Market Concentration, Labor Quality, and Efficiency: Evidence from Barriers in the Real Estate Industry

NOTE: THIS IS A PRELIMINARY DRAFT

For the most updated version, please click here.*

Jeanna Kenney

November 13, 2023

Abstract

This paper estimates the trade-off between quality and concentration due to occupational licensing and the consequences of entry restriction for market efficiency in the real estate industry. I exploit a policy reform in Texas in 2012 which provides quasi-exogenous variation in the cost of entry that real estate salespeople face in order to become brokers. A future increase in licensing cost should lead to an anticipatory spike in entry before a supply restriction, which has different implications in the short vs. long term. Using a novel dataset of the universe of licensees in a number of states matched with property listing data, I compare counties in Texas with similar counties in other states in a synthetic difference-in-difference estimation. The reform generates an unintended short-term increase in entry, leading to a 12% increase in the stock of brokers in the average county, before broker supply is restricted in the long term. The increased barrier does not increase broker quality, while market concentration persistently decreases by more than 25%, due to the unintended entry effect, suggesting minimal benefits to consumers of restricting entry. However, efficiency, i.e., average broker listing volume, decreases due to entry. This simple definition does not capture distributional effects of barriers: I find that costlier licensing leads to a smaller share of entering female and Hispanic brokers.

*Jeanna Kenney: jhkenney@wharton.upenn.edu, Wharton School, University of Pennsylvania.

I would like to thank my dissertation committee, Fernando Ferreira, Benjamin Keys, Todd Sinai, and Judd Kessler for their advice on this project. I would also like to thank Jessie Handbury, Maisy Wong, Nitzan Tzur-Ilan, Yichen Su, Lauren Lambie-Hanson, Leonard Nakamura, Giri Parameswaran, and seminar participants at the Wharton Urban Lunch, Federal Reserve Bank of Dallas, Urban Economics Association Summer School, Haverford College, and Swarthmore College for their insights. Finally, I thank Devika Hazra, Claes Backman, and Thomas Davidoff for helpful discussions of earlier versions of this paper.

1 Introduction

Occupational licensing is a common tool that trade associations use to protect consumers from untrained participants or to protect incumbents from competitors. A benign view of licensing suggests that increased training leads to a better quality workforce. However, economists often consider a trade-off between a possible increase in the quality of workers and an increase in market concentration; by increasing the cost of entry, licensing potentially provides market power to incumbent workers. This trade-off is particularly understudied in apprenticeship industries, in which entry-level employees must initially be trained by professionals with more experience (e.g., counselors, plumbers, and real estate agents). Further, economic theory generally posits that free entry is most efficient for markets; thus, any policy which restricts entry into an industry or progression to the next license should have consequences for productivity and output.[1]

This paper estimates both the trade-off between quality and market concentration due to occupational licensing and the consequences of entry costs for market efficiency. I study these issues in the setting of the U.S. real estate industry, where there are generally two types of real estate licenses: a "salesperson," the entry level, who must work under the employment of a "broker," the professional level. The real estate industry is not only a massive consumer industry, but it is also a large labor market. Over 90% of home sellers used a real estate agent in 2020, and there are over 3 million licensees in the U.S.[2] Agents are crucial to the home selling and buying process: not only do they facilitate matches between buyers and sellers (Salant, 1991), but they also advise households how to buy and sell what for many is their largest asset and potentially steer clients towards certain financial products (Lopez et al., 2019).[3]

Establishing connections between apprenticeship licensing and various market outcomes is challenging for two primary reasons. The first is measurement; particularly when studying apprenticeships, the timing of entry and exit between the levels of employment *within* the industry, not

[1] Friedman and Kuznets (1945) first study these consequences in the setting of increasing entry standards for doctors. Stigler (1971) and Shapiro (1986) also provide early work on the effects of regulating labor market entry.

[2] National Association of Realtors (NAR)

[3] See Bernheim and Meer (2013) for additional examples of the bundle of services real estate agents offer.

just in and out of the industry, must be observed. Further, it is often difficult to measure productivity and performance in many service industries. The second challenge is identification; there are many factors which could explain worker entry into an industry and consumer outcomes in that industry. Therefore, to isolate the effect of labor market entry on worker productivity and industry outcomes, an exogenous change in the cost of entry is needed. However, there is often minimal geographic variation in licensing policy or enforcement.

To overcome these challenges, I first compile a novel dataset of the universe of real estate licensees in a number of states, matched with productivity information in the form of listings data from several Multiple Listings Services (MLS). These new data allow me to separately identify licensees by type and connect licensees to their own output, which prior studies have been unable to leverage.

Second, to deal with identification challenges, I utilize a research design exploiting a policy reform which provides quasi-exogenous variation in licensing cost. Specifically, in 2012 the Texas Real Estate Commission (TREC) announced that the eligibility requirements to upgrade to a broker license, the professional level in the industry, would significantly increase.[4] A notable feature of the policy is that it was announced a year before it was set to go into effect, thereby allowing for an anticipatory response.

To understand the possible impacts of such a policy reform, I develop a conceptual framework enhancing the stylized model of Hsieh and Moretti (2003). Using a static model of wages for homogeneous agents, the authors show that, in a market such as real estate where commissions are generally fixed, free entry of agents is socially inefficient. The inefficiency occurs because average agent productivity, i.e., listings per agent, decreases and new entrants dissipate any additional market commissions. In my model, which allows for heterogeneity in quality, brokers have a dynamic wage equation and are forward looking. Thus, when there is an unexpected announcement of a future increase in the licensing costs, an anticipatory influx of brokers is predicted in the short term before a long-term decrease in entry. The model illustrates that a future increase in an entry

[4]The requirements to become a salesperson, the entry-level license, remained unchanged.

barrier should predict a decrease in market concentration in the short term and an increase in the average quality of entering brokers in the long term. While *listings* per broker will decrease in both the short and long term, the effect on efficiency as defined by *sales* per broker is ambiguous.

In order to test these predictions, I compare county-level markets in Texas with similar control counties in other states before and after the policy reform. Using an event study design, I first establish that, because it was announced a year before its effective date, the policy generates a significant anticipatory influx of entry in the short run, before suppressing entry in the long-term. In turn, the stock of licensed brokers increased for about three years by 12% in the average county, and then afterwards decreased relative to control counties. Thus, this policy is valuable for studying the effects of entry costs in apprenticeship industries because it generates two phases of variation: a short-term inducement of entry due to current low barriers, and a long-term restriction of entry due to higher barriers. To estimate the short- and long-term impacts of entry barriers on quality, concentration, and efficiency, I then utilize a novel synthetic difference-in-difference framework that weights control counties and time periods based on their similarity to treated counties.[5] This is one of the first papers to my knowledge to apply the synthetic difference-in-difference method to a specific setting.

Ultimately, I provide evidence of a licensing setting where practitioner quality is unchanged in the long term while concentration decreases due to an unintended effect of the policy. In the short term, I find that the new brokers who rush to enter under the easier regime are of lower quality than prior cohorts of brokers, as measured by probability of sale and aggregate listing outcomes (i.e., sale price and days on the market). Brokerage market concentration decreases by around 9% due to the unintended supply influx, driven by markets with the smallest stock of brokers before the policy. In the long term, entering brokers are higher quality in that they are more adept at completing sales, but they do so by generating lower return to listings. Further, the increased training does not translate into a measurable increase in the quality of salespeople, suggesting there are minimal gains in the ability of brokers to train their apprentices. The market de-concentration

[5] See Arkhangelsky et al. (2021).

amplifies to around 27%.

Taken together, these results suggests that there are minimal benefits to consumers of restricting entry via higher licensing costs. On the contrary, unintended entry leads to a decrease in market efficiency, as defined by either the listings or sales per broker. This result suggests that higher licensing costs are desirable in order to decrease entry in the market. However, this narrow definition of efficiency does not capture the full distributional consequences of restricting entry. I find that costlier licensing leads to a 24% decrease in the share of entering female brokers and 47% decrease for Hispanic brokers, for whom the entry barriers may be particularly costly.

This paper contributes to the broader licensing literature by presenting a setting in which both an entry increase and an entry decrease due to licensing can be separately considered. The effects of more stringent licensing on quality and concentration have been studied in a wide array of industries, beginning with the work of Friedman and Kuznets (1945), which showed that increasingly restrictive entry into medical school led to growth in doctors' wages. In more recent work, Kleiner and Soltas (2023) estimate a model of labor market equilibrium showing that licensing leads to a welfare loss borne more so by laborers than consumers. Few studies find improved quality effects of licensing; one recent exception is Anderson et al. (2020), which shows that midwifery licensing laws in the early 1900s led to a decrease in maternal and infant mortality.[6] In their review of the recent licensing literature, Kleiner and Timmons (2020) highlight that research into the effects of occupational licensing across industries has historically been limited due to the high cost and lack of availability of regulatory data. This paper overcomes this challenge by compiling new data relevant to licensing: the types of licenses within an industry, linked to output for those licensees.

A smaller subset of this literature has focused specifically on licensing in the real estate industry. For instance, Shilling and Sirmam (1988) provide early empirical evidence that higher barriers to entry, as measured by the licensing examination pass rate set by the state licensing board, improve agent quality, as measured by fewer formally filed complaints, but have anti-competitive effects as well. These studies have focused largely on licensing exam pass rates as a measure of

[6] See also Angrist and Guryan (2008), Blair and Chung (2019), Zapletal (2019), Yelowitz and Ingram (2021), and Bowblis and Smith (2021).

barriers to entry and formally filed complaints against agents as a measure of agent quality.[7] In more recent work studying an increase in the training required for both salespeople and brokers in Illinois, Chung (2022) finds that increased renewal cost led to an one-time spike in exits with no long-term increase in agent quality, also as measured by complaints.[8]

This paper additionally contributes to the literature by focusing on professional- as opposed to entry- level barriers. Specific to the context of real estate, this is the first paper to my knowledge to directly consider how the distinction between salespeople and brokers impacts broader market outcomes and to use licensee data across multiple states to do so. Using data from Virginia, Turnbull et al. (2022) find that, because brokers have stronger reputational concerns in that their income is dependent on the performance of their whole office, they are more likely to sell a clients' home as they would their own.[9] Using licensee information from Nevada, Lopez (2021) finds that licensees (and their family members) sell their properties at higher prices. The author controls for whether the licensee holds a broker or salesperson license, though a difference in performance of the two is not the focus of the study. Turnbull and Waller (2018) similarly controls for the type of license while studying the added value that "top-tier" agents, as defined by total market share, bring to transactions. These papers control for the license type but do not analyze outcomes separately by type. Gilbukh and Goldsmith-Pinkham (2019) make a distinction between "experienced" and "inexperienced" agents.[10] The authors show that more inexperienced agents have a lower probability of sale, even when controlling for the brokerage they work for.[11] They write a model estimating the effect of the experience distribution of agents on housing market liquidity, and show that

[7] Jud and Winkler (2000) put forth a model suggesting that the stringency of the licensing exam pass rate and the educational requirements should decrease the numbers and increase the incomes of real estate agents. For related empirical work, see Guntermann and Smith (1988) and Johnson and Loucks (1986).

[8] Few of these papers link occupational licensing to the housing outcomes obtained by agents. One exception to the former is Carroll and Gaston (1979), which uses the amount of time a listed home is vacant as a measure of the quality of agents. The authors find that increased licensing requirements, as measured by exam pass rates and educational requirements, in fact have an adverse effect on quality.

[9] This result complements the work of Levitt and Syverson (2008), which initially showed that agents in general (i.e., when not distinguishing their type) sell their own home at higher prices than those of their clients due to agency costs.

[10] However, the authors do not use information of licensing or type of agent, and so experience is a noisy proxy for the salesperson and broker distinction in this study.

[11] Similarly, Waller and Jubran (2012) find that "rookie" agents, i.e., those with a salesperson license for less than two years, sell for less and have longer market duration.

increasing entry costs, which should in turn lead to better experienced agents, increases liquidity.

The paper will proceed as follows: Section 2 describes the real estate industry licensing structure and the Texas policy change. Section 3 presents a conceptual framework regarding the effect of a future increase in entry costs for professional-level workers. Section 4 describes the new data used for this study, while Section 5 presents the empirical strategies used to test the framework's predictions. Section 6 discusses the trade-off between quality and competition, as well as the consequences for market efficiency. Finally, Section 7 concludes.

2 Institutional Context

All U.S. states require a license in order to represent others in the sale or purchase of real estate. There are generally two types of licenses for real estate agents in the U.S., typically called a "salesperson" and a "broker."[12] Salespeople and brokers can both represent clients in buying and selling property. However, a salesperson must do this under the supervision and employment of a broker. A broker, on the other hand, is eligible to work independently and to hire other real estate agents to work for them.

A key distinction between the two types of licensees is that only brokers can legally contract with clients. Thus, even if a salesperson obtains a listing on her own, the broker's name must be on the contract, and it is the broker who is paid and provides payment to the salesperson. Brokers typically share the commission on the sales of all the salespeople they employ. The most common payment scheme is the commission split model, in which the broker and salesperson split the gross commission from all transactions the salesperson completes.[13]

There are relatively low barriers to obtaining a salesperson license. An individual typically mus take a certain number of credit-hours of education, either in person or via online instruction,

[12]This licensing process is overseen by a state real estate commission or board, traditionally made up of a number of local real estate professionals, who is charged with establishing the licensing requirements, monitoring real estate activity, and enforcing real estate rules and regulations.

[13]Another common, though less prevalent, scheme is office fee model, in which the salesperson pays the broker a regular fee but keeps all commissions.

focused on issues such as contract terminology, disclosure laws, and anti-discrimination polices, and must then pass two exams. One is focused on state-specific rules and regulations, and one is a broader national-focused exam. If the individual passes the exams, they may then officially apply for a license. For that license to be active and valid, the licensee must find a licensed broker who will sign the license and display it in the office. In other words, a salesperson cannot begin representing clients until they are employed by a broker. The total fees to become a salesperson can range from about $400 to $1,000.[14] This is a relatively small entry cost: For context, a 3% commission to a listing agent on a listing sold for $100,000 is $3,000, and even if the agent splits 50% of that with the broker, they still make $1,500.

A primary reason to upgrade to a broker's license is the ability to work for oneself and therefore not have to share any commission. Additionally, a broker can hire other agents and earn money off of the transactions of those agents. A broker is also eligible to be the "managing" or "designated" broker for a brokerage firm, which may come with added compensation for running the day-to-day operations of a firm.[15]

The eligibility requirements to upgrade to a broker license are more restrictive than those for a salesperson license, though similarly low in direct financial costs. To be eligible to upgrade to a broker's license, an agent must be actively licensed as a salesperson for a given number of years, depending on the state.[16] Before applying, the agent must take additional broker-specific coursework and pass another set of exams. In certain states, the application must include a resume of prior listings signed off by the current employing broker. For instance, a requirement of precisely this type was added as part of the 2012 Texas policy change, which will be discussed below.[17]

[14] See this example from online education provider, VanEd. This includes the fee for the background check, a course registration fee, an exam fee, and an application fee.

[15] While each brokerage firm is required to have one broker as the "managing," more than one broker can work at a firm. These brokers are typically called an "associate broker" or a "broker-salesperson."

[16] The modal experience requirement across states is two years, according to the Association or Real Estate Licensing Law Officials (ARELLO).

[17] The tests for both levels of licenses are not particularly prohibitive to pass. According to ARELLO, of the 15 states that reported exam data in 2021, the modal passing score for the state and national exams for both brokers and salespeople is 70%. A license of either type then typically has to be renewed every two to three years, depending on the state, to remain active. The renewal process requires an additional amount of continuing education coursework during the renewal period and a renewal fee to the state. Failure to comply with the renewal process will result in an inactive, and ultimately expired, license.

While the main financial cost of becoming licensed as a broker, e.g., the additional coursework, exam, and application fees, are similar to that at the salesperson level, there are additional non-monetary costs which may dissuade an agent from upgrading. For instance, there are startup costs of setting up an office for brokers who choose to open their own brokerage firm. Similarly, brokers who manage other salespeople typically spend fewer hours working directly with clients and more time managing and training their supervisees. A broker also assumes the liability of their employees; in other words, if a broker's salesperson does something unethical on the job, the broker could lose their license or suffer other disciplinary action.

2.1 Texas Real Estate Commission Policy Change

Prior to 2012 in Texas, an agent had to be licensed as a salesperson for at least two years to be eligible to apply to be a broker. Before applying, the agent was required to take 270 credit-hours broker-specific education and then pass another set of exams similar in structure to those for salespeople.

In June 2011, the Texas Real Estate Commission (TREC) Act was amended via Texas Senate Bill 747. The bill stated that, effective January 1, 2012, an aspiring broker applicant would need to hold a salesperson's license for at least four years before applying, doubling the previous standard, and, if the applicant did not have a degree from an accredited university, take an additional 630 course-hours of education, tripling the previous standard. Additionally, the bill noted that TREC would have to write new rules instituting a transactions experience requirement for brokers.

These new rules were announced in October 2011, again to be made effective on January 1, 2012. The most substantial change was that applicants would have to show evidence of completed transactions (i.e., representing either a seller or buyer in a completed sale) in each of the prior four years of holding a salesperson's license. Prior to this, broker applicants did not have to provide any proof of transactions, and could essentially apply to be a broker after two years of holding a salesperson's license even if the agent had never been involved in an actual transaction.[18]

[18]The organization "Texas REALTORS" described the motivating force of the policy change as "focused on better

To measure experience, TREC introduced a point system which assigned various values to different types of transactions (e.g., residential vs. commercial, rentals vs. sales, etc.). Essentially, the point system amounted to proving the agent had transacted at least one property in each of the four years prior and around twelve transactions total across the four years. This record of experience would have to be verified by the applicant's employing broker.

Anyone who was already licensed as a broker and had active status before the announced policy change was made effective was grandfathered in by the policy, but needed to meet the new education requirements by the next renewal. A previous broker whose license had been expired more than two years could only apply for reinstatement of the license if the applicant met all the new requirements to apply. If not, the individual had to start over as a salesperson. Note that the policy referred to being eligible to *apply* to be a broker. Applications can take up to a year to become official; thus, entry effects are delayed and for empirical purposes I consider January 2013 to be the effective date of the policy.[19] An overview of the policy change is displayed in Table 1. Note that, to this day, Texas remains the most stringent state in terms of broker licensing.[20]

Table 1: Broker Requirements in Texas Before and After TREC Change

	Before	After
Wait Period	2 Years	4 Years
Education	270 Hrs	270 Hrs*
Exams	Yes	Yes
Transactions Experience	0	12**

*Broker responsibility course now required; additional 630 hours of elective education needed if no college degree
**Applicants needed at least one transaction in each of the preceding four years and about twelve transactions total across those four years.

This policy reform provides an unexpected future change to broker licensing cost. I therefore exploit this quasi-exogenous variation in the cost in order to analyze the economics consequences

preparing license holders to represent consumers in real estate transactions and ensuring education for applicants and license holders is targeted and of the highest quality."

[19]In Texas, an applicant can apply first before meeting all the requirements, and the application is only approved once (and only if) the requirements are met within a year of submitting. This gap in application and approval is why the data show a run up in brokers in late 2012 (as opposed to 2011), because the application was submitted just before the policy went into effect in late 2011.

[20]The only change for salespeople in Texas at the same time was the removal of exemptions from course work for those who had college credit in a real estate-related field.

of licensing barriers at the professional level in an industry.

3 Conceptual Framework

In this section, I develop a conceptual framework to study the impact of future increases in entry barriers on the employment decisions for forward-looking brokers. This framework enhances the stylized setting of Hsieh and Moretti (2003), who posit that free entry is inefficient in a setting such as real estate when the price of the service is generally fixed.[21] They write a static wage equation for real estate agents and show that, when the price of land in a city increases, the number of real estate agents increases, the productivity of an average real estate agent decreases, and real wage of typical agent remains unchanged, as this should be directly proportional to land prices. Decreasing average broker productivity points to social inefficiency because this suggests that there are agents who could be engaged in profitable activities in other industries.

In this framework, brokers are forward-looking and face some sort of entry cost. They have a dynamic, instead of static, wage equation which leads to changing equilibrium broker employment over time. The framework is then extended to predict how differing short- and long-term patterns in entry will affect quality, competition, and efficiency in the market. I use quasi-exogenous variation in entry, as opposed to cross-sectional variation in land prices, to test these predictions.

3.1 Labor Entry Decision

I begin by modelling the entry decision for professionals in an industry (in the context of real estate, licensed brokers). Assume that all potential (i.e., those at the entry level) and current professionals are forward-looking.

[21] Barwick and Wong (2019) note that commissions have been fairly constant across both time and geography (at around 6%), are higher in the U.S. than in other countries, and are higher in real estate than in other industries; meanwhile the market share of the largest brokerage firms remains considerably high. See also Barwick and Pathak (2015) and Han and Hong (2011).

3.1.1 Case 1: Perfect Foresight with Unchanging Cost

Consider the case of the broker entry decision when facing a fixed entry cost, F, which is exogenously set by a policymaker or trade association. Let N_t be the total number of brokers in the labor market in period t.

First, let us consider the period-specific wage a broker will make. Let the total number of listings in a market be denoted by X_t and the average price in the market by P_t. Assume that listings are randomly assigned to brokers, such that each broker performs $\frac{X_t}{N_t}$ listings. Further, assume there is some fixed commission rate c.[22] Finally, assume that a share ρ of listings are sold. For now, we will assume this is constant across agents.[23] Thus, each broker earns the same wage, denoted by w_t, where:

$$w_t = \frac{\rho \cdot c \cdot P_t \cdot X_t}{N_t},$$

If a constant interest rate r and a constant exit rate, δ, are assumed, then the discounted present value of working at the professional level in period t can be written as:

$$V_t = \int_t^{\infty} w_s e^{-r(s-t)} \cdot e^{-\delta(s-t)}\, ds,$$

which can be rewritten as:

$$V_t = \rho \cdot c \int_t^{\infty} \frac{P_t X_t}{N_t} e^{-(r+\delta)(s-t)}\, ds \tag{1}$$

Therefore, the value of being a professional evolves according to the schedule:

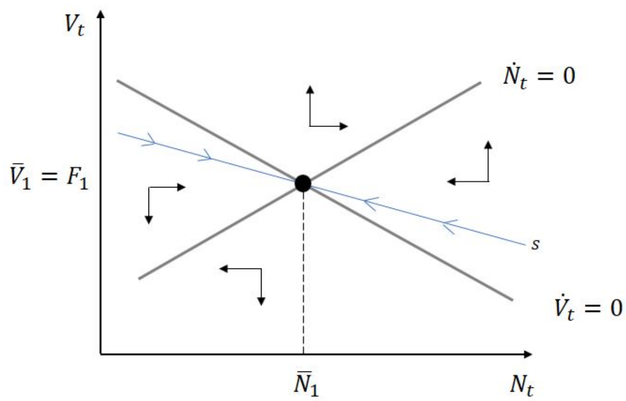
$$\dot{V}_t = (r+\delta)V_t - g(N_t), \tag{2}$$

[22] Some subset of these listings will be performed by that broker's salespeople, for which the broker will earn some share of the commission c. Let's assume this share of listings performed by salespeople is constant across brokers.

[23] Hsieh and Moretti (2003) theorize a fixed number of *sales*, which they denote S. This would be equivalent to $\rho \cdot X_t$ in my setting. I make this change so that I can still maintain that volume is fixed in a period, while relaxing an assumption that agents are homogeneous (here, I can vary ρ across agents).

where $g'(N_t) < 0$. In other words, the value of being licensed as a broker decreases in the number of total brokers, because the transaction volume a broker can attain decreases if there are more competitors. The $\dot{V}_t = 0$ schedule is downward sloping, because any rise in V_t would require the number of workers (N_t) to fall in order to preserve $\dot{V}_t = 0$. Note that the dynamics of V_t are unstable. Holding N_t constant, if $\dot{V}_t > 0$, V_t will rise and if $\dot{V}_t < 0$, V_t will fall.

Figure 1: Fixed Entry Cost Dynamics



The flow of brokers will be determined by the amount of exits in a given period, determined by δ, and the number of entrants.[24] Assume that the rate of broker entry (i.e., upgrading from a salesperson to a broker license) is proportional to the difference between V_t and the entry cost, F. In other words, if there is a larger gap, there will be more entry at the professional level. Therefore, the number of brokers evolves as:

$$\dot{N}_t = -\delta N_t + \gamma(V_t - F) \tag{3}$$

Note that $\dot{N}_t = 0$ schedule is upward sloping. As N_t increases, larger values of V_t will be needed to preserve zero entry. See Figure 1.

[24]In this setup, the assumption of a constant exit rate broadly means that a broker only exits the labor market when they retire die. This is generally supported by the data; in real estate, a license expires if an agent fails to complete the renewal process in a given time frame. While salespeople are more likely to do this, official broker exits are much more uncommon.

Setting $\dot{V}_t = \dot{N}_t = 0$, this model will have a steady-state equilibrium. In this equilibrium, $\bar{V} = F$ (the average lifetime value of being a professional is equal to the fixed cost), and $\bar{N} = f^{-1}((r+\delta)F)$ (the average number of professionals is a function of the discounted fixed cost, following from equations 2 and 3). At any given number of brokers N_t, the corresponding equilibrium V_t is determined by the labor market adjusting along the saddle path s.

3.1.2 Case 2: Unanticipated Increased Entry Cost

Now suppose a policymaker or trade association wants to increase the fixed cost F in an effort to increase the qualifications and quality of the overall pool of brokers.

Suppose it is announced unexpectedly that the entry cost will increase at some future period $t+k$. The dynamics should evolve as follows:

- If there is no change in the initial number of professionals from the time of the announcement to the time of being effective (i.e., from time t to time $t+k$), the lifetime value of being a professional immediately rises with the announcement due to the decreased entry in the future.

- Since F has not yet increased, and V_t immediately increases with the announcement, by equation 3, there will be an immediate entry of professionals.

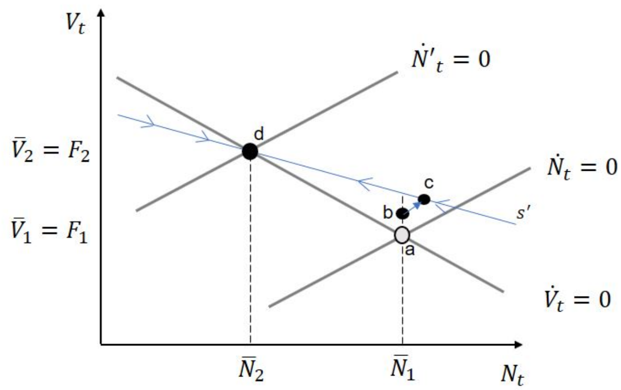
- Because F only increases later at $t+k$, there should be an initial increase in professional entry, followed by a decline at time $t+k$.

The specifics of these dynamics depend on the size of the immediate initial increase in V_t. This new movement is depicted in Figure 2. Two conditions must hold:

1. V and N are governed by the initial system (F_1) through time $t+k$, and the new system (F_2) from $t+k$ onward.

2. There are no other future increases in V that can be anticipated by the labor market participants (i.e., they do not expect any other future changes in requirements).

Figure 2: Unanticipated Increase in Entry Cost



From the initial equilibrium a, when the entry cost is announced to eventually increase from F_1 to F_2, there is an immediate increase in the value of being a professional to some point b. Note that even at point b, the entry path is still governed by the original F_1 because F_1 is still in effect.

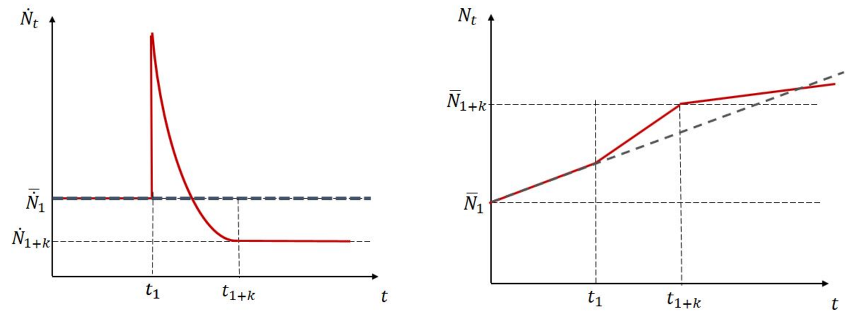
The difference between F_1 and F_2 will determine how close b is to the new point c along the new saddle path s', because, due to condition (1) above, the path must reach c at precisely period $t+k$. Once on the path s', the dynamics are governed by the higher F_2 until the new equilibrium point d is reached.

Figure 3 Panel A displays the flows of brokers under this setting, while Panel B displays the total stock of brokers. The solid red lines represent a market where the increased cost as described occurs; the dashed grey lines represent a market which sees no change in entry cost. The strong increase in flows at t_1 translates to a relatively faster-growing stock between t_1 and t_{1+k}. The stock then grows relatively more slowly after t_{1+k}. Note that the way it has been drawn in Figure 3 Panel B assumes that the new equilibrium flow rate under the new licensing regime is greater than zero. Eventually the stock may be relatively lower at some point in the future determined by the new equilibrium flow rate $\dot{N}_t$ (which is dependent on δ and γ).

As such, a policy of this type is predicted to generate two phases: In the short term, entry and stock will increase from from t_1 to t_{1+k}. In the long term, entry will decrease, and stock will

Figure 3: Flows and Stock of Brokers over Time

(a) Flows (b) Stock



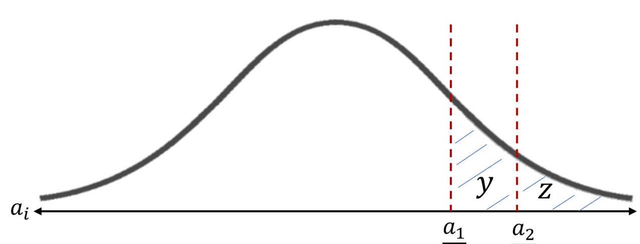
potentially decrease from t_{1+k} onward depending on the parameters δ and γ.

3.2 Quality and Competition

3.2.1 Quality

Now, let us assume that ρ, the share of listings that an agent sells, is a function of an agent's quality, a_i. If an agent has a higher quality, she will sell more listings, and therefore earn more profit; thus, it is more profitable to be a higher quality agent. Thus, for any fixed entry cost F, there is some minimum quality cutoff $\underline{a}$ such that it is profitable for a salesperson to upgrade to a broker. Let us denote this cutoff at the original fixed cost F_1 as $\underline{a_1}$, and the announced future fixed cost F_2 to be $\underline{a_2}$, where $\underline{a_2} \geq \underline{a_1}$. Assume some hypothetical distribution of a_i, as depicted in Figure 4.

Figure 4: Agent Quality Distribution



Note that under the F_1 licensing regime, salespeople from both region y and region z are eligible to enter as a broker, whereas under the higher cost F_2 regime, salespeople from region y will be excluded. As hypothesized above, the announcement will induce a certain number of salespeople to become brokers in the short run.

Since salespeople from region y will be excluded once the cost increases, it is likely to be the case that there will be more entrants from y than from z in the short term, whereas there will only be entrants from z in the long term. Therefore, in the short run, average quality of entering brokers should decrease. In the long run, average quality of entering brokers should increase.

3.2.2 Market Power

Let us measure the concentration in the market at a given time t using a Hirschman-Herfindahl Index (HHI) (see Rhoades (1993) for an overview of the development and applications of the HHI). The index is calculated as:

$$h_t = \sum_{i=1}^{N_t} s_i^2,$$

where s_i is a broker i's share of the total listings. Note that smaller values of h_t represent a market that is *less* concentrated (i.e., exhibiting less market power). In this stylized setting where listings are randomly assigned, s_i is the same for all brokers and decreasing in N_t.

Thus, in the short term, when the policy induces entry, h_t will decrease. In the long term, h_t will still be lower than before the policy announcement, but it will decrease at a slower rate after the policy is effective and stock still increases but at a slower rate.

3.3 Efficiency

I now revisit the idea of efficiency as it is discussed in Hsieh and Moretti (2003). In their framework, the total volume of sales is fixed. Social efficiency is defined by the productivity of the average agent, or $\frac{Sales}{N}$. Thus, anything that increases N decreases social efficiency, because that suggests that there are agents in the real estate industry who could be involved in profitable activities elsewhere.

In contrast, in this setting, total sales ($\rho \cdot X_t$) are not fixed; however, total listings (X_t) are. Therefore, I can separate the concept of sales per broker from listings per broker. While listings per broker should unambiguously decrease when the stock of brokers increases, it is not obvious in this setting that sales per broker (i.e., the preferred measure of broker productivity in this stylized setting) will decrease. This is because is it possible that higher quality brokers (i.e., those with a larger a_i) may close more sales.

3.4 Summary of Predictions

A future increase in the cost of entry for brokers predicts:

In the short term:

1. An increase in broker entry and broker stock;
2. An ambiguous effect of average entering broker quality (dependent on y and z);
3. A decrease in market concentration;
4. An ambiguous effect on efficiency.

And in the long term:

1. A decrease in broker entry and an ambiguous effect on broker stock (dependent on the new equilibrium $\dot{N}_t$);
2. An increase in average entering broker quality;
3. An ambiguous effect on market concentration (dependent on the effect on stock);
4. An ambiguous effect on efficiency.

4 Data and Summary Statistics

4.1 Primary Data Sources

This paper leverages two primary sources of data. The first source is licensing records for real estate agents of both levels in a number of states. Few papers have addressed the distinction between real estate brokers and salespeople, because there is no administrative dataset or standardized central repository of licensees nationwide. While licenses for many industries are generally public record, these records are maintained individually by a state commission. Each state records and maintains professional licenses differently, and thus these data must be collected and cleaned individually state by state. An additional difficulty of compiling these data is that each state has different policies regarding the transition from one license to another, how to record out-of-state licensees, and which dates are relevant to the licensee (e.g., the first time the individual was ever licensed vs. when that current license held was first made active). Therefore, compiling any one state's data requires a large amount of institutional knowledge regarding the state's licensing application process.

Using these public records, I collect a novel dataset of the license information for all real estate licensees, both current and inactive, in Texas, Florida, Louisiana, Ohio, and Connecticut dating back to at least 2000. This is the first paper to my knowledge to study salespeople and broker licensing combining multiple states of records. Having multiple states is beneficial because it is then possible to draw from many counties to create suitable controls for the counties in Texas, all of which were treated by this policy change.

Real estate licenses across states generally identify an individual's type of license (e.g., broker vs. salesperson), when she received it, whether (and when) the license expires, and for whom she works if currently active. These records allow one to identify the type of license an agent has and also reconstruct the employment network of salespeople and their employing brokers for those that are currently active.[25] This also makes it possible to track entry and exit of both tiers of licensees over time.

The second data source is property listings from various Multiple Listing Service (MLS) databases. MLS data includes any information that would show in a property listing such as list price, property characteristics, and crucially, the listing and buying agents' names. The agent name

[25] To my knowledge, no state provides a record of the last employing broker and inactive/expired salesperson worked for, or a history of brokers an agent has worked for. In other words, for all states, the only employment information publicly available is the current broker an active salesperson is working for at that point in time.

connects these listings data with the licensee data. This therefore allows me to match productivity with a licensee and analyze outcomes separately by type.

4.2 Sample Construction

The final analysis sample includes all counties present in the MLS listings from 2009-2019 with at least three listings per quarter for all quarters. Only residential properties for sale are considered. Further, listings with a list price above the 99th or below the 1st percentile for that county-quarter are omitted, as well as listings on the market for more than two years. The sample retains only the counties which have fully populated secondary covariate data for all quarters in the sample period as well. These data include county-by-quarter employment indicators from the Quarterly Census of Employment and Wages (QCEW) and Zillow's Home Value Index (ZHVI).

Table 2 compares the counties in Texas which are used for the analysis sample against the remaining Texas counties which are not included due to data limitations. There are 250 counties in the licensee data (and, since both the Census and QCEW cover a universe, 250 counties in those as well). Note that there are two ways channels through which a county may be excluded from the sample: the county does not appear in the MLS (either because it is not a database accessible through CoreLogic or there are not at least three listings in each quarter of analysis), or the county does not have Zillow ZHVI data. There are 40 counties which have all three sources populated. The biggest restriction is MLS access.

Note that the data in the final analysis sample skew towards larger counties. The 88 counties without Zillow or MLS data have a mean population of less than 10,000 and are about 70% rural; thus, this study is more informative about larger, more urban areas. However, there is still a mix of rural areas in the sample; the average county in the sample is 36% rural. Additionally, housing dynamics such as the share of listings sold and the mean sale price are similar across counties with MLS coverage that are and are not in the final sample.

Matched Listings Sample: Most of the analyses in this paper can be performed with either the licensee or the MLS data alone. However, the MLS listings data do not provide the listing or

Table 2: Balance Table - Means of Key Variables for Texas Counties

Variable	**Source**	**In Final Sample**	**In All 3, Not Enough MLS**	**In MLS, no ZHVI**	**In ZHVI, No MLS**	**No ZHVI, No MLS**
Total Listings	MLS	1,016.78	41.00	83.38	.	.
Share Sold	MLS	0.54	0.44	0.49	.	.
Mean Sale Price	MLS	153,259.30	115,235.20	100,495.40	.	.
ZHVI	Zillow	123,515.70	86,929.59	.	95,762.18	.
Stock of Brokers	License	471.15	33.00	28.63	64.86	5.75
Stock of Salespeople	License	1,697.60	95.15	98.63	233.91	14.67
Earnings	QCEW	3,456.15	3,228.05	2,916.88	3,319.06	3,323.55
Employment	QCEW	152,294.10	14,026.80	8,157.38	22,452.57	2,090.69
Population	Census	469,554.20	66,823.90	42,985.63	83,985.64	8,860.09
Pct Rural	Census	36.46	63.13	73.82	46.36	69.49
N	**250**	**40**	**20**	**8**	**94**	**88**

Note: Data is shown for 2011Q4. There are 250 total counties in the public record licensee data (and, since QCEW and the Census cover a universe, all 250 of those counties have these variables as well). The column "In Final Sample" represents the 40 Texas counties used in my final sample of analysis (that have both MLS and Zillow ZHVI coverage). "In All 3, Not Enough MLS" refers to counties that have Licensee, MLS, and Zillow data but do not have enough MLS coverage either in terms of number of listings or times periods. "In MLS, No ZHVI" has MLS data but not Zillow ZHVI index, while "In ZHVI, no MLS" are the counties without MLS data. "No ZHVI, No MLS" has neither Zillow nor listing data. Note that for the "Mean Sale Price" row, all sale prices are winsorized to the 3rd and 97th percentile for a year-quarter across all states.

buying agent's type of license. Therefore, to connect listing volume and productivity to the type of license, the licensing records must be matched to the listings data using the agent's name (and other identifying information such as the county they are licensed in).[26] There are many reasons to expect that not all of the licensed agents will be found in a record in the MLS. Primarily, the MLS data do not cover entire states.[27] Further, many agents will obtain a license but then never actually perform a listing, while some agents work exclusively with rentals, commercial, timeshares, etc., which are all excluded from the sample.[28]

[26]To be sure the agents I am analyzing are cleanly matched to the office for which they work, I further drop agents with listings attached to multiple office identification numbers in the same quarter. This restriction amounts to about 11% agents and 35% listings.

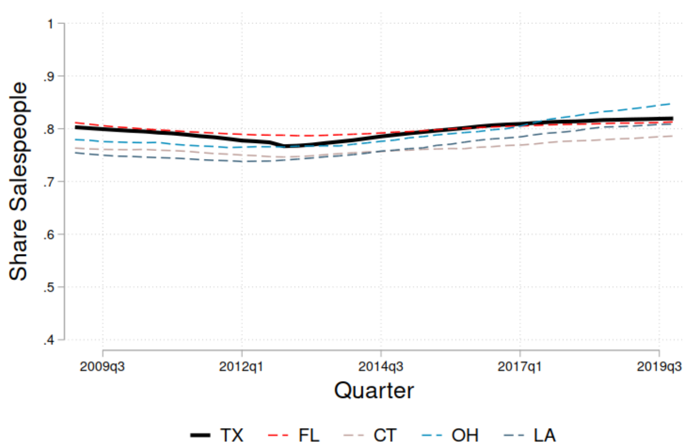
[27]For instance, in my data there are six of the MLS systems in Texas. The two largest MLS'es in Texas, the Houston Association of Realtors (HAR) MLS, which covers the larger Houston area, and the North Texas Real Estate Info Systems (NTREIS) MLS, which covers the larger Dallas area, are both in my data. By way of comparison, I consider the CoreLogic MLS data in relation to the Texas A&M (TAMU) Real Estate Center Data, which provides county-level aggregate housing market measures for over 50 MLS systems in Texas. In Appendix A Figure A1 I compare the total number of sold listings according to the CoreLogic MLS data with the total number of sold listings in the TAMU data in the year 2017 by county. Bluer counties represent counties in which the CoreLogic data, which I use, have more listings. Of note, the CoreLogic data have better coverage in the Dallas and Houston areas, while coverage is lacking in the greater Austin area.

[28]Appendix B Table 1 compares the listings in the MLS data in the sample counties to the listings which ultimately

4.3 Summary Statistics

These novel licensing data illustrate that the share of licensees that are salespeople is consistent both over time and geography at around 80% (see Figure 5). The thick black line represents the share of salespeople in Texas, while the remaining dashed lines represent the other states in the licensee data. Note that across the four states the share is quite similar, ranging only from about 75% to 80%. Further, this share has stayed consistent across all states over the last decade. While the number of overall agents has been rapidly rising, the overall composition of the industry has remained fairly constant.

Figure 5: Salesperson Share of Agents



Note: Figure reports the share of all licensed agents in each state who have a salesperson's license.

The ratio of salespeople-to-brokers is fairly consistent across geography despite the overall stock of both types relative to the population varying across states. Table 3 reports the mean stock and entry of brokers and salespeople per 1,000 residents across counties in each of the states in the sample. Note that Florida has the largest real estate labor market with over seven salespeople and two brokers per 1,000 people, even though Texas has a larger population. The smallest of the states, Connecticut, falls in the middle with just over one broker and almost four salespeople per 1,000. In the empirical work, I will show that once controlling for various county and time trends,

end up matched to agents licensed in those counties.

the stock of both brokers and salespeople is indistinguishable across states in the years before the TREC policy announcement, suggesting this heterogeneity in the number of workers is largely due to heterogeneity in housing markets. Despite these differences, the mean salesperson-to-broker ratio ranges only from 2.9 to 3.6.

Table 3: 2011Q4 Means per 1,000 Residents

State	New Brokers	Stock Brokers	New Salespeople	Stock Salespeople	S:B Ratio	n
CT	0.005	1.269	0.042	3.719	2.933	8
FL	0.022	2.221	0.104	7.322	3.269	43
LA	0.004	0.811	0.023	2.576	3.532	7
OH	0.001	0.420	0.037	1.371	3.459	63
TX	0.012	0.881	0.051	3.121	3.600	40

Note: Means are for 2011Q4 for sample counties in all states. All numbers are per 1,000 county residents.

5 Empirical Methods

5.1 Event Study Specification

To test the hypothesis of an anticipatory entry increase due to the future increased barrier, I first utilize an event study framework which compares counties in Texas with control counties in the other states before and after the announcement of the future increase in broker licensing cost. I estimate the following equation:

$$Y_{jt} = \alpha + \sum \beta_{1t} Quarter_t * TEXAS_j + X_{jt} + \pi_{jt} + \lambda_j + \gamma_t + \varepsilon_{jt} \tag{4}$$

In this setting, Y_{jt} is an outcome in county j in quarter t. $Quarter_t$ represents dummy variables for year-quarters from 2009Q1 to 2019Q4. $TEXAS_j$ is an indicator equal to 1 if county j is in Texas, and was therefore treated by the policy announcement. λ_j are county fixed effects and γ_t are year-quarter fixed effects.

To address seasonality most prevalent in specifications using listings data as outcomes, the specification includes state-by-quarter of year fixed effects (π_{jt}). In order to leverage these fixed effects maximally, four $Quarter_t$ dummies are omitted from the specification, namely, quarters one through four of 2011 (the policy was announced in 2012Q1). Additionally, the γ_t fixed effect for quarter four of 2011 is omitted. Thus, coefficients can be thought of as relative to the average over the year before treatment.

X_{jt} represents a vector of county-by-quarter characteristics. These include controls for employment dynamics using the Quarterly Census of Employment and Wages (QCEW) from the Census. These variables include quarterly hirings, separations, total employment, and total earnings at the county-level. I also control for house price dynamics using the Zillow Home Value Index (ZHVI). Finally, I construct measures of housing market dynamics using the MLS data; these include the total listings sold in a county-quarter, the share of all listings sold, and the median days on the market for sold listings.[29] The specification includes one-quarter lagged values of these housing market variables in all specifications, unless noted. All specifications are weighted by a county's population in 2010 using the Census and standard errors are clustered at the state level.

The event study approach allows me to establish the two phases of variation in employment predicted by the conceptual framework in Section 3, as it generates a point estimate for each quarter individually. However, a possible concern is that counties in the four control states are not suitable controls when given equal weighting. Appendix A Figure A3 displays data from the counties in each state separately across four dimensions. Panels A and B plot the mean earnings as reported by the QCEW across all counties per quarter, and the median ZHVI. Panels C and D plot the total (i.e., all counties added together) stock of brokers and salespeople, respectively, in the five sample states. Broadly speaking, the five states on aggregate appear to evolve similarly in the quarters before the policy announcement. However, along the dimensions of home values and broker employment, certain states exhibit different growth rates.

Thus, to estimate the short- and long-term consequences for quality, concentration, and efficiency, I employ a synthetic difference-in-difference design which allows me to compares counties

[29]Before calculating the median, the days on the market for all listings is winsorized to the 3rd and 97th percentile of a year-quarter across all counties in the sample.

in Texas against counties across the four untreated states while putting more weight on untreated counties which evolve most similarly to Texas counties.

5.2 Synthetic Difference-in-Difference

The synthetic difference-in-difference approach, put forth by Arkhangelsky et al. (2021), combines tools from both difference-in-difference and synthetic controls to provide causal estimates in settings with multiple treated units but a potentially insufficient control group. The advantage of the synthetic difference-in-difference (herein SDiD) approach over a standard difference-in-difference is that it is easier to restore the parallel trends assumption with this re-weighting. The advantage over synthetic controls is that the pre-trends do not need to match exactly; it just requires parallel trends after re-weighting. Additionally, this approach allows for multiple, as opposed to just one, treated unit as in synthetic controls. This is particularly helpful for leveraging variation in housing markets across counties and for studying heterogeneous effects across different types of counties.

The basic idea behind a SDiD design is to assign both unit and time period weights to non-treated units to better match the treated units. Therefore, more weight will be put in the control group on time periods in which the treated and non-treated counties are more similar and on the non-treated counties which are more similar to Texas counties along the dimensions of housing and employment dynamics. These weights are algorithmically selected based on the pre-treatment values of the outcome variable and selected inputs. I use the covariates described above in vector X_{jt} of Equation 4.

The SDiD approach assumes a balanced panel with N units and T time periods. The first N_{co} control units are never treated, while the last N_{tr} units are treated after time T_{pre}. Let Y_{jt} be an outcome Y for county j in quarter t and W_{jt} denote the binary treatment exposure. Further, let α_j be a unit fixed effect and β_t be a time fixed effect. SDiD uses both unit weights to align pre-exposure trends in the outcome of untreated units with treated units (as in synthetic controls) and time weights to balance pre-exposure time periods with post-exposure time periods. Denote the unit weights $\omega^{\hat{sdid}}$ and time weights $\lambda_t^{\hat{sdid}}$.

The SDiD approach can be thought of an alternative two-way fixed effect regression to estimate the causal effect of exposure to some treatment. Denoting this effect as τ, the estimator can be written as follows:

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \arg\min_{\tau,\mu,\alpha,\beta} \{\sum_{j=1}^{N}\sum_{t=1}^{T}(Y_{jt} - \mu - \alpha_j - \beta_t - W_{jt}\tau)^2 \hat{\omega_j^{sdid}} \hat{\lambda_t^{sdid}}\}$$

By comparison, a standard difference-in-difference estimator is:

$$(\hat{\tau}^{did}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \arg\min_{\tau,\mu,\alpha,\beta} \{\sum_{j=1}^{N}\sum_{t=1}^{T}(Y_{jt} - \mu - \alpha_j - \beta_t - W_{jt}\tau)^2\}$$

These two estimators are similar, with the exception that the difference-in-difference implicitly uses unit and time weights equal to one. Finally, a synthetic controls estimator is:

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\beta}) = \arg\min_{\tau,\mu,\beta} \{\sum_{j=1}^{N}\sum_{t=1}^{T}(Y_{jt} - \mu - \beta_t - W_{jt}\tau)^2 \hat{\omega_j^{sc}}\}$$

While the synthetic controls approach does use the unit weights, it does not use unit fixed effects. The unit fixed effects allow for the flexibility of parallel trends in the pre-period, as opposed to identical matching. Further, the synthetic controls estimator does not utilize time weights either.[30]

To construct an estimate that is as parallel to the event study (and standard difference-in-difference) coefficients as possible, I begin by regressing an outcome variable on all of the controls as above in Equation 4 and predict residuals, as following:

$$Y_{jt} = \alpha + X_{jt} + \pi_{jt} + \lambda_j + \gamma_t + \varepsilon_{jt}. \tag{5}$$

I then use the covariates in vector X_{jt} to construct the unit weights for the synthetic difference-in-difference, with the residualized Y_{jt} (i.e., ε_{jt}) as the outcome variable.

A caveat to the SDiD approach is that it requires a balanced panel. In this context, in areas

[30] See Appendix C for construction of the unit and time weights.

where there may not be any broker entry or listing outcomes in a given quarter, that county would not be included at all in the estimation. Thus, for all results, I will report in the Appendix the corresponding standard difference-in-difference estimates using all county-quarters available in the sample (i.e., not forcing a balanced sample).

6 Results

6.1 Effects of Policy Change on Entry

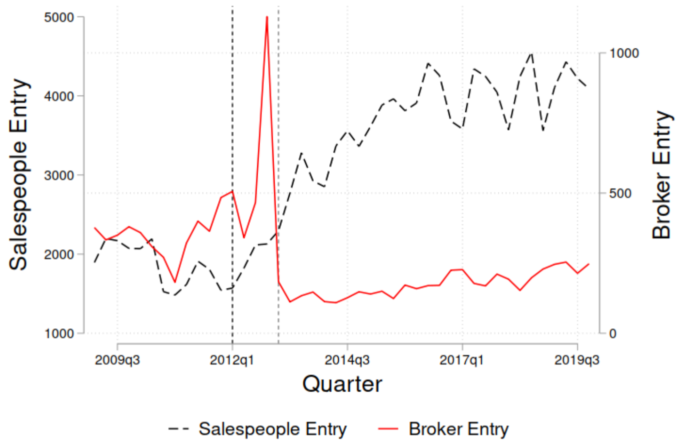
Recall that the conceptual framework described in Section 3 hypothesizes that when there is a future increase in the entry cost, there should be an anticipatory effect in which there is a large influx of entry of licensees in the short term.

The TREC policy change was quite salient to potential brokers when it was announced. Figure 6 plots the number of entering salespeople in the dashed black line (with the scale on the left-hand y-axis) and brokers in the solid red line (with the scale on the right-hand y-axis). While the entry of salespeople tracks a fairly smooth path over this time, there are three distinct phases of broker entry. First is a period of consistent entry of about 200-300 new brokers per quarter, then a sharp increase of about 800 new brokers in the quarter before the more stringent broker licensing is set to become effective, and then many quarters of depressed entry of less than 200 new entrants per quarter in the six-plus years following the policy change. Thus, while the policy does not appear to have an obvious or immediate effect on the decisions of potential salespeople, it is evidently quite relevant to potential brokers.

To establish a sense of the effects of this policy change on the career trajectories of prospective brokers, I compare the cohorts of salespeople who took advantage of "grandfathering" into the cheaper licensing with those who did not hold a salesperson's license for long enough at the announcement to immediately apply for a broker license.

Recall that prior to the policy, a salesperson only needed to hold a salesperson's license for two years to be eligible to upgrade to a broker, while after the policy was effective the salesperson

Figure 6: TX Quarterly Entry



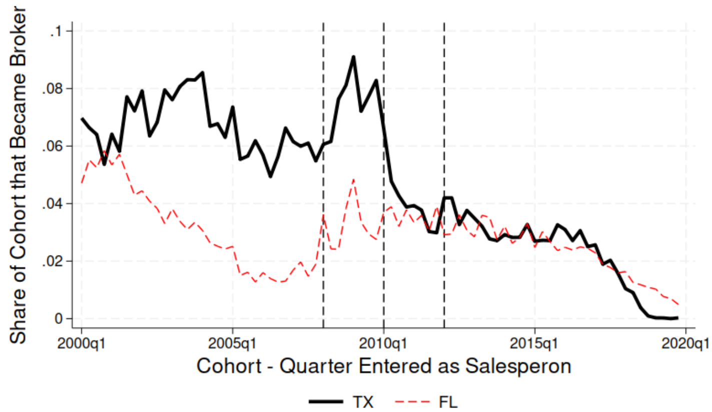
Note: Figure reports the number of newly licensed agents per year in Texas by agent type. The first black vertical line represents the policy announcement, and the second grey vertical line represents the policy effective date.

would need wait four years. Thus, any agent with a salesperson's license for two to four years when the policy was announced was capable of quickly upgrading to a broker before the requirement became too stringent. Conversely, any salesperson with a license for zero to two years was unable to quickly upgrade, therefore needing to wait the full four years at minimum. I call the two-to-four years cohort the "grandfather-eligible" cohort, and the zero-to-two years cohort the "grandfather-ineligible" cohort.

Figure 7 plots the share of each entry cohort in Texas (where a cohort is defined as all of the agents being licensed as salespeople, i.e., beginning their real estate career, in a quarter) that eventually upgraded to a broker within six years of entering the industry. In the red dashed line, I also plot this for salesperson cohorts in Florida, where this is no change in policy, for comparison.

There is a noticeably higher share of the grandfather-eligible cohort upgrading to brokers compared to the earlier cohorts, even though both groups faced the same similarly minimal requirements, suggesting that the policy did indeed encourage more people to become brokers than might have absent any change. Furthermore, while the share was constant between 4% to 8% for all cohorts prior to 2010, the share drops considerably for cohorts beginning in 2010, who would be grandfather-ineligible. This pattern suggests that the new policy, once binding, had the effect of

Figure 7: Share Broker within Six Years



Note: Figure displays the total share of each salesperson entry cohort that ever upgraded to a broker within six years of entry. A cohort is defined as all the agents in a given quarter who entered as a salesperson. The first red vertical line represents cohorts who would have four years of salesperson experience when the TREC policy change would go effective. The second line represents cohorts who would have two years of experience. The third line is when the policy when effective.

preventing people who might have otherwise become brokers absent the change from doing so eventually.

Those who are in the grandfather-eligible cohorts were also less likely to become a broker after the initial grandfathering period. Appendix A Figure A2 plots the density of the number of brokers entering. Panel A displays the quarter of entry for all brokers who entered the industry as a salesperson from 2008-2010, while Panel B represents those who entered as a salesperson from 2010-2012. The overwhelming majority of the grandfather-eligible cohort upgraded to broker before 2012, with very few remaining becoming brokers after. In other words, most people in that cohort who became a broker either did it before the more stringent rules were in place or not at all.[31]

[31] On the contrary, as shown in Panel B of Appendix A Figure A2, those who are grandfather-ineligible become a broker at a much smoother distribution. The largest mass occurs four years after the policy change, but overall there is a more equal distribution of years of experience before upgrading to a broker for this two-year cohort.

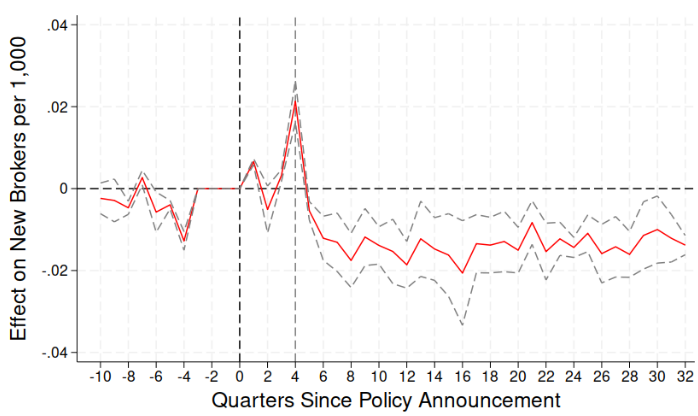
6.2 Broker Entry

To formalize these patterns, I first use the event study specification described above, because, as noted, it generates a point estimate for each quarter individually. This allows me to discuss the evolution of effects of the policy on entry over time.

I begin by showing that a future increase in professional licensing cost increases the current stock of employees at the professional level in the short term. In the long term, this effect reverses due to restricted entry. Figure 8 displays event study coefficients for estimating Equation 4 with the total number of licensed brokers per 1,000 county residents as the outcome.

As the conceptual framework predicts, a future increase in entry cost induces a great deal of entry in the short term. The quarter the higher entry cost was set to go into effect saw an increase of 0.02 brokers per 1,000 residents in Texas counties, which is double the mean entry in the year before the policy was announced. As soon as the policy is effective, however, entry decreases relative to the pre-period, and still eight years later does not return to pre-period entry levels. This supply restriction is similarly about double the mean entry in Texas counties in the year before.

Figure 8: New Brokers per 1,000 Residents

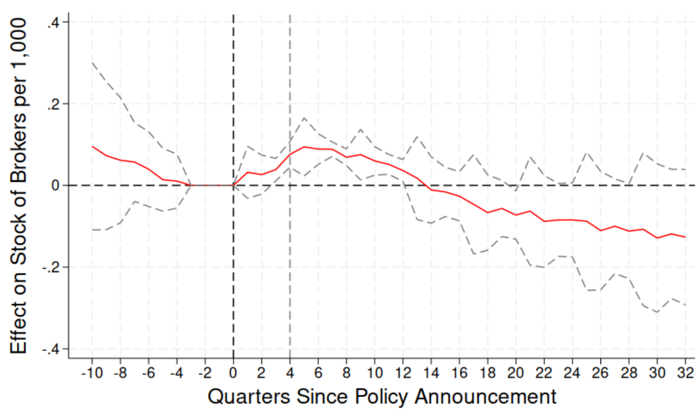


TX Pre-Year Mean: 0.011
DiD Coeff: -0.0075*

Note: Figure displays event study coefficients estimating Equation 4. The outcome variable is the total number of entering brokers per 1,000 county residents. The red solid line represents the coefficients, while the dashed gray lines represent the 95% confidence interval. The first black vertical line represents the policy announcement, while the second grey vertical line represents the policy effective date. This result uses licensee data in sample counties only.

While the intended effect of the policy may have been to decrease entry, thereby ultimately restricting supply, the initial inducement of entry leads to a larger stock of brokers in the short term (see results displayed in Figure 9). This effect lasts nearly three years.[32] This initial increase in stock is non-trivial. At its highest point about five quarters after announcement, the stock of brokers increased by about 0.1 per 1,000 residents. This is an 11.5% increase over the pre-period quarterly mean across Texas counties of 0.865 brokers per 1,000 people. For context, the average Texas county in the sample has a population of 469,554 in 2010. This would mean an increase from about 406 brokers to 453 brokers in the average county.

Figure 9: Stock of Brokers per 1,000 Residents



TX Pre-Year Mean: 0.865
DiD Coeff: -0.0427

Note: Figure displays event study coefficients estimating Equation 4. The outcome variable is the total stock of brokers per 1,000 county residents. The red solid line represents the coefficients, while the dashed gray lines represent the 95% confidence interval. The first black vertical line represents the policy announcement, while the second grey vertical line represents the policy effective date. This result uses licensee data in sample counties only.

In the long term, however, the intended effect of the policy becomes evident. Starting about twelve quarters after the policy was announced, the total stock of brokers on average in Texas counties is lower relative to the control counties. This effect is persistent all the way through 2020. While the changing barrier at the professional level induced an increased stock of brokers, it did not change employment at the entry level for salespeople. Appendix A Figure A5 displays event

[32] In Appendix A Figure A4, I consider the effect of the policy on broker exits which evidently is minimal. This is sensible given that renewal requirements are generally not costly relative to the entry requirements.

study coefficients for the employment dynamics of salespeople. Panel A considers entry while Panel B considers stock. While on average, entry, and therefore stock, is decreasing over time, we cannot reject that this decrease is significantly different relative to counties in the other states without a policy change.

The policy also creates two distinct phases of labor market composition. Because this is an apprenticeship industry, it is important to consider how the overall workforce composition is impacted by licensing requirements. My primary measure of workforce composition is the salesperson-to-broker ratio (more broadly, the apprentice-to-professional ratio). Figure 10 displays estimating results for the evolution of the county-level licensed salesperson-to-broker ratio. In the short term, the changing licensing cost decreases the amount of salespeople per brokers by about one-quarter.

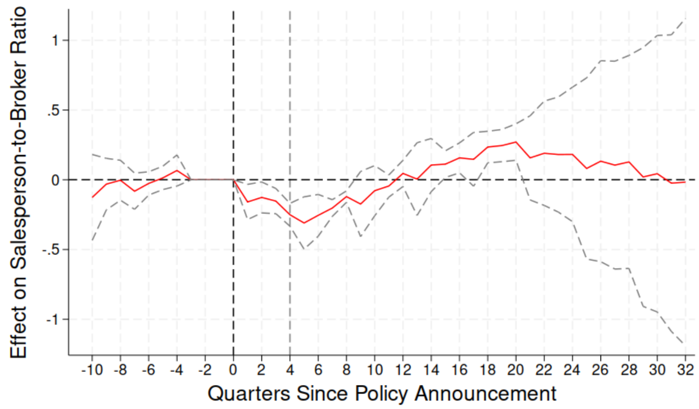
In the long term, however, the labor market becomes more concentrated in Texas counties relative to control markets. Five years after the policy is effective (quarter 20), the salesperson-to-broker ratio in Texas counties increased by about 0.25 from a pre-policy mean of 3.67, amounting to a 7% increase in the amount of salespeople "under" brokers in the employment structure.

These patterns are robust, though less precisely estimated, when considering the employment of *active* agents; i.e., salespeople and brokers who are identifiable in the MLS data. Appendix A Figure A6 utilizes the matched listings sample to consider the number of brokers and salespeople that are attached to a listing in a given quarter. First note that there is a much smaller number of brokers used per 1,000 residents than there are brokers licensed. The mean in Texas counties the year before the policy was announced is 0.165 brokers per 1,000.[33] The number of brokers identifiable in the MLS data does generally trend upwards in Texas counties relative to the control counties for about three years, decreasing by about double the pre-period mean four years out.

It is evident that the changing entry costs are salient and have a measurable impact on employment choices and labor force composition. Also evident is that the decision to announce the policy

[33]Note that, as described in Section 4, this will be an under count of total brokers, as these represent only the brokers on listings who could be matched to their Texas licensing record. Further, this does not account for brokers who may be overseeing the listing process in a managerial sense while not actually being named in the listing itself.

Figure 10: Salesperson-to-Broker Ratio



TX Pre-Year Mean: 3.670
DiD Coeff: 0.019
Note: Figure displays event study coefficients estimating Equation 4. The outcome variable is the total number of licensed salespeople divided by the total number of licensed brokers in a county. The red solid line represents the coefficients, while the dashed gray lines represent the 95% confidence interval. The first black vertical line represents the policy announcement, while the second grey vertical line represents the policy effective date. This result uses licensee data in sample counties only.

a year before its effective date generates two phases: an initial *increase* in broker entry before a decrease in the long term. Thus, the question becomes whether this delayed effect is strong enough to outweigh the inducement of brokers generated by the policy announcement.

Because of these two distinct phases of employment trends as a result of the varying entry cost, I will herein investigate separately short- and long-term effects on quality and market concentration. "Short term" will be considered as quarters 0-12, when the stock of brokers in Texas counties is increasing relative to other counties, and "long term" is quarters 12-32, when the stock ceases to increase. I shift to the SDiD approach to generate causal point estimates for both phases. In Table 4, I re-estimate these employment effects using the SDiD for consistency.

Note that the effects on broker entry (Column 1) are negative in the short term despite the entry influx because this new definition of "short term" includes the announcement period *and* two years after, when entry was already being suppressed. For this reason, for all results in the paper I also report the analogous results re-defining the "short term" as only quarters 0-4, when entry was higher in Texas counties before the policy was effective, and the "long term" as quarters

5-32, when the policy was effective and entry was relatively lower. These results are all reported in Appendix Section B.1.

Table 4: Entry and Stock per 1,000 County Residents

	(1) New Brokers	(2) Stock Brokers	(3) New Salespeople	(4) Stock Salespeople
(1) Short Term	-0.003***	0.060***	0.000	0.175***
	(0.001)	(0.009)	(0.007)	(0.045)
(2) Long Term	-0.009***	0.063**	-0.019*	0.469**
	(0.001)	(0.030)	(0.010)	(0.194)
TX Pre-Year Mean	0.011	0.865	0.054	3.118
N Total	8480	8480	8480	8480

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Note that the outcome variables are all measured per 1,000 county residents. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information.

As expected, there is an increased stock of brokers (Column 2) in the short term, which levels out and gets no larger in the long term due to the decreased entry. Also note that the evidence from the SDID framework, which provides a better suited control group, suggests that the policy leads to a larger stock of salespeople both in the short and long term.[34]

6.3 Broker Quality vs. Market Concentration

The evidence above illustrates that a salient future increase in the cost of entry leads to an increase in the stock of brokers in the short term, without affecting employment at the entry level. This suggests that future brokers are forward looking while future salespeople are not as much. Further, the event study also illustrates the two different phases of variation the policy generates (short-term cheap entry and long-term costly entry).

Recall that if the announcement of a higher entry cost induces more entry from agents with a quality that would only be profitable under the old regime, the conceptual framework in Section 3

[34]Note that in Appendix B.2 Table 1, only the effects of the policy on broker entry are significant when running a naive difference-in-difference.

hypothesizes that in the *short term*, the average quality of entering brokers should decrease, and in the *long term*, the average quality of entering brokers should increase.

Further, the short-term anticipatory effect should lead to a short-term decrease in market concentration, which will continue to decrease in the long term assuming the flow of brokers is still positive in equilibrium under the new regime. I turn now to considering this trade-off.

6.3.1 Broker Quality

The TREC policy being studied is unique in that, unlike many licensing policies which aim to improve the quality of entry-level laborers, it instead focuses on the quality of managers. By increasing the entry barrier at the professional level, the policy potentially changes who is a desirable supervisor and trainer. Therefore, it is of interest whether brokers are better quality themselves, specifically at the time they become brokers and are eligible to manage, and whether they are better at training salespeople.

There is no consensus in the literature, however, regarding how to measure the "quality" of an agent.[35] As discussed, prior studies of licensing in real estate have used data on formally filed complaints against agents in order to quantify the number of "bad" agents. However, because these complaints are filed by a presumably wronged client, they are quite subjective, highly selected, and likely to capture only extreme wrongdoing.[36]

To quantify quality in a way that is more objective than complaints and less subject to selection, I turn to the match of licensees to listings output, which prior papers could not leverage. It can reasonably be assumed that a primary reason that a home seller would hire an agent is to, on the extensive margin, successfully sell the home, and, on the intensive margin, sell it for a higher price and at a quicker pace. Therefore, to calculate entering broker quality, I consider two sets of

[35]There is similarly little consensus in the literature over what specifically is the primary function of a real estate agent, broadly speaking, in a transaction. For a thorough review of the micro-structure of housing markets, including the role of the agent in transactions, see Han and Strange (2015). Empirically, Aiello et al. (2022) use exogenous variation in the likelihood of agent attention to show that the primary function of agents is to facilitate search (as opposed to provide information).

[36]These complaints are also rare. For instance, in all of Dallas in calendar year 2022, there were seven total complaints filed with TREC.

measures; probability of sale (extensive margin) and sale price and time to sell (intensive margin). Specifically, I calculate these measures over the four year window *prior* to entering as a broker, in order to capture the quality of brokers when they are immediately eligible to supervise and train salespeople. Broadly, a justification for licensing is to ensure that consumers in any industry have a better experience in their relevant transaction. Thus, these measures capture broker quality to the extent that they capture a broker performing better on behalf of their client.

To measure the probability of sale, for each broker, I calculate the overall share of their sale-side listings which result in a sale in the four years prior to becoming a broker. I similarly calculate the share that sell in less than 30 days, and the share that sell in less than 90 days. I then take the mean of these shares across all brokers entering in a given entry cohort, such that results should capture whether upgrading brokers are better quality after the policy than cohorts who upgraded before. Results are displayed in Table 5. As noted above, the SDiD approach requires a balanced panel. However, in many counties, there are a handful of quarters in which there are zero brokers entering, or none who have enough data over the prior four years. Thus, for this estimation in particular, too much information is lost by using the SDiD estimation. Hence, in this section only I report standard difference-in-difference estimates with the caveat that the sample panel is unbalanced.[37]

I find that the policy induces entry of brokers who are less adept at selling homes quickly. Column 3 of Table 5 displays that entering broker cohorts in the short term after the policy have a lower average share of homes that sell within 90 days. However, these brokers are generally no less worse at selling homes at all; there is no significant difference in the overall share sold in the short term.

In the long term, entering broker cohorts after the policy outperform those entering before the policy along the extensive margin; they sell a higher share of homes overall and a higher share within 90 days (though in neither the short nor long term do entering brokers get better at selling homes most quickly, i.e., within 30 days). Broker cohorts entering in the long term have an average

[37]As noted, for all other results in the paper using the SDiD approach, I also report the standard difference in difference coefficients in Appendix B.2, such that the results in this section are the only results for which only one of the two methods is reported, by necessity.

share sold of 0.06 percentage points higher relative to entering brokers in control counties, which amounts to about 15% of the mean of cohorts entering in the year before the policy in Texas.

Table 5: Broker Quality - Extensive Margin

	(1) Share Sold	(2) Share Sold $<$ 30 Days	(3) Share Sold $<$ 90 Days
(1) Short Term	0.008	-0.038	-0.098***
	(0.034)	(0.024)	(0.014)
(2) Long Term	0.064*	0.031	0.121***
	(0.029)	(0.023)	(0.010)
TX Pre-Year Mean	0.433	0.125	0.289
N Total	2276	2276	2276

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). All outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Results using the alternative short and long term are reported in Appendix B.1 Table 2. This result uses the matched listing-to-licensee sample.

While brokers may be improving at securing a sale in the long term, this is only beneficial (i.e., of good quality) to the consumer to the extent that the consumer is satisfied with the price and duration of the sale. To capture the ability of brokers to generate betters returns on listings in terms of price and speed, I utilize the sample of listings matched to a licensee. I estimate a hedonic model for both price and time to sale by regressing the natural log of the sale price or the days on the market before the sale (DOM) for a given listing on a number of property characteristics, as well county and year-quarter fixed effects.[38] The residual of this regression can then be considered as the additional unexpected "return to the listing" (in terms of price or time). The residual is then averaged over all listings performed by brokers from the same entering cohort in the four years prior to broker licensing. This measure in turn captures whether listings performed by brokers from a given cohort are generating different returns from the listings performed by salespeople or brokers who upgraded at a different time. Results are displayed in Table 6.

While entering brokers may be more adept at securing a sale, they are doing so at lower sale prices and a slower pace. In the short term, the average unexpected return on a listing for the log

[38] Characteristics include the square footage of the property, total living area, the year the property was constructed, total number of bathrooms, total number of bedrooms, and indicators for whether the property has a fireplace, a pool, and a garage.

of price decreases and for the log of days on the market increases (i.e., worsens). This means that broker cohorts entering in Texas counties in the short term after the policy are of worse quality compared to broker cohorts that entered in the years before, relative to cohorts in the untreated counties. In the long term, the effect on DOM goes away, while broker cohorts entering are still generating lower sale prices. These results could potentially be explained by the incentives of the policy; because the policy institutes a minimum transactions requirement, salespeople intending to upgrade may be inclined to use smaller sale prices and wait longer to secure a sale in order to hit the minimum.[39] This is supported by the results in Appendix B.1 Table 3, which replicates these results using the alternative definition of short and long term. When the short term is defined as only the four quarters between policy announcement and policy effective date, where there is no need to complete a certain amount of transactions, the entering cohorts do not generate worse returns to the listings in terms of price.[40]

I conclude that brokers are not of better quality due to the policy because consumers are no better off. This is due to the fact that, to the extent that consumers want to optimize successfully selling a home at a high price in a small amount of days, to gain on one dimension, they must sacrifice on at least one other. In this context, while broker cohorts sell a higher share of homes in the long term, they are doing so by generating lower returns in terms of prices.

Even though evidence suggests that new brokers after the policy are not of higher quality along the dimension of listing outcomes, it may still be the case that they are better managers than before. However, I find no effects of the policy on the quality of entering salespeople either in the short

[39] Appendix B Table 2 displays results testing whether the transactions requirement is indeed binding. In Column 1, I calculate the average total number of sale-side listings in the four years before broker for all brokers in an entering cohort. As expected, there is no effect in the short term (when the policy does not yet require a minimum number of transactions), and a strong positive effect in the long term. Column 2 calculates the total *sold* listings, which counts as a "transaction" by the policy's definition, and this is again positive in the long term. Columns 3 and 4 calculate sale-side and buy-side listings separately, suggesting that potential brokers attempt to meet the requirement by representing more sellers.

[40] Appendix B Table 3 displays these intensive margin results calculates for broker cohorts over the four years *after* upgrading. This measure captures whether the costlier licensing requirements led to future improvements in quality. I similarly find that in both the short and long term, entering broker cohorts generate lower returns on prices in their first four years as brokers relative to entering cohorts in control counties (see Column 1). However, the evidence in Column 2 suggests that the entering cohorts in the short term are able to generate quicker sales in their initial years as brokers.

Table 6: Broker Quality Intensive Margin

	(1) Mean Return ln(Price)	(2) Mean Return ln(DOM)
(1) Short Term	-0.185*	0.359***
	(0.077)	(0.039)
(2) Long Term	-0.152**	-0.124
	(0.051)	(0.062)
TX Pre-Year Mean	0.137	0.020
N	753	740

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). All outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Results using the alternative short and long term are reported in Appendix B.1 Table 3. This result uses the matched listing-to-licensee sample.

term or long term. To consider this effect for salespeople, I look at the analogous quality measures for salespeople in their first four years *after* being licensed. This measure is intended to capture the early-career training newly licensed salespeople are receiving from their employing brokers who were licensed under the stricter regime. Results are displayed in Appendix B Table 4.

To the extent that licensing policies are passed in order to provide a better experience to the clients of practitioners, I do not find evidence this is occurring. Regarding the predictions of the conceptual framework, this result suggests that the announcement does indeed induce more brokers of a lower quality, while in the long term there are not enough higher quality entering brokers to counteract this. However, the other reason insiders often desire licensing policies is to provide more protection for incumbents, and in turn, market power. Thus, I turn now to considering the effects of the changing entry cost on market concentration.

6.3.2 Market Concentration

Recall that, if the effect of the induced entry outweighs the effect of the long-term restricted entry, the industry should see a decrease in concentration. To test for the effect on market concentration, I calculate a Hirschman-Herfindahl Index (HHI) to quantify the concentration of listings across brokerage offices, as describe above in Section 3.

The HHI for a given county-quarter is calculated as follows:

$$HHI = \sum_i m_i^2,$$

where i indexes a listing office as denoted in the MLS, and m_i represents that office's share of the total listings in the county that quarter. Higher values of the HHI indicate a less concentrated, more monopolistic market.

In the short term, I find that brokerage market concentration decreases, an effect that is driven by markets with the smallest number of brokers to begin with. I also find that this effect accelerates in the long term, again seeing no change in the larger markets (as defined by number of pre-existing brokers). Results are displayed in Table 7. In Column 1, the effect is estimated on all counties in the sample. Note that the brokerage market in Texas was already quite unconcentrated at the county level even before this policy was announced; an HHI below 0.15 is generally considered unconcentrated, and the mean HHI across Texas counties in 2011 was 0.08. However, due to the influx of new brokers, this reduced in the short term by 0.008, a decrease of 10%. Also notable is that this is a sustained effect; even in the long term once entry is costlier for brokers, the county-level HHI is still lower in Texas counties relative to control counties. The long-term decrease in concentration is about 27.5% relative to pre-policy levels. While it may seem counterintuitive that the de-concentration is larger in the long term, recall that in the SDiD both the short- and long-term peiods are being compared to the same pre-period. The de-concentration in the long term can be slowing down relative to the short-term period, but is still at a lower level relative to the pre-period than in the short term. I plot the re-weighted data using the SDiD algorithm in the short term in Appendix A Figure A7. As evident, brokerage markets in Texas counties are slightly less concentrated though similarly stable to markets in other counties, then see a sharp decrease in HHI in the quarters after the policy is announced and entry is induced.

Columns 2-5 of Table 7 display the SDiD estimation for four different market sizes, as defined

Table 7: Brokerage Market Concentration

	(1) All	(2) < 40 Brokers	(3) 40-90	(4) 90-300	(5) > 300
(1) Short Term	-0.008***	-0.023**	0.003	0.006	-0.004
	(0.003)	(0.012)	(0.005)	(0.010)	(0.004)
(2) Long Term	-0.022***	-0.043**	-0.009	-0.010	-0.015
	(0.006)	(0.020)	(0.009)	(0.014)	(0.012)
TX Pre-Year Mean	0.08	0.134	0.074	0.086	0.038
N Total	8480	2808	1566	1404	2646

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Results using the alternative short and long term are reported in Appendix B.1 Table 4. Results of a re-estimation using a standard difference-in-difference are reported in Appendix B.2 Table 2. This result uses MLS data in sample counties only.

by the stock of brokers in the county in quarter 4 of 2011.[41] Notably, smaller markets are the most concentrated, ranging from an HHI of 0.134 in markets with fewer than 40 brokers, to an HHI of 0.038 in markets with over 300 brokers. Further, the effect of increased concentration is driven solely by the smallest markets. It is in markets with a pre-existing small number of brokers where a spike of entry impacts the market landscape.

This effect is driven by smaller markets because the effect of depressed entry, and therefore stock, is occurring only in large markets instead. In Appendix A Figure A8, I replot the event studies estimating the effect of the policy on the stock of brokers per capita using the same market breakdown as in Table 7. As evident, it is the markets with the largest brokers in which the stock of brokers per 1,000 residents ultimately decreases relative to control counties. Thus, while the induced entry is affecting all markets, the restricted entry is only affecting the largest markets, where market power is least likely to be changed.

Two alternative measures of market power are considered for robustness. Appendix B Table 5

[41]These size selections are loosely based on quartiles of the distribution of brokers in Texas counties in 2011q4.

Column 1 investigates the HHI from the buy-side of the market (e.g., each firm's share of buyer representation across all sales in a given quarter). This measure is similarly decreasing, though the effect is not significant in the short term. Column 2 measures the share of transactions that are considered "dual transactions." These are transactions in which both the seller's and the buyer's agents have the same office identification number in the MLS. The share of dual transactions increases by 1.6 percentage points after the policy in the short term, suggesting less competition. This effect reverses and nearly doubles in the long term, though, suggesting again that an entry influx outweighs long-term decreased entry, leading to a higher share of transactions spread across firms. Therefore to summarize, I find that the policy decreases market concentration persistently, suggesting that the short-term induced entry outweighs the long-term restricted entry.

6.4 Efficiency

The prior results have analyzed the trade-off between quality and competition in a licensing setting where both a short-term increase and long-term decrease in entry are generated. I find that brokers are of no better quality along the dimension of consumer outcomes due to the licensing cost. I also find a *decrease* in market concentration, because the magnitude of the unintended induced entry in the short term outweighs the long-term restricted entry. Together, these results suggest that professional-level licensing is not beneficial to consumers, in the sense that it does not increase quality in the long term. Further, entry leads to de-concentration, suggesting that restricting entry would increase market power in the industry.

I now consider what the unintentional induced entry and the long-term supply restriction can tell us about the effects of licensing on efficiency in the market. The simple static model of Hsieh and Moretti (2003) shows that if indeed commissions are fixed, then when barriers to entry are low, entry of agents into cities with high housing prices will be socially inefficient. This is because commission payments are dissipated amongst new agents. In other words, the "inefficiency" here is that the productivity of the average agent, $(\frac{\rho \cdot X}{N})$, decreases. This is not efficient because these agents could be engaging in profitable activities in other industries. The authors also note that at the cross-sectional level, real wages of agents are unchanged, because they should be directly proportional to house prices in that city.

Using CPS data across about 300 MSAs, they find that when land prices in a city increase, the fraction of real estate agents increases, the productivity of an agent falls, and the real wage is unchanged.[42] These conclusions are supported in a more causal setting by Ingram and Yelowitz (2019), who gather information on entry-level training requirements across all states in the U.S. and find that, while an increase in house prices leads to an influx of new agents, this effect is dampened by more stringent licensing. Both of these settings rely on cross-sectional price increases to generate entry into the real estate market. The setting of my study, however, allows us to revisit these predictions with an exogenous shock to entry, through the unintended influx of brokers due to the TREC policy announcement.

To test the assumptions of my updated framework, I first show that the policy does not change the housing market at large. Table 8, Columns 1 and 2 consider market quantities via the total number of listings and the total number of sales per 1,000 residents in a county. Note that Column 1 corresponds to X_t and Column 2 corresponds to $\rho \cdot X_t$. Column 3, which corresponds to P_t, estimates the effect of the policy on mean sale prices.[43]

The evidence suggests that listings and sales are exogenous to broker entry. However, house prices may not be: Column 3 displays a short-term increase in county-level sale prices in response to broker entry, though this is small at about 3%. Note that this is a positive effect, though, meaning that the Hsieh and Moretti (2003) setting using price variation and this setting using entry variation should still generate the same conclusions regarding the predictions.

Given that the evidence suggests that prices and quantities are generally orthogonal to the licensing policy which generates entry, it should follow that productivity decreases. The conceptual

[42]Note that the authors do not distinguish between salespeople and brokers; they call anyone an agent who identifies their occupation as "real estate sales occupation" in the Census. This is notable, as agents enter the real estate market (largely freely) as salespeople and not brokers.

[43]Because these are housing outcomes, I do not want to include the MLS housing market controls in the specifications which residualizes the outcome variable, due to exogeneity concerns. In other words, using Equation 5, I omit the MLS market-level controls (lagged total listings sold, share sold, and median days on the market). These variables are included, however, in the algorithm to choose how to weight control counties, as they still speak to pre-period similarity.

Table 8: Housing Market

	(1) Listings per 1,000 Residents	(2) Sales per 1,000 Residents	(3) ln(Mean Sale Price)
(1) Short Term	-0.101	-0.054	0.034**
	(0.086)	(0.061)	(0.014)
(2) Long Term	0.191	0.081	0.010
	(0.153)	(0.108)	(0.026)
TX Pre-Year Mean	2.98	1.52	161,729.2
N Total	8480	8480	8480

Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Results using the alternative short and long term are reported in Appendix B.1 Table 5. Results of a re-estimation using a standard difference-in-difference are reported in Appendix B.2 Table 3. This result uses MLS data in sample counties only.

framework predicts that listings per broker should unambiguously decrease in the short and long term, while the effect on *sales* per broker is ambiguous in my setting.

As a result of this quasi-exogenous shock to entry, brokers, in fact, have fewer listings and sales, an effect that is present even in the short term but accelerates in the long term. Results are in Table 9; note that that Column 1 corresponds to $(\frac{X_t}{N_t})$ and Column 2 to $(\frac{\rho \cdot X_t}{N_t})$. Thus, the market is more inefficient as a result of extra entry, as defined by listings and sales per broker. In other words, brokers in general are worse off, because the size of the pie has not changed and the induced entry effect dominates the restricted entry.

In the Hsieh and Moretti (2003) setting, a result of this decreased productivity due to entry is that *real* wages should be unchanged as they will necessarily decrease with entry and be proportional to house prices. The authors find supportive evidence of this in the cross-section. The analogue in this setting is to measure the change of commissions over time within-area. If it is true that agents cannot compete on commission price (and are not competing elsewhere), then within a county the wages of brokers should decrease due to increased entry. However, I do not find significant evidence that this is the case.

To test this, I use the listings data to calculate an estimate of the average commissions on

Table 9: Broker's Share

	(1) Listings per Broker	(2) Sales per Broker	(3) Mean ln(Broker Commission)
(1) Short Term	-0.532***	-0.350***	-0.027
	(0.124)	(0.082)	(0.041)
(2) Long Term	-0.678***	-0.886***	-0.059
	(0.209)	(0.259)	(0.067)
TX Pre-Year Mean	3.687	1.899	16,939.39
N Total	8480	8480	6642

Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Results using the alternative short and long term are reported in Appendix B.1 Table 6. Results of a re-estimation using a standard difference-in-difference are reported in Appendix B.2 Table 4. Column 1 and 2 use MLS data in sample counties only for the numerator and licensee data only for the denominator. Column 3 uses the matched listing-to-licensee sample.

listings that brokers across a county earn in a given quarter. Because neither the commission on the sale and how this is divided amongst the selling and buying agent, nor the split between a supervising broker and salesperson are observable, I make a number of simplifying assumptions.[44] Once I have a total quarterly commission for each broker, I take the average across all brokers in a county-quarter who are making a positive amount (such that this measures commissions for "active" brokers), and then compute the natural log of this average.[45]

With this measure, displayed in Column 3, I find that broker commissions are unchanged in both the short and long term. Note that in the long term broker entry is restricted and yet commissions are still unchanging; this further supports the notion of persistent de-concentration due to induced entry. This also suggests that, while brokers are not necessarily able to compete on price,

[44]The first is that each sale has a 6% commission, split evenly between the buying and selling agent (i.e., for any given listing, whether on the sale- or buy-side, an agent is taking home 3% for the firm). The second is that brokers retain 100% of this commission for themselves, while salespeople retain 80% and give the remaining 20% to brokers at the firm. Because many firms have multiple brokers attached to them in the listing data and I cannot observe which specific brokers supervise which salespeople, the calculation evenly splits the 20% of the firm's total salespeople commissions amongst all brokers at the firm.

[45]Note that there are multiple reasons this will likely be an undercount of broker revenues. This is not accounting for any annual fees a salesperson may have to pay to a broker in order to offset overhead costs. Further, this does not account for any salaried income a broker may earn, particularly at a larger firm.

because the mean broker's wage is not changing despite an entry influx, brokers may be competing on other dimensions such as volume or number of supervised salespeople.

Evidence on the quality-concentration trade-off implies that broker licensing is not beneficial to consumers. In contrast, these results regarding efficiency suggest that broker productivity, as defined by both listings and sales per broker, decreases as a result of extra entry. Therefore, by this stylized definition, entry into the real estate industry is indeed inefficient. However, there are alternative ways to conceptualize efficiency aside from listing productivity and wages, particularly in the context of an apprenticeship industry where licensing is affecting the professionals, as opposed to the entry-level laborers.

The ostensible purpose of the TREC policy is to increase the managerial ability of brokers by requiring them to be better qualified to train and supervise salespeople. Thus, an alternative way to think about "efficiency" from broker entry is whether the larger pool of brokers are behaving more as managers. I provide evidence that the higher barrier to entry led brokers to shift to a more managerial role over time in the long term. Table 10 considers the distribution of the workload within the mean firm. Column 1 measures the mean total number of sale-side listings across firms in a county-quarter, while Columns 2 and 3 measure how many of these are performed by brokers or salespeople, respectively, as the listed agent in the MLS. Finally, Column 4 measures the share of firms in the MLS in a county in which there is no broker who performs a listing for that firm that quarter (i.e., all of that office's listings are performed by licensed salespeople). I call these "broker-manager" firms.

In the short term, brokers begin to perform fewer listings on average for the firm. Further, there is a 3 percentage point decrease in the share of broker-manager firms. In the long term, the evidence suggests that the mean firm is performing fewer listings, driven by fewer listings by brokers. Note that, in the year before the policy, the mean firm has brokers performing 1.2 sale side listings. Thus, a decrease of about 0.4 listings per quarter is around a third of the quarterly productivity and amounts to 0.12 fewer listings in a year. The increase in broker-manager firms doubles in the long term by about 17% relative to the pre-period mean.

Table 10: Shifting Broker Role

	(1) Mean Firm Listings	(2) Broker Listings	(3) Salesperson Listings	(4) Share Broker-Manager
(1) Short Term	-0.251	-0.242*	-0.170	0.034**
	(0.189)	(0.126)	(0.124)	(0.017)
(2) Long Term	-0.484**	-0.393***	-0.227	0.075***
	(0.205)	(0.124)	(0.153)	(0.027)
TX Pre-Year Mean	3.1	1.2	1.9	0.43
N (Total)	8100	8100	8100	8100

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.0011$

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Results using the alternative short and long term are reported in Appendix B.1 Table 7. Results of a re-estimation using a standard difference-in-difference are reported in Appendix B.2 Table 5. This result uses the matched listing-to-licensee sample.

An additional alternative for efficiency is the number of salespeople supervised per broker, as opposed to the number of listings per broker, to the extent that the purpose of a broker is not to sell listings directly but rather to oversee salespeople. In Table 11, I investigate the effects of the policy on firm structure and particularly the overall composition of salespeople and brokers. Columns 1 and 2 measure the mean number of brokers and salespeople, respectively, at a firm across all firms in the county. Note that in the short term, the mean firm looks the same as before the policy relative to untreated counties; however, in the long term, the mean firm has fewer brokers, despite the fact that the policy induced entry. This effect is in line with the result in Table 10 Column 4, which suggests that more firms have brokers who are not performing listings. In Column 3, I show suggestive evidence that this leads to a higher salesperson-to-broker ratio at the mean firm; however, this result is imprecisely estimated. Finally, Column 4 tests the effect of the policy on the county-level ratio of licensed salespeople to licensed brokers; this ratio decreases in both the short and long term necessarily due to the persistent effects of the anticipatory broker entry. Thus, while the aggregate productivity in terms of salespeople-per-broker has decreased in the long term, there is suggestive evidence that at the mean firm, brokers are more productive.

Table 11: Firm Management

	(1) Mean Firm Broker	(2) Mean Firm Salespeople	(3) Mean Firm S:B	(4) County S:B Ratio
(1) Short Term	-0.016	-0.042	0.084	-0.154***
	(0.027)	(0.055)	(0.101)	(0.052)
(2) Long Term	-0.147***	-0.174	0.254	-0.298*
	(0.033)	(0.128)	(0.181)	(0.168)
TX Pre-Year Mean	0.72	1.91	1.22	3.67
N Total	8100	8100	7668	8480

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Results using the alternative short and long term are reported in Appendix B.1 Table 8. Results of a re-estimation using a standard difference-in-difference are reported in Appendix B.2 Table 6. Columns 1-3 use the matched listing-to-licensee sample. Column 4 uses licensee data in sample counties only.

A final consideration of the "efficiency" of broker entry is the distributional effect of an increased licensing cost. It has been established that licensing barriers could have differential effects for different types of potential entrants. For instance, Angrist and Guryan (2008) find that increasing the stringency of licensing for teachers led to fewer new Hispanic teachers. Specific to real estate, Ingram and Yelowitz (2019) find that the impact of licensing is stronger for females. In Table 12, I consider how this policy affected female and minority entry into the broker market, as the policy specifically changed to barrier of upgrading to a broker.[46]

It should first be noted that, before the policy was announced, the majority (about 60%) of brokers are male; while real estate is often cited in the popular media as a female-majority industry, it is at the salespeople level where women make up the majority of agents (also coming in at around 60%). It follows, then, that the mean share of female entrants is just 30% in the year before the policy change. In the long term, the policy decreases this entry share by 7.2 percentage points, a decrease of nearly 24%. The policy also has a negative effect on the entry of Hispanic brokers. Texas has a large Hispanic population, and as such the largest non-White category of entering

[46]To estimate the gender of agents, I use the algorithm from genderize.io which uses a large number of social media profiles to predict gender based on first name. To estimate race, I use the NamePrism algorithm. See Ye et al. (2017) and Ye and Skiena (2019) for more information.

brokers pre-policy change is Hispanic, with a mean share across counties of 10.9% of entrants. The policy has an immediate negative effect on entry which amplifies in the long run, amounting to a decrease in 5.1 percentage points. This effect is nearly half the pre-period entry share. I do not find significant evidence of a differential effect for Black brokers; however, the baseline Black share of entrants is relatively small at only 1.3%.

Therefore, as the intended effect of the policy, to restrict broker entry, takes hold, fewer women and minorities enter the market. This result further suggests that free entry is indeed "efficient," in the sense that it allows for a more equal chance for everyone to participate in the market.

Table 12: Broker Entry - Gender and Race

	(1) Share Entrants Female	(2) Share Entrants Hispanic	(3) Share Entrants Black
(1) Short Term	-0.005	-0.031*	-0.010
	(0.029)	(0.016)	(0.013)
(2) Long term	-0.072**	-0.051***	-0.018
	(0.030)	(0.015)	(0.014)
TX Pre-Year Mean	0.300	0.109	0.013
N	8424	8424	8424

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Results using the alternative short and long term are reported in Appendix B.1 Table 9. Results of a re-estimation using a standard difference-in-difference are reported in Appendix B.2 Table 7. This result uses licensee data in sample counties matched with the genderize.io (Column 1) or NamePrism (Columns 2 and 3) algorithm.

7 Conclusion

Policymakers and trade associations often use occupational licensing as a tool to ostensibly ensure better service for consumers; however, the entry costs associated with licensing requirements come with a trade-off between increasing practitioner quality and also increasing market concentration. These effects are particularly relevant yet understudied in apprenticeship industries, where there are multiple licenses to progress through within an industry, providing different implications for quality spillovers and market structure changes. Additionally, any policy which restricts entry into a labor market is likely to have consequences for market efficiency.

This paper uses the real estate industry as the setting to study the economic consequences of entry costs due to occupational licensing, with a particular focus on costs at the professional level of an apprenticeship industry. It is important to understand these effects in the real estate industry, as it is both a consequential intermediary market for consumers, with financial implications at both the household and macro-level, and a massive labor market with a large number of participants across the US. This paper estimates the trade-off between quality and concentration and the consequences for market efficiency when the future licensing barriers for brokers changes. To estimate these effects, I exploit a policy change in Texas in 2012 that provides quasi-exogenous variation in the licensing cost for brokers in a synthetic difference-in-difference research design.

The policy reform causes an unintended increase in broker entry, leading to a persistently higher stock of brokers, before entry is restricted in the long term. The results illustrate that the increased licensing cost does not lead to an increased quality of entering brokers in the long term. Further, market concentration decreases due to the unintended effects of the policy. This suggests that restricting entry via higher licensing costs is not desirable for consumers. On the contrary, the evidence also shows that the unintended induced entry leads to decreased broker productivity in the labor market. This would instead suggest that restricting entry has benefits for market efficiency. However, this narrow definition of efficiency does not capture the full distributional effects of higher licensing costs. I find that costlier licensing leads to a smaller share of entering female and Hispanic brokers in the long term, such that unrestricted entry is more equitable.

The results suggest many directions for future research. In particular, there are other industries which have a similar apprenticeship licensing structure which have not been analyzed. Specific to real estate, there is notable heterogeneity across states in the allowance of reciprocal agreements. For instance, Texas does not allow reciprocal licensing, whereas Florida has reciprocity with Alabama, Arkansas, Connecticut, Georgia, Illinois, Mississippi, Nebraska, and Rhode Island.

There is little research on the origins of these reciprocal agreements and what the spillovers may be in terms of cost of entry and employment structure. Finally, there are significant differences in how real estate intermediaries are licensed and monitored in other countries potentially leading to different industry outcomes.

References

Aiello, Darren, Mark J Garmaise, and Taylor Nadauld. 2022. "What Problem Do Intermediaries Solve?" *Available at SSRN 4105923*.

Anderson, D Mark, Ryan Brown, Kerwin Kofi Charles, and Daniel I Rees. 2020. "Occupational licensing and maternal health: Evidence from early midwifery laws." *Journal of Political Economy* 128 (11): 4337–4383.

Angrist, Joshua D, and Jonathan Guryan. 2008. "Does teacher testing raise teacher quality? Evidence from state certification requirements." *Economics of Education Review* 27 (5): 483–503.

Arkhangelsky, Dmitry, Susan Athey, David A Hirshberg, Guido W Imbens, and Stefan Wager. 2021. "Synthetic difference-in-differences." *American Economic Review* 111 (12): 4088–4118.

Barwick, Panle Jia, and Parag A Pathak. 2015. "The costs of free entry: an empirical study of real estate agents in Greater Boston." *The RAND Journal of Economics* 46 (1): 103–145.

Barwick, Panle Jia, and Maisy Wong. 2019. "Competition in the real estate brokerage industry: A critical review." *Urban Development* 1.

Bernheim, B Douglas, and Jonathan Meer. 2013. "Do real estate brokers add value when listing services are unbundled?" *Economic inquiry* 51 (2): 1166–1182.

Blair, Peter Q, and Bobby W Chung. 2019. "How much of barrier to entry is occupational licensing?" *British Journal of Industrial Relations* 57 (4): 919–943.

Bowblis, John R, and Austin C Smith. 2021. "Occupational licensing of social services and nursing home quality: A regression discontinuity approach." *ILR Review* 74 (1): 199–223.

Carroll, Sidney L, and Robert J Gaston. 1979. "State Occupational Licensing Provisions and Quality Provisions and Quality of Service: The Real Estate Business." *Rsch. in L. & Econ.* 1 1.

Chung, Bobby W. 2022. "The costs and potential benefits of occupational licensing: A case of real estate license reform." *Labour Economics* 76 102172.

Friedman, Milton, and Simon Kuznets. 1945. "Income from Independent Professional." *New York*.

Gilbukh, Sonia, and Paul S Goldsmith-Pinkham. 2019. "Heterogeneous real estate agents and the housing cycle." *Available at SSRN 3436797*.

Guntermann, Karl, and Richard Smith. 1988. "Licensing requirements, enforcement effort and complaints against real estate agents." *Journal of Real Estate Research* 3 (2): 11–20.

Han, Lu, and Seung-Hyun Hong. 2011. "Testing cost inefficiency under free entry in the real estate brokerage industry." *Journal of Business & Economic Statistics* 29 (4): 564–578.

Han, Lu, and William C Strange. 2015. "The microstructure of housing markets: Search, bargaining, and brokerage." *Handbook of regional and urban economics* 5 813–886.

Hsieh, Chang-Tai, and Enrico Moretti. 2003. "Can free entry be inefficient? Fixed commissions and social waste in the real estate industry." *Journal of Political Economy* 111 (5): 1076–1122.

Ingram, Samuel J, and Aaron Yelowitz. 2019. "Real estate agent dynamism and licensing entry barriers." *Journal of Entrepreneurship and Public Policy* 10 (2): 156–174.

Johnson, Linda L, and Christine Loucks. 1986. "The effect of state licensing regulations on the real estate brokerage industry." *Real Estate Economics* 14 (4): 567–582.

Jud, G Donald, and Daniel T Winkler. 2000. "A note on licensing and the market for real estate agents." *The Journal of Real Estate Finance and Economics* 21 (2): 175–184.

Kleiner, Morris M, and Evan J Soltas. 2023. "A welfare analysis of occupational licensing in US states." *Review of Economic Studies* rdad015.

Kleiner, Morris M, and Edward J Timmons. 2020. "Occupational licensing: Improving access to regulatory information." *Journal of Labor Research* 41 (4): 333–337.

Levitt, Steven D, and Chad Syverson. 2008. "Market distortions when agents are better informed: The value of information in real estate transactions." *The Review of Economics and Statistics* 90 (4): 599–611.

Lopez, Luis A, Shawn McCoy, and Vivek Sah. 2019. "Steering Consumers to Affiliated Financial Services: Evidence from Pre-Approvals and the Cost of Credit." *Available at SSRN 3365347*.

Lopez, Luis Arturo. 2021. "Asymmetric information and personal affiliations in brokered housing transactions." *Real Estate Economics* 49 (2): 459–492.

Rhoades, Stephen A. 1993. "The herfindahl-hirschman index." *Fed. Res. Bull.* 79 188.

Salant, Stephen W. 1991. "For sale by owner: When to use a broker and how to price the house." *The Journal of Real Estate Finance and Economics* 4 (2): 157–173.

Shapiro, Carl. 1986. "Investment, moral hazard, and occupational licensing." *The Review of Economic Studies* 53 (5): 843–862.

Shilling, James, and C Sirmam. 1988. "The effects of occupational licensing on complaints against real estate agents." *Journal of Real Estate Research* 3 (2): 1–9.

Stigler, George J. 1971. "The theory of economic regulation." *The Bell journal of economics and management science* 3–21.

Turnbull, Geoffrey K, and Bennie D Waller. 2018. "(What) do top performing real estate agents deliver for their clients?" *Journal of Housing Economics* 41 142–152.

Turnbull, Geoffrey K, Bennie D Waller, and Scott A Wentland. 2022. “Mitigating agency costs in the housing market.” *Real Estate Economics* 50 (3): 829–861.

Waller, Bennie, and Ali Jubran. 2012. “The impact of agent experience on the real estate transaction.” *Journal of Housing Research* 21 (1): 67–82.

Ye, Junting, Shuchu Han, Yifan Hu, Baris Coskun, Meizhu Liu, Hong Qin, and Steven Skiena. 2017. “Nationality classification using name embeddings.” In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 1897–1906.

Ye, Junting, and Steven Skiena. 2019. “The secret lives of names? name embeddings from social media.” In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3000–3008.

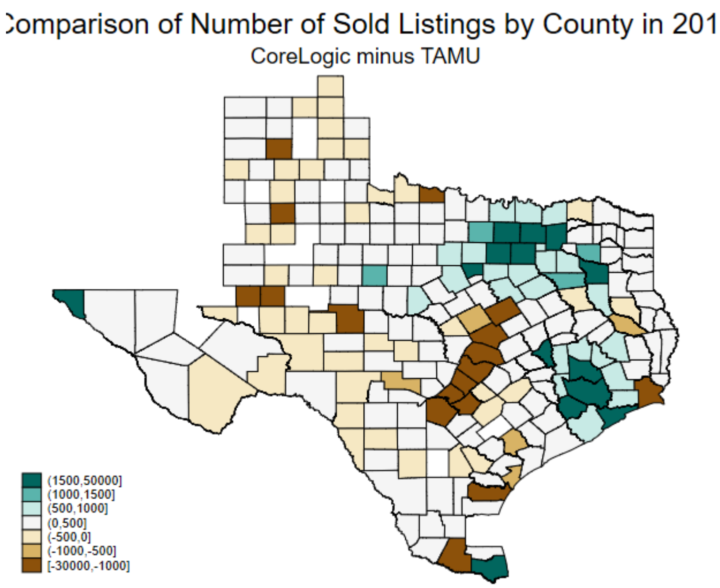
Yelowitz, Aaron, and Samuel J Ingram. 2021. “How does occupational licensing affect entry into the medical field? An examination of emergency medical technicians.” *Southern Economic Journal*.

Zapletal, Marek. 2019. “The effects of occupational licensing: evidence from business-level data.” *British Journal of Industrial Relations* 57 (4): 894–918.

Appendix

A Figures

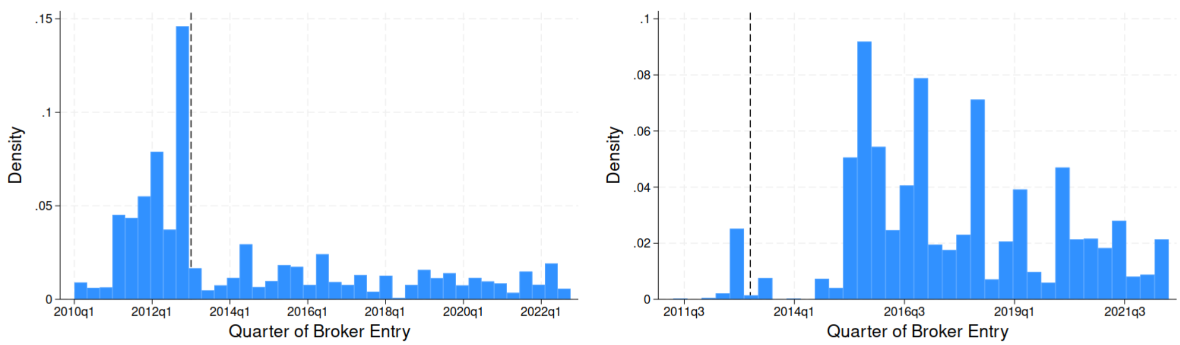
Figure A1: Comparing CoreLogic to TAMU Real Estate Center



Note: Figure displays the number of total listings sold by county in 2017 in the CoreLogic data vs. the Texas A&M Real Estate Center data. Bluer shades reflect better coverage in CoreLogic, while orange shades represent better coverage by TAMU.

Figure A2: Entry Year by Group

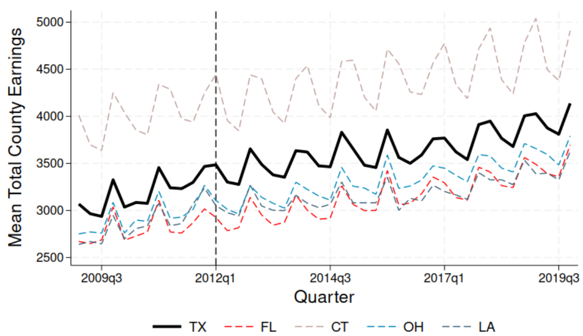
(a) Grandfather Eligible (b) Grandfather Ineligible



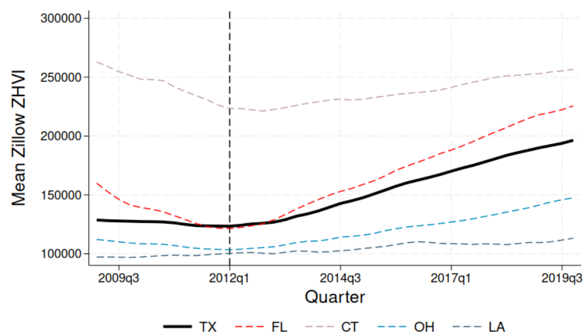
Note: Figure displays a distribution of the quarter in which an agent upgrades to a broker for two separate groups. In Panel A, I plot this distribution for the group of agents who were “Grandfather Eligible,” i.e., those agents who became licensed as salespeople from 2008-2010, and had a least two years experience as a salesperson when the stricter licensing policy was announced. Panel B plots this for the “Grandfather Ineligible” group, i.e, the agents who were licensed from 2010-2012, and would not have two years experience when the impending stricter policy was announced.

Figure A3: Raw Data Trends for All States

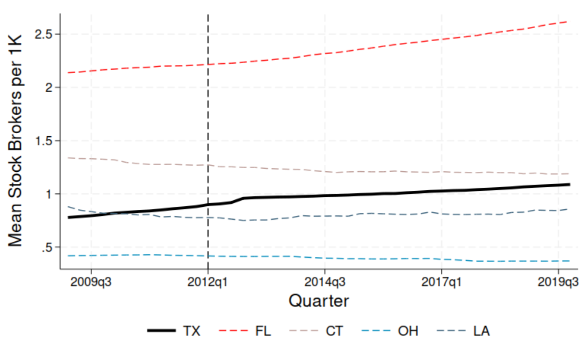
(a) Median Earnings



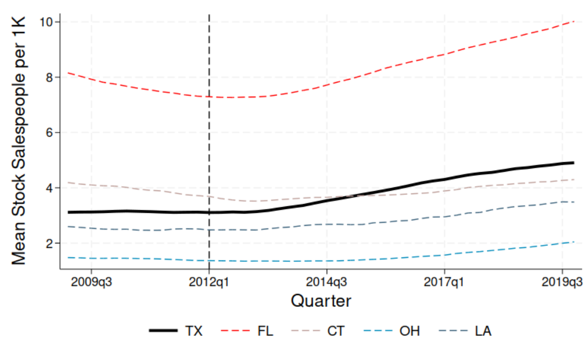
(b) Median ZHVI



(c) Total Stock of Brokers



(d) Total Stock of Salespeople



Note: Figure displays raw data for counties in Texas (the blue solid line) and counties in my four control states. Panel A displays the Median Earnings across all counties is a given state over time using QCEW data. Panel B displays the median of Zillow's county-level ZVHI. Panel C displays total brokers added up across all counties, and Panel D displays total salespeople. The vertical line represents the policy announcement.

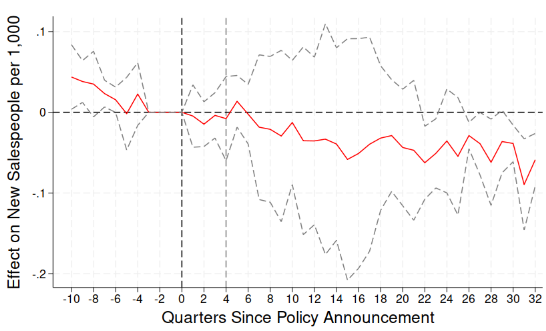
Figure A4: Exiting Brokers per 1,000 Residents

TX Pre-Year Mean: 0.001
DiD Coeff: 0.001

Note: Figure displays event study coefficients estimating Equation 4. The outcome variable is the total number of exiting brokers per 1,000 county residents. The blue solid lines represent the coefficients, while the dashed gray lines represent the 95% confidence interval. The first black vertical lines represent the policy announcement, while the second grey vertical lines represent the policy effective date. This result uses licensee data in sample counties only.

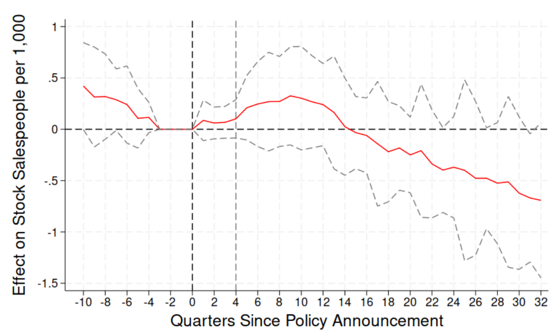
Figure A5: Salespeople Stock and Entry

(a) New Salespeople per 1,000 Residents



TX Pre-Year Mean: 0.054
DiD Coeff: -0.046

(b) Stock Salespeople per 1,000 Residents

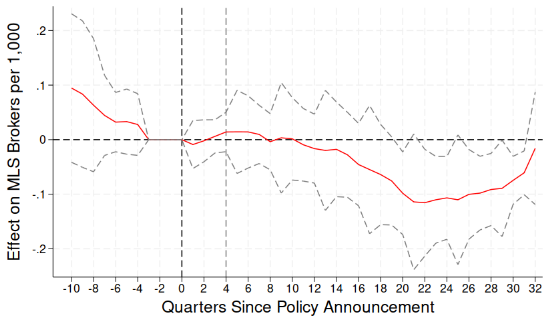


TX Pre-Year Mean: 3.118
DiD Coeff: -0.238

Note: Figure displays event study coefficients estimating Equation 4. The outcome variable is the total number of entering salespeople (Panel A) and the total stock of salespeople (Panel B) per 1,000 county residents. The red solid lines represent the coefficients, while the dashed gray lines represent the 95% confidence interval. The first black vertical lines represent the policy announcement, while the second grey vertical lines represent the policy effective date. This result uses licensee data in sample counties only.

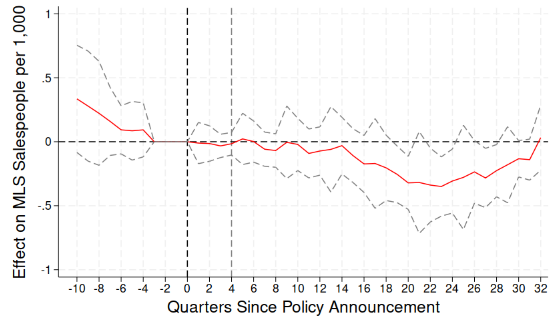
Figure A6: Active Brokers and Salespeople per 1,000 Residents

(a) Active Brokers



TX Pre-Year Mean: 0.165
DiD Coeff: -0.072

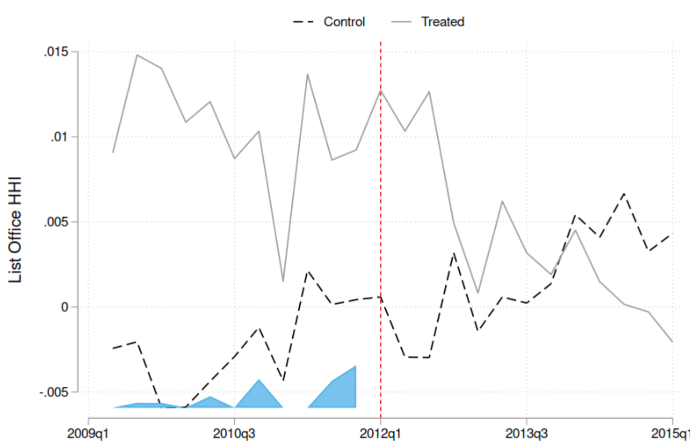
(b) Active Salespeople



TX Pre-Year Mean: 0.503
textbfDiD Coeff: -0.207

Note: Figure displays event study coefficients estimating Equation 4. The outcome variable is the total number of brokers (Panel A) and the total number of salespeople (Panel B) performing at least one listing in the MLS per 1,000 county residents. The red solid lines represent the coefficients, while the dashed gray lines represent the 95% confidence interval. The first black vertical lines represent the policy announcement, while the second grey vertical lines represent the policy effective date. This result uses the matched listing-to-licensee sample.

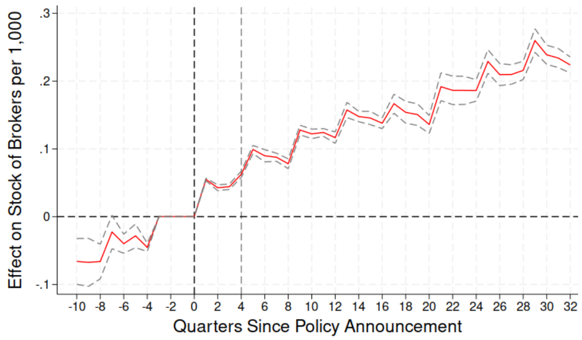
Figure A7: SDiD Weighted Data



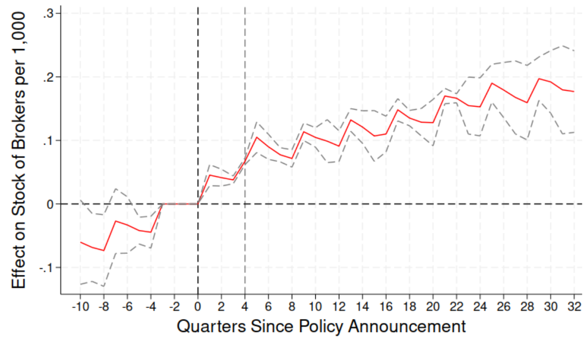
Note: Figure displays trends in broker employment for counties in Texas ("Treated," solid line) and the weighted average of counties in the control states ("Control," dashed line). The vertical line represents the policy announcement. The weights used to average the pre-treatment time period are displayed in blue at the bottom of the graphs. See Arkhangelsky et al. (2021) and specifically their Figure 1 for more information.

Figure A8: Broker Stock per 1,000 Residents

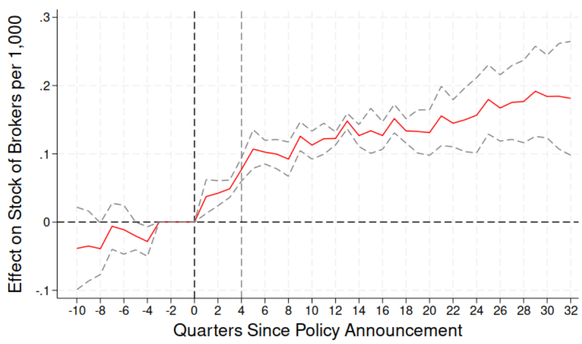
(a) $<$ 40 Brokers



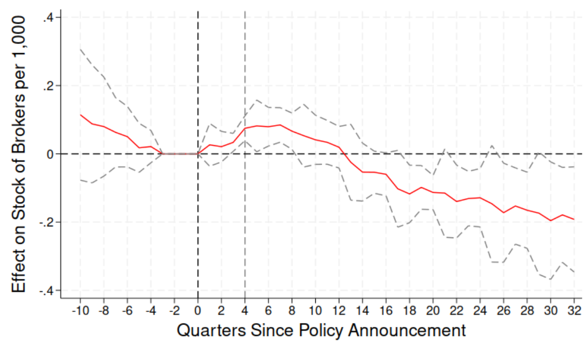
(b) 40-90 Brokers



(c) 90-300 Brokers



(d) $>$ 300 Brokers



Note: Figure displays event study coefficients estimating Equation 4. The outcome variable is the total stock of brokers per 1,000 county residents. Panel A displays this on the sample of counties with less than 40 brokers in 2011Q4; Panel B counties with 40-90 brokers; Panel C 90-300 brokers; Panel D more than 300 brokers. The red solid lines represent the coefficients, while the dashed gray lines represent the 95% confidence interval. The first black vertical lines represent the policy announcement, while the second grey vertical lines represent the policy effective date. This result uses licensee data in sample counties only.

B Tables

Table 1: Constructing the Matched Listings Sample

State	In Sample Counties	Dropping Moving Agents	Matched to Listing Agent	Percent of Total
CT	717,361	527,375	274,721	38.3
FL	4,466,908	2,831,099	1,271,317	28.5
LA	64,089	34,747	17,708	27.63
OH	1,655,363	1,147,988	609,408	36.8
TX	2,627,731	1,533,963	641,952	24.5
Total	9,531,452	6,075,172	2,815,106	29.5

Note: Table reports number of MLS listings by state. "In Sample Counties" reports the total number of MLS listings in the sample counties by state. "Dropping Moving Agents" reports the total number of listings remaining when I drop listings that are attached to an agent who works with multiple firms in the same quarter. "Matched to a Listing Agent" reports the total number of listings remaining which can be matched to a licensee in my licensing records. "Percent of Total" reports the share of the total listings remaining.

Table 2: New Broker Quality - 4 Year Prior Experience

	(1) Total Listings	(2) Total Sold	(3) Buy Side Listings	(4) Sale Side Listings
(1) Short Term	-3.114	-1.996	-0.401	-2.713
	(3.864)	(2.457)	(0.799)	(3.099)
(2) Long Term	12.963***	6.654***	4.063***	8.900***
	(1.464)	(1.132)	(0.570)	(1.553)
TX Pre-Year Mean	9.89	3.76	3.47	6.41
N	3145	3145	3145	3145

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. All outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). This result uses the matched listing-to-licensee sample.

Table 3: New Broker Quality - 4 Year Post - Intensive Margin

	(1) Mean Return ln(Price)	(2) Mean Return ln(DOM)
(1) Short Term	-0.206***	-0.184***
	(0.020)	(0.031)
(2) Long Term	-0.211***	0.061
	(0.017)	(0.111)
TX Pre-Year Mean	0.073	0.022
N	753	740

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. All outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). This result uses the matched listing-to-licensee sample.

Table 4: New Salespeople Quality - 4 Year Post - Intensive Margin

	(1) Mean Return ln(Price)	(2) Mean Return ln(DOM)
(1) Short Term	-0.049	0.050
	(0.053)	(0.091)
(2) Long Term	-0.087	0.267
	(0.055)	(0.221)
TX Pre-Year Mean	0.041	-0.128
N	1404	1404

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. This result uses the matched listing-to-licensee sample.

Table 5: Brokerage Market Concentration

	(1) Buy Office HHI	(2) Share Dual Transactions
(1) Short Term	-0.010	0.014*
	(0.008)	(0.007)
(2) Long Term	-0.024*	-0.023**
	(0.013)	(0.011)
TX Pre-Year Mean	0.108	0.298
N	8424	8370

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. This result uses MLS data in sample counties only.

Table 6: Salesperson's Share

	(1) Listing per Salesperson	(2) Sales per Salesperson	(3) ln(Mean Sales Commission)
(1) Short Term	-0.120***	-0.128***	-0.086
	(0.034)	(0.028)	(0.062)
(2) Long Term	0.011	-0.098***	-0.104
	(0.054)	(0.031)	(0.093)
TX Pre-Year Mean	1.00	0.512	8,442.03
N	8480	8480	7261

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Column 1 and 2 use MLS data in sample counties only for the numerator and licensee data only for the denominator. Column 3 uses the matched listing-to-licensee sample.

B.1 Alternative Short and Long Term

Table 1: Entry and Stock per 1,000 County Residents

	(1) New Brokers	(2) Stock Brokers	(3) New Salespeople	(4) Stock Salespeople
(1) Short Term	0.006***	0.031***	0.001	0.017
	(0.001)	(0.006)	(0.004)	(0.016)
(2) Long Term	-0.009***	0.072***	-0.007	0.433***
	(0.001)	(0.022)	(0.010)	(0.142)
TX Pre-Year Mean	0.011	0.865	0.054	3.118
N Total	8480	8480	8480	8480

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD as described in Section 5. Note that the outcome variables are all measured per 1,000 county residents. Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. This result uses licensee data in sample counties only.

Table 2: New Broker Quality - Extensive Margin

	(1) Sold	(2) Sold < 30	(3) Sold < 90
(1) Short Term	-0.008	-0.033	-0.132**
	(0.037)	(0.047)	(0.033)
(2) Long Term	0.038	0.073***	0.009
	(0.026)	(0.021)	(0.003)
TX Pre-Year Mean	0.433	0.125	0.289
N Total	2276	2276	2276

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a standard DiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. All outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter This result uses the matched listing-to-licensee sample.

Table 3: New Broker Quality - Intensive Margin

	(1) Mean Return ln(Price)	(2) Mean Return ln(DOM)
(1) Short Term	-0.135	0.472***
	(0.093)	(0.101)
(2) Long Term	-0.223**	-0.036
	(0.060)	(0.038)
TX Pre-Year Mean	0.137	0.020
N	753	740

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a standard DiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. All outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter. This result uses the matched listing-to-licensee sample.

Table 4: Brokerage Market Concentration

	(1) All	(2) $<$ 40 Brokers	(3) 40-90	(4) 90-300	(5) $>$ 300
(1) Short Term	-0.001	-0.007	0.012**	0.005	0.002
	(0.003)	(0.010)	(0.005)	(0.009)	(0.004)
(2) Long Term	-0.019***	-0.040**	-0.008	-0.003	-0.013
	(0.005)	(0.018)	(0.008)	(0.013)	(0.011)
TX Pre-Year Mean	0.08	0.134	0.074	0.086	0.038
N Total	8480	2808	1566	1404	2646

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. This result uses MLS data in sample counties only.

Table 5: Housing Market

	(1) Listings per 1,000 Residents	(2) Sales per 1,000 Residents	(3) ln(Mean Sale Price)
(1) Short Term	0.092*	0.055	0.014
	(0.055)	(0.035)	(0.012)
(2) Long Term	0.068	0.032	0.020
	(0.131)	(0.096)	(0.022)
TX Pre-Year Mean	2.98	1.52	161,729.2
N Total	8480	8480	8480

Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. This result uses MLS data in sample counties only.

Table 6: Broker's Share

	(1) Listings per Broker	(2) Sales per Broker	(3) Mean ln(Broker Commission)
(1) Short Term	-0.130*	-0.073	-0.024
	(0.076)	(0.053)	(0.035)
(2) Long Term	-0.681***	-0.758***	-0.053
	(0.179)	(0.222)	(0.060)
TX Pre-Year Mean	3.687	1.899	16,939.39
N Total	8480	8480	6642

Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Column 1 and 2 use MLS data in sample counties only for the numerator and licensee data only for the denominator. Column 3 uses the matched listing-to-licensee sample.

Table 7: Shifting Broker Role

	(1) Mean Firm Listings	(2) Broker Listings	(3) Salesperson Listings	(4) Share Broker-Manager
(1) Short Term	0.255	-0.068	0.181	0.018
	(0.171)	(0.093)	(0.118)	(0.016)
(2) Long Term	-0.517**	-0.370***	-0.264*	0.064***
	(0.201)	(0.122)	(0.140)	(0.024)
TX Pre-Year Mean	3.1	1.2	1.9	0.43
N (Total)	8100	8100	8100	8100

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.0011$

Note: Table displays results estimating an SDiD as described in Section 5.Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. This result uses the matched listing-to-licensee sample.

Table 8: Firm Management

	(1) Mean Firm Broker	(2) Mean Firm Salesppl	(3) Mean Firm S:B	(4) County S:B Ratio
(1) Short Term	0.039	0.105***	0.051	-0.178***
	(0.024)	(0.035)	(0.093)	(0.031)
(2) Long Term	-0.116***	-0.168	0.211	-0.265**
	(0.029)	(0.104)	(0.154)	(0.124)
TX Pre-Year Mean	0.72	1.91	1.22	3.67
N Total	8100	8100	7668	8480

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. Columns 1-3 use the matched listing-to-licensee sample. Column 4 uses licensee data in sample counties only.

Table 9: Broker Entry - Gender and Race

	(1) Share Entrants Female	(2) Share Entrants Hispanic	(3) Share Entrants Black
(1) Short Term	0.067*	-0.002	0.001
	(0.037)	(0.014)	(0.010)
(2) Long Term	-0.065**	-0.050***	-0.018
	(0.029)	(0.016)	(0.014)
TX Pre-Year Mean	0.300	0.109	0.013
N	8424	8424	8424

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating an SDiD as described in Section 5. Row (1) reports the coefficient on Texas*Post in the short term, which is now defined as the first four quarters after the policy was announced. Row (2) reports this for the long term now defined as quarters 5-32. Before taking the residual as described is Equation 5, all outcome variables are first winsorized to the 97th and 3rd percentiles for the state-year-quarter Standard errors are bootstrapped. See Arkhangelsky et al. (2021) for more information. This result uses licensee data in sample counties matched with the genderize.io (Column 1) or NamePrism (Columns 2 and 3) algorithm.

B.2 Corresponding Difference-in-Difference Output

Table 1: Entry and Stock per 1,000 County Residents

	(1) New Brokers	(2) Stock Brokers	(3) New Salespeople	(4) Stock Salespople
(1) Short Term	-0.004*	0.020	-0.041	0.021
	(0.002)	(0.040)	(0.028)	(0.069)
(2) Long Term	-0.011***	-0.108	-0.062	-0.473
	(0.002)	(0.063)	(0.030)	(0.280)
TX Pre-Year Mean	0.011	0.865	0.054	3.118
N	3744	3744	3744	3744

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Note: Table displays results estimating a standard DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). This result uses the matched listing-to-licensee sample.

Table 2: Brokerage Market Concentration

	(1) All	(2) $<$ 40 Brokers	(3) 40-90	(4) 90-300	(5) $>$ 300
(1) Short Term	-0.006**	-0.006	0.004	-0.015***	-0.003
	(0.002)	(0.003)	(0.002)	(0.001)	(0.004)
(2) Long Term	-0.016***	-0.031**	-0.020**	-0.022**	-0.011**
	(0.001)	(0.006)	(0.004)	(0.005)	(0.003)
TX Pre-Year Mean	0.08	0.134	0.074	0.086	0.038
N Total	8480	2808	1566	1404	2646

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). This result uses MLS data in sample counties only.

Table 3: Housing Market

	(1) Listings per 1,000 Residents	(2) Sales per 1,000 Residents	(3) ln(Mean Sale Price)
(1) Short Term	0.013	0.008	-0.017
	(0.062)	(0.029)	(0.038)
(2) Long Term	0.391**	0.144*	-0.042
	(0.106)	(0.057)	(0.030)
TX Pre-Year Mean	2.98	1.52	161,729.2
N Total	8480	8480	8480

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). This result uses MLS data in sample counties only.

Table 4: Broker's Share

	(1) Listings per Broker	(2) Sales per Broker	(3) Mean ln(Broker Commission)
(1) Short Term	-0.144	-0.146	0.059***
	(0.069)	(0.232)	(0.007)
(2) Long Term	0.083	-0.209	0.165***
	(0.156)	(0.339)	(0.025)
TX Pre-Year Mean	3.687	1.899	16,062.02
N Total	8480	8480	7847

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Column 1 and 2 use MLS data in sample counties only for the numerator and licensee data only for the denominator. Column 3 uses the matched listing-to-licensee sample.

Table 5: Shifting Broker Role

	(1) Mean Firm Listings	(2) Broker Listings	(3) Salesperson Listings	(4) Share Broker-Manager
(1) Short Term	0.055	0.054	-0.012	0.120**
	(0.026)	(0.054)	(0.006)	(0.028)
(2) Long Term	-0.009	-0.070	0.037**	-0.068
	(0.117)	(0.089)	(0.009)	(0.192)
TX Pre-Year Mean	4.0	1.2	3.1	0.43
N (Total)	8253	8253	8253	8253

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.0011$

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). This result uses the matched listing-to-licensee sample.

Table 6: Firm Management

	(1) Mean Firm Broker	(2) Mean Firm Salesppl	(3) Mean Firm S:B	(4) County S:B Ratio
(1) Short Term	0.062**	-0.190***	0.037	XX
	(0.018)	(0.041)	(0.043)	XX
(2) Long Term	-0.109***	-0.283	0.360***	XX
	(0.017)	(0.157)	(0.043)	XX
TX Pre-Year Mean	0.72	1.90	1.19	3.67
N Total	8253	8253	8151	8480

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). Columns 1-3 use the matched listing-to-licensee sample. Column 4 uses licensee data in sample counties only.

Table 7: Broker Entry - Gender and Race

	(1) Share Entrants Female	(2) Share Entrants Hispanic	(3) Share Entrants Black
(1) Short Term	0.004	-0.069***	-0.034
	(0.012)	(0.012)	(0.022)
(2) Long Term	-0.062**	-0.103***	-0.055*
	(0.022)	(0.011)	(0.025)
TX Pre-Year Mean	0.300	0.109	0.013
N	8424	8424	8424

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Note: Table displays results estimating a DiD using the time-invariant version of Equation 4. Row (1) reports the coefficient on Texas*Post in the short term (the first 12 quarters after the policy was announced). Row (2) reports this for the long term (quarters 13-32). This result uses licensee data in sample counties matched with the genderize.io (Column 1) or NamePrism (Columns 2 and 3) algorithm.

C SDiD Weight Construction

C.0.1 Unit Weight Construction

Unit weights are chosen such that:

$$(\hat{\omega}_0, \hat{\omega^{sdid}}) = \arg \min_{\omega_0 \in \mathbb{R}, \omega \in \Omega} \ell_{unit}(\omega_0, \omega),$$

where:

$$\ell_{unit}(\omega_0, \omega) = \sum_{t=1}^{T_{pre}} (\omega_0 + \sum_{j=1}^{N_{co}} \omega_j Y_{jt} - \frac{1}{N_{tr}} \sum_{j=N_{co}+1}^{N} Y_{jt})^2 + \zeta^2 T_{pre} \|\omega\|_2^2$$

and:

$$\Omega = \{\omega \in \mathbb{R}_+^N : \sum_{j=1}^{N_{co}} \omega_j = 1,\, \omega_j = N_{tr}^{-1}\, for\, all\; j = N_{co}+1, ..., N\}$$

ζ is a “regularization parameter” chosen to match the size of a typical one-period outcome change (Δ_{jt}) for unreated units in the pre-period, and then scaled:

$$\zeta = (N_{tr} T_{post})^{\frac{1}{4}} \hat{\sigma},$$

where:

$$\hat{\sigma}^2 = \frac{1}{N_{co}(T_{pre}-1)} \sum_{j=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} (\Delta_{jt} - \bar{\Delta})^2$$

Note that there are two key differences from the synthetic controls unit weights. The first is the inclusion of the intercept term ω_0, which allows for the unexposed pre-trends to only need to be parallel. The second is this regularization parameter, which ensures the uniqueness of the weights.

C.0.2 Time Weight Construction

The time weights are constructed such that:

$$(\hat{\lambda}_0, \hat{\lambda^{sdid}}) = \arg \min_{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda} \ell_{time}(\lambda_0, \lambda),$$

where:

$$\ell_{time}(\lambda_0, \lambda) = \sum_{j=1}^{N_{co}} (\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{jt} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{jt})^2$$

and:

$$\Lambda = \{\lambda \in \mathbb{R}_+^T : \sum_{j=1}^{T_{pre}} \lambda_t = 1,\, \lambda_t = T_{post}^{-1}\, for\, all\, t = T_{pre}+1, ..., T\}$$