Consumer Protection in an Online World: An Analysis of Occupational Licensing *

Chiara Farronato† Andrey Fradkin‡ Bradley J. Larsen§ Erik Brynjolfsson¶

July 12, 2019

Abstract

We study the effects of occupational licensing on consumer choices and market outcomes in a large online platform for residential home services. We exploit exogenous variation in the time at which licenses are displayed on the platform to identify the causal effects of licensing information on consumer choices. We find that consumers do not value the platform-verified licensing status of a professional, while they are heavily influenced by reviews and prices. We confirm this result in an independent consumer survey. We also use variation in regulation stringency across states and occupations to measure the effects of licensing on aggregate market outcomes on the platform. Our results show that more stringent licensing regulations are associated with less competition and higher prices and not with any improvement in customer satisfaction as measured by review ratings and the propensity to use the platform again.

---

*We thank Stone Bailey, Felipe Kup, Ziao Ju, Rebecca Li, Jessica Liu, Ian Meeker, Hirotaka Miura, Michael Pollmann, and Chuan Yu for outstanding research assistance. We thank the company employees for sharing data and insights and participants at ASSA 2018, Boston University, Collegio Carlo Alberto, FTC Microeconomics Conference, INFORMS Revenue Management and Pricing Conference, Institute for Industrial Research Stockholm, Lehigh University, Marketing Science Conference, Platform Strategy Research Symposium, SITE 2019 Occupational Licensing Conference, SOLE 2019, WISE 2018, and ZEW ICT for comments. We acknowledge support from grants through the Hellman Foundation, the Laura and John Arnold Foundation, the Russell Sage Foundation, and the MIT Initiative on the Digital Economy.

†Harvard Business School and NBER, cfarronato@hbs.edu

‡Boston University, fradkin@bu.edu

§Stanford University, Department of Economics and NBER, bjlarsen@stanford.edu

¶MIT Sloan and NBER, erikb@mit.edu

1 Introduction

Heated debates over the effects of occupational licensing date back hundreds of years, with a long treatise on the subject contained in *The Wealth of Nations* (Smith (1776)), and continue intensely today.[1] Occupational licensing is a restriction placed on who is allowed to perform certain types of services, requiring that practitioners meet licensing requirements in order to legally practice. These laws affect nearly 30% of the US labor force (Kleiner and Krueger 2010), a larger proportion of workers than are in unions or covered by minimum wage laws, and over 1,100 occupations are licensed in at least one state (Kleiner and Krueger 2010). These occupations include electricians, contractors, interior designers, and many more, although the stringency of the licensing requirements—and the range of specific tasks within a service category requiring or not requiring a license—varies widely from state to state.

Occupational licensing laws have come into scrutiny particularly recently with an antitrust suite in North Carolina (*NC State Board of Dental Examiners v. FTC*) and with specific attention from the Federal Trade Commission and Department of Justice to the role licensing laws may play in protecting consumers or in restricting competition.[2] On the one hand, licensing may protect consumers from poor service outcomes, guaranteeing at least some minimum standards of quality and safety for consumers (as in the model of Leland 1979). On the other hand, these laws may restrict competition, raise consumer prices, and increase rents for licensed professionals. The model of Shapiro (1986) demonstrates that the benefits of occupational licensing to some consumers may come at costs to other consumers who face higher prices due to licensing. The relative magnitude of these costs and benefits is unknown, and has been an important question for policymakers across the political spectrum.

In this paper we study a new dataset from a large online labor market where consumers can hire professionals for home improvement services. On this platform, a consumer can

[1] See, for example, recent discussions in the *New York Times* (Cohen 2016), *Wall Street Journal* (Zumbrun 2016), and *Forbes* (Millsap 2017).

[2] Given its stated mission of "Protecting consumers and competition" (https://www.ftc.gov/about-ftc), the FTC has dedicated recent roundtable discussions to the topic of occupational licensing, as has the DOJ.

post a request for a particular job. Professionals respond to this request with a quote. On this quote, the consumer can see the proposed price, several measures of the professional's online reputation (such as a 1–5 star average rating, the number of reviews, and the number of previous hires), and a flag indicating that the professional is licensed. This flag is only displayed if the professional has uploaded proof of licensure to the platform and the platform has independently verified this proof using state-level databases. Depending on the specific project needs or the required professional qualifications, a service provider may need a license in some jurisdictions but not others.

These features of the market and our transaction-level data from the platform provides an unprecedented opportunity to study several aspects of occupational licensing we address herein. First, we study how much weight consumers place on professionals' licensing status when deciding whom to hire, and we compare this effect to how much consumers value a professional's reputation and price. We also collect new survey evidence on how consumers think about these hiring decisions and how consumers value licensing. Finally, we use state-by-occupation level variation in licensing laws to measure how market equilibrium outcomes—entry of service providers, prices, and service quality—vary with licensing stringency.

The data examined in this paper consists of over a million requests by consumers in hundreds of distinct service categories throughout the United States for over eight months.[3] It comes directly from the company's databases, and allows visibility into most dimensions of the search and exchange process occurring through the platform. We discuss the data and institutional setting in section 2.

In section 3 we analyze how consumers decisions depend on the characteristics of professionals (their verified licensing status and online reputation) and their bids (prices). We begin with event studies that analyze consumers' probability of hiring a professional surrounding the exact date on which the professional's uploaded licensing status is verified by the platform. We exploit a unique feature of our data that allows us to identify the causal effect on consumers' decisions from displaying the professional's verified licensing status.

[3]The actual number of requests, the actual time frame, and the name of the company are not revealed to protect company's confidential information.

Professionals choose to upload proof of licensure, but this information is not displayed to consumers until some exogenous point in the future when the platform verifies the licensure. In the data, we see the timestamp for the original uploading of licensure proof by the professional and the timestamp for the platform's verification. We exploit this information in our event studies and we find no statistically significant change in the probability that a consumer hires a professional surrounding this verification date. In contrast, we find a discontinuous positive jump in the probability of hiring a professional following the first time that professional receives a review, suggesting that consumers respond to online reputation characteristics of professionals and not to indicators of licensure. We also examine whether, around the time of their license verification or first review, professionals themselves change their behavior in terms of prices they charge or types of requests on which they bid, and we find no evidence of these changes.

We then analyze consumer choices in a regression framework, where we regress consumers' choices to hire or not on our indicator of whether the professional has a verified licensing status and controlling for an indicator for whether the professional has uploading licensure proof, again allowing us to obtain the causal effect of the verified licensing signal. We also include price and online reputation measures in this regression (average star rating, number of previous reviews, and number of previous hires). These variables may be correlated with unobservable characteristics of the job request and the professionals' quality. We address this concern through a number of additional bid-level controls, request and professional fixed effects, and a novel instrumental variables strategy. In our regression framework we find similar results to our event studies: consumers appear to value professionals' reputation much more than their licensing status. We also find that consumers value prices much more than licensing status.

In section 4, we present the results of an original survey we conducted using a nationally representative panel of individuals who purchased a home improvement service within the past year. We find that the survey respondents typically list prices and reviews when they are asked about the factors that influenced their decision to hire a particular professional. In contrast, fewer than 1% of these respondents mentioned licensing status among the top 3 reasons for why they hired a given service professional—further evidence that consumers may care more about prices and online reputation than licensing status. This finding and our above findings using consumer decision data may simply reflect consumers thinking that all professionals are licensed. We asked survey respondents whether they knew the licensing status of the professional they ended up hiring. Only 61% of consumers were sure that their service provider was licensed and, of those, a majority only found out when they signed their contract rather than during their search, suggesting that most consumers may not be particularly knowledgeable of professionals' licensing at the time of their hiring decision.

In section 5 we consider the market level effects of licensing regulations. We exploit the large heterogeneity in regulatory stringency across occupations and states to measure the effect of licensing regulation—rather than the effect of licensing signals—on market equilibrium. To do this, we create a measure of licensing stringency at the level of each state and occupation using data on education, training, and other requirements of state licensing regulation. We study whether licensing stringency is associated with market level outcomes when controlling for a rich set of observables including time and geographic controls, service category, and request details submitted by the customer. We find that more stringent licensing laws are associated with less competition and higher prices, but have no detectable effect on two proxies of customer satisfaction: a customer's online rating of the service provider and their propensity to use the platform again.

Our paper points to the importance of digital technologies for the design of regulation. Online platforms (such as Uber) allow many occasional providers to offer their services, with little scrutiny of their licensing status. At the same time online markets make it easy to rate providers through online reviews. Friedman (1962) and Shapiro (1986) argued that a well-functioning feedback system can be an effective substitute for licensing by reducing the need for upfront screening or quality certification. The advent of online reputation mechanisms may be providing just such a system. If low-quality service providers can be easily and quickly identified by consumers' past experiences, the cost and benefit trade-off of occupational licensing might tip towards reducing licensing regulation. Our work suggests that, at least for the setting of residential home improvement services, consumers pay much more attention to reputation measures than licensing signals and more stringent licensing laws impose costs on consumers in terms of higher prices without corresponding benefits in terms of customer satisfaction. We should note that our customer satisfaction metrics—online ratings and return rates to the platform—are unlikely to take into account quality dimensions that are unobservable to the consumer during the transaction, or that may impact consumer safety in the long-run. We may also not have enough statistical power to detect extremely rare but costly mistakes made by service professionals.

Our paper is directly related to research studying whether occupational licensing laws protect consumers from poor service outcomes. Leland (1979) proposes a model of occupational licensing, where licensing restrictions aid in overcoming an Akerlof (1970) lemons problem faced by consumers of professional services. Perhaps surprisingly, empirical work has largely found non-positive effects of increased occupational licensing stringency on consumer outcomes. Previous studies found non-positive effects of more stringency licensing requirements on quality for electricians (Carroll and Gaston 1981), contractors (Maurizi 1980), dentists (Kleiner and Kudrle 2000; Carroll and Gaston 1981), accountants (Barrios 2019), and physicians (Kugler and Sauer 2005). Some positive effects are documented in Larsen (2015) in the market for teachers, with these effects accruing primarily to high-income areas. We are unable to detect quality benefits of higher licensing stringency for a wide variety of professions.

In contrast to most of the previous literature, which focuses on the effects of licensing on the labor market, our approach allows us to also measure how consumers respond to seeing a signal indicating that a professional is licensed across a variety of professions. Demand-side analysis related to licensing has received limited attention; exceptions include the work of Harrington and Krynski (2002) and Chevalier and Scott Morton (2008) on funeral homes, the structural welfare analysis in Kleiner and Soltas (2019), and the structural auction analysis of an online labor market in Krasnokutskaya et al. (2018). We focus on consumer choices within the digital platform setting, which is already an important channel for finding service professionals and is likely to become more important over time. Our empirical strategies are enabled by the fact that digital platforms collect transaction data, which allows us to control for detailed characteristics of desired jobs and professionals in order to study individual effects of licensing signals on consumer decisions and the effects of licensing stringency on market outcomes.

Our work also relates to research on the effect of licensing laws on competition and rents for professionals in the licensed occupation. A near-universal finding (confirmed in our study) is that increased stringency of licensing requirements raises wage in the occupation. Occupational licensing requirements have also been shown to harm professionals who are unable to meet licensing standards but whose services are nonetheless desirable to consumers, such as Vietnamese manicurists (Federman et al. 2006) or the recent case of hair braiders in South Dakota (Sibilia 2017). Work by Law and Marks (2009) and Blair and Chung (2018) suggests that licensing laws can also aid minority workers in signaling quality to potential customers. State-level licensing requirements may also impose limitations on labor mobility (Johnson and Kleiner 2017, Kleiner and Xu 2019, Buonanno and Pagliero 2019, and White House report in DOT 2015) and decrease productivity growth through reallocation as a result. DePasquale and Stange (2016) provide some evidence against this hypothesis for nurses.

2 Institutional Details

The data comes from a large US-only online platform which operates in all 50 states and offers consumers access to professional service providers in a variety of categories, such as interior design, home renovation, and painting. The platform allows customers to submit a project request. Several professionals are then allowed to submit quotes, consisting of a price and textual details of the service. The quoted price is not binding, and the actual payment takes place off the platform.

A nontrivial fraction of service providers bidding on the platform have submitted information on their occupational license in at least one service category, and a large fraction of the services require a license in at least some jurisdictions. All of these features together—the nature of physical tasks often requiring occupational licenses, the prevalence of licensed professionals, and the bidding process—make this platform an ideal market for studying whether and how the knowledge of occupational licenses matter in markets where reputation and other information about professionals is readily available to consumers.

This marketplace is distinct from other websites, such as Yelp (Luca 2016), that solely provide a directory of businesses and professionals with crowd-sourced reviews. It also differs from platforms matching consumers to professional freelancers providing digital services, such as Freelancer and Upwork (Pallais 2014), since projects on this platform are nearly all physical tasks. Finally, it differs from platforms such as Instacart or Amazon Mechanical Turk, which match consumers to service providers for tasks that require less professional training—typically physical tasks such as grocery pickup/delivery for Instacart, and virtual tasks such as image identification for Mechanical Turk (Cullen and Farronato 2015; Chen and Horton 2016).[4]

The platform works as follows. Interested professionals can join the platform and create a profile containing information about themselves and their services. They can also submit proof of a license to be verified by the platform. The platform then takes some time to verify the license. This process typically takes a few days with some variation across professionals. According to conversations with platform employees, this variation in time-to-verification is not dependent on the characteristics of the professionals during our study period and is as good as random.[5] Time stamps for both the initial license submission and the subsequent verification are contained in our sample.

An individual consumer requests a quote for a particular type of service, describing their needs using pre-specified fields as well as some additional open-ended fields. Professional service providers in the appropriate occupation who have profiles on the platform are then notified of the procurement request and may then place bids for the contract. A limited number of professionals are allowed to bid, and bids are passed on to the consumer on a first-come, first-priority basis. The professionals pay a fee to submit bids. As bids are submitted, the consumer can look up information about each of the bidders, and then may, if she chooses, select a service provider from among those bidders.

The information available to the consumer about each of the professionals submitting

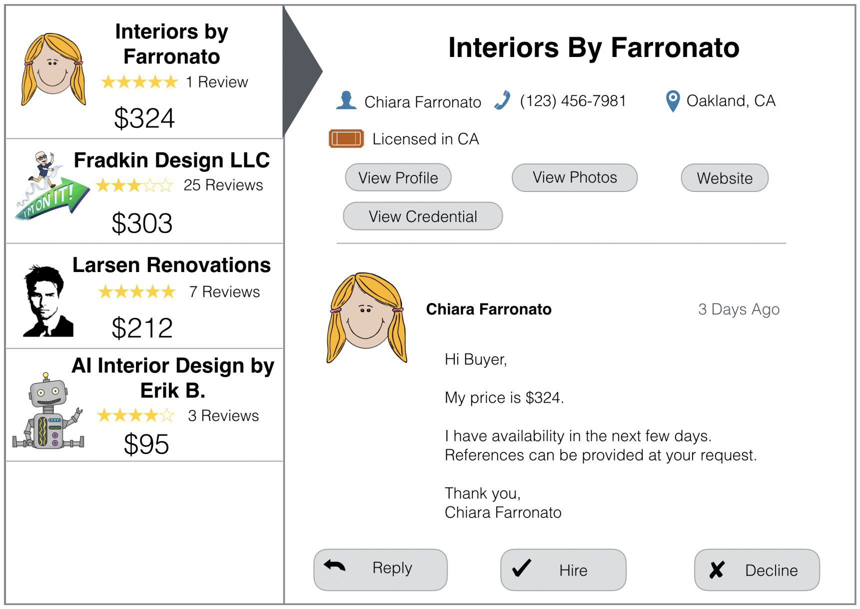
[4]See Horton (2010) for further discussion of online labor markets.

[5]Note that the verification process has changed over time within the platform. Our description reflects this process during the period for which we have data. Furthermore, professionals on the platform who do not display a license may still have a license but have just not reported it to the platform. The licensing effect we will measure in section 3 will be specifically the effect of displaying the license verification signal to consumers.

quotes varies by bidder, and may contain photos or detailed descriptions of the kind of work the professional has performed in the past. To some extent, the amount and type of information available depends on what the professional decides to share on the platform. A stylized depiction of a consumer's interface for choosing a professional is available in Fig. 1. Importantly for our study, for each bidder, the consumer is able to see any licensing information reported by the bidder. This licensing information is prominently visible if it has been verified by the platform. The consumer is also able to see any reviews of the professional's past work for other consumers, along with a 1–5 star average rating, the number of the previous reviews, and the number of previous times the professional has been hired through this platform. In particular, reviews can come from services exchanged on or off the platform. If the review is submitted by a consumer who hired the professional through the platform it is denoted an *on-platform* review. Otherwise, it is an *off-platform* review. More detailed information is available if the customer clicks on the professional's profile.

We define professionals as *verified licensed professionals* if they choose to upload proof of license to the platform and the platform verifies the validity of the license. There is a high degree of variation in the fraction of professionals who report a license to the platform, which is key to our empirical strategy. It is important to make two remarks. First, a professional who does not display the licensing badge might nonetheless have a professional license, and would be able to disclose this to consumers via other means—e.g. private messages or by including it in the body of text on their profile. Second, depending on the profession, an unlicensed professional may still legally provide services, but might be restricted in how they refer to the services they offer. For example, in the case of interior designers in Florida, a professional is legally not allowed to refer to themselves as an "interior designer" without a license, and will often instead describe their work using terms like "interior decorator", "interiors" or "organize your place". However, within the data, these professionals can still be identified as providing services similar to interior design. Unlicensed professionals may also provide services within a profession that typically requires a license if the project satisfies certain characteristics. For example, some states require professionals to have a license for commercial work—e.g., electrical work in a public building—but not for work

Figure 1: Stylized Representation of Platform



This is a reproduction of the information about professionals displayed on the platform. The layout and the identity of the people displayed are products of the authors' imagination.

in a private home. Or for general contractors in California, a license is only required if the payment for the services is over $500. We provide an analysis of this latter rule in Appendix B. Our results regarding verified licenses in section 3 should be interpreted as relating to the signaling value of the licensing information rather than the entire value of licensing regulation.

The sample that we use contains the following restrictions. We first limit the sample to an eight-month period during 2015 for which we could observe both the timing of license submission by the professional and the license verification by the platform. We then drop any requests containing hourly price quotes (as opposed to bids containing a fixed price quote for the job or no price quote); requests with any price quotes less than $20 or greater than $3,000; requests with missing date/time stamps; requests in service categories that never contain licenses; and a small number of requests in which more than one professional is recorded as having been hired (which are likely misrecorded) or requests that are extreme outliers in terms of having a very high number of bids.

Table 1 displays summary statistics at the bid level for requests in our selected sample. Beginning with the licensing related variables, we see that 12% of bids are by professionals with a verified occupational license and 14% are by professionals who have uploaded proof of license. It is possible for professionals to signal their licensing status in ways other than the structured platform verification, such as through the text of their profile or the text of their quote, both of which the consumer can observe. We do not observe this information in our primary data sample; our analysis of consumer response to the licensing status of professionals will only focus on the formal licensing verification signal provided by the platform. In Appendix A we discuss an independent data sample we construct (through a web crawl of the platform) in which we do see professionals' profile text. There we find that about 10% of professionals mention a license in their profile text and 6% have a license status verified by the platform.

Table 1 demonstrates that the median bid comes from a professional with 4 reviews, a rating of 4.8 stars, and a fixed price of $200. 6% of bids result in a recorded hire and hired bids are made by professionals with more reviews and higher ratings, lower prices, and similar licensing-related variables as the typical bid. The platform relies on either customers or professionals to voluntarily mark a job as hired. This means that not all hires resulting from the platform will be present and that some hires may not be accurately logged. We return to some of these issues in section 3 where we discuss our empirical specification.

3 The Determinants of Consumer Choice

In this section we study how professionals' licensing status, prices, and online ratings affect consumer choices of whom to hire. We offer two ternative approaches to analyze consumer sensitivity to licensing and reputation information: an event study approach and a regression analysis. Both approaches lead us to conclude that consumer choices are affected by online ratings and prices much more than by occupational licensing information.

Table 1: Summary Statistics: Bid Level

Variable	**min**	**median**	**max**	**mean**	**sd**
License Validated	0	0	1	0.12	0.32
License Submitted	0	0	1	0.13	0.34
Pro Reviews	0	4	386	9.90	19.13
Pro Rating	1	4.9	5	4.73	0.49
Hired	0	0	1	0.07	0.26
Price	20	199	3000	380	513
License Verified Given Hire	0	0	1	0.11	0.31
License Submitted Given Hire	0	0	1	0.12	0.33
Pro Reviews Given Hire	0	6	344.0	14.6	25.4
Pro Rating Given Hire	1	4.9	5	4.80	0.35
Price Given Hire	20	125	3000	252	363

Summary statistics at the bid level.

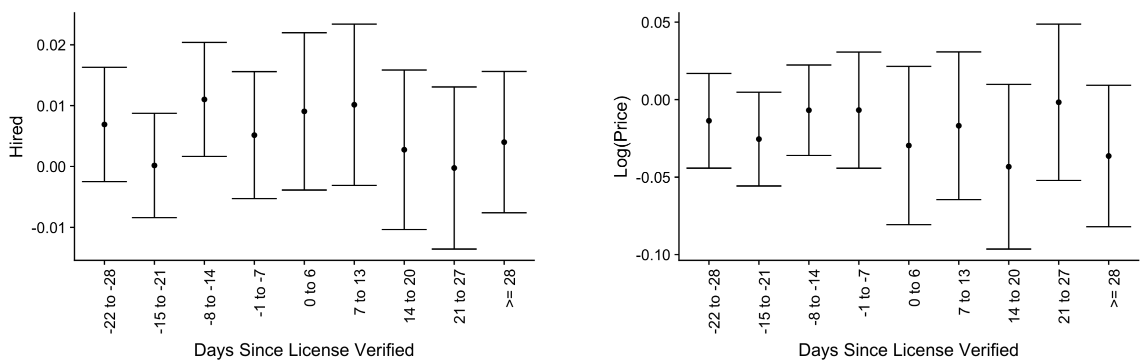
3.1 Event Study

We start by describing our event study. Our platform data allows us to measure each opportunity that a professional has to get hired, as well as the hiring outcome. We consider the probability with which a professional is hired for a job to which she submitted a bid around the time of license verification. If license verification positively affects consumer choices then bids submitted a few days before license verification should have a lower chance of being chosen than bids submitted just after the license is verified (and thus visible to consumers). More formally, we regress an indicator for whether a professional was hired on dummy variables for the leads and lags relative to the license verification day. We also add provider fixed effects to control for unobserved heterogeneity across professionals, and request fixed effects to control for the particular request and amount of competition. Our specification is the following:

$$hired_{jr} = \sum_{t=-4}^{5} \beta_t * \mathbf{1}\{diff_{jr} = t\} + submitted_{jr} + \gamma_j + \mu_r + \epsilon_{jr}, \tag{1}$$

where $diff_{jr}$ is the difference (in weeks) between the date of professional j's bid on request r and the date professional j's license was verified by the platform. Week $t = 0$ starts on the day the license is verified. We consider weeks within an eight-week interval around platform verification, and include a dummy variable for whether a bid was submitted more than four weeks after license verification. The variable *submitted* is an indicator for whether professional j has uploaded a license at the time of request r. Request fixed effects are denoted μ_r, and professional fixed effects are denoted γ_j. The β_t coefficients should be interpreted as hiring probabilities relative to the probability of being hired for a bid submitted more than four weeks prior to license verification.

Figure 2: Event Study Estimates—License Verification



(a) Hired

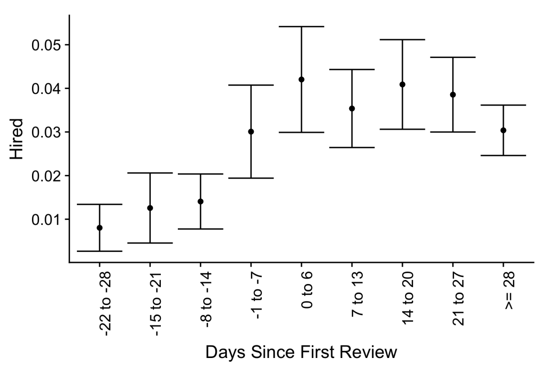
(b) Log(Price)

Figure 2a displays the estimated coefficients β_t from Equation 1. We find no significant differences in the probability of being hired as a function of when the bid was placed relative to the time of license verification. The estimated coefficients also show no significant pre-trend in the likelihood that a professional is hired prior to the license verification date, consistent with our assumption that the timing of verification is exogenous. Overall, the results suggest that consumers' decisions of whom to hire are not influenced by the visibility of licensing information, although the 95% confidence interval does not exclude sizable effects. In subsection subsection 3.2 we use an alternative identification strategy which yields more precise estimates and confirms that knowing about a professional's licensing status does not affect consumer decisions. We also investigated whether there may be a positive effect of the license signal for professionals without a prior hire. We find suggestive but imprecise evidence of such an effect (see Appendix C).

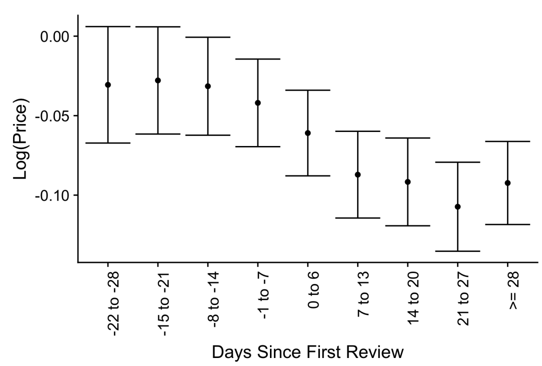
One potential threat to the identification of the effect of displaying licensing information is that professionals may adjust their bidding behavior around the time of license verification. We examine this by repeating the estimation of Equation 1 using the professional's quoted price as the left-hand-side variable of interest (Figure 2b). We find no significant differences in bid prices across these time periods, suggesting that professionals do not appear to be bidding differently in anticipation of or after license verification. We also find no changes in the *types* of requests professionals bid on and the *timing* of these bids as a result of the license verification (see Appendix C).

To show that our empirical strategy would be able to pick up important determinants of customer choices, we repeat the above exercise using the first review received by a professional as the relevant event. We use the first review because it is typically a 5-star review, so we do not need to differentiate between good and bad ratings.[6] To be more precise, we estimate the same specification as in Equation 1 but substitute the timing relative to license verification with the timing relative to the submission of the first review. We exclude bids that lead to the first review in the specification so that there is no mechanical relationship between first review and hire.

Figure 3: Event Study Estimates—Reviews



(a) Hired



(b) Log(Price)

Figures 3a and 3b display the estimated coefficients β_t from the event study of the first review. We can see that there is a 10 percentage point jump in hiring rates around the time of the first review and a smooth decline in prices around the focal date. It is worth noting

[6]We find similar results when we instead consider the first five-star review received by the professional.

that there seems to be a pre-trend in 3a with an increase in the hiring rate in the seven days preceding the first review. Our hypothesis for this effect is that customers may take some time to decide whom to hire, and their final decision for a given request may occur *after* the first review is revealed; if this is true, consumers' hire rates would appear to react to reviews several days before the arrival of the review.

To investigate this hypothesis, we re-estimate the event study using a closer approximation to the time at which the customer made a choice: rather than comparing the arrival time of the first review to the arrival time of the bid, we compare it instead to the time the consumer first messaged the professional about this request. We also limit the sample to cases where the 'hired' button was clicked by the customer rather than by the professional, since the professional might be more strategic in timing when to click "hired".[7] Note that this sample is substantially smaller and we consequently get wider confidence intervals. The results are displayed in Figure 4a, where we find much less of a pre-trend, and we see a sharp and similarly sized increase in the hire probability following the display of the first review. There is no similar discontinuity in professionals' quoted prices around the time of the first review (Figure 4b), suggesting that professionals are not discontinuously changing their pricing behavior surrounding this event. Appendix D discusses further event studies which suggest that the effect of reviews on hiring does not seem to be driven by supply side responses. It also shows that the effect of the first review is driven by first reviews with high ratings.

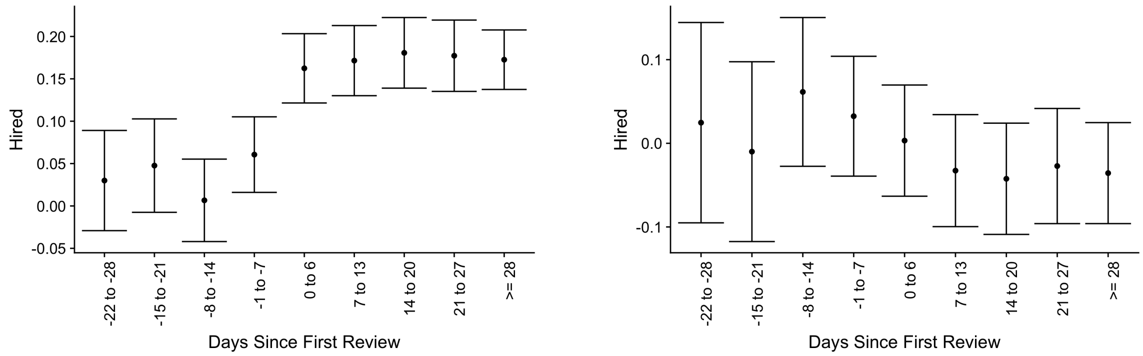
3.2 Choice Regressions

We now present a regression framework for measuring the effects of displaying licensing status and the effects of professionals' prices and online reputation on consumer choices. For professional j's bid on request r, we specify the indicator for whether j gets hired as follows:

$$\begin{aligned} hired_{jr} = {} & \beta_0 + submitted_{jr}\beta_1 + verified_{jr}\beta_2 + log(price_{jr}+1)\beta_3 + \\ & log(reviews_{jr}+1)\beta_4 + previous\_hires_{jr}\beta_5 + avg\_rating_{jr}\beta_6 + X'_{jr}\beta_7 + W'_{jr}\beta_8 + \varepsilon_{jr}, \end{aligned} \tag{2}$$

[7]Either the professional or the customer is allowed to click the "hired" button.

Figure 4: Event Study Estimates—Reviews (Subsample)



(a) Hired

(b) Log(Price)

Figure uses subsample of requests in which hired button was not clicked by the professional but by the customer, and measures time since first review relative to the time the customer first messaged the professional.

As in the event study, we control for the license submission decision (*submitted*). This indicator is visible to us in the data, but the professional's reported license status is not visible to the consumer until verified ($verified_{jr} = 1$). We can then interpret the coefficient β_2 on the verified variable as the causal effect of knowing that a professional is licensed on the hiring probability. The variable $price_{jr}$ is professional j's quoted price for request r; $reviews_{jr}$ and $previous\_hires_{jr}$ represent the number of reviews and number of previous hires the professional has received before submitting a quote on request r; and avg_rating_{jr} is j's average star rating (1–5) at the time of submitting the bid on request r. The vector X_{jr} includes at least an indicator for whether the quote is missing a price (in which case $price_{jr}$ is also set to zero), an indicator for whether $reviews_{jr} = 0$, an indicator for whether $previous\_hires_{jr} = 0$, and a flexible set of controls for the time of a request relative to the license submission time.

The vector W_{jr} differs depending on our specification. In our simplest specification, W_{jr} is omitted. In our next set, we include in W_{jr} a quadratic term for the time the professional has been registered on the platform; the character length of the text of the professional's quote; indicators for whether the professional has a business license submitted and whether this business license is validated (a business license is distinct from an occupational license);

indicators for whether the professional's profile has pictures, has a website link, lists the number of employees, and provides a date of establishment of the business; indicators for the arrival order of the bids for the request. In our most flexible specifications, W_{jr} includes (in addition to the previous variables) request fixed effects, or both request fixed effects and professional fixed effects.

Column 1 of Table 2 displays the results of estimating (2) without any additional controls W. Consistent with our event study analysis, we find no effect the licensing signal on the hiring choice (the coefficient on $verified$ is a precise zero). We do find significant positive impacts for each of the reputation measures (previous hires, average rating, and number of reviews) and significant negative effects of prices. Each of these variables is potentially correlated with the difficulty of the job the consumer needs done. In Column 2, we account for this possibility by including request fixed effects, in addition to the other controls listed above. We find virtually the same results. Professionals' price quotes and their online reputation may also be correlated with other features of the quality of the professional that are observable to the consumer but not to us. We account for time-invariant unobserved quality of the professional by adding professional fixed effects in Column 3. All results remain similar except for the $previous\_hires$ coefficient, which flips signs.

Prices are likely also correlated with unobservables at the bid level not accounted for with our additional controls, such as time-varying dimensions of professional quality. To address this concern, in columns 4 and 5, we instrument for price in (2) using the geographic distance between the consumer's zip code and the professional's zip code. The majority of the services on the platform would require working in the home of the consumer. This location requirement implies that the professional's costs should be increasing in this distance, but this distance is unlikely to directly affect the consumer's willingness to pay for the service.[8]

[8] More precisely, we instrument for price with this distance measure, along with a dummy for whether the professional and consumer are in the same location and a dummy for whether the professional and consumer are more than 100 miles apart. One may be concerned that customers prefer professionals located very close the them, and indeed the survey evidence presented in section 4 suggests that customers care about whether the professional is "local". Adding to the second stage a dummy variable for whether the professional is located in the same zip code as the consumer and a dummy variable for whether they are located more than 100 miles away does not change our results.

Column 4 displays the results with request fixed effects, and column 5 with request and professional fixed effects.[9] In each case we find a much larger magnitude for the price coefficient, consistent with price endogeneity. We continue to find significant positive effects for most reputation measures. We do see a marginally significant (at the 10% level) positive effect for the license verified signal in column 4, but any significance disappears when we add professional fixed effects in column 5.

Similar to prices, the reputation measures in Equation 2 may be correlated with time-varying unobservables that relate to the quality of the professional. This could hinder a comparison of the license verified effect to the effects of online reviews. We therefore propose, as an additional specification, an instrumental variables strategy based on the work of Chen (2018). Specifically, we instrument for a professional's current rating using the ratings that the focal professional's raters (i.e. those who have rated the focal professional until now) assigned to *other* professionals. Similarly, we instrument for the professional's current number of reviews using the propensity of the focal professional's previous hirers (i.e. those who have hired the focal professional) to leave reviews on others whom they hired. Finally, we instrument for the professional's number of previous hires using the propensity of the previous consumers this professional has bid on to hire vs. not hire. We describe the construction of these instruments in Appendix E. Column 6 displays the results using these instruments in addition to the price instruments. There we again see a large negative effect of price, and a significant and positive effect of most reputation measures. In this specification, the $reviews$ effect is negative and insignificant and the $previous\_hires$ effect is positive and significant; we believe the sign changes for these two variables in Table 2 are likely due to a high degree of correlation between the two measures.

This last specification, which only includes request fixed effects, not professional fixed effects, displays a positive and marginally significant (at the 10% level) effect of licensing. The bulk of the evidence from Table 2 suggests that this effect is non-existent; however, even a generous interpretation of the licensing effect—taking it to be a positive 0.0112 effect on the hired probability (from column 6)—does not imply a very economically meaningful

[9] First stage estimates corresponding to columns 4 and 5 are found in Table E.1 and first stage estimates for column 6 are found in Table E.2.

effect. In particular, comparing this effect to the price coefficient from column 6 suggests that the licensing signal is worth a drop in price from \$200 (the median) to about \$195. And comparing to the average rating coefficient suggests that the licensing signal is worth about 0.03 of a star. To summarize, we fail to find substantial effects of knowing about a professional's licensing status on consumer choices. In contrast, we find that reputation measures and quoted prices have important effects on hiring probabilities.[10]

4 Survey Evidence

To dive deeper into how consumers think about (or don't think about) licensing when making choices, we conducted a survey of a nationally representative sample of customers about their choices regarding home improvement professionals. Our survey panel was created by the service ProdegeMR and consisted of 12,760 respondents, of whom 5,859 hired a home services professional within the past year and 5,219 of those fulfilled additional validation criteria to be considered a reliable response. The survey questions are available in Appendix F.

We first asked respondents about the service they purchased. The most common word stems included 'paint' (10.1%), 'replac' (8.4%),'plumb' (8.3%), 'repair' (7.6%), 'instal' (7.5%), and 'roof' (6.5%). Broadly, the services purchased by the survey respondents mirror the services purchased on the platform. When we categorize the responses according to occupations, we find that the most common occupations include HVAC contractors (20%), plumbers (19%), and painting contractors (10%).

Many consumers find their service providers online, validating the importance of studying consumer choices in online platforms. The modal way through which consumers find service providers is still word of mouth through a friend (53%), but Google and Yelp are used by 25% of the respondents, and 16% say they used a platform like the one we study. Note that for those consumers who said they used Google, the exchange may in fact have

[10]We find very similar effects to those shown in Table 2 when we use a conditional logit model instead of linear probability model. Marginal effects from this model are shown in Table E.3, where we estimate regressions corresponding to columns 2, 4, and 6 of Table 2. We do not estimate professional fixed effect versions of the conditional logit as this is infeasible in a nonlinear model.

Table 2: Consumer Choice Regressions: Outcome = Hired

	(1) OLS	(2) OLS	(3) OLS	(4) Price IVs	(5) Price IVs	(6) All IVs
License Submitted	-0.254***	0.0132**	-0.0056	0.00696	-0.00729	0.00387
	(0.00529)	(0.00537)	(0.00586)	(0.00749)	(0.0102)	(0.00695)
License Verified	-0.003	0.00224	0.0046	0.0121*	0.0118	0.0112*
	(0.00441)	(0.00455)	(0.00495)	(0.00633)	(0.00839)	(0.00581)
Log(Previous Hires + 1)	0.0370***	0.0194***	-0.0704***	-0.00203**	-0.0792***	0.0166***
	(0.00042)	(0.000562)	(0.00157)	(0.000970)	(0.00253)	(0.00491)
Average Rating	0.0246***	0.0313***	0.0131***	0.0331***	0.0109***	0.390***
	(0.000399)	(0.000548)	(0.00115)	(0.000811)	(0.00202)	(0.0104)
Log(Reviews + 1)	0.00484***	0.0118***	0.0287***	0.00740***	0.0261***	-0.00546
	(0.000491)	(0.000649)	(0.00149)	(0.000883)	(0.00242)	(0.00688)
Log(Price + 1)	-0.0447***	-0.0543***	-0.0531***	-0.710***	-1.204***	-0.523***
	(0.000462)	(0.000650)	(0.0008)	(0.0183)	(0.0413)	(0.0129)
N	2603635	2603635	2603635	2603635	2603635	2603635
Other controls	No	Yes	Yes	Yes	Yes	Yes
Request FE	No	Yes	Yes	Yes	Yes	Yes
Pro FE	No	No	Yes	No	Yes	No

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Regression results from Equation 2. The first three columns use OLS and progressively add fixed effects to control for request and professional unobservable characteristics. The next two columns instrument for price and include request and then professional fixed effects. The final column instruments for price as well are reputation measures and includes request fixed effects.

been intermediated by digital platforms like the one we study. Overall, the shares suggest that the internet is an important way to find home improvement professionals.[11]

Survey respondents also care more about prices and reputation—online or word-of-mouth—than knowing about whether a professional is licensed. When asked to list up to 3 reasons for why consumers selected a particular professional, respondents' answers included the word stems 'price' (50%), 'cost' (14%),[12] 'quality' (14%), 'review' (13%), 'recommend' (13%), and 'friend' (12%).[13] Fewer than 40 respondents listed licensing in their top 3 reasons for hiring a professional.

Respondents do not seem very knowledgeable about the occupational licensing status of their providers, at least during the search process. Indeed, 61% of the respondents say they knew that their chosen provider was licensed for the service requested, but 52% of those only found out when they signed the contract, and 33% found out from the professional telling them. More people found out about a professional's licensing status on a platform like the one we study (9%) than on an official government website (6%).

It may be the case that the customers who hired non-licensed professionals or did not know about the licensing status of the professional did so because they knew that a license was not required for the service. To examine this possibility we asked respondents whether a license was required by law for the service requested. 51% of the respondents said they were not sure (37%) or that it was not required (14%). This suggests that a large share of customers choose professionals without knowing about the relevant regulations.[14] One reason for this may be that customers simply trust that the existing regulations and their enforcement are enough to guarantee acceptable quality standards. We do find some support for this, with 53% of the respondents in favor of licensing regulation, and 16% against it. In

[11] 15% of the respondents selected the 'Other' category, but then mentioned family and friends, Facebook, neighborhood, and professionals they hired previously as the way in which they found the current professional.

[12] An additional 6% of the responses included the words 'cheap' and 'afford'.

[13] An additional 13% of the responses included 'refer', and an additional 9% of the responses included 'reput'.

[14] We were not able to confirm whether customers' beliefs in answering this question corresponded to reality; i.e. if a customer said that a license was required for the service requested, we could not always verify whether that was indeed the case.

the next section we quantify how these regulations affect aggregate demand, competition, and equilibrium prices.

5 Effect of Occupational Licensing Regulation on Market Outcomes

In this section we study the aggregate effects of licensing stringency on competition, prices, and customer satisfaction. Aggregate effects from licensing—for both supply and demand—may arise even if, as our above results suggest, consumers do not consider the licensing status of individual providers in their choices. On the supply side, licensing laws may affect the decisions of individuals to become professionals in an occupation, and their bidding strategies. Non-licensed professionals may also change their quality due to competition from licensed professionals. Overall, to the extent that licensing regulation increases entry barriers, we should expect a negative effect on competition. It is also possible that consumers do not take into account licensing information when choosing a service provider but rely instead on the existing regulatory framework to prevent unqualified individuals from operating. This seems in line with our survey evidence. If regulation increases consumer trust in service providers, it may also increase aggregate demand for the services covered by occupational licensing.

The advantage of studying how occupational licensing affects market outcomes in our context is twofold. First, the platform facilitates matching across a wide range of service categories and US states. To our knowledge, this level of breadth is unique in the literature on occupational licensing. Second, since the platform tracks requests, quotes, and to some extent matches (i.e. cases where a consumer hires a professional) and consumer evaluation of service quality, we can measure the effect of occupational licensing regulation on multiple stages of the consumer-professional *exchange funnel*: ex-ante search, transaction, and ex-post satisfaction. We also observe many details about the job requests not available in prior studies on occupational licensing.

To evaluate the extent to which licensing regulation affects demand and supply, we exploit variation in the stringency of licensing requirements across states and service categories. We consider a *market* to be defined by a service category and zip code.[15] Within each market, we form a measure of licensing stringency by combining data on occupational licensing regulation from the Institute for Justice with our own manually collected data. The Institute for Justice "License to Work" database contains several dimensions of licensing requirements across all 50 states and the District of Columbia for 102 lower-income occupations.[16] For occupations not covered by this study, such as plumbers and electricians, we collect analogous information online and by phone from state government agencies.

The dimensions of licensing recorded in this data are fees, number of required exams, minimum grade for passing an exam, minimum age required before practicing, education requirements (expressed in years or credit hours), and experience requirements (in years). We reduce these dimensions to a one-dimensional stringency score for each state-occupation pair by taking the first element of a principal component on the full set of requirements. A higher score corresponds to more stringent regulation. We will refer to this score as *licensing stringency*. Table 3 displays the correlation between our measure of licensing stringency and each regulatory dimension included in the principal component analysis. The table shows that our measure of licensing stringency is indeed positively correlated with all dimensions of regulation, but especially with the number of required exams and fees.[17] The first principal component explains 47% of the variation in the dimensions of licensing regulation.

Using our measure of licensing stringency and the data scraped from the platform, Figure H.3 shows that our measure of licensing stringency is positively correlated with the share of professionals with a license on the site—either verified or mentioned in their profile. In this figure, we divide state-occupation pairs into deciles according to the distribution of the licensing stringency variable. The y-axis shows the share of bidders within a

[15] Most occupational licensing regulation is defined at the state level, but additional requirements may be added at the county or city level. For example, the states of Colorado, New York, Texas, and Wyoming do not have state-level licensing requirements for many occupations, but instead allow cities and counties to set their own standards.

[16] `http://ij.org/report/license-work-2/`

[17] "Days Lost" is an estimate of how many days of work a professional loses to satisfy the occupational licensing requirements. It is included in the Institute for Justice database but not in the additional occupations for which we collected data. Adding it or removing it from the analysis does not change our results.

Table 3: Licensing Regulation and Dimensionality Reduction.

Licensing Stringency	Correlation
Fees	0.845
Days Lost	0.853
Exams	0.815
Min Grade	0.290
Min Age	0.746
Education (Years)	0.082
Education (Credits)	0.071
Experience (Years)	0.556

Correlation matrix between the first principal component and the dimensions of occupational licensing regulation used in the principal component analysis. 'Days Lost' is an estimate of how many days of work a professional loses to satisfy the occupational licensing requirements. It is computed by the Institute for Justice, so we do not have it for all occupations.

state-occupation who have a verified license by the end of our sample period. As licensing stringency increases, the share of licensed professionals also increases. The correlation is relatively modest, implying that there are other factors which influence whether professionals choose to inform online consumers about their licensing status.

Before describing our estimation strategy and results, we discuss how our proxy for stringency regulation affects data selection. There are almost 400 home improvement categories defined by the platform, ranging from gutter cleaning and maintenance to pest control. We manually reduce the number of categories to 164, by combining categories for similar services. For example, "solar panel installation", "solar panel repair", and "solar panels" are combined into a single meta-category. We then associate each service *meta-category* to a corresponding *occupation*. For example, "toilet installation and repair" and "shower/bathtub installation and repair" are meta-categories associated with plumbers. Because our estimation relies on cross-state variation in licensing regulation for a single occupation we remove all meta-categories that are not covered by occupational licensing regulation in any state, such as "gardening" or "packing and unpacking". Since a few occupations without state licensing regulation have local regulation (e.g. at a county or city level) which is hard to collect, we remove all state-occupation pairs without any state regulation. We further limit the sample to meta-categories with at least 100 posted tasks in at least 10 states. We trim quotes to be between \$2 and \$2,000 for hourly prices, and between \$20 and \$50,000 for fixed prices. Specifically, we keep all quotes, but we code the quoted prices outside of these intervals as missing. We are left with 203,688 markets defined the zip code and the meta-category. At the state-occupation level, we have 396 groups, covering 45 states and 20 separate occupations.

To illustrate this stringency measures, we highlight some examples. Three occupations in Oregon have a licensing stringency measures equal to the average value of 0.14—mason, painting, and sheet metal contractors—all of which require professionals to be at least 18 years old, pay \$385 in fees, attend 16 clock hours of class instruction,[18] and pass 1 exam. One standard deviation above the mean of the stringency measure yields a level of regulation corresponding to electricians in Connecticut. They have to be at least 18 years old, pay \$705, pass 3 exams, and have two years of experience. Subtracting one standard deviation means reducing the level of regulation to the laws covering cement finishing or painting contractors in Massachusetts, who only need to pay \$250 to be able to work.

Table 4 shows task-level descriptive statistics for the market equilibrium variables at the search, matching, and post-transaction phase. They include the number of quotes received by each task and the average quoted price for those tasks with fixed price bids, the match rate and the transaction price, the probability that the buyer gives the provider a 5-star review after matching, and the buyer's probability to post again on the platform. We observe a large degree of heterogeneity across service categories. With an average licensing stringency variable equal to 0.42, tasks on the platform tend to be posted in states and occupations with more stringent requirements than the average state-occupation level (0.14).

We let z denote a zip code, c denote a meta-category, and t denote a month-year. Licensing stringency is defined at the state-occupation level, which is coarser than our market definition. For example, services within the "air conditioning" category in zip code 02139 and services within the "duct/vent" category in zip code 02138 are subject to the

[18] 37.5 clock hours are equivalent to one semester credit, so 16 clock hours are equivalent to o.43 semester credit (`http://www.acics.org/news/content.aspx?id=4419`).

Table 4: Descriptive Statistics on Licensing Stringency and Equilibrium Outcomes.

Variable	Observations	Mean	Standard Deviation	10th Pctl.	Median	90th Pctl.
Licensing Stringency	1,035,717	0.42	1.78	-1.82	0.44	2.42
Nr. Quotes	1,035,717	2.01	1.55	0	2	5
Avg. Fixed Quote ($)	414,511	723.5	1,956	67.5	199	1,750
Match Probability	848,947	0.16	0.37	0	0	1
Fixed Sale Price ($)	64,818	324.7	949.8	55	125	600
5-Star Review	140,571	0.48	0.5	0	0	1
Request Again	140,571	0.24	0.42	0	0	1
Request Again Diff. Cat.	140,571	0.23	0.42	0	0	1

Row 1 and 2 include all tasks submitted in categories and states with some level of occupational licensing regulation. The following rows focus on a subset of these observations. Row 3 restricts attention to tasks with at least one fixed price quote. Row 4 focuses on any task with at least one offer. Row 5 focuses on the successful tasks whose winning bid includes a fixed price quote. Row 6 and 7 focus on all successful tasks.

same regulation for HVAC contractors in Massachusetts.

We first run regressions to evaluate the effect of licensing stringency on aggregate *demand*. Here we test the possibility that, if regulation increases consumer trust in service providers by raising service quality, it may increase demand for the services provided by professionals covered by more stringent licensing regulation. The equation we estimate is:

$$\begin{split} log(posted\_requests_{occupation(c)zt}) = \alpha licensing\_stringency_{state(z)occupation(c)} + \\ \mu_z + \mu_{occupation(c)} + \mu_t + \epsilon_{occupation(c)zt}. \end{split} \tag{3}$$

Results are presented in Table 5. The estimated effect is a relatively precise zero, suggesting that consumers do not post more requests on the platform for services that are covered by more stringent licensing regulation. This finding that aggregate demand on the platform does not appear to change with licensing stringency suggests that we can appropriately examine request-level outcomes as we study effects on supply-side market outcomes (such as competition and prices) below.

To study effects on supply-side factors, we run regressions of the following form:

$$y_{rctz} = \mu_z + \mu_c + \mu_t + \beta licensing\_stringency_{state(z)occupation(c)} + \beta X_{rctz} + \epsilon_{rctz}, \tag{4}$$

In this regression, X_{rsctz} includes controls for how the customer was acquired (e.g. organic

Table 5: Cross-State Regression Estimates—Aggregate Demand

	Dependent variable:			
	Number of Posted Requests (log)			
	(1)	(2)	(3)	(4)
Licensing Stringency	0.004	0.013***	0.002	0.002
	(0.004)	(0.002)	(0.003)	(0.003)
Mean of Y:	0.162	0.162	0.162	0.162
Month-Year FE	No	No	No	Yes
Zip Code FE	No	No	Yes	Yes
Occupation FE	No	Yes	Yes	Yes
Observations	3,405,204	3,405,204	3,405,204	3,405,204
R^2	0.0003	0.032	0.213	0.288
Adjusted R^2	0.0003	0.032	0.208	0.284
Residual Std. Error	0.381 (df = 3405202)	0.375 (df = 3405183)	0.339 (df = 3385293)	0.322 (df = 3385261)

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors clustered on occupation-state reported in parentheses

Regression results for aggregate demand (Equation 3). An observation is a month-year-zipcode-occupation, and the outcome of interest is the number of posted requests. We augment the data to include all observations with no posted requests. columns 1 through 4 increasingly add controls. Controls occupation, zip code, and month-year fixed effects.

search or online advertising) and request characteristics (e.g. text length of the request). The variable y_{rctz} is one of many possible outcome measures: at the *bidding* stage, our outcome variables include the number of quotes received and the average (log) quoted price for quotes with a fixed price bid; at the *matching* stage, we use a dummy for whether a match (i.e. a hire) was confirmed on the platform and the (log) transacted price for matches where the winning quote had a fixed price bid; at the *post-transaction* stage, we use a dummy for whether the consumer leaves a five star review and a dummy for whether the consumer posts another request at least one week after the current request.[19] Nosko

[19]The one-week delay is to avoid confounding buyer's choice to post again on the platform with buyer's decision to re-post an identical request. We can also restrict attention to posting again, but in a different

and Tadelis (2015) showed that consumers draw conclusions about the quality of a platform from individual transactions; we take the propensity to post again on the platform as a signal of consumer satisfaction about the service received by the hired professional.

Baseline regression results are in Table 6. On average across all categories, increases in occupational licensing stringency are associated with increases in quoted and transaction prices. The coefficient estimates imply that a one standard deviation increase in licensing stringency (1.78) decreases the number of quotes by 0.05, or 2.4%, and increases quoted prices by 3.2% and transacted prices by 2.5%. Licensing stringency does not significantly affect the match probability. More stringent licensing is also not associated with higher customer satisfaction, as measured by ratings or customer returns. If anything the coefficients are negative, although the point estimates are not economically significant.

The above analysis does not rule out possible compositional differences in the nature of tasks requested across states and occupations. For example, it might be the case that painting jobs in Arizona are for much bigger houses than in Massachusetts, and so some of the price differences that we capture with our licensing stringency measure might in fact be a function of different task requests. To control for this possibility, we make use of the large set of questions that customers answer before posting a job, and flexibly control for their answers using the double machine learning estimator developed by Chernozhukov et al. (2018). This estimator predicts both our licensing stringency variable and our outcome variables as a function of all our observables. For this prediction, we use LASSO regressions, and set the penalty parameter using 10-fold cross validation. For this prediction we split the data in two groups, train the model on one group to predict on the other. Then we use the predictions to regress the residual of our outcome variables on the residual of our licensing stringency variable. We do this 100 times ('splits'), and use the distribution of the resulting coefficients to get at our final estimate and standard errors.

The results displayed in Table 7 show the median estimated coefficients, and confirm the main conclusions drawn from Table 6. Furthermore, because these regressions use additional information from requests, they result in lower standard errors. The estimates from the double-ML procedure are nearly identical to our estimates from the fixed effects

service category. Results do not change.

regressions. Even with the additional precision, we are not able to detect a statistically significant effect on measures of customer satisfaction.[20]

[20] Appendix G contains an exploratory analysis of licensing stringency by occupation.

Table 6: Cross-State Regression Estimates—Matching Outcomes

	Nr. Quotes	Avg FP Quote (log)	Match	Matched Quote (log)	5-Star Review	Post Again	Post Again Diff. Cat.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Licensing Stringency	-0.027^{**}	0.018^{***}	-0.001	0.014^{**}	0.001	-0.003^{**}	-0.003^{**}
	(0.014)	(0.007)	(0.001)	(0.006)	(0.001)	(0.001)	(0.001)
Included Tasks	All	With FP Quotes	With Quotes	Matched to FP Quote	Matched	Matched	Matched
Mean of Y:	2.01	5.5	0.16	5.02	0.48	0.24	0.23
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zip Code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sub-Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,035,717	414,511	848,947	64,818	140,571	140,571	140,571
R^2	0.507	0.522	0.073	0.575	0.105	0.129	0.129

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Regression results of equation 4, where we estimate one α equal across all categories. For category-specific estimates, see Figures G.1 through G.4. Column (1) includes all tasks submitted in categories and states with some level of occupational licensing regulation. The following columns focus on a subset of these observations. Column (2) restricts attention to tasks with at least one fixed price quote. Column (3) focuses on any task with at least one offer. Column (4) focuses on the successful tasks whose winning bid includes a fixed price quote. Column (5) and (6) focus on all successful tasks. Standard errors are clustered at the state-occupation level.

Table 7: Cross-State Regression Estimates - Double Machine Learning

	Nr. Quotes	Avg Quote Price (log)	Match	Transaction Price (log)	5-Star Review	Post Again	Post Again Diff. Cat.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Licensing Stringency	-0.0245***	0.0172***	-0.0012***	0.0142***	0.0000	-0.0020**	-0.0019*
SE (median)	[0.0011]	[0.0012]	[0.0004]	[0.0025]	[0.0012]	[0.0010]	[0.0010]
SE (median method)	(0.0011)	(0.0012)	(0.0004)	(0.0025)	(0.0012)	(0.0010)	(0.0010)
Included Tasks	All	With FP Quotes	With Quotes	Matched to FP Quote	Matched	Matched	Matched
Mean of Y	2.01	5.50	0.16	5.02	0.48	0.24	0.23
Month-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zip Code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sub-Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,035,717	414,511	848,947	64,818	140,571	140,571	140,571
R^2	0.0005	0.0004	0.0000	0.0005	0.0000	0.0000	0.0000
Adjusted R^2	0.0005	0.0004	0.0000	0.0005	-0.0000	0.0000	0.0000
Lasso Y on X Nonzero FE (%)	52.53	11.26	5.02	5.69	4.80	1.65	1.61
Lasso Y on X Nonzero Non-FE (%)	2.28	2.30	0.99	2.58	0.95	0.29	0.30
Lasso D on X Nonzero FE (%)	24.14	19.93	24.09	5.86	10.05	10.05	10.05
Lasso D on X Nonzero Non-FE (%)	1.65	1.83	1.82	1.36	1.43	1.43	1.43

Double Machine Learning estimates of equation 4 (Chernozhukov et al. (2018)), where we use lasso to predict both treatment and outcome variable as a function of our explanatory variables. Our explanatory variables include features constructed from geography, time, category, and task request details. To run these estimates, we use 100 splits, with each split having 2 cross-fits. To set the lasso penalty parameter we use 10-fold cross validation. Point estimates are based on the median across all splits. Standard errors are clustered by state abbreviation-occupation groups. Standard errors reported in brackets are based on the median across all splits. Standard errors reported in parentheses are based on the median method. Significance determined using point estimates and standard errors based on the median across all splits: *$p<0.1$; **$p<0.05$; ***$p<0.01$. *Mean of outcome variable,* R^2, *adjusted* R^2, *and nonzero FE/Non-FE Lasso coefficient measures are all based on the median across all splits. Machine learning method employed is the Lasso regression. Initial optimal regularization parameter lambda is selected based on minimum mean squared-error calculated from conducting 10-fold cross-validations done over a grid ranging from* 2^{-30} *to* 2^{30} *in increments of base 2. After the initial coarse search, an additional search is done in the neighborhood of the initial optimal lambda by defining a grid equal to the sequence of numbers from 0.6 to 1.9 in increments of 0.1 multiplied by the initial optimal lambda.*

Before concluding this section, we explore heterogeneity of the effects of licensing regulation for different degrees of job complexity. Some states only regulate professionals if they perform jobs above a certain price threshold,[21] and a natural dimension along which to measure heterogeneity in how licensing stringency affect outcomes is the expected price of a job. We construct a proxy for the expected price of a given request by using a machine learning approach to predict whether the average quote submitted would be above a certain dollar threshold. Specifically, we use only job requests that have at least one fixed price quote and we split the observations into 5 groups. For each group we train a model to predict the average quoted price on the remaining 80% of the sample, and use this prediction on the focal group. We also use the entire sample of requests with at least one fixed quote to train a model that predicts the average quoted price for jobs with no bids or no fixed price quotes. We do this separately for price thresholds at $200, $500, and $1,000. For this prediction we use only request-level features such as the answers that the customer submits to the platform questionnaire when posting their request. Appendix Table H.3 demonstrates that our predictive accuracy for this measure is very high.

We now use the predicted price for each job to see whether the effect of regulation on competition and prices is driven by larger and more complex jobs—defined as those jobs whose predicted price is above a given dollar amount.[22] Table 8 presents estimates of Equation 4 with licensing stringency interacted with a dummy variable for whether the job has a predicted price that is higher than a given threshold ($200 for the top panel, $500 for the middle panel, and $1,000 for the bottom panel). The reduction in number of quotes seem similar across low- and high-priced jobs, but the price increase is mostly driven by the more complex jobs. Looking at column (4), we see that the interaction coefficient increases (and stays significant) as the price threshold increases. A one standard deviation increase in licensing stringency raises the price of jobs above $200 by 7.2%, it raises the price of jobs above $500 by 12.3%, and it raises the price of jobs above $1,000 by 17.3%. Thus, increases in licensing stringency are associated with higher prices *especially* for expensive tasks.

[21] As stated earlier, in California, painters are required to have an occupational license only if they perform jobs above $500.

[22] As in the overall dataset, we do not see any effect of regulation stringency on aggregate demand for jobs whose price is higher than $200, $500, or $1,000.

Table 8: Heterogeneity by Price Point

	Nr. Quotes	Avg FP Quote (log)	Match	Matched Quote (log)	5-Star Review	Post Again	Post Again Diff. Cat.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Licensing Stringency	−0.020	0.004	0.001	0.003	0.0001	−0.003*	−0.003**
	(0.015)	(0.009)	(0.002)	(0.007)	(0.002)	(0.001)	(0.001)
Licensing Stringency*≥ $200	−0.014	0.035**	−0.004**	0.041***	0.002	0.0003	0.001
	(0.016)	(0.015)	(0.002)	(0.013)	(0.002)	(0.002)	(0.002)
R^2	0.507	0.522	0.073	0.576	0.105	0.129	0.129
Licensing Stringency	−0.030**	0.011	−0.001	0.006	0.001	−0.003**	−0.003**
	(0.014)	(0.007)	(0.001)	(0.006)	(0.002)	(0.001)	(0.001)
Licensing Stringency*≥ $500	0.014	0.039*	−0.002	0.069***	0.001	0.001	0.002
	(0.020)	(0.021)	(0.002)	(0.016)	(0.003)	(0.002)	(0.002)
R^2	0.507	0.522	0.073	0.576	0.105	0.129	0.129
Licensing Stringency	−0.029**	0.013*	−0.001	0.009	0.0004	−0.003**	−0.003**
	(0.014)	(0.007)	(0.001)	(0.006)	(0.002)	(0.001)	(0.001)
Licensing Stringency*≥ $1,000	0.012	0.045	−0.002	0.097***	0.004	0.001	0.001
	(0.025)	(0.032)	(0.002)	(0.028)	(0.004)	(0.003)	(0.003)
R^2	0.507	0.522	0.073	0.576	0.105	0.129	0.129
Included Tasks	All	With FP Quotes	With Quotes	Matched to FP Quote	Matched	Matched	Matched
Observations	1,035,717	414,511	848,947	64,818	140,571	140,571	140,571

Note: *p<0.1; **p<0.05; ***p<0.01

Three sets of regressions where the licensing stringency variable is interacted with a dummy variable for whether the predicted job price is above $200 (top panel), $500 (middle panel), or $1,000 (bottom panel). Zip code, month-year, and sub-category fixed effects are included as controls.

6 Conclusion

In this paper we took advantage of new data collected by a digital platform to take a fresh look at the role of occupational licensing laws on individual choices and market outcomes. First, we examine whether information about the professional's licensing status has an effect on consumer choice of whom to hire. We find causal evidence that consumers care about online reviews more than about occupational licensing signals available on the platform. The result could be consistent with a world where consumers believe that a professional complies with relevant state regulation when bidding for a particular service, either by being licensed or by providing services of comparable quality to licensed professionals.

To the extent that consumers do rely on occupational licensing regulation to screen out low quality providers, policy makers would want licensing regulations to improve service quality. Our cross-state analysis exploits the large heterogeneity in licensing requirements across home improvement categories and across US states to estimate the effect of licensing stringency on market equilibrium. We find that licensing stringency is associated with fewer quotes and higher transaction prices but not better service, at least as measured through online reviews and propensity to use the platform again.

The paper has a number of limitations. In particular, services differ in the degree to which consumers have visibility into the dimensions of quality relevant for their safety. We proxy customer satisfaction with the propensity to use the platform again, and to positively rate the service providers. Both are likely to be noisy measures of provider quality. In addition, some consumer safety or professional quality issues may take time to manifest themselves. For example, a consumer might not realize that a roofing contractor ignores basic safety measures when repairing a roof until much later—and potentially when it is too late to submit a review.

Another limitation of the study is that we address the population of primarily residential consumers who purchase online. If online consumers are less sensitive to licensing credentials, and service providers sort between online and offline customers accordingly, the effects measured in this paper do not immediately extend to offline transactions. For similar reasons, we cannot say anything about the importance of licensing for commercial tasks relating to construction and home improvement.

Both regulators and platforms have an interest in protecting consumers and ensuring service quality. Our results have implications for the design of licensing regulation and of digital platforms for services. In particular, the availability of alternative signals of quality, such as online reviews, has probably reduced the level of regulatory stringency needed to ensure the same service quality as in a world without online reviews. For the platform, the lack of attention paid to verified licensing status by consumers suggests that the disclosure of status is not important. The broader ramifications of our findings for platform design hinge on whether the lack of consumer attention to licensing is caused by misinformed consumers, by the redundancy of licensing with reputational signals, or by the inability of licensing to ensure quality. Measuring the importance of these explanations requires more targeted research designs. We leave these questions for future investigation.

References

Akerlof, G. A. (1970): "The Market for" Lemons": Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 488–500.

Barrios, J. M. (2019): "Occupational Licensing and Accountant Quality: Evidence from the 150-Hour Rule," Becker Friedman Institute for Research in Economics Working Paper 2018-32.

Blair, P. Q. and B. W. Chung (2018): "Job Market Signaling through Occupational Licensing," NBER Working Paper 24791.

Buonanno, P. and M. Pagliero (2019): "Occupational Licensing, Labor Mobility, and the Unfairness of Entry Standards," Working Paper, Colegio Carlo Alberto.

Carroll, S. and R. Gaston (1981): "Occupational Restrictions and the Quality of Service Received: Some Evidence," *Southern Economic Journal*, 959–976.

Chen, D. L. and J. J. Horton (2016): "Research noteAre online labor markets spot markets for tasks? A field experiment on the behavioral response to wage cuts," *Information Systems Research*, 27, 403–423.

Chen, Y. (2018): "User-Generated Physician Ratings: Evidence from Yelp," .

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018): "Double/debiased machine learning for treatment and structural parameters," *The Econometrics Journal*, 21, C1–C68.

Chevalier, J. A. and F. Scott Morton (2008): "State Casket Sales Restrictions: A Pointless Undertaking?" *The Journal of Law and Economics*, 51, 1–23.

Cohen, P. (2016): "Moving to Arizona Soon? You Might Need a License," *New York Times*.

Cullen, Z. and C. Farronato (2015): "Outsourcing Tasks Online: Matching Supply and Demand on Peer-to-Peer Internet Platforms," Working Paper, Harvard University.

DePasquale, C. and K. Stange (2016): "Labor Supply Effects of Occupational Regulation: Evidence from the Nurse Licensure Compact," NBER Working Paper 22344.

DOT (2015): "Occupational Licensing: A Framework for Policymakers," Prepared by the Department of the Treasury Office of Economic Policy, the Council of Economic Advisers, and the Department of Labor. Available at www.whitehouse.gov.

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

Friedman, M. (1962): *Capitalism and Freedom*, University of Chicago Press Chicago.

Harrington, D. E. and K. J. Krynski (2002): "The Effect of State Funeral Regulations on Cremation Rates: Testing for Demand Inducement in Funeral Markets," *The Journal of Law and Economics*, 45, 199–225.

Horton, J. J. (2010): "Online Labor Markets," in *International Workshop on Internet and Network Economics*, Springer, 515–522.

Johnson, J. E. and M. M. Kleiner (2017): "Is Occupational Licensing a Barrier to Interstate Migration?" NBER Working Paper 24107.

Kleiner, M. and E. Soltas (2019): "A Welfare Analysis of Occupational Licensing in U.S. States," Working Paper, MIT.

Kleiner, M. and M. Xu (2019): "Occupational Licensing and Labor Market Fluidity," Working Paper, Queens University.

Kleiner, M. M. and A. B. Krueger (2010): "The Prevalence and Effects of Occupational Licensing," *British Journal of Industrial Relations*, 48, 676–687.

Kleiner, M. M. and R. T. Kudrle (2000): "Does Regulation Affect Economic Outcomes? The Case of Dentistry," *Journal of Law and Economics*, 43, pp. 547–582.

Krasnokutskaya, E., K. Song, and X. Tang (2018): "The Role of Quality in Internet Service Markets," Working Paper, Rice. Forthcoming *Journal of Political Economy.*

Kugler, A. and R. M. Sauer (2005): "Doctors without Borders? Relicensing Requirements and Negative Selection in the Market for Physicians," *Journal of Labor Economics*, 23, 437–465.

Larsen, B. (2015): "Occupational Licensing and Quality: Distributional and Heterogeneous Effects in the Teaching Profession," Working Paper, Stanford University.

Law, M. T. and M. S. Marks (2009): "Effects of Occupational Licensing Laws on Minorities: Evidence from the Progressive Era," *Journal of Law and Economics*, 52, pp. 351–366.

Leland, H. (1979): "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards," *Journal of Political Economy*, 1328–1346.

Luca, M. (2016): "Reviews, Reputation, and Revenue: The Case of Yelp.Com," Working Paper, Harvard University.

Maurizi, A. (1980): "The Impact of Regulation on Quality: The Case of California Contractors," in *Occupational Licensure and Regulation*, ed. by S. Rottenberg, Washington, DC: American Enterprise Institute for Public Policy Research, 299–333.

Millsap, A. (2017): "Some Progress On Occupational Licensing But Much More Needed," *Forbes*.

Nosko, C. and S. Tadelis (2015): "The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment," NBER Working Paper 20830.

Pallais, A. (2014): "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review*, 104, 3565–3599.

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

Sibilia, N. (2017): "South Dakota Just Legalized Braiding Hair Without A License," *Forbes*.

Smith, A. (1776): *An Inquiry into the Nature and Causes of the Wealth of Nations*.

Zumbrun, J. (2016): "Occupational Licenses May Be Bad for the Economy, But Good for Workers Who Have Them," *Wall Street Journal*.

APPENDIX

A Additional Data and Analysis from Crawling Platform

In order to compile a sample of web pages to crawl, in 2018 we found the largest three cities for each state in terms of unique professionals who are potentially licensed, and joined that list with the top 100 cities in terms of site activity. We excluded cities with fewer than 10 professionals in the city. For each category and city, we found the corresponding landing page for the platform. We then obtained information about all professionals displayed on the landing page and their reviews. This information included the professional license status, ranking, name, number of hires, years in business, whether they passed the platform's background checks without any negative information, photos, zip code, city, and an indicator of high engagement with the platform (similar to the "Superhost" badge on Airbnb). We also obtained the text that the professional added to their profile, and their answers to commonly asked questions. Lastly, for each provider, we obtained all review text, dates, and ratings.

In total, the crawl found 79,111 professionals whose profiles were displayed on at least one of the URLs corresponding to the landing page for an occupation in a given city. Table A.1 displays summary statistics for professionals in the sample. The median professional in the sample had no hires, and 1 off-platform review. Conditional on being in the top 5 results for at least one URL, the median professional had 19 hires, 14 reviews, of which 12 were on-platform reviews, and a median average rating of 4.9. 10% of professionals mentioned a license in their profile and 6% had a verified a license. Overall, 14% of professionals mentioned an occupational license in their profile, had it verified by the platform, or both.[23] Many professionals who mention a license in their online profile do not have it verified by the platform. This could be due to the professional intentionally not submitting their license for verification, due to the fact that some licenses are issued at a local level and the platform only verifies state-issued licenses, or due to some licenses being submitted but not

[23]Note that differences in the rates of verification between the crawl and platform sample can occur for many reasons including the fact that professionals differ in their propensity to bid and that the crawl was conducted during a different time period from the platform sample.

yet verified.[24] Professionals also mentioned certifications (7% of the time) and insurance (12% of the time).

Table A.1: Summary Statistics (Across Professionals)

Variable	min	q25	median	q75	max	mean	sd
License Text	0.00	0.00	0.00	0.00	1.00	0.11	0.31
License Verified	0.00	0.00	0.00	0.00	1.00	0.06	0.24
Either License	0.00	0.00	0.00	0.00	1.00	0.15	0.35
Certification Text	0.00	0.00	0.00	0.00	1.00	0.07	0.25
Insurance Text	0.00	0.00	0.00	0.00	1.00	0.12	0.32
Background Check	0.00	0.00	0.00	0.00	1.00	0.17	0.37
Avg. Rating	0.00	0.00	3.00	4.90	5.00	2.42	2.39
Num. Reviews	0.00	0.00	1.00	9.00	1327.00	10.77	31.75
Total Hires	0.00	0.00	0.00	9.00	2912.00	15.94	56.22

Table H.1 and Table H.2 display breakdowns of these statistics for the top 20 categories in terms of number of professionals and in terms of the share of licensed professionals. 18% of professionals in the top category, 'General Contracting', mention a license in their online profile, and 12% have a verified license. Categories that are more technical such as plumbing, home inspection, electrical wiring, and pest extermination top the list of the categories with the highest share of professionals with any licensing information. However, even in these categories fewer than 50% of professionals disclose any credential and fewer than 28% mention a license.

In order to evaluate the similarity of the quality signals from licensing information and online ratings, we use the data scraped from the website (see details in Appendix A). Table A.2 displays regressions of review related variables on a professional's status of licensing verification and background checks. In particular, we use two dummy variables, the first equal to 1 if the license was verified by the platform, and the second equal to 1 if the professional passed the platform's background checks without any negative information.[25] In the regressions we weigh each observation by one over the number of times a particular

[24] In a manual investigation using websites of state licensing boards, we found it difficult to verify the validity of licenses of professionals who mentioned them in their profile. This could happen because the registered name of the professional differed from the name on the platform, because the license had expired, or because the professional held a different type of license than the one we were searching for.

[25] Background checks are voluntary. If the professional chooses to undergo them, these background checks

professional appears in the search results of our scrape. This is because an observation is a professional profile appearing in a given category and zip code, and a particular professional can show up on the search results corresponding to multiple categories and zip codes.

Columns (1) and (2) display the results for average ratings. Professionals who passed background checks and have their license verified tend to also have higher average ratings. However, when the outcome variable is the average on-platform ratings[26] of the pro, the relationships with license verification and with successful background checks reverse. Lastly, professionals who passed the platform's background checks and had their license verified also have more hires. These regressions suggest that professionals who are committed to heavily rely on the platform to find customers accumulate positive off-platform ratings, submit their license for platform verification, and indeed find more jobs on the platform.

B Analysis of California General Contractor $500 Job Threshold

One reason why professionals may not submit proof of their license for platform verification is that they are bidding on just the projects for which a license is not required. To test for this, we use the data that the platform shared with us. We focus on general contractors in California, who are allowed to perform general contractor jobs without a license as long as those jobs are below $500. Figure H.2 displays the distribution of bids separately for professionals who had the platform verify their license, and for professionals who did not. The majority of bids for both types of professionals is below $500. However, both platform-verified verified and never-verified professionals also bid above the $500 threshold. This is consistent either with those professionals having a license that is not observable to us, or those professionals skirting some occupational licensing laws. Given our data, we cannot distinguish between these two alternatives.

are performed by a separate firm hired by the platform and the results are then evaluated by the platform.

[26] As we described at the beginning of this section, on-platform ratings are those associated with a hire made on the platform, whereas off-platform ratings can be submitted by anyone. The distribution for off-platform ratings is more positive than for on-platform ratings (Appendix Figure H.1). This is consistent with some professionals selectively asking previous customers to leave reviews.

Table A.2: Cross-Sectional Regression of Reviews Ratings on Licensing Status

	Dependent variable:					
	Avg. Rating		Avg. On-Platform Rating		Log(Hires + 1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Background Check	0.058***	0.049**	−0.015***	−0.010	0.380***	0.350***
	(0.008)	(0.022)	(0.003)	(0.009)	(0.017)	(0.047)
Lic. Verified	0.068***	0.071*	−0.017***	−0.032**	0.270***	0.250***
	(0.013)	(0.041)	(0.005)	(0.016)	(0.027)	(0.088)
Lic. Text	−0.008	0.026	−0.005	0.002	−0.100***	−0.048
	(0.011)	(0.037)	(0.004)	(0.014)	(0.024)	(0.076)
Category and Zip Code FE	Yes	No	Yes	No	Yes	No
Category by State Code FE	No	Yes	No	Yes	No	Yes
First Review Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Text Length Bins	Yes	Yes	Yes	Yes	Yes	Yes
Observations	131,541	131,541	131,541	131,541	131,541	131,541
R^2	0.330	0.920	0.340	0.920	0.390	0.920

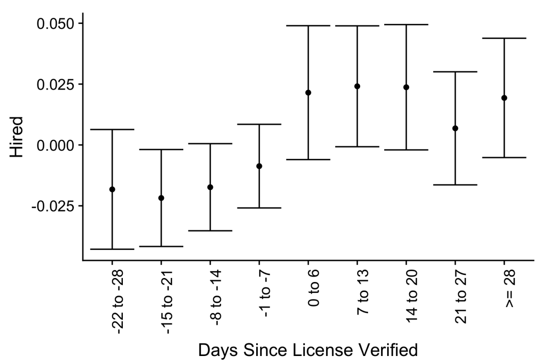
Note: *p<0.1; **p<0.05; ***p<0.01

The regression results show correlations between professionals' outcomes on the platform three variables: whether a processional uploaded a license which was verified, whether the professional volunteered to participate in a background check which they passed, and whether they mentioned a license in their profile text. An observation is a professional who is shown on the platform landing page when a consumer searches for a specific occupation in a given city.

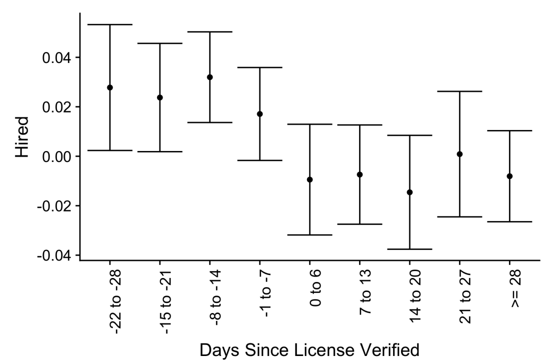
C Additional Event Studies Related to License Verification

In this section we discuss additional results regarding the event study design for license verification. We first investigate the possibility of heterogeneous treatment effects by whether the professional had a hire at the time of license verification. Professionals with a hire may have found other ways to signal quality reducing the need for the licensing signal or they presence of a prior hire may serve a substitute for licensing information. Figure C.1 displays the event study results where the time since license validation is interacted with whether the professional had a hire prior to the time of the bid. Although the results are imprecise, the point estimates suggest a positive effect of licensing when there's no hire and a negative effect when there is a hire.

Figure C.1: Licensing Event Study - Heterogeneity



(a) Licensing Effect for Pros Without a Prior Hire



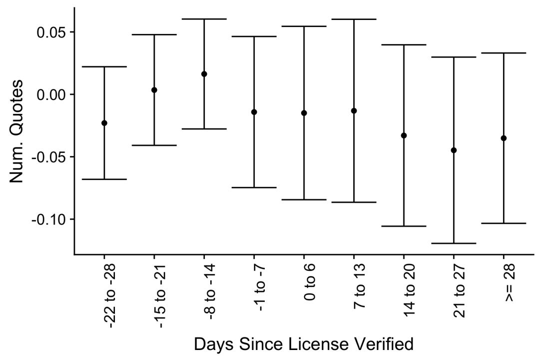
(b) Licensing Effect for Pros With a Prior Hire

One reason that there may not find be effects of licensing on hiring is that professionals adjust their bidding behavior around the time of the license validation. We've already shown that there is no evidence of this for the price that professionals bid. Below, we consider other margins of adjustment using the specification in Equation 1.

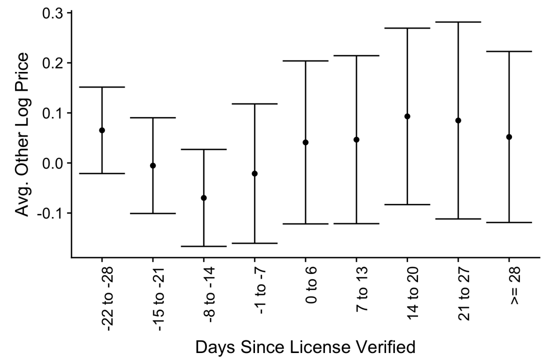
Figure C.2a displays the number of quotes received by the requests that the professional bid on and C.2b displays the average log prices of competitors faced by the professional. Both of these outcomes, which relate to the types of requests professionals bid on don't vary with verified license status. Figure C.2c displays an event study where the outcome is the order (relative to other bidders) in which a professionals bid arrived for a given request. There is no detectable effect of license validation status in the speed at which professionals bid on a request.

Lastly, we consider the quantity of bids submitted by professionals. Figure C.2d displays the number of bids sent by a professional in the weeks around license validation when including a control for whether the license was uploaded to the platform. We find that professionals decrease platform participation by .6 bids relative to a mean of 3.7 bids around the time of license validation. When we exclude the control for license validation (Figure C.2e) we find that bidding temporarily increases prior to license validation, when the license was submitted and decreases afterwards. This is consistent with professionals increasing using of the platform at the same time temporarily on several and then reverting back to a baseline.

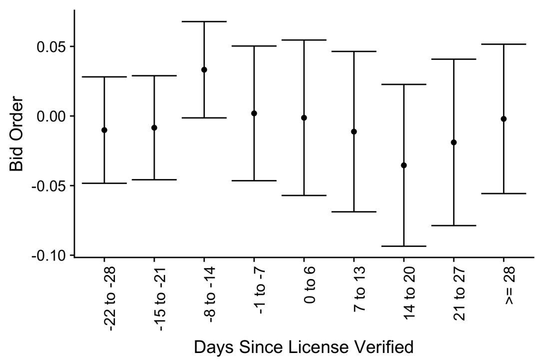
Figure C.2: Licensing Event Study - Supply Side Responses



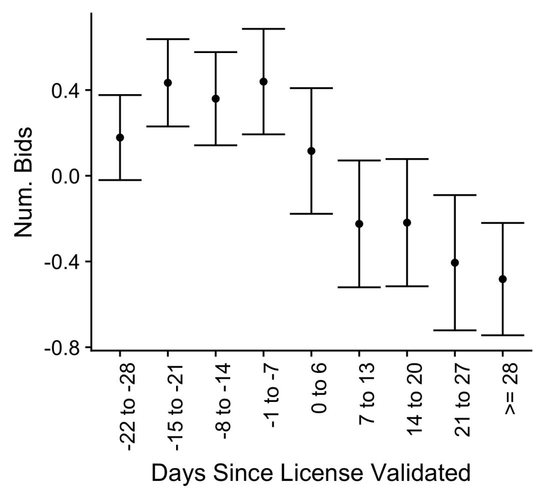
(a) Number of Other Bids on Request



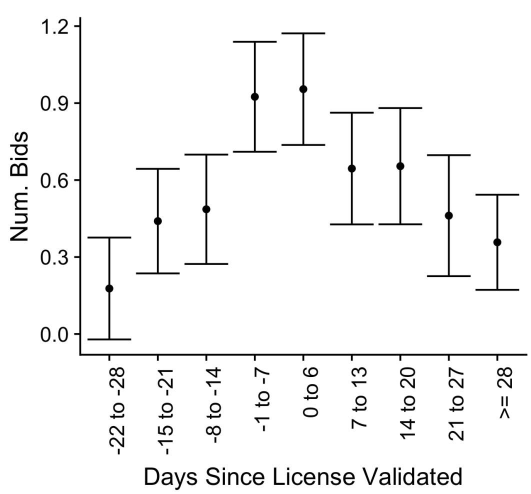
(b) Average Log Prices of Other Bidders on Request



(c) Order of Bid Timing on a Request



(d) Number of Bids by Professional



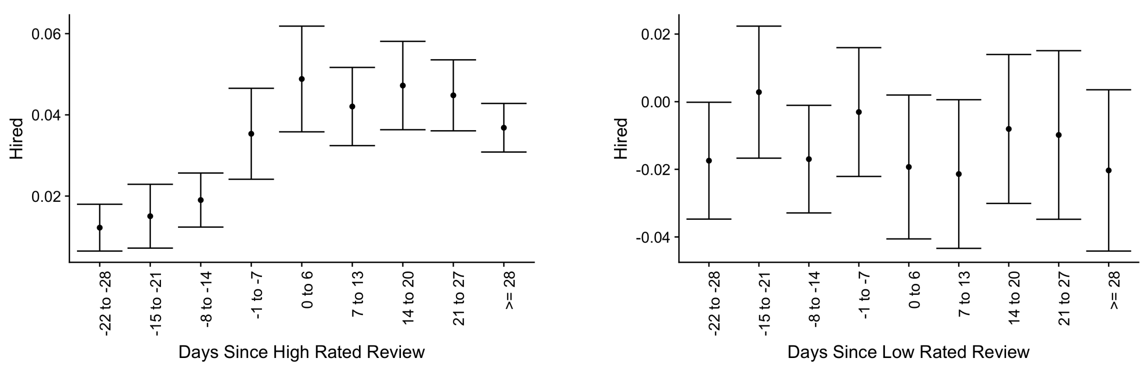
(e) Number of Bids by Professional - No Control For License Submission

D Additional Event Studies Related to First Reviews

In this section we discuss additional results regarding the event study design for the first review. We first investigate the possibility of heterogeneous treatment effects by whether the review had a high rating and by whether the review was on-platform. As a prior, we expect that the positive effect of first reviews on hiring of a first review comes from first reviews associated with high ratings. Furthermore, we would expect on-platform reviews to be more credible to consumers than off-platform reviews, and thus to have larger effects. Professionals with a hire may have found other ways to signal quality reducing the need for the licensing signal or they presence of a prior hire may serve a substitute for licensing information.

Figure D.1 displays the event study results for high and low rated first reviews. We find a large positive effect for high rated reviews and no effect on hiring rates for low rated reviews. We hypothesize that the lack of negative effect of low rated reviews is due to the fact that the baseline hiring rate of pros without reviews is already close to 0. Figure D.2 displays a similar contrast for on-platform reviews. There is a bigger and sharper jump in hiring rates for on-platform reviews, although the differences across the two review types are not statistically significant.

Figure D.1: First Review Event Study - Ratings

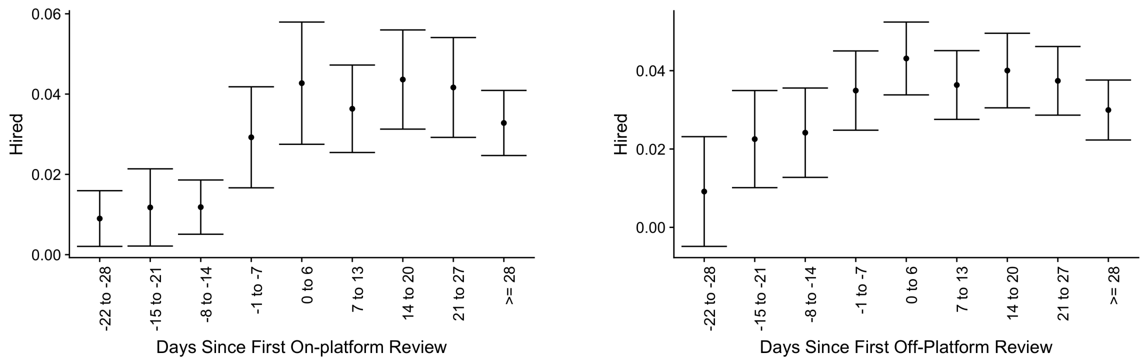


(a) Effect of First Reviews With Rating > 3

(b) Effect of First Reviews With Rating < 4

We now investigate whether the positive effect of a first review is driven by supply or demand side responses. We've already shown that there is no evidence of this for the

Figure D.2: First Review Event Study - On-platform vs Off-platform



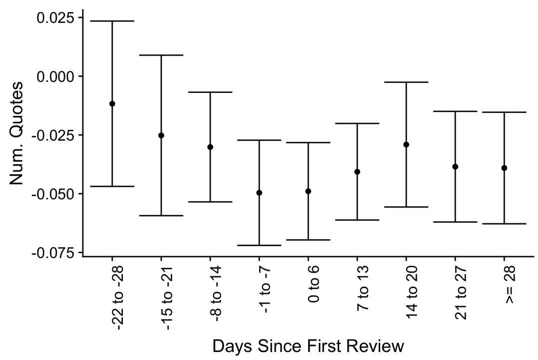
(a) Effect of On-platform First Reviews (b) Effect of Off-platform First Reviews

price that professionals bid. Below, we consider other margins of adjustment using the specification in Equation 1.

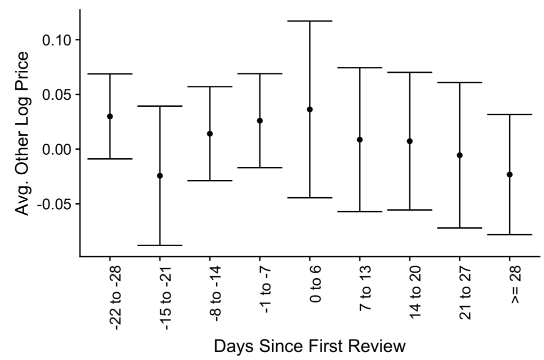
Figure D.3a displays the number of quotes received by the requests that the professional bid on and D.3b displays the average log prices of competitors faced by the professional. Both of these outcomes, which relate to the types of requests professionals bid on don't vary with verified license status. Figure D.3c displays an event study where the outcome is the order (relative to other bidders) in which a professionals bid arrived for a given request. There is no detectable effect of license validation status in the speed at which professionals bid on a request immediately after a first review.

Lastly, we consider the quantity of bids submitted by professionals. Figure D.3d displays the number of bids sent by a professional in the weeks around license validation when including a control for whether the license was uploaded to the platform. We find that professionals greatly increase bidding activity after obtaining a first review. This is, in principle, not a problem for our interpretation of the review effect on hiring being due to consumer demand. The reason is that although professionals increase their bidding frequency, the types of requests that are bid on and the prices of their bids do not greatly change due to a first review.

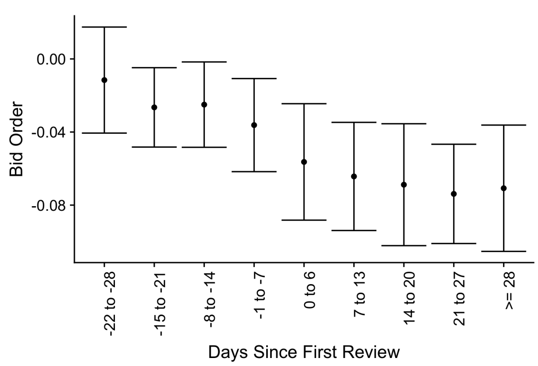
Figure D.3: Review Event Study - Supply Side Responses



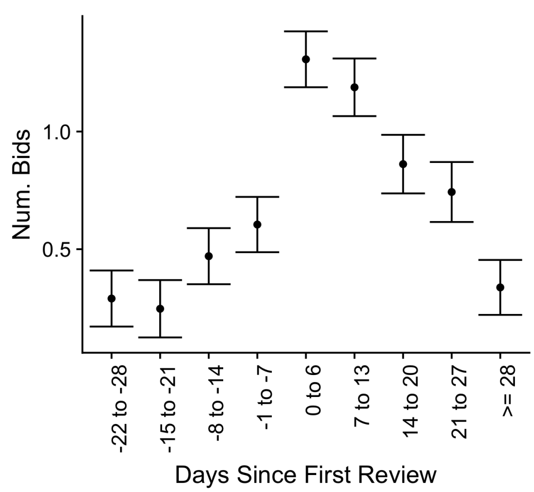
(a) Number of Other Bids on Request



(b) Average Log Prices of Other Bidders on Request



(c) Order of Bid Timing on a Request



(d) Number of Bids by Professional

E Construction of Ratings/Reviews Instruments

Following Chen (2018), we instrument for a professional's rating using a measure of how lenient that professional's previous reviewers tend to be when rating *other* professionals (in any service category on the platform, not just those for home improvement services) . The exclusion restriction this instrument must satisfy is that the leniency of a professional's previous reviewers does not directly affect the current customer's decision to hire the professional *except* through its effect on the professional's current rating. This argument appears to be quite reasonable in this context. For example, customers are unlikely to directly take such leniency into account, as it would require a great deal of searching on the platform for an individual user to learn how a given previous user rates other professionals. A violation of the exclusion restriction might occur if more lenient reviewers are attracted to professionals who are of higher (unobservable to the econometrician) quality, and hence are more likely to be hired by the current customer, but this case also seems unlikely to us in this context.

The leniency measure is constructed as follows. Let R_{jr} represent the set of requests in which users rated professional j before request r is listed. For each request $\tilde{r} \in R_{jr}$, let $i(\tilde{r})$ (or simply "i" for short) represent the identity of the consumer who rated j on request $\tilde{r}$, and let $R_{i(\tilde{r}),-j}$ represent the set of requests on which user i rated some professional *other than* j. We compute the average rating that consumer i gives to professionals other than j as

$$other\_pro\_rating_{i(\tilde{r}),-j} = \frac{1}{\#\{R_{i(\tilde{r}),-j}\}} \sum_{s \in R_{i(\tilde{r}),-j}} indiv\_rating_{i(\tilde{r}),s}$$

where, in the summand, $indiv\_rating_{i(\tilde{r}),s}$ is the actual integer rating i left on some request s, and the notation $\#\{\cdot\}$ represents the count of the elements in a set. We then construct the leniency instrument by averaging over all of these individual consumers' average ratings given to other professionals:

$$leniency_{jr} = \frac{1}{\#\{\tilde{r} \in R_{jr} : R_{i(\tilde{r}),-j} \neq \emptyset\}} \sum_{\tilde{r} \in R_{jr} : R_{i(\tilde{r}),-j} \neq \emptyset} other\_pro\_rating_{i(\tilde{r}),-j}$$

In the case where a professional j has no previous raters who are observed rating some other professional $-j$ at some point, we set $leniency_{jr} = 0$. Our instruments for $ratings_{jr}$

are $leniency_{jr}$ and a dummy for whether professional j has any previous raters who have also rated other professionals (that is, a dummy for whether the leniency measure can be constructed).

We form an instrument for the number of previous reviews using a similar approach to that of the leniency instrument: we construct the *propensity to review* of consumers who have previously hired professional j. Let H_{jr} represent the set of requests on which some user hired professional j before request r is listed. We wish to construct an instrument that captures the expected number of previous reviews we would predict for professional j on request r from knowing who j's previous hirers have been—and how likely they have been to review others whom they have hired. In constructing this instrument, we will take into account that some previous hirers may be slower than others in leaving reviews; that is, even if a previous hirer has not yet left a review, she may do so at some point.[27]

Similar to the instrument for average ratings, for each $\tilde{r} \in H_{jr}$, let $i(\tilde{r})$ represent the identity of the consumer who *hired* j on request $\tilde{r}$, and let $H_{i(\tilde{r}),-j}$ represent the set of requests on which user i hired some professional *other than* j. Also, let $t_{r,\tilde{r}}$ represent the amount of time (in days) between when the hired bid was posted for request $\tilde{r}$ and when j's bid on request r was posted. Let

$$P^r_{i(\tilde{r}),\tilde{r},-j} = \frac{1}{\#\{H_{i(\tilde{r}),-j}\}} \sum_{s \in H_{i(\tilde{r}),-j}} 1\{i(\tilde{r}) \text{ leaves review within } t_{r,\tilde{r}} \text{ days}\}$$

We then construct the propensity-to-review instrument by averaging over all of these individual consumers' propensity to review other professionals they have hired:

$$propensity\_to\_review_{jr} = \frac{1}{\#\{\tilde{r} \in H_{jr} : H_{i(\tilde{r}),-j} \neq \emptyset\}} \sum_{\tilde{r} \in H_{jr} : H_{i(\tilde{r},-j)} \neq \emptyset} P^r_{i(\tilde{r}),\tilde{r},-j}$$

In the case where a professional j has no previous hirers who are observed hiring some other professional at some point, we set $propensity\_to\_review_{jr} = 0$. Our instruments for $review_{jr}$ are then given by $log(propensity\_to\_review_{jr} + 1)$, a dummy for $propensity\_to\_review_{jr}$ being equal to 0, and a dummy for whether professional j has any previous hirers who have

[27] A related point is that some consumers who hire a professional may not have time to leave a review before our main sample period ends. Fortunately, we observe data on ratings and reviews (but not bids and hiring decisions) for a full year following the ending of our main sample period, so this is not a concern.

also hired other professionals. The argument for the validity of this instrument is similar to that of the *leniency* measure: *propensity* will be a valid instrument unless consumers with a higher propensity to review are attracted to professionals who are of higher or lower quality (in a way that cannot be observed to the econometrician). We argue that this exclusion restriction is plausible in our context.

We instrument for the number of previous hires by constructing a measures of consumers' *propensity to hire*. We build this instrument as follows. Fix some professional j and some request r, and suppose we want to calculate the value of the instrument for professional j at the time that request r was listed. Let B_{jr} represent the set of requests for which professional j has bid on before request r was listed. For each request $\tilde{r}$, let $i(\tilde{r})$ represent the identity of the customer who originally created the request $\tilde{r}$, and let $B_{i(\tilde{r}),-j}$ represent the set of requests that customer $i(\tilde{r})$ created in which professional j was not a bidder. We compute customer $i(\tilde{r})$'s own propensity to hire a professional other than j as

$$\mathcal{H}_{i(\tilde{r}),-j} = \frac{1}{\#\{B_{i(\tilde{r}),-j}\}} \sum_{s \in B_{i(\tilde{r}),-j}} 1\{i(\tilde{r}) \text{ hires someone in request } s\}$$

We then construct the propensity-to-hire instrument by averaging over all of these individual consumers propensity to hire other professionals:

$$\text{propensity\_to\_hire}_{jr} = \frac{1}{\#\{\tilde{r} \in B_{jr} : B_{i(\tilde{r}),-j} \neq \emptyset\}} \sum_{\tilde{r} \in B_{jr} : B_{i(\tilde{r}),-j} \neq \emptyset} \mathcal{H}_{i(\tilde{r}),-j}$$

In the case where a professional j has no previous requests for which the creator of that request ever made another request in which professional j was not a bidder, we set $\text{propensity\_to\_hire}_{jr} = 0$. Our instruments for previous hires are then given by the $log(\text{propensity\_to\_hire}_{jr} + 1)$, a dummy for $\text{propensity\_to\_hire}_{jr}$ being equal to 0, and a dummy for whether the $\text{propensity\_to\_hire}_{jr}$ instrument is calculable.

Finally, we also include as instruments the log of $previous\_bids + 1$, where $previous\_bids$ is the number of previous bids this professional has placed, and a dummy for whether this quantity is zero. We also include two instruments for *reviews* being positive and for *previous_hires* being positive. These are given by

$$1-(1-propensity\_to\_review)^{propensity\_to\_hire*previous\_bids}$$

and

$$1-(1-propensity\_to\_hire)^{previous\_bids}$$

Table E.1: First Stage Results For Price IV Regressions

	(4)	(5)
Request FE	Request FE, Pro FE	
Far Pro Distance	-0.0592***	0.00347
	(0.00363)	(0.00495)
Same Location	0.0496***	0.0312***
	(0.00273)	(0.00249)
Log(Distance + 1)	0.0290***	0.0198***
	(0.000574)	(0.000603)
Observations	2603635	2603635

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table E.2: First Stage Results For All IVs Regression

	(1) Log(Price + 1)	(2) Log(Prev Hires + 1)	(3) Prev Hires > 0	(4) Has Rating	(5) Avg Rating	(6) Log(Reviews + 1)
Far Pro Dist	-0.0574***	0.0424***	0.0229***	0.0292***	0.139***	0.0911***
	(0.00364)	(0.00446)	(0.00193)	(0.00233)	(0.0116)	(0.00497)
Same Location	0.0486***	0.00485	-0.00765***	-0.0172***	-0.0800***	-0.00708
	(0.00274)	(0.00332)	(0.00154)	(0.00185)	(0.00921)	(0.00382)
Log(Dist + 1)	0.0284***	-0.0129***	-0.00691***	-0.0101***	-0.0474***	-0.0147***
	(0.000575)	(0.000707)	(0.000313)	(0.000381)	(0.00189)	(0.000809)
Log(Hire Propensity + 1)	-0.258***	1.023***	-0.558***	-0.490***	-2.587***	0.856***
	(0.00672)	(0.00737)	(0.00425)	(0.00499)	(0.0247)	(0.00863)
Hire Propensity Calculable	-0.0479***	0.167***	-0.159***	-0.124***	-0.612***	0.204***
	(0.00529)	(0.00488)	(0.00363)	(0.00415)	(0.0205)	(0.00602)
Hire Propensity =0	0.0291***	-0.473***	0.274***	0.188***	0.969***	-0.442***
	(0.00520)	(0.00500)	(0.00376)	(0.00419)	(0.0207)	(0.00613)
Log(Prev Bids + 1)	-0.0138***	0.601***	0.0469***	0.0534***	0.184***	0.469***
	(0.000482)	(0.000638)	(0.000243)	(0.000316)	(0.00158)	(0.000763)
Prev Bids > 0	0.00916*	-0.620***	0.0728***	-0.0187***	-0.00601	-0.522***
	(0.00412)	(0.00303)	(0.00176)	(0.00256)	(0.0127)	(0.00380)
Instrument for Positive Prev Hires	0.0943***	-1.214***	0.660***	0.494***	2.567***	-1.045***
	(0.00557)	(0.00568)	(0.00379)	(0.00440)	(0.0217)	(0.00695)
Instrument for Has Rating	-0.0203*	0.536***	-0.318***	0.0905***	0.490***	0.665***
	(0.00964)	(0.00970)	(0.00379)	(0.00710)	(0.0350)	(0.0122)
Leniency	-0.00407***	0.0311***	0.00400***	0.00431***	0.0469***	0.0316***
	(0.000979)	(0.00132)	(0.000330)	(0.000519)	(0.00266)	(0.00145)
Leniency Calculable	0.00204	0.258***	-0.0248***	0.0782***	0.319***	0.439***
	(0.00476)	(0.00643)	(0.00162)	(0.00263)	(0.0134)	(0.00709)
Log(Review Propensity + 1)	0.127***	-0.768***	0.0592***	-0.0282***	-0.253***	-0.359***
	(0.00365)	(0.00493)	(0.00120)	(0.00197)	(0.0100)	(0.00554)
Review Propensity	-0.0651***	0.431***	0.589***	Calculable 0.122***	0.563***	-0.288***
	(0.00914)	(0.00893)	(0.00360)	(0.00685)	(0.0338)	(0.0113)
Review Propensity =0	0.0469***	0.00876	-0.231***	0.0424***	0.224***	0.325***
	(0.00923)	(0.00905)	(0.00359)	(0.00691)	(0.0340)	(0.0115)
Observations	2603635	2603635	2603635	2603635	2603635	2603635

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table E.3: Conditional Logit Choice Regression Marginal Effects

	(1) No IVs	(2) IV Price	(3) IV All
License Submitted	0.0171**	0.00761	0.00628
	(0.00798)	(0.00641)	(0.00699)
License Verified	0.00668	0.0151***	0.0152***
	(0.00623)	(0.00501)	(0.00546)
Log(Previous Hires + 1)	0.0244***	-0.00155**	0.0450***
	(0.000753)	(0.000748)	(0.00473)
Average Rating	0.0618***	0.0516***	0.612***
	(0.00137)	(0.00114)	(0.0110)
Log(Reviews + 1)	0.00660***	0.00105	-0.0369***
	(0.000838)	(0.000677)	(0.00638)
Log(Price + 1)	-0.0411***	-0.679***	-0.516***
	(0.000615)	(0.00948)	(0.00859)
N	2603635	2603635	2603635

Table displays results marginal effects from a conditional logit version of the choice regressions. Each regression contains the same variables as the request fixed effects columns in Table 2 (columns 2, 4, and 6), and the request level is the grouping for the conditional logit model. The IV columns are estimated by first performing a first-stage regression of the endogenous variable(s) on the instruments and then controlling for the corresponding residuals from the first stage in the second stage.

F Survey Questions

Below is the set of questions asked in the survey of customers. The order of the numbered answers was randomized. Furthermore, the order of the licensing questions was randomized by blocks. The three blocks randomized consisted of questions 8 and 9, question 10, and question 11.

Q1. Hired

Have you hired someone to do home improvement services on your home in the past year? (For example painting, plumbing, electric services, interior design, heating or AC services, etc.)

Yes (1) No (2)

Display This Question:

If Hired = 1

Q2. HiredDate

When was the improvement done during the past year Year Month

Q3. ServiceType

What type of home improvement service did you need help with? Describe in a few words:

Q4. HomeLocation

Where was the home needing improvement located? Alabama (1) ... Home is/was not located in the United States (53)

Q5. OwnHome

Did you own or jointly own the home where you needed the home improvement service? Yes (1) No (2) Other. Please specify: (3)

Q6. FindProvider

How did you find the service provider? Select ALL that apply: Referral from a friend (1) Google (2) Yelp (3) Angie's List (4) Yellow Pages (5) HomeAdvisor (6) Thumbtack (7) Other. Please specify: (8)

Q7. Reasons

What are two or three reasons why you chose this service provider over other providers? List the reasons from most important to least important. Most important (1) Second most important (2) Third most important (3)

Q8 Approximately how much in total did you pay for this service?

Q9 Approximately how many hours did the job take?

Q10. Licensed

Did the service provider you hired have an occupational license? Yes (1) No (2) Not sure (3)

Display This Question:

If Licensed = 1

Q10b. DiscoverLicenseInfo

How did you know whether the service provider you hired had an occupational license? It was in the contract I signed (1) He/She told me (2) I saw it on Yelp, or a similar website (3) I verified on a government website (4)

Q11. LicenseRequired

Does the service provider you hired work in a profession for which occupational licensing is required by law in your geographic area? Yes (1) No (2) Not sure (3)

Q12. LicensingDifficulty

Do you think obtaining an occupational license in your geographic area for the service you requested is: Easy, requiring little training and post-secondary education (1) Moderately difficult, requiring some training and post-secondary education (2) Difficult, requiring a lot of training and post-secondary education (3) Not sure (4)

Q13a. LicensingOpinion

Suppose laws were to change so that an occupational license is no longer required for the home improvement services you requested. What would be your opinion of this change? In favor (1) Opposed (2) Indifferent (3)

Q13b. LicensingOpinion

Suppose laws were to change so that an occupational license is required for the home improvement services you requested. What would be your opinion of this change? In favor (1) Opposed (2) Indifferent (3)

Q13c. LicensingOpinion

What would be your opinion of a law requiring occupational licensing for the home improvement services you requested? In favor (1) Opposed (2) Indifferent (3)

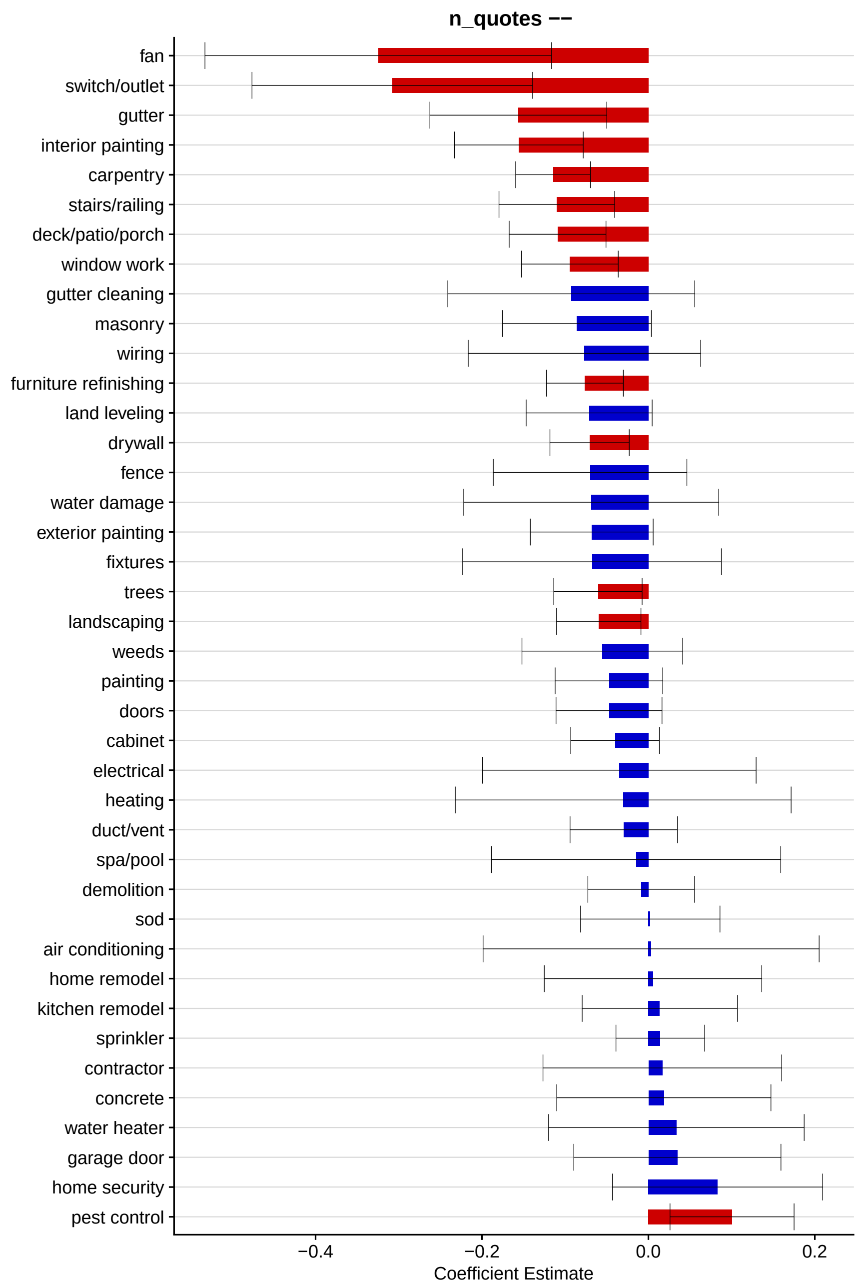
G Heterogeneity of Licensing Effects

In this section, we estimate the effect of licensing stringency on equilibrium outcomes separately for each category. Figures G.1 through G.4 plot the coefficients estimates for different outcome variables. There is a high degree of heterogeneity in the coefficient estimates. At the bidding stage (fig. G.1 and G.2), we find that stricter licensing requirements limit competition and increase quoted prices for the majority of the service categories. The first figure shows that stringency reduces the number of quotes received by each request in most service categories. The second figure implies that for most categories stringency increases the quoted prices. The coefficients range in magnitude and significance, but many imply sizable effects of licensing regulation.

At the matching stage (fig. G.3), few coefficient estimates are statistically significant. Similarly, Figure G.4 does not provide conclusive evidence on how licensing regulation affects customer satisfaction at a category level. When we look at the propensity to return to the platform and post another request, stricter regulation mostly decreases buyer retention but we are lacking in statistical power to precisely measure these estimates.

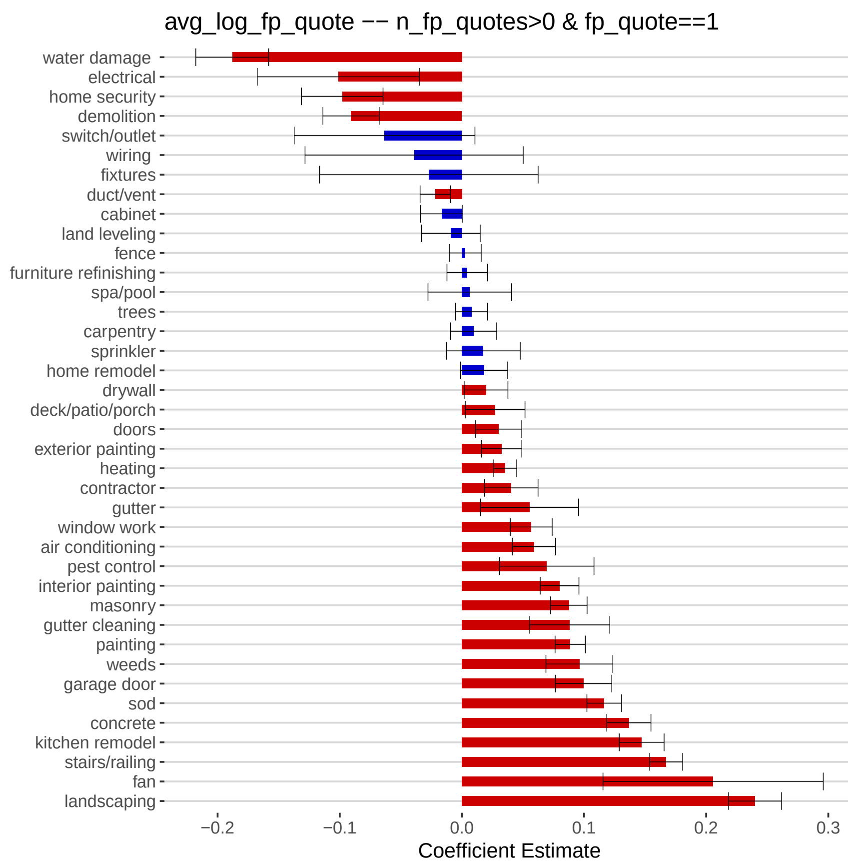
In this section we have shown that licensing stringency increases prices faced by consumers, with no positive average effects on post-transaction customer satisfaction. Our data does not allow us to precisely measure category specific effects of licensing for match and post-transaction outcomes. However, we are able to detect heterogeneity at the category level for competition effects—i.e. number of quotes submitted for each task—and quoted prices.

Figure G.1: Market-Level Effects at the Bidding Stage - Num. Quotes



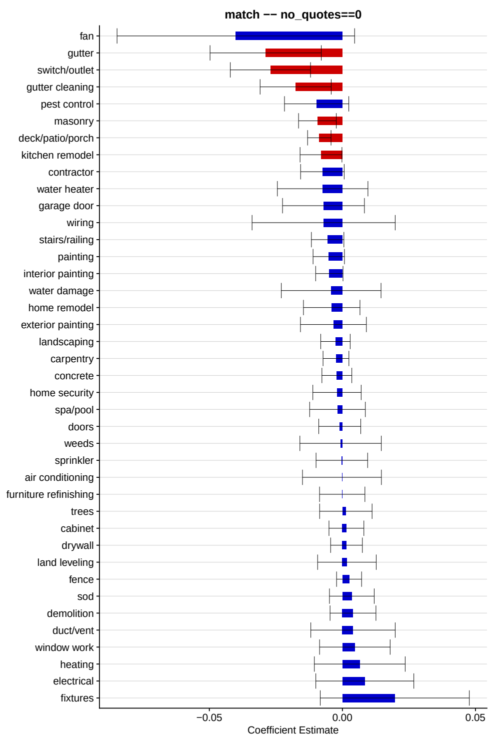
The figures plot the estimates of α_c*, one per category, from equation 4. The dependent variable is the number of quotes received by request* r*. The coefficient estimates are identified by the red bars (blue if not statistically significant), while 95% confidence intervals are identified by the black lines.*

Figure G.2: Market-Level Effects at the Bidding Stage - Average Log(Bid Price)

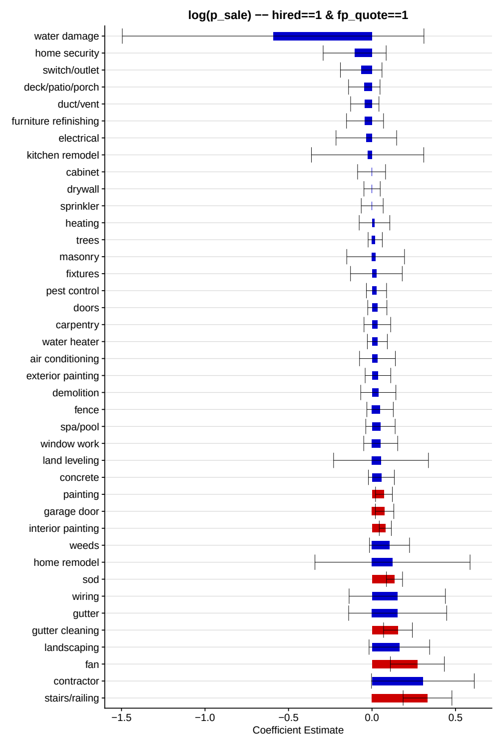


The figure plots the estimates of α_c, one per category, from equation 4. The dependent variable is the average (log) price of fixed price quotes received by request r. The coefficient estimates are identified by the red bars (blue if not statistically significant), while 95% confidence intervals are identified by the black lines.

Figure G.3: Market-Level Effects at the Matching Stage



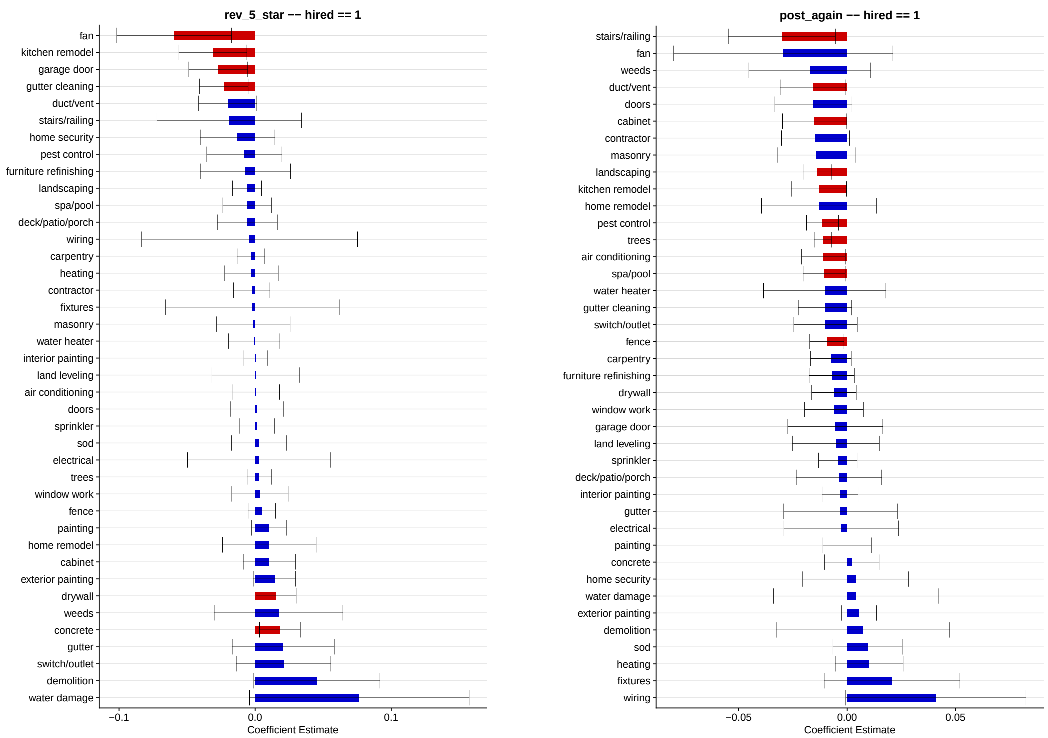
(a) Outcome: Match



(b) Outcome: Winning Price

The figures plot the estimates of α_c, one per category, from equation 4. On the left panel, the dependent variable is the a dummy for whether a professional was hired for request r, conditional on receiving at least one quote. On the right panel, the dependent variable is the (log) price of the winning quote for request r, when this quote was submitted with a fixed price bid. The coefficient estimates are identified by the red bars (blue if not statistically significant), while 95% confidence intervals are identified by the black lines.

Figure G.4: Market-Level Effects at the Post-Transaction Stage



(a) Outcome: 5 Star Review

(b) Outcome: Customer Returns to Platform

The figures plot the estimates of α_c, one per category, from equation 4. On the left panel, the dependent variable is a dummy for whether a consumer left a five star review to the professional hired for request r. On the right panel, the dependent variable is a dummy for whether a consumer posted another request at least one week after posting the matched request r. The coefficient estimates are identified by the red bars (blue if not statistically significant), while 95% confidence intervals are identified by the black lines.

H Additional Figures and Tables

Figure H.1: Distribution of Ratings

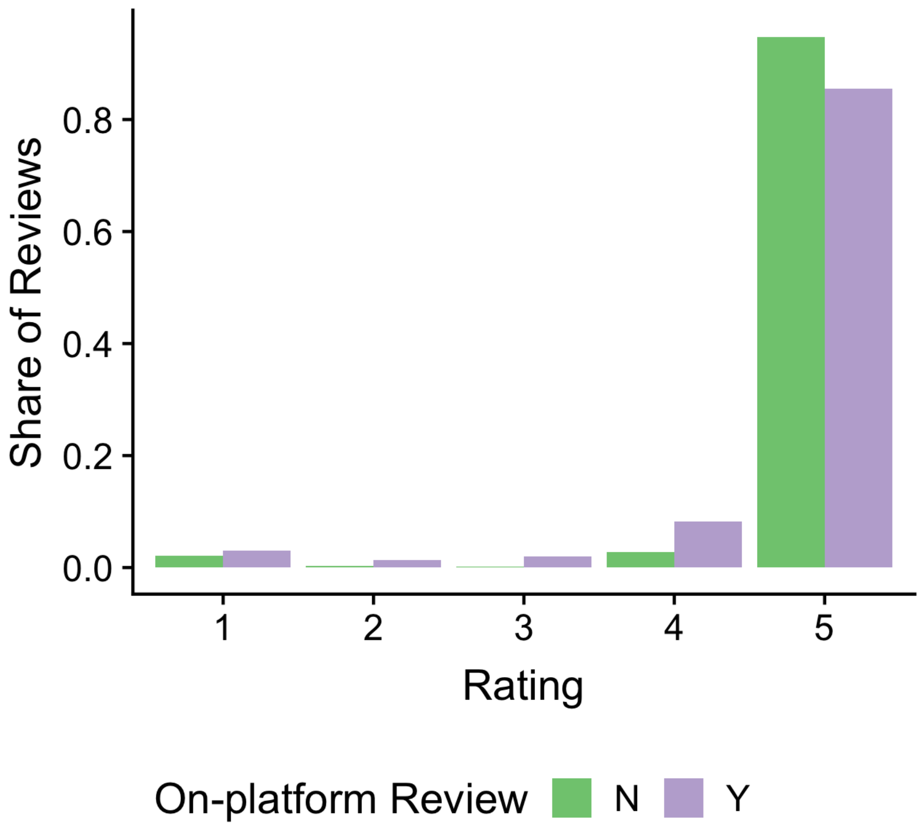
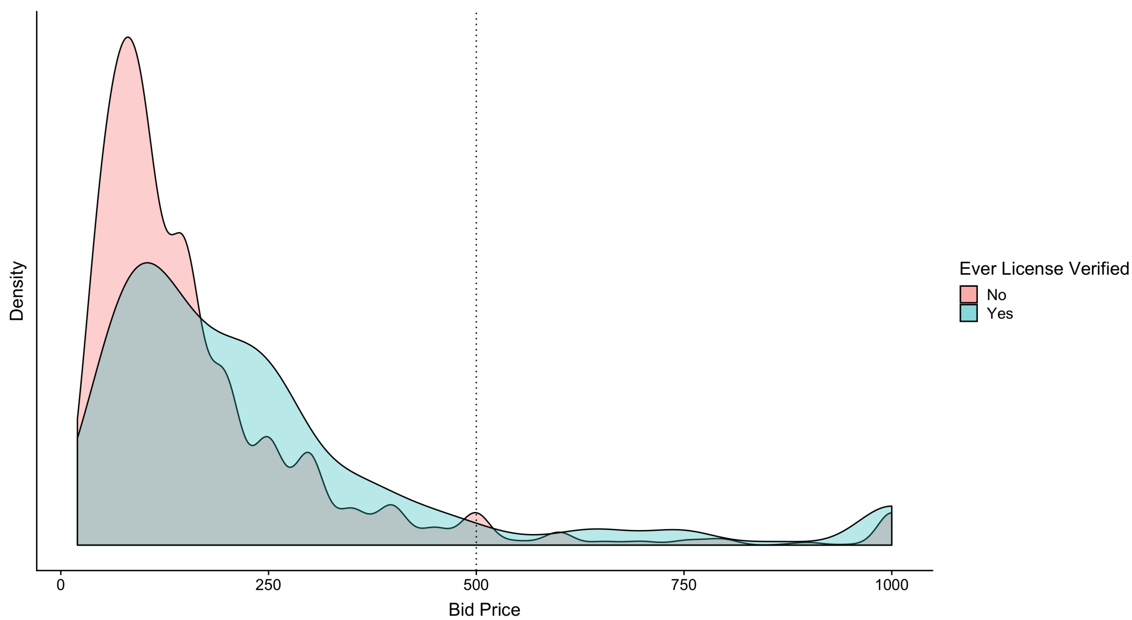
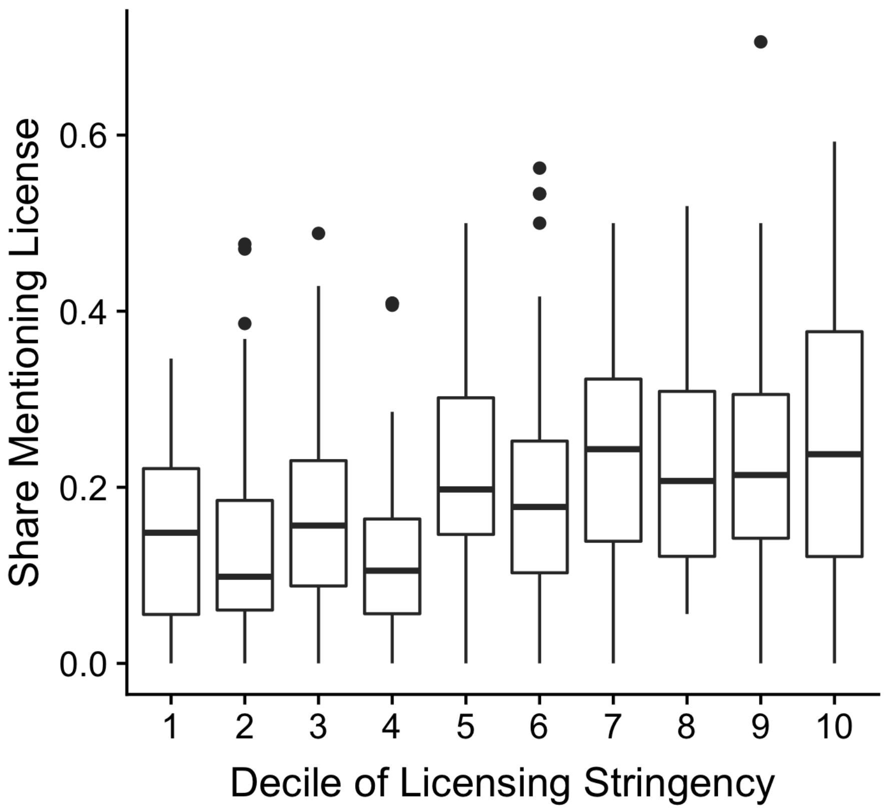


Figure H.2: General Contractor Bids By Verified License Status (California)



This figure presents the distribution of fixed prices bid for General Contractor tasks in California. Ever licensed is a binary variable taking the value of 1 if we ever observe the professional having a platform verified license in the data. Prices are censored at 1000 to improve readability.

Figure H.3: Licensing Stringency and Share of Licensed Professionals



The figure plots the average licensing stringency and share of licensed professional by quantile of the stringency variable. The licensing stringency variable is obtained from data on occupational licensing regulation (see section 5 for details), while the share of licensed professionals comes from the data obtained by scraping the website (see Appendix A for details). The stringency variable is grouped into deciles, the x-axis plots the average within each quantile, and the y-axis plots the median and interquartile range of the shares of professionals who disclose their licensing status on the platform, either by submitting proof of license which is then verified by the platform or by mentioning their license on their profile. Individual points are outliers exceeding 1.5 times the interquartile range.

Table H.1: Top Categories by Number of Professionals

Category	Text Lic.	Verified Lic.	Both Lic.	Cert.	Insurance	Credential	Background	Num. Pros
General Contracting	0.180	0.120	0.042	0.055	0.170	0.330	0.140	3,242
Handyman	0.084	0.045	0.016	0.038	0.100	0.180	0.170	2,285
Electrical and Wiring Issues	0.230	0.120	0.056	0.068	0.160	0.350	0.170	2,211
Roof	0.160	0.120	0.039	0.110	0.250	0.400	0.160	1,952
Carpet Cleaning	0.058	0.005	0.002	0.120	0.100	0.200	0.140	1,892
Home Inspection	0.230	0.180	0.073	0.240	0.160	0.500	0.190	1,802
Interior Design	0.044	0.039	0.011	0.058	0.022	0.120	0.180	1,801
Property Management	0.140	0.180	0.066	0.038	0.063	0.300	0.140	1,766
Interior Painting,Painting	0.090	0.069	0.018	0.048	0.150	0.240	0.210	1,615
Commercial Cleaning	0.076	0.006	0.003	0.039	0.150	0.190	0.170	1,445
Welding	0.031	0.010	0.003	0.140	0.037	0.170	0.064	1,411
Home Staging	0.052	0.025	0.007	0.072	0.036	0.150	0.160	1,398
Pressure Washing	0.093	0.025	0.005	0.042	0.180	0.240	0.220	1,394
General Carpentry	0.074	0.045	0.013	0.028	0.091	0.170	0.100	1,347
Architectural Services	0.140	0.120	0.036	0.035	0.029	0.250	0.100	1,345
Fence Related	0.091	0.051	0.014	0.043	0.110	0.210	0.180	1,317
Central AC	0.170	0.120	0.043	0.110	0.130	0.330	0.200	1,288
Flooring	0.095	0.059	0.020	0.057	0.120	0.230	0.160	1,276
Concrete Installation	0.100	0.066	0.017	0.044	0.130	0.230	0.160	1,249
Window Cleaning	0.081	0.010	0.002	0.035	0.180	0.210	0.210	1,242

Table H.2: Top Categories by % Mentioning Licensing in Profile Text

Category	Text Lic.	Verified Lic.	Both Lic.	Cert.	Insurance	Credential	Background	Num. Pros
Plumbing	0.280	0.190	0.095	0.087	0.150	0.440	0.290	576
Home Inspection	0.230	0.180	0.073	0.240	0.160	0.500	0.190	1, 802
Electrical and Wiring Issues	0.230	0.120	0.056	0.068	0.160	0.350	0.170	2, 211
Bed Bug Extermination	0.220	0.150	0.061	0.120	0.120	0.380	0.220	1, 139
Animal and Rodent Removal	0.210	0.100	0.042	0.110	0.110	0.340	0.200	424
Fixtures	0.190	0.110	0.044	0.056	0.120	0.310	0.190	681
Ceiling Fan,Fan Installation	0.180	0.120	0.059	0.065	0.120	0.300	0.330	493
General Contracting	0.180	0.120	0.042	0.055	0.170	0.330	0.140	3, 242
Central Air Conditioning Repair or Maintenance	0.170	0.120	0.043	0.110	0.130	0.330	0.200	1, 288
Land Surveying	0.160	0.140	0.045	0.210	0.074	0.410	0.066	470
Central Air Conditioning Installation	0.160	0.083	0.035	0.110	0.120	0.280	0.110	942
Roof Installation or Replacement	0.160	0.120	0.039	0.110	0.250	0.400	0.160	1, 952
Lighting Installation	0.160	0.110	0.059	0.063	0.140	0.290	0.260	494
Mold Inspection and Removal	0.150	0.085	0.035	0.310	0.250	0.470	0.180	1, 091
Local Moving	0.150	0.120	0.052	0.029	0.180	0.280	0.240	445
Property Management	0.140	0.180	0.066	0.038	0.063	0.300	0.140	1, 766
Architectural Services	0.140	0.120	0.036	0.035	0.029	0.250	0.100	1, 345
Long Distance Moving	0.140	0.120	0.042	0.038	0.160	0.290	0.190	818
Switch and Outlet Installation,Tile Installation	0.140	0.054	0.020	0.041	0.077	0.210	0.110	607
Tree Planting	0.130	0.029	0.012	0.088	0.220	0.300	0.150	907

Table H.3: Confusion Matrices for Price Predictions

$200 threshold

Actual/Predicted	0	1	Total
0	181, 119	36, 147	217, 266
1	44, 100	153, 145	197, 245
Total	225, 219	189, 292	414, 511

$500 threshold

Actual/Predicted	0	1	Total
0	288, 699	20, 513	309, 212
1	37, 560	67, 739	105, 299
Total	326, 259	88, 252	414, 511

$1,000 threshold

Actual/Predicted	0	1	Total
0	337, 334	13, 538	350, 872
1	27, 829	35, 810	63, 639
Total	365, 163	49, 348	414, 511

Confusion matrices for price predictions. The top panel shows the number of requests with at least one fixed price quote, and divide them based on whether the actual fixed price quote is above $200, and whether the predicted fixed price quote is above $200. On the diagonal we have jobs for which the prediction matches reality. The middle panel does the same for a $500 threshold, and the bottom panel for a $1,000 threshold. AUC (area under the curve) accuracy measures are 0.882 (95% C.I. 0.881-0.883), 0.916 (95% C.I. 0.915-0.917), and 0.927 (95% C.I. 0.926-0.928) for the three thresholds respectively.