Mixed Signals? Occupational Licensing and Intermittent Labor Force Participation

Matthew Brown

Advisor: Dr. Abigail Wozniak, University of Notre Dame

March 30, 2019

Abstract

Workers who attempt to return to the labor force after an interruption to labor force participation receive lower wages and are less likely to be employed than their peers. The quality signal of occupational licensing may give interrupted workers an opportunity to overcome the negative signal of their intermittent participation, but licensing also may create barriers to entry which are especially costly for interrupted workers. In this paper, I investigate the differential effects of licensing on interrupted individuals. I focus on broad measurements of licensing prevalence rather than individual licensing status to avoid bias which arises from the unobserved differences between licensed and unlicensed individuals. There is evidence that the presence of licensing has disproportionately negative effects on employment for interrupted workers, but I do not find a differential wage premium. I find that results which demonstrate the benefits of licensing for individuals who hold licenses do not extend to show benefits for individuals who live in highly licensed environments.

Acknowledgements: First and foremost, this thesis would not have been possible in any way without the guidance of Abigail Wozniak, who has also been a fantastic mentor. Gratitude is also due to Dante Domenella and Michael Smith for their insightful conversation and assistance with programming, and to Isabella Penola for her comments and proofreading.

I. Introduction

Roughly a fourth of American workers are required to hold an occupational license to practice their profession, a fraction which has been growing steadily since the 1950s. As licensing has become more common, it has increasingly become a subject of legislative focus and public discussion.[1] Economists have traditionally discussed licensing in terms of efficiency and the large-scale tradeoffs between benefits for license holders and costs for society, but recent research has focused on the distributional effects of licensing. Blair and Chung (2018a) find that the gender and racial wage gaps are about 40% smaller for licensed people. They explain this finding by arguing that the signal of licensing is especially helpful for members of groups that are negatively impacted by statistical discrimination. Other differential effects of licensing may arise from this framework. People who are attempting to return to work after an interruption to labor force participation have lower wages and are less likely to be employed than other labor force participants. It is possible that the signal effects of licensing allow interrupted workers to overcome the negative signal of their intermittent labor force participation. On the other hand, the barriers raised by licensing may be especially costly for interrupted workers.

This paper attempts to determine the effect of occupational licensing on the penalty from intermittent labor force participation. Because licensing acts as a signal and as a barrier, the direction of the effect is not immediately obvious. If the signal effect dominates, we would expect licensing to improve outcomes for interrupted workers, but if the barrier effect is stronger, then licensing would make outcomes worse. It is also possible that the two effects balance each other out and there is no effect on outcomes. Any effect would have consequences for the

---

[1] For an overview of some groups who have produced policy proposals, speeches, or academic papers, see the brief article "The Future of Occupational Licensing Reform," available at https://www.brookings.edu/opinions/the-future-of-occupational-licensing-reform/ . The bipartisan representation in this collection is remarkable.

distributional effect of licensing, since intermittent labor force participation is more common in some demographic groups than in others.

To study the relationship between licensing and intermittent labor force participation, I use data from two nationally representative surveys, the Survey of Income and Program Participation (SIPP) and the Current Population Survey (CPS), which include questions about licensing. I begin by comparing state and occupation level licensing prevalences calculated from the two datasets with each other and with previous estimates of licensing prevalence drawn from proprietary phone interview surveys. While national statistics appear similar, the correlation between state-level measurements of licensing prevalences from the SIPP and the CPS is surprisingly low. I eliminate some possible reasons for the difference, but do not discover an explanation. The differences between these two measurements and the previous, proprietary estimate of state-level licensing prevalence from Kleiner and Vorotnikov (2017) are even more striking. These discrepancies cast some doubt on past research which relied on data from similar small, proprietary surveys.

I construct empirical models to isolate the differential impact of licensing on interrupted workers, one for each survey. Throughout my analysis, I modify Blair and Chung's framework to focus on the effects of the *prevalence* of licensing in states and occupations rather than the effects of the licensing *status* of individuals. This allows me to eliminate many of the confounding factors which are associated with individual licensing decisions. In the first model, I sort a sample of labor force participants from late 2012 in the SIPP into interrupted participants and noninterrupted participants. Using this sample, I regress wage and employment dependent variables on an interrupted dummy, state licensing prevalence, and an interaction term. The coefficient on the interaction term corresponds to the differential effect of licensing on interrupted workers. For my second model, I use the CPS to remedy two deficiencies in this SIPP model. First, I use the CPS's large sample size to calculate licensing prevalence at the more specific state-occupation level. Second, I calculate licensing prevalences at different points in time and write a differenced model, based on group employment shares, which eliminates cross-sectional sources of bias. For both models, I implement alternative specifications using groups whose members are likely to have faced interruptions (mothers of young children and near retirees) as identifiers of intermittent labor force participation. This strategy accounts for the endogeneity of the labor force participation decision, and it also allows me to study the effects of licensing on labor force participation.

My regression results are suggestive, but I am unable to draw definitive conclusions. As expected, interrupted workers are less likely to be employed and earn lower wages that their peers. I find some evidence that the presence of licensing has a disproportionately negative effect on employment for interrupted workers. This suggests that the barriers to entry created by licensing for interrupted workers outweigh the positive effects that come from licensing's power as a signal. My analysis of near retirees supports the idea of powerful barriers to re-entry, as individuals seem to retire later in states with high licensing prevalence, possibly because they are accounting for the fact that it will be more difficult for them to return to work. I do not find differential licensing wage premia for interrupted workers in any of my tests.

I find that many effects of licensing do not persist when I focus on broad measures of licensing prevalence instead of individual licensing status. Specifically, I find large and significant employment and wage benefits for individuals who hold licenses, but there is no evidence that living in states with high licensing prevalences provides any positive effect. This may indicate that individuals who choose to pursue licenses are different in unobserved ways from individuals who do not, which biases estimates of the benefits of licensing upwards. It is also possible that licensing does provide significant benefits, but state level aggregation produced a measure of licensing prevalence that was too broad to find any significant results. Indeed, these two possibilities are not mutually exclusive.

II. Defining and Measuring Licensing

Occupational Licensing in the United States

There is no universal framework for occupational regulation in the United States. In fact, there is significant variation in regulatory regimes across both states and occupations. I begin by defining the main types of regulations that this paper studies. Workers are required to hold *licenses* to practice their jobs legally. In a licensed profession, any unlicensed worker commits a crime and may be punished accordingly by the state. Less commonly,[2] occupations may have *certifications* available to workers. These may be issued by the government or by a private organization. The fundamental difference between the two systems is that in licensed professions, the government uses licenses to grant certain individuals the "right to practice," while certifications serve only as a signal of quality – individuals who do not hold a certification are still permitted to work legally. Using the interpretation of Berk and Van Binzbergen (2017), certification can be characterized as voluntary "information disclosure" and licensing as a mandatory "minimum standard." For this reason, occupational certification is much less restrictive than occupational licensing.

Both regulatory systems are designed to allow workers to demonstrate that they have achieved a certain level of competence. There is often an examination associated with attaining a

[2] Gittleman et. al (2015) estimate that 20% of American workers are licensed, while 8% are certified.

license or certification. However, the difficulty of fulfilling licensing and certification requirements varies significantly by profession. It takes on average over 4 and a half years of education and experience to become a licensed preschool teacher, but a taxidermist's license can be obtained with no experience at all (Institute for Justice, 2012). Even within an occupation, the burdens of licensing vary significantly by state. In Tennessee, for example, 4 years of education, 3 years of work experience, and five passed examinations are required for a preschool teacher to obtain a license, while Ohio issues licenses for preschool teachers after only 2 years of education and one passed examination (ibid). There are also monetary costs: fees paid to the entity providing the license vary greatly.

No matter how it is defined, occupational licensing affects a significant, growing portion of the American labor force. A national labor force survey conducted in 2008 estimated that 35 percent of American workers hold either a license or a certification (Kleiner and Krueger, 2013). Kleiner and Krueger's analysis of census data shows that this high level of licensing is a recent phenomenon: they estimate that in the 1950s, less than 5 percent of the American labor force was licensed (2013).

There is a dearth of high-quality data related to occupational regulation. Until recently, no nationally representative survey collected information about licensing and certification at the individual level. Researchers were forced to base their analyses on classifications of occupations by licensing status on the state level. Such studies were necessarily "low-resolution," and they also missed any licensing which occurred on the federal or local levels. Kleiner and Krueger (2010 and 2013) conducted the first nationally representative individual-level surveys of licensing in the late 2000s, with sample sizes in the thousands. Two nationally representative surveys, the Survey of Income and Program Participation (SIPP) and the Current Population

Survey (CPS), have recently introduced questions pertaining to licensing. The sample sizes in these government surveys are an order of magnitude larger than those in the surveys conducted by Kleiner and Krueger, and I use both of them in my empirical analysis.

When analyzing survey data, I define a "licensed occupation" as an occupation in which some sort of license or certification is "required" to practice – whether this is a legal obligation or a practical one. I do not consider differences in the burdens of attaining licenses. It is important to note that this is a dramatic simplification of reality, but one which is made necessary by the limitations of available data.

Theories of Licensing and Its Adoption

The study of occupational licensing is at least as old as the field of economics itself – Adam Smith discusses the regulation of occupations in *The Wealth of Nations.* Smith, writing in the late 18th century, sharply criticized the contemporary system of apprenticeship, which had its origins in the medieval guild system. In some professions, novices were forced to submit themselves to the training of masters for several years before they were legally allowed to practice a craft. Smith argued that these regulations had a detrimental effect on the public and served only to benefit the legally established members of the profession. In the following passage, he considers the possible implications of the repeal of apprenticeship laws:

> The master, indeed, would be a loser. He would lose all the wages of the apprentice, which he now saves, for seven years together. In the end, perhaps, the apprentice himself would be a loser. In a trade so easily learnt he would have more competitors, and his wages, when he came to be a complete workman, would be much less than at present. The same increase of competition would reduce the profits of the masters as well as the wages of the workmen. The trades, the crafts, the mysteries, would all be losers. But the public would be a gainer, the work of all artificers coming in this way much cheaper to market (Smith 1776).

For Smith, apprenticeship laws artificially restrict the labor supply and increase the wages of the privileged few who hold the right to practice. Moreover, the laws deny workers the right to move freely between professions. Smith calls the apprenticeship “a manifest encroachment upon the just liberty both of the workman, and of those who might be disposed to employ him” (ibid). Many economic thinkers through history have taken up this theme and condemned licensure, both as a system which inefficiently allocates resources and as an affront to liberty. Famously, Milton Friedman devoted a full chapter of *Capitalism and Freedom* to occupational licensing. Many professionals in licensed occupations argue that regulations are in fact beneficial, because they ensure that only high-quality service will be provided. Friedman is skeptical that licensing has this effect, but he argues that *even if* licensing imposes a quality floor, it still may be detrimental to society. He makes an analogy between the labor market and the market for goods. Suppose that a law is passed which requires any car that is sold to be at least of the quality of a Cadillac. This law would effectively abolish a large segment of the market in which consumers demand lower-quality cars. Similarly, if licensing ensures that only “high quality” workers may practice, then employers or clients who only need cheaper, less skilled workers to achieve their ends are forced to pay more than what they would in the absence of regulation. This has the general effect of driving up wages for those skilled few who obtain licenses as “lower quality” workers are prohibited from selling their services (Friedman, 1962).

Because of this supply restriction, high-quality workers have an interest in obtaining and maintaining occupational regulations for their profession. An interesting question arises which is relevant to this work and all investigations of the impact of licensing: Under what conditions are these groups likely to succeed? The optimistic *public interest theory* holds that licensing is the resolution of a market failure. The government may intervene in markets when informational asymmetries make it impossible for consumers to determine the difference between trained professionals and charlatans (Graddy 1991). This is almost always the basis on which licensing is justified. A classic example is the medical profession – it may be in the public interest for an organization to thoroughly vet physicians so that there are no “bad actors” peddling false medical knowledge. This theory runs into difficulty, however, when we notice state-level variation in licensing within an occupation. Auctioneers, for example, are licensed in 30 states (Institute for Justice). Is there a compelling reason that the public interest would require the regulation of auctioneers in North Dakota, but not in South Dakota?

There may be more insidious motives at work. In his landmark paper on economic regulation, Stigler (1971) recognizes that industries who understand the potential benefits of regulation actively lobby governments for it. Because the burdens of the labor supply restriction are widely spread, it is often difficult to muster a cohesive opposition to these lobbies. As Maurizi puts it (emphasis mine):

> Occupational licensing has been justifiable in the view of legislatures on the grounds that it protects the public interest; often, however, *it is the producers of the good or service who present this argument to the state legislatures*. This is hardly surprising, since the typical consumer is likely to suffer too small a wealth loss (in the form of higher prices) … The end result is the promotion of the interests of the producer group rather than those of the public. (Maurizi 1974)

In this *interest group theory*, regulations arise out of a political process in which organized groups capture governmental bodies and use their influence to extract surplus from the economy. In the specific case of licensing regulations, this occurs through an artificial restriction of the labor supply. There is some evidence that interest groups play an important role in the licensing process. Kleiner (2000) notes that even when licensing boards are created by the state, they are generally dominated by industry members. However, Graddy’s (1991) analysis suggests that many diverse factors play a role in the origin of licensing regimes. It is difficult to disentangle the public interest from the private interest of small groups in the messy political process.

Licensing Premia and Distributional Effects of Licensing

While there is still uncertainty about the mechanisms by which licensing requirements arise, recent research has convincingly demonstrated the existence of a licensing premium. Kleiner and Krueger (2010, 2013) use data from two telephone surveys to calculate a licensing premium. In the first study they find that licensing is associated with a 15% increase in wages, and in the second they see an 18% premium. Gittleman, Klee, and Kleiner (2018) examine data from the SIPP, and they find that licensed workers earn 23.6% more than unlicensed workers. However, once they control for occupational fixed effects, the licensing premium falls to 6.5%. There are certainly significant benefits to holding a license, but it is important that we do not confuse the licensing premium with the differential which naturally arises if licensed occupations are different from other occupations for reasons which are unrelated to the license itself. In general, estimates of the licensing premium which account for demographic, educational, and occupational differences between licensed and unlicensed workers are smaller than “raw” estimates.

Because license holders are not evenly distributed across demographic groups, this premium can have distributional effects. For example, licensed workers are more likely to have completed a bachelor’s degree than their unlicensed peers, and they are less likely to be black (Kleiner and Krueger, 2013). To the extent that licensing artificially increases wages after controlling for differences in human capital, this differential results in lower incomes for groups which have lower licensing prevalences. There is some evidence that these differentials are not only due to differences in demand for licenses and that licensing requirements are more burdensome for some groups than for others. Dorsey (1980) finds that black and nonnative individuals are more likely to fail cosmetology examinations than white natives, even after controlling for levels of education and training. Because these rejected applicants are not evidently less competent than successful applicants, he concludes that the examinations themselves may be biased. Licensing also raises barriers to entry (Stiglitz 1971) for all workers, and it may be more difficult for underprivileged people to take on the sunk costs associated with obtaining a license.

However, there is also some evidence that licensing reduces inequality between demographic groups. Law and Marks (2009) find that in the Progressive Era (late 19th to early 20th century), licensing tended to increase the representation of blacks and women in traditionally white male occupations, especially in occupations where worker quality was difficult to ascertain. Redbird (2017) finds a similar trend when studying modern changes in licensing policy. She argues that members of disadvantaged groups may face informal barriers in unlicensed environments, in which employment is often the result of social networking or employer preferences. These people would disproportionately benefit from licensing, because it creates a codified path of entry which allows them to bypass these informal barriers. Blair and Chung (2018a) examine heterogeneity in the licensing premium using recent survey data. They show that the licensing premium is larger for black workers and females than for white workers and males. The gender and racial wage gaps for licensed workers are roughly 40% smaller than the gaps for unlicensed workers.

Data Sources

The two main data sources for this paper are the 2008-2012 panel of the Survey of Income and Program Participation (SIPP) and the Current Population Survey (CPS). The SIPP and the CPS are the two largest nationally representative surveys which asked respondents questions about occupational licensing. The SIPP and the CPS have different survey structures, so the two surveys capture different types of information. The longitudinal nature of the SIPP makes it possible to track a sample of respondents over a period of four years. The CPS only tracks respondents over a period of 16 months, but its larger sample size and consistent structure allows me to study specific licensing environments more closely.

In this subsection, I provide some background on these surveys. I discuss the language used in specific questions, and I explain how I identify licensed individuals and individuals with intermittent labor force participation in the data set. I report summary statistics of variables related to licensing. I compare the results from the CPS and the SIPP both with each other and with aggregate measures from an earlier proprietary phone survey. One of the contributions of this paper is a close examination of the comparability of licensing measurements from different surveys. I find that measurements of licensing prevalence vary considerably across the surveys, especially on the state level.

Survey Characteristics and Language of Licensing Questions

The SIPP is a nationally representative survey which follows households over a 4 year period.[3] Households are contacted in 16 “waves” (every four months) to answer a set of questions about demographic characteristics, work experience, earnings, program participation, and income called the core survey. The SIPP is longitudinal by nature – it is possible to track the

[3] More detailed information can be found in the SIPP User’s Guides, available at https://www.census.gov/programs-surveys/sipp/guidance/users-guide.html

answers of households and individuals to the core survey questions across the waves. Many waves also include a one-time-only topical module as a supplement to the core survey. These topical modules usually ask several questions about one specific area, and they allow researchers to conduct detailed analysis for a variety of subjects. I use data from wave 1 and wave 13 in this analysis. At wave 1, the 2008 SIPP collected responses from 105,663 individuals, and at wave 13 there were 76,034 individuals.

The 2008-2012 panel of the SIPP, which began in the second half of 2008, was the first and only version to include questions about occupational licensing. In the Professional Certifications and Educational Certificates topical module, which was administered in late 2012 as part of wave 13 of the SIPP, respondents who were over the age of 16 were asked several questions about the licensing status of people in the household, including:

(i) Did [name] have a professional certification or state or industry license? (Answer options: Yes, No)

(ii) Who awarded this certification or license? (Answer options: Federal Government; State Government; Local Government; Industry; Business, Company, or Non-Profit Organization; Professional Organization; Other)

(iii) Is this certification or license required for [name]'s current or most recent job? (Answer options: Yes, No)

I classify individuals as *licensed* if they answered “Yes” to (iii); that is, if they say that their certification or license is required. I link these variables from the topical modules to the core modules, which includes data on employment and demographics. Using this merged data set, I construct a measure of the *licensing prevalence* in a given occupation or state. Licensing prevalence is defined to be the fraction of members of a group in the labor force that are licensed. For example, the licensing prevalence in Arizona is equal to the number of people in the Arizona labor force who are licensed divided by the number of people in the Arizona labor force. I define members of the labor force to be individuals who either had a job or looked for a job during the last four months. For calculations of licensing prevalence, I use the SIPP sample weights to adjust for sampling error.

The CPS is a monthly nationally representative survey conducted by the Bureau of Labor Statistics. Each household in the sample is surveyed a total of eight times. There are interviews in each of the first four months after the first contact, then eight months of no interviews, and then four more months of interviews. A new rotation group is added to the CPS in each month, so in any given month, the proportion of people in each month of their rotation in the CPS sample is roughly equal. That is, of the roughly 128,000 CPS respondents in each month, approximately 16,000 are interviewed for the first time, 16,000 for the second time, and so on. Compared to the SIPP, the CPS has the advantage of larger sample size, but it follows its respondents for a shorter period.

Starting in 2016, the CPS began to include three questions about occupational licensing in the first and fifth surveys of its rotation.[4] Each respondent over the age of 16 answers these questions. There are roughly 14,000 labor force participants between the ages of 18 and 64 who respond to these questions in each month. The relevant questions are:

(i) Do you have a currently active professional certification or a state or industry license? Do not include business licenses, such as a liquor license or vending license. (Answer options: Yes, No)

[4] Source: https://www.bls.gov/cps/certifications-and-licenses-faqs.htm#questions

(ii) Were any of your certifications or licenses issued by the federal, state, or local government? (Answer options: Yes, No)

(iii) Earlier you told me you had a currently active professional certification or license. (Is/was) your certification or license required for your (job/main job/job from which you are on layoff/job at which you last worked)? (Answer options: Yes, No)

While the exact wording differs, these questions are very similar to the three questions I chose from the SIPP topical module. Indeed, in the hopes of generating comparable results, both surveys constructed their sections about occupational certification with reference to a report about data collection published by the Interagency Group on Expanding Measures of Enrollment and Education from the U.S. Department of Education.[5] As before, I define an individual to be *licensed* if they answer yes to (iii). Just as I did in the SIPP, I generate measurements of licensing prevalence for states and occupations by calculating the weighted means of the licensed variables for members of the labor force in a state or occupation. For the CPS, I define members of the labor force as people who either were holding a job or seeking work during the reference period of the past month.

While the SIPP only asked questions about occupational licensing at one point in time, the CPS has continued to ask questions about occupational licensing in each month from early 2016 until the present day. This continuity allows me to see how state and occupation level licensing prevalence measurements changed over time in the CPS. I choose January-June 2016 as a starting point because these were the first months which included questions about licensing.

[5] More detail on this report can be found at: https://nces.ed.gov/surveys/GEMEnA/surveys.asp

My end point is January-June 2018, because this was the latest year for which these questions were available at the time of writing.

The CPS and SIPP both determine the occupations of respondents. The CPS identifies the occupation of any respondent over the age of 15 who has been in the labor force within the last 5 years, while the SIPP only identifies the occupation of individuals who held a job during a reference period. They both use 4-digit Census occupation codes to categorize occupations. In 2010, the Census Bureau modified its occupation code scheme. The CPS years that I am examining use the new scheme, but the 2008 SIPP used an earlier scheme. When I compare specific occupations across surveys, I implement a crosswalk published by the Census Bureau to harmonize the definitions.[6]

Proxy Responses

The SIPP and the CPS include proxy responses. Both surveys would prefer individuals over the age of 15 to answer all interview questions personally. However, due to subject uncooperativeness or unavailability, an interview is often not possible with every member of a household. In such situations, the interviewer attempts to conduct a proxy interview with a responsible adult member of the household who is knowledgeable about the labor force activities of the absent household members. Data collected from such an interview are known as *proxy responses,* as opposed to *self responses.* Some adult members of a household may not know that one of the other members of the household holds an occupational license. If this is frequently the case, then measures of licensing prevalence garnered from proxy responses will understate the true licensing prevalence. Moreover, if the SIPP and CPS have different proxy procedures, bias

[6] This crosswalk is published at: https://www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html

from proxy responses could cause measurements of licensing prevalence to differ between the surveys.

In the CPS, 50 percent of responses are proxy responses, while in the SIPP the number is 39 percent. In Tables 1 and 2, I report the results of t-tests which compare the characteristics of proxy responses and self responses for the two surveys. In both surveys, proxy responses are more likely to be male, have lower educational attainment, are younger, and are less likely to have an occupational license. These differences are statistically significant, with large t-statistics. In both surveys, there is not a large difference in employment, but proxy responses are more likely to be have experienced an interruption to their labor force participation.[7]

These tables provide evidence that that self responses and proxy responses differ in systematic ways. However, I cannot conclude from the raw difference in means that proxy responses underreport licensing relative to self responses *ceteris paribus.* The demographic differences between the two groups depress the expected fraction of licensed individuals in the proxy group. Because males are less likely to be licensed, young people are less likely to be licensed, and so on, we would expect to see a lower fraction of licensed individuals in the proxy group even if there was no systematic underreporting. To isolate the effect of being a proxy response from demographic effects, I regress the licensing dependent variable on a proxy dummy variable and a vector of demographic characteristics for both the SIPP and the CPS. The coefficient $\hat{\beta}$ on the proxy dummy is the corrected estimate of the reporting bias from proxy responses. One interpretation of this estimate is that for a given sample, I would expect the

[7] I go into detail about my definition of interruption in Section IV.

difference between the number of people who would be reported as licensed if all responses were proxy responses and the fraction of people who would self-report as licensed to be $\hat{\beta}$.

The corrected estimate of reporting bias is -.002 in the SIPP and -.008 in the CPS, and neither effect is statistically significant. This analysis suggests that demographic differences account for almost all of the difference in licensing reports, and underreporting accounts for at most a negligible fraction. I conclude that any differences between the SIPP and the CPS measurements of licensing are not the results of different proxy procedures.

Summary Statistics and Comparisons Between Surveys

The SIPP and the CPS use similar language in their questions, which allows us to define licensing in a way which should be at least roughly comparable across the two surveys. In Table 3, I report summary statistics from the two surveys for all workers and for licensed workers. I draw this group of workers from the same sample that I use in my empirical analysis. We can draw several insights from Table 3. At least at this broad level, the two surveys give similar measurements of national licensing prevalence among workers, which lie between 20 and 25%. Licensed workers are more likely to be female, be well-educated, or belong to a union than the average worker. They also earn higher wages. In both surveys, licensing prevalence varies significantly from state to state. Using SIPP estimates, Alaska is the most-licensed state, with 46.9% of labor force participants reporting that they are licensed. The lowest licensed state is New Mexico, where only 11.3% of labor force participants are licensed. Figure 1 shows this variation using state level licensing prevalences from the SIPP. States with high and low licensing prevalence are found in all regions of the country, which suggests that levels of licensing are not the results of geography or regional culture, but of idiosyncratic state policy.

I compare my two data sources to proprietary survey data which has been used for analysis in the past. Before the SIPP and the CPS asked questions about licensing, the most advanced data sources about the topic came from telephone surveys conducted by Kleiner and Krueger in 2006 with the Westat corporation. While I do not have access to the raw survey data, I obtain measurements of state-level licensing prevalence from this survey, published in Kleiner and Vorotnikov (2017). I use the Westat results as a comparison group for the SIPP and CPS measurements.

In the Westat survey, 2,513 adult members of the labor force were interviewed as part of a random digit dial sample design. The survey asked two questions relevant to this paper:

(i) Do you have a license or certification that is required by a federal, state or local government agency to do your job? (Answer options: Yes, No)

(ii) Would someone who does not have a license or certificate be legally allowed to do your job? (Answer options: Yes, No)

In their analysis of the survey data, Kleiner and Krueger (2013) define workers to be *licensed* if they answer yes to (i) and no to (ii), and *certified* if they answer yes to both (i) and (ii). To generate a licensing prevalence variable which is comparable to the one I define for the CPS and the SIPP, I simply add the fraction of people who are licensed to the fraction of people who are certified. This definition differs from the ones I use for the SIPP and the CPS in important ways. First, the Westat survey attempts to identify only those people for whom a license is "required," while this characteristic is a different question in the SIPP and CPS. Perhaps most importantly, the SIPP and CPS make no mention of whether a "government agency" issued the license.

While the language in the Westat survey differs from that in the SIPP and the CPS, we would still hope to see broadly similar characteristics of licensing prevalence across the three surveys. As a consistency check, I compare state licensing prevalence in states that I generate in the SIPP and the CPS to each other and the reported Westat values.

In general, the measurements of licensing prevalence across the three surveys are only weakly correlated at the state level. The correlation coefficient between measurements of state-level licensing prevalence from the SIPP and those from the CPS is $\rho = 0.29$. When comparing these measurements to the Westat survey, the correlation decreases: $\rho = 0.19$ between the Westat survey and the SIPP, and $\rho = -0.02$ between the Westat survey and the CPS. Figures 2, 3, and 4 show the relationship between these three measurements in graphical form. The fact that ρ is highest when the CPS and SIPP measurements are compared to each other is comforting; it makes sense that Westat survey results would be dissimilar from the CPS and SIPP results because of differences in the question language. Nevertheless, the large differences between these measurements give us pause. Table 4 presents one particularly alarming result: there is very little overlap between the top five and bottom five states by licensing prevalence across our three different measurements. Nevada, for example, appears in the bottom five of both the CPS and SIPP measurements and the top five of Westat.[8]

I conduct a similar exercise to check consistency between surveys of measurements of licensing prevalence across occupations. Because I do not have occupation level data from the Westat survey, I limit my analysis to the CPS and the SIPP. I am interested in the consistency of the measurements both over time (though the SIPP only measures licensing prevalence at one

[8] Because of the preceding section, I conclude that these differences are not due to proxy responses. Proxies did not have large impacts in the CPS and SIPP, and the Westat survey did not include any proxy responses.

point in time) and across the two surveys. In Table 5, I report three measurements of the fraction of licensed workers for eight large commonly licensed occupations: one from the SIPP in late 2012, one from the CPS in early 2016, and one from the CPS in early 2018.

Many of the occupations listed in Table 5 "should" report licensing fractions of 1, or at least very close to 1. Lawyers, for example, must receive a license from their state's bar association to practice law. We would expect there to be very few (if any) lawyers in our data set who do not have a license. Yet in the CPS we observe around 15% of lawyers reporting that they do not hold a license, and in the SIPP the number is over 30%. This fact suggests that both surveys likely under-count licensed workers. The lawyer example also demonstrates another trend: the SIPP appears to count fewer licenses than the CPS. In every occupation listed, the fraction of workers licensed in the SIPP is higher than the fraction of workers licensed in the CPS. This result is peculiar, as national summary statistics in Table 3 demonstrate that the SIPP and CPS both calculate the fraction of licensed workers at close to the same level. Figure 5 is an overlapping histogram of licensing prevalences at the occupational level. It displays each of the three measurements from Table 5. We see that the SIPP measurement includes fewer occupations at the right end of the distribution with licensing prevalence of >.8.

III. Framework: Licensing as a Signal and Barrier

Signaling Effects, Intermittent Labor Force Participation, and Licensing

In Friedman's Cadillac analogy, restrictions on the sale of low-quality vehicles make everyone worse off. Friedman assumes perfect information – consumers are able to tell the difference between a good car and a bad car before they purchase it. His argument precedes George Akerlof's famous paper about "The Market for Lemons" (Akerlof 1970). Akerlof demonstrates that if consumers are not able to determine a car's quality, it is in the salesman's interest to sell only bad cars. Eventually, all those selling good cars will exit the market, leaving only the so-called "lemons." This is an example of market failure called *adverse selection* which is common when one side of the transaction does not have complete information. To alleviate this concern, sellers of good cars will look for ways to signal that their cars are high quality and retain their market niche.[9] In the labor market, where quality is often highly uncertain, individuals (the "sellers of labor") behave similarly.

When Spence (1973) introduced the concept of job market signaling, he used education as an illustrative example. Spence demonstrates that even if education does not improve the productivity of a worker in any way, it is possible for a wage gap to arise between educated workers and non-educated workers. Specifically, if it is assumed that college is more costly[10] for low-productivity workers to complete than for high-productivity workers, then employers who are uncertain about worker quality may view college-educated workers as more likely to be high-quality than other workers and compensate them accordingly. Licensing fits neatly into this signaling framework. Even if the process of licensing added nothing to a worker's productivity, because of the associated tests and costs, licenses may provide a signal of worker quality.

There are often multiple signals of worker quality operating in a labor market. When hiring workers, employers seek signals which provide information about applicants with the greatest possible certainty. For example, Wozniak (2007) examines the signal effects of drug testing. In labor markets where drug testing became prominent, labor market outcomes for low-

[9] Recalling Friedman's example, Akerlof notes that branding is one of these signals – Cadillac has staked its reputation on selling high-quality cars.
[10] Thinking of cost not necessarily in terms of dollars, but of effort expended – it is harder for a low-skilled person to complete a task than for a high-skilled person

skilled black males improved relative to other workers. It appears that employers had been statistically discriminating on the basis of race, but they discarded this low-certainty signal once they had access to the high-certainty signal of drug testing.

The wage penalty for intermittent labor force participation has been convincingly established by several studies (Hotchkiss and Pitts, 2005; Hotchkiss and Pitts, 2007; Stratton, 1995). A portion of this penalty is driven by losses in human capital associated with exit from the labor force. Individuals who leave the labor market are less likely to receive on-job training which makes them more productive, and they are mechanically less likely to attain high-paid senior positions which are associated with long tenures. As an analogy, consider what happens to professional athletes when they stop practicing their sport – their skill level slowly erodes. Interrupted workers may also see the value of skills depreciate over time, particularly in the case of long absences. Consider, for example, the case of a computer programmer who has been out of the labor market for many years – it is possible that the areas in which she has expertise are no longer in high demand. I call these and other effects of interruptions on productivity the human capital effects of intermittency. However, to focus only on the effects of interruptions on workers ignores half of the labor market. Interruptions to labor force participation also provide a signal effect, as employers infer information about quality from applicants' histories.

There are two primary channels through which this signal may depress outcomes for interrupted workers. First, employers may statistically discriminate against interrupted workers on the grounds that, on average, they are more likely than non-interrupted workers to have suffered the sorts of human capital losses discussed above. This signaling effect is subtly different from the human capital effect, because it affects *all* interrupted workers, not only those who have actually seen their job skills depreciate (the signaling effect is detrimental even to a computer programmer who has independently remained up to date in his coding skills during an extended absence from the labor market). Second, employers may be unwilling to commit resources to unattached workers. That is, if an individual has exited the labor market once, employers may reason that he is likely to do so again in the future and behave accordingly. The first channel is related to employers' beliefs about the person's productivity while he is working; the second is related to beliefs about the chances of the person leaving the job.

It is difficult to disentangle the signaling effect from the human capital effect. Some resume audit studies provide evidence supporting the existence of a signal. Farber et al. (2018) find that job applicants who had not been working for more than a year have a lower callback rate than similar applicants who had not been working for shorter periods of time. Another resume study showed that applicants who were unemployed for 18 months were perceived as "less qualified" than comparable applicants who had been unemployed for shorter periods of time (Shore and Taschian, 2013).

An occupational license has the potential to "undo" the negative signal of intermittent labor force participation. The first channel through which the signal operates is fundamentally a quality signal: employers may believe that uninterrupted workers are better than interrupted ones because they have not experienced the human capital losses associated with intermittency. Yet licensing also provides a quality signal, since a worker would (presumably) be unable to pass a licensing test if she were not competent in her occupation. In this sense, a license is an informational tool: it allows high quality interrupted workers to *prove* that they are indeed high quality. The existence of such a tool means that employers are not forced to rely so heavily on the imperfect practice of statistical discrimination against workers with labor force interruptions. Of course, interrupted workers are not a monolithic group. In fact, there is a potential for divergent outcomes within the set of interrupted workers. Interrupted workers who are unable to receive a license may be at an even greater disadvantage than they were before – they now send *two* negative signals to employers. Licensing allows high quality interrupted workers to separate themselves from their peers.

The costs associated with licensing may also serve as a commitment guarantee, which helps to undo the second channel through which intermittent labor force participation provides a negative signal. In this case, the important fact about licensing is not that only high *quality* workers can pass a licensing test, but that only highly *committed* workers would be willing to take on the costs of attaining a license. Obtaining a license is expensive both in dollars (fees paid to the licensing organization) and in time (educational and experience requirements). The costs of licensing are sunk costs, since there is no way to obtain a refund for licenses. A worker who assigns a high probability to leaving the labor force within a short period of time is unlikely to devote the money and time necessary to obtain a license. It only makes sense for him to impose the sunk costs of licensing on himself if he believes that he will see a significant payoff to his investment, probably over a long period of time. Understanding this dynamic, employers may assume that workers who have obtained licenses are more committed to long-term labor force participation than those who have not. The existence of this commitment mechanism mitigates the effect of the second channel, because interrupted workers are able to demonstrate to employers that they are unlikely to exit the labor force. This commitment mechanism has a greater effect when the costs of licensing are high, because higher sunk costs allow workers to prove their commitment with more certainty.

Licensing as a Barrier

I have demonstrated with economic reasoning that licensing may "undo" the negative signal of intermittent labor force participation. There is some empirical evidence which supports this hypothesis, most notably the finding that the gender pay gap is smaller for licensed individuals than for unlicensed individuals (Blair and Chung, 2018a). However, there may also be negative aspects of licensing for interrupted workers. While Redbird (2017) notes that licensing may help disadvantaged groups through the removal of informal barriers to entry, licensing does introduce legal barriers to entry. The burdens of these new barriers may fall disproportionately on interrupted individuals, because they may be less able to navigate the arcane bureaucracies associated with licensing.

Occupations may be licensed on the federal, state, or local level. Sometimes tests are administered by professional organizations, sometimes by government agencies. The byzantine structure of regulatory agencies and licensing boards often make the path to obtaining a license unclear. The stresses and time costs associated with navigating this complex system deter workers from seeking new licenses. Many studies have shown that licensing reduces interstate migration, as workers who hold licenses in their home state avoid becoming licensed in a new state (Peterson et al., 2014; Kleiner et al., 1982; Johnson and Kleiner, 2017). Individuals who are deciding whether to return to the labor market are highly sensitive to these burdens. For example, the Learning Policy Institute conducted a study on former public school teachers who were considering re-applying for certification. 41% of individuals cited "easier and less costly renewal of teacher certification" and "state certification reciprocity" as an "extremely" or "very important" factor in their decision (Learning Policy Institute, 2016).

In general, current labor force participants are more likely to have professional networks and institutional knowledge than those who are applying for work after an absence. They may have mentors in their chosen profession who encourage them to become licensed and guide them through the onerous process. By contrast, interrupted individuals have less social capital in the profession, which may force them to navigate the system on their own. Along with these informational barriers, licensing often explicitly privileges current workers through "grandfather clauses." These provisions are common when licensing laws are passed, and they exempt current members of the occupation from the licensing requirement. In the case of grandfather clauses, the burden of obtaining a license falls exclusively on workers who are trying to enter the profession.

A priori, it is not obvious whether licensing would improve or worsen the outcomes of interrupted workers relative to their uninterrupted peers. Licensing may allow interrupted workers to overcome employers' negative beliefs about intermittent labor force participation. It is also possible that licensing introduces barriers to entry which are more difficult for interrupted workers to overcome. The empirical models in this paper are designed to answer this question by isolating the differential effects of licensing on interrupted workers.

IV. Empirical Methods and Identification

In this section, I construct two models which allow licensing to have different impacts on interrupted and uninterrupted workers. The models exploit the unique features of each dataset. The longitudinal nature of the SIPP allows me to focus closely on interrupted individuals, while the larger sample size and temporal consistency of the CPS allows me to more accurately identify changes in specific licensing environments. In the SIPP model, I identify individuals who have experienced any interruption in the labor force in the last 4 years and study the differential impacts of cross-sectional differences in state licensing prevalence on these individuals. The CPS model, contrastingly, measures changes in licensing prevalences on the state-occupation level, but it does not follow specific individuals over long periods of time. Rather, it examines how changes in licensing prevalence affect aggregate labor market outcomes for interrupted and uninterrupted people. My measure of interruption is endogenous, so I also examine two "more exogenous" groups whose members are likely to have experienced intermittent labor force participation: mothers with young children and near retirees. These analyses also allow me to examine differential effects of licensing on labor force participation.

Defining Interrupted Workers

To detect intermittent labor force participation for an individual, I must look for periods of labor market non-participation in the worker's past. But before I examine past data, I must decide on a reference point – to define the past, I need to call something the present. In the following subsection, I describe how I define this reference point and create a category of interrupted workers.

In the SIPP, my reference point is wave 13, because this is the wave which received the topical module questions about licensing. I define interrupted individuals to be people who reported being out of the labor force for any week between waves 1 and 12 and were members of the labor force at wave 13. With this definition, I capture any interruption in labor force participation from the reference point late 2012 back to when the survey began in late 2008, a period of 4 years. For consistency, I must restrict my sample to individuals whose records exist in each of the 13 waves.

Defining interruption for CPS is a bit more complicated. The licensing question is asked in month 1 and month 5 of the 8-month CPS survey, but there is no way for me to identify interrupted workers who are interviewed in month 1, because I do not have information about their past labor force participation. When looking at interruption, then, I must restrict my sample to month 5 respondents. This cuts the sample size in half. I define interrupted workers as people who reported being out of the labor force in any of the 4 previous survey months, which took place 12 to 16 months before the month 5 survey is administered. Because of this method, I can only examine individuals for whom I can link data between months 1-4 and month 5. There is significant attrition associated with this linking, so this restriction cuts the sample size further by a third. This definition of interruption looks at a relatively short window of 4 months which is a year removed from the reference point, so it is a narrower definition than the one for the SIPP.

I compare the characteristics of interrupted workers to those of the general population in Table 6. The sample for this table is restricted to employed individuals. The group of interrupted workers is more female than the general working population, earns lower wages, is younger, and has a higher concentration of minority racial groups. The SIPP has a higher fraction of interrupted workers than the CPS – this makes sense, because the SIPP tracks labor market history three years into the past from the date of measurement, while the CPS only goes back one year. Interrupted workers are slightly less likely to hold licenses than uninterrupted ones; the two traits are negatively correlated. The table also reports statistics for mothers with children under the age of 8, who are much more likely to be interrupted than the general population.[11] I exploit this fact when I discuss identification.

I discovered an interesting discrepancy in the survey data when comparing the licensing prevalences of interrupted and uninterrupted workers within occupations. When we think of

[11] While mothers with young children are also more likely to be licensed than other workers, this difference disappears completely once I control for occupational choice.

licensing laws as part of a regulatory regime, we naturally assume that licensing regulations apply to all workers equally. All workers in a licensed occupation must hold a license, whether they are male or female, interrupted or uninterrupted. Differences in licensing prevalences between interrupted and uninterrupted workers on the national and state level are expected, but we assume that they arise from differences in occupational choice – perhaps interrupted workers select out of licensed occupations, but within occupations there should not be differences in licensing prevalence between interrupted and uninterrupted workers.

At least when it comes to survey responses, this assumption appears to be incorrect. Consider, for example, elementary and middle school teachers in the CPS, one of the largest commonly licensed occupations. I partition employed teachers into two groups: those who have experienced an interruption to their labor force participation in the last year, and those who have not. 78% of the uninterrupted group report that they hold a license, compared to only 58% of the interrupted group. Using the CPS, I calculated occupation level licensing prevalences for interrupted workers and noninterrupted workers, again restricting the sample to employed individuals. Across all occupations, the average difference between the licensing prevalences for the two groups was -.057. This implies that only about 40% of the raw difference in licensing prevalence between interrupted and uninterrupted workers is due to occupational choice, while the other 60% persists after conditioning on occupation. This surprising fact proves crucial later in the section when I define licensing prevalence as an explanatory variable.

Model 1: Exploiting Longitudinal Qualities of the SIPP

The SIPP's comparative advantage over the CPS comes from its longitudinal structure. Because I follow individuals over four years, I am able to confidently identify workers who have faced an interruption to labor force participation and compare them to their peers. However, because the SIPP only asked the question about licensing once, I do not have access to detail about licensing over time. This first model, then, examines the outcomes of interrupted individuals. Both the signal effect and the barrier effect of licensing will impact the relative outcomes of interrupted and uninterrupted workers. If the signal effect is more powerful, we would expect competent interrupted workers to do relatively better where there is greater access to licensing, while if the barrier effect dominates, we would expect them to do relatively worse.[12] Another way to think about this is that the employment and wage gaps between interrupted and noninterrupted workers may be smaller or larger in environments with greater access to licensing, depending on the direction of the effect. I use state-level variation in licensing prevalence as a measurement of "access to licensing," and examine employment and wage gaps at the state level.

It is important to remember that this statement is about relative outcomes between groups in the labor market. In the aggregate, workers who return to the labor market after an absence have worse outcomes than their peers who did not experience a labor force participation interruption. I study the impact of occupational licensing on this penalty. I must focus sharply on the differential effect of licensing on interrupted workers, because high access to licensing may also have general, non-differential labor market effects. If I only examined the outcomes of interrupted workers without comparing them to their non-interrupted peers, I would be unable to separate the specific effect of licensing on the intermittency penalty from general effects of licensing which impact all workers equally.

[12] As I mentioned before, the caveat to this is that the signal effect of licensing may have a negative effect on *low*-quality interrupted workers.

My empirical model uses uninterrupted workers as a control group to identify the differential effects of access to licensing. Using the SIPP dataset of labor force participants in 2012, I implement the following model:

$$(1): D_i = \beta_0 + \beta_1 I_i + \beta_2 L_s + \beta_3 L_s \cdot I_i + \Gamma \cdot X_i + \theta_j + \epsilon_i$$

Where D_i is the dependent variable (I primarily examine wages and employment), I_i is a dummy variable indicating whether the worker was interrupted (as defined in the data section), L_s is the licensing prevalence in the worker's state, X_i is a vector of control variables, θ_j is the fixed effect for the two digit occupation code for the occupation j of individual i, and ϵ_i is an error term. β_1 estimates the penalty of being interrupted on D_i, without taking licensing effects into account. The impact of licensing prevalence on D_i for uninterrupted workers is β_2, while the impact of licensing on D_i for interrupted workers is $\beta_2 + \beta_3$. I obtain the differential impact of licensing by subtracting the effect on uninterrupted workers from the effect on interrupted workers: $\beta_2 + \beta_3 - \beta_2 = \beta_3$. This is the coefficient of interest. If the signal effect dominates the barrier effect, then β_3 will be positive for employment and wage dependent variables.

As I discussed in the data section, there are many demographic differences between interrupted and uninterrupted workers. For example, interrupted workers are much more likely to be female. If we implemented the model without including any control variables, then β_3 would be biased insofar as licensing also has differential effects on women absent any labor force interruptions. I include control variables to absorb other differential effects of licensing which may be correlated with labor force interruptions.[13] The occupational fixed effects work similarly.

[13] Though I exploit longitudinal data to construct the indicator variable I, my model is cross-sectional, as it only examines the sample members at a single point in time. Thus, the model is unaffected by any time-trends related to labor force participation, employment, licensing, or even sample composition. Zabel (1998) shows that in the 1990

Model 2: Examining Changes in Licensing Environments with the CPS

The CPS only follows individuals for a period of 16 months, so it is not as effective as the SIPP at detecting long-term labor force participation intermittency. However, the larger sample size in the CPS allows me to obtain detailed information on licensing prevalence for state-occupation pairs, which is not possible with the SIPP. Moreover, the CPS asks licensing questions every month from 2016 to 2018, which allows me to track licensing prevalence over time. I use these advantages of the CPS to examine the impact of changes in licensing environments on aggregate employment outcomes. To the extent that differences in licensing prevalence represent changes in the availability of licenses, we would expect an increase in licensing prevalence for a state-occupation to be accompanied by a change in the intermittency penalty. The direction of this change depends on the relative importance of the signal effect and the barrier effect. Because the CPS follows individuals for only 16 months, I am unable to track people with intermittent labor force participation between 2016 and 2018. I can, however, examine aggregated information about groups of CPS respondents in this time frame. If the signal effect is stronger than the barrier effect, we would expect occupations that increased in licensing prevalence to attract more interrupted workers over time. If the barrier effect is stronger, we would expect increases in licensing prevalence to be associated with decreases in the concentration of interrupted workers.

I implement the following regression on a sample of state-occupation pairs:

SIPP, individuals who had been unemployed or out of the labor force in a given wave were more likely to be missing from the sample in following waves. The first wave of the SIPP includes records for 105,663 individuals, while the thirteenth wave includes records for only 76,988. This difference implies an attrition rate of 27%. If I began with a group of workers in Wave 1 and attempted to track them through the time, I would end up with a biased sample. However, the construction of I looks backwards from an already existing sample of Wave 13 workers. Thus, I avoid these issues of sample attrition.

$$(2): \Delta_{t,t-1}\delta_{j,s,i=1} - \Delta_{t,t-1}\delta_{j,s,i=0} = \beta_0 + \beta_1\Delta_{t,t-1}L_{j,s} + \Gamma \cdot \Delta_{t,t-1}X_{j,s} + \theta_j + \alpha_s + \epsilon_{j,s}$$

Where $\delta_{j,s,i=k}$ indicates the fractions of workers in state s and of type k who are employed in occupation j, $L_{j,s}$ indicates the fraction of workers in state s and occupation j who are licensed, $X_{j,s}$ is a vector of average demographic characteristics of the state-occupation pair, $\Delta_{t,t-1}$ indicates that I am measuring the difference between measurements in early 2018 and early 2016 for all of these variables, θ_j is an occupation fixed effect, α_s is a state fixed effect, and $\epsilon_{j,s}$ is an error term.

This model requires a bit of interpretation. $\delta_{j,s,i=k}$ is the *employment share* in job j for workers of type k in state s. For example, if 1% of all interrupted workers in Illinois are schoolteachers, then $\delta_{Schoolteacher,Illinois,i=Interrupted} = .01$. This employment share measurement captures the concentration of individuals of a certain type in a state-occupation. $\Delta_{t,t-1}\delta_{j,s,i=0}$ measures the change in the employment share over time for each state-occupation pair for group 0, while $\Delta_{t,t-1}\delta_{j,s,i=1}$does the same for group 1. The final dependent variable $\Delta_{t,t-1}\delta_{j,s,i=1} - \Delta_{t,t-1}\delta_{j,s,i=0}$ is a difference in differences: it measures the difference between the two employment trends. On the right side, the key explanatory variable is $\Delta_{t,t-1}L_{j,s}$, which is the change in *licensing prevalence* for the state-occupation pairs. The coefficient β_1 estimates the impact of changes in licensing prevalence on the dependent variable. Because the differenced dependent variable isolates the difference between two time trends, β_1 can be interpreted as the differential effect of licensing on whatever group we define to have $i = 1$, relative to the group $i = 0$. I define the group $i = 1$ to be interrupted workers, and the group $i = 0$ to be uninterrupted workers, using the definitions from the data section.

One of the downsides of this definition is that I must restrict the sample in the left-side variable to individuals whose labor market history is available in the data – that is, month 5 respondents whose prior responses are available. However, if I used only this restricted sample for the whole analysis, many state-occupation pairs would not have enough observations for me to calculate a licensing prevalence on the right-hand side with acceptable precision. Fortunately, I do not need to examine labor market history to determine an individual's licensing status. To increase the number of pairs with acceptable measurements of licensing prevalence, I use the entire sample of month 1 and month 5 individuals to calculate licensing prevalences for state-occupation pairs. I drop all state-occupation pairs which have fewer than 8 observations for either time period in this larger sample. I also weight the regression by the total number of individuals in a pair in both time periods.

I measure the change in licensing prevalence $\Delta_{t,t-1}L_{j,s}$ using only uninterrupted (type 0) workers. The best way to see that this restriction is necessary is to consider what would happen to the model if I included all workers in my calculation of licensing prevalence. As I showed in the data section, licensing status is negatively correlated with labor force interruption, even after conditioning on occupations. Suppose, then, that a state-occupation pair has a relative increase in its share of interrupted workers (that is, the dependent variable $\Delta_{\mathrm{t,t-1}}\delta_{j,s,i=1} - \Delta_{\mathrm{t,t-1}}\delta_{j,s,i=0}$ increases). Mechanically, we would expect to see a corresponding decrease in the explanatory variable $\Delta_{t,t-1}L_{j,s}$ because a higher fraction of workers in the state-occupation pair are interrupted and therefore less likely to hold licenses. This link between the dependent and independent variable obscures the causal effect of changes in licensing prevalence on the dependent variable that the model is constructed to capture. Excluding type 1 workers from the calculation of $\Delta_{t,t-1}L_{j,s}$ eliminates this concern.

In occupations that experience a change in licensing policy, one would expect to observe a large change in measured licensing prevalence from a fraction close to 0 to a fraction close to 1.[14] The majority of state-occupation pairs did not experience a policy change between 2016 and 2018. We will certainly observe small changes in licensing prevalence in these environments, but these may be attributed to noisy features of the data – people move in and out of occupations, people answer questions differently at different times, et cetera. If this description accurately characterizes the patterns in the data, then $\Delta_{t,t-1}L_{j,s}$ may have a skewed distribution, with many state-job pairs which do not see a policy change clustered around 0, and a few outliers near ± 1[15]. However, a histogram of $\Delta_{t,t-1}L_{j,s}$ over state-occupation pairs with more than seven observations (Figure 6) demonstrates that this is not the case: changes in licensing prevalence appear to be more or less normally distributed.

I do not expect small changes in licensing prevalence to have much explanatory power in my regression. Such changes are probably caused by "noise" and therefore unlikely to indicate the policy changes that my model aims to capture. Because occupations with these small changes are so numerous, I implement a second specification of the model which captures the effects of dramatic changes in licensing prevalence more concretely:

$$(3)\text{:}\ \Delta_{\mathrm{t,t-1}}\delta_{j,s,i=1} - \Delta_{\mathrm{t,t-1}}\delta_{j,s,i=0} = \beta_0 + \beta_1 P_{j,s} + \beta_2 M_{j,s} + \theta_j + \alpha_s + \epsilon_{\mathrm{j,s}}$$

This equation is identical to (2), but I have replaced the continuous explanatory variable $\Delta_{t,t-1}L_{j,s}$ with dummy variables $P_{j,s}$ and $M_{j,s}$ which partition the dataset. $P_{j,s}$ should be 1 for

[14] Grandfather clauses would make the changes look less dramatic, especially over short periods of time.
[15] We would expect to observe more outliers on the positive end for two reasons. First, it is usually easier to pass new regulations than to repeal old ones. Second, if an occupation lost its official license requirement, it is possible that workers will continue to hold their license even if it is not mandatory – the occupation may move from "licensed" to "certified."

observations which had a change in licensing prevalence $\Delta_{t,t-1}L_{j,s}$ greater than some threshold k, while $M_{j,s} = 1$ when $L_{j,s} < -k$. The dummy variables allow me to take the mean of $\Delta_{\mathrm{t,t-1}}\delta_{j,s,i=1} - \Delta_{\mathrm{t,t-1}}\delta_{j,s,i=0}$ across three groups while still correcting for state and occupational fixed effects. The base group is composed of state-occupation pairs which did not see a large change in licensing prevalence. β_1 estimates the difference in means of the dependent variable between the group of jobs which had a large increase in licensing prevalence and the base group. Similarly, β_2 estimates the difference in means of the dependent variable between jobs which had a large decrease in the licensing prevalence and the base group. The goal of this method is to distinguish environments which actually experienced a policy change from the many unchanged environments whose licensing prevalences fluctuate slightly, but without meaning. If the signal effect dominates the barrier effect, I expect β_1 to be positive and β_2 to be negative, and vice versa if the barrier effect dominates.

Problems with Identification and Alternative Specifications

My models are concerned with two key phenomena: interruption and licensing. Unfortunately, both of these are endogenous factors. I do not have a random shock which distributes workers into interrupted and uninterrupted groups, nor do I randomly assign some environments to be licensed and others to be unlicensed. Rather, both features are correlated with other factors in the regression, which may generate biased regression coefficients. I argue in this subsection that the political economy of licensing eliminates many concerns related to the endogeneity of licensing, and I describe two alternative specifications of my SIPP model which mitigate the endogeneity of interruption.

State licensing prevalence is certainly more exogenous than the licensing status of individuals, which is deeply intertwined with characteristics like work ethic, intelligence, and personality. As a baseline test for my SIPP model, I implement a regression which uses individual licensing status as the explanatory variable. In other words, I replace L_s (the percent of labor force participants who are licensed in a person's state) with L_i (a binary indicator of licensing status) in equation (1). Keeping with the literature, I expect to find a licensing premium when I implement this regression – that is, I expect that β_2 will be positive when I use wage dependent variables. However, it is impossible to tell if this result is truly an effect of licensing, or if it arises from unobserved differences between licensed people and unlicensed people. Any examination of licensing must deal with these unobserved differences, but another issue arises when I specifically consider the interaction of licensing and interruption. As I have defined it, interruption is backwards-looking (it depends on responses to past surveys), while licensing is measured at the time of sampling. Because of this discrepancy, it is possible that those interrupted people who obtained licenses were simply those who had already seen the most success in the work force. However, to the model it will seem that licensing caused the positive outcomes, rather than the other way around, so my differential effects of licensing on interrupted people would be biased upwards. I eliminate these confounding individual-level relationships from my analysis by examining people's access to licensing as opposed to their actual licensing statuses.

Licensing requirements typically arise as the result of votes, either in legislative bodies or in professional organizations. The members of these bodies vote for or against licensing depending on their personal motivations and the interests of their constituents. State licensing prevalence, then, is not strictly speaking exogenous. Rather, it may be a function of state characteristics. However, there are reasons to believe that licensing policy is "random enough" to create a natural experiment. A cursory examination of Figure 1 establishes at least that geography does not determine state licensing prevalence. Moreover, Table 4 suggests that partisan lean is not an indicator: both the top 5 and bottom 5 are dominated by "red" states. In section II, I discussed theories of licensing's adoption. If the "interest group" theory of licensing holds, then we can think of licensing prevalences as the result of a process that is at most tangentially related to broader state labor market conditions.

People are able to move freely between licensed states and unlicensed states, which poses another problem for my models. Even if people do not influence over their local government's licensing policies, they can simply move to another location where the levels of licensing are more in line with their preferences. If enough people behaved this way and state licensing policies remained relatively constant, then over time, the population would sort itself to some extent by its preferences about licensing policy. If people with preferences for high licensing prevalence are different from people with low licensing prevalence in other important ways, this phenomenon would introduce bias to the model.

If licensing is in fact endogenous, the two models are affected in different ways. To illustrate this, suppose that high levels of licensing are correlated with the existence of state programs which assist those who are re-entering the labor force. Then any effects of these programs will be most powerful in states with high licensing prevalences. These effects will appear to the first model as if they were effects of occupational licensing. If these programs do have a positive effect on the employment of interrupted workers, then the estimate β_3 of the differential impact of licensing on interrupted workers will be biased upwards. Contrastingly, if we assume that these programs do not change over time, then the second model captures such phenomena in the state fixed effect terms α_s. The situation which would cause bias for the second model would be if these programs were often adopted alongside changes in licensing policy. For a factor to cause bias in the second model's estimate, it must vary over state-occupation pairs *and* over time. In other words, factors which are correlated with cross-sectional licensing measurements no longer concern us, because only *other changes* which are correlated with the *changes* in licensing environments may cause bias. In this sense, the CPS model is superior to the SIPP model. Studying changes in licensing prevalence is preferable to studying cross-sectional differences because it allows us to be more certain that licensing itself is the driving factor behind any effects we observe. Moreover, because the CPS model studies changes over a short period (three years), it avoids the issue of people sorting themselves by preferences for licensing between states. This sorting process, if it occurs, would likely take place over a long period of time.

Like licensing, interruption as I have defined it cannot be thought of as a random shock. On a basic level, the labor force participation decision and employment/wage outcomes are both functions of choices that individuals make.[16] This introduces a confounding factor into my analysis. It is possible that effects on employment/wage which seem to be caused by interruption may in fact be caused by correlations with these other choices. In both models, the control vector X absorbs some potential bias. But issues related to omitted variables bias persist, since I cannot observe every possible confounding factor.

It is worth looking for an alternative specification which avoids these issues completely. I exploit the fact that intermittent labor force participation varies across demographic groups.

[16] Thinking of choices such as: whether to pursue opportunities to gain human capital, where to live, how much to prioritize career vs. other facets of life, etc.

Mothers with young children are disproportionately likely to experience a labor force interruption. In the SIPP, 36% of female labor force participants with children under the age of 8 have experienced an interruption within the last 4 years, compared to 20% of all labor force participants (Table 6). Similarly, interruption becomes more common as people approach retirement age.[17] The interruption rate of labor force participants increases from 18% for individuals aged 60-64, to 22% for individuals aged 65-69, to 29% for individuals aged 70-75.

I implement two alternative specifications. For the first, I use mothers with young children as a "likely licensed" group. In equation (1), this means that mothers with children under the age of 8 are indicated by the variable I, and in equation (2), the group $i = 1$ corresponds to mothers with children under the age of 8. To examine retiring individuals, I limit the sample to individuals aged 60 to 75, partition the sample into younger, medium, and older individuals, and let the effect of licensing prevalence vary across these groups. Admittedly, neither of these solutions completely eliminates concerns of endogeneity. Having children is not a random shock, and 60-year-olds are different from 75-year-olds in many ways other than their proximity to retirement age. However, both these analyses add a degree of separation between the labor force participation decision and the explanatory variable. Additionally, for the SIPP model, I am able to implement regressions with labor force participation as the dependent variable, which was previously impossible because of the mechanical link between labor force participation and interruption under my definition.

V. Results

Model 1 – SIPP

[17] It may be that individuals frequently get "bored" with retirement and decide to return to the work force.

Figure 7 illustrates the raw correlation between the dependent variable and independent variable of interest in equation (1). I plot state licensing prevalences against interrupted-noninterrupted employment gaps.[18] The correlation coefficient between the two variables is $\rho = -.245$, and a negative slope is evident in the best fit line. Figure 8 tells a similar story. I group individuals by the quintile of their state licensing prevalence and graph employment for interrupted and uninterrupted workers. We see that the lowest two quintiles of individuals by licensing prevalence have the smallest employment gaps between interrupted and uninterrupted workers. Figure 9 is similar to Figure 7, but now I illustrate the relationship between state licensing prevalence and state labor force participation. The trend line is quite dramatic: $\rho = .462$. This relationship is similar to the one depicted in Blair and Chung (2018b).

I report the results from my baseline test which measures licensing at the individual level in Table 7. This specification is equation (1) where L_s the state licensing prevalence is replaced with the licensing status L_i. In this individual-level specification, the coefficient on licensing is interpreted as the effect of holding a license, and the coefficient on the interaction term shows how this effect changes for interrupted individuals – recall that in equation (1), the impact of licensing on uninterrupted people is β_2 and the impact on interrupted people is $\beta_2 + \beta_3$. Columns (1) and (2) use an employment dummy variable for the dependent variable D_i, while (3), (4), and (5) use hourly wage. Column (5) restricts the sample to employed individuals. Therefore, the results in column (5) correspond to effects on wages for people who are actually receiving wages. Recall that the sample here is restricted to labor force participants, so the employment dependent variable can also be interpreted as (1 – the unemployment rate). Columns

[18] This state-level analysis only includes the 33 states with more than 200 respondents, because including states with few respondents creates outliers which skew the trend in a way that the individual level model does not.

(2), (4), and (5) include occupational fixed effects θ_j, which categorize occupations based on the first two digits of the 4 digit occupational codes.[19] We interpret results in these columns as effects on workers who hold the same occupation.

The results in Table 7 are as expected. In columns (1) and (2) we see that interruption is associated with a large and statistically significant penalty to the probability of being employed. Being licensed improves employment outcomes, with a larger effect for interrupted workers than uninterrupted workers. In columns (3), (4), and (5), we observe a licensing premium of just over $2.00 per hour and a large wage penalty for interruption. In column (3), there is evidence of a differential effect of licensing on wages for unemployed workers, $p<.05$. However, this result loses its significance after controlling for occupational choice and eliminating unemployed workers from the sample in column (5). All of this is consistent with the literature which establishes the intermittent labor force participation penalty and the licensing premium, and the results suggest that licenses may be more valuable for interrupted workers than for uninterrupted ones. As I discussed in the identification subsection, though, these results are subject to bias because the individual's decision to obtain a license is highly endogenous.

The results of the regression from equation (1) which use state licensing prevalence as the explanatory variable are in Table 8. The difference between the interpretation of these results and those in Table 7 is the difference between the "impact of holding a license" and "the impact of living in an environment where lots of licenses are available." Columns (1) through (5) are defined just as they were in Table 7. As before, the coefficients on the interrupted variable are negative in all columns. Column 2 tells us that interrupted labor force participants are 6.4

[19] Modifying this definition to three-digit occupational codes does not substantially alter the results.

percentage points less likely to be employed than their uninterrupted peers. In column 5, we see that even when interrupted labor force participants do find work, they receive lower wages than their uninterrupted peers. However, we see a difference from Table 7 when we look at the effects of licensing. I do not find significant effects of state licensing prevalence on the labor force as a whole for either employment or wages. The licensing employment advantage and the licensing premium both disappear when we measure licensing on the state level rather than the individual level. Th reason may be that the bias which results from unobserved differences between licensed individuals and unlicensed individuals was eliminated when we measured at the state level, but it is also possible that my state-level measurements were simply too broad to capture the effects of licensing.

The coefficient on the interaction term measures the differential effect of licensing on interrupted workers. To interpret the magnitudes of the coefficients, we must understand the distribution of licensing prevalences across states. The 90th percentile of licensing prevalence by state is .298, while the 10th percentile is .175. This means that the difference between a state with a lot of licensing and a state with very little licensing may be around .123. Using the coefficient of -.251 on the interaction term from column (2), this difference corresponds to a change in the employment gap between interrupted and uninterrupted labor force participants in the most and least licensed states of (-.251) * (.123) = -.031. Because the negative effect is applied to interrupted workers, this result implies that high licensing prevalences *increase* the penalty for intermittency. This result is economically significant. The national employment gap between uninterrupted and interrupted workers is equal to .156, so the magnitude of this difference between states is substantial – about 20% of the national average. This same exercise can help demonstrate the disappearance of the licensing premium: the coefficient -1.385 in column (5) corresponds to a licensing premium of (-1.385) * (.123) = -.17 per hour as we move from a state with low licensing to high licensing. This small estimate, coupled with high standard errors, means that I cannot reject the hypothesis that state licensing prevalence has no effect on wages.

Table 9 provides the results from this model on the same SIPP dataset using an alternative specification in which mothers with young children are the group indicator. I restrict the sample to only include women aged 26 to 45, so that the comparison group is other women who do not have children rather than the general population. Columns (1) through (5) are set up just as they are in Tables 7 and 8. In these columns, we do not see significant results on the mother coefficient. This is perhaps less surprising when we realize that the sample is restricted to mothers who are in the work force. In columns (1) and (2), I do not find significant evidence of an effect of state-level licensing prevalence on employment for labor force participants. The coefficients on the licensing prevalence variable in regressions with the wage dependent variable are positive and large. The estimate of 8.845 from column (5) corresponds to a difference in wages between workers in the most licensed states and workers in the least licensed states of (8.845) * (.123) = 1.08 per hour. However, this licensing premium is not statistically significant (p=.288). There are no significant effects on the interaction term for the wage or employment dependent variables.

Because I use mothers as the group indicator, I can run a regression using a labor force participation dummy as a dependent variable on a sample of individuals which is not restricted to members of the labor force, which is reported in column (6).[20] In this column, the mother coefficient is large and negative, albeit with only marginal significance. The effect of licensing

[20] It would be misleading to use occupational fixed effects with the labor force coefficient, because individuals who are not members of the labor force do not report an occupation (or rather, they would all be coded as having "missing occupation"). Thus, I do not include a column with a labor force dependent and occupational fixed effects

on labor force participation is positive, but without statistical significance. I estimate that the difference between labor force participation rates for this sample in highly licensed states and states with low licensing levels is (.130) * (.123) = .016. The descriptive relationship from Figure 9 is apparently much weaker when I restrict the sample to women of childbearing age and control for other factors.[21] The coefficient on the interaction term in column (6) is essentially zero, which means that the effect of licensing on labor force participation is not different for mothers with young children.

Table 10 reports the results of my analysis on older individuals in the SIPP dataset. The dependent variables in columns (1) through (6) are identical to those in Table 8. The sample includes only workers between the ages of 60 and 75 (inclusive). I separate these workers into three groups by age: The “old” workers are between 70 and 75, the “medium” workers are between 65 and 69, and the “young” workers are between 60 and 65. As I mentioned in section IV, interruption rates conditional on labor force participation are increasing for this sample. In the table, the young workers are the base group.

In columns (1) to (5), we do not see many significant effects. Once again, when we note that the employment rate for 70- to 75-year-olds is not significantly different from the rate for 60- to 65-year-olds, we must remember that these analyses only include labor force participants: if a worker over the age of 70 is having difficulty finding a job, she may simply decide to leave the labor force. This is particularly problematic for regressions with a wage dependent variable,

[21] If I run this regression without restricting to women aged 26 to 45 (i.e. measure the effect of licensing on labor force participation for all individuals), the coefficient on licensing prevalence is .231 (p=.138). This is larger and more significant, and it suggests that the trend in figure 7 is driven mostly by men.

as the wage is strongly linked to the retirement decision. A high wage means that retirement carries a high opportunity cost.

Examining column (6), we see that labor force participation decreases as the groups get older. The coefficient on licensing prevalence is the effect of licensing prevalence on labor force participation for the base group (60- to 64-year-olds). For this group, there is a strong relationship between licensing prevalence and labor force participation: (.412) * (.123) = .051, which is close to what we see in Figure 9. However, there is some evidence that the positive relationship decays with age. The coefficients on both interaction terms are negative, and the one for 70- to 75-year-olds is larger than the one for 65- to 69-year-olds. For 65- to 69-year-olds, more than half of the effect disappears, and for 70- to 75-year-olds, the positive relationship is almost completely wiped out. This squares well with the licensing-as-barrier framework. If licensing raises barriers to reentering a job, then individuals who are close to the retirement age may remain in the labor force longer in highly licensed states, because they know that after they leave the labor force, it will be difficult for them to return. Figure 10 depicts this phenomenon graphically. I segment the sample into thirds by state licensing prevalence, and plot labor force participation against age. The high licensing group oscillates between having clearly the highest participation and being tied with the others until workers hit their late 60s, after which the relationship becomes less clear.

Model 2 - CPS

Table 11 displays the results of my analysis of CPS data. Columns (1) and (2) use the continuous independent variable as in equation (2), while columns (3), (4), and (5) use discrete indicators as in equation (3). I vary the cutoff k from the definition of $P_{j,s}$ and $M_{j,s}$ between these columns. Columns (2) through (5) control for trends over time in race, gender, and educational attainment. The top panel is the model as I introduced it, where the dependent variable is the difference in employment share time trends between interrupted and uninterrupted workers. The bottom panel uses the alternative specification where the group $i = 1$ corresponds to mothers with children under the age of 8 – that is, the dependent variable is the difference in employment share time trends between these mothers and other those of other workers, as I discussed in the identification section. As before, licensing prevalence is given as a decimal between 1 and 0, so a 1 percentage point change in the continuous licensing prevalence variable corresponds to a change in the dependent variable by the coefficient divided by 100.

This table has no significant results – there is only one coefficient with a p-value of less than .4. The coefficients on the continuous dependent variable have the expected sign for the lower panel, but they move in the opposite direction in the upper panel. In both panels, the “gain license” variable has the same sign as the “lose license” dependent variable, which is also unexpected. The magnitudes of all coefficients are very small. Taken as a whole, these results do not provide any evidence of differential effects of licensing on employment.

VI. Discussion

This paper examines the impact of occupational licensing on individuals who have experienced interruptions to their labor force participation. The signal effects of licensing may improve outcomes for interrupted workers relative to their peers, but licensing may also impose barriers to entry whose costs fall disproportionately on these workers. I construct empirical models to isolate this differential effect. To avoid the issues arising from the endogeneity of the individual licensing decision, I use state and state-occupation level licensing prevalence as an explanatory factor rather than individual licensing status. Another contribution of this paper is to study similarities and differences between modern government surveys which capture data about licensing and earlier data sources.

I find some evidence that licensing depresses outcomes for people with intermittent labor force participation. The employment gap between interrupted and uninterrupted people is largest in highly licensed states. This finding suggests that licensing creates barriers to entry in the labor market, as interrupted workers may be particularly susceptible to these barriers. A strong barrier effect could also account for the positive relationship between state-level licensing prevalence and state labor force participation – if labor force participants understand that re-entry is costly, they will be more likely to remain in the labor force than to exit it if they face a marginal decision. Another observation which supports this idea is that retirement appears to take place later in states with high licensing prevalences. A deeper investigation of the behavior of retiring individuals in licensed occupations could provide interesting insights about the barriers to entry imposed by licensing.

I confirm the literature by demonstrating the existence of benefits for individual holders of licenses. However, I do not find evidence of these same benefits when I measure licensing on the state level rather than the individual level. This may indicate that my individual level measurements are biased by unobserved characteristics of people who hold licenses. The state-level variation in licensing prevalence is also small, and there are other important differences between states which may dominate any effect of licensing.

The comparison between data from different surveys raises serious concerns about the consistency of measurements of licensing prevalence. While measurements of licensing prevalence from the SIPP and the CPS are weakly correlated with each other, I still find important differences between the two surveys. I eliminate bias from proxy respondents as a possible cause of these differences, but I cannot identify the true cause. The two government surveys ask similar questions, but these vary significantly from the questions asked in earlier phone surveys. I show that the state licensing prevalences calculated from the earlier Westat phone survey have large differences from those calculated in the government surveys. More work is needed to understand the differences between these surveys. Researchers who study licensing must keep in mind that the wording of survey questions and survey design may have large effects on measurements of licensing.

The results from my CPS model were disappointing, as I essentially failed to find any significant effects. It is likely that this failure stems from an inability to accurately identify changes in licensing policy in the data. At the time of writing, the CPS had only been asking questions about licensing for 3 years. Because of grandfather clauses, it may take a long time for changes in policy to manifest themselves as changes in licensing prevalence. Even if I could identify these changes well, 3 years is a short time, and it is likely that there were not many real policy changes between 2016 and 2018. As the CPS continues to ask its licensing questions, these concerns will become less relevant. It will also become possible to pool respondents over longer periods of time. Once the CPS has been asking questions for 10 or 15 years, my method for identifying policy changes may become much more accurate, which may warrant a return to the model.

References

Akerlof, George. 1970. “The Market for Lemons: Qualitative Uncertainty and the Market Mechanism.” *Quarterly Journal of Economics*. 84: 488–500.

Baum, Charles, 2002. "The Effect of Work Interruptions on Women's Wages." *Labour,* 16: 1-36.

Berk, John and Jules H. van Binsbergen, 2017, ‘Regulation of charlatans in high-skill professions.’ NBER Working Papers, no. 23696.

Blair, Peter, and Bobby W. Chung, 2018a. “Job Market Signaling Through Occupational Licensing.” NBER Working Papers, no. 24791.

Blair, Peter, and Bobby W. Chung, 2018b. “How Much of Barrier to Entry is Occupational Licensing?” NBER Working Papers, no. 25262.

Dorsey, Stuart, 1980. “The Occupational Licensing Queue.” *Journal of Human Resources*, 15(3): 424–34.

Farber, Henry S., Chris M. Herbst, Dan Silverman, and Till von Watcher, 2018. “Whom Do Employers Want? The Role of Recent Unemployment Status and Age.” *Journal of Labor Economics*, 37(1): (forthcoming)

Friedman, Milton, 1962. *Capitalism and Freedom.* Chicago: University of Chicago Press.

Gittleman, Maury, Mark Klee, and Morris Kleiner, 2018. “Analyzing the labor market outcomes of occupational licensing.” *Industrial Relations*: A Journal of Economy and Society, 57 (1): 57–100.

Graddy, Elizabeth, 1991. “Toward a General Theory of Occupational Regulation.” *Social Science Quarterly*, 72:(4): pp. 676–95.

Hotchkiss, Julie and M. Melinda Pitts, 2005. "Female Labor Force Intermittency and Current Earnings: A Switching Regression Model with Unknown Sample Selection." *Applied Economics*, 37: 545-60

Hotchkiss, Julie and M. Melinda Pitts, 2007. "The Role of Labor Market Intermittency in Explaining Gender Wage Differentials." *American Economic Review*, 97(2): 417-21

Institute for Justice, 2012 “License to Work: A National Study of the Burdens of Occupational Licensing”

Johnson, Janna and Morris Kleiner 2017. “Is Occupational Licensing a Barrier to Interstate Migration?” NBER Working Papers, no. 24107.

Kleiner, Morris, 2000. “Occupational Licensing.” *The Journal of Economic Perspectives*, 14(4): 189-202

Kleiner, Morris and Alan B. Krueger, 2010. "The Prevalence and Effects of Occupational Licensing." *British Journal of Industrial Relations*, 48(4): 676-687

Kleiner, Morris and Alan B. Krueger, 2013. “Analyzing the Extent and Influence of Occupational Licensing on the Labor Market.” *Journal of Labor Economics*, 31(2): S173-S202.

Kleiner, Morris and Evgeny Vorotnikov, 2017. “Analyzing Occupational Licensure Across the States.” *Journal of Regulatory Economics*, 52(2): 132-158

Kleiner, Morris, Robert Gay and K. Greene, 1982. “Barriers to Labor Migration: The Case of Occupational Licensing.” *Industrial Relations*, 21(3): pp. 383–91

Law, Marc T. and Mindy S. Marks, 2009. “Effects of Occupational Licensing Laws on Minorities: Evidence from the Progressive Era.” *The Journal of Law and Economics* 52(2): 351-366

Learning Policy Institute, 2016. “Solving the Teacher Shortage How to Attract and Retain Excellent Educators”

Maurizi, Alex, 1974. “Occupational Licensing and the Public Interest.” *Journal of Political Economy*. 82, Part 1, pp. 399–413

Peterson, Brenton, Sonal Pandya and David Leblang, 2014. “Doctors with borders: occupational licensing as an implicit barrier to high skill migration.” *Public Choice* 160(1/2): 45-63

Redbird, Beth, 2017. “The New Closed Shop? The Economic and Structural Effects of Occupational Licensure.” *American Sociological Review* 82(3): 600-624

Shore, Ted, and Armen Tashchian, 2013. “Perceptions Of Unemployed Workers: Unemployment Duration, Volunteerism, And Age.” *The Journal of Applied Business Research*, 29(4): 983-990

Smith, Adam, 1976 [1776]. *An Inquiry into the Nature and Causes of the Wealth of Nations* (edited by Edwin Canaan). Chicago, IL: University of Chicago Press.

Spence, Michael, 1973. “Job Market Signaling.” *The Quarterly Journal of Economics*, 87(3): 355-374

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

Stratton, Leslie, 1995. "The Effect of Interruptions in Work Experience Have on Wages." *Southern Economic Journal*: 955-70.

Wozniak, Abigail, 2015. “Discrimination and the Effects of Drug Testing on Black Employment.” *Review of Economics and Statistics*, 97(3): 548-566

Table 1: Mean Characteristics of Proxy Responses and Self Responses - SIPP

Variable	Self Responses	Proxy Responses	T-test
Male	0.457	0.597	-18.822
Education			
Less than High School	0.062	0.103	-10.180
High School	0.195	0.245	-8.036
Some College	0.138	0.136	0.536
College	0.604	0.517	11.811
Age			
Age	45.558	44.581	5.874
30 or under	0.112	0.144	-6.409
31 to 54	0.632	0.622	1.389
55 or over	0.256	0.234	3.345
Race			
White	0.817	0.833	-2.694
Black	0.111	0.087	5.267
Asian	0.040	0.051	-3.333
Hispanic	0.100	0.144	-9.081
Hours Worked per Week	32.833	33.377	-2.030
Hourly Wage	20.468	19.617	1.936
Union Member	0.116	0.115	0.194
Licensed	0.240	0.207	5.229
Interrupted	0.215	0.236	-3.293
Employed	0.926	0.921	1.143
N	11527	7139	

Notes: This table includes all labor force participants aged 24 to 64 in 2012 who responded to the survey for each of the first 13 waves from late 2008 to 2012. Characteristics refer to the individual's status at wave 13. Hours worked per week and hourly wage are calculated for the main job only. Interruption is defined in section IV.

Table 2: Mean Characteristics of Proxy Responses and Self Responses - CPS

Variable	Self Responses	Proxy Responses	T-test
Male	0.503	0.541	-11.286
Education			
Less than High School	0.058	0.088	-16.428
High School	0.240	0.293	-17.433
Some College	0.302	0.288	4.368
College	0.400	0.331	20.753
Age			
Age	42.804	39.616	36.531
30 or under	0.194	0.307	-38.224
31 to 54	0.593	0.519	21.768
55 or over	0.213	0.174	14.245
Race			
White	0.807	0.813	-2.313
Black	0.111	0.091	9.855
Asian	0.051	0.060	-6.139
Hispanic	0.123	0.156	-13.695
Hours Worked per Week	38.929	37.987	10.239
Licensed	0.227	0.193	12.069
Interrupted	0.069	0.103	-17.644
Employed	0.958	0.956	1.444
N	42010	41856	

Notes: This table includes all labor force participants who responded to months 1 through 5 of the CPS and whose month 5 survey occurred in January to June of 2016 or January to June of 2018. All characteristics refer to an individual's status in month 5. Hours worked per week are calculated for the main job. Interruption is defined in section IV.

Table 3: Characteristics of Licensed Workers

	SIPP Analysis Sample		CPS Analysis Sample	
	All Workers	Licensed	All Workers	Licensed
Male	0.528	0.484	0.521	0.460
Education				
Less than High School	0.0749	0.0223	0.0701	0.0216
High School	0.207	0.121	0.261	0.145
Some College	0.141	0.0958	0.295	0.277
College	0.577	0.761	0.373	0.557
Age				
Age	44.50	45.04	41.41	43.49
30 or Under	0.137	0.115	0.243	0.163
31 to 54	0.639	0.647	0.561	0.622
55 or Over	0.224	0.237	0.196	0.215
Race				
White	0.820	0.831	0.814	0.845
Black	0.102	0.0872	0.0979	0.0810
Asian	0.0456	0.0482	0.0558	0.0491
Hispanic	0.145	0.0837	0.138	0.0816
Hours Worked per Week	34.69	34.33	40.18	42.22
Hourly Wage	22.53	26.05		
Union Member	0.126	0.201		
Mother of Young Child	0.139	0.172	0.146	0.189
Licensed	0.233	1	0.216	1
N	17244	4062	80282	17326

Notes: The SIPP analysis sample includes all employed individuals aged 24 to 64 in 2012 who responded to the survey for each of the first 13 waves from late 2008 to 2012. Characteristics refer to the individual's status at Wave 13. The CPS analysis sample includes all employed individuals who responded to months 1 through 5 of the CPS and whose month 5 survey occurred in January to June of 2016 or January to June of 2018. Characteristics refer to the individual's status at month 5. In the SIPP, hours per week and hourly wage are reported only for the worker's main job. Mothers of young children are mothers who have children younger than 8 years of age.

Table 4: Top and Bottom 5 States by Licensing Prevalence

SIPP			
Top 5		Bottom 5	
State	Percent Licensed	State	Percent Licensed
Alaska	46.9%	Delaware	16.5%
Idaho	38.5%	Mississippi	16.1%
Wyoming	35.2%	Nevada	15.0%
Nebraska	30.5%	Rhode Island	14.4%
Montana	29.9%	New Mexico	11.3%
CPS			
Top 5		Bottom 5	
State	Percent Licensed	State	Percent Licensed
Maine	26.8%	Hawaii	17.7%
Montana	25.4%	Nevada	17.6%
Rhode Island	25.0%	Georgia	17.5%
Wyoming	24.6%	District of Columbia	17.2%
Alaska	23.7%	California	16.6%
Westat			
Top 5		Bottom 5	
State	Percent Licensed	State	Percent Licensed
Iowa	33.2%	Kansas	14.9%
Nevada	30.7%	Indiana	14.9%
Washington	30.5%	New Hampshire	14.7%
Florida	28.7%	Rhode Island	14.5%
Kentucky	27.8%	South Carolina	12.4%

Notes: In the SIPP, the sample is labor force participants aged 24 to 64 and who have data from SIPP waves 1-13. In the CPS, the sample is month 1 and month 5 labor force participants aged 18 to 64 from January 2016 to November 2018. The Westat results, published in Kleiner and Vorotnikov (2017), are from a phone survey of labor force participants aged 18 and older.

Table 5: National Statistics for Commonly Licensed Occupations

	SIPP		CPS Early 16		CPS Early 18	
Occupation Name	Fraction Licensed	Count	Fraction Licensed	Count	Fraction Licensed	Count
Elementary and Middle School Teachers	0.704	517	0.757	1912	0.726	1880
Registered Nurses	0.763	448	0.864	1920	0.839	1884
Driver/sales workers and truck drivers	0.351	387	0.346	2033	0.287	1919
Nursing, psychiatric, and home health aides	0.487	304	0.520	1233	0.490	1057
Secondary school teachers	0.751	205	0.771	686	0.762	584
Lawyers	0.685	111	0.867	668	0.818	685
Physicians and Surgeons	0.735	83	0.876	557	0.868	530
Hairdressers, Hair Stylists, and Cosmetologists	0.774	53	0.774	433	0.718	422

Notes: The sample for the SIPP is all labor force participants in the appropriate occupation who also have data available for the first 13 waves of the SIPP. The sample for the CPS is all labor force participants in the appropriate occupation who are in either month 1 or month 5 of the survey. A crosswalk was used to match SIPP occupational codes to CPS occupational codes.

Table 6: Characteristics of Interrupted Workers

	SIPP Analysis Sample			CPS Analysis Sample		
	All Workers	Interrupted	Mother of Young Child	All Workers	Interrupted	Mother of Young Child
Male	0.528	0.439	0	0.521	0.411	0
Education						
Less than High School	0.0749	0.100	0.0678	0.0701	0.107	0.0574
High School	0.207	0.233	0.146	0.261	0.279	0.204
Some College	0.141	0.134	0.153	0.295	0.344	0.304
College	0.577	0.533	0.632	0.373	0.270	0.434
Age						
Age	44.50	40.36	34.83	41.41	37.32	34.41
30 or Under	0.137	0.282	0.269	0.243	0.399	0.300
31 to 54	0.639	0.552	0.731	0.561	0.429	0.688
55 or Over	0.224	0.166	0.000677	0.196	0.172	0.0120
Race						
White	0.820	0.777	0.795	0.814	0.774	0.780
Black	0.102	0.128	0.122	0.0979	0.125	0.118
Asian	0.0456	0.0572	0.0520	0.0558	0.0615	0.0611
Hispanic	0.145	0.187	0.172	0.138	0.185	0.159
Hours Worked per Week	34.69	31.24	33.16	40.18	34.11	37.03
Hourly Wage	22.53	16.36	19.72			
Union Member	0.126	0.0689	0.113			
Mother of Young Child	0.139	0.225	1	0.146	0.210	1
Licensed	0.233	0.200	0.293	0.216	0.135	0.282
Interrupted	0.200	1	0.356	0.0775	1	0.119
N	17244	3340	1487	80282	6223	8115

Notes: The SIPP analysis sample includes all employed individuals aged 24 to 64 in 2012 who responded to the survey for each of the first 13 waves from late 2008 to 2012. Characteristics refer to the individual's status at wave 13. The CPS analysis sample includes all employed individuals who responded to months 1 through 5 of the CPS and whose month 5 survey occurred in January to June of 2016 or January to June of 2018. Characteristics refer to the individual's status at month 5. In the SIPP, hours per week and hourly wage are reported only for the worker's main job. Mothers of young children are mothers who have children younger than 8 years of age.

Table 7: Regression Results - Individual Specification

Dependent Variable:		Employment (Mean = .924)		Wage (Mean = 20.14)		
		(1)	(2)	(3)	(4)	(5)
		No Fixed Effects	Occupational Fixed Effects	No Fixed Effects	Occupational Fixed Effects	Occupational Fixed Effects, Employed Individuals Only
Interrupted	Beta	-0.163	-0.129	-8.343	-4.603	-4.488
(Mean = .222)	Standard Error	(0.013)	(0.009)	(0.507)	(0.357)	(0.432)
	P-value	0.000	0.000	0.000	0.000	0.000
Licensed	Beta	0.013	0.028	2.155	2.295	2.018
(Mean = .227)	Standard Error	(0.003)	(0.004)	(0.598)	(0.508)	(0.526)
	P-value	0.000	0.000	0.001	0.000	0.000
(Licensed) x						
(Interrupted)	Beta	0.079	0.053	2.177	-0.100	0.865
	Standard Error	(0.015)	(0.012)	(1.010)	(0.856)	(0.969)
	P-value	0.000	0.000	0.036	0.907	0.376
N		18666	18666	15177	15177	13755
R^2		0.071	0.182	0.085	0.177	0.151

Notes: This table reports results from the model defined by equation (1), with "interrupted" as the group indicator. The sample includes all labor force participants in Wave 13 of the SIPP who are aged 26-64 have valid data for the first 12 waves. The employment dependent variable is a binary employment indicator, and the wage dependent variable is the individual's hourly wage at his or her main job (coded to 0 for unemployed workers). The regressions all include control variables for gender, race, educational status, and age. Columns (2), (3), and (5) include fixed effects for 2-digit SOC classifications. Column (5) restricts the sample to include only workers who are earning nonzero wages. Standard errors are clustered at the state level. The discrepancy between the mean of the individual licensing variable and the licensing prevalence variable is due to rounding.

Table 8: Regression Results - SIPP Model (Interrupted Indicator)

Dependent Variable:		Employment (Mean = .924)		Wage (Mean = 20.14)		
		(1)	(2)	(3)	(4)	(5)
		No Fixed Effects	Occupational Fixed Effects	No Fixed Effects	Occupational Fixed Effects	Occupational Fixed Effects, Employed Individuals Only
Interrupted	Beta	-0.088	-0.064	-6.395	-4.622	-6.390
(Mean = .222)	Standard Error	(0.049)	(0.042)	(3.259)	(2.885)	(3.209)
	P-value	0.078	0.136	0.055	0.115	0.052
Licensing Prevalence	Beta	0.023	0.034	-0.176	0.562	-1.385
(Mean = .224)	Standard Error	(0.051)	(0.055)	(12.789)	(11.393)	(12.300)
	P-value	0.659	0.545	0.989	0.961	0.911
(Licensing Prevalence) x						
(Interrupted)	Beta	-0.270	-0.251	-7.193	-0.281	9.077
	Standard Error	(0.218)	(0.192)	(14.354)	(12.534)	(13.888)
	P-value	0.221	0.197	0.618	0.982	0.516
N		18666	18666	15177	15177	13755
R^2		0.067	0.178	0.083	0.176	0.150

Notes: This table reports results from the model defined by equation (1), with "interrupted" as the group indicator. The sample includes all labor force participants in Wave 13 of the SIPP who are aged 26-64 have valid data for the first 12 waves. The employment dependent variable is a binary employment indicator, and the wage dependent variable is the individual's hourly wage at his or her main job (coded to 0 for unemployed workers). The regressions all include control variables for gender, race, educational status, and age. Columns (2), (3), and (5) include fixed effects for 2-digit SOC classifications. Column (5) restricts the sample to include only workers who are earning nonzero wages. Standard errors are clustered at the state level.

Table 9: Regression Results - SIPP Model (Mothers With Young Children vs. Other Women)

Dependent Variable:		Employment (Mean = .906)		Wage (Mean = 16.36)			Labor Force Participation (Mean = .766)
		(1)	(2)	(3)	(4)	(5)	(6)
		No Fixed Effects	Occupational Fixed Effects	No Fixed Effects	Occupational Fixed Effects	Occupational Fixed Effects, Employed Individuals Only	No Fixed Effects
Mother with Young Child	Beta	0.018	-0.016	2.221	2.978	2.597	-0.120
(Mean = .366)	Standard Error	(0.061)	(0.060)	(2.842)	(2.402)	(2.733)	(0.078)
	P-value	0.772	0.787	0.438	0.221	0.347	0.131
Licensing Prevalence	Beta	0.091	-0.010	6.473	8.646	8.845	0.130
(Mean = .225)	Standard Error	(0.140)	(0.125)	(8.187)	(7.381)	(8.227)	(0.251)
	P-value	0.519	0.935	0.433	0.247	0.288	0.606
(Licensing Prevalence) x							
(Mother with Young Child)	Beta	-0.116	0.023	-3.029	-9.024	-5.668	0.007
	Standard Error	(0.267)	(0.263)	(11.780)	(10.378)	(11.913)	(0.339)
	P-value	0.665	0.931	0.798	0.389	0.636	0.985
N		4351	4351	3760	3760	3351	5675
R^2		0.030	0.251	0.115	0.300	0.250	0.080

Notes: This table reports results from the model defined by equation (1), with "mother with young children" as the group indicator. For columns (1) through (5), the sample includes all female labor force participants who are aged 26 to 45 in Wave 13 of the SIPP who have valid data for the first 12 waves. Column (6) does not restrict the sample to labor force participants. The employment dependent variable and labor market participation dependent variables are binary indicators, and the wage dependent variable is the individual's hourly wage at his or her main job (coded to 0 for unemployed workers). The regressions all include control variables for gender, race, educational status, and age. Columns (2), (3), and (5) include fixed effects for 2-digit SOC classifications. Column (5) restricts the sample to include only workers who are earning nonzero wages. Standard errors are clustered at the state level.

Table 10: Regression Results - SIPP Model (Retirement-aged)

Dependent Variable:		Employment (Mean = .913)		Wage (Mean = 19.18)			Labor Force Participation (Mean = .393)
		(1)	(2)	(3)	(4)	(5)	(6)
		No Fixed Effects	Occupational Fixed Effects	No Fixed Effects	Occupational Fixed Effects	Occupational Fixed Effects, Employed Individuals Only	No Fixed Effects
Age in [70,75]	Beta	-0.065	-0.037	-6.615	-2.841	-9.720	-0.278
Overall Mean: .290	Standard Error	(0.119)	(0.123)	(9.940)	(9.198)	(10.287)	(0.089)
Labor Force Mean: .142	P-value	0.590	0.765	0.509	0.759	0.349	0.003
Age in [65, 69]	Beta	0.017	0.040	1.644	4.821	5.307	-0.166
Overall Mean: .325	Standard Error	(0.068)	(0.072)	(8.293)	(6.985)	(7.103)	(0.059)
Labor Force Mean: .291	P-value	0.801	0.577	0.844	0.493	0.458	0.007
Licensing Prevalence	Beta	0.005	0.027	26.302	33.170	28.190	0.412
Overall Mean: .224	Standard Error	(0.239)	(0.231)	(22.304)	(20.418)	(23.310)	(0.289)
Labor Force Mean: .225	P-value	0.983	0.908	0.244	0.111	0.232	0.160
(Age in [70, 75]) x (Licensing Prevalence)	Beta	0.354	0.313	21.499	7.145	37.209	-0.386
	Standard Error	(0.516)	(0.533)	(44.984)	(40.567)	(44.931)	(0.394)
	P-value	0.496	0.560	0.635	0.861	0.412	0.331
(Age in [65, 69]) x (Licensing Prevalence)	Beta	-0.025	-0.094	-15.759	-28.268	-29.456	-0.246
	Standard Error	(0.308)	(0.324)	(36.312)	(30.248)	(30.714)	(0.279)
	P-value	0.934	0.772	0.666	0.355	0.342	0.382
N		3423	3423	2385	2385	2088	8721
R^2		0.010	0.065	0.057	0.179	0.158	0.138

Notes: This table reports results from the model defined by equation (1), with two indicators: one for individuals aged 70 to 75, and one for individuals 65 to 69 (the base group is individualss aged 60 to 64). For columns (1) through (5), the sample includes all labor force participants who aged 60 to 75 in Wave 13 of the SIPP who have valid data for the first 12 waves. Column (6) does not restrict the sample to labor force participants. The interruption rates for labor force participants are: .176 for the youngest group, .219 for the middle group, .297 for the oldest group. The employment dependent variable and labor force participation dependent variables are binary indicators, and the wage dependent variable is the individual's hourly wage at his or her main job (coded to 0 for unemployed workers). The regressions all include control variables for gender, race, and educational status. Columns (2), (3), and (5) include fixed effects for 2-digit SOC classifications. Column (5) restricts the sample to include only workers who are earning nonzero wages. Standard errors are clustered at the state level.

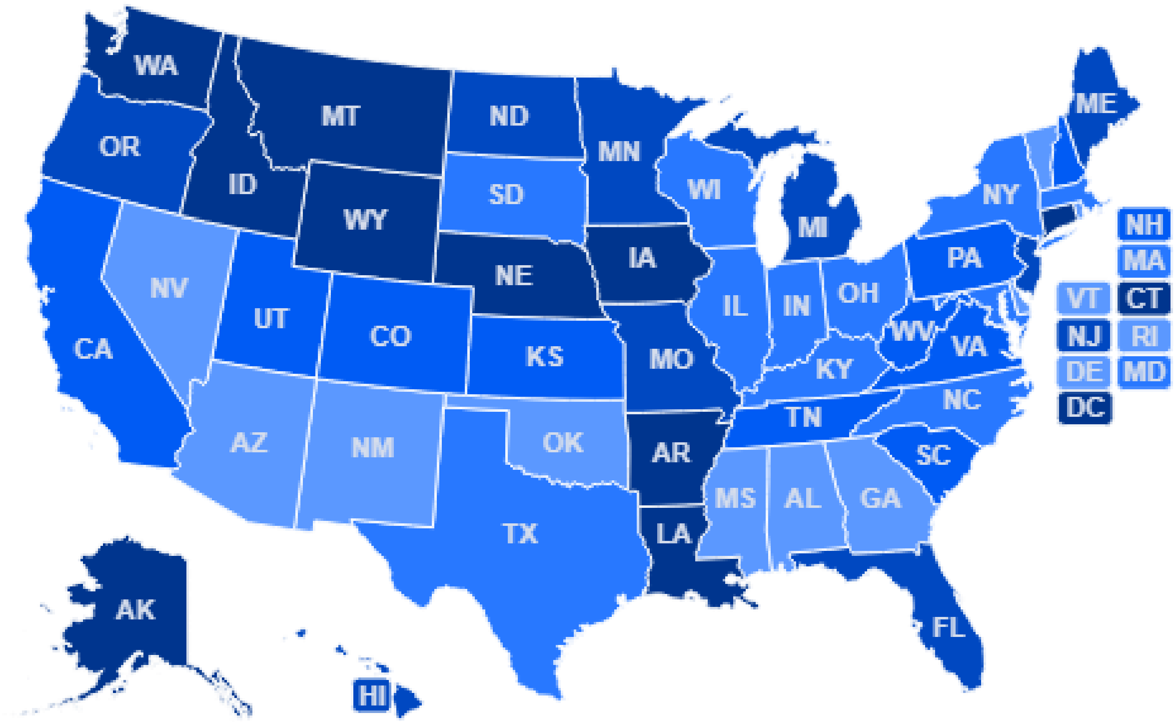
Table 11: Regression Results - CPS Model

		Dependent Variable: Difference in Difference, Interrupted vs. Noninterrupted Mean of Dependent Variable = -.000835, Standard Deviation =.0202				
		(1)	(2)	(3)	(4)	(5)
	Independent Variables	Continuous Independent Variable	Continuous Independent Variable	Dummy Variable - .10 Cutoff	Dummy Variable - .15 Cutoff	Dummy Variable - .25 Cutoff
	Control Variables?	No	Yes	Yes	Yes	Yes
Change in Licensing Prevalence	Beta	-0.0047	-0.0045			
	Standard Error	(0.0068)	(0.0069)			
	P-Value	0.493	0.510			
Gain License Dummy	Beta			-0.0003	0.0019	-0.0040
	Standard Error			(0.0022)	(0.0027)	(0.0048)
	P-Value			0.897	0.473	0.405
Lose License Dummy	Beta			-0.0042	-0.0046	-0.0049
	Standard Error			(0.0082)	(0.0082)	(0.0082)
	P-Value			0.605	0.578	0.552
N		762	762	762	762	762
R^2		0.120	0.127	0.128	0.128	0.128

		Dependent Variable: Difference in Difference, Mothers with Young Children vs. Others Mean of Dependent Variable = -.000212, Standard Deviation =.02162				
		(1)	(2)	(3)	(4)	(5)
	Independent Variables	Continuous Independent Variable	Continuous Independent Variable	Dummy Variable - .10 Cutoff	Dummy Variable - .15 Cutoff	Dummy Variable - .25 Cutoff
	Control Variables?	No	Yes	Yes	Yes	Yes
Change in Licensing Prevalence	Beta	0.0013	0.0002			
	Standard Error	(0.0071)	(0.0071)			
	P-Value	0.849	0.972			
Gain License Dummy	Beta			0.0020	0.0065	0.0040
	Standard Error			(0.0025)	(0.0035)	(0.0059)
	P-Value			0.436	0.062	0.504
Lose License Dummy	Beta			0.0007	0.0009	0.0033
	Standard Error			(0.0022)	(0.0027)	(0.0049)
	P-Value			0.769	0.733	0.500
N		762	762	762	762	762
R^2		0.146	0.156	0.172	0.176	0.172

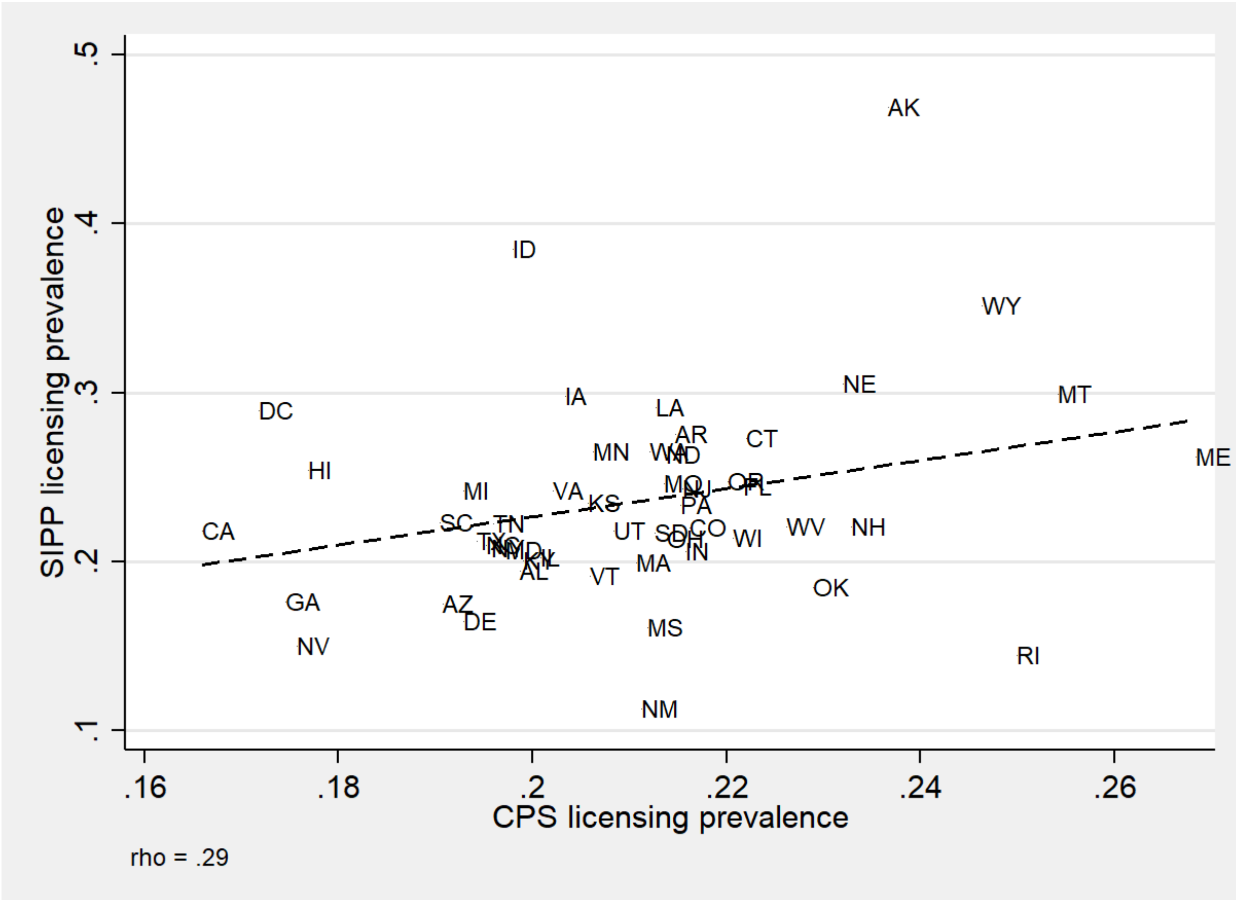
Notes: This table reports regression results from the model defined by equations (2) and (3). In the top panel, group i=1 is interrupted, and group i=0 is uninterrupted. In the bottom panel, group i=1 is mothers who have children who are younger than 8, and group i=0 is all other workers. Columns (1) and (2) report results using a continuous dependent variable for the change in licensing prevalence over time, while columns (3) through (5) report results which separate the dependent variable into categorical dummies with various cutoffs to define groups with large changes in licensing prevalence. All regressions are weighted by the total number of workers in the state-occupation. The sample is restricted to state-occupation pairs with 8 or more uninterrupted workers in each time period.

Figure 1: Heatmap of State Licensing Prevalences



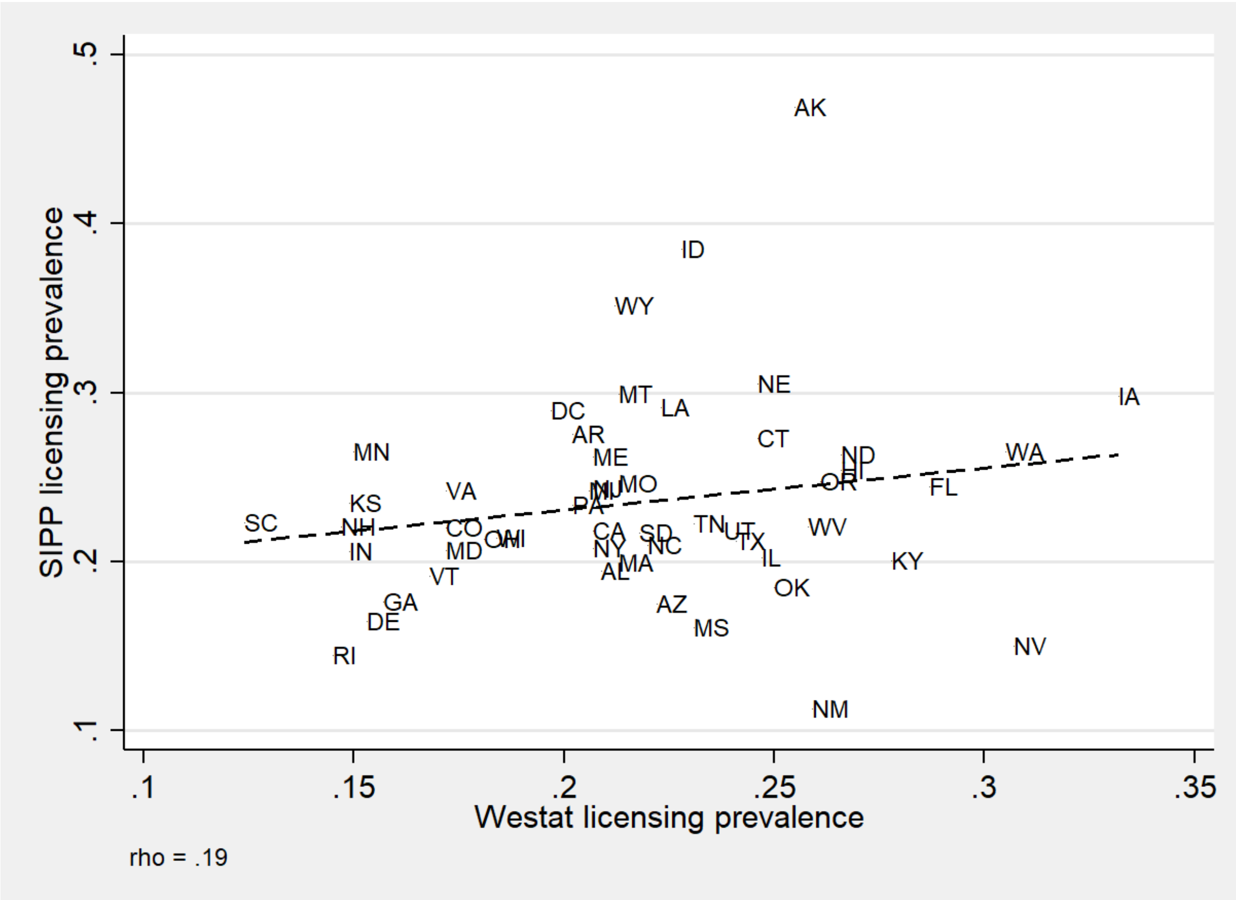
Notes: Measurements are from the SIPP. Licensing prevalence is defined as the fraction of labor force participants in a state from age 24 to 64 who report holding a license. The sample is respondents who have data available from waves 1-13. In the heatmap, ark colors correspond to licensing prevalences.

Figure 2: State Licensing Prevalence Comparison – SIPP vs CPS



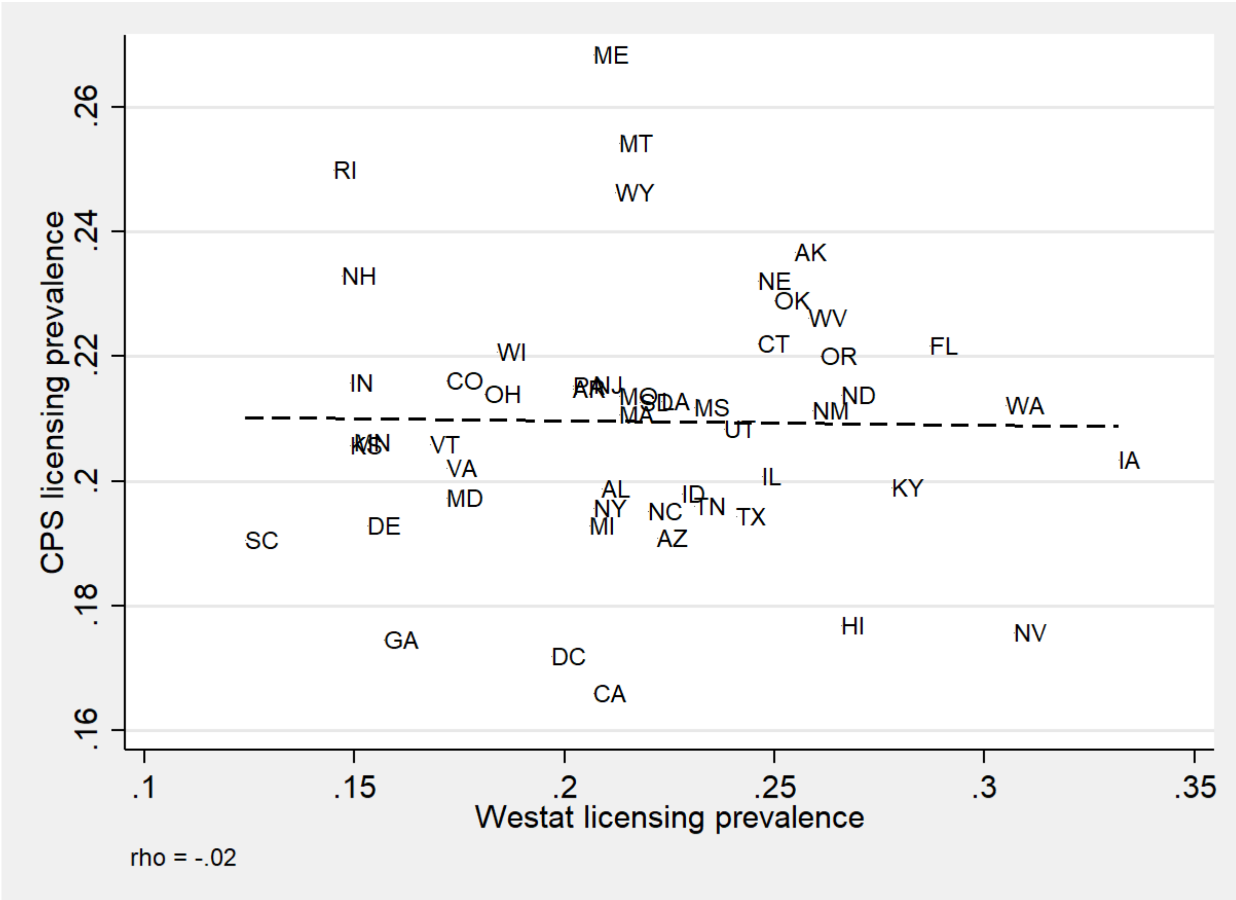
Notes: In the SIPP, licensing prevalence is defined as the fraction of labor force participants in a state from age 24 to 64 who report holding a license, and the sample is restricted to respondents who have data available from waves 1-13. In the CPS, licensing prevalence is the fraction of labor force participants from age 18 to 64 who report holding a license, and the sample is restricted to month 1 and month 5 labor force participants from January 2016 to November 2018.

Figure 3: State Licensing Prevalence Comparison – SIPP vs Westat



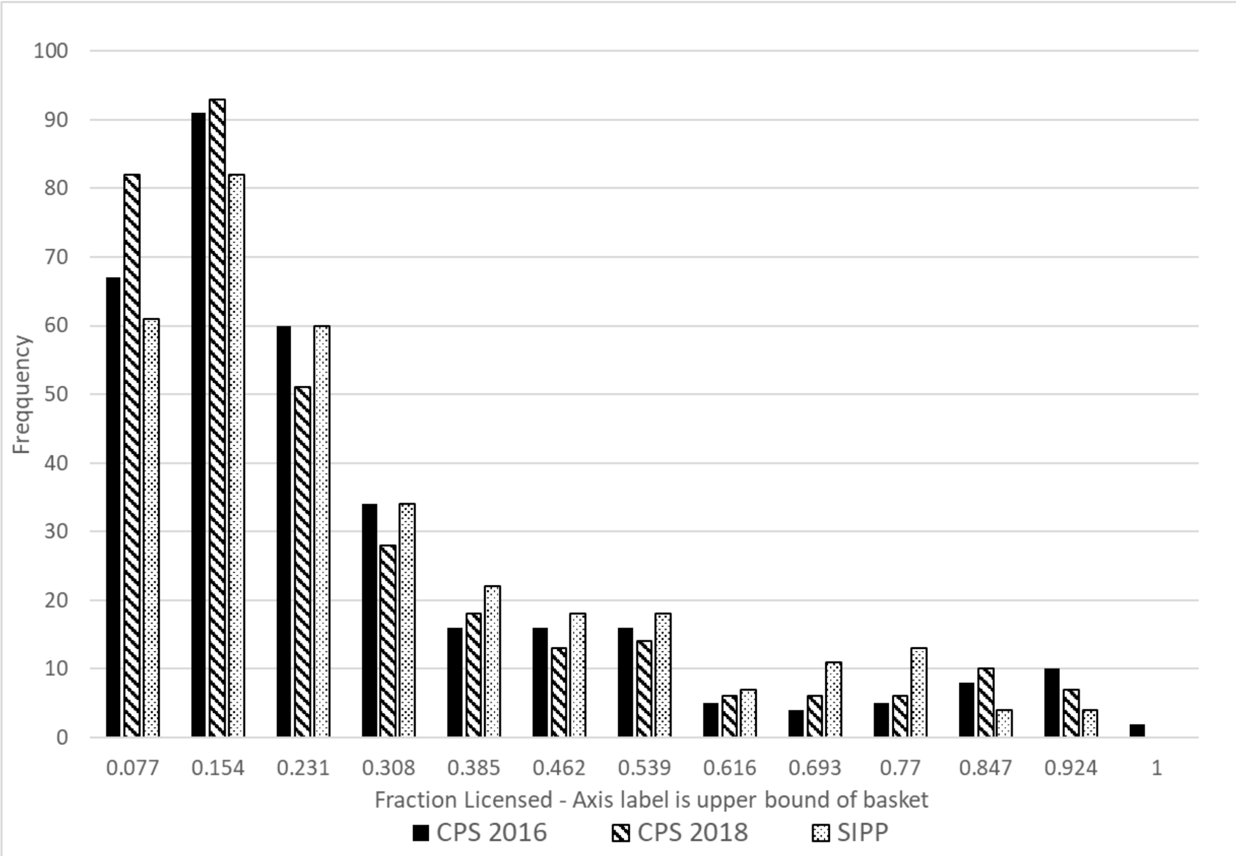
Notes: In the SIPP, licensing prevalence is defined as the fraction of labor force participants in a state from age 24 to 64 who report holding a license, and the sample is restricted to respondents who have data available from waves 1-13. The Westat results, published in Kleiner and Vorotnikov (2017), are from a phone survey of labor force participants aged 18 and older.

Figure 4: State Licensing Prevalence Comparison – CPS vs. Westat



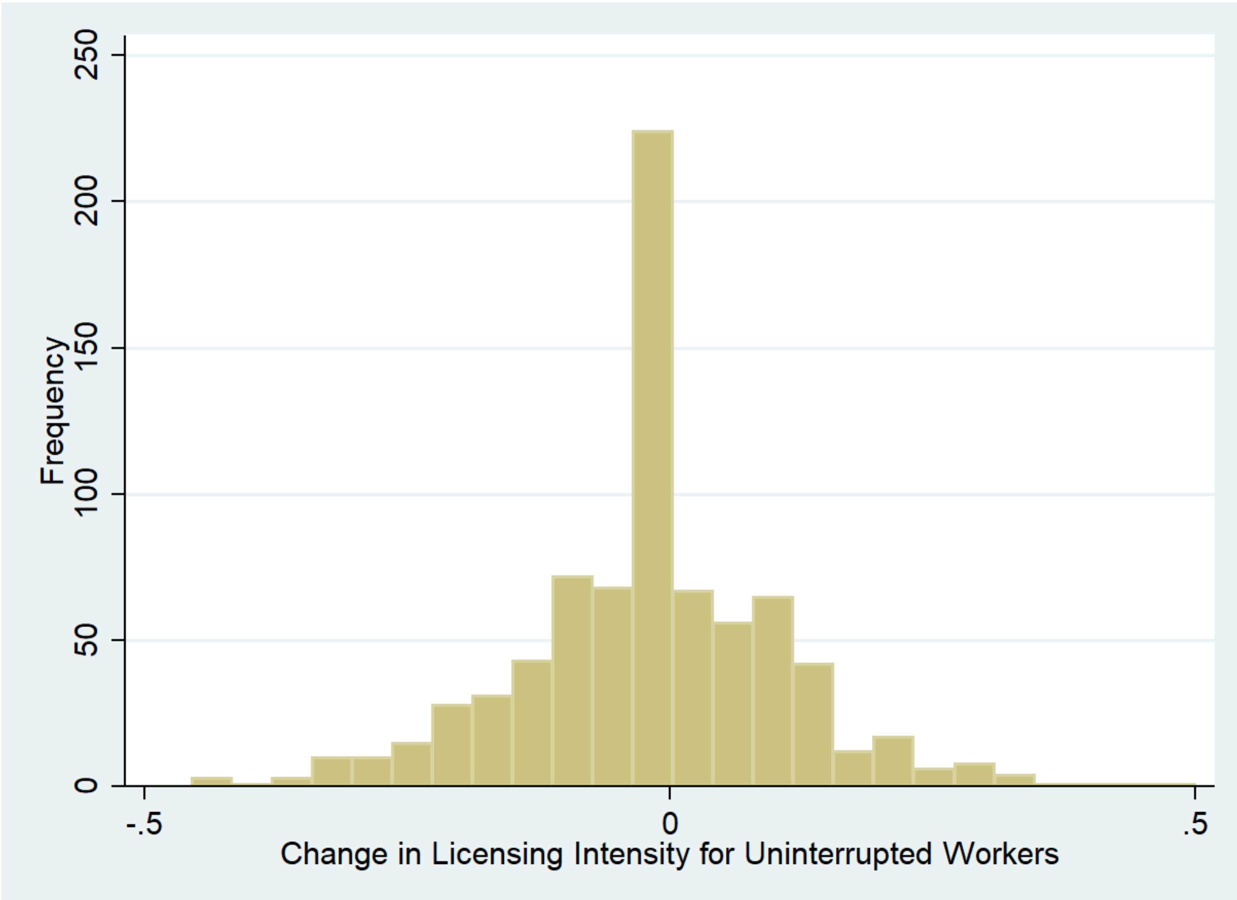
Notes: In the CPS, licensing prevalence is the fraction of labor force participants from age 18 to 64 who report holding a license, and the sample is restricted to month 1 and month 5 labor force participants from January 2016 to November 2018. The Westat results, published in Kleiner and Vorotnikov (2017), are from a phone survey of labor force participants aged 18 and older.

Figure 5: Overlapping Histogram of Licensing Prevalences



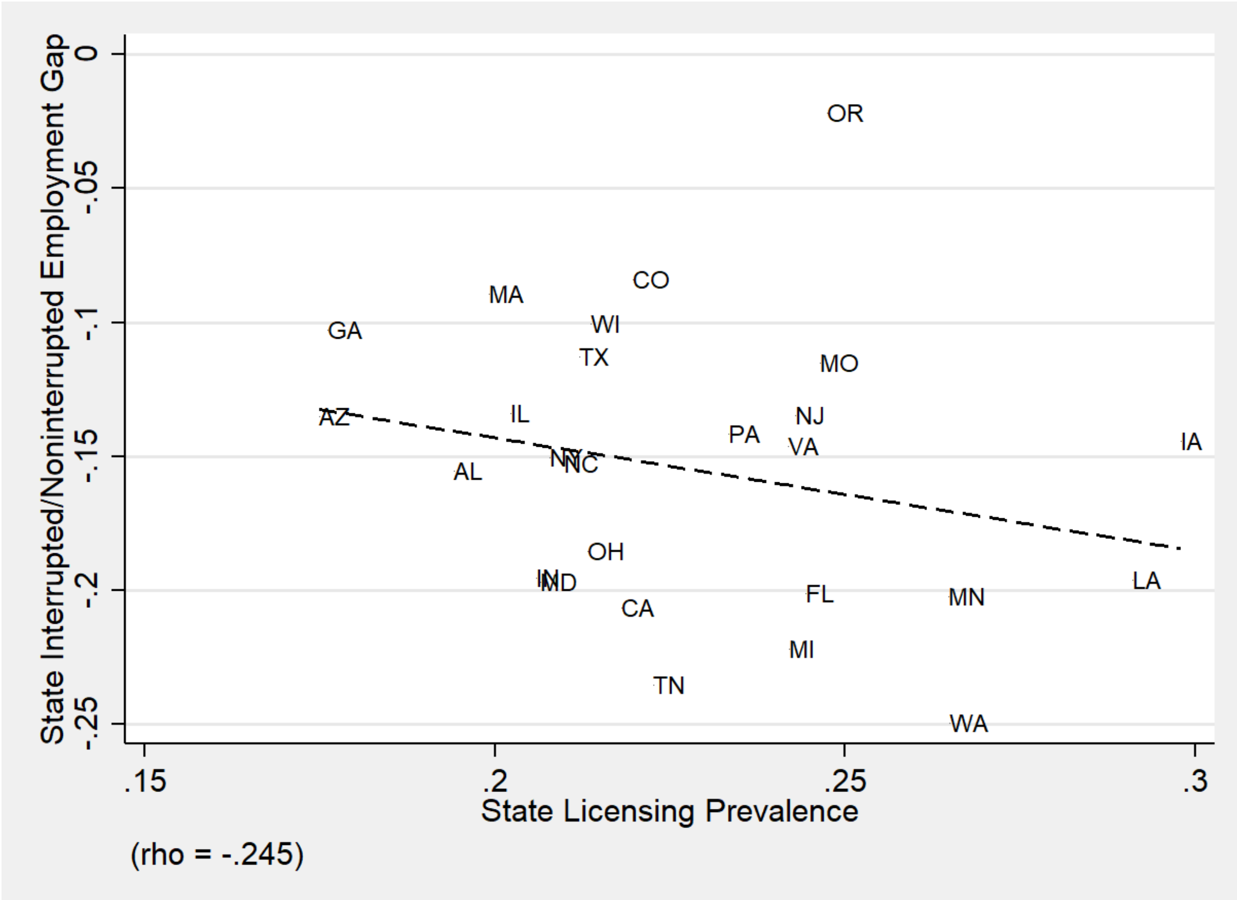
Notes: In the CPS, licensing prevalence is the fraction of labor force participants from age 18 to 64 who report holding a license, and the sample is restricted to month 1 and month 5 labor force participants. In the SIPP, licensing prevalence is defined as the fraction of labor force participants in a state from age 24 to 64 who report holding a license, and the sample is restricted to respondents who have data available from waves 1-13. Frequency corresponds to “number of occupations” whose licensing prevalence falls within a given range. I drop occupations with 5 or fewer respondents in either the SIPP or one of the two CPS periods. N = 336.

Figure 6: Histogram of $\Delta_{t,t-1}L_{j,s}$



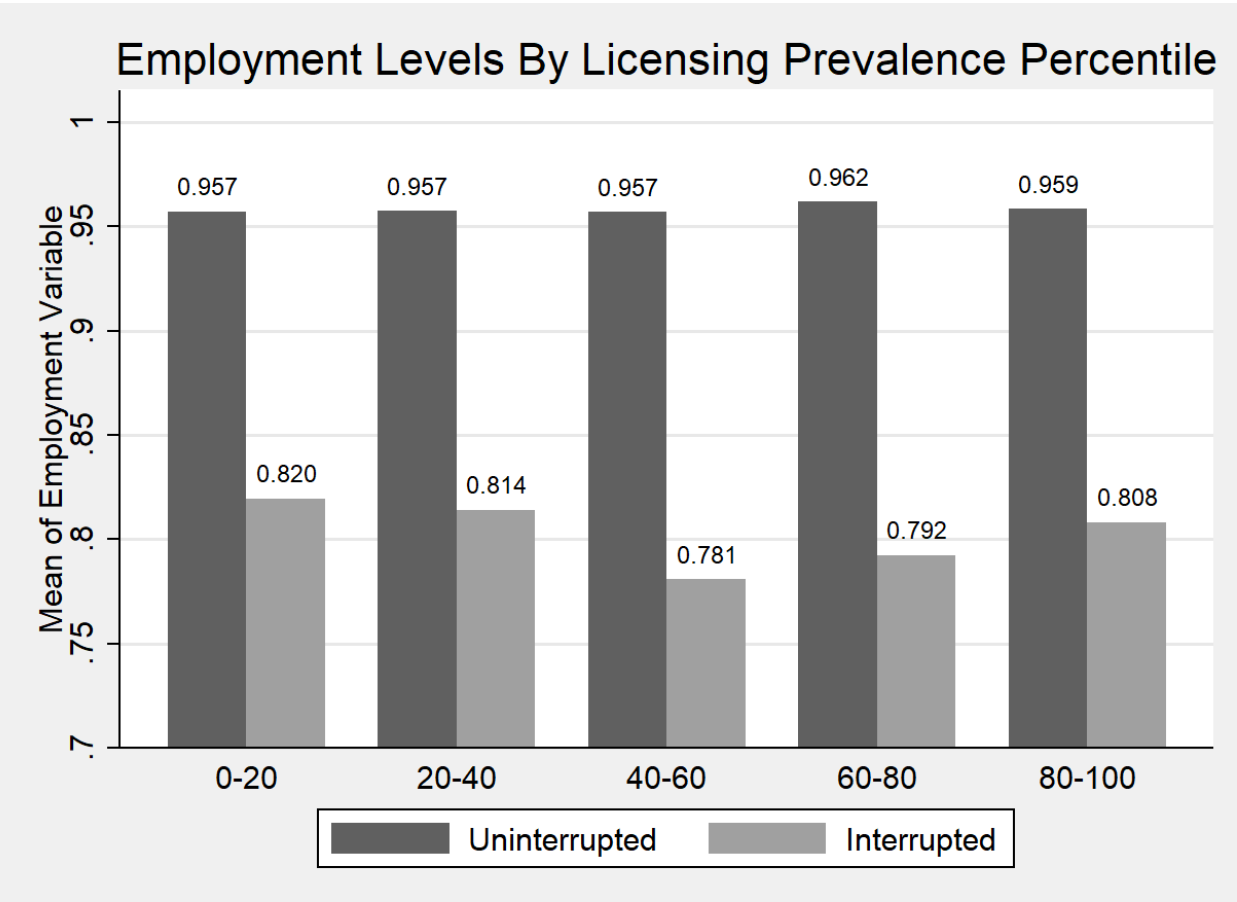
Notes: N = 789. The change in licensing intensity for uninterrupted workers is $\Delta_{t,t-1}L_{j,s}$. The construction of this variable is outlined in section IV. The sample includes all employed individuals who responded to months 1 through 5 of the CPS and whose month 5 survey occurred in January to June of 2016 or January to June of 2018. The sample is restricted to state-occupation pairs with more than 7 respondents in both 2016 and early 2018

Figure 7: Licensing Prevalence vs. Interrupted/Noninterrupted Employment Gap



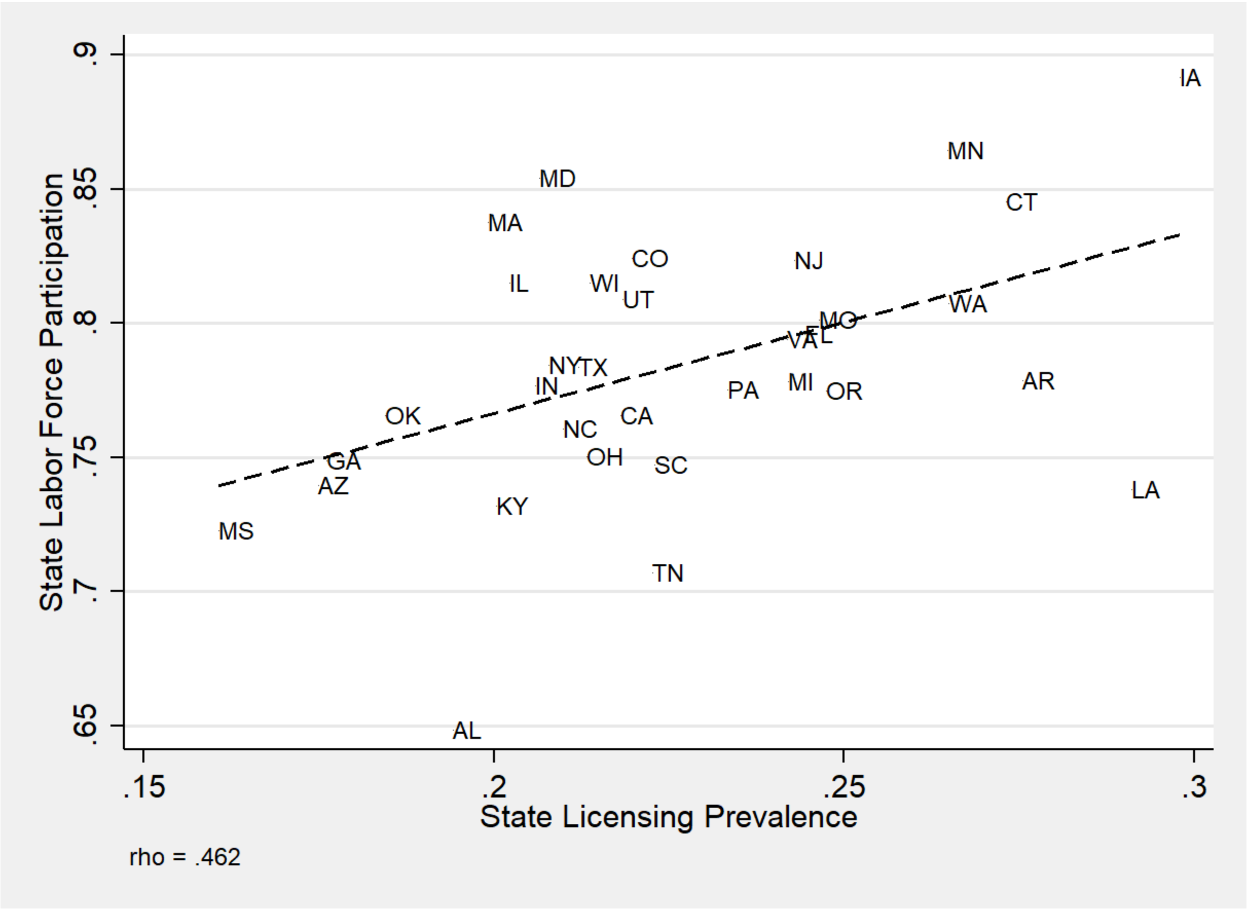
Notes: State licensing prevalence is the fraction of labor force participants in a state who report that they are licensed. The interrupted/noninterrupted employment gap is (fraction of interrupted labor force participants who are employed – fraction of noninterrupted labor force participants who are employed). The sample includes all labor force participants who are aged 26 to 64 in Wave 13 of the SIPP who have valid data for the first 12 waves. States with fewer than 200 total respondents are excluded.

Figure 8: Interrupted/Noninterrupted Employment Gap by Percentile of State Licensing Prevalence



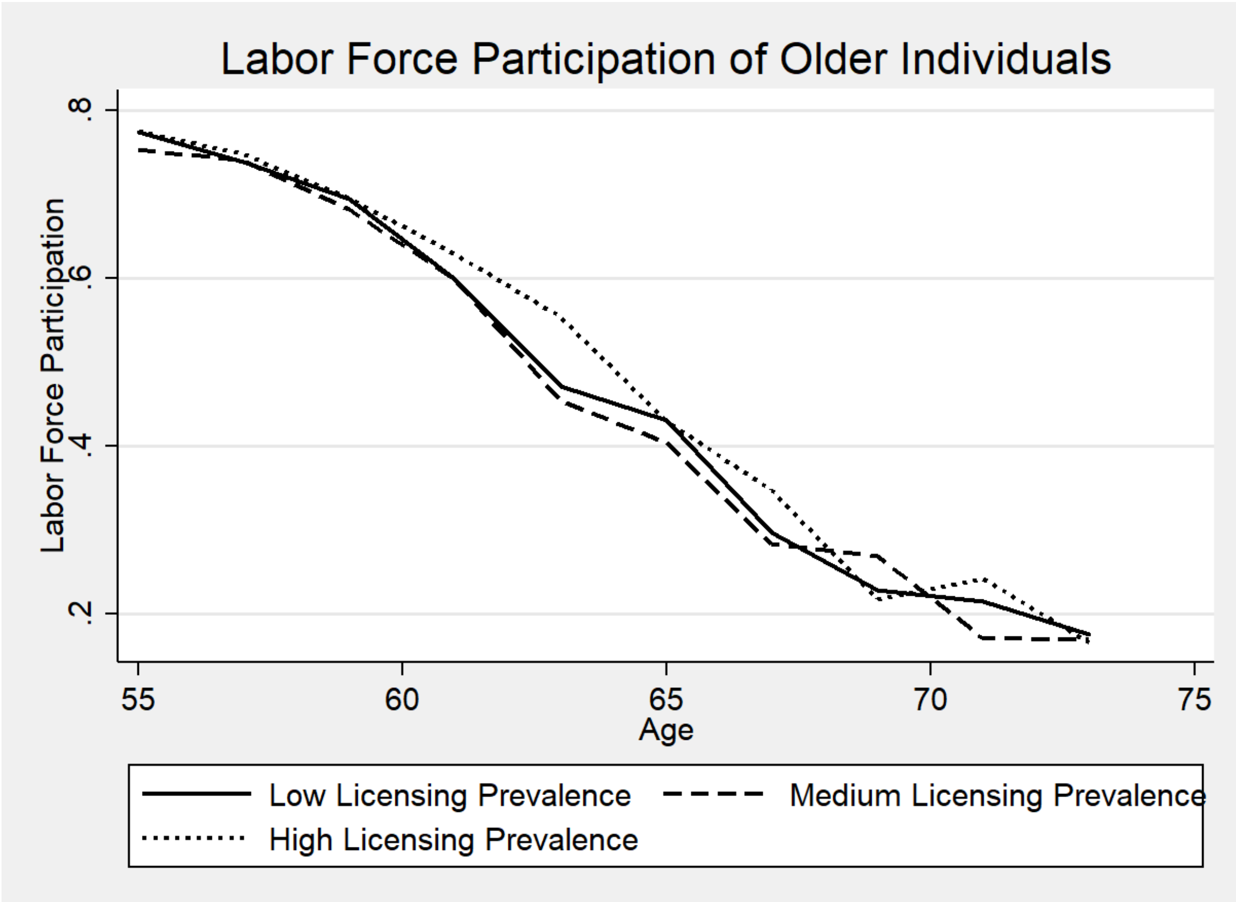
Notes: State licensing prevalence is the fraction of labor force participants in a state who report that they are licensed. The sample includes all labor force participants who are aged 26 to 64 in Wave 13 of the SIPP who have valid data for the first 12 waves. Percentile cutoffs are: 0-20 <= .206 <= 20-40 <= .214 <= 40-60 <= .223 <= 60-80 <= .245 <= 80-100.

Figure 9: State Licensing Prevalence vs. State Labor Force Participation



Notes: State licensing prevalence is the fraction of labor force participants in a state who report that they are licensed. State Labor Force Participation is the fraction of individuals who had a job or looked for a job in the last 4 months. The sample includes all individuals who are aged 26 to 64 in Wave 13 of the SIPP who have valid data for the first 12 waves.

Figure 10: Labor Force Participation of Older Individuals by Age and Licensing Prevalence



Notes: State licensing prevalence is the fraction of labor force participants in a state who report that they are licensed. State Labor Force Participation is the fraction of individuals who had a job or looked for a job in the last 4 months. Ages are grouped in bins of 2 (55/56, 57/58…) to increase the sample size for each bin. The low licensing prevalence line corresponds to respondents in the bottom third of state licensing prevalence, the medium line is the middle third, and the high line is the top third. The sample includes individuals who had valid data from the first 13 waves of the SIPP