# The Importance of Networks for Career Success: Evidence from the Real Estate Industry*

Soon Hyeok Choi† Patrick S. Smith‡ Ali Tosyali§

Current Draft: March 2026
First Draft: August 2025

[ latest version here ]

**Abstract**

Identifying the causal effect of networks on career success is challenging due to endogenous network formation. Leveraging quasi-random matching among residential real estate agents, we show that transacting with a well-connected counterparty agent increases a new agent's short-term probability of survival, transaction volume, and career length. A one-standard-deviation increase in initial counterparty centrality increases a new agent's probability of surviving one year by more than 4 percentage points and lengthens average career duration by 37%. A dynamic Bayesian network model formalizes the underlying mechanism, demonstrating how initial network capital shapes long-run career trajectories.

Keywords: Career Success, Centrality, Mentorship, Networks, Real Estate

JEL Classifications: D85, J24, J44, R31

*We thank Murillo Campello, Konhee Chang, Jorge De la Roca, Will Doerner, Ram Dubey, Marco Giacoletti, Jennifer Grootegoed, Blake Jackson, Gi Heung Kim, Jeanna Kenney, Matthew Khan, Amrutha Manjunath (discussant), Fikri Pitsuwan, Michael Reher, Alberto Rossi, Lawrence White, So Hye Yoon, and seminar and conference participants from the MIT Center for Real Estate, Baruch College, 19th North American Meeting of the Urban Economics Association and 3rd ODU Real Estate Research Symposium for constructive feedback.

†Saunders College of Business, Rochester Institute of Technology. Email: **schoi@saunders.rit.edu**.
‡Charlotte Belk College of Business, UNC Charlotte. Email: **patrick.smith@charlotte.edu**.
§Saunders College of Business, Rochester Institute of Technology. Email: **atosyali@saunders.rit.edu**.

# 1 Introduction

Various factors influence an individual's career success, including personal abilities, organizational environments, and external opportunities. Recent research, which we review later in this section, has increasingly emphasized the importance of networks. However, isolating the impact of an individual's network on their career success is difficult due to the endogenous nature of network formation. For example, an individual may possess a strong network as a result of career success. Conversely, a strong network may also facilitate the individual's achievement of career success. We address this endogeneity issue by exploiting the quasi-random matching of *new* real estate agents with counterparty agents to examine the effect of network connections on career success in the U.S. residential real estate brokerage market.

Three characteristics of the residential real estate brokerage market make it an ideal setting for examining the role of networks in shaping new agents' career success. First, there are relatively few barriers to entry into the industry (Hsieh and Moretti, 2003). Individuals who want to obtain a real estate agent license must take an exam, but the requirements to pass the exam are limited. Second, the relatively low barrier to entry results in high turnover and a low survival rate. Only a fraction of new agents (ranging from 56 to 74% in our data from 2004 through 2021) remain active for more than one year.[1] Third, and most importantly, a new agent's initial match with a counterparty agent is quasi-randomly assigned. We exploit these characteristics to examine how new agents' initial trading network connections affect their career success.

Our identifying assumption that new agents enter the trading network through a quasi-random matching process is highly plausible given the institutional structure of residential brokerage and the nature of housing markets. Housing markets are notoriously thin and subject to substantial matching frictions (Bayer et al., 2008; Han and Strange, 2015): transactions are infrequent, buyers and sellers have heterogeneous preferences, and agents can only interact with the counterparties involved in the listings their clients choose to pursue. These frictions leave new agents with very limited scope to strategically select their initial counterparty. Instead, the identity and network position of a new agent's first match are largely determined by which incumbent agents happen to be active (i.e., representing a buyer or seller) at the time

[1] This finding aligns with a recent industry report indicating that nearly half of new real estate agents who closed their first transaction in 2022 were unable to close another transaction in 2023 (Relitix, 2024).

of the transaction, combined with the idiosyncratic search behavior of their clients. Taken together, these features imply that, conditional on local market structure and contemporaneous incumbent agent activity, the network connectedness of a new agent's first counterparty is plausibly quasi-random.

To formally assess the quasi-randomness of the initial counterparty match, we conduct balance tests using the new agent's first transaction characteristics. We compare entrants within the same zip-by-year cell to test whether those matched to more central counterparties begin on observably different deals. The evidence shows only limited differences on a few margins, and these do not explain the main relationship between initial counterparty centrality and career survivability once we control for first transaction characteristics. The results indicate that the sample is broadly consistent with quasi-random matching across markets.

To motivate our empirical analysis, we develop a dynamic Bayesian network model in which entrant agents are quasi-randomly matched to counterparty incumbents with heterogeneous network centrality. Each new agent internalizes the centrality of their initial counterparty through a strictly concave internalization technology, generating future network opportunities and compounding spillover benefits. The model captures how career paths diverge based solely on the initial network position of one's first counterparty and formalizes how early matching dynamics influence future connectivity, transaction volume, and survival.

The model yields two central predictions. First, new agents who are initially matched with well-connected counterparty agents should exhibit higher survival rates and longer career durations than those matched with less-connected counterparty agents. This first prediction underscores the role of early network embeddedness in shaping career trajectories. Second, the marginal benefit of counterparty centrality declines as agents accumulate experience and further develop their own networks. This second prediction highlights the significant role that network capital plays at the outset of a career, when agents are unestablished and most dependent on their network position.

We test these predictions using transaction-level Multiple Listing Service (MLS) data from ten large U.S. cities between January 2001 and June 2023. The data includes over 6 million residential real estate transactions. Each record captures a completed sale involving a listing agent and a buyer agent, allowing for the construction of agent-to-agent trading networks for each of the ten cities. To capture the dynamic

evolution of professional ties, we segment the dataset into rolling three-month intervals and construct backward-looking three-year networks for each window. Within each network, nodes represent agents and directed edges represent transactions from listing to buyer agents. Repeated interactions form a directed multi-graph as depicted in Figure 1, which we then collapse into a weighted directed graph that aggregates transaction volume, allowing for a richer relational characterization. Figure 2 illustrates an example of the reduced-form directed real estate agent network for Charlotte, North Carolina, constructed from transactions in the second quarter of 2006. We compute eigenvector centrality as our primary network centrality measure, which captures the idea that new agents become more influential when they are connected to other agents who themselves occupy prominent positions.

We find empirical evidence consistent with our theoretical predictions: new real estate agents who transact with well-connected counterparty agents in their first deal are significantly more likely to survive than new agents who transact with less-connected counterparty agents. In our baseline specifications, we find that a one-standard-deviation increase in the network centrality of the initial counterparty agent increases the probability of surviving one year by 4.1 to 4.8 percentage points. This translates to a roughly 6 to 7% increase from the baseline one-year survival rate. These results are robust to additively and multiplicatively separable sale year and zip code fixed effects, ensuring comparisons are drawn among new agents who enter the profession around the same time and work in the same neighborhood.[2]

We interpret the higher survival rate as evidence that new agents benefit from transacting with well-connected counterparty agents. It seems intuitive that transacting with a well-connected agent would positively impact the early career success of a new agent more than a less-connected counterparty agent. The benefits of a network connection may be explicit; a well-connected counterparty agent can provide mentorship and offer access to their professional network. Or the benefits of the network connection could be implicit in that the new agent observed the actions and behavior of the well-connected counterparty agent and mirrored them in subsequent transactions. Regardless, the new agents' initial transaction with the well-connected counterparty increases their likelihood of survival.

[2]In Section 5.2, we provide robustness results under the survival analysis framework using the Cox proportional hazard models (Cox, 1972).

Typically, when examining networks' impact on an outcome, other factors that might be correlated with position in the network and the outcome itself must be carefully accounted for. We are not overly concerned about these factors in our study because the new agents we focus on are, by definition, not yet part of the trading network. They enter the trading network via a quasi-random matching process - where the counterparty's position determines theirs. One could argue that new agents already have professional (but not trading) connections to the other agents in their brokerage office. We mitigate this potential confounding effect by including fixed effects for the brokerage office, thereby controlling for differences in organizational environments (e.g., training, equipment, within-office mentorship, etc.). We continue to find that new real estate agents who transact with well-connected counterparty agents in their first deal are significantly more likely to survive than new agents who transact with less-connected counterparty agents.

Initial network connections shape not only whether new agents survive but also how productive they are in the early stages of their careers. We find that new agents who transact with well-connected counterparties have a significantly higher first-year and longer-term transaction volume, become more connected in the trading network, and ultimately have longer real estate careers. These patterns are consistent with the rapid accumulation of network capital in which early exposure to well-connected agents expands access to information and deal flow, helping new entrants become more central themselves. This early embeddedness, in turn, is a strong predictor of longer-term survival, underscoring the persistent and economically meaningful role of initial network ties in shaping new agents' career trajectories.

Our study contributes to several strands of research at the intersection of network, labor, talent, and financial economics. A substantial literature examines how networks shape career outcomes (e.g., Montgomery 1991; Calvó-Armengol and Jackson 2004; Beaman and Magruder 2012; Kramarz and Skans 2014; Cattan et al. 2025), yet causal interpretation remains challenging due to the endogenous nature of network formation (e.g., Manski 1993; Bala and Goyal 2000; Jackson 2008). We address this challenge by exploiting the quasi-random matching of new real estate agents to their first network connection. This novel setting allows us to provide new evidence that networks have a causal effect on career success.

Our findings contribute to a growing literature that emphasizes the compounding benefits of early opportunities and mentorship (e.g., Athey et al. 2000; Cunha and

Heckman 2007; Hilmer and Hilmer 2007; Blau et al. 2010). In line with recent evidence on the power of mentorship and peer interventions (e.g., Resnjanskij et al. 2024; Falk et al. 2026), we show that early network capital exerts persistent, long-term effects on the new entrant's survival and performance. Our findings demonstrate how early advantages - even those that are randomly bestowed, such as an initial transactional tie to a well-connected counterparty agent - can have a sizable impact on an individual's career trajectory.

Our study contributes to the literature on how network structures shape financial outcomes and deal flow. Prior research establishes that network centrality confers significant informational and performance advantages. For instance, Di Maggio et al. (2019) demonstrate that central brokers in the equity market facilitate information diffusion from informed order flows, enabling their best clients to capture superior returns. Similarly, Rossi et al. (2018) find that central pension fund managers earn higher risk-adjusted returns and attract greater capital inflows, attributing this outperformance to superior access to investment opportunities. We distinguish our work by shifting the focus from information transmission within established networks to the path-dependent effects of network entry. Focusing on the entry margin, we show that initial exposure to a well-connected counterparty leads new entrants to internalize incumbent centrality spillovers, generating persistent outcome heterogeneity reflected in survival trajectories and cumulative productivity.

## 2 Theoretical Framework

In this section, we develop a dynamic Bayesian network model to study how real estate agents' initial professional connections shape short-term survival and longer-term career success. New agents are randomly matched to established agents with varying network centrality and internalize their counterpart's influence through a strictly concave internalization technology. Those matched with high-centrality counterparty agents benefit from spillover effects that compound over time, producing heterogeneous trajectories from symmetric entry.

## 2.1 The Model

### 2.1.1 Set-up

Suppose two new agents enter the trading network simultaneously. There is an existing network (graph) $G = (V^\theta, E; \Theta)$, where node $v_i^\theta \in V^\theta$ represents agent $i \in \{a, b\}$ in the sub-network $V^\theta$ where $\theta \in \Theta$ with $\Theta = \{H, L\}$. The agents in the existing network are partitioned into two mutually exclusive groups with the network centrality measure such that $V^H \cup V^L = V$ and

$$c_i = \frac{1}{\lambda} \sum_{j=1}^{n} \alpha_{ij} c_j \tag{2.1}$$

$$v_i^\theta \in \begin{cases} V^H & \text{if } c_i \geq c^* \\ V^L & \text{otherwise.} \end{cases} \tag{2.2}$$

Given agents $i$ and $j$ are in $V^\theta$, the network centrality of agent $i$ is denoted as $c_i$ with constant $\lambda$ and relative centrality score $c_j$ where $A = [\alpha_{ij}]$ denote its adjacency matrix where $\alpha_{ij} = 1$ if there is an edge from node $i$ to node $j$, and $\alpha_{ij} = 0$ otherwise.[3] For an exogenous critical threshold $c^*$, an agent with a network centrality measure exceeding $c^*$ is classified as belonging to the high-centrality group, otherwise, the agent is considered part of the low-centrality group. The partitioning of the network reflects a key feature of the real estate brokerage industry, which is markedly bifurcated between well-connected, highly successful agents and those who are less established with weaker links.

### 2.1.2 Initial Connection & Centrality Internalization

In the first stage, nature assigns each entrant $v_i^\theta$ with a unique link $a_{ij}$ to an incumbent node $v_j^\theta \in V^\theta$ via a draw from $U[0, 1]$ with $j \in \{1, ..., n\}$.[4] The entrant then observes whether the matched counterparty is of high or low centrality.[5] Each entrant is endowed with a strictly concave, continuously differentiable internalization technology $\varphi_i \in [0, 1]$,

[3] In the context of our analysis, we use the weighted eigenvector centrality measure. Importantly, the use of alternative centrality metrics does not affect the model's key outcomes.

[4] Edge direction is not first-order in the model; empirically, we use direction to distinguish listing versus buying agents.

[5] We assume the information revealed includes the counterparty's centrality type.

chosen after observing the counterparty type and symmetric in equilibrium.[6] Upon forming the first connection, the entrant internalizes the counterparty's centrality, yielding payoffs

$$\pi_i = \varphi_i c_j, \tag{2.3}$$

which converts a fraction of the initially connected counterparty agent $j$'s centrality. In other words, the new agent $i$ internalizes the spillover effect of its initial counterparty $j$'s network connectedness.

### 2.1.3 Uncertainty: Network Evolution and Agent Survival

Following the initial connection, we model the new agent's subsequent career trajectory as a dynamic Bayesian network formation process unfolding over discrete time periods $t \in \mathbb{N}$. Two sources of uncertainty shape this dynamic path: (1) the stochastic evolution of the agent's network position and (2) the agent's own survival process.

We formalize the agent's network evolution using a time-homogeneous Markov chain. Given the partition of the network into high-centrality ($V^H$) and low-centrality ($V^L$) groups, the transition probabilities are defined as:

$$\Pr\left(v_{i,t}^{\theta} \in V^{s'} \mid v_{i,t-1}^{\theta} \in V^{s}\right) = \xi_{ss'}, \qquad s, s' \in \{H, L\}. \tag{2.4}$$

The key feature of this structure is that agents previously matched with high-centrality nodes are more likely to encounter high-centrality counterparts in the future (i.e., $\xi_{HH} > \xi_{LH}$). This captures a positive feedback loop from early exposure to central agents, generating compounding network benefits over time.

Second, we model the new agent's survival as a discrete-time Markov process. The probability that agent $i$ survives to period $t+1$ conditional on being active in period $t$ is given by $\mathbb{P}(i \text{ survives at } t+1 \mid i \text{ survives at } t) = \rho_t$ with $\rho_t \in (0,1)$. The cumulative survival probability up to period $t$ is:

$$S_i(t) = \prod_{\tau=1}^{t} \rho_\tau. \tag{2.5}$$

[6] Agent $i$'s action is equivalent to the internalization technology. Specifically, $\phi_i = g(e_i)$ where $e_i$ is the $i$'s effort level and $g(\cdot)$ is the technology. For simplicity, it is represented as $\phi_i$.

This formulation implies monotonicity of the survival function since $\rho_{t+1} < 1$. Therefore, under standard assumptions, survival probability is weakly decreasing over time. Thus, these two layers of uncertainty, dynamic shifts in network access and declining survival probability, jointly determine the agent's lifetime payoff and the stochastic evolution of their career success.

### 2.1.4 Initial Expected Payoff

We begin by characterizing the agent's initial expected payoff, which depends on the probability that the assigned counterparty $j$ belongs to the high-centrality subgraph, along with the associated conditional expected payoff, and similarly for the low-centrality subgraph. Since new agents are rational and forward-looking, the symmetric equilibrium in which each agent maximizes its expected payoff conditional on the counterparty's network centrality information entails full commitment of effort at entry.

**Proposition 1.** *In a symmetric Bayesian Nash equilibrium, each entrant fully commits to internalize the counterparty's network centrality. In particular,*

$$\varphi_i^* = 1 \quad \textit{with} \quad \mathbb{E}[\pi_i^*] = \frac{\varphi_i^*}{2}. \tag{2.6}$$

**Proof.** *See Appendix A.*

### 2.1.5 Persistent Impact of Counterparty Centrality

We extend the baseline model by allowing the new agent's survival probability to depend not only on the independent survival probability but also on the initial counterparty's network centrality and the new agent's first period transaction volume. This extension captures the stylized fact that early deal flow is a key determinant of career longevity. Let $V_i^{(1)}$ denote the total number of transactions new agent $i$ completes in the first period, and let $c_{j,0}$ denote the centrality of the counterparty in agent $i$'s initial match at $t = 0$. We define the expected lifetime payoff as

$$\mathbb{E}[\tilde{W}_i] = \varphi_i \sum_{t=1}^{\infty} \delta^t \, S_i(t) \, \mathbb{E}[c_{j,t}], \qquad S_i(t) \equiv \left(1 - e^{-\nu \, c_{j,0}(V_i^{(1)}+1)}\right)\rho^t, \tag{2.7}$$

where $S_i(t)$ is the probability that agent $i$ remains active at period $t$, $\nu > 0$ governs the strength of entry spillovers, and $\rho \in (0,1)$ is a constant per-period retention probability (i.e., $\rho_t \equiv \rho$).[7]

**Proposition 2.** *Initial exposure to high-centrality counterparties exerts a persistent positive effect on agent outcomes. Specifically, an increase in initial counterparty centrality $c_{j,0}$ strictly increases the survival probability in all future periods $t \in \mathbb{N}$, thereby raising the agent's expected lifetime payoff. Formally, we have $\frac{\partial S_i(t)}{\partial c_{j,0}} > 0$, which implies $\frac{\partial \mathbb{E}[\tilde{W}_i]}{\partial c_{j,0}} > 0$.*

**Proof.** *See Appendix A.*

This proposition formalizes the idea that early exposure to well-connected counterparties confers durable career advantages by enhancing network access, generating reputational spillovers, and facilitating knowledge transfer. High-centrality agents act as gateways to more valuable parts of the trading network, enabling new agents to establish credibility and access further opportunities more rapidly.

### 2.1.6 Diminishing Marginal Returns to Network Capital

While an initial match with a high-centrality counterparty confers a meaningful survival advantage, the marginal benefit of this network capital is not constant. We investigate the curvature of the return function to determine if the advantage exhibits diminishing returns.

**Proposition 3.** *The agent's expected lifetime payoff exhibits diminishing marginal returns with respect to the initial counterparty's centrality. That is, while better connections strictly increase expected welfare, the incremental gain decreases as centrality increases:*

$$\frac{\partial^2 \mathbb{E}[\tilde{W}_i]}{\partial c_{j,0}^2} < 0. \tag{2.8}$$

**Proof.** *See Appendix A.*

[7] In this set-up, if agent $i$ transacts as a dual agent in its first deal (i.e., $c_{j,0} = c_{i,0} = 0$), the survival probability is zero. Although initial dual-agent transactions provide short-term monetary benefits, we empirically find that higher commission earned in the dual agent transaction does not offset the initial loss of network connectivity. Hence, for simplicity, we maintain this structure in the extension.

## 2.2 Theoretical Hypotheses

We summarize the model's key mechanisms into testable predictions about how entry-time network exposure shapes early-career outcomes. We focus on two implications: (i) whether being initially matched with a more central incumbent improves a new agent's survivability, and (ii) whether these network-based advantages attenuate as the entrant builds its own network base.

**Prediction 1.** *New agents who are initially matched with high-centrality counterparties will exhibit superior survivability compared to those matched with low-centrality counterparties.*

This prediction is grounded in Proposition 2, which shows that initial exposure to central counterparties increases the survival probability. Consistent with prior evidence (Rossi et al., 2018; Di Maggio et al., 2019; Han et al., 2024), high-centrality agents serve as conduits of information and deal flow. A new agent linked to a well-connected peer enters a denser part of the transaction network, accelerating their access to future opportunities and increasing visibility among potential collaborators.

**Prediction 2.** *The marginal returns to counterparty centrality decline as entrants become established.*

This prediction is a direct implication of Proposition 3. The marginal returns of network capital internalized through initial match with an incumbent declines with time. While network spillovers are critical during early career stages, agents eventually generate their own connections and reputation, reducing reliance on their initial counterpart's network position (Kramarz and Skans, 2014; Cattan et al., 2025).

# 3 Data and Networks

This section outlines the dataset and the methodology used to build the agent networks and the new agent dataset from housing transaction records.

## 3.1 Data Overview

Our analysis draws on MLS housing transaction data from Cotality for ten major U.S. cities, covering the period from January 2001 through June 2023. The data

include detailed information on the unit type (single-family detached, condominium, or townhouse), physical characteristics (e.g., bedrooms, bathrooms), location (e.g., zip code), and transaction details (e.g., sale date, sale price). The data also contains fields - most importantly, listing agent and buyer agent identifiers - which enable us to construct agent networks within each city. These networks are inherently city-specific because real estate agents almost always operate exclusively within their local markets. For example, agents active in Boston do not typically also work in Houston.[8]

We selected the ten cities using the following criteria. First, each city has continuous data available from 2001 to 2023. Second, the city has at least 200,000 transactions during the study period. Third, a single MLS operated in the MSA during the study period. This criterion facilitates the creation of within-city agent networks because agent membership and IDs differ across MLS platforms. To verify this condition, we measure the share of duplicate listings appearing across multiple MLS platforms. We set a threshold of less than 5%, but in most cases, fewer than 1% of listings appear on more than one MLS platform across the ten selected cities. Fourth, listing and buyer agent IDs must be reliably populated, as these fields are essential for network construction. Finally, we prioritize cities that offer geographic diversity and whose urban cores do not cross state lines. Based on these criteria, the selected cities are Boston, MA (BOS); Cleveland, OH (CLE); Charlotte, NC (CLT); Columbus, OH (COL); Detroit, MI (DET); Houston, TX (HOU); Jacksonville, FL (JAX); Minneapolis, MN (MIN); Oklahoma City, OK (OKC); and Tucson, AZ (TUC).

After selecting the ten cities, we filter the transaction data for each city to remove outliers (e.g., list or sale prices above $2 million) and distressed transactions, as well as to exclude records missing key information, including cases where the listing agent ID is not available. The internet appendix provides a detailed list of the filters. The 10-city combined filtered dataset contains just over 6 million housing transactions involving 572,253 unique real estate agents. Panel A of Table 1 displays summary statistics for the filtered housing transaction dataset and Table 2 provides the variable definitions.

[8]The internet appendix provides evidence supporting this conjecture. Specifically, we show that there is no overlap in agent IDs and very little overlap in agent names and cell phone numbers across the ten cities.

## 3.2 Constructing Networks

We construct an agent network for each of the ten cities separately. To capture the evolving structure of agent interactions, we construct a series of dynamic networks over rolling time windows. We segment the full sample period into sequential three-month intervals, starting in January, April, July, and October of each calendar year. For each time window, we define a backward-looking three-year observation period and include all transactions completed during that window to construct the agent network. Finally, we aggregate the panel of dynamic network series for each of the 10 cities.

In each resulting network, nodes represent individual real estate agents, and directed edges represent housing transactions, oriented from the listing agent to the buyer agent. Because agents may transact multiple times with the same counterparty agent, the network takes the form of a directed multigraph, allowing multiple edges between agent pairs. Each edge reflects a distinct transaction and is labeled with the corresponding sale price, thereby capturing both the frequency and economic value of interactions.

Figure 1 illustrates an example of our constructed network. The left graph shows the raw directed multigraph where repeated interactions between agents are represented as multiple edges. This detailed representation enables a fine-grained view of market behavior, such as patterns of repeated cooperation or agent specialization. The right graph presents a reduced-form version of the multigraph, in which multiple edges between the same agent pair are aggregated into a single directed edge. Each edge in the collapsed graph is annotated with the number of transactions ($n$), total transaction volume ($TP$), and average transaction price ($AP$), summarizing the intensity and economic scale of the relationship. Figure 2 displays the reduced-ford directed graph for Charlotte, NC in Q2 2006, illustrating what the actual agent network look like at scale in our sample.

While the multigraph preserves the granularity of individual interactions, we use the reduced-form directed graph as the input for centrality analysis. This transformation allows us to capture both the presence and intensity of agent relationships, facilitating the computation of weighted network measures. Economically, frequent and high-value repeated transactions between two agents may signal trust, shared clientele, or strategic alignment. These relational dynamics, in turn, shape agents' access to information, client flows, and future opportunities, all of which are critical for understanding influence within the marketplace.

## 3.3 Measuring Network Centrality

The centrality of a node in a network is a micro-level structural feature that facilitates comparisons among nodes by quantifying their relative importance within the network (Jackson, 2008). Various centrality measures exist, each capturing a different notion of importance, such as degree, closeness, or betweenness. In this study, we employ eigenvector centrality to assess an agent's influence, not only through its direct connections but also through the centrality of the agents to which it is connected. This recursive structure is particularly relevant in the residential real estate brokerage setting, where indirect connections and the prominence of counterparty agents' neighbors can enhance visibility, information access, or deal flow.

Our network is a directed multigraph, with edges oriented from listing agents to buyer agents, a natural direction that mirrors the flow of the property from seller to buyer. To reflect this structure, we compute three measures of eigenvector centrality: in-centrality, out-centrality, and combined centrality.

*In-centrality* is computed based on incoming edges; that is, an agent is central if it frequently transacts as a buyer agent with other highly central listing agents. The in-centrality of agent $a$ is defined as:

$$C_a^{in} = \frac{1}{\lambda^{in}} \sum_{b \in \mathcal{N}^-(a)} A_{ba} C_b^{in}, \tag{3.1}$$

where $\mathcal{N}^-(a)$ denotes the set of agents with directed edges pointing to agent $a$ (i.e., its in-neighbors), $A_{ba}$ is the adjacency matrix entry indicating the presence and weight of an edge from agent $b$ to agent $a$, and $\lambda^{in}$ is the largest eigenvalue of the adjacency matrix. This formulation emphasizes that an agent's in-centrality is recursively determined by the centrality of those agents with whom they transact as listing agents.

*Out-centrality* is computed based on outgoing edges; that is, an agent is central if it frequently transacts as a listing agent with other highly central buyer agents. The out-centrality of agent $a$ is defined as:

$$C_a^{out} = \frac{1}{\lambda^{out}} \sum_{b \in \mathcal{N}^+(a)} A_{ab} C_b^{out}, \tag{3.2}$$

where $\mathcal{N}^+(a)$ denotes the set of agents with directed edges originating from agent $a$ (i.e., its out-neighbors), $A_{ab}$ is the adjacency matrix entry indicating the presence and weight of an edge from agent $a$ to agent $b$, and $\lambda^{out}$ is the largest eigenvalue of the transposed adjacency matrix.

*Combined centrality* captures an agent's overall network position by averaging the in- and out-centrality measures:

$$C_a^{combined} = \frac{C_a^{in} + C_a^{out}}{2}. \tag{3.3}$$

Economically, in-centrality reflects an agent's prominence as a buyer-side representative who transacts with well-connected listing agents, while out-centrality reflects prominence as a listing agent who transacts with well-connected buyer agents. Combined centrality provides a holistic measure of an agent's influence across both sides of the market.

We expand the unweighted specification by computing the weighted eigenvector centrality in which the adjacency matrix entries $A_{ba}$ (for in-centrality) and $A_{ab}$ (for out-centrality) are defined by the number of transactions between agents. This approach captures not only the presence of a connection but also its intensity. Specifically, repeated interactions between the same pair of agents - such as multiple transactions initiated by a listing agent toward the same buyer agent - are weighted more heavily in the centrality calculation.

Incorporating transaction frequency as a weight provides richer economic interpretation. In real estate markets, repeated collaborations may reflect underlying information channels, professional trust, or brokerage networks that facilitate more efficient deal-making. Agents who are repeatedly selected by other central agents are likely to have stronger reputational capital, broader market reach, or superior client matching capabilities. By capturing both the structure and depth of relationships, the weighted version of eigenvector centrality offers a more nuanced measure of influence within the agent network.

### 3.4 New Agents

We define a new agent as one with no prior transactions in the data. Although we have transaction data from January 2001 through June 2023, our empirical analysis uses only data between January 2004 and June 2021. This allows us to (i) identify new agents with no transactions from 2001 through 2003, (ii) construct the backward-looking three-year network centrality measures for the first counterparty agent that the new agents transact with, and (iii) determine whether the new agent survived one year. We identify 217,837 new agents who entered the real estate trading network between January 2004 and the first half of 2021.

A challenge arises because a new agent's first transaction may not involve a "traditional" full-service counterparty agent. This issue is particularly relevant when the new agent represents the buyer, as the listing agent may instead be an entry-only or flat-fee listing service, a limited-service provider, or a team of several agents that are responsible for posting the property. These "non-traditional" agent IDs often generate spuriously high transaction volumes and, thus, high network centrality measures, but they do not correspond to individual, socially connected agents in the conventional network sense. Because the MLS data does not include a field that identifies these non-traditional agent IDs, we mitigate this concern by trimming agent IDs based on their counterparty network centrality, removing the top 5% within each city for every rolling three-month window. We also drop new agents with more than 20 transactions in their first year to remove non-traditional new agents.[9]

Panel B of Table 1 presents summary statistics for 181,766 new agents after trimming the counterparty agent network data to remove non-traditional agent IDs. The first three rows of Panel B display summary statistics for the three counterparty network centrality measures: in, out, and combined. These rolling quarterly centrality measures are normalized to account for differences in the timing of agents' entry into the network and variation across the ten cities.

The next five rows of Panel B in Table 1 display summary statistics for a set of mutually exclusive control variables that classify whether the agents involved in a transaction

[9]The internet appendix demonstrates that our results are robust to trimming the top 1% or 3% of counterparty agent IDs based on their network centrality as well as the inclusion of new agents with more than 20 transactions. The internet appendix also provides results showing that our findings are similarly unaffected when, instead of trimming, we flag the top 5% of non-traditional counterparty agent IDs with an indicator variable and include it as an additional control.

operate at arm's length. *Dual Agent* is an indicator variable, $Dual_{i,t}$, that equals 1 if new agent $i$ represents both the buyer and the seller in the transaction at time $t$; otherwise, it equals 0. Because no counterparty agent $j$ exists in such cases, the counterparty network centrality measure, $Centrality_{j,t}$, is set to 0 when $Dual_{i,t} = 1$. *In-house Agent* is an indicator variable, $Inhouse_{i,j,t}$, that equals 1 if new agent $i$ and counterparty agent $j$ work for the same brokerage office; otherwise, it equals 0. By construction, $Inhouse_{i,j,t} = 0$ when $Dual_{i,t} = 1$. Together, these two variables capture non-arm's-length relationships between the agents involved in the transaction. The summary statistics indicate that 7% of new agents' first transactions involve dual agency, and an additional 7% occur within the same brokerage office, reflecting a nontrivial incidence of non-arm's-length relationships in initial market entry.

The next three control variables identify arm's-length relationships between the agents involved in the transaction. *Listing Agent* is an indicator variable, $Listing_{i,j,t}$, that equals 1 if new agent $i$ represents the seller in an arm's-length transaction with buyer agent $j$ who is a MLS member, and 0 otherwise. *Counterparty Non-member* is an indicator variable, $CA_nonmember_{i,j,t}$, that equals 1 if new agent $i$ represents the seller in an arm's-length transaction with buyer agent $j$ who is not a MLS member, and 0 otherwise.[10] When $CA_nonmember_{i,j,t} = 1$, the counterparty agent $j$ is not part of the MLS network, so their network centrality, $Centrality_{j,t}$, equals 0. The final relationship control variable, *Buyer Agent Member* is an indicator variable, $Buyer_{i,j,t}$, that equals 1 if new agent $i$ represents the buyer in an arm's-length transaction with listing agent $j$ who is a MLS member, and 0 otherwise. The summary statistics indicate that new agents are much more likely to provide buyer-side representation in their first transaction than seller-side representation (i.e., listing agent).

The remaining rows of Panel B in Table 1 display summary stats for the unit types and outcome variables related to the new agents' survival rate and transaction volume. The unit type summary statistics indicate that 85% of new agents' first transactions involved a single-family detached house, 13% a condominium, and 2% a townhome, which is very similar to the unit type distribution for the full transaction sample in Panel A. The survival rate summary statistics indicate that 67% of the new agents survive for more than 1 year, 58% for more than 2 years, and only 48% for more than 3 years. The

[10] Only real estate agents that are members of the MLS can list properties on the MLS, so there are no *Counterparty Non-member* transactions in which new agent $i$ represents the buyer in an arm's-length transaction with listing agent $j$ who is not a MLS member.

average survival length of new agents in the data is 1,686 days, or roughly four and a half years. Figure 3 highlights how new agents' survival length varies over time with market conditions. Finally, the transaction volume summary statistics indicate that new agents average 3.43 transactions in their first year and 23.06 total transactions in the data.

## 3.5 Case Study: Evolution of the Networks

To understand how early network positioning influences long-term connectivity, we analyze the structural evolution of agents' networks over time, with a particular focus on treatment and control groups defined by new agents' initial exposure to well-connected counterparty agents. As a case study, we begin by identifying new agents whose first recorded transaction occurs within a randomly selected reference window (Q2 2006) in Union County (Charlotte, NC) as depicted in Figure 2. Using these transactions, we construct a baseline agent network and compute centrality measures for all counterparty agents involved. Agents whose first transaction is with a counterparty in the top 25th percentile of eigenvector centrality are classified as the treatment group, while all others are assigned to the control group.

Figure 4 illustrates the longitudinal evolution of average degree, in-degree, and out-degree for treatment and control agents. In the baseline period, both groups exhibit similar levels of network connectivity. However, over time, treatment agents demonstrate a systematically higher degree growth compared to control agents. This divergence is particularly pronounced for in-degree, suggesting that treatment agents become increasingly prominent as recipients of transactions, possibly due to reputational spillovers or preferential attachment mechanisms. These dynamics imply that early access to well-connected agents may serve as a catalyst for sustained integration, influence, and deal flow within the network.

Figure 5 provides a complementary visualization of network evolution by displaying snapshots of the agent network at four time points: Q2 2006, 2012, 2018, and 2023. Red nodes represent treatment agents, green nodes represent control agents, and gray nodes correspond to all other agents. Node size is scaled by degree to reflect each agent's connectivity in the network. In the initial period (Q2 2006), treatment and control agents exhibit relatively similar node sizes, indicating comparable levels of engagement. However, as time progresses, treatment agents show visibly larger node sizes compared to control agents, reflecting a more substantial increase in transaction volume and

network participation. This growth suggests that treatment agents become more active over time, while control agents tend to exhibit more modest changes in degree.

Together, these patterns highlight the persistent effect of early network positioning. Agents who begin their careers transacting with central market players appear more likely to build and sustain a broad and active set of connections. These findings support the notion that early network capital plays a critical role in shaping professional success in decentralized markets such as real estate. We formally test this notion in the following section.

# 4 Methodology and Results

## 4.1 Empirical Strategy

We exploit the quasi-random matching of real estate agents in housing transactions to examine whether new agents who transact with well-connected counterparty agents have a higher survival rate than those who transact with less-connected counterparty agents. We estimate agent-level regressions of the following form:

$$Survive_{i,j,t} = \beta_1 Centrality_{j,t} + \beta_X X_{i,j,t} + \beta_3 Inhouse_{j,t} + \lambda_t + \varepsilon_{i,j,t} \tag{4.1}$$

The dependent variable in Equation 4.1, $Survive_{i,j,t}$, is a binary indicator identifying whether real estate agent $i$ remains active - that is, has another transaction in the data - after a specified duration following their initial transaction with counterparty agent $j$ at time $t$. Our baseline analyses focus on a one-year survival period, for which $Survive_{i,j,t}$ = 1 if the new agent completes at least one additional transaction more than 365 days after their first transaction and 0 otherwise. We also examine longer survival periods, including a continuous measure of $Survive_{i,j,t}$ that denotes the new agent's survival length as the number of days between their first and last transactions in the data.

In Equation 4.1, the primary variable of interest, $Centrality_{j,t}$, is one of the three time-varying network centrality measures (in, out, or combined) for the first counterparty agent $j$ that the new agent $i$ transacts with in the data. We estimate Equation 4.1 with and without a set of agent relationship and unit type controls, $X_{i,j,t}$. The five mutually exclusive agent relationship controls identify whether the agents

involved in the transaction have an arm's-length relationship. The unit type controls indicate whether the new agent's first transaction involved a single-family detached house, a condominium, or a townhome. To avoid perfect multicollinearity, we omit one of the agent relationship (Buyer Agent) and unit type (Single-family Detached) controls. Finally, Equation 4.1 also includes additively or multiplicatively separable zip code and sale year fixed effects, $\lambda_t$, for the new agent's first transaction at time $t$. Including location and time fixed effects allows us to isolate the impact of network connections on the early career success of new real estate agents who work in the same area and enter the profession at the same time.

## 4.2 New Agent Survival

Table 3 reports the regression estimates from Equation 4.1, which assess the extent to which a new agent's initial network connection influences their early career survival. Panel A reports estimates from our baseline estimation of Equation 4.1 that does not include the agent representation and unit type controls, whereas Panel B displays estimates from a version that includes the additional controls. Columns 1 and 2 report estimates for the *Counterparty Centrality In* measure, columns 3 and 4 for the *Counterparty Centrality Out* measure, and columns 5 and 6 for the *Counterparty Centrality Combined* measure. The dependent variable in each column is an indicator variable identifying whether the new agent survived one year.

The counterparty network centrality coefficient estimates are positive and statistically significant in all six columns in Panel A of Table 3. The counterparty centrality coefficient estimates indicate that the probability of survival for new agents increases with the connectedness of the first counterparty agent they transact with. Specifically, a one-standard-deviation increase in the counterparty agents' network centrality increases new agents' one-year survival rate by 2.7 to 6.6 percentage points. This translates to approximately a 4 to 10 percent increase from the baseline one-year survival rate of 67 percent.

After including the additional agent relationship and unit type controls, the counterparty network centrality coefficient estimates are positive in all six columns and statistically significant in five of the six columns in Panel B of Table 3. Although the Counterparty Centrality Out coefficient estimate is statistically insignificant when we include the more restrictive location-by-time fixed effects in column 4, it is important to

note that the majority of the new agents provide buyer-side representation in their first transaction. This implies that Counterparty Centrality In - which captures edges pointing from listing agents to buyer agents - or Counterparty Centrality Combined - which captures edges pointing in both directions - are more appropriate measures. Accordingly, we will focus on the Counterparty Centrality Combined measure and report results only for it going forward. In columns 5 and 6 of Panel B, a one-standard-deviation increase in the counterparty agents' network centrality increases new agents' one-year survival rate by 4.1 to 4.8 percentage points. This translates to a roughly 6 to 7 percent increase from the baseline one-year survival rate.

Consistent with Prediction 1, the results in Panels A and B of Table 3 indicate that new agents benefit from forming network connections with well-connected agents. The benefits may be explicit - through mentorship and direct access to the counterparty agents' network connections - or implicit, as new agents learn by observing and emulating the more experienced counterparty agents' behavior. In either case, transactions with well-connected counterparty agents significantly increase new agents' chances of short-term survival.

Panel B of Table 3 also reports the coefficient estimates for the agent relationship and unit type controls in Equation 4.1. The unit type controls are statistically insignificant across all six columns, indicating that new agents whose first transaction is a condominium or townhome are just as likely to survive as those whose first transaction is a single-family detached house. The rest of this section focuses on the two non-arm's-length (dual agent, in-house agent) and arm's length (listing agent, counterparty non-member) agent relationship controls. As noted in Section 4, the buyer agent control is omitted from Equation 4.1, so the coefficient estimates discussed below are all relative to new agents who provide buyer-side representation in transactions involving counterparty agents from unaffiliated brokerage offices.

The control for dual agent transactions highlights an interesting tradeoff that new agents face when representing both the buyer and seller in their first transaction. On the one hand, the new agent earns a larger commission for representing both the buyer and seller. The larger commission should arguably increase the new agent's probability of survival. On the other hand, the new agent does not develop a network connection with a counterparty agent. Assuming networks matter - which is the crux of this paper - the lack of network connectivity will decrease the new agent's probability of survival relative to new agents that transact with well-connected counterparty agents.

In Panel B of Table 3, the coefficient estimates on dual agent transactions are uniformly negative and statistically significant across all six specifications. These estimates indicate that new agents whose first transaction does not involve a counterparty agent are substantially less likely to survive than those who provide buyer-side representation with counterparty agents in other offices. In terms of survival, this implies that the higher commission earned in the dual agent transaction does not offset the initial loss of network connectivity. Consistent with our main findings, this result indicates that new agents' survival probability increases with the network connectedness of their initial counterparty.

One could argue that new agents already possess non-trading network ties within their brokerage office, which could influence their survival probability. Consequently, the brokerage office a new agent joins may shape their survival prospects. To partially address this issue, Equation 4.1 includes an in-house agent indicator that captures within-office relationships. This variable also helps mitigate concerns that new agents may benefit from referrals from more experienced colleagues. We examine this mechanism in greater detail later in the section.

The coefficient estimates for in-house agent transactions are positive across all six specifications in Panel B, but they become statistically insignificant in the even-numbered columns that incorporate zip-code-by-year fixed effects. This pattern indicates that new agents whose initial transactions involve counterparties within the same brokerage office are not significantly more likely to survive than those who provide buyer-side representation with counterparty agents in other offices. Taken together, the results suggest that new agent survival is primarily shaped by the network connectedness of the initial counterparty agent, rather than by shared brokerage affiliation. More broadly, although cross-office transactions may expose new agents to a wider segment of the trading network, the evidence is not consistent with the view that transacting across brokerage offices is itself a key determinant of early-career survival.

The coefficient estimates for the indicator for transactions in which the new agent represents the seller, i.e., serves as the listing agent, are negative and statistically significant across all six specifications in Panel B of Table 3. This finding indicates that agents who begin their real estate career on the seller side are significantly less likely to survive for one year than those who first represent buyers. Specifically, we find that new listing agents have a one-year survival rate that is 2.4 to 2.6 percentage points lower than new buyer agents. A plausible interpretation is that early buyer-side work

provides broader network exposure and more frequent interaction with experienced, well-connected counterparty agents that have direct control over local deal flow through their listings. We provide evidence consistent with this interpretation by plotting the average number of previous transactions, which serves as a rough proxy for experience and network connectivity, that listing agents have relative to buyer agents. Figure 6 shows that listing agents typically have two to three times as many prior transactions as buyer agents.[11]

Panel B of Table 3 also reports coefficient estimates for Counterparty Non-member transactions, which are similar to Listing Agent transactions in that the new agent represents the seller but differ because the counterparty buyer agent is not an MLS member. By definition, these counterparties are not part of the MLS network, so their network centrality is zero. Consistent with this lack of network connectivity, the coefficient estimates on the Counterparty Non-member indicator are negative and statistically insignificant across all six specifications. The magnitude of the estimates is economically significant: new listing agents whose first transaction involves a non-member counterparty have a one-year survival rate that is 9.8 to 10.4 percentage points lower than comparable new agents who begin on the buyer side. This finding is consistent with Prediction 1 in that initial transactions with unconnected non-member counterparty agents provide no direct opportunity for further network formation, thereby limiting subsequent deal flow and lowering short-run survival.

## 4.3 Arm's Length Agent Relationships

While the results in Panel B of Table 3 highlight the importance of controlling for agent relationships within MLS trading networks, they also raise concerns that non-member and non-arms-length relationships could distort the estimated effects of counterparty centrality. In particular, counterparty centrality is mechanically set to zero for all Counterparty Non-member and Dual Agent transactions, which could bias the counterparty centrality coefficient estimates. Table 4 addresses these concerns by removing these transactions from the sample. Columns 1 and 2 exclude Counterparty Non-member transactions, and columns 3 to 6 apply an additional filter that removes Dual Agent and In-house transactions, thereby focusing the analysis on arms-length

[11] The transaction data starts in 2001, so the average prior transaction volume is censored on the frontend of the study for both agent types.

agent interactions where the counterparty agents' network centrality is meaningfully defined.[12]

After restricting the sample to address the concerns outlined above, the estimates for counterparty network centrality remain positive and statistically significant across all six specifications in Table 4. The network centrality coefficient estimate in columns 3 and 4 of Table 4 is comparable to the estimate for the full sample in columns 5 and 6 of Panel B in Table 3, suggesting the subsample of arms-length agent transactions is similar to that of the full sample. Once again, these results indicate that new agents' one-year survival probabilities increase with the network connectedness of their initial counterparty agent. A one-standard-deviation increase in counterparty centrality combined is associated with a 3.7 to 4.5 percentage point increase in the likelihood of surviving one year, which corresponds to roughly a 6 to 7 percent improvement relative to the baseline survival rate in the data. This finding highlights the important role of network connections in shaping the short-run career success of new real estate agents.

## 4.4 Brokerage Office Fixed Effects

The coefficient estimates for in-house agent transactions are positive and statistically significant when we include additively separable time and location fixed effects in Equation 4.1, suggesting that within-brokerage office trading connections may enhance the likelihood of new agents' survival. However, this indicator likely understates the true effect of such connections, as it only captures cases where a new agent's first transaction involves a counterparty from the same office. To investigate this relationship further, we incorporate brokerage office fixed effects into Equation 4.1. The brokerage office fixed effects control for pre-existing brokerage office network connections as well as differences in organizational environments (e.g., training, equipment, within-office mentorship, etc.).

Columns 5 and 6 of Table 4 incorporate brokerage office fixed effects alongside additively or multiplicatively separable zip code and sale year fixed effects. The fixed effects allow us to compare new agents who begin their careers in the same year

[12]For brevity, Table 4 reports only the estimates for Counterparty Centrality Combined. The internet appendix provides the corresponding results for the Counterparty Centrality In and Out measures, which show that the arm's-length subsample yields similar results to the full sample estimates reported in Table 3.

and geographic area, while controlling for differences in organizational environments and access to resources. Introducing these fixed effects reduces the magnitude of the counterparty centrality coefficient estimate, yet it remains economically and statistically significant. A one-standard-deviation increase in the counterparty agent's network centrality increases a new agent's probability of surviving one year by 2.6 to 3.2 percentage points. This translates to a roughly 4 to 5% improvement relative to the baseline survival rate in the data. This finding underscores the importance of network connections: even when comparing new agents who share the same local market conditions and brokerage office support, those who transact with better-connected initial counterparty agents are more likely to survive. Therefore, the results indicate that initial network connections are a key determinant of new agents' short-term career success.

## 4.5 New Agent Transaction Cohorts

The results in the preceding sections indicate that the network connectedness of the first counterparty agent a new agent transacts with influences the likelihood of that agent's short-term survival. Additional factors - such as the number of transactions a new agent completes in their first year - are also likely to affect their survival. Intuitively, a higher transaction count in the first year should increase a new agent's survival chances by increasing both commission earnings (holding transaction prices constant) and connections within the counterparty agent network. We analyze the impact of first-year transaction volume on agent survival using two approaches.

First, we modify Equation 4.1 to include indicator variables for first-year transaction cohorts that identify new agents with 3-5, 6-10, and 11-20 first-year sales. We omit the transaction cohort indicator variable for new agents with 1-2 first-year sales. Columns 1 and 2 of Table 5 display the regression results from the modified equation using the full sample of new agents. As expected, the coefficient estimates for the first-year transaction cohorts monotonically increase with the new agents' first-year transaction volume. New agents with 3-5 first-year sales are 31% more likely to survive their first year than new agents with 1-2 first-year sales. Whereas new agents with 6-10 or 11-20 first-year sales are 43-47% more likely to survive their first year than those with 1-2 first-year sales. Importantly, the counterparty network centrality coefficient estimate remains statistically and economically significant. A one-standard-deviation

increase in the counterparty agents' network centrality increases new agents' one-year survival rate by 2.2 percentage points.

Second, we modify Equation 4.1 by introducing a continuous measure of first-year transaction volume for each new agent and an interaction between this variable and the counterparty centrality measure. Columns 3 and 4 of Table 5 report the resulting estimates for the full sample of new agents. The coefficients on both counterparty centrality and first-year transactions are positive and statistically significant, indicating that new agents' probability of survival increases with both the connectedness of the initial counterparty and the agent's early transaction activity. However, the coefficient on the interaction term is negative and statistically significant in both specifications, implying that the marginal effect of counterparty centrality declines as first-year transaction volume increases. This pattern is consistent with Prediction 2, in which early network connections are most valuable for agents with few transactions, but become less important once new agents begin generating sufficient endogenous transaction volume to sustain themselves independent of their initial counterparty's network connectedness. That said, the network centrality of these counterparties likely contributed to the higher transaction volumes to begin with. We examine this relationship in greater detail in the next section.

## 4.6 New Agent Transaction Volume

The preceding sections establish the important role that network connections play in new agents' survival in the real estate industry. Because real estate agents are compensated solely through commissions (i.e., they only earn income when a transaction closes), an important next step is to examine how these early network connections affect subsequent transaction volume. If new agents derive benefits and learn from transacting with well-connected counterparties, one natural channel through which these benefits may operate is an increase in subsequent transaction volume. Higher transaction volume not only provides additional income but also facilitates greater agent and client exposure, repeat-interaction opportunities, and overall network expansion, all of which increase the probability that new agents will survive.

Table 6 examines the effect of counterparty agents' network connectedness on the new agents' transaction volume. Each column regresses the log of new agents' transaction volume over the specified time horizon on counterparty agents' network centrality,

relationship controls, and zip-code-by-year fixed effects. We initially focus on new agents' first-year transaction volume, then expand the analysis to the longer term. Column 1 uses the full new agent sample in Panel B of Table 1. The remaining columns restrict the sample to new agents that enter the profession between 2004 and 2019. This restriction ensures a minimum 3.5-year post-entry window for evaluating longer-term transaction volumes.

The counterparty network centrality coefficient estimates are positive and statistically significant in columns 1 and 2 of Table 6. This finding indicates that new agents' transaction volume increases with the connectedness of the first counterparty agent they transact with. Specifically, a one-standard-deviation increase in the counterparty agents' network centrality increases new agents' first-year transaction volume by 8.8 to 9.2 percentage points.

The short-term effect on new agents' transaction volume in columns 1 and 2 is unsurprising, given that new agents' survival rate and transaction volume are interrelated. The longer a new agent survives, the more transactions they can complete. The more transactions a new agent completes, the longer they will likely survive. The results in columns 3 to 5 of Table 6 support this conjecture, showing that the magnitude of the network centrality coefficient estimate increases with the transaction window length. In column 3, a one-standard-deviation increase in counterparty agents' network centrality increases new agents' three-year transaction volume by 11.7 percentage points. Similarly, in column 4, a one-standard-deviation increase in counterparty agents' network centrality increases new agents' total transaction volume by 12.5 percentage points. These findings jointly reinforce Prediction 1 in that the long-run effects of initial network connections are not transitory and compound over time.

The final column in Table 6 adjusts for differences in new agents' survival length using the log of transactions per year as the dependent variable. As in the preceding columns, we find that new agents' average annual transaction volume increases with the network centrality of the first counterparty they transact with. A one-standard-deviation increase in the network centrality of counterparty agents increases new agents' average annual transaction volume by 5.3 percentage points.

The coefficient estimates for the agent relationship controls in Table 6 closely align with those reported in Table 3. The estimates for Dual Agent, Listing Agent, and Counterparty Non-member transactions are uniformly negative, indicating that new

agents entering the network via these relationship types generate lower subsequent transaction volume than those who enter as Buyer Agents. This reduced transaction volume is consistent with, and likely contributes to, the earlier finding that these agents exhibit lower short-run survival rates. In contrast, the coefficient estimates for In-house Agent transactions are positive, suggesting that new agents who transact with counterparty agents inside their brokerage office experience higher early-career transaction activity, which reinforces their higher observed survival probability. Taken together, these patterns highlight the extent to which initial relationship structures shape both the volume and survival of new agents early in their careers.

## 4.7 Longer Term Survival

Table 7 extends this analysis by evaluating the extent to which these relatively short-term effects persist over the longer term. Once again, we restrict the sample to new agents that enter the trading network between 2004 and 2019. Columns 1 and 2 of Table 7 report results from a regression using a modified version of Equation 4.1 where the dependent variable is the log of new agent $i$'s survival length. The new agents' survival length equals the number of days between their first and last transaction date in the data. In column 1, a one-standard-deviation increase in the counterparty agents' network centrality increases their survival length by 36-37.4%. This finding indicates that transacting with well-connected counterparty agents significantly increases new agents' career length. These findings suggest that the benefits of initial exposure to well-connected counterparties persist throughout the agent's career.

Columns 3 through 6 of Table 7 examine this dynamic in more detail by re-estimating Equation 4.1 with a two-year survival indicator as the dependent variable. Columns 3 and 4 use the entire sample of new agents that enter the profession between 2004 and 2019, whereas columns 5 and 6 further restrict the sample to agents who survived their first year. In columns 3 and 4, the coefficient on counterparty centrality is positive and statistically significant, indicating that early exposure to a more connected counterparty is associated with a higher probability of survival over two years.

Once the sample is conditioned on surviving the first year, however, the estimated effects in columns 5 and 6 decline in magnitude and become statistically insignificant. This attenuation suggests that the influence of initial counterparty connectedness operates primarily through early-career dynamics, such as accelerated learning, faster

integration into the local trading network, or improved access to deal flow. In other words, after the first year, survival no longer appears to be driven by the network centrality of the new agents' first network connection. This finding is intuitive, given that new agents likely become less reliant on their initial network connection as they complete subsequent transactions and become more embedded in the network.

This pattern does not suggest that network connections no longer matter as agents advance in their careers. Instead, it suggests that the relative importance of the initial connection diminishes as agents gain more experience and expand their trading networks. We investigate this dynamic in more detail by examining how the initial counterparty agent's connectivity affects the new agents' position within the network at the end of their first year. In Panel A of Table 8, the dependent variable is the new agent's end-of-year-one (EOY1) network centrality, which we regress on the network centrality of the first counterparty agent with whom they transact. Columns 1 and 2 of Panel A use the entire new agent sample, whereas columns 3 and 4 restrict the sample to new agents that enter the profession between 2004 and 2019 and survive their first year. The counterparty centrality coefficient estimates are positive and statistically significant in all four columns, indicating that new agents who begin their careers by transacting with more central counterparties tend to occupy more central positions themselves by the end of their first year. This pattern is consistent with persistence in early network advantages: initial ties to well-connected market participants appear to accelerate a new agent's integration into the broader trading network.

Next, we investigate how a new agent's position within the network at the end of their first year affects their longer-term survival. In Panel B of Table 8, the dependent variable is an indicator identifying whether the new agent survived, which we regress on their EOY1 network centrality. Columns 1 and 2 revisit the results in Table 7 by examining a two-year survival period, whereas columns 3 and 4 examine a three-year survival period. In both cases, we use a subsample of new agents that survived the previous year. The coefficient estimates in all four columns of Panel B are positive and statistically significant, indicating that new agents' EOY1 network centrality is a strong predictor of their survival the following year. Taken together, these results imply that initial network connectedness facilitates new agents' early-career network formation, which, in turn, increases their probability of survival over the longer term.

# 5 Robustness Tests

In this section, we present two robustness checks. First, we examine whether new entrants matched to more central counterparties are simply sorting into observably different first transactions by conducting balance tests and coefficient-stability exercises with increasingly richer and flexible controls. Second, we re-estimate the main relationship in a duration framework using hazard models, which exploits the full survival-length information and explicitly accounts for right-censoring near the end of the sample.

## 5.1 Balance Test

To assess whether initial counterparty centrality is capturing selection into observably different first transactions, we conduct a set of balance tests. The objective is to address the concern that new agents matched to more central counterparties may simply begin on systematically different deals, rather than benefiting from network exposure itself. Based on (4.1), for each new agent $i$ and initial transaction at time $t$ with counterparty agent $j$, we observe $Centrality_{j,t}$ and the first-transaction characteristics $X_{i,j,t}$, including sale price, property type, bedrooms, bathrooms, living area bin, house age bin, lot size bin, and transaction structure indicators.[13] The balance-test sample uses the same arm's-length MLS member first transactions as in the main analysis (excluding dual agency cases, non-member counterparties, and observations without an identified counterparty). Throughout, we compare new agents within the same local market and cohort using zip-by-year fixed effects.

First, we estimate a flexible balance specification that regresses initial counterparty centrality on first-transaction observables:

$$Centrality^{z}_{j,t} = \alpha + f\left(X_{i,j,t}\right) + \lambda_{z(i)\times y(t)} + u_{i,j,t}, \tag{5.1}$$

[13] We model first-transaction characteristics flexibly with indicator bins rather than linear terms. Controls include 100-square-foot living area bins, 10-year house age bins, bedroom and bathroom categories, lot size bins, property type, and an in-house transaction indicator. To limit tail sparsity, we top-code bedrooms and bathrooms at five or more and pool the largest living area bins. With zip-by-year comparisons, this specification flexibly captures nonlinear differences in first-transaction environments. We follow a similar approach from Goldsmith-Pinkham and Shue (2023) and Choi et al. (2025).

where $Centrality^{z}_{j,t}$ is the standardized value of the initial counterparty's centrality, $f(X_{i,j,t})$ is a saturated set of indicator variables for first-transaction characteristics, and $\lambda_{z(i)\times y(t)}$ denotes zip-by-year fixed effects based on the location and year of new agent $i$'s first transaction.[14]

Rather than interpreting individual coefficients, we use family-level joint tests on the components of $f(X_{i,j,t})$ to assess whether initial counterparty centrality is systematically related to groups of first-transaction characteristics. Limited or statistically weak associations support the view that the initial counterparty match is approximately quasi-random conditional on local market-by-cohort fixed effects.

Second, we estimate a coefficient-stability specification to examine whether the estimated effect of initial counterparty centrality on new agent's survival is sensitive to increasingly rich controls for first-transaction observables. Specifically, we estimate:

$$Survive_{i,j,t} = \alpha + \beta\, Centrality_{j,t} + g(X_{i,j,t}) + \lambda_{z(i)\times y(t)} + \varepsilon_{i,j,t}, \tag{5.2}$$

where $Survive_{i,j,t}$ is an indicator equal to one if new agent $i$ remains active after a specified duration following the initial transaction with counterparty agent $j$ at time $t$, and zero otherwise, and $Centrality_{j,t}$ denotes the initial counterparty's centrality measured at time $t$.[15] We begin with a baseline specification and then sequentially augment it with increasingly rich controls for first-transaction observables, denoted by $g(X_{i,j,t})$.

In the parsimonious specification, $g(X_{i,j,t})$ includes a coarse set of first-transaction controls entered in simple form, including bedrooms, bathrooms, property type, and in-house transaction status. In the flexible specification, these controls are replaced with a richer nonlinear set of indicator variables for living area bins, house age bins,

[14] We standardize initial counterparty centrality within the arm's-length MLS member first-transaction sample used in the balance test. Let $Centrality_{j,t}$ denote the initial counterparty centrality for new agent $i$'s first transaction at time $t$. We define

$$Centrality^{z}_{j,t} = \frac{Centrality_{j,t} - \overline{Centrality}}{\sigma(Centrality)},$$

where $\overline{Centrality}$ and $\sigma(Centrality)$ are the sample mean and sample standard deviation of $Centrality_{j,t}$ in the estimation sample. This scaling expresses variation in standard deviation units and facilitates interpretation of coefficients and joint tests.

[15] We estimate this specification using one-year, two-year, and three-year survival indicators as the dependent variable, consistent with the main outcome definitions.

bedroom and bathroom categories, lot size bins, property type, and in-house transaction status. We then assess whether the estimated coefficient on $Centrality_{j,t}$ remains stable across these specifications.

Table 9 provides evidence that the initial match to a more central counterparty is not simply capturing broad sorting into observably different first transactions. In Panel A, the flexible balance test allows first transaction characteristics to enter in a nonlinear way through bins and category indicators while conditioning on zip-by-year fixed effects. The joint tests show some residual sorting on living area, age, and in house status, but much weaker evidence for bedrooms, bathrooms, property type, and lot size. This pattern is important because it indicates that any observable imbalance is limited and concentrated in a few margins, rather than reflecting systematic differences across the full set of first transaction characteristics.

Panel B reports coefficient stability checks for one year, two year, and three year survival outcomes. In each row corresponding to the survival outcomes, the estimated coefficient on initial counterparty centrality remains highly stable as we move from the baseline specification to models with coarse first transaction controls and then to a flexible nonlinear control set. The effect is largest for one year survival, remains positive and statistically significant for two year survival, and attenuates but remains positive for three year survival. These findings are consistent with the main results in Section 4.2 and Prediction 2: counterparty network prominence has its strongest effect on first-year survival, with the effect attenuating over longer horizons. The stability of the coefficient within each survival outcome indicates that the main relationship is not driven by observable differences in the new agent's first transaction environment, which supports the quasi-random environment of the initial counterparty match.

## 5.2 Survival Analysis

We conduct a robustness check of the long-term survival results in Section 4.7 by re-estimating the relationship between a new entrant's counterparty network embeddedness and career longevity within a formal survival analysis framework (Kaplan and Meier, 1958; Cox, 1972). We measure duration as the number of days between the entrant's first and last observed transactions and treat agents with a last observed transaction within 180 days of the administrative sample end as right-censored.

In Appendix D, we provide additional results from the hazard patterns using the Kaplan-Meier survival curves.

Let $T_i$ denote the observed survival duration (in days) for new entrant $i$, measured from the first transaction to the last observed transaction. We estimate hazard models of the form

$$h_i(t) = h_0(t) \exp\left(\beta_1 \text{Centrality}_{j,t} + \beta' X_{i,j,t}\right), \tag{5.3}$$

where $h_i(t)$ is the exit hazard at duration $t$, $\text{Centrality}_{j,t}$ is the network centrality of entrant $i$'s first counterparty $j$, and $X_{i,j,t}$ includes the same relationship and transaction controls as in the baseline regressions (dual-agent status, in-house status, entrant role, counterparty non-member, property characteristics, and market-time controls). Our primary specification is a Cox proportional hazards model with zip-by-year stratification. As a parametric robustness check, we also estimate a Weibull proportional hazards model. In both hazard specifications, a negative coefficient implies a lower exit hazard and therefore longer survival.

Table 10 provides the survival analysis results that reinforce the main finding that a new entrant's initial counterparty network embeddedness is predictive of career longevity. Column 1 reports the estimates from the Cox proportional hazards specification with zip-by-year stratification. The coefficient estimate on initial counterparty centrality implies higher embeddedness is associated with a lower exit hazard of the new entrant. Interpreted in hazard-ratio terms, a one-unit increase in the centrality measure corresponds to approximately a 2.9% reduction in the instantaneous exit hazard. The Weibull proportional hazards model delivers the same sign with a slightly smaller magnitude and lower precision, which is consistent with the Cox result and suggests that the finding is not driven by the semi-parametric baseline-hazard specification. Economically, these estimates indicate that entrants who begin by transacting with more central counterparties remain active longer, consistent with early access to information, referral pathways, and repeated deal opportunities that reduce early-career attrition.

The control variables on arm's length type and agent relationship also show economically meaningful patterns that are consistent with our results in Section 4.2. New entrants whose first transaction is a dual-agent deal face substantially higher exit risk corresponding to roughly a 9–11% higher hazard of exit; this is consistent

with dual-agency first deals providing less opportunity to form durable inter-agent relationships. Similarly, the coefficient estimates on in-house counterparty is positive in both hazard specifications implying a higher exit hazard conditional on the full set of controls, even though the raw in-house relationship may capture short-run brokerage cultural and resource support. The estimates for counterparty non-member status are especially large implying roughly a 24–31% higher exit hazard, which is economically substantial and consistent with the idea that first transactions involving non-member counterparties generate weaker network spillovers and fewer opportunities to build embeddedness in the MLS-based professional network. Hence, the survival analysis is consistent with both our new-agent survival (Table 3) and long-term survival results (Table 7), reinforcing the robustness of our main findings.

# 6 Conclusion

We find that initial network connections play a critical role in shaping the early career trajectories of new real estate agents. By leveraging the quasi-random matching of new agents to counterparty agents, we establish causality and show that network ties to well-connected agents significantly increase new agents' probability of survival, transaction activity, and subsequent network embeddedness. These empirical results corroborate the theoretical model's central prediction: initial network capital compounds over time, generating persistent differences in productivity and career longevity.

These findings provide important insights into how careers unfold in information-sensitive markets. They reveal that seemingly minor, random early differences, such as transacting with a well-connected counterparty, can greatly impact one's career trajectory. By formalizing this mechanism through a Bayesian framework, we provide a micro-founded account of how network capital is internalized, reinforcing the path-dependent nature of career success. In doing so, our study bridges theories of human capital, mentorship, and network economics within a structural, data-rich setting.

More broadly, the findings highlight the importance of network connections in professions where early career success depends on information flow, mentorship, reputation, and repeated interactions. Importantly, the residential real estate brokerage setting we examine not only provides clean identification, but the underlying mechanisms we examine (i.e., access to deals, connectivity within a professional network,

and the advantages of early embeddedness) and corresponding results are likely to generalize to other highly intermediated, relationship-driven labor markets. Recognizing that initial network ties help shape professional outcomes provides actionable insights for firms and policymakers seeking to design more effective onboarding systems or targeted interventions to reduce early attrition.

# References


Athey, Susan, Christopher Avery, and Peter Zemsky, 2000, Mentoring and Diversity, *American Economic Review* 90, 765–786.

Bala, Venkatesh, and Sanjeev Goyal, 2000, A Noncooperative Model of Network Formation, *Econometrica* 68, 1181–1229.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa, 2008, Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes, *Journal of Political Economy* 116, 1150–1196.

Beaman, Lori, and Jeremy Magruder, 2012, Who Gets the Job Referral? Evidence from a Social Networks Experiment, *American Economic Review* 102, 3574–3593.

Blau, Francine D, Janet M Currie, Rachel T A Croson, and Donna K Ginther, 2010, Can Mentoring Help Female Assistant Professors? Interim Results from a Randomized Trial, *American Economic Review Paper and Proceedings* 100, 348–352.

Calvó-Armengol, Antoni, and Matthew O. Jackson, 2004, The Effects of Social Networks on Employment and Inequality, *American Economic Review* 94, 426–454.

Cattan, Sarah, Kjell G. Salvanes, and Emma Tominey, 2025, First-Generation Elite: The Role of School Social Networks, *American Economic Review* 115, 4369–4403.

Choi, Soon Hyeok, Adam Nowak, Patrick S. Smith, and Alexei Tchistyi, 2025, The Winner's Curse in Housing Markets, *SSRN Electronic Journal* .

Cox, D. R., 1972, Regression Models and Life-Tables, *Journal of the Royal Statistical Society. Series B (Methodological)* 34, 187–220.

Cunha, Flavio, and James Heckman, 2007, The Technology of Skill Formation, *American Economic Review Papers and Proceedings* 97, 31–47.

Di Maggio, Marco, Francesco Franzoni, Amir Kermani, and Carlo Sommavilla, 2019, The Relevance of Broker Networks for Information Diffusion in the Stock Market, *Journal of Financial Economics* 134, 419–446.

Falk, Armin, Fabian Kosse, and Pia Pinger, 2026, Mentoring and schooling decisions: Causal evidence, *Journal of Political Economy* 134, 366–396.

Goldsmith-Pinkham, Paul, and Kelly Shue, 2023, The Gender Gap in Housing Returns, *The Journal of Finance* 78, 1097–1145.

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.

Han, Munhee, Sanghyun (Hugh) Kim, and Vikram K Nanda, 2024, Institutional Brokerage Networks: Facilitating Liquidity Provision, *The Review of Financial Studies* 37, 2903–2935.

Hilmer, Christiana, and Michael Hilmer, 2007, Women Helping Women, Men Helping Women? Same-gender Mentoring, Initial Job Placements, and Early Career Publishing Success for Economics PhDs, *American Economic Review Paper and Proceedings* 97, 422–426.

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, 1076–1122.

Jackson, Matthew O., 2008, *Social and Economic Networks*, volume 3 (Princeton University Press).

Kaplan, E. L., and Paul Meier, 1958, Nonparametric Estimation from Incomplete Observations, *Journal of the American Statistical Association* 53, 457–481.

Kramarz, Francis, and Oskar Nordström Skans, 2014, When Strong Ties are Strong: Networks and Youth Labour Market Entry, *The Review of Economic Studies* 81, 1164–1200.

Manski, Charles F., 1993, Identification of Endogenous Social Effects: The Reflection Problem, *The Review of Economic Studies* 60, 531–542.

Montgomery, James D., 1991, Social Networks and Labor-Market Outcomes: Toward an Economic Analysis, *American Economic Review* 81, 1408–1418.

Relitix, 2024, The Alarming Failure Rate of Recent New Real Estate Agents: A Closer Look.

Resnjanskij, Sven, Jens Ruhose, Simon Wiederhold, Ludger Woessmann, and Katharina Wedel, 2024, Can Mentoring Alleviate Family Disadvantage in Adolescence? A Field Experiment to Improve Labor Market Prospects, *Journal of Political Economy* 132, 1013–1062.

Rossi, Alberto G., David Blake, Allan Timmermann, Ian Tonks, and Russ Wermers, 2018, Network Centrality and Delegated Investment Performance, *Journal of Financial Economics* 128, 183–206.

# Tables

Table 1. Summary Statistics

| | Mean | St. Dev. | Pctl(25) | Median | Pctl(75) |
|---|---|---|---|---|---|
| **Panel A: MLS Transaction Data** | | | | | |
| List Price (000s) | 289.06 | 226.83 | 149.90 | 228.80 | 350.00 |
| Sale Price (000s) | 278.10 | 215.97 | 145.00 | 220.00 | 340.00 |
| Days on Market | 73.04 | 113.99 | 13 | 37 | 91 |
| Single-family Detached | 0.86 | 0.34 | 1 | 1 | 1 |
| Condominium | 0.12 | 0.32 | 0 | 0 | 0 |
| Townhome | 0.02 | 0.13 | 0 | 0 | 0 |
| Age (Years) | 37.39 | 30.43 | 13 | 31 | 55 |
| Bedrooms | 3.21 | 1.62 | 3 | 3 | 4 |
| Bathrooms | 2.41 | 1.07 | 2 | 2 | 3 |
| Observations | 6,055,235 | | | | |
| Unique Agents | 572,253 | | | | |
| **Panel B: New Agent Data** | | | | | |
| Counterparty Centrality In | 0.14 | 0.23 | 0.0001 | 0.02 | 0.16 |
| Counterparty Centrality Out | 0.01 | 0.06 | 0.00 | 0.0000 | 0.001 |
| Counterparty Centrality Combined | 0.07 | 0.12 | 0.0002 | 0.01 | 0.09 |
| Dual Agent | 0.07 | 0.25 | 0 | 0 | 0 |
| In-house Agent | 0.07 | 0.26 | 0 | 0 | 0 |
| Listing Agent | 0.24 | 0.43 | 0 | 0 | 0 |
| Counterparty Non-member | 0.03 | 0.16 | 0 | 0 | 0 |
| Buyer Agent | 0.59 | 0.49 | 0 | 1 | 1 |
| Single-family Detached | 0.85 | 0.36 | 1 | 1 | 1 |
| Condominium | 0.13 | 0.34 | 0 | 0 | 0 |
| Townhome | 0.02 | 0.13 | 0 | 0 | 0 |
| Survive One Year | 0.67 | 0.47 | 0 | 1 | 1 |
| Survive Two Years | 0.58 | 0.49 | 0 | 1 | 1 |
| Survive Three Years | 0.48 | 0.50 | 0 | 0 | 1 |
| Survival Length (Days) | 1,685.78 | 1,859.65 | 122 | 1,023 | 2,602 |
| First Year Transactions | 3.43 | 3.23 | 1 | 2 | 4 |
| First Three Years Transactions | 8.03 | 10.60 | 2 | 4 | 10 |
| Total New Agent Transactions | 23.06 | 60.82 | 2 | 6 | 20 |
| New Agent Count | 181,766 | | | | |

**Notes:** Table 1 displays summary stats for the data used in the empirical analysis. Panel A displays summary stats for the MLS transaction data used to construct the real estate agent trading networks. Panel B displays summary stats for the new agent data.

Table 2. Variable Definitions

| | |
|---|---|
| **Panel A: Counterparty Centrality** | |
| In-Centrality | Eigenvector centrality based on incoming edge (buyer agent) |
| Out-Centrality | Eigenvector centrality based on outgoing edge (listing agent) |
| Combined Centrality | The average of in-centrality and out-centrality |
| **Panel B: Agent & Transaction Type** | |
| Dual Agent | 1 if new agent represents both sides in the first transaction |
| In-house Agent | 1 if new agent and counterparty are in the same brokerage office |
| Listing Agent | 1 if new agent is a listing agent, counterparty is an MLS member |
| Buyer Agent | 1 if new agent is a buyer agent, counterparty is an MLS member |
| Counterparty Non-member | 1 if new agent is a listing agent, counterparty is not an MLS member |
| **Panel C: Property Type** | |
| Single-family Detached | 1 if transaction property is single-family detached |
| Condominium | 1 if transaction property is condominium |
| Townhome | 1 if transaction property is townhome |
| **Panel D: Survival & Volume** | |
| Survive 1/2/3 Years | 1 if the agent completes at least one additional transaction more than 365/730/1095 days after the first transaction |
| Survival Length (Days) | Days between first and last observed transaction of the agent |
| First Year Transactions | Number of transactions within 365 days after first transaction. |
| First Three Years Transactions | Number of transactions within 1095 days after first transaction |
| Total New Agent Transactions | Total number of transactions over the agent's observed sample window |

**Notes:** Table 2 reports the variable definitions used in the empirical analysis. Panel A defines the three variants of eigenvector centrality computed for a new agent's first counterparty following the measurement procedure in Section 3.3. Panel B defines the agent and transaction type variables. Panel C defines the property types. Panel D provides variables related to agent survival length and transaction volume. All binary variables are indicators that equal 1 if the stated condition holds and 0 otherwise.

Table 3. New Agent Survival Rate

| | Survive One Year | | | | | |
|---|---|---|---|---|---|---|
| | In | | Out | | Combined | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Baseline estimation | | | | | | |
| Counterparty Centrality | 0.029*** | 0.027*** | 0.066*** | 0.055*** | 0.061*** | 0.056*** |
| | (0.005) | (0.005) | (0.019) | (0.021) | (0.010) | (0.010) |
| Observations | 181,766 | 181,766 | 181,766 | 181,766 | 181,766 | 181,766 |
| Adjusted $R^2$ | 0.023 | 0.030 | 0.023 | 0.030 | 0.023 | 0.030 |
| Panel B: Baseline estimation with additional controls | | | | | | |
| Counterparty Centrality | 0.023*** | 0.020*** | 0.048** | 0.034 | 0.048*** | 0.041*** |
| | (0.005) | (0.005) | (0.019) | (0.021) | (0.010) | (0.010) |
| Dual Agent | −0.014*** | −0.016*** | −0.017*** | −0.018*** | −0.014*** | −0.016*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| In-house Agent | 0.009** | 0.006 | 0.009** | 0.006 | 0.009** | 0.007 |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Listing Agent | −0.024*** | −0.026*** | −0.024*** | −0.026*** | −0.024*** | −0.026*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Counterparty Non-member | −0.099*** | −0.104*** | −0.101*** | −0.106*** | −0.098*** | −0.104*** |
| | (0.008) | (0.009) | (0.008) | (0.009) | (0.008) | (0.009) |
| Condominium | −0.006 | −0.004 | −0.006 | −0.004 | −0.006 | −0.004 |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Townhome | −0.010 | −0.007 | −0.010 | −0.007 | −0.010 | −0.007 |
| | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) |
| Observations | 181,766 | 181,766 | 181,766 | 181,766 | 181,766 | 181,766 |
| Adjusted $R^2$ | 0.025 | 0.032 | 0.025 | 0.032 | 0.025 | 0.032 |
| Zip+Year FE | ✓ | | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ | | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** Table 3 presents regression estimates from Equation 4.1 examining whether the connectedness of the first counterparty agent that new agents transact with affects their probability of surviving one year. Panel A reports estimates excluding the agency relationship and unit type controls. Panel B reports the corresponding estimates when these additional controls are included. The dependent variable in each column is a one-year survival indicator variable. Columns 1 and 2 report results for the Counterparty Centrality In measure, columns 3 and 4 for the Counterparty Centrality Out measure, and columns 5-6 for the combined counterparty network centrality measure. Odd-numbered specifications include additively separable zip code and sale year fixed effects, whereas even-numbered specifications include multiplicatively separable zip code and sale year fixed effects. Standard errors are clustered at the zip code level.

Table 4. Arm's Length and Brokerage Office Fixed Effects

| | Survive One Year | | | | | |
|---|---|---|---|---|---|---|
| | Members Only | | Arm's Length | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Counterparty Centrality Combined | 0.045*** | 0.038*** | 0.042*** | 0.037*** | 0.026** | 0.032*** |
| | (0.010) | (0.010) | (0.010) | (0.011) | (0.011) | (0.012) |
| Dual Agent | −0.014*** | −0.015*** | | | | |
| | (0.005) | (0.005) | | | | |
| In-house Agent | 0.010** | 0.007 | | | | |
| | (0.005) | (0.005) | | | | |
| Listing Agent | −0.024*** | −0.026*** | −0.024*** | −0.026*** | −0.019*** | −0.019*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Condominium | −0.006 | −0.004 | −0.004 | −0.002 | −0.002 | 0.001 |
| | (0.004) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Townhome | −0.010 | −0.008 | −0.007 | −0.006 | −0.010 | −0.009 |
| | (0.009) | (0.009) | (0.010) | (0.011) | (0.011) | (0.011) |
| Observations | 176,783 | 176,783 | 151,412 | 151,412 | 151,412 | 151,412 |
| Adjusted $R^2$ | 0.022 | 0.030 | 0.023 | 0.032 | 0.080 | 0.088 |
| Zip+Year FE | ✓ | | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ | | ✓ |
| Office FE | | | | | ✓ | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** Table 4 presents regression estimates from Equation 4.1 examining whether the connectedness of the first counterparty agent that new agents transact with affects their probability of surviving one year. The dependent variable in each column is a one-year survival indicator variable. Columns 1 and 2 report estimates for a subset of new agents whose first transaction was with a counterparty agent that was an MLS member. Columns 3 through 6 report the corresponding estimates for a subset of new agents whose first transaction was with an arm's-length counterparty that was an MLS member. Odd-numbered columns include additively separable zip code and sale year fixed effects. Even-numbered columns include multiplicatively separable fixed effects for zip code and sale year. Columns 5 and 6 also include brokerage office fixed effects. Standard errors are clustered at the zip code level.

Table 5. First Year Transaction Cohorts

| | Survive One Year | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Counterparty Centrality Combined | 0.022** | 0.022** | 0.066*** | 0.071*** |
| | (0.009) | (0.009) | (0.014) | (0.015) |
| Dual Agent | −0.009** | −0.010** | −0.012*** | −0.013*** |
| | (0.004) | (0.005) | (0.005) | (0.005) |
| In-house Agent | −0.014*** | −0.015*** | −0.008* | −0.010** |
| | (0.004) | (0.005) | (0.004) | (0.005) |
| Listing Agent | −0.021*** | −0.022*** | −0.022*** | −0.023*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Counterparty Non-member | −0.057*** | −0.060*** | −0.071*** | −0.075*** |
| | (0.007) | (0.008) | (0.007) | (0.008) |
| Condominium | −0.009** | −0.007* | −0.008* | −0.007 |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| Townhome | −0.009 | −0.008 | −0.009 | −0.008 |
| | (0.008) | (0.008) | (0.008) | (0.009) |
| 3-5 First-Year Sales | 0.311*** | 0.310*** | | |
| | (0.003) | (0.003) | | |
| 6-10 First-Year Sales | 0.431*** | 0.428*** | | |
| | (0.003) | (0.003) | | |
| 11-20 First-Year Sales | 0.469*** | 0.471*** | | |
| | (0.003) | (0.004) | | |
| First Year Transactions | | | 0.052*** | 0.052*** |
| | | | (0.0004) | (0.0005) |
| Centrality × Transactions Interaction | | | −0.012*** | −0.013*** |
| | | | (0.002) | (0.002) |
| Observations | 181,766 | 181,766 | 181,766 | 181,766 |
| Adjusted $R^2$ | 0.173 | 0.179 | 0.141 | 0.147 |
| Zip+Year FE | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** Table 5 presents regression estimates from a modified version of Equation 4.1 examining whether the connectedness of the first counterparty agent that new agents transact with affects their probability of surviving their first year. The dependent variable in each column is a one-year survival indicator variable. Columns 1 and 2 incorporate a set of first-year transaction-cohort indicators. Columns 3 and 4 replace these indicators with a continuous measure of first-year transaction volume and its interaction with counterparty centrality. Odd-numbered columns include additively separable zip code and sale year fixed effects. Even-numbered columns include multiplicatively separable fixed effects for zip code and sale year. Standard errors are clustered at the zip code level.

Table 6. New Agent Transaction Volume

| | Log(Number of Transactions) | | | | |
|---|---|---|---|---|---|
| | Year 1 | Year 1 | Years 1-3 | Total | Per Year |
| | (1) | (2) | (3) | (4) | (5) |
| Counterparty Centrality Combined | 0.092*** | 0.088*** | 0.117*** | 0.125*** | 0.053** |
| | (0.019) | (0.020) | (0.027) | (0.036) | (0.022) |
| Dual Agent | −0.015* | −0.009 | −0.055*** | −0.109*** | −0.076*** |
| | (0.009) | (0.010) | (0.013) | (0.018) | (0.010) |
| In-house Agent | 0.098*** | 0.097*** | 0.121*** | 0.086*** | 0.083*** |
| | (0.009) | (0.009) | (0.013) | (0.018) | (0.010) |
| Listing Agent | −0.010** | −0.0002 | −0.043*** | −0.098*** | −0.049*** |
| | (0.005) | (0.005) | (0.008) | (0.010) | (0.006) |
| Counterparty Non-member | −0.199*** | −0.198*** | −0.360*** | −0.515*** | −0.276*** |
| | (0.016) | (0.016) | (0.023) | (0.030) | (0.018) |
| Condominium | 0.018** | 0.018** | 0.011 | −0.004 | 0.004 |
| | (0.008) | (0.008) | (0.011) | (0.015) | (0.009) |
| Townhome | 0.002 | −0.004 | −0.011 | −0.022 | 0.003 |
| | (0.017) | (0.018) | (0.026) | (0.033) | (0.020) |
| Observations | 181,766 | 163,494 | 163,494 | 163,494 | 163,494 |
| Adjusted $R^2$ | 0.072 | 0.076 | 0.087 | 0.056 | 0.069 |
| Zip×Year FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Data Filter | | 2004-19 | 2004-19 | 2004-19 | 2004-19 |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** Table 6 presents regression estimates from a modified version of Equation 4.1 examining whether the connectedness of the first counterparty agent that new agents transact with affects their transaction volume. The dependent variable in every column is the log of the new agent's transaction count over the specified time horizon. Columns 1 and 2 examine the number of transactions during the new agent's first year, column 3 examines the new agent's first three years, column 4 examines the new agent's total number of transactions, and column 5 examines the new agent's transactions per year. Column 1 uses the full new agent data sample from 2004 through June 2021, while columns 2 through 5 use new agents that enter the profession between 2004 and 2019. Each column includes multiplicatively separable zip code and sale year fixed effects. Standard errors are clustered at the zip code level.

Table 7. New Agent Longer-Term Survival

| | Survival Length | | Two-Year Survival | | | |
|---|---|---|---|---|---|---|
| | | | Unconditional on Surviving First Year | | Conditional on Surviving First Year | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Counterparty Centrality Combined | 0.374*** | 0.360*** | 0.054*** | 0.053*** | 0.006 | 0.009 |
| | (0.066) | (0.069) | (0.010) | (0.011) | (0.008) | (0.009) |
| Observations | 163,494 | 163,494 | 163,494 | 163,494 | 111,163 | 111,163 |
| Adjusted $R^2$ | 0.027 | 0.041 | 0.025 | 0.034 | 0.013 | 0.012 |
| Zip+Year FE | ✓ | | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ | | ✓ |
| Data Filter | 2004-19 | 2004-19 | 2004-19 | 2004-19 | 2004-19 | 2004-19 |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** Table 7 presents regression estimates from Equation 4.1 examining whether the connectedness of the first counterparty agent that new agents transact with affects their longer-term survival. The dependent variable in columns 1 and 2 is a continuous variable measuring the new agent's survival length as the number of days between their first and last transactions in the data. The dependent variable in columns 3-6 is a two-year survival indicator. Columns 3 and 4 examine whether the new agent survives two years, regardless of whether they survived the first year. Columns 5 and 6 differ in that they examine whether the new agent survives two years using a subsample of new agents that survived their first year. Odd-numbered specifications include additively separable zip code and sale year fixed effects, whereas even-numbered specifications include multiplicatively separable zip code and sale year fixed effects. Standard errors are clustered at the zip code level.

Table 8. New Agent End of Year Network Centrality

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: New Agent EOY1 Centrality Combined | | | | |
| Counterparty Centrality Combined | 0.057*** | 0.045*** | 0.056*** | 0.042*** |
| | (0.003) | (0.003) | (0.004) | (0.004) |
| Observations | 181,766 | 181,766 | 111,163 | 111,163 |
| Adjusted $R^2$ | 0.024 | 0.042 | 0.028 | 0.050 |
| Panel B: Survival Conditional on Surviving Previous Year | | | | |
| | Two Years | | Three Years | |
| New Agent EOY1 Centrality Combined | 0.096*** | 0.096*** | 0.103*** | 0.107*** |
| | (0.009) | (0.010) | (0.008) | (0.009) |
| Observations | 111,163 | 111,163 | 97,259 | 97,259 |
| Adjusted $R^2$ | 0.014 | 0.013 | 0.021 | 0.012 |
| Zip+Year FE | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** Panel A of Table 8 presents regression estimates from a modified version of Equation 4.1 that examines whether the connectedness of the first counterparty agent that new agents transact with affects their network position at the end of their first year. The dependent variable in each column of Panel A is the new agents' end-of-year one (EOY1) network centrality combined measure. Columns 1 and 2 of Panel A use the full new agent sample, while columns 3 and 4 use a subsample of new agents who entered the profession from 2004-2019 and survived their first year. Panel B examines whether the new agents' EOY1 connectedness affects their longer-term survival probability. Columns 1 and 2 of Panel B examine whether the new agent survives their second year using a subsample of new agents who entered the profession from 2004-2019 and survived their first year. Columns 3 and 4 of Panel B examine whether the new agent survives their third year using a subsample of new agents who entered the profession from 2004-2019 and survived their second year. Odd-numbered specifications include additively separable zip code and sale year fixed effects, whereas even-numbered specifications include multiplicatively separable zip code and sale year fixed effects. Standard errors are clustered at the zip code level.

Table 9. Balance Tests

| Panel A: Joint Test (Initial Counterparty Centrality & First Transaction Observables) | |
|---|---|
| Living Area | 0.0029 |
| Age | 0.0006 |
| Bedroom | 0.2252 |
| Bathroom | 0.3567 |
| Lot Size | 0.0529 |
| Property Type | 0.1132 |
| In-House Status | 0.0000 |
| Observations | 158,500 |
| Adjusted $R^2$ | 0.112 |

| Panel B: Robustness to First Transaction Controls ($z$-score) | | | |
|---|---|---|---|
| | (1) Base | (2) Coarse | (3) Flexible |
| One-year Survival | 0.004092*** | 0.004172*** | 0.004164*** |
| | (0.001282) | (0.001283) | (0.001287) |
| Two-year Survival | 0.003988*** | 0.004037*** | 0.004035*** |
| | (0.001296) | (0.001296) | (0.001300) |
| Three-year Survival | 0.002426* | 0.002436* | 0.002406* |
| | (0.001306) | (0.001305) | (0.001307) |
| Observations | 158,500 | 158,500 | 158,500 |
| Adjusted $R^2$ (1-year) | 0.029 | 0.030 | 0.030 |
| Adjusted $R^2$ (2-year) | 0.046 | 0.046 | 0.046 |
| Adjusted $R^2$ (3-year) | 0.105 | 0.105 | 0.105 |
| Zip-by-year FE | ✓ | ✓ | ✓ |
| Arm's Length MLS Member | ✓ | ✓ | ✓ |
| Coarse Controls | | ✓ | |
| Flexible Controls | | | ✓ |

*p<0.10; **p<0.05; ***p<0.01

**Notes:** Table 9 reports balance diagnostics for the new entrant's first transaction sample and a robustness check for the main outcome regression. The sample includes one observation per new agent's first recorded transaction, restricted to arm's length MLS member transactions with an identified counterparty, excluding dual agency and non member counterparties. All specifications include zip-by-year fixed effects and zip clustered standard errors. Panel A reports family level joint test $p$ values from a flexible nonlinear balance specification where the dependent variable is the $z$ score of initial counterparty centrality. Regressors include indicators for living area bins (100 square feet), age bins (10 years), bedroom and bathroom categories, lot size bins, property type, and in house status. Panel B reports the coefficient on initial counterparty centrality from the main survival regressions (one, two, and three year survival) under progressively richer controls for first transaction observables. Column (1) is the baseline, Column (2) adds coarse controls (bedrooms, bathrooms, property type, in house status), and Column (3) adds the full flexible nonlinear controls used in Panel A.

Table 10. Survival Analysis

| | (1) | (2) |
|---|---|---|
| Initial Counterparty Centrality | -0.029* | -0.019 |
| | (0.017) | (0.015) |
| Dual Agent | 0.105*** | 0.088*** |
| | (0.016) | (0.013) |
| In-house Agent | 0.053*** | 0.038*** |
| | (0.015) | (0.013) |
| Listing Agent | 0.105*** | 0.096*** |
| | (0.009) | (0.007) |
| Counterparty Non-member | 0.271*** | 0.241*** |
| | (0.027) | (0.024) |
| Property Type | ✓ | ✓ |
| Property Characteristics | ✓ | ✓ |
| Location/Time FE | Zip-Year Strata | Year + MSA FE |
| Right-censor Buffer | 180 days | 180 days |
| Observations | 144,949 | 144,949 |

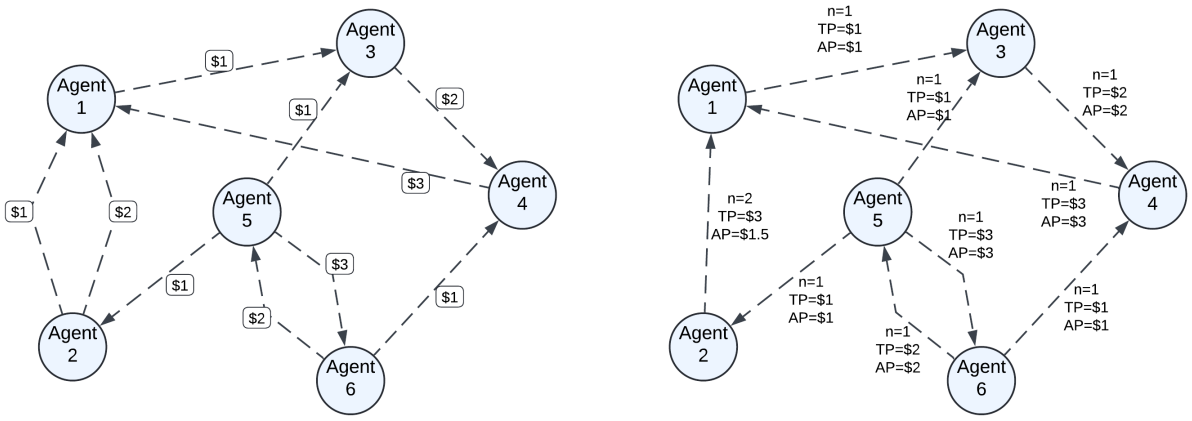
Standard errors in parentheses.

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

**Notes:** The dependent variable is survival duration (days) from a new entrant's first to last observed transaction. Column (1) reports a Cox proportional hazards model with zip-by-year stratification. Column (2) reports a Weibull proportional hazards model with year and MSA fixed effects. All specifications include the same relationship and transaction controls as in the baseline regressions. Standard errors are clustered at the zip-code level. The estimates are reported in log hazard scale.

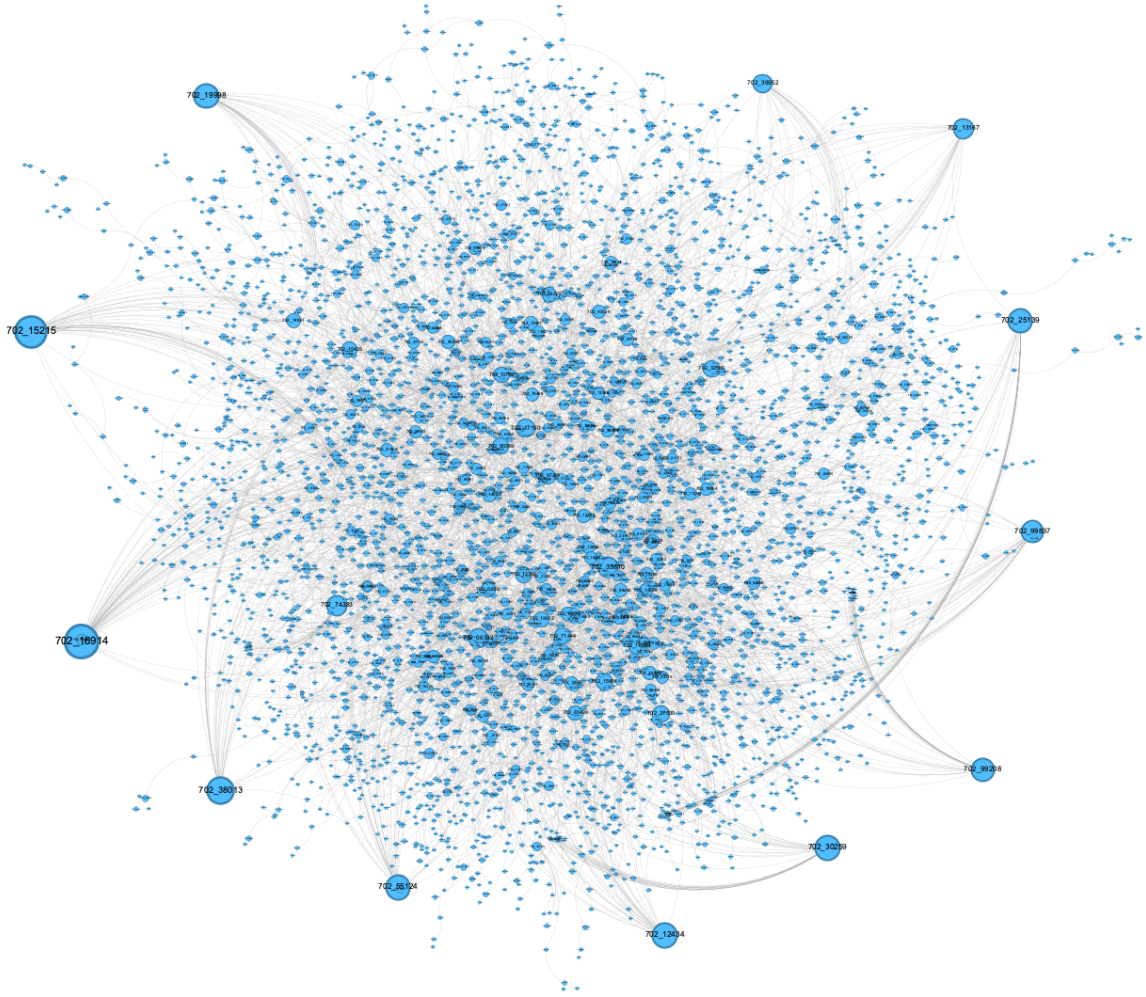
# Figures

Figure 1. Raw Network Construction & Reduced-Form Graph



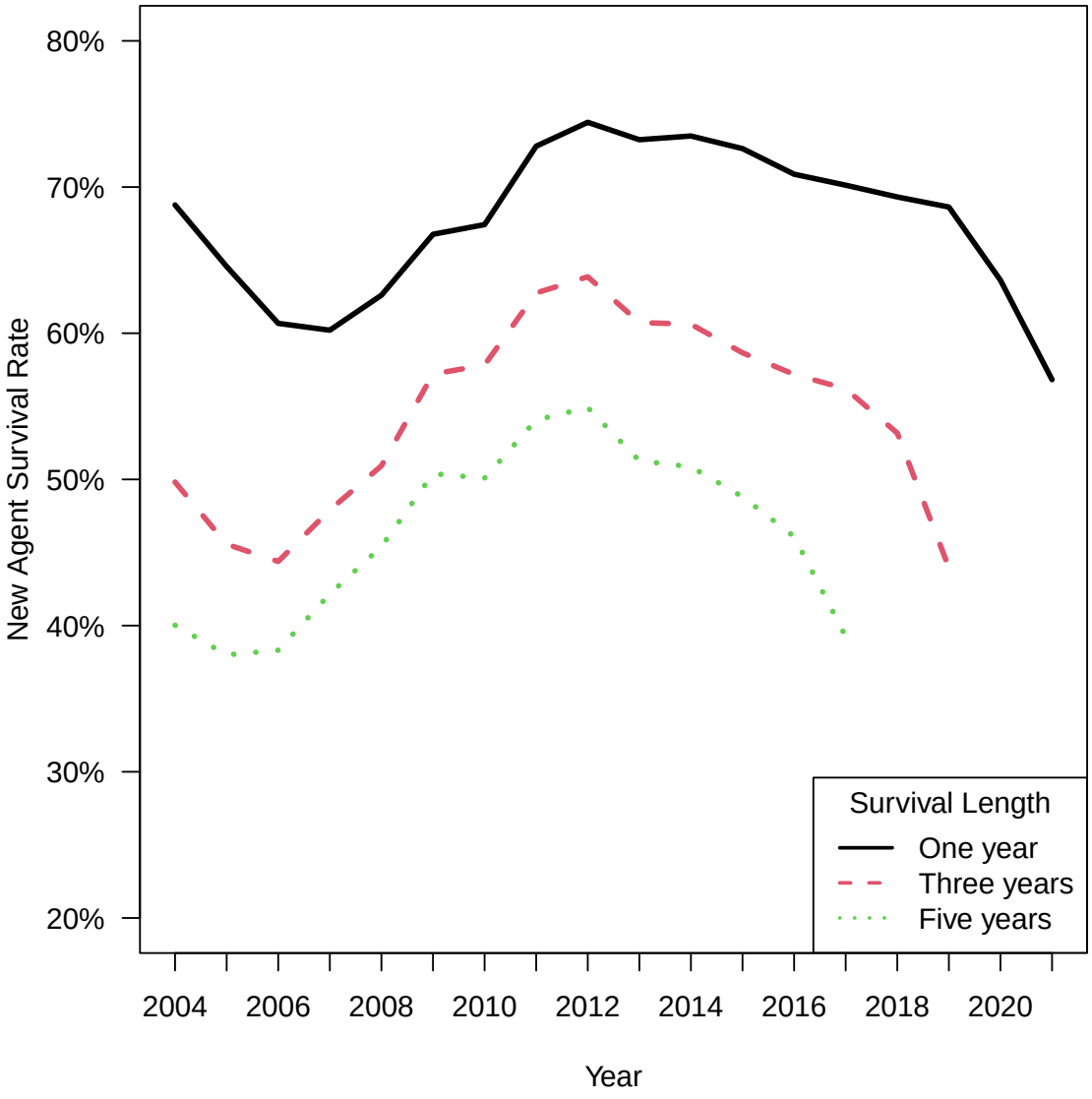

*Note*: Figure 1 plots an illustration of agent transaction networks in the real estate market. The left panel depicts a directed multigraph, where nodes represent agents and each directed edge represents a single transaction from the listing agent to the buyer agent. Edges are labeled with transaction sale prices. The right panel shows the corresponding reduced-form directed graph, where multiple transactions between the same pair of agents are aggregated into a single edge. Each edge is annotated with the number of transactions (n), total transaction price (TP), and average price (AP), summarizing the transactional relationship between agents over the observation window. This collapsed representation supports feature extraction while preserving the directionality and intensity of agent interactions.

Figure 2. Reduced-Form Directed Graph of Charlotte for Q2 2006.



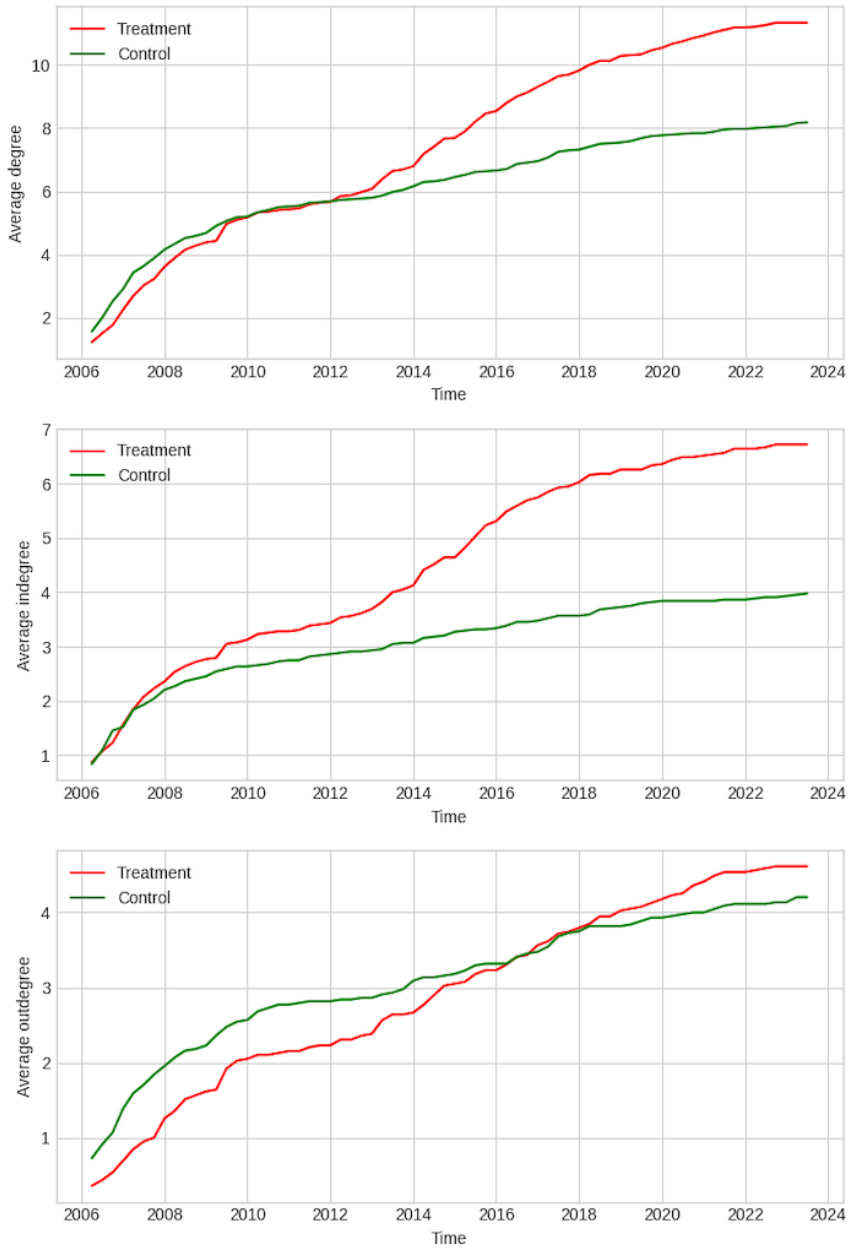

*Note*: Figure 2 displays the reduced-form directed graph of the real estate agent network for Charlotte, NC, constructed using transactions from Q2 2006. Each node represents an individual real estate agent, with node size scaled proportionally to the agent's degree (i.e., transaction volume). Directed edges represent aggregated transaction relationships from listing agents to buyer agents, where multiple transactions between the same agent pair are collapsed into a single weighted edge.

Figure 3. Fraction of New Agents that Survive Each Year



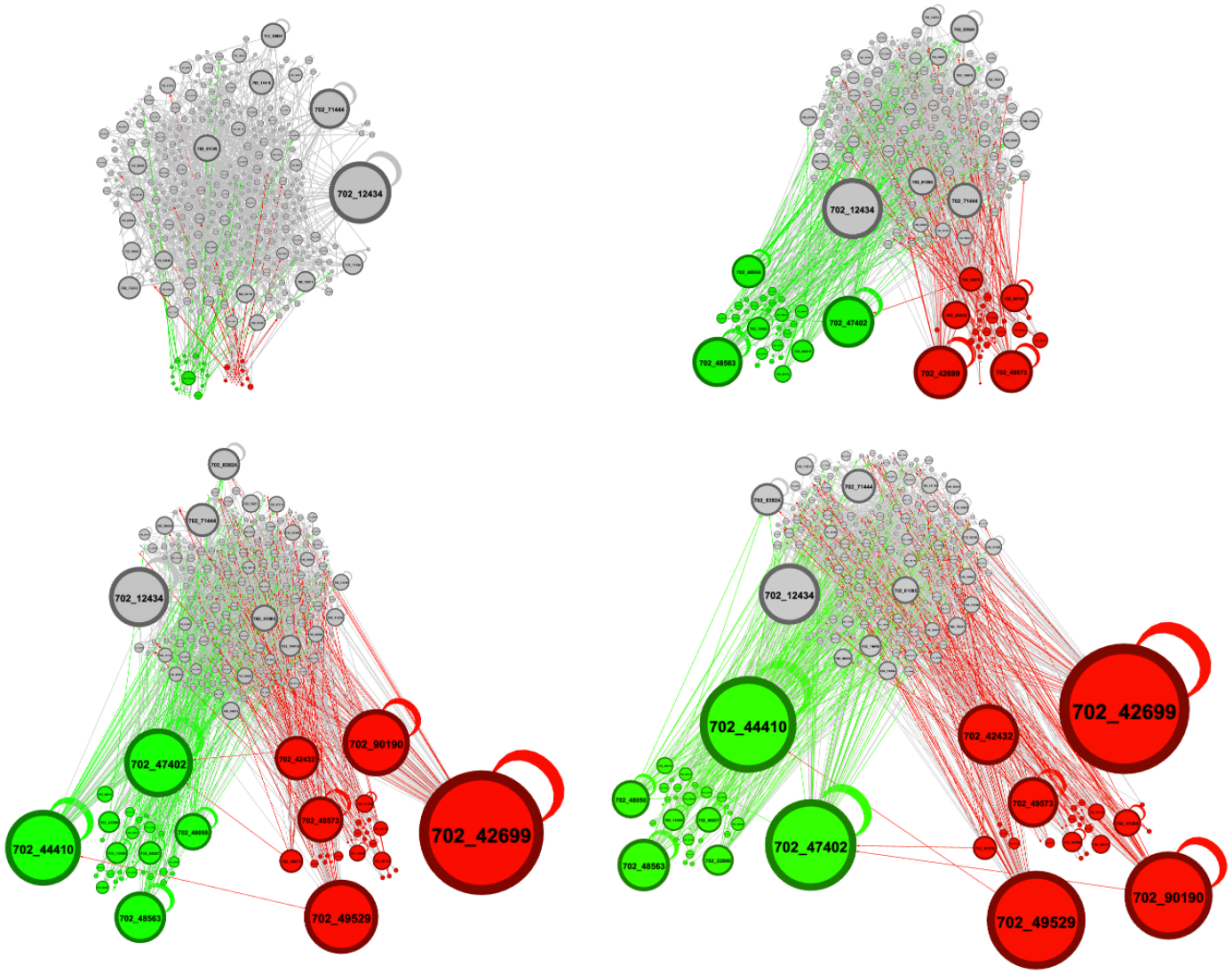

*Note*: Figure 3 plots the percent of new agents that survive for more than one, three, and five years based on the year the new agent first transacts in the data.

Figure 4. Treatment vs. Control Group (Avg Deg, In-deg, Out-deg)



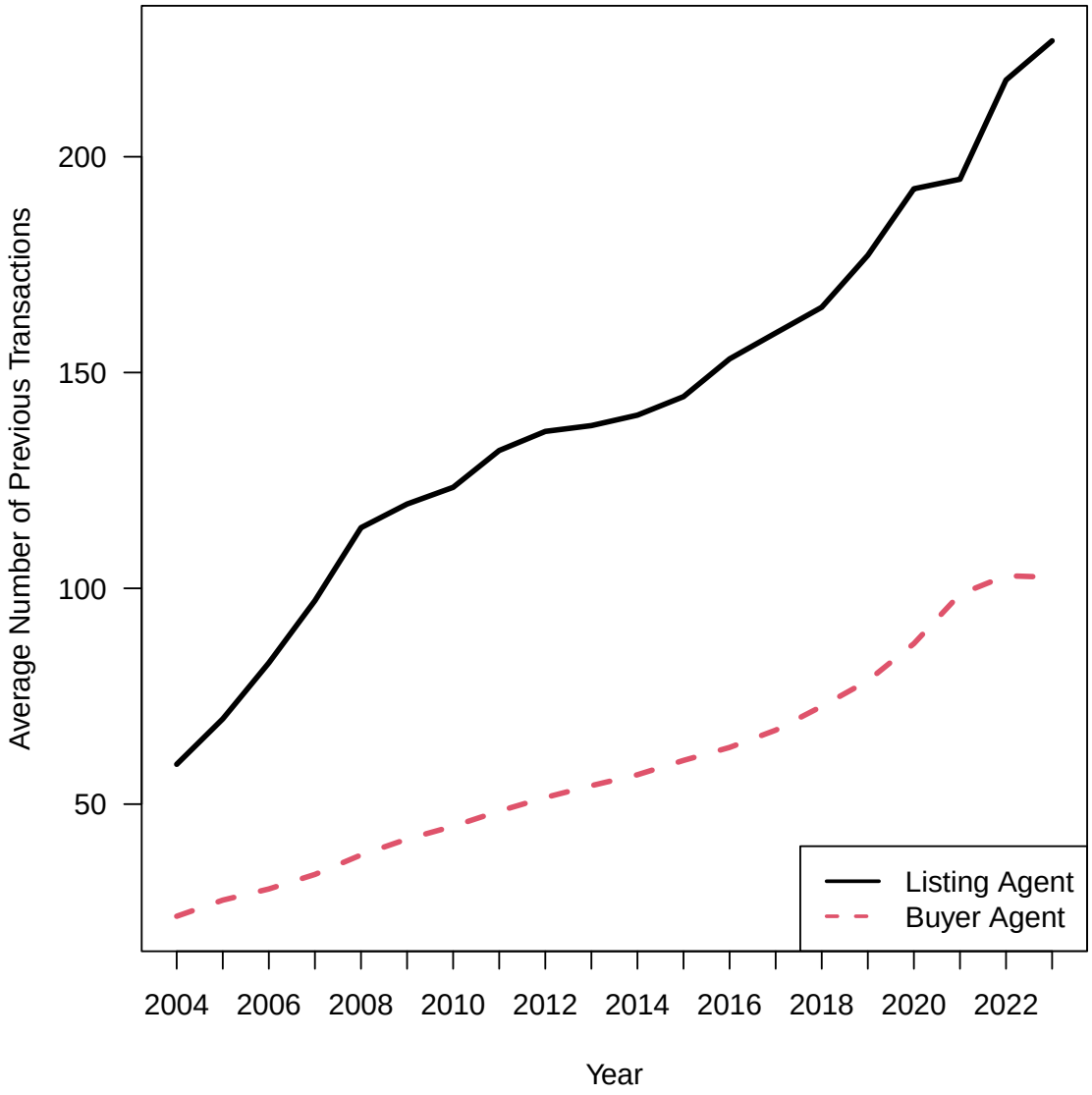

*Note*: Figure 4 plots the evolution of average degree, in-degree, and out-degree for treatment and control group agents. We define a baseline network using transactions from Union County in Q2 2006 and identify agents whose first deal occurs in this window. Agents are assigned to the treatment group if their first counterparty ranks in the top 25% of network centrality; others form the control group. We track degree, in-degree, and out-degree over time by sequentially adding transactions and retaining only those involving treatment or control agents, allowing us to analyze the evolution of network positions post-entry.

Figure 5. Visualization of The Treatment & Control Groups from Figure 4.




*Note*: Figure 5 displays the baseline agent network and its snapshots at successive time points, illustrating the evolution of network structure for treatment and control groups. This visualization corresponds to the setup described in Figure 4, highlighting how agents' positions change over time relative to their initial classifications. The panels display snapshots of the agent network at four time points: Q2 2006 (top left), 2012 (top right), 2018 (bottom left), and 2023 (bottom right). Red nodes represent agents in the treatment group, green nodes represent agents in the control group, and gray nodes represent all other agents in the network. Node size is scaled proportionally to the agent's degree in the corresponding time-specific network, reflecting each agent's level of connectivity at that point in time.

Figure 6. Previous Transactions by Agent Type




*Note*: Figure 6 plots the average number of previous transactions completed by listing and buyer agents over time.

Online Appendix to

# The Importance of Networks for Career Success: Evidence from the Real Estate Industry

Soon Hyeok Choi & Patrick S. Smith & Ali Tosyali

## A Proofs

### A.1 Proof of Proposition 1

*Proof.*

$$\pi_i = \Pr\left(v_j^\theta \in \theta^H\right) \mathbb{E}[\varphi_i c_j \mid v_j^\theta \in \theta^H] + \Pr\left(v_j^\theta \in \theta^L\right) \mathbb{E}[\varphi_i c_j \mid v_j^\theta \in \theta^L] \tag{A.1}$$

$$= (1-c^*)\mathbb{E}[\varphi_i c_j \mid c_j \in [c^*,1]] + c^*\mathbb{E}[\varphi_i c_j \mid c_j \in [0,c^*]] \tag{A.2}$$

$$= (1-c^*)\varphi_i\mathbb{E}[c_j \mid c_j \in [c^*,1]] + c^*\varphi_i\mathbb{E}[c_j \mid c_j \in [0,c^*]] \tag{A.3}$$

$$= (1-c^*)\varphi_i\frac{c^*+1}{2} + c^*\varphi_i\frac{c^*}{2} \tag{A.4}$$

$$= \varphi_i\left[(1-c^*)\frac{c^*+1}{2} + c^*\frac{c^*}{2}\right] \tag{A.5}$$

$$= \varphi_i\left[\frac{(1-c^*)(c^*+1)+(c^*)^2}{2}\right] \tag{A.6}$$

$$= \frac{\varphi_i}{2} \tag{A.7}$$

Since $\pi_i(\varphi_i) = \frac{\varphi_i}{2}$ is strictly increasing in $\varphi_i$ on $[0,1]$, the unique best response is $\varphi_i^* = 1$, yielding $\mathbb{E}[\pi_i^*] = \frac{1}{2}$. □

### A.2 Proof of Proposition 2

*Proof.* Recall that the survival probability is given by:

$$S_i(t) = \left(1 - e^{-\nu\, c_{j,0}(V_i^{(1)}+1)}\right)\rho^t. \tag{A.8}$$

$$\frac{\partial S_i(t)}{\partial c_{j,0}} = \rho^t \frac{\partial}{\partial c_{j,0}} \left(1 - e^{-\nu\, c_{j,0}(V_i^{(1)}+1)}\right) \tag{A.9}$$

$$= \rho^t \left[\nu(V_i^{(1)}+1)\, e^{-\nu\, c_{j,0}(V_i^{(1)}+1)}\right]. \tag{A.10}$$

Since $\nu > 0$, $V_i^{(1)} \geq 0$, $\rho \in (0,1)$, it follows that:

$$\frac{\partial S_i(t)}{\partial c_{j,0}} > 0 \quad \forall\, t \in \mathbb{N}. \tag{A.11}$$

Differentiating the expected lifetime payoff with respect to $c_{j,0}$:

$$\frac{\partial \mathbb{E}[\tilde{W}_i]}{\partial c_{j,0}} = \varphi_i \sum_{t=1}^{\infty} \delta^t \underbrace{\frac{\partial S_i(t)}{\partial c_{j,0}}}_{>0} \underbrace{\mathbb{E}[c_{j,t}]}_{\geq 0}. \tag{A.12}$$

$$\frac{\partial \mathbb{E}[\tilde{W}_i]}{\partial c_{j,0}} > 0. \tag{A.13}$$

□

## A.3 Proof of Proposition 3

*Proof.* Recall

$$\mathbb{E}[\tilde{W}_i] = \varphi_i \sum_{t=1}^{\infty} \delta^t\, \mathbb{E}[c_{j,t}]\, \underbrace{\rho^t \left(1 - e^{-\nu c_{j,0}(V_i^{(1)}+1)}\right)}_{S_i(t)}. \tag{A.14}$$

$$\frac{\partial S_i(t)}{\partial c_{j,0}} = \rho^t \frac{\partial}{\partial c_{j,0}} \left(1 - e^{-\nu c_{j,0}(V_i^{(1)}+1)}\right) \tag{A.15}$$

$$= \rho^t \left[\nu(V_i^{(1)}+1)\, e^{-\nu c_{j,0}(V_i^{(1)}+1)}\right]. \tag{A.16}$$

$$\frac{\partial^2 S_i(t)}{\partial c_{j,0}^2} = \rho^t\, \nu(V_i^{(1)}+1) \left[-\nu(V_i^{(1)}+1)\, e^{-\nu c_{j,0}(V_i^{(1)}+1)}\right] \tag{A.17}$$

$$= -\rho^t \left[\nu(V_i^{(1)} + 1)\right]^2 e^{-\nu c_{j,0}(V_i^{(1)}+1)}. \tag{A.18}$$

Since $\rho, \nu, (V_i^{(1)} + 1) > 0$, it follows that $\frac{\partial^2 S_i(t)}{\partial c_{j,0}^2} < 0$. Finally, applying this to the total expected payoff:

$$\frac{\partial^2 \mathbb{E}[\tilde{W}_i]}{\partial c_{j,0}^2} = \varphi_i \sum_{t=1}^{\infty} \delta^t \, \mathbb{E}[c_{j,t}] \underbrace{\frac{\partial^2 S_i(t)}{\partial c_{j,0}^2}}_{<0} < 0. \tag{A.19}$$

Since the weights $\varphi_i \delta^t \mathbb{E}[c_{j,t}]$ are non-negative and the second derivative of the survival function is strictly negative, the total expected payoff is strictly concave in $c_{j,0}$. □

# B Data Overview

## B.1 Data Availability Statement

We purchased the Multiple Listing Service (MLS) data used in this study from Cotality. Our agreement with Cotality prohibits the posting or redistribution of the data. Researchers interested in replication can obtain the MLS data directly from Cotality. Additional information is available on the Cotality website.

## B.2 Data Filters

Prior to running the empirical analysis, we apply several filters remove outliers and records with missing information in key fields. Below is a list of the filters we apply to the transaction data before creating the agent networks.

1. filter 1: unit type = single family detached | condo | townhome
2. filter 2: listing successfully sold
3. filter 3: 2001 $\leq$ sale year $\leq$ 2023
4. filter 4: \$20,000 $\leq$ sale price $\leq$ \$2,000,000
5. filter 5: \$20,000 $\leq$ list price $\leq$ \$2,000,000
6. filter 6: age $\geq$ 1
7. filter 7: remove duplicate listings
8. filter 8: number of sales in zip code $\geq$ 50
9. filter 9: remove if listing agent or buyer agent ID missing

# C Robustness Checks

## C.1 Additional Results

Table C.1. New Agent Survival Rate - Trim Top 3%

| | Survive One Year | | | | | |
|---|---|---|---|---|---|---|
| | In | | Out | | Combined | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Baseline estimation | | | | | | |
| Counterparty Centrality | 0.031*** | 0.030*** | 0.042** | 0.027 | 0.061*** | 0.058*** |
| | (0.005) | (0.006) | (0.019) | (0.021) | (0.010) | (0.011) |
| Observations | 193,024 | 193,024 | 193,024 | 193,024 | 193,024 | 193,024 |
| Adjusted $R^2$ | 0.023 | 0.029 | 0.023 | 0.029 | 0.023 | 0.029 |
| Panel B: Baseline estimation with additional controls | | | | | | |
| Counterparty Centrality | 0.025*** | 0.023*** | 0.023 | 0.006 | 0.049*** | 0.043*** |
| | (0.005) | (0.006) | (0.019) | (0.021) | (0.010) | (0.011) |
| Dual Agent | −0.015*** | −0.017*** | −0.018*** | −0.020*** | −0.015*** | −0.017*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| In-house Agent | 0.012*** | 0.010** | 0.012*** | 0.010** | 0.012*** | 0.010** |
| | (0.004) | (0.005) | (0.004) | (0.005) | (0.004) | (0.005) |
| Listing Agent | −0.025*** | −0.026*** | −0.024*** | −0.026*** | −0.024*** | −0.026*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Counterparty Non-member | −0.101*** | −0.106*** | −0.103*** | −0.108*** | −0.101*** | −0.106*** |
| | (0.008) | (0.009) | (0.008) | (0.009) | (0.008) | (0.009) |
| Condominium | −0.007 | −0.005 | −0.007 | −0.006 | −0.007 | −0.005 |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Townhome | −0.007 | −0.004 | −0.007 | −0.004 | −0.007 | −0.004 |
| | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) |
| Observations | 193,024 | 193,024 | 193,024 | 193,024 | 193,024 | 193,024 |
| Adjusted $R^2$ | 0.025 | 0.031 | 0.025 | 0.031 | 0.025 | 0.031 |
| Zip+Year FE | ✓ | | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ | | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** The results in this table correspond to Table 3 except that we trim the top 3% of counterparty agent IDs instead of the top 5%.

Table C.2. New Agent Survival Rate - Trim Top 1%

| | Survive One Year | | | | | |
|---|---|---|---|---|---|---|
| | In | | Out | | Combined | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Baseline estimation | | | | | | |
| Counterparty Centrality | 0.028*** | 0.022*** | 0.044** | 0.028 | 0.055*** | 0.041*** |
| | (0.007) | (0.008) | (0.017) | (0.019) | (0.012) | (0.013) |
| Observations | 206,408 | 206,408 | 206,408 | 206,408 | 206,408 | 206,408 |
| Adjusted $R^2$ | 0.022 | 0.028 | 0.022 | 0.028 | 0.022 | 0.028 |
| Panel B: Baseline estimation with additional controls | | | | | | |
| Counterparty Centrality | 0.024*** | 0.016** | 0.026 | 0.010 | 0.044*** | 0.028** |
| | (0.007) | (0.008) | (0.017) | (0.019) | (0.012) | (0.013) |
| Dual Agent | −0.016*** | −0.018*** | −0.017*** | −0.019*** | −0.016*** | −0.018*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| In-house Agent | 0.015*** | 0.014*** | 0.015*** | 0.014*** | 0.015*** | 0.014*** |
| | (0.004) | (0.005) | (0.004) | (0.005) | (0.004) | (0.005) |
| Listing Agent | −0.024*** | −0.025*** | −0.024*** | −0.025*** | −0.024*** | −0.025*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Counterparty Non-member | −0.105*** | −0.108*** | −0.105*** | −0.109*** | −0.105*** | −0.108*** |
| | (0.008) | (0.009) | (0.008) | (0.009) | (0.008) | (0.009) |
| Condominium | −0.008** | −0.007* | −0.009** | −0.007* | −0.008** | −0.007* |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Townhome | −0.009 | −0.007 | −0.009 | −0.007 | −0.009 | −0.007 |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| Observations | 206,408 | 206,408 | 206,408 | 206,408 | 206,408 | 206,408 |
| Adjusted $R^2$ | 0.024 | 0.030 | 0.024 | 0.030 | 0.024 | 0.030 |
| Zip+Year FE | ✓ | | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ | | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** The results in this table correspond to Table 3 except that we trim the top 1% of counterparty agent IDs instead of the top 5%.

Table C.3. New Agent Survival Rate - Include Non-traditional Agent IDs

| | Survive One Year | | | | | |
|---|---|---|---|---|---|---|
| | In | | Out | | Combined | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: Baseline estimation | | | | | | |
| Counterparty Centrality | 0.025*** | 0.019*** | 0.025*** | 0.026*** | 0.036*** | 0.033*** |
| | (0.005) | (0.005) | (0.006) | (0.006) | (0.007) | (0.007) |
| Non-traditional Counterparty | 0.014*** | 0.010*** | 0.015*** | 0.015*** | 0.015*** | 0.015*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.002) | (0.003) |
| Observations | 217,837 | 217,837 | 217,837 | 217,837 | 217,837 | 217,837 |
| Adjusted $R^2$ | 0.022 | 0.028 | 0.022 | 0.028 | 0.022 | 0.029 |
| Panel B: Baseline estimation with additional controls | | | | | | |
| Counterparty Centrality | 0.019*** | 0.012** | 0.012** | 0.012** | 0.024*** | 0.020*** |
| | (0.005) | (0.005) | (0.006) | (0.006) | (0.007) | (0.007) |
| Non-traditional Counterparty | 0.010*** | 0.006* | 0.006** | 0.005* | 0.009*** | 0.008*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.002) | (0.003) |
| Dual Agent | −0.013*** | −0.015*** | −0.014*** | −0.014*** | −0.011** | −0.012** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| In-house Agent | 0.014*** | 0.013*** | 0.014*** | 0.013*** | 0.014*** | 0.013*** |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Listing Agent | −0.023*** | −0.025*** | −0.022*** | −0.024*** | −0.022*** | −0.024*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Counterparty Non-member | −0.101*** | −0.105*** | −0.102*** | −0.104*** | −0.099*** | −0.102*** |
| | (0.008) | (0.009) | (0.008) | (0.009) | (0.008) | (0.009) |
| Condominium | −0.008** | −0.007* | −0.008** | −0.007* | −0.008** | −0.007* |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Townhome | −0.008 | −0.006 | −0.008 | −0.006 | −0.008 | −0.006 |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| Observations | 217,837 | 217,837 | 217,837 | 217,837 | 217,837 | 217,837 |
| Adjusted $R^2$ | 0.023 | 0.030 | 0.023 | 0.030 | 0.023 | 0.030 |
| Zip+Year FE | ✓ | | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ | | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** The results in this table correspond to Table 3 except that instead of trimming the top 5% of counterparty agent IDs to remove non-traditional agent IDs, we set their counterparty centrality equal to 0 and flag them with a Non-traditional Counterparty indicator variable.

Table C.4. Arm's Length and Brokerage Office Fixed Effects - In Measure

| | Survive One Year | | | | | |
|---|---|---|---|---|---|---|
| | Members Only | | Arm's Length | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Counterparty Centrality In | 0.022*** | 0.018*** | 0.020*** | 0.018*** | 0.012** | 0.014** |
| | (0.005) | (0.005) | (0.005) | (0.006) | (0.006) | (0.006) |
| Dual Agent | −0.014*** | −0.016*** | | | | |
| | (0.005) | (0.005) | | | | |
| In-house Agent | 0.010** | 0.007 | | | | |
| | (0.005) | (0.005) | | | | |
| Listing Agent | −0.024*** | −0.026*** | −0.024*** | −0.026*** | −0.019*** | −0.019*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Condominium | −0.006 | −0.004 | −0.004 | −0.002 | −0.002 | 0.001 |
| | (0.004) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Townhome | −0.010 | −0.008 | −0.007 | −0.006 | −0.010 | −0.009 |
| | (0.009) | (0.009) | (0.010) | (0.011) | (0.011) | (0.010) |
| Observations | 176,783 | 176,783 | 151,412 | 151,412 | 151,412 | 151,412 |
| Adjusted $R^2$ | 0.022 | 0.030 | 0.023 | 0.032 | 0.080 | 0.088 |
| Zip+Year FE | ✓ | | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ | | ✓ |
| Office FE | | | | | ✓ | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** The results in this table correspond to Table 4 except for the use of the Counterparty Centrality In measure instead of the Counterparty Centrality Combined measure.

Table C.5. Arm's Length and Brokerage Office Fixed Effects - Out Measure

| | Survive One Year | | | | | |
|---|---|---|---|---|---|---|
| | Members Only | | Arm's Length | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Counterparty Centrality Out | 0.045** | 0.032 | 0.037* | 0.032 | 0.039* | 0.045** |
| | (0.019) | (0.021) | (0.020) | (0.022) | (0.021) | (0.022) |
| Dual Agent | −0.017*** | −0.018*** | | | | |
| | (0.005) | (0.005) | | | | |
| In-house Agent | 0.010** | 0.007 | | | | |
| | (0.005) | (0.005) | | | | |
| Listing Agent | −0.024*** | −0.026*** | −0.024*** | −0.026*** | −0.019*** | −0.019*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Condominium | −0.006 | −0.005 | −0.004 | −0.002 | −0.002 | 0.001 |
| | (0.004) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Townhome | −0.010 | −0.008 | −0.007 | −0.006 | −0.010 | −0.009 |
| | (0.009) | (0.009) | (0.010) | (0.011) | (0.011) | (0.010) |
| Observations | 176,783 | 176,783 | 151,412 | 151,412 | 151,412 | 151,412 |
| Adjusted $R^2$ | 0.022 | 0.030 | 0.023 | 0.032 | 0.080 | 0.088 |
| Zip+Year FE | ✓ | | ✓ | | ✓ | |
| Zip×Year FE | | ✓ | | ✓ | | ✓ |
| Office FE | | | | | ✓ | ✓ |

*p<0.1; **p<0.05; ***p<0.01

**Notes:** The results in this table correspond to Table 4 except for the use of the Counterparty Centrality Out measure instead of the Counterparty Centrality Combined measure.

## C.2 Cross-Market ID & New Agent Classification

Our analyses inherently assumes that an agent's first transaction in a local MLS trading network represents their entry into the profession. A natural concern is that we may classify an agent with prior industry experience in a previous market as a new agent in their new market if they relocate across regions. Empirically, this is not a major concern because the new agent, with prior experience in a different market, will still be randomly assigned to a counterparty agent with unknown network centrality in the new market. Moreover, the new agent with prior experience will still need to build network connections to succeed in the new market, and the counterparty agent they first connect with will affect their probability of success.

Although we do not view this as a major concern, this section demonstrates that there is very little overlap among real estate agents across the ten MLS platforms in our study. Because the data are assembled from ten non-overlapping local MLS platforms, and we find no universally consistent identifier across them, we conduct a series of high-level diagnostic tests to assess the extent to which agents appear in more than one market. These diagnostics include comparing agent names across MLS platforms and examining the frequency with which identical agent identifiers (cell phone numbers) appear across them. Across all approaches, we find little evidence of cross-platform duplication. The incidence of potential duplication is extremely small, and the few cases that do arise are consistent with common names rather than true cross-market changes. Taken together, these tests reinforce that our classification of new agents is not meaningfully contaminated by agents with prior experience elsewhere, and that our empirical setting effectively captures new entrants into the real estate brokerage profession. The following subsections document the diagnostics we perform, their methodology, and key findings.

### C.2.1 Agent IDs as National Realtors Database System

In Cotality MLS data, the agent identifier variable is defined as

> "Specific unique identifier provided by the Multiple Listing Service (MLS) to the buyer's agent. May be the same as the National Realtors Database System (NRDS) ID."

This naturally leads us to first examine whether the agent identifier in the MLS data behaves like an NRDS ID. We stack listing- and buyer-side agent records into an

agent-level file and flag IDs observed across multiple markets. Using the full set of agent records across the ten MLS platforms, we find no evidence of cross-market overlap in IDs: every agent ID appears in exactly one market (both in row-weighted counts and at the unique-ID level). This implies that either the MLS agent identifier is market-specific, or, if some are NRDS IDs, this subset is unique (i.e., there is no overlap).

### C.2.2 Agent Names

Since IDs do not link across markets, we next assess name-based overlap. We standardize agent names by applying common cleaning operations (e.g., trimming whitespace, converting to uppercase, and removing punctuation/hyphens) to create a clean name variable. Using the agent-level file, we compute for each cleaned name the number of distinct markets in which it appears, and define an indicator variable to examine whether the cleaned name appears in two or more markets. We report both (i) row-weighted shares (treating each agent-record as an observation) and (ii) unique-name shares (treating each distinct as one observation).

Table C.6. Distribution of Cleaned Agent Names Across Markets

| Markets per Name | Unique Names | Percent (%) | Cumulative (%) |
|---|---|---|---|
| 1 Market | 273,650 | 97.15 | 97.15 |
| 2 Markets | 6,595 | 2.34 | 99.49 |
| 3 Markets | 1,051 | 0.37 | 99.87 |
| 4 Markets | 285 | 0.10 | 99.97 |
| 5+ Markets | 94 | 0.03 | 100 |
| Total | 281,675 | 100 | |

The results in Table C.6 show that cross-market name ambiguity is uncommon in our sample: 97.15% of distinct cleaned names appear in only one market, and only 2.85% (8,025 out of 281,675) appear in two or more markets. Moreover, the small amount of overlap is concentrated in exactly two-market matches (2.34% of names), with three- or more-market overlaps extremely rare. Overall, this distribution suggests that potential misclassification of first appearance due to cross-market name duplications is unlikely to be a first-order driver of our results.

### C.2.3 Agent Cell Phone Numbers

As an additional test, we examine the frequency with which agent cell phone numbers overlap across MLS platforms. In doing so, we assume that an agent's cell phone number is a stable personal identifier that does not change if they relocate or operate in more than one market, even when other identifiers - such as last names or agent IDs (see above) - may change. Because cell phone numbers are portable and rarely shared across unrelated individuals, they provide an additional way to detect whether the same person appears in multiple MLS systems. That said, it is important to note that the cell phone number field is not required (i.e., it can be blank) and is unpopulated for every agent in two of ten (Boston and Houston) MLS datasets.

Similar to the agent name analysis above, we standardize agent cell phone numbers by applying common cleaning operations (e.g., trimming whitespace and dropping non-numeric characters) to create a clean cell phone number variable. We also require the cell number to have ten digits, which has a minimal effect on the observation count. We then tabulate how many of the eight MLS platforms the cell phone number appears in.

Table C.7. Distribution of Agent Cell Phone Numbers Across Markets

| Markets Per Cell # | Unique Cell # | Percent (%) | Cumulative (%) |
|---|---|---|---|
| 1 Market | 140,139 | 99.667 | 99.667 |
| 2 Markets | 450 | 0.320 | 99.987 |
| 3 Markets | 9 | 0.006 | 99.994 |
| 4 Markets | 4 | 0.003 | 99.996 |
| 5+ Markets | 5 | 0.004 | 100 |
| Total | 140,607 | 100 | |

If agents were routinely operating in multiple cities or relocating across cities, we would expect to observe nontrivial overlap in cell phone numbers across platforms. Instead, the results in Table C.7 show that the incidence of shared cell phone numbers is very rare (less than 0.5%), reinforcing our assertion in the paper that agents overwhelmingly operate within a single local market and that cross-platform agent mobility is minimal.

# D Kaplan–Meier Survival Curves & Hazard Patterns

This appendix presents nonparametric survival evidence using Kaplan–Meier (KM) curves. These figures complement Section 5.2 by providing a model-free description of how new entrant survival evolves over time across economically relevant groups. In particular, the KM plots are useful for visualizing (i) the timing of attrition (e.g., whether separation occurs early or gradually), (ii) the magnitude of raw survival differences across groups, and (iii) whether the ordering of survival profiles is stable over the duration distribution.[1]

## D.1 Specification

Let $T_i$ denote the observed survival duration (in days) for new entrant $i$, measured as the number of days between the entrant's first and last observed transactions. Let $C_i$ denote the censoring time implied by the administrative end of the sample (June 2023). The observed duration is $Y_i = \min\{T_i, C_i\}$, and the failure indicator is $\Delta_i = \mathbb{1}\{T_i \leq C_i\}$. As in Section 5.2, new entrants whose last observed transaction occurs within 180 days of June 2023 are treated as right-censored.

For the nonparametric analysis, we estimate the Kaplan–Meier survivor function,

$$\hat{S}(t) = \prod_{\tau \leq t} \left(1 - \frac{d(\tau)}{n(\tau)}\right) \tag{D.1}$$

where $d(\tau)$ is the number of exits at duration $\tau$ and $n(\tau)$ is the number of entrants still at risk immediately before $\tau$. Intuitively, the survival function represents the probability that a new entrant remains active more than $t$ days after entry, conditional the new agent belongs to a group (e.g., those with top 25% counterparty centrality, dual agent).

[1] Since the nonparametric estimator requires no functional-form assumptions on the baseline hazard, the KM figures are descriptive. They do not condition on first-transaction characteristics, relationship controls, or market-time controls, and are included to summarize raw survival patterns and motivate the multivariate hazard analysis in Section 5.2.

## D.2 Results

### D.2.1 Initial Counterparty Centrality

Figure D.1 plots KM survivor functions by whether the new entrant's first counterparty is in the top 25% of the centrality distribution (top quartile) or the bottom 75%. In the KM figures by centrality, entrants are grouped using a pooled (full-sample) cutoff for initial counterparty centrality, rather than market-time-cell-specific cutoffs. The two panels present the same comparison over short (0–60 day) and longer (0–500 day) horizons, which helps distinguish immediate attrition differences from gradual separation over time. The key descriptive pattern is that entrants matched to more centrally connected counterparties exhibit higher survival probabilities throughout most of the observed duration range. The level differences are modest but persistent in direction, which is consistent with the negative hazard-model coefficient estimates on initial counterparty centrality reported in Section 5.2.

### D.2.2 Dual-Agent v. Non-Dual First Transactions

Figure D.2 compares KM survivor functions for entrants whose first transaction is a dual-agent deal versus a non-dual-agent deal. The short-horizon panel highlights a significantly lower survival probability implying early-career attrition, while the longer-horizon panel shows whether the survival paths continue to diverge over time. The descriptive pattern is consistent with the regression evidence in Table 3. New entrants whose first transaction is a dual-agent deal tend to exhibit lower survival probabilities, suggesting that dual-agency entry may provide fewer opportunities to establish durable inter-agent ties at the start of a career. The hazard regressions in Section 5.2 confirm this relationship after conditioning on first-transaction characteristics and market-time controls.

### D.2.3 In-House v. Different-Brokerage First Counterparties

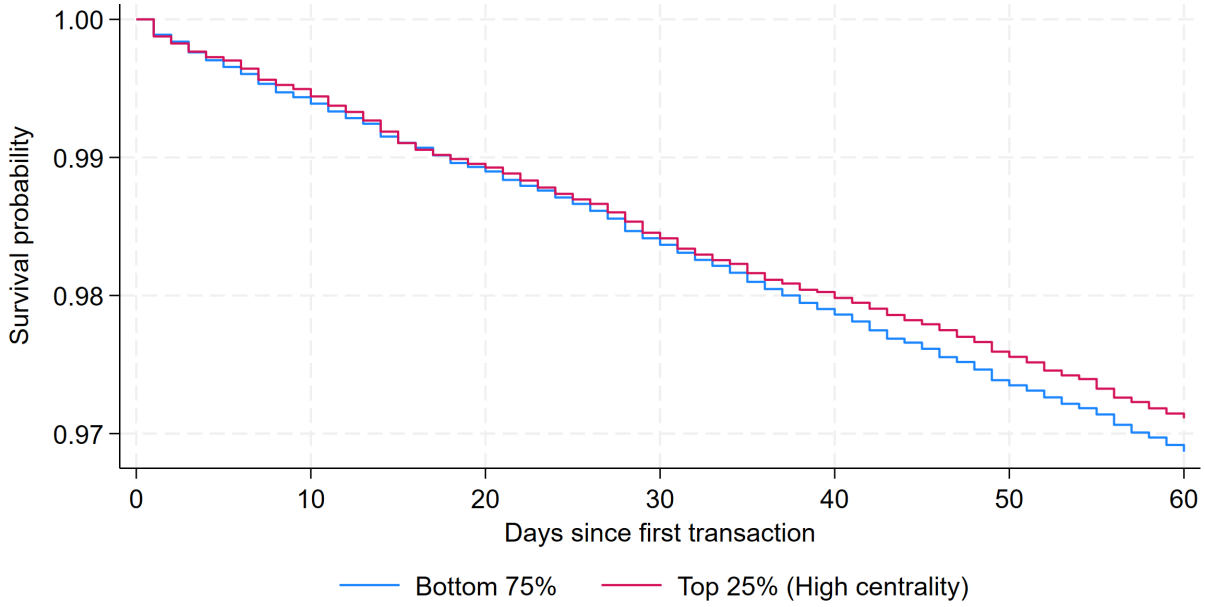
Figure D.3 reports KM survivor functions for entrants whose first counterparty is in the same brokerage office (in-house) versus a different office (different-brokerage). This comparison is restricted to non-dual first transactions so that a distinct counterparty is observed. In the 60-day horizon since the new entrant's first transaction, we do

not see strong evidence of diverging survival probability paths for two groups. This pattern is consistent with our finding in Section 4.2 that the new agent survival is not driven primarily by shared brokerage affiliation. In Panel (B), we document an interesting visual pattern. The new agent who transacts first with an in-house counterparty exhibits a meaningful lower survival probability than the one who transacts first with a counterparty at a different brokerage office. Intuitively, the first in-house counterparty transaction might limit that very initial external exposure a new entrant needs to reinforce its career longevity.

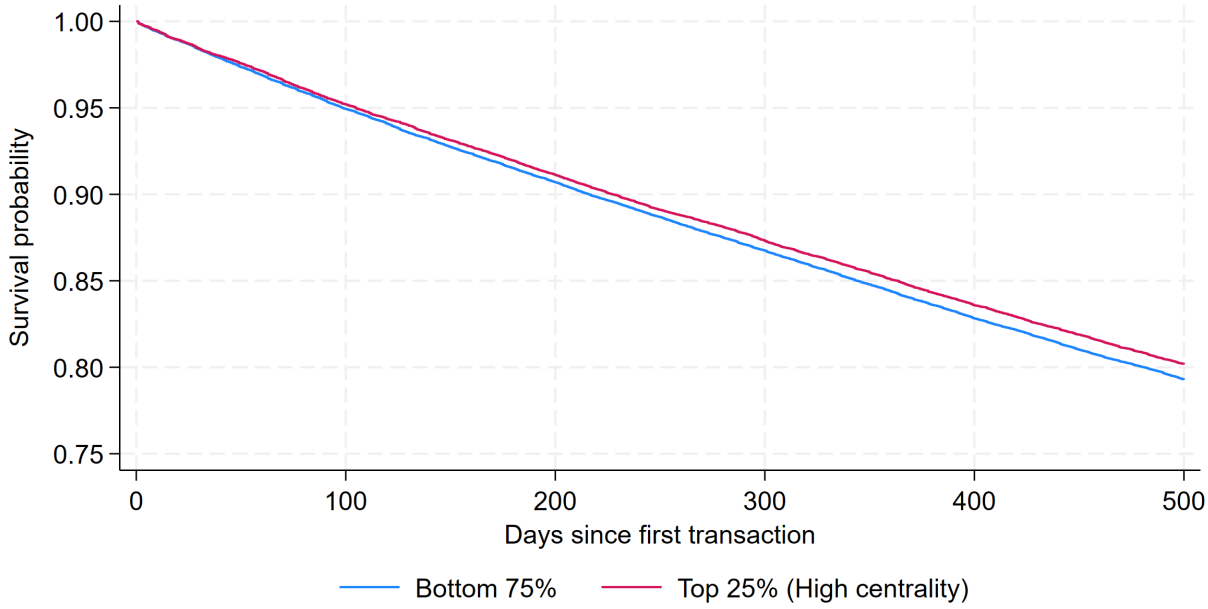
#### D.2.4 Entrant Role in the First Transaction

Figure D.4 compares KM survivor functions by entrant role in the first transaction, distinguishing entrants who first appear as buyer agents from those who first appear as listing agents. Consistent with Figure 6 and our main regression analysis, we see a significantly lower survival probability for the new entrants who transact on their first deal as a listing agent. In Panel (B), we see this divergence to be persistent. This divergence confirms the intuition that buyer-side and listing-side entry may reflect different on-boarding pathways, opportunity sets, and relationship-formation margins.

Figure D.1. Kaplan–Meier Survival by Initial Counterparty Centrality



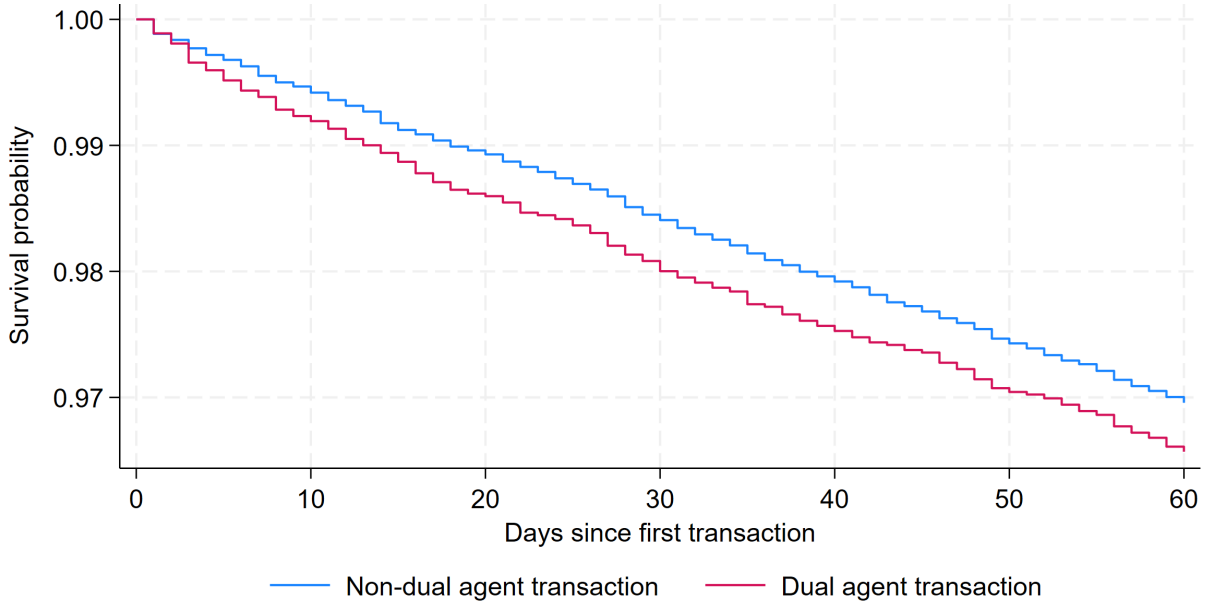

(A) 0–60 days



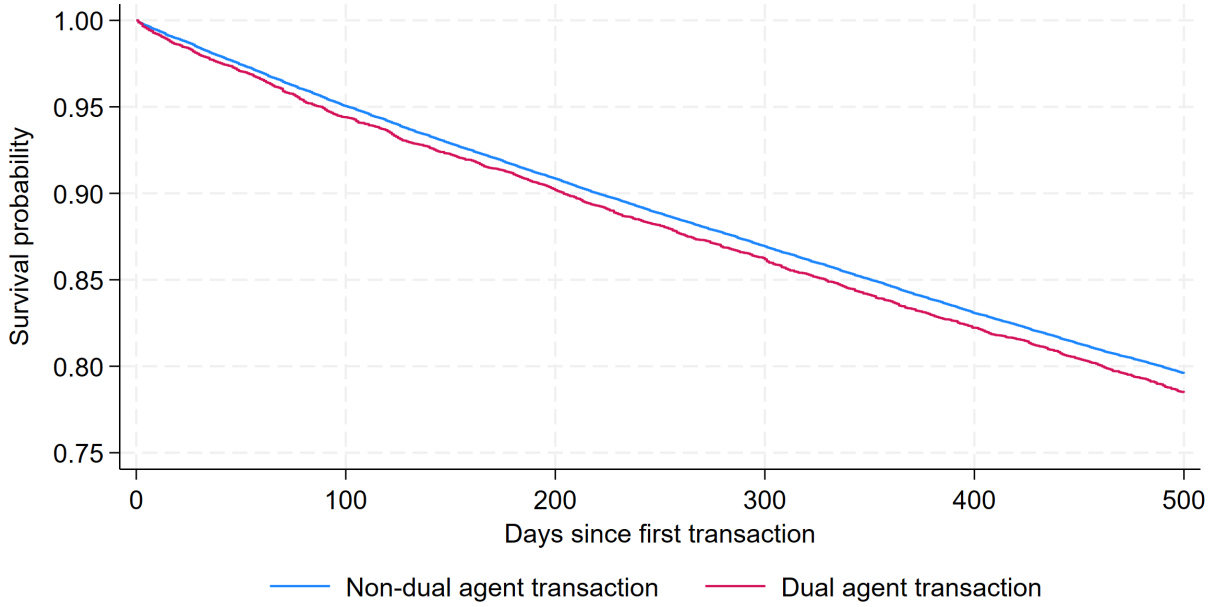

(B) 0–500 days

**Notes:** This figure plots Kaplan–Meier survivor functions based on the estimator in Equation (D.1) where the red line represents a group of counterparty agents with the top 25% centrality (pooled) and the blue line represents the rest. The horizontal axis measures days since the new entrant's first transaction, and the vertical axis reports the estimated survival probability. Panels (A) and (B) present the same comparison over the 0–60 day and 0–500 day horizons, respectively.

Figure D.2. Kaplan–Meier Survival by Dual-Agent Status of First Transaction



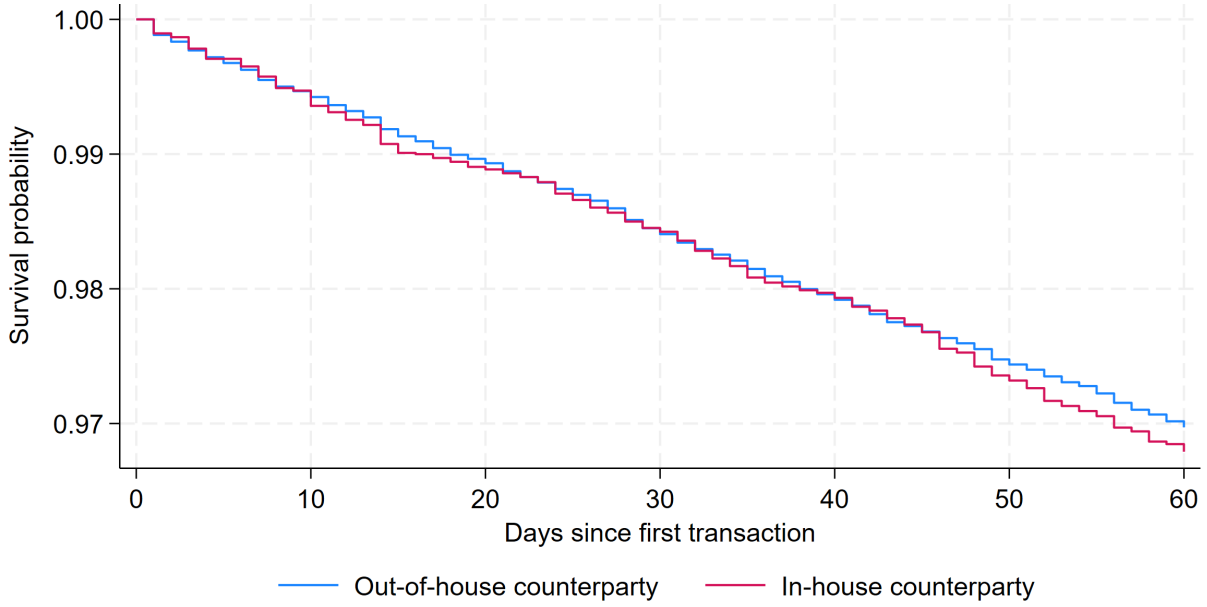

(A) 0–60 days



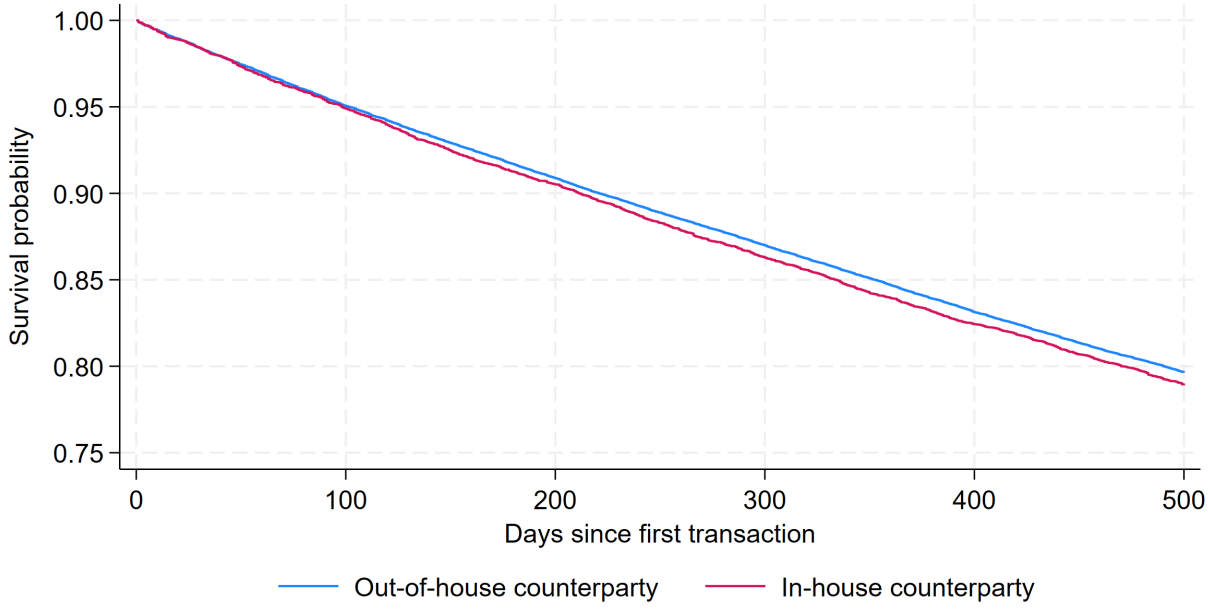

(B) 0–500 days

**Notes:** This figure plots Kaplan–Meier survivor functions based on the estimator in Equation (D.1) where the red line represents a group of new entrant agents whose first transactions are dual-agent deals and the blue line represents a group of new entrant agents whose first transactions are not dual-agent deals. The horizontal axis measures days since the new entrant's first transaction, and the vertical axis reports the estimated survival probability. Panels (A) and (B) present the same comparison over the 0–60 day and 0–500 day horizons, respectively.

Figure D.3. Kaplan–Meier Survival by In-House Counterparty Status



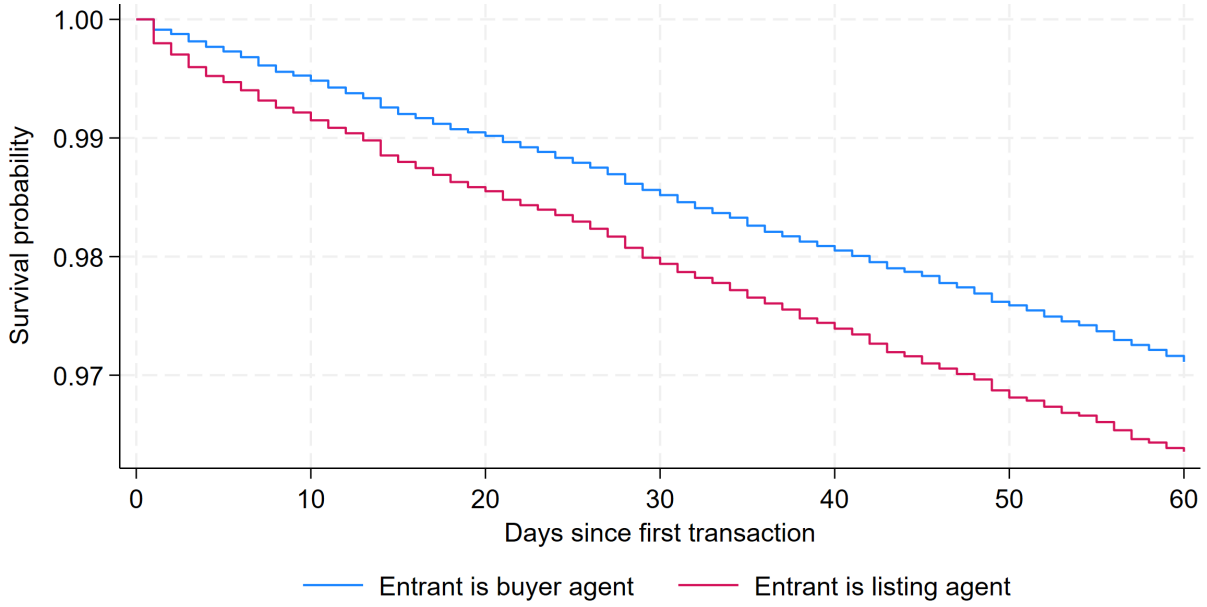

(A) 0–60 days



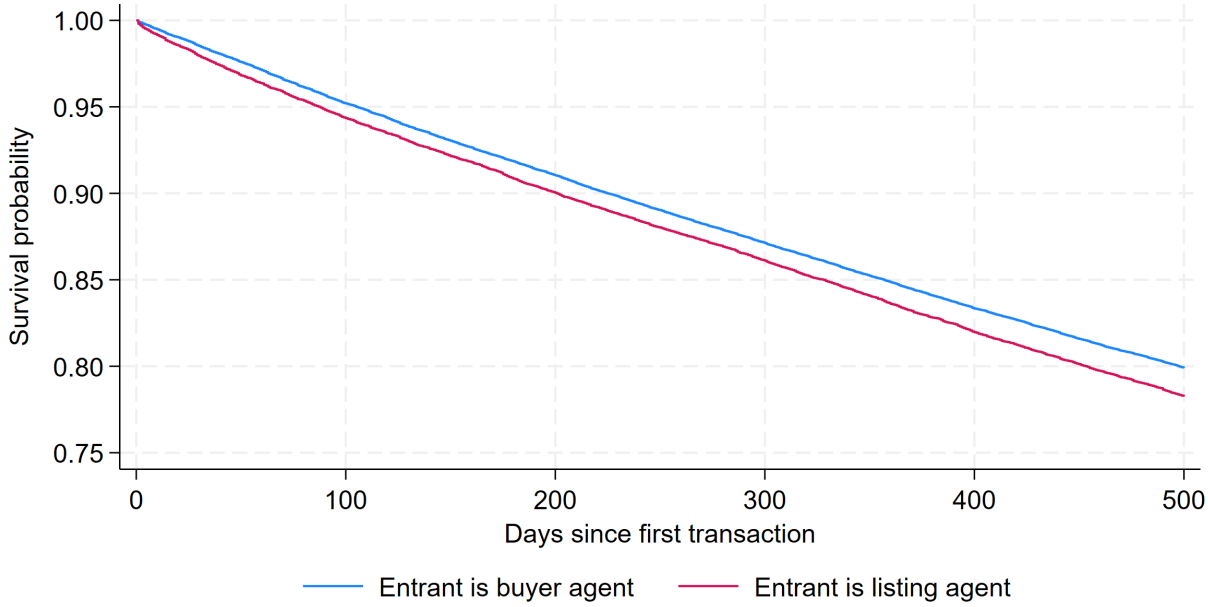

(B) 0–500 days

**Notes:** This figure plots Kaplan–Meier survivor functions based on the estimator in Equation (D.1). The horizontal axis measures days since the new entrant's first transaction, and the vertical axis reports the estimated survival probability. Panels (A) and (B) compare entrants with in-house versus (red) versus out-of-house (blue) first counterparties over the 0–60 day and 0–500 day horizons, respectively. Here, the out-of-house implies a different brokerage office. The comparison is restricted to non-dual first transactions so that a distinct counterparty is observed.

Figure D.4. Kaplan–Meier Survival by Entrant Role in First Transaction




(A) 0–60 days




(B) 0–500 days

**Notes:** This figure plots Kaplan–Meier survivor functions based on the estimator in Equation (D.1). The horizontal axis measures days since the new entrant's first transaction, and the vertical axis reports the estimated survival probability. Panels (A) and (B) compare entrants who first appear as buyer agents (blue) versus listing agents (red) over the 0–60 day and 0–500 day horizons, respectively.