s_{j'} - \frac{\sum_{j:j \neq j'} s_j^2}{1 - s_{j'}} \right) - \frac{1 + \eta}{\eta} \frac{\partial \log P}{\partial \tau_{j'}},
\]

which has a rich economic interpretation. We have characterized the welfare effect of licensing occupation \( j \) on employment in occupation \( j \) even in a model with nonnegligible spillovers across occupations, and it reflects changes in employment in the licensed occupation and in the price level. Second, the normalized Herfindahl index of employment shares summarizes the extent of these cross-occupation spillovers. This is, to the the best of our knowledge, a novel theoretical connection between the normalized Herfindahl index and the relevance of spillovers to welfare.

In the limit \( H_j = \sum_j s_j^2 \to 0 \) in which the effective number of occupations approaches infinity, spillovers become negligible, and we obtain a particularly stark welfare result:
\[
\frac{\partial \log \bar{W}}{\partial \tau_j} = \frac{s_j \partial \log s_j}{\sigma} - \frac{1 + \eta}{\eta} \frac{\partial \log P}{\partial \tau_j}.
\]

In the paper, we perform several manipulations on this result. First, we define occupational surplus as the difference in social welfare, holding all other \( \{\tau_j\} \) constant, between the equilibrium with \( \tau_j = 0 \) (no licensing) and the equilibrium \( \tau_j \to \infty \) (occupation banned).
\[
\bar{W}_j = \bar{W}(0, \{\tau_j\}) - \lim_{\tau_j \to \infty} \bar{W}(\tau_j, \{\tau_{j'}\}).
\]

Then the above rewrites to
\[
\frac{\partial \log \bar{W}_j}{\partial \tau_j} = \frac{1}{\sigma} \frac{\partial \log s_j}{\partial \tau_j} - \frac{1 + \eta}{\eta s_j} \frac{\partial \log P}{\partial \tau_j}.
\]
Furthermore, we can obtain the partial derivative of $P$ with respect to $\tau_j'$:

$$
\frac{\partial \log P}{\partial \tau_j'} = \frac{1}{1 - \varepsilon} \cdot \frac{\partial}{\partial \tau_j'} \log \sum_j q_j w_j^{1 - \varepsilon}
$$

$$
\approx \frac{1}{1 - \varepsilon} \sum_j s_j \frac{\partial}{\partial \tau_j'} [\varepsilon \log q_j + (1 - \varepsilon) \log w_j]
$$

$$
= \frac{1}{1 - \varepsilon} \sum_j s_j \left[ \varepsilon \left( \frac{\partial \log q_j}{\partial \tau_j'} - \frac{\partial \log w_j}{\partial \tau_j'} \right) + \frac{\partial \log w_j}{\partial \tau_j'} \right].
$$

From Equation 8, we have

$$
\varepsilon \left( \frac{\partial \log q_j}{\partial \tau_j'} - \frac{\partial \log w_j}{\partial \tau_j'} \right) = \frac{\partial \log s_j}{\partial \tau_j'} + \frac{\partial \log h_{i,j=i=j}}{\partial \tau_j'},
$$

and so by substitution,

$$
\frac{\partial \log P}{\partial \tau_j} = \frac{1}{1 - \varepsilon} \sum_j s_j \left( \frac{\partial \log s_j}{\partial \tau_j'} + \frac{\partial \log h_{i,j=i=j}}{\partial \tau_j'} + \frac{\partial \log w_j}{\partial \tau_j'} \right)
$$

$$
= \frac{1}{1 - \varepsilon} \sum_j s_j \frac{\partial \log w_j H_j}{\partial \tau_j'}.
$$

A similar argument as above applies to the off-diagonal terms, yielding the approximation

$$
\frac{\partial \log P}{\partial \tau_j} = \frac{s_j}{1 - \varepsilon} \frac{\partial \log w_j H_j}{\partial \tau_j'},
$$

which in turn implies

$$
\frac{\partial \log W_j}{\partial \tau_j} = \frac{1}{\sigma} \frac{\partial \log s_j}{\partial \tau_j} + \frac{1 + \eta}{\eta(\varepsilon - 1)} \frac{\partial \log w_j H_j}{\partial \tau_j'}.
$$

We have now mapped the social welfare effect of licensing into two reduced-form comparative statics that are, in principle, estimable from only labor market data—the effects of licensing on the own-occupation employment share and wage bill—and three structural parameters. These structural parameters all have known sign ($\sigma > 0, \eta > 0, \varepsilon - 1 > 0$), and thus we can view the welfare effect as a weighted sum of these two reduced-form responses.

A second manipulation of the welfare result is to define respectively changes in worker and consumer surplus as

$$
\frac{\partial \log W^L}{\partial \tau_j} = s_j \frac{\partial \log s_j}{\partial \tau_j}
$$

$$
\frac{\partial \log W^C}{\partial \tau_j} = \frac{1 + \eta}{\eta} \frac{\partial \log P}{\partial \tau_j}.
$$

We introduce this terminology to think intuitively about the incidence of licensing: Workers bear the
costs of licensing insofar as licensing reduces the present value of nominal income in an occupation (and thus spurs workers to exit the occupation on the margin), whereas consumers bear the costs of licensing insofar as licensing raises the price level, reducing the real income of all workers, including those not in the licensed occupation. Using Proposition 1, we express licensing’s effects on worker and consumer surplus in terms of structural parameters:

\[
\frac{\partial \log W_L}{\partial \tau_j} = s_j \cdot \frac{(1 + \eta)(\alpha \varepsilon - \rho(\varepsilon + 1/\eta))}{1 + \sigma(1 + \eta) + \eta \varepsilon},
\]

\[
\frac{\partial \log W_C}{\partial \tau_j} = s_j \cdot \frac{(1 + \eta)^2}{\eta^2(\varepsilon - 1)} \cdot \frac{\alpha \varepsilon(1 + \sigma \eta) - \rho \sigma(1 - \eta \varepsilon)}{1 + \sigma(1 + \eta) + \eta \varepsilon}.
\]

Taken together, and rescaled into occupational surplus, we obtain

\[
\frac{\partial \log W_j}{\partial \tau_j} = \frac{1 + \eta}{\eta} \left( \frac{\alpha \varepsilon}{\varepsilon - 1} - \rho \right).
\]

We can also use the model to analyze incidence. First, we may write the share of licensing costs that are offset for workers by increases in wages fully in terms of primitives:

\[
\frac{1}{\rho} \frac{\partial \log w_j}{\partial \tau_j} = \frac{\sigma(1 + \eta)\alpha \varepsilon - \rho(\varepsilon + 1/\eta)}{\rho(1 + \sigma(1 + \eta) + \eta \varepsilon)},
\]

Next, we can write the effect of licensing on the WTP-adjusted price in terms of primitives:

\[
\frac{\partial \log(q_j^{\varepsilon} p_j^{1-\varepsilon})}{\partial \tau_j} = \frac{\alpha + (\alpha - \rho)\sigma(1 + \eta)}{1 + \sigma(1 + \eta) + \eta \varepsilon},
\]

and then we can calculate the share of the price increase offset by increases in WTP in terms of primitives:

\[
\frac{\partial \log(q_j^{\varepsilon} p_j^{1-\varepsilon})}{\partial \tau_j} / \frac{\partial \log p_j}{\partial \tau_j} = \frac{\rho \sigma(1 + \eta) + \alpha \varepsilon \eta}{\alpha + (\alpha - \rho)\sigma(1 + \eta)}.
\]

B.9 Proofs of Propositions

Proposition 1

Section B.7 presents a detailed derivation.
Proposition 2

Take the partial derivative with respect to $\alpha$:

$$
\frac{\partial}{\partial \alpha} \begin{bmatrix}
\frac{\partial \log s_j}{\partial \tau_j} \\
\frac{\partial \log h_{i,j} - j}{\partial \tau_j} \\
\frac{\partial \log w_j}{\partial \tau_j} \\
\frac{\partial \log H_j}{\partial \tau_j} \\
\frac{\partial \log q_j}{\partial \tau_j}
\end{bmatrix}
= \frac{1}{1 + \sigma(1 + \eta) + \eta \varepsilon} \cdot \frac{\partial}{\partial \alpha} \begin{bmatrix}
\sigma(1 + \eta)(\alpha \varepsilon - \rho(\varepsilon + 1/\eta)) \\
\rho \sigma(1 + \eta)/\eta + \alpha \varepsilon \\
\rho \sigma(1 + \eta) + \alpha \varepsilon \eta \\
(1 + \eta)(\rho \sigma(1 + \eta) + \alpha \varepsilon)/\eta \\
\alpha(1 + \sigma(1 + \eta) + \eta \varepsilon)
\end{bmatrix}
$$

One immediately sees the claimed sign on all crosspartials.

Proposition 3

Section B.8 presents a detailed derivation.

Proposition 4

Proposition 3 proves that the social welfare effect of licensing, in terms of the percentage change in occupational surplus, is

$$
\frac{\partial \log W_j}{\partial \tau_j} = \frac{1}{\sigma} \frac{\partial \log s_j}{\partial \tau_j} + \frac{1 + \eta}{\eta(\varepsilon - 1)} \frac{\partial \log w_j H_j}{\partial \tau_j},
$$

and substituting in comparative statics from Proposition 1, we obtain

$$
\frac{\partial \log W_j}{\partial \tau_j} = \frac{(1 + \eta)(\alpha \varepsilon - \rho(\varepsilon + 1/\eta)) + \frac{(1 + \eta)^2}{\varepsilon}(\frac{(1 + \sigma)\alpha\varepsilon}{\varepsilon - 1} - \rho \sigma)}{1 + \sigma(1 + \eta) + \eta \varepsilon},
$$

and since we wish to test that $\partial \log W_j/\partial \tau_j < 0$, we divide by the common factor $\frac{1 + \eta}{1 + \sigma(1 + \eta) + \eta \varepsilon}$ and obtain the test

$$
\alpha \varepsilon - \rho(\varepsilon + 1/\eta) + \frac{1 + \eta}{\eta} \left(\frac{(1 + \sigma)\alpha\varepsilon}{\varepsilon - 1} - \rho \sigma\right) < 0,
$$

which, for $\varepsilon > 1$, simplifies to

$$
\rho > \frac{\alpha \varepsilon}{\varepsilon - 1}.
$$
Proposition 5

Incidence ($\gamma^L$). We divide our formula for the worker welfare effect by the social welfare effect:

$$\gamma^L = \frac{\Delta W^L}{\Delta W} = \frac{(1+\eta)(\alpha - \rho(\varepsilon - 1/\eta))}{1+\sigma(1+\eta) + \eta \varepsilon} \cdot \frac{\alpha \varepsilon}{\varepsilon - \rho} \cdot \frac{(\alpha - \rho)\eta \varepsilon - \rho(\varepsilon - 1)}{(\alpha - \rho)(\varepsilon + \rho)(1 + \sigma(1 + \eta) + \eta \varepsilon)}.$$ 

Conditions for $\Delta W^L < 0 < \Delta W^C$. First, using the worker welfare formula, we obtain

$$\Delta W^L < 0 \iff \alpha \varepsilon - \rho(\varepsilon - 1/\eta) < 0 \iff \alpha < \rho \left(1 - \frac{1}{\eta \varepsilon}\right).$$

Next, using the consumer welfare formula, we obtain

$$\Delta W^C > 0 \iff \alpha \varepsilon(1 + \sigma) - \rho \sigma (\varepsilon - 1) > 0 \iff \alpha > \frac{\rho \sigma}{1 + \sigma} \left(1 - \frac{1}{\varepsilon}\right).$$

Thus

$$\Delta W^L < 0 < \Delta W^C \iff \alpha \in \left(\frac{\rho \sigma}{1 + \sigma} \left(1 - \frac{1}{\varepsilon}\right), \rho \left(1 - \frac{1}{\eta \varepsilon}\right)\right).$$

B.10 Constructive Proof of Identification

In this subsection we show constructively that the vector of reduced-form empirical moments $\hat{\beta} = \{\hat{a}_i, \hat{w}_j, \hat{h}_i, \hat{J}_{i=j}, \hat{s}_j\}$ just-identify the vector of structural parameters $\theta = [\rho, \eta, \alpha, \tau]$ with the calibration of $\sigma$ and $\varepsilon$. The structural parameters may be recovered by:

$$\eta = \frac{\hat{w}_j}{\hat{h}_i}$$

$$\tau = \hat{a}_i$$

$$\alpha = \frac{\hat{w}_j}{\varepsilon} (\hat{s}_j + \hat{h}_i)$$

$$\rho = \frac{\hat{w}_j \hat{s}_j}{\sigma(\hat{w}_j + \hat{h}_j)}.$$
Table A6: Reduced-Form Worker Effects of Occupational Licensing, Including Universally Licensed Occupations

<table>
<thead>
<tr>
<th>Panel A: Years of Education</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Licensed = 1</strong></td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>(2)</td>
</tr>
<tr>
<td>(3)</td>
</tr>
<tr>
<td><strong>0.375</strong>* 0.449*** 0.388*****</td>
</tr>
<tr>
<td>(0.010) (0.054) (0.052)</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>2,149,992 2,149,992 2,149,992</td>
</tr>
<tr>
<td>Clusters</td>
</tr>
<tr>
<td>21,890 21,890 21,890</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Years of Age</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observations</strong></td>
</tr>
<tr>
<td>1.289*** 1.737*** 1.715***</td>
</tr>
<tr>
<td>(0.035) 0.266 0.264</td>
</tr>
<tr>
<td>Clusters</td>
</tr>
<tr>
<td>811,117 811,117 811,117</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Log Hourly Wage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observations</strong></td>
</tr>
<tr>
<td>0.154*** 0.200*** 0.149***</td>
</tr>
<tr>
<td>(0.005) (0.024) (0.023)</td>
</tr>
<tr>
<td>Clusters</td>
</tr>
<tr>
<td>20,273 20,273 20,273</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D: Log Weekly Hours Per Worker</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observations</strong></td>
</tr>
<tr>
<td>0.045*** 0.049*** 0.036***</td>
</tr>
<tr>
<td>(0.002) (0.010) (0.010)</td>
</tr>
<tr>
<td>Clusters</td>
</tr>
<tr>
<td>2,149,992 2,149,992 2,149,992</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel E: Log Employment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observations</strong></td>
</tr>
<tr>
<td>-0.179***</td>
</tr>
<tr>
<td>(0.061)</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from Equation 10 of effects of licensing on outcomes of interest which correspond to reduced-form moments of the model. The estimate in Column 1 refers to individual-worker licensing status, whereas those in Columns 2 and 3 refer to the state-occupation cell licensed share of workers. In Columns 1 and 3, we include strata fixed effects for predetermined demographic observables. All specifications include fixed effects for occupation, state, industry, and month, except in Panel E, which has only state and occupation fixed effects. Standard errors are clustered at the level of the state-occupation cell. "*** = p < 0.01."
Table A7: Reduced-Form Effects of Occupational Licensing, ACS Sample

<table>
<thead>
<tr>
<th>Panel A: Years of Age</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Licensed</td>
<td>0.642***</td>
<td>0.642***</td>
</tr>
<tr>
<td></td>
<td>(0.191)</td>
<td>(0.191)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,326,484</td>
<td>1,326,484</td>
</tr>
<tr>
<td>Clusters</td>
<td>19,187</td>
<td>19,187</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Log Hourly Wage</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Licensed</td>
<td>0.101***</td>
<td>0.101***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Observations</td>
<td>4,032,135</td>
<td>4,032,135</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,124</td>
<td>20,124</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Log Weekly Hours Per Worker</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Licensed</td>
<td>0.020**</td>
<td>0.020**</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Observations</td>
<td>4,032,135</td>
<td>4,032,135</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,124</td>
<td>20,124</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D: Log Employment</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Licensed</td>
<td>-0.247***</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
</tr>
<tr>
<td>Observations</td>
<td>20,230</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from Equation 10 of effects of licensing on outcomes of interest which correspond to reduced-form moments of the model. The data is the 5-year sample (2010–2015) of the American Community Survey. In Column 2, we include strata fixed effects for predetermined demographic observables. All specifications include fixed effects for occupation, state, industry, and month, except in Panel D, which has only state and occupation fixed effects. Standard errors are clustered at the level of the cell. *** = p < 0.01.
Table A8: Robustness Checks, Including Universally Licensed Occupations

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unions &amp; Cert.</td>
<td>Household FE</td>
<td>Occ. &amp; Demo. Mix</td>
<td>State–Occ. Group FE</td>
<td>Div–Occ. FE</td>
<td>Drop ± 1 SD</td>
</tr>
<tr>
<td>% Licensed</td>
<td>0.426***</td>
<td>0.102**</td>
<td>0.378***</td>
<td>0.335***</td>
<td>0.276***</td>
<td>0.385***</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.044)</td>
<td>(0.052)</td>
<td>(0.049)</td>
<td>(0.049)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,149,992</td>
<td>2,124,590</td>
<td>2,144,001</td>
<td>2,149,992</td>
<td>2,149,989</td>
<td>440,886</td>
</tr>
<tr>
<td>Clusters</td>
<td>21,890</td>
<td>21,867</td>
<td>21,015</td>
<td>21,890</td>
<td>21,887</td>
<td>9,075</td>
</tr>
</tbody>
</table>

Panel A: Years of Education

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Licensed</td>
<td>1.751***</td>
<td>0.351***</td>
<td>1.752***</td>
<td>1.718***</td>
<td>1.256***</td>
<td>1.240***</td>
</tr>
<tr>
<td></td>
<td>(0.269)</td>
<td>(0.123)</td>
<td>(0.267)</td>
<td>(0.250)</td>
<td>(0.266)</td>
<td>(0.230)</td>
</tr>
<tr>
<td>Observations</td>
<td>811,117</td>
<td>790,531</td>
<td>809,150</td>
<td>811,117</td>
<td>811,090</td>
<td>154,122</td>
</tr>
<tr>
<td>Clusters</td>
<td>19,266</td>
<td>19,095</td>
<td>18,814</td>
<td>19,266</td>
<td>19,239</td>
<td>7,219</td>
</tr>
</tbody>
</table>

Panel B: Years of Age

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Licensed</td>
<td>0.108***</td>
<td>0.093***</td>
<td>0.147***</td>
<td>0.135***</td>
<td>0.119***</td>
<td>0.165***</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.026)</td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Observations</td>
<td>365,261</td>
<td>273,904</td>
<td>364,221</td>
<td>365,260</td>
<td>365,163</td>
<td>75,322</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,273</td>
<td>19,099</td>
<td>19,068</td>
<td>20,272</td>
<td>20,175</td>
<td>7,927</td>
</tr>
</tbody>
</table>

Panel C: Log Hourly Wage

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Licensed</td>
<td>0.036***</td>
<td>0.037***</td>
<td>0.036***</td>
<td>0.033***</td>
<td>0.029***</td>
<td>0.028***</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.009)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,149,992</td>
<td>2,124,590</td>
<td>2,144,001</td>
<td>2,149,992</td>
<td>2,149,989</td>
<td>440,886</td>
</tr>
<tr>
<td>Clusters</td>
<td>21,890</td>
<td>21,867</td>
<td>21,015</td>
<td>21,890</td>
<td>21,887</td>
<td>9,075</td>
</tr>
</tbody>
</table>

Panel D: Log Weekly Hours Per Worker

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Licensed</td>
<td>-0.197***</td>
<td>-0.139**</td>
<td>-0.052</td>
<td>-0.125**</td>
<td>-0.112*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.057)</td>
<td>(0.054)</td>
<td>(0.051)</td>
<td>(0.059)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>22,098</td>
<td>21,026</td>
<td>22,098</td>
<td>22,008</td>
<td>8,311</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from variations on Equation 10 as explained in the text. All estimates refer to the coefficient on the licensed share of workers in the state–occupation cell. All specifications include fixed effects for occupation, state, industry, and month, except in Panel E, which has only state and occupation fixed effects. Standard errors are clustered at the level of the state–occupation cell. * = p < 0.10, ** = p < 0.05, *** = p < 0.01.
<table>
<thead>
<tr>
<th>Panel</th>
<th>(1) State–Occ. Group FE</th>
<th>(2) State-Demog. &amp; Occ-Demog. FE</th>
<th>(3) Flexible Licensed Share</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Years of Education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>0.319***</td>
<td>0.312***</td>
<td>0.364***</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.053)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,865,209</td>
<td>1,865,172</td>
<td>1,865,209</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,321</td>
<td>20,319</td>
<td>20,321</td>
</tr>
<tr>
<td><strong>Panel B: Years of Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>1.070***</td>
<td>0.665***</td>
<td>0.964***</td>
</tr>
<tr>
<td></td>
<td>(0.239)</td>
<td>(0.203)</td>
<td>(0.244)</td>
</tr>
<tr>
<td>Observations</td>
<td>722,168</td>
<td>722,128</td>
<td>722,168</td>
</tr>
<tr>
<td>Clusters</td>
<td>17,842</td>
<td>17,830</td>
<td>17,842</td>
</tr>
<tr>
<td><strong>Panel C: Log Hourly Wage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>0.139***</td>
<td>0.133***</td>
<td>0.163***</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.023)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Observations</td>
<td>317,141</td>
<td>316,764</td>
<td>317,142</td>
</tr>
<tr>
<td>Clusters</td>
<td>18,752</td>
<td>18,601</td>
<td>18,753</td>
</tr>
<tr>
<td><strong>Panel D: Log Weekly Hours Per Worker</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Licensed</td>
<td>0.032***</td>
<td>0.027***</td>
<td>0.028***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.010)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Observations</td>
<td>18,65209</td>
<td>1865172</td>
<td>1865209</td>
</tr>
<tr>
<td>Clusters</td>
<td>20,321</td>
<td>20,319</td>
<td>20,321</td>
</tr>
<tr>
<td><strong>Panel E: Log Employment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.097</td>
<td></td>
<td>-0.297***</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td></td>
<td>(0.058)</td>
</tr>
<tr>
<td>Observations</td>
<td>20,524</td>
<td></td>
<td>20,524</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from Equation 10. All estimates refer to the coefficient on the licensed share of workers in the state–occupation cell. All specifications include fixed effects for occupation, state, industry, and month, except in Panel E, which has only state and occupation fixed effects. Standard errors are clustered at the level of the cell. *** = \( p < 0.01 \).
Table A10: State–Occupation Licensed Shares and Local Political Determinants

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Licensed</td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2)</td>
</tr>
<tr>
<td>%Rep_o × Slant_s</td>
<td>-0.006 (0.008)</td>
<td></td>
</tr>
<tr>
<td>%Indep_o × Polarization_s</td>
<td>0.004 (0.016)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>18,245</td>
<td>18,245</td>
</tr>
<tr>
<td>R2 (within)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from Equation 13, which tests for political determinants of licensing at the state–occupation level that reflect either local political economy or occupation-specific political position. Variables are defined in text. Both specifications include fixed effects for occupation and state. Standard errors are clustered at the level of the state–occupation cell. * = p < 0.10, * = p < 0.05, *** = p < 0.01.
C Further Results

C.1 Supplementary Robustness Checks

See Appendix Table A9. More discussion of these results will be added in a subsequent draft.

C.2 Educational Attainment

Note: This section is out-of-date and will be updated in a subsequent draft.

Occupational licensing regulations commonly specify a minimum required educational credential (Gittleman et al., 2018). Here we seek to recover the relevant credential for each occupation when it is licensed, and splitting occupations by these credentials, estimate distinct effects of licensing on the distribution of educational attainment. We view these results as providing our most credible evidence that licensing policy has a causal effect on educational attainment: That is, we claim that, absent licensing requirements, workers would not obtain such educational credentials.

Motivated by the results in Figure 1, we posit that licensing schemes divide into two types: one that requires associates’ degrees or similar, and another requiring more than a bachelor’s degree. We argue the former is consistent with licensed occupations with relatively low average level of education and the latter with licensed occupations with relatively high average level of education. We implement this division by $k$-means clustering: we compute the share of workers with each detailed level of education by occupation using sample weights and then use the $k$-means algorithm to divide occupations for $k = 2$. We find that these clusters split occupations into intuitively low- and high-education groups: See Appendix Figure A2. In addition, our results are robust to alternative approaches, such as splitting occupations at the median by average years of education.

Appendix Figure A3 displays the results. Consistent with our hypothesis, occupational licensing has sharply heterogeneous effects on the education distribution in low- and high-education occupations. In low-education occupations, we see a large (6.0 p.p.) decline in the share of workers whose highest level of education is a high school diploma and a large (8.2 p.p.) increase in the share of workers with vocational associate’s degrees. By contrast, in high-education occupations, the effects are concentrated in a large (6.4 p.p.) decline in the share of workers with bachelor’s degrees and a concomitant (8.2 p.p.) increase in the share of workers with master’s degrees. These point estimates are precisely estimated, with standard errors in the range of 0.1–1 p.p. of workers. We can easily reject equality of coefficients for the effects of licensing in low- versus high-education occupations, for most individual education levels and jointly across all education levels. These results establish a notably direct link between the specific educational requirements likely required when an occupation is licensed and the actual changes in the distribution of educational attainment within that occupation.

Does licensing affect the distribution of education by inducing selection or investment? That is,
do workers respond to licensing by entering and exiting occupations, or by changing their human capital decisions? We step beyond our main research design to consider this question. In particular, we allow licensing in a state–occupation cell to affect the state share of workers with that education by removing state fixed effects from our specification. The decomposition

\[ \beta_{\text{StateFE}} = \beta_{\text{Selection}} + \beta_{\text{Investment}} \]

\[ \beta_{\text{NoStateFE}} = \beta_{\text{Investment}} \]

allows for an economic interpretation of the differences between estimates with and without state fixed effects.\(^{23}\) Appendix Figure ?? displays the results. Overall, they indicate that the observed

---

\(^{23}\)Removing state fixed effects eliminates the within-state selection component (\(\beta_{\text{Selection}}\)) because, holding educa-
Figure A3: Distribution of Educational Attainment, by Occupation Cluster

<table>
<thead>
<tr>
<th>Educational Attainment</th>
<th>Percent of Workers in Clustered Occupations</th>
</tr>
</thead>
<tbody>
<tr>
<td>None or preschool</td>
<td>-20%</td>
</tr>
<tr>
<td>Grades 1, 2, 3, or 4</td>
<td>-15%</td>
</tr>
<tr>
<td>Grades 5 or 6</td>
<td>-10%</td>
</tr>
<tr>
<td>Grades 7 or 8</td>
<td>-5%</td>
</tr>
<tr>
<td>Grade 9</td>
<td>0%</td>
</tr>
<tr>
<td>Grade 10</td>
<td>5%</td>
</tr>
<tr>
<td>Grade 11</td>
<td>10%</td>
</tr>
<tr>
<td>12th grade, no diploma</td>
<td>15%</td>
</tr>
<tr>
<td>High school diploma or equivalent</td>
<td>20%</td>
</tr>
<tr>
<td>Some college but no degree</td>
<td></td>
</tr>
<tr>
<td>Associate’s degree, occupational</td>
<td></td>
</tr>
<tr>
<td>Associate’s degree, academic</td>
<td></td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td></td>
</tr>
<tr>
<td>Master’s degree</td>
<td></td>
</tr>
<tr>
<td>Professional school degree</td>
<td></td>
</tr>
<tr>
<td>Doctorate degree</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This figure presents estimates of the effects of occupational licensing on the cell shares of workers by detailed level of educational attainment, in which we split the effects based on whether the occupation is assigned to the low- or high-education cluster by a k-means procedure described in Appendix C.

Changes in education occur due to changes in investment rather than selection of workers by education into or out of licensed occupations. Regressing the investment coefficients on the total coefficients, we find investment explains 86 percent and selection an insignificant 14 percent of the total response to licensing; see Appendix Figure ??.

C.3 Robustness to Political Confounds

Do local political determinants of regulation including, but extending beyond, occupational licensing confound our identification strategy? For example, it may be that occupations whose workers tend to vote for Republicans (Democrats) also tend to be more heavily licensed in states that gen-

---

Notes: This figure presents estimates of the effects of occupational licensing on the cell shares of workers by detailed level of educational attainment, in which we split the effects based on whether the occupation is assigned to the low- or high-education cluster by a k-means procedure described in Appendix C.

Changes in education occur due to changes in investment rather than selection of workers by education into or out of licensed occupations. Regressing the investment coefficients on the total coefficients, we find investment explains 86 percent and selection an insignificant 14 percent of the total response to licensing; see Appendix Figure ??.
erally vote Republican (Democrat). To evaluate this and related hypotheses, we use data on the political ideology of workers by occupation from the 1972–2016 Cumulative Datafile of the U.S. General Social Survey (GSS) as well as the ideology of politicians in state legislatures from Shor and McCarty (2011).

The GSS asks participants for their occupation as well as their political party affiliation. Occupations are classified as in the CPS. The GSS asks about party affiliation with the question: “Generally speaking, do you usually think of yourself as a Republican, Democrat, Independent, or what?” We coded individuals who responded they were a “strong” or “not strong” Republican or Democrat as their respective parties. Remaining respondents identified as either independents or members of another party and were coded as a third category. The pooled sample includes 62,644 responses and 534 unique occupations. To reduce sampling error in the Republican and Democratic shares of workers in each occupation, we estimated a mixed-effects logistic regression model, with occupation random effects nested within random effects for 23 Census detailed occupation groups. The following analysis uses the model-based predicted Republican share of the two-party vote by occupation. For state-level variation, we use ideal-point estimates from Shor and McCarty (2011) of the average ideology of each U.S. state legislature in 2014, taking the simple average of the upper and lower legislative bodies in each state, as well as the distance between the median Republican and median Democratic legislator. For ease of interpretation, we then standardized these state-politics variables to be mean zero and unit standard deviation.

We estimate variations on the following specification, which interacts a GSS occupation-level variable with a Shor and McCarty (2011) state-level variable:

\[
\%\text{License}_{os} = \alpha_o + \alpha_s + \beta \cdot (\text{OccupationPolitics}_o \times \text{StatePolitics}_s) + e_{os}
\]  

(13)

We keep the state–occupation licensed share as the dependent variable, cluster at the state–occupation cell level, and include in state- and occupation fixed effects. To the extent a coefficient is significant, this may raise concerns that the state–occupation licensed share is correlated with other regulations and policies that vary among states and occupations.

Appendix Table A10, however, finds no evidence of associations of occupation- and state-level political variable interactions with the licensed share. We try plausible specifications that might reveal local political determinants of licensing. In Column 1, we interact the occupation Republican share with the average left-right slant of the state legislature. Column 2 uses instead the occupation Democratic share in the interaction. These two results suggest that Republican- and Democratic-leaning legislatures do not respectively differentially treat Republican- and Democratic-leaning occupations with licensing. Column 3 uses the share of workers who are either Republicans or Democrats and interacts this with the distance between party medians. The insignificant result suggests that polarized state legislatures do not differentially treat occupations that are relatively more or less politically independent with licensing.

Though this exercise does not rule out all possible local political explanations, it does suggest that patterns of licensing across U.S. states and occupations are relatively idiosyncratic and not
easily explained by local politics.

D Econometric Extensions

This Appendix is in two parts. In Sections 1 and 2, we explain the beta–binomial model we use to reduce sampling error in the licensed share. First, we estimate cell-level standard errors by frequentist and Bayesian methods. Section 3 of this Appendix develops two controls we use in Section 5 of the main text as robustness checks.

D.1 Estimating Cell-Level Standard Errors

In this subsection we present both Bayesian and frequentist approaches to obtaining a formula for the mean and the standard error of the leave-out state–occupation licensed share. Throughout this subsection, we define for notational convenience

\[ L_{os} = \sum_{i \in W_{os}} L_i, \]

where \( L_i = 1 \) if worker \( i \) is licensed and equals zero otherwise, \( s \) indexes states, \( o \) indexes occupations, and worker \( i \) is in \( W_{os} \) if they are in state \( s \) and occupation \( o \). \( L_o \) is defined analogously.

**Frequentist Approach.** The leave-out licensed share of worker is

\[ \%L_i = \frac{L_{os} - L_i}{N_{os} - 1}, \]

and using the formula for the variance of a Bernoulli random variable, we obtain the variance

\[ \sigma^2_{ui} = \frac{\%L_i(1 - \%L_i)}{N_{os} - 1}. \]

Two considerations weigh against a frequentist approach in our measurement error correction. First, we do not exploit information from licensed shares of workers in other states but the same occupation to reduce error. Second, the estimated cell-level measurement error is zero when all or no workers are licensed in the cell.

**Empirical Bayes Approach.** Following common practice in Bayesian statistics (Bolstad and Curran, 2016, Ch. 8), we propose to model the distribution of licensed and unlicensed workers across state–occupation cells as

\[ p_o \sim \text{Beta}(\alpha_o, \beta_o) \]
\[ L_{os} \sim \text{Binom}(N_{os}, p_o). \]

The first step is to calibrate \( \alpha_o \) and \( \beta_o \), the occupation-specific parameters of the prior distribution of the licensed share across state–occupation cells. We use the beta distribution because, as the conjugate distribution to the binomial, conditioning on the binomial count data of licensed and
unlicensed workers will yield a posterior that is also a beta distribution, a result we provide below. We estimate by maximum likelihood where possible and by method of moments where the likelihood estimation does not converge.

In the MLE procedure, we estimate a beta-binomial regression of $L_{os}$ on a constant, given observations $N_{os}$, using the canonical logit link function. For 164 of 483 occupations, this procedure yields negative estimates of $\alpha_o$ or $\beta_o$, particularly when there are relatively few licensed or total workers in an occupation. In these cases, we discard the estimates and re-estimate the parameters of the beta distribution by method of moments:

$$\hat{\alpha}_o = \frac{\mu^2_{1o} - \mu^3_{1o} - \mu_{1o}\mu_{2o}}{\mu_{2o}}$$
$$\hat{\beta}_o = -\frac{\mu^2_{1o} - \mu^3_{1o} - \mu_{1o}\mu_{2o}}{\mu^2_{2o} - \mu^3_{1o} - 2\mu_{1o}\mu_{2o}},$$

where $\mu_{1o} = L_o/N_o$ and $\mu_{2o} = \frac{1}{N_{os}} (L^2_{os} - L^2_o)$. This procedure nevertheless fails for 4 of 483 occupations. For these occupations, we assume the uninformative prior $\alpha_o = \beta_o = 1/2$ for the state–occupation licensed share.

We now use Bayes’ theorem to update the beta prior with the count data. Our assumption that counts of licensed and unlicensed workers in a state–occupation cell are drawn from a cell-specific binomial distribution implies:

$$p(L_{os}|N_{os}, \theta_{os}) = \binom{N_{os}}{L_{os}} \theta_{os}^{L_{os}} (1 - \theta_{os})^{N_{os} - L_{os}}.$$ 

With a constant $k$, our prior is

$$p(\theta_{os}) = k \theta_{os}^{\hat{\alpha}_o - 1} (1 - \theta_{os})^{\hat{\beta}_o - 1}.$$ 

By Bayes’ theorem,

$$p(\theta_{os}|(L_{os}, N_{os})) = k' \theta_{os}^{\hat{\alpha}_o + L_{os} - 1} (1 - \theta_{os})^{\hat{\beta}_o + N_{os} - L_{os} - 1}.$$ 

The posterior distribution for the state–occupation licensed share is therefore

$$\theta_{os}|(L_{os}, N_{os}) = \text{Beta}(\alpha_o + 1 + L_{os}, \beta_o - 1 + N_{os} - L_{os}).$$ 

The posterior mean is

$$\frac{\alpha_o + L_{os}}{\alpha_o + \beta_o + N_{os} - L_{os}}.$$ 

61
Figure A4: Bayesian Adjustment Affects Only Very Small State–Occupation Cells

and the posterior variance is

$$\frac{(\alpha_o + L_{os})(\beta_o + N_{os} - L_{os})}{(\alpha_o + \beta_o + N_{os})^2(\alpha_o + \beta_o + 1 + N_{os})}.$$  

The leave-out results in the text follow immediately. As the mean of the prior distribution is \(\hat{\alpha}_o/(\hat{\alpha}_o + \hat{\beta}_o)\), and the licensed share is \(L_{os}/N_{os}\), the empirical Bayes estimate of the licensed share is a convex combination of the prior mean and the licensed share, with the relative weight on the licensed share increasing in the number of observations in the state–occupation cell. Notably, as the sample \(N_{os}\) becomes large, the weights in the posterior shift away from the prior and towards the data.

D.2 Applying the Correction

We document the consequences of the empirical Bayes adjustment of cell licensed shares. As the number of observations in a cell increases, the implied weight on the prior declines to zero. In Figure A4, we see that the adjustment is generally small, and only of consequence for cells with very few workers. For cells with more than 10 workers, the average absolute difference between the raw leave-out-mean and the empirical Bayes estimate is about 0.03. We have truncated Figure A4 at 500 workers to make the small cells visible.
D.3 Additional Controls Used in Robustness Checks

Here we explain the occupation-mix and demographic-mix controls we use in our robustness checks in Section 5 of the main text, specifically in Table 3.

**Occupation Mix Control.** To explain our procedure, let $M$ be a matrix of employment shares whose columns are occupations and rows are states. Find the first $k$ principal components of the submatrix $M_{o^*,s^*}$, which deletes column $o^*$ and row $s^*$. Then, by this rotation, predict the principal component scores for all occupations but $o^*$ in the hold-out state $s$, and augment the matrix of principal component scores with these predicted scores. Using this augmented matrix, estimate the regression

$$s_{o^*s} = \sum_k \beta_k p_{ks} + e_s,$$

where $s_{o^*s}$ is the share of workers from state $s$ in occupation $o^*$ and $p_{ks}$ is the value of the $k$th principal component in $s$. For the hold-out observation $(o^*, s^*)$, predict $\hat{s}_{o^*s^*}$ by Equation 14. Repeat for all $(o, s)$ and use the log predicted value as a control. The resultant data capture the predictable variation in occupational employment shares across states from employment in other occupations in that state and correlations across occupations' employments in other states. For example, if some states with relatively many (few) farmers also tend to have relatively many (few) loggers, we would expect other states to respect this rural-urban pattern and would want to rule out the possibility that such patterns are used to identify causal effects of licensing. Our method is a “leave-out” strategy for predicting relative employment from such correlations.

We set $k = 5$, and Figure A5 depicts the results. Each panel of the figure assigns states to equal-frequency bins according to each of their principal component scores. We see strong regional and thematic patterns. PC1 is strongly correlated with population density, PC2 is East versus West, PC3 is North versus South, PC4 is high in the Pacific Coast and Deep South but low elsewhere, and PC5 is high in the Mid-Atlantic and Southwest but low elsewhere. Our control explains 18 percent of the “within” variation in log employment after state and occupation fixed effects. As reported in Section 5, we find broadly the same effects of licensing as in our baseline specification. This confirms that estimated employment effects are not confounded by correlations with broad features of the state occupational mix.

**Demographic Mix Control.** We predict state–occupation employment levels using a Bartik-like technique that combines the national occupational employment shares of a demographic group $d \in \{1, \ldots, K\}$ and the state shares of population of these demographic groups. For standard reasons, this predicted employment is formed via a “leave-self-out” method.

Let $L_{osd}$ be the employment count in occupation $o$ and state $s$ for workers of demographic type $d$. Let $L_{sd} = \sum_o L_{osd}$, $L_{od} = \sum_s L_{osd}$, $L_d = \sum_o L_{od}$ and $L_s = \sum_d L_{sd}$. Then our control is

$$\hat{L}_{os} = \sum_d L_{sd} \left( \frac{L_{od} - L_{osd}}{L_d - L_{sd}} \right).$$
Figure A5: Principal Component Scores from Occupational Employment Shares

Notes: This figure depicts the principal component scores for state shares of employment by occupation, therefore extracting the low-dimensional patterns in states’ employment mixes. In each of the five panels, states are ranked and colored according to their respective principal component score. The colors are in five equal-frequency bins.
This control explains about 11 percent of the residual variation in employment after removing state and occupation fixed effects. Together with the occupation-mix control, about 25 percent of the residual variation in employment is explained.

D.4 Bias Correction in Estimating Total Variation Distance

With \( k = 1, \ldots, K \) denoting a level of educational attainment, we define a treatment effect \( \beta_k \) as the percentage-point change in the share of workers with education \( k \) that is the causal effect of licensing. Total variation distance is defined as

\[
\text{TVD} = \sum_k |\beta_k|.
\]

Computing \( \hat{\text{TVD}} \) from estimates \( \hat{\beta}_k \) will be biased upwards, with the bias increasing in the standard error \( \sigma_k \) and decreasing in the absolute value \( |\beta_k| \). This is immediate from the case of \( \beta_k = 0 \) for all \( k \) but estimates \( \hat{\beta}_k \) estimated with any error: Estimated total variation distance is positive when true total variation distance is zero. Using the truncated normal distribution and unbiased estimators \( \hat{\beta}_k \) and \( \hat{\sigma}_k \), the analytical expression for this bias is

\[
\mathbb{E}[\hat{\text{TVD}} - \text{TVD}] = \sum_k \frac{\phi(|\hat{\beta}_k|/\hat{\sigma}_k)}{\Phi(|\hat{\beta}_k|/\hat{\sigma}_k)} \hat{\sigma}_k.
\]

In our application, we estimate \( \hat{\text{TVD}} = 0.1194 \) and \( \mathbb{E}[\hat{\text{TVD}} - \text{TVD}] = 0.0122 \), therefore \( \mathbb{E}[\text{TVD}] = 0.1072 \). Our bias-corrected estimate is therefore that 10.72 percent of workers obtained a different level of educational attainment due to licensing than they would have attained absent licensing requirements. Our uncorrected estimate is biased upwards by a factor of 1.11, implying that our estimate of total variation distance only slightly inflated by the effect of sampling variation.
References for Appendices

