How Much of Firm Performance Is Human Capital? A Variance Decomposition Below the Business Unit

Timothy Gubler
Brigham Young University

Gordon Scott
Washington University in St. Louis

Abstract: Variance decomposition research rarely extends below the business unit, leaving it unclear how much of firm performance reflects the people who do the work. Using 1.9 million real estate transactions involving 113,670 agents across 15,368 offices over seventeen years, we separate individual and organizational contributions to production. Individuals account for 44% to 48% of market-adjusted production variance after correcting for estimation noise, compared with approximately 6% for offices when accounting for individual differences. Individual performance persists across organizations, while moves to stronger offices are associated with additional gains. Gains also partly reflect when people move. These findings distinguish organizations that assemble productive people from those that make their people more productive.

Managerial Summary: Do successful organizations make their people more productive, or do they have more productive people? We study seventeen years of real estate transactions across three states, following agents as they change offices. Much of the difference in office performance reflects who works there. Individual productivity persists over time and across organizations. Agents moving to weaker offices show no decline in production, while those moving to stronger offices improve substantially. For businesses where production depends heavily on individuals and their client relationships, these findings highlight the importance of identifying, attracting, and retaining productive people. They also suggest that managers should evaluate organizational support by whether it helps the same people perform better.

Keywords: Strategic Human Capital, Microfoundations, Variance Decomposition, Employee Mobility, Firm Performance

INTRODUCTION

Strategy research has long sought to explain why some organizations persistently outperform others (Barney, 1991; Peteraf, 1993; Rumelt et al., 1994). For four decades, variance decomposition has been a principal tool in answering this question. By separating performance differences across the levels at which organizations are nested, this research has established that industry, corporate, and business unit effects each contribute, but that the business unit is the largest by a wide margin (McGahan & Porter, 1997; Rumelt, 1991; Schmalensee, 1985), a result that became central empirical support for resource-based explanations of advantage (Barney, 1991). Subsequent work extended the accounting upward to the corporate parent, the country, the strategic group, the CEO, and the alliance network (Bowman & Helfat, 2001; Kumar et al., 2022; Mackey, 2008; Makino et al., 2004; Quigley & Graffin, 2017; Short et al., 2007). Notably, almost every level added lies at or above the business unit, and the only individual level routinely examined is the CEO. Yet organizations produce through people, and a business unit effect cannot distinguish a unit with better routines from a unit with better people. The standard approach has therefore left open how much of the performance attributed to a unit reflects the people who do the work.

Strategy theory has meanwhile moved in the opposite direction. Microfoundations research argues that organizational constructs are incompletely explained until the individuals and interactions that produce them are specified (Barney & Felin, 2013; Felin & Foss, 2005; Felin et al., 2015), the knowledge-based view's firm-level locus of knowledge (Grant, 1996; Kogut & Zander, 1992) has been challenged on the grounds that collective heterogeneity may simply reflect the prior heterogeneity of the individuals nested within it (Felin & Hesterly, 2007), and strategic human capital

research holds that human capital is a source of competitive advantage that resides in individuals rather than the firm (Campbell, Coff, et al., 2012; Coff, 1997; Nyberg et al., 2014; Ployhart & Moliterno, 2011). Together, these literatures make individual action and human capital central to explanations of firm performance, but few studies have measured their contribution directly. Human capital is typically measured through inputs such as education and tenure and related to performance at the unit level (Crook et al., 2011; Ray et al., 2023), and worker-firm decompositions in labor economics examine wages rather than output (Abowd et al., 1999). Two studies come closest. Mollick (2012) shows that individual managers explained more variance in project revenue than the organization, and Bhaskarabhatla et al. (2021) find that inventor human capital explains far more of the variance in inventor output than firm capabilities. Both show that individual effects can rival organizational effects, but neither examines whether those effects travel across organizations. The empirical relationship is also not obvious. For instance, research on surgeons, equity analysts, and external hires finds individual performance to be substantially organization-specific and to decline on separation (Bidwell, 2011; Groysberg et al., 2008; Huckman & Pisano, 2006). We consequently know that individuals matter to firm performance, but it remains unclear how much is driven by the people versus the firm itself.

This gap matters because an organization-level performance difference is consistent with at least three accounts. An organization may create a context that makes its members more productive, it may assemble more productive people, or particular people and organizations may be especially productive together. The three generate identical firm-level variance but imply different sources of advantage, and current variance decompositions are unable to tell them apart. In this paper, we extend the decomposition to the individual level and separate individual contributions from organizational contributions. We ask how much variation in output is attributable to individuals,

how much is attributable to the organizations in which they work, and whether the individual component travels when people change organizations.

To answer these questions, we assemble a seventeen-year panel of residential real estate transactions recorded in the Multiple Listing Services of Utah, Arizona, and Idaho from 2007 through 2023, comprising 1,898,013 completed home sales produced by 113,670 agents affiliated with 15,368 brokerages. The setting is well suited because every transaction is attributed to named agents who are observed repeatedly, and mobility events are clear. We decompose the variance in agents' annual sales (net of local market conditions and tenure) across the agent, the office, and the franchise brand, and apply two independent corrections for the estimation noise that mechanically inflates variance components (Andrews et al., 2008; Kline et al., 2020). We also study agents who change offices, including 2,399 agents displaced by office closures, to determine whether performance travels with the agent across brokerages.

We find several important results. Our conventional decomposition, which does not account for individuals, attributes 39.7% of performance variance to the office. However, once we hold agents constant, the office component falls to 6.4%, and its covariance with the agent component is not distinguishable from zero, providing no evidence that productive agents sort into offices with larger organizational effects. Offices therefore differ mainly because they employ different people, not because they make people more productive, although the production that remains is worth more to weaker agents than to stronger ones. The individual component, by contrast, is large and concentrated. Agents account for 43.6% of performance variance, nearly five times the office share, while the franchise brand explains about 1%, and two independent corrections for estimation noise place the individual share between 44% and 48%. It is also stable, with performance in one half of an agent's career correlating at .74 with the other half and at .89 among the most precisely measured agents, and a small stable group accounts for more than a quarter of all above-average output. The

individual component also travels. Agents who move to weaker offices do not lose output, and neither do agents displaced when their office closes, so the decline that firm-specific human capital would predict does not occur. Most of what travels is the agent's own component, although leaving an office costs slightly less than arriving at an equally good one gains. Gains from moving nonetheless rise with the quality of the destination, and 69 offices show evidence of raising the performance of the agents who join them. Finally, most of the apparent return to changing offices reflects when agents move rather than the move itself. Voluntary movers improve by 0.161 standard deviations but move after a downturn in their own performance, while agents displaced by closures gain 0.045 from a flat baseline. The move thus accounts for just under a third of the gain and timing for the rest.

Our study makes three important contributions. First, we show that in one important context most of what variance decomposition has measured as a firm effect is actually a human capital effect. Building on Mollick (2012) and Bhaskarabhatla et al. (2021), we separate composition from organizational value added and find that composition dominates. The average firm-level production premium is small, and we find no evidence that productive agents and productive offices are concentrated together. The mobility results are most consistent with agents carrying their own capability, with a small additional part of what a prior employer contributed. Thus, assembling and retaining productive people may itself be an important capability, particularly in professional services, although a small tail of offices do improve agent performance. But selection and organizational value added are distinct mechanisms of advantage with different sources of imitability, and they should not be inferred from the same variance component.

Second, we provide direct evidence on the contribution of individual human capital to organizational performance, which the strategic human capital literature has had little opportunity to measure. Existing studies relate unit-level performance to unit-level aggregates of individual

characteristics (Crook et al., 2011; Ray et al., 2023), placing aggregates created by the same sorting process on both sides of the estimating equation. By analyzing more carefully what individuals produce, we find that in our setting individuals account for 44% to 48% of the variance in output, roughly seven times the organization's own contribution. This component is durable and follows individuals across organizations without loss. This helps answer the call to ground human capital-based advantage in individual-level evidence (Coff & Kryscynski, 2011; Wright et al., 2014).

Finally, our findings help clarify how organizations matter when the dominant performance resource is portable. Human capital theory emphasizes firm specificity, co-specialization, and mobility constraints as the mechanisms linking individual contribution to organizational advantage (Campbell, Coff, et al., 2012; Coff, 1997; Groysberg et al., 2008). In our setting those constraints are weak, and the organizational effect that survives is small on average and concentrated in a tail. Moving to a better office raises output modestly, and a small number of offices consistently improve the agents who arrive. Where production is individual and mobility is cheap, identifying, attracting, and retaining productive people matters more than transforming them after hiring. Development can matter a great deal, but the capability to make it matter appears scarce.

THEORETICAL DEVELOPMENT

The Seeds of Organizational Effects in Strategy Research

Variance decomposition occupies a central place in strategy because it connects theories of competitive advantage to empirically observed differences in performance. By partitioning performance across industries, firms, business units, and other levels, it asks where persistent performance heterogeneity resides. The answer matters because the level at which performance resides is also the level at which theory must locate the resources and mechanisms that produce it (Rumelt, 1991; McGahan & Porter, 1997; Bowman & Helfat, 2001).

Strategy has traditionally looked to organizations for that explanation. Business unit effects were interpreted as evidence of the imperfectly imitable resources emphasized by the resource-based view (Barney, 1991; Rumelt, 1991; Wernerfelt, 1984). Additional organization-level theories explained persistent performance differences through routines that coordinate individual actions (Nelson & Winter, 1982), knowledge that becomes productive when integrated with other knowledge (Grant, 1996; Kogut & Zander, 1992), and complementary assets that condition the value of a capability (Teece, 1986). Similarly, work on strategic human capital resources has argued that firm-level resources emerge through interaction and complementarities among members rather than through a simple sum of individual knowledge and skills (Ployhart & Moliterno, 2011). In each case, the organization contributes something its people could not produce on their own.

These theories share an observable implication. If the organization makes its people more productive, then an individual's performance should fall when the individual leaves, because the routines, colleagues, and assets that augmented it stay behind. The most direct evidence on performance heterogeneity below the business unit tests exactly this implication by following people across organizations, and some studies support the organizational view. For instance, cardiac surgeons improve with their volume at a particular hospital rather than with their total volume across hospitals (Huckman & Pisano, 2006). Star equity analysts suffer performance declines after changing employers, especially when they leave without colleagues (Groysberg et al., 2008; Groysberg & Lee, 2009). External hires underperform internal promotes for roughly two years despite receiving higher pay (Bidwell, 2011). Prior work therefore gives us good reason to expect an organizational component. But a business unit effect does not show which mechanism is operating. It may reflect what the organization does, whom it employs, or how particular people and organizations work together.

The Missing Individual Level

The standard approach leaves individuals outside the decomposition. Firm-level studies observe the performance of a business unit or firm, rather than the output of the people who produce it. The business unit is therefore treated as the smallest entity worth explaining, while differences among the people inside it remain part of the unit effect.

Microfoundations research challenges this approach by arguing that organizational explanations are incomplete unless they specify the individuals and interactions that produce them (Barney & Felin, 2013; Felin & Foss, 2005; Felin et al., 2015). Felin and Hesterly (2007) make the point most directly. Collective heterogeneity may simply reflect the prior heterogeneity of the individuals within the collective. What appears to be an organizational resource may therefore be the aggregate of individual knowledge, skill, and judgment. Organizations still coordinate and combine these contributions, and that coordination may create value beyond their simple sum. The point is not that organizational capabilities reduce to individual attributes. It is that their contribution cannot be assessed without first observing the people whose knowledge and actions produce them.

Human capital theory makes a related claim about the location of productive resources. Knowledge and skills are embedded in individuals, who decide where and how to deploy them (Becker, 1964; Coff, 1997). Firms must therefore attract and retain human capital to gain access to resources they cannot own in the same way that they own physical or legal assets (Campbell, Coff, et al., 2012; Chadwick, 2017). The relevant resource is also broader than ability. It can include effort, reputation, social ties, and relationships with clients. The individual component is thus a bundle of productive attributes and relationships, rather than innate talent alone.

This distinction matters because people can change organizations. Evidence from employee mobility shows that departures can reduce the performance of source firms, particularly when high-earning employees leave or when employees move together with colleagues (Campbell, Ganco, et al., 2012; Wezel et al., 2006). Knowledge and relationships can also move with individuals across

organizational boundaries (Almeida & Kogut, 1999; Song et al., 2003; Somaya et al., 2008). These findings do not imply that all human capital is general or fully portable. They imply that at least some productive resources are embodied in people and can travel with them, even though organizations determine how effectively those resources are combined with other assets.

Two conditions should determine whether individual differences remain visible in unit-level performance. First, individual output must not simply average away. If performance were normally distributed, most people would cluster near the mean, and differences among individuals would become less important as units grew larger. But individual performance is often skewed, with a small number of people producing far more than the rest (O'Boyle & Aguinis, 2012). Research on star employees reaches a similar conclusion (Aguinis & O'Boyle, 2014; Call et al., 2015). When output is concentrated in a few people, gaining or losing one of them can materially change a unit's total production. Workforce composition then becomes an important source of differences across units.

Second, individual output must be observable. Effort and knowledge are difficult for firms to observe directly, which creates the familiar problems of monitoring and team production (Alchian & Demsetz, 1972; Holmström, 1982). When output is attributed to a named producer and compensation is tied to that output, however, individual contribution becomes easier to measure (Lazear, 2000). These conditions also suggest where an individual component should be most visible: where production is individual rather than joint, where the relationships that generate business attach to the person rather than the firm (Broschak, 2004), where mobility is common enough to reveal portability (Campbell, Coff, et al., 2012; Marx et al., 2009), and where units are small enough that averaging cannot hide individuals. Where these conditions fail, the individual share should be smaller and the organizational share correspondingly larger.

These arguments do not imply that organizations are unimportant. They imply that organizational performance may reflect both what an organization contributes and the people it

employs. The key empirical questions are therefore how much output variation is attributable to individuals, how much is attributable to organizations after individual differences are accounted for, and whether the individual component travels when people change organizations. Two studies have taken a similar decomposition to the individual level. Mollick (2012) finds that individual middle managers explain roughly 22% of the variance in game revenue, compared with roughly 21% explained by the developing firm. Bhaskarabhatla et al. (2021) find that inventor human capital explains five to ten times more variation in inventor output than firm capabilities. Neither study examines whether the individual component is durable, whether it survives moves between organizations, or what the organizational component captures once individuals enter the accounting.

What Business Unit Effects Currently Capture

Once individuals enter the accounting in a variance decomposition, a business unit effect can represent at least three mechanisms. First, the organization may make a given individual more productive than that individual would be elsewhere through routines, complementary assets, training, or coordination (production mechanism). Second, the organization employs individuals who are more productive wherever they work (composition mechanism). Finally, particular people and particular organizations are productive together in ways neither would be separately (match mechanism).

Only the production mechanism supports the inference from a business unit effect to a firm resource in the sense intended by the resource-based view. Only under production does the organization contribute something its people did not bring with them. A conventional business unit effect therefore combines organizational value added, workforce composition, and matching.

Prior decompositions cannot separate these accounts because they observe unit performance rather than individual output and do not observe the same individuals producing in more than one organization. Labor economics addressed a similar problem in separating worker and firm

contributions to wages by observing workers across employers (Abowd et al., 1999; Card et al., 2013). The same logic applies to productivity. Mobility makes it possible to distinguish what individuals carry with them from what organizations add.

With individuals held constant, the remaining organizational component is the part most consistent with organizational value added. The difference between that component and the conventional business unit effect reflects workforce composition and sorting. The covariance between individual and organizational effects reveals whether productive people are concentrated in organizations that add the most to them. This is evidence of assortative sorting, not by itself of a match effect, which would appear as output from particular pairings beyond what the two effects together predict. Identifying either requires observing how the same individuals perform across organizations.

The estimates also require correction for sampling noise. Individual effects typically rest on fewer observations than firm effects, so their estimates contain more noise. That noise inflates the apparent contribution of the level at which the effect is estimated and biases the estimated covariance between worker and firm effects downward (Andrews et al., 2008; Kline et al., 2020). A real individual component should remain after this correction and should persist across independent periods, with the agreement strengthening as measurement error falls. Otherwise, it may reflect estimation noise or accumulated luck rather than durable human capital (Barney, 1986; Denrell, 2004). A durable component could still be diffuse, with small differences spread evenly across people, in which case the individual level would absorb variance without offering firms an actionable lever. If it instead concentrates in a small and identifiable group (Call et al., 2015; O'Boyle & Aguinis, 2012), talent identification becomes a first-order strategic activity and the loss of specific individuals a first-order risk.

This distinction also clarifies the role of selection. A firm may argue that assembling productive people is itself an organizational capability. We accept that possibility but distinguish it from the production mechanism, because the two imply different theories of how advantage is sustained. Production-based advantage depends on what the firm enables its people to do, and it is protected to the extent that mechanism is causally ambiguous or socially complex (Barney, 1991; Lippman & Rumelt, 1982). Composition-based advantage depends on finding and retaining people who are productive wherever they work, and it is protected only to the extent competitors cannot bid the same people away, which makes it a function of labor market frictions rather than organizational inimitability (Campbell, Coff, et al., 2012; Chadwick, 2017). The two mechanisms may both create value, but they should not be inferred from the same variance component. Notably, a small average production component does not mean that no organization improves performance. The resource-based view expects a production-based advantage to be rare, so once composition is held constant the place to look for it is the tail of the organizational distribution. An organization in that tail should be distinguishable from a lucky or well-staffed organization in two ways: the same people should produce more there than they do elsewhere, and the difference should persist across independent periods.

Mobility as a Test of Where Performance Resides

Mobility provides the clearest test of whether organizations produce individual productivity or assemble it. It also determines whether firms can retain the value their people create (Chadwick, 2017; Coff, 1999). If productivity is jointly produced with firm-specific assets, routines, or colleagues, performance should decline when the individual leaves and recover only as the individual becomes familiar with the destination (Bidwell, 2011; Groysberg et al., 2008; Huckman & Pisano, 2006). If performance rests primarily on general human capital, it should be preserved across moves.

In that case, firms retain value in the face of employee mobility by offering a superior context, providing attractive terms, or limiting mobility (Campbell, Coff, et al., 2012; Kryscynski et al., 2021).

A third possibility looks similar at the moment of separation. An organization may build general capability in the people it employs, so that performance survives departure because the organization's contribution has become the individual's own portable human capital. These explanations can be distinguished by how the origin and destination enter the change in output around a move. Under pure portability, the individual loses the origin's context and gains the destination's, so the change should depend on the difference in their quality and the two should enter with equal and opposite weight. Under portable capability building, part of the origin's contribution stays with the individual, so the origin should weigh less than the destination. This comparison of origin and destination effects is the empirical test that distinguishes portable capability building from pure portability.

Organizational context may add to a portable individual component. If context complements what the individual brings, moving to a stronger organization should raise output, while moving to a weaker one should not reduce it (Milgrom & Roberts, 1990; Teece, 1986). The distinction from firm specificity is whether moves in the opposite direction produce a loss.

The timing of mobility creates a final problem. People choose when to make voluntary moves, and they often move after a downturn in their own performance. A similar pattern appears among entrants to training programs, whose earnings dip before entry and recover afterward (Ashenfelter, 1978; Heckman & Smith, 1999). A before-and-after comparison then combines the effect of the move with a recovery that would have occurred anyway, and because the dip can build over several years, a flat pre-trend over a short window does not rule it out. Separating the two requires moves whose timing the individual did not choose. The difference between chosen and unchosen moves

indicates how much of the apparent return to mobility reflects the timing of the move rather than the organization that receives the individual.

Taken together, these arguments yield three questions. How much variation in output is attributable to individuals? How much is attributable to the organizations in which they work? And does the individual component travel when people change organizations? The first compares the relative importance of the two levels. The second separates organizational value added from composition and sorting. The third distinguishes portable human capital from productivity that is specific to the organization. We investigate these questions empirically.

EMPIRICAL SETTING

Answering these questions requires output attributed to individuals and observed repeatedly over long careers. Individuals need to produce in more than one organization, including some whose moves they did not choose, and paid on that output so the record measures production rather than evaluation. Residential real estate meets all four requirements. It also meets the conditions under which an individual component should be most visible, since production is individual, client relationships attach to the agent, mobility is inexpensive, and offices are small. These features make this setting useful for investigating our research questions.

Agents are licensed professionals who must affiliate with a brokerage in order to transact, but who often operate similar to independent contractors. The brokerage supplies a brand, office space, supervision by a principal broker, training, and sometimes referral flow. The agent supplies nearly everything else, finding clients, marketing listings, and managing the transaction. Compensation is almost entirely commission-based, although agents often pay brokerages a flat desk fee. The commission on a sale is split between the listing and buyer agents and then between each brokerage. An agent who closes nothing earns nothing. For each home sale captured by the local multiple listings service credit is attributed to both sides of the transaction.

Clients generally choose an agent rather than a brokerage. Most buyers and sellers find their agent through referral or prior use (DiMaggio & Louch, 1998; Gubler, 2019; National Association of Realtors, 2023), and an agent's access to transactions depends heavily on the agent's own social and relational capital (Gubler, 2019; Gubler & Cooper, 2019). Reputation likewise accumulates to the agent and to a lesser extent the office. When an agent changes offices, many of the relationships that generate business move with the agent. Importantly, a listing agreement is a legal contract with the brokerage, so brokerages can retain active listings when an agent departs. This makes our portability estimates conservative.

Mobility is relatively cheap and common in this context. A license belongs to the agent and is re-affiliated with a new brokerage through a transfer at nominal cost. Non-compete agreements are rare and agents maintain their reputation, transaction history, and client base. Of the agents observed long enough for our mobility analyses, 14,099 move exactly once and 18,365 never move. Brokerages also sometimes close, and we identify 249 closures that displaced 2,399 agents in our data. These closures allow us to observe more exogenous moves whose timing the agent did not fully choose.

The brokerage is our focal organizational unit because it is where agents affiliate and where supervision and local resources are organized. Brokerages range from a single agent to well over one hundred. Many belong to a national or regional franchise (e.g., ReMAX, Coldwell Banker, etc.) that standardize branding and share best practices across offices. About 48% of agents work at a franchise-affiliated office at some point in our data. Because offices in our setting are small, individual differences are less likely to be averaged away in office-level performance. This feature makes the setting well suited to measuring individual heterogeneity, but it also limits how broadly the results should generalize, which we address later.

DATA AND EMPIRICAL APPROACH

Data and Sample Construction

Our data come from the ATTOM national MLS listing-analytics extract for Utah, Arizona, and Idaho, covering residential transactions from 2007 through August 2024. The raw data contain 14,208,564 listing-status records. Because each status record reports a single state of a listing, we collapse the panel to one row per listing, yielding 4,206,982 distinct listings geocoded to five-digit ZIP codes. Restricting the data to sold listings attributed to a valid listing office, excluding records flagged as existing solely for appraisal or comparable purposes, and removing duplicates, listings before 2007, and listings without a usable ZIP code leaves 2,472,825 sold listings. We then narrow the sample to completed sales of homes and drop leases and rentals, non-residential property types, listings credited to placeholder or non-member agent identifiers, sale prices below $10,000, impossible marketing durations, properties recorded twice by overlapping listing services, and sales in 2024 (the datasets final partial year). This leaves 1,898,013 sold transactions. We credit each transaction to both its listing agent and buyer agent because brokerages earn revenue on both sides of a deal. The buyer agent is identified for 90.2% of sold transactions, with coverage of 95.5% in Idaho, 95.6% in Arizona, and 81.5% in Utah. Coverage is lower in part because several listing services credit buyer agents who are not members to a single non-member identifier, which we treat as unidentified rather than as an agent (Appendix E.1). Each MLS assigns agents a single member identity across both roles. The resulting sample contains 113,670 agents affiliated with 15,368 offices. Sales transacted outside a listing service do not appear in the data. Appendix E details each restriction with data counts.[1]

[1] One restriction applies only to the transaction-level price decomposition in Appendix A. Because that analysis regresses sale price on physical property characteristics, it uses the 810,471 listings that report complete information on those characteristics, or 29.4% of the transaction population. All analyses reported in the main text use the full sample because they do not require property characteristics.

For the individual-level analyses we build an agent-year panel by crediting each transaction to both its listing and buyer agent and then collapsing the data to one observation per agent per calendar year. The panel contains 571,295 agent-years across 113,670 agents and is the unit of analysis for everything that follows.

Descriptive Statistics

Table 1 reports descriptive statistics for our sample. Three features of the data are important for the analyses that follow. First, individual output is highly skewed, consistent with the heavy-tailed performance distributions documented by O'Boyle and Aguinis (2012). The median agent closes three transactions a year (mean of 6) and closes 7 transactions over their career (mean of 31). The most active agents close thousands, although the extreme upper tail is institutional rather than individual: a member identifier denotes a license, and the most active belong to iBuyers, homebuilders, and distressed-asset disposition contracts that transact through one agent of record (Appendix E.3). Second, mobility is common. Agents work at 1.71 offices on average and 41.1% change office at least once, which is what identifies the two-way and event-study models below. Third, most offices are small. The median office-year has two agents and three quarters have four or fewer, although because large offices hold a disproportionate share of agents, the mean agent works alongside 35 colleagues. The median size of these firms is far below the median size of firms in the other firm-level decompositions, so individual heterogeneity, if present, should be visible in office performance rather than averaged away. The sample also spans a full market cycle, from the 2009 trough during the housing bust through the expansion the following decade to the cooling after 2021. The market adjustment described below is constructed so that these aggregate swings are not mistaken for individual performance. Appendix E reports the supporting distributions.

[Insert Table 1 about here]

Dependent Variable: Sales Performance

Our main dependent variable uses agent transactions (homes bought or sold) each year. However, a raw measure of yearly sales confounds individual performance with market conditions. An agent working in an expensive high demand area will see large dollar volumes regardless of skill, while an agent in a small town might have much lower volume with higher skill. To account for this we control for market conditions, residualizing each agent's log annual sales volume in a ZIP-by-year market, controlling for tenure. We use the following model:

$$\ln(\text{Volume})_{it} = f(\text{Tenure}_{it}) + \lambda_{z(i,t)\times t} + \varepsilon_{it} \qquad (1)$$

where $\lambda_{z\times t}$ is a fixed effect for each ZIP-year, absorbing how favorable a market was in a particular year, and $f(\text{Tenure})$ controls for agent experience. Each agent-year is assigned its modal ZIP but our results are similar using a weighted measure (Appendix E.2). Together the market and tenure terms explain about a third of the variance of log annual volume. The residual ε is the part of an agent's annual volume that market conditions and experience do not explain. We divide it by its standard deviation so that it is measured in standard deviation units, which puts agents in expensive and inexpensive markets on the same scale and lets every effect in the paper be read the same way. A value of 0.10 means an agent produced one tenth of a standard deviation more in that year than the market and experience alone would predict. Throughout the paper, market-adjusted performance refers to this residual.

Importantly, annual sales volume can be separated into two parts, because an agent's dollar volume in a year equals the number of transactions closed times the average price per transaction. In logs the product becomes a sum, ln(Volume) = ln(Transactions) + ln(Average price), so market-adjusted performance divides exactly into a component reflecting how many deals an agent closed and a component reflecting the prices of those deals. The decomposition below uses this split to ask whether individuals differ in the number of transactions they generate or in the prices they achieve.

Empirical Strategy

Our empirical strategy proceeds in three parts. The first is the variance decomposition. Because the agent, the office, the franchise brand, and the ZIP market overlap, any sequential attribution depends on the order in which levels are entered. We therefore use a Shapley decomposition, crediting each level its average marginal contribution across all orderings, so that shares do not depend on order (Sharapov et al., 2021).

The second part corrects for estimation noise. Each agent effect is estimated from a finite number of years, and each office effect from a finite number of agents, so every estimate equals the unit's true contribution plus whatever good or bad fortune happened to occur in the years we observe. A decomposition that fails to account for this credits each level with its own sampling error and does so most generously at the level that is measured least precisely (Andrews et al., 2008). We remove the noise in two ways. The first (sample-splitting correction) splits each agent's career into two halves that share no years. Transitory shocks do not repeat from one half to the other, but ability does, so the covariance between the two half-career estimates recovers the real part of the individual component. The second is the leave-out correction of Kline et al. (2020), which applies the same logic one observation at a time inside a two-way model with agent and office effects. It also produces the covariance between agent and office effects, which measures whether productive agents are concentrated in productive offices. We test the assumptions this estimator requires below.

The third part uses mobility to separate production from composition and to test whether individual performance is portable. We estimate event studies that trace an agent's market-adjusted performance in the years before and after a change of office. The sample includes agents who move exactly once and agents who never move. The never-movers are needed because, with agent and year fixed effects in the model, a sample made up only of movers cannot separate the effect of the move from the effect of calendar time, since every mover experiences both (Borusyak et al., 2024; Sun & Abraham, 2021). We use this model:

$$y_{it} = \alpha_i + \delta_t + \sum_k \beta_k \, D_{it}^k + \varepsilon_{it} \qquad (2)$$

where y is market-adjusted performance from equation (1), α_i and δ_t are agent and year fixed effects, and the D^k variables are indicators for years relative to the move. The year before the move (k=-1)is the omitted baseline, and the endpoint indicators (k=-5 and k=5) absorb all years five or more before or after the move, so no observations are dropped. We cluster standard errors by agent. Because the first and last calendar years an agent appears are mechanically partial, our main specification drops each agent's first and last observed year.[2] Because agents who move voluntarily choose when to do so, and tend to choose a moment after a downturn in their own performance, we also estimate the same event study for agents displaced by office closures, whose timing was not their own. The gap between the two estimates measures how much of the apparent return to moving is timing rather than the move itself. Finally, we apply the same design to arrivals at the small set of offices that pass our screens for a genuine organizational effect.

RESULTS

Main Effect Results From the Variance Decomposition

Table 2 presents results from the Shapley decomposition. Of the variance in the agents' market-adjusted annual production, the individual agent accounts for 43.6% and the brokerage office for 9%, a ratio of nearly five to one. This is the ordering the individual-level view predicts and the organizational view does not. The remaining 47% is year-to-year variation within agents that neither the agent's own average nor the office explains. The ordering does not depend on the sample. The agent share is 47% using all agents, 44.7% among agents observed at least three years, and 43.6% among agents observed at least five. The office share moves little across the three. We report the

[2] Section 6.1 shows the results survive and are similar including these years

five-year sample as our main specification because its agents effects are the best measured, and the correction results below remove the remaining noise formally rather than by dropping observations. Adding the franchise brand, which can be estimated only among franchise-affiliated offices, leaves the agent share higher at 48.7%, the office at 4.5%, and the brand at 1.2%.

The individual share is this large because annual volume depends mostly on how many deals an agent closes, and deal count is where individual agents differ most. Splitting production into its count and price components, the agent accounts for 46.0% of the variance in deal count but only 25.8% of the variance in price (see Table 2, Panel B). What an individual brings to a brokerage is overwhelmingly the ability to generate and close transactions rather than to obtain a higher price for a home. Two checks on the appendix support this. At the level of a single sale the individual's share is smaller but still exceeds the office's, at 16.3% against 10.5% (see Appendix A, Table A2). And the deal-count advantage is not an artifact of where agents work or what they specialize in, since market conditions explain about 15% of the variation in deal counts and observable specialization under 9% (see Appendix B, Figure B1).

[Insert Table 2 about here]

Testing Production, Composition, and Sorting

A conventional decomposition shows that offices differ, but not why. Separating production from composition requires observing the same individual at more than one office, so that the individual stays fixed while the organization changes. Whatever moves with the individual is their effect, and whatever stays at the office is the office effect.

Table 3 shows this separation in four steps. Row 1 is the conventional variance decomposition. It compares offices only with other offices, which is what prior firm-level studies have done because they observe what a unit produced without observing the people who produced it (McGahan & Porter, 1997; Rumelt, 1991). On that basis, offices account for 39.7% of performance variance, and

whatever the workforce contributes is folded into the office effect. This is the number the tradition would report as the firm effect in our setting. Row 2 applies the sample-splitting correction to the office effects, since an office's estimated effect contains the luck of the agents who happened to work there, and leaves 32.4%. Row 3 adds agents to the model, so that each office is judged against its own agents' output at the other offices where they worked. This removes composition and shows the office's contribution falls to 9.9%. Row 4 applies the leave-out correction to the two-way estimates. In this model the noise comes mainly from the few movers who connect each office to the others, and removing it leaves 6.4%. These results show that roughly three quarters of the conventional firm effect is composition and another tenth is estimation noise. Offices differ because of who works in them, not because of what they do for the people who work there. The 9.0% office share in Table 2 is a Shapley share rather than one of these four estimates, for the reason explained in Appendix C.4.

Sorting is also present but does not change the story much. Good agents cluster together, with a correlation of .30 between an agent's corrected effect and the average quality of the agent's officemates. However, the covariance between agent and office effects, which measures whether productive agents are concentrated in the offices that add the most to them, is essentially zero at negative 1.8% of variance. Talent concentrates, but we find no evidence that assembling it produces anything beyond the talent itself.

The agent column of Table 3 provides the test the theory posed. If differences among agents were mostly sampling error, then removing this should make those differences largely disappear. It does not. Compared only with other agents and uncorrected, agents account for 54.6% of performance variance (row 1). Removing noise by estimating each agent's effect separately on two halves of the career that share no years, so that ability repeats across halves but luck cannot, leaves 48.4% (row 2). About seven eighths of the spread among agents therefore reappears in data it was

not measured on. Adding offices to the model leaves 52.7% (row 3), and removing noise from that estimate on the connected set of 154,434 agent-years leaves 44.4% (row 4). The two corrections rest on different assumptions and show a corrected individual share of 44% to 48%. The office share thus falls by a factor of six as composition and noise are removed, while the individual share loses less than a fifth and remains the largest component.

[Insert Table 3 about here]

Are the Corrected Estimates Trustworthy?

The corrected estimates in Table 3 are important to the paper's central claims, so we next test the assumptions behind them following the diagnostics that Bonhomme et al. (2026) recommend for two-way worker and firm models. We report the intuition and the result for each test here and leave the construction to Appendix C.

The most serious concern is that the office effect rests on little data, since offices are compared only through agents who worked at more than one of them, and a sparsely connected market can manufacture an office effect out of the luck those few movers happened to have (Appendix C.1). We therefore thin the data deliberately and watch what each estimator does, following the sub-sampling diagnostic of Andrews et al. (2012). Discarding agents at random across sub-samples that span 4.4 to 12.1 movers per office, the uncorrected office share moves from 22.1% to 9.9%, more than halving purely because of how well connected the data are. The corrected office share also moves, from 12.7% to 6.4%, so the correction removes about half of that effect rather than all of it, while the corrected individual share is far steadier at 47.6% to 44.4%. Because the remaining bias falls as the number of movers per office rises, we report estimates from the full sample, which has the most movers per office and therefore the least bias. And because the remaining bias inflates the office share, the corrected office component of about 6% is an upper bound on how much the

average office adds, which strengthens rather than weakens the conclusion that composition dominates.

Two further assumptions matter. The first is additivity, meaning that an office adds roughly the same amount to every agent. If it does not, the office effect in the two-way model is an average across agents rather than an increment that any particular agent receives. Among movers, a one standard deviation improvement in office quality is worth 0.396 SD to a weaker agent but only 0.176 SD to a stronger one, a difference of -0.220 (SE 0.073, $p = 0.003$). Additivity is therefore rejected, and the violation runs toward substitution rather than complementarity. A good office does most for the agents who bring the least. The office shares in Table 3 should thus be seen as averages across agents, and the organizational contribution is concentrated among weaker agents. A second test, comparing moves between the extremes of the office quality distribution, does not reject additivity, but has far less power than the direct test and we report it for completeness (Appendix C.2). The second assumption is that agents do not move in response to a temporary slump, since a rebound after such a move would masquerade as an organizational effect. Here we find that agents moving to "worse" offices were already declining in performance, while agents moving to "better" ones were not. The resulting bias runs against our argument rather than for it, because attributing an agent's own decline to the destination office makes organizations look more consequential than they are. The office closure design in the mobility analysis also does not depend on this assumption

Is the Individual Component a Resource?

The theory identified two properties the individual component must have before it can be called a resource rather than accumulated good fortune (Barney, 1986; Denrell, 2004). It must persist, and it must be concentrated enough that particular people matter.

To test persistence, we estimate each agent's performance separately on the odd-numbered and even-numbered years of their career, so that the two estimates share no observations. Across the

52,525 agents observed in both halves, the two estimates correlate at .737 (rank correlation .717). More telling, a real trait measured with less error should correlate more strongly with itself, while luck should not. We find the correlation climbs steadily with the number of transactions per half, from 0.170 among the most thinly observed agents to 0.886 among agents with more than 25 transactions per half (Figure 1). In concrete terms, 72.0% of agents in the top fifth of their odd years are in the top fifth again in their even years, and only 3.5% fall to the bottom two fifths. The test can only use agents observed in both halves, but that restriction covers 79.4% of agents and 98.3% of all listings, and demanding still more survival raises the correlation only slightly, which is not how a selection artifact behaves (Appendix B).

[Insert Figure 1 about here]

Concentration also matters because a durable component spread evenly across everyone would support the microfoundational claim while implying that no single individual matters much to the firm. We find it is not spread evenly. Summing each above-average agent's market-adjusted performance above the mean, the top 10% of agents account for 47.9% of that total and the top 5% for 28.5%, the concentration that heavy-tailed distributions predict (Call et al., 2015; O'Boyle & Aguinis, 2012). Removing the top 5% of agents reduces the variance between agents by 25.5%, five times what removing 5% of people would cost if performance were evenly spread, and career volume has a Gini coefficient of 0.628. The top performers are also the same people over time. The top quintile is the most persistent part of the distribution, with a split-half correlation of .682 against .172 to .291 in the middle three quintiles (Figure 2). Because the same individuals are both the high performers and the most stable, this is an identifiable group rather than a rotating set of agents with temporarily high performance, and the loss of specific people is a first-order risk rather than ordinary attrition.

[Insert Figure 2 about here]

DOES INDIVIDUAL PERFORMANCE TRAVEL?

The decomposition results show that performance variation primarily attaches to individuals. Mobility provides the more direct test of whether organizations drive that productivity or assemble it, because it reveals what happens when the organization changes while the individual does not. Prior evidence predicts a decline on separation (Bidwell, 2011; Groysberg et al., 2008; Huckman & Pisano, 2006). Testing that prediction requires accounting for selection, because agents who move voluntarily tend to move after a downturn in their own performance, which confounds the move with a recovery that would have happened anyway. We document the problem, then address it with comparisons among movers leveraging unexpected office closures to capture moves where the timing was less likely to be the agent's choice.

Voluntary Moves and the Timing Challenge

Figure 3 plots the coefficients, each measuring performance in a given year relative to the year before the move. Performance is 0.144 SD higher in the first full year at the new office and rises to 0.151 and 0.165 SD higher in the two years after that. The interaction-weighted estimator of Sun and Abraham (2021), which guards against bias from comparing movers who moved at different times, gives nearly identical estimates. Taken at face value, changing offices raises production by about a sixth of a standard deviation.

[Insert Figure 3 about here]

However, the years before the move show why that interpretation is incorrect. Every coefficient measures performance relative to the year before the move, so the design assumes that year is representative of the agent. It is not. Five years before the move these agents were performing 0.104 SD above that baseline, and their performance declined steadily from there (pre-trend flatness rejected, $p < .001$). Movers experienced a downturn and then moved. Because the baseline year is the low point, any recovery toward the agent's long-run productivity appears as a gain from moving

even if the move did nothing. The post-move gain therefore describes what surrounds a voluntary move, but it cannot be interpreted as the move's effect. Specification checks in Appendix D, including the treatment of partial first and last career years, do not change these estimates.

Moves Up and Moves Down

The timing problem does not affect comparisons among movers, because a downturn shared by all movers cancels out when one group of movers is compared with another. We therefore split movers by whether they moved to a better or a worse office, measuring each office's quality as the average market-adjusted performance of its other agents (excluding the mover's own outcomes from both the origin and the destination). The two groups behave very differently (Figure 4). The 4,765 agents who moved to a better office gain 0.220, 0.231, and 0.242 SD in the three years after the move, each with a standard error below 0.020, from a flat pre-move window ($p = .608$). The 4,830 agents who moved to a worse office lose nothing either, with post-move coefficients of 0.035, 0.054, and 0.059 SD that are jointly different from zero ($p = 0.007$) but roughly a quarter the size of the gains above. If performance depended on firm-specific human capital, agents moving to weaker offices should decline, but we do not find evidence of this. Agents moving to stronger offices improve performance, so organizations add to what the individual brings rather than substituting for it.

[Insert Figure 4 about here]

A sharper test asks what exactly travels with the agent. We regress each mover's change in performance around the move on the quality of the origin office and the quality of the destination office, entered separately. If agents carry only their own ability, then leaving an origin brokerage should cost exactly what arriving at a destination brokerage of equal quality gains, so the two coefficients should be equal in size and opposite in sign. If agents also carry capability their previous office built in them, part of the origin's contribution stays with them when they leave, so the origin coefficient should be smaller in size than the destination coefficient. The estimates are 0.330 for

destination quality and -0.266 for origin quality, and equality in size is rejected ($p = .037$). The origin office coefficient is the smaller of the two, so most of what agents carry is their own component while a small part of what the origin office built appears to leave with them. The gap is 0.064 SD against coefficients of roughly 0.3. Combining the two, the post-move gain rises by 0.300 SD for each standard deviation of improvement in office quality. Alternative constructions that guard against a depressed baseline, mismatched measurement windows, and error in the quality measure place this slope between 0.15 and 0.30 SD (Appendix D). Because the pre-move decline among agents moving to worse offices would inflate rather than depress the slope, the lower end of the range is the more conservative estimate.

Moves the Agent Did Not Choose: Forced Mobility From Office Closures

These mobility estimate share the weakness that agents chose when to move, and on average they move after a downturn in performance. Office closures supply moves whose timing was set by the firm rather than the agent. We identify closures with four screens requiring an operating history, a non-trivial final-year headcount, dispersal to at least two destinations, and a closure date before the last fully observed year (Appendix D.2). Two hundred and forty-nine brokerage closures pass these screens, spread across 2009 to 2021, displacing 2,399 agents.

We find that agents displaced by closures show no pre-move performance trend ($p = .192$), which is exactly what the voluntary movers lacked. Their post-move gain averages 0.045 SD (SE 0.022, $p = 0.035$), compared with 0.161 SD (SE 0.010) for voluntary movers over the same three-year horizon. The difference of 0.115 SD has a standard error of 0.023 ($p < .001$), and the two profiles differ jointly ($p < .001$).

[Insert Table 4 about here]

Three conclusions follow from these results. First, individual performance appears to be portable. Agents who move to weaker offices do not lose performance, and neither do agents

displaced when their office closes, so the decline that firm-specific human capital predicts does not occur. This is the best supported of the three conclusions, because it rests on the absence of a decline rather than the size of a gain. Second, organizations add to what the individuals bring. Moving to a better office raises output by 0.15 to 0.30 SD per standard deviation of office quality, and the near-symmetry of the origin and destination coefficients shows that most of what travels is the agent's own ability, with a small part attributable to capability absorbed from a prior office. Third, moving itself raises production only modestly. Agents who chose neither whether nor when to move gain about 0.045 SD, just under a third of the 0.161 SD that a comparison of voluntary movers reports. The rest reflects when agents chose to move rather than the move itself.

DO ANY OFFICES MAKE THEIR AGENTS BETTER?

The corrected office share of about 6% is an average across thousands of offices, and an average cannot distinguish a world in which no office adds much from one in which a few offices add a great deal and the rest add little. The resource-based view expects the second, since what sustains advantage is rare and hard to imitate (Barney, 1991; Lippman & Rumelt, 1982). We therefore ask whether there are outlier offices that make the agents who join them more productive.

An office can look exceptional for two reasons that have nothing to do with developing its agents. The first is luck, since a small office can post an extreme average by chance (Denrell, 2004). We screen this out by requiring an office's quality to reproduce itself across the two halves of the panel, as we did for agents. Among 2,552 offices with at least five agent-years in each half, the two halves correlate at .886, so office quality is a stable attribute. The second is sorting, since an office may simply employ good agents. We screen this out by scoring each office on how its own agents performed there relative to their average at the other offices where they worked, leaving out each agent's own contribution (Appendix F). Because the comparison is within person, hiring good people cannot produce it. Intersecting the two screens leaves 69 offices, about 4% of the 1,672 for

which both can be computed. A within-agent boost could still reflect timing rather than development if these offices recruit agents at a low point, but an arrival event study rules this out. Agents who join one of the 69 offices gain 0.589, 0.543, and 0.523 SD over the three following years against 0.033 to 0.047 SD at ordinary offices, and the gain remains large when arrivals with unusually low pre-arrival performance are dropped (Appendix F).

These results establish that such offices exist, but not how large their effect is. The 69 offices were selected because their agents outperformed there, and the event study is estimated on those same outcomes. An office that passed the screen partly on luck will carry that luck into the event study as well, so the estimated gain is an upper bound on what these offices contribute. A cleaner estimate would identify the offices using odd years and estimate arrival effects using even years, and we leave this extension for future work. The result complements the decomposition rather than contradicts it. The typical office is close to interchangeable, which is why the average office share is about 6%, but a few dozen identifiable offices substantially raise the performance of the agents who join them. The organizational capability to develop individuals exists, but it is scarce.

DISCUSSION

Our theory argued that a firm effect combines production, composition, and match mechanisms, and that variance decompositions stopping at the business unit level cannot separate them. Our results in this paper separate them in one important context. Our unit-level decomposition attributes 39.7% of performance variance to the brokerage, similar to prior work, but we find almost all of this reflects composition. Once we hold agents constant, the average production component falls to about 6%, and the covariance between agent and office effects is indistinguishable from zero, so productive agents are not concentrated in the offices that add the most. The individual component, in contrast, is large at 44% to 48% after correcting for estimation noise. It is also durable, concentrated in a small and stable group of top performers, and portable, since neither voluntary

moves to weaker offices nor displacement by office closures reduces it. The mobility results show that roughly 70% of the apparent return to changing offices reflects when agents choose to move rather than the move itself. Organizations still matter in two ways, however. Agents who move to better offices produce more, so organizations add to what the individual brings, and a few dozen offices demonstrably improve the agents who join them.

These findings have two implications for the variance decomposition tradition. The first concerns interpretation. Most of what a firm-level decomposition in this setting would report as a firm effect is workforce composition, and composition is a different source of advantage than the organizational resources the tradition has inferred from firm effects. Production-based advantage is protected by causal ambiguity and social complexity, while composition-based advantage is protected only by labor market frictions (Campbell, Coff, et al., 2012; Lippman & Rumelt, 1982). Studies that interpret firm effects as evidence of firm resources may therefore be measuring, in large part, the outcome of a competition for people. The second concerns estimation. Estimated effects contain measurement error, so adding a level whose effects are imprecisely estimated produces a component for that level even when its true contribution is zero. Decompositions that report uncorrected shares, without a correction or a diagnostic showing one is unnecessary, cannot distinguish a real effect from the precision with which it was measured.

For strategic human capital research, our results provide the direct individual-level estimate the literature has called for and qualify the firm-specificity mechanism that much of this research assumes. In a setting where production is individual and clients attach to the person, individual human capital crosses organizational boundaries intact, in contrast to the decline on separation found among surgeons, analysts, and external hires (Bidwell, 2011; Groysberg et al., 2008; Huckman & Pisano, 2006). The origin and destination effects are close to symmetric, which indicates that most of what travels is the agent's own component, with a small part attributable to capability built by a

prior office. For mobility research, our results show that when workers choose when to move, event-study estimates capture the timing of the decision along with its consequences. In our data the timing component accounts for roughly 70% of the apparent gain, and a test for flat pre-trends over a short window would not have detected it.

Our findings also have implications for managers. Where production is individual and clients follow the person, performance belongs largely to the person rather than the firm. It is durable, portable, and concentrated in a few people who can be identified from their track records. A firm that hires these people gains what they were already producing, and a firm that loses them loses it. Managers in these settings should therefore focus on attracting and retaining productive people rather than relying on developing them after hiring, and should evaluate candidates on their long-run record rather than recent performance, since people tend to move after a downturn and much of what appears to be post-hire improvement is recovery. Development can matter a great deal, but the capability to make it matter appears scarce, and understanding what the few offices that improve their agents do differently is a promising direction for future research.

Our study has several limitations. First, the office closure estimate rests on 249 closures, and displaced agents worked at offices that failed, so a larger sample of closures would strengthen the causal interpretation. Second, office quality is measured as the average performance of an office's other agents, so it contains sampling error, and the return to office quality is therefore a range of 0.15 to 0.30 SD across constructions rather than a point estimate (Appendix D). We also find that an office does not add the same increment to every agent. Office quality is worth roughly twice as much to a weaker agent as to a stronger one (Appendix C.2), so the average office share describes the typical agent rather than every agent. Third, a member identifier denotes a license rather than a person, so a single identifier may cover a team working under its lead agent or an institutional seller transacting through one agent of record. Removing every identifier that ever closes more than one

sale per calendar day (38 identifiers holding 2.4% of transactions) moves the individual share from 43.6% to 43.5% (Appendix E.3). Fourth, we observe revenue rather than costs, profitability, or effort, and we observe which agent a transaction is credited to rather than how it was produced. Co-production within an office would therefore appear as individual output, although the within-agent mobility results limit how much of the individual component such crediting could explain.

In considering the generalizability of our findings, residential brokerage has the features under which our theory argued an individual component should be most visible, so the individual share we estimated applies to contexts that share these features. However, three points support the potential broader relevance of the results. First, these features did not make the office effect small, since our conventional decomposition still attributed 39.7% of performance variance to the office. Second, the results carrying the most theoretical weight are not those a setting favorable to individuals would produce on its own. A few offices generate a production component, and the timing result reduces rather than inflates the individual's share. Third, individual attribution, output-based pay, client relationships that follow the person, and low switching costs describe much of professional services and sales, not real estate alone. Our theory predicts a smaller individual share where production is joint, technology is firm-specific, clients attach to the organization, mobility is constrained, or establishments are large, and testing these predictions across settings is an important next step.

Overall, in a setting where individual output can be observed directly, most of what four decades of variance decomposition would have recorded as a firm effect reflects the people the firm employs. Organizations are not irrelevant—they add to what individuals bring, and a small number of them make their people better. But composition dominates, which implies that in many important settings competition determining organizational performance is largely a competition for people.

AI Use Statement. *The authors used Anthropic's Claude (Claude Opus, accessed via Claude Code, 2026) for assistance with constructing and debugging the data pipeline, running statistical diagnostics and robustness checks, formatting figures, and improving the clarity of written text. All AI-generated content was reviewed, verified for accuracy, and edited by the authors. The authors take full responsibility for the accuracy and integrity of all material in the manuscript.*

REFERENCES

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2), 251–333.

Aguinis, H., & O'Boyle, E. (2014). Star performers in twenty-first century organizations. *Personnel Psychology*, 67(2), 313–350.

Alchian, A. A., & Demsetz, H. (1972). Production, information costs, and economic organization. *American Economic Review*, 62(5), 777–795.

Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, 45(7), 905–917.

Andrews, M. J., Gill, L., Schank, T., & Upward, R. (2008). High wage workers and low wage firms: Negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society Series A: Statistics in Society*, 171(3), 673–697.

Andrews, M. J., Gill, L., Schank, T., & Upward, R. (2012). High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias. *Economics Letters*, 117(3), 824–827.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *Review of Economics and Statistics*, 60(1), 47–57.

Barney, J. B., & Felin, T. (2013). What are microfoundations? *Academy of Management Perspectives*, 27(2), 138–155.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.

Barney, J. B. (1986). Strategic factor markets: Expectations, luck, and business strategy. *Management Science*, 32(10), 1231–1241.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. Columbia University Press.

Bhaskarabhatla, A., Cabral, L., Hegde, D., & Peeters, T. (2021). Are inventors or firms the engines of innovation? *Management Science*, 67(6), 3899–3920.

Bidwell, M. (2011). Paying more to get less: The effects of external hiring versus internal mobility. *Administrative Science Quarterly*, 56(3), 369–407.

Bonhomme, S., Lamadon, T., & Manresa, E. (2026). A users' guide to uncovering worker and firm effects: The ABC of AKM. *Journal of Economic Perspectives*, 40(2), 195–214.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285.

Bowman, E. H., & Helfat, C. E. (2001). Does corporate strategy matter? *Strategic Management Journal*, 22(1), 1–23.
Broschak, J. P. (2004). Managers' mobility and market interface: The effect of managers' career mobility on the dissolution of market ties. *Administrative Science Quarterly*, 49(4), 608–640.
Call, M. L., Nyberg, A. J., & Thatcher, S. M. B. (2015). Stargazing: An integrative conceptual review, theoretical reconciliation, and extension for star employee research. *Journal of Applied Psychology*, 100(3), 623–640.
Campbell, B. A., Coff, R., & Kryscynski, D. (2012). Rethinking sustained competitive advantage from human capital. *Academy of Management Review*, 37(3), 376–395.
Campbell, B. A., Ganco, M., Franco, A. M., & Agarwal, R. (2012). Who leaves, where to, and why worry? Employee mobility, entrepreneurship and effects on source firm performance. *Strategic Management Journal*, 33(1), 65–87.
Card, D., Cardoso, A. R., & Kline, P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *Quarterly Journal of Economics*, 131(2), 633–686.
Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *Quarterly Journal of Economics*, 128(3), 967–1015.
Chadwick, C. (2017). Toward a more comprehensive model of firms' human capital rents. *Academy of Management Review*, 42(3), 499–519.
Coff, R., & Kryscynski, D. (2011). Drilling for micro-foundations of human capital-based competitive advantages. *Journal of Management*, 37(5), 1429–1443.
Coff, R. W. (1997). Human assets and management dilemmas: Coping with hazards on the road to resource-based theory. *Academy of Management Review*, 22(2), 374–402.
Coff, R. W. (1999). When competitive advantage doesn't lead to performance: The resource-based view and stakeholder bargaining power. *Organization Science*, 10(2), 119–133.
Crook, T. R., Todd, S. Y., Combs, J. G., Woehr, D. J., & Ketchen, D. J., Jr. (2011). Does human capital matter? A meta-analysis of the relationship between human capital and firm performance. *Journal of Applied Psychology*, 96(3), 443–456.
Denrell, J. (2004). Random walks and sustained competitive advantage. *Management Science*, 50(7), 922–934.
DiMaggio, P., & Louch, H. (1998). Socially embedded consumer transactions: For what kinds of purchases do people most often use networks? *American Sociological Review*, 63(5), 619–637.
Felin, T., & Foss, N. J. (2005). Strategic organization: A field in search of micro-foundations. *Strategic Organization*, 3(4), 441–455.
Felin, T., & Hesterly, W. S. (2007). The knowledge-based view, nested heterogeneity, and new value creation: Philosophical considerations on the locus of knowledge. *Academy of Management Review*, 32(1), 195–218.
Felin, T., Foss, N. J., & Ployhart, R. E. (2015). The microfoundations movement in strategy and organization theory. *Academy of Management Annals*, 9(1), 575–632.
Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17(S2), 109–122.
Groysberg, B., & Lee, L. E. (2009). Hiring stars and their colleagues: Exploration and exploitation in professional service firms. *Organization Science*, 20(4), 740–758.
Groysberg, B., Lee, L. E., & Nanda, A. (2008). Can they take it with them? The portability of star knowledge workers' performance. *Management Science*, 54(7), 1213–1230.
Gubler, T. (2019). Connected, but qualified? Social affiliations, human capital, and service professional performance. *Organization Science*, 30(5), 912–936.

Gubler, T., & Cooper, R. (2019). Socially advantaged? How social affiliations influence access to valuable service professional transactions. *Strategic Management Journal*, 40(13), 2287–2314.
Heckman, J. J., & Smith, J. A. (1999). The pre-programme earnings dip and the determinants of participation in a social programme: Implications for simple programme evaluation strategies. *Economic Journal*, 109(457), 313–348.
Holmström, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, 13(2), 324–340.
Huckman, R. S., & Pisano, G. P. (2006). The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science*, 52(4), 473–488.
Kline, P., Saggio, R., & Sølvsten, M. (2020). Leave-out estimation of variance components. *Econometrica*, 88(5), 1859–1898.
Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, 3(3), 383–397.
Kryscynski, D., Coff, R., & Campbell, B. (2021). Charting a path between firm-specific incentives and human capital-based competitive advantage. *Strategic Management Journal*, 42(2), 386–412.
Kumar, P., Liu, X., & Zaheer, A. (2022). How much does the firm's alliance network matter? *Strategic Management Journal*, 43(8), 1433–1468.
Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, 90(5), 1346–1361.
Lippman, S. A., & Rumelt, R. P. (1982). Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *Bell Journal of Economics*, 13(2), 418–438.
Mackey, A. (2008). The effect of CEOs on firm performance. *Strategic Management Journal*, 29(12), 1357–1367.
Makino, S., Isobe, T., & Chan, C. M. (2004). Does country matter? *Strategic Management Journal*, 25(10), 1027–1043.
Marx, M., Strumsky, D., & Fleming, L. (2009). Mobility, skills, and the Michigan non-compete experiment. *Management Science*, 55(6), 875–889.
McGahan, A. M., & Porter, M. E. (1997). How much does industry matter, really? *Strategic Management Journal*, 18(S1), 15–30.
Milgrom, P., & Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review*, 80(3), 511–528.
Mollick, E. (2012). People and process, suits and innovators: The role of individuals in firm performance. *Strategic Management Journal*, 33(9), 1001–1015.
National Association of Realtors. (2023). *Profile of home buyers and sellers*. National Association of Realtors.
Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Harvard University Press.
Nyberg, A. J., Moliterno, T. P., Hale, D., Jr., & Lepak, D. P. (2014). Resource-based perspectives on unit-level human capital: A review and integration. *Journal of Management*, 40(1), 316–346.
O'Boyle, E., Jr., & Aguinis, H. (2012). The best and the rest: Revisiting the norm of normality of individual performance. *Personnel Psychology*, 65(1), 79–119.
Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*, 14(3), 179–191.
Ployhart, R. E., & Moliterno, T. P. (2011). Emergence of the human capital resource: A multilevel model. *Academy of Management Review*, 36(1), 127–150.
Quigley, T. J., & Graffin, S. D. (2017). Reaffirming the CEO effect is significant and much larger than chance: A comment on Fitza (2014). *Strategic Management Journal*, 38(3), 793–801.
Ray, C., Essman, S., Nyberg, A. J., Ployhart, R. E., & Hale, D. (2023). Human capital resources: Reviewing the first decade and establishing a foundation for future research. *Journal of Management*, 49(1), 280–324.

Rumelt, R. P. (1991). How much does industry matter? *Strategic Management Journal*, 12(3), 167–185.
Rumelt, R. P., Schendel, D. E., & Teece, D. J. (Eds.). (1994). *Fundamental issues in strategy: A research agenda*. Harvard Business School Press.
Schmalensee, R. (1985). Do markets differ much? *American Economic Review*, 75(3), 341–351.
Sharapov, D., Kattuman, P., Rodriguez, D., & Velazquez, F. J. (2021). Using the SHAPLEY value approach to variance decomposition in strategy research: Diversification, internationalization, and corporate group effects on affiliate profitability. *Strategic Management Journal*, 42(3), 608–623.
Short, J. C., Ketchen, D. J., Jr., Palmer, T. B., & Hult, G. T. M. (2007). Firm, strategic group, and industry influences on performance. *Strategic Management Journal*, 28(2), 147–167.
Somaya, D., Williamson, I. O., & Lorinkova, N. (2008). Gone but not lost: The different performance impacts of employee mobility between cooperators versus competitors. *Academy of Management Journal*, 51(5), 936–953.
Song, J., Almeida, P., & Wu, G. (2003). Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Science*, 49(4), 351–365.
Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, 15(6), 285–305.
Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171–180.
Wezel, F. C., Cattani, G., & Pennings, J. M. (2006). Competitive implications of interfirm mobility. *Organization Science*, 17(6), 691–709.
Wright, P. M., Coff, R., & Moliterno, T. P. (2014). Strategic human capital: Crossing the great divide. *Journal of Management*, 40(2), 353–370.

FIGURES AND TABLES

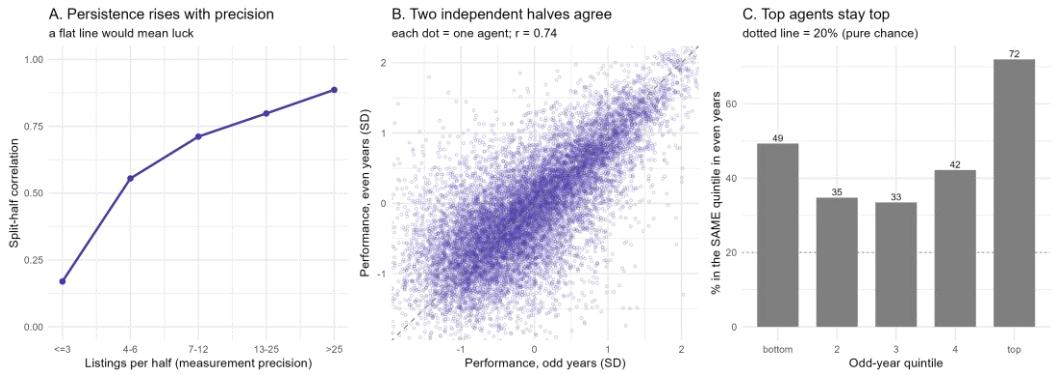


Figure 1. The agent effect is durable. Market-adjusted performance estimated separately on odd- and even-numbered years for the 52,525 agents observed in both. (A) Split-half correlation by number of transactions per half. (B) Odd-year versus even-year estimates. (C) Even-year quintile of each odd-year quintile.

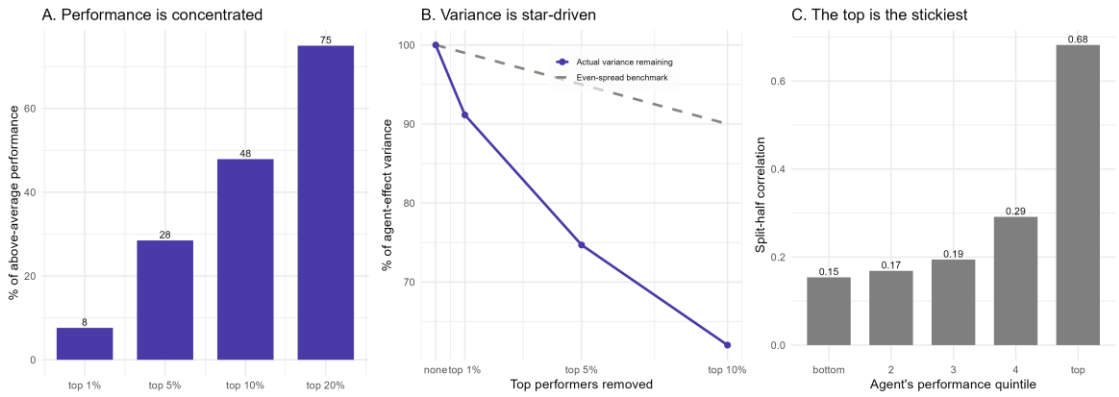


Figure 2. Performance is concentrated in a stable group of top performers. Agents observed at least three years. (A) Share of total above-average performance accounted for by the top k% of agents. (B) Between-agent variance remaining as top performers are removed, against the benchmark of evenly spread performance. (C) Split-half correlation within each performance quintile.

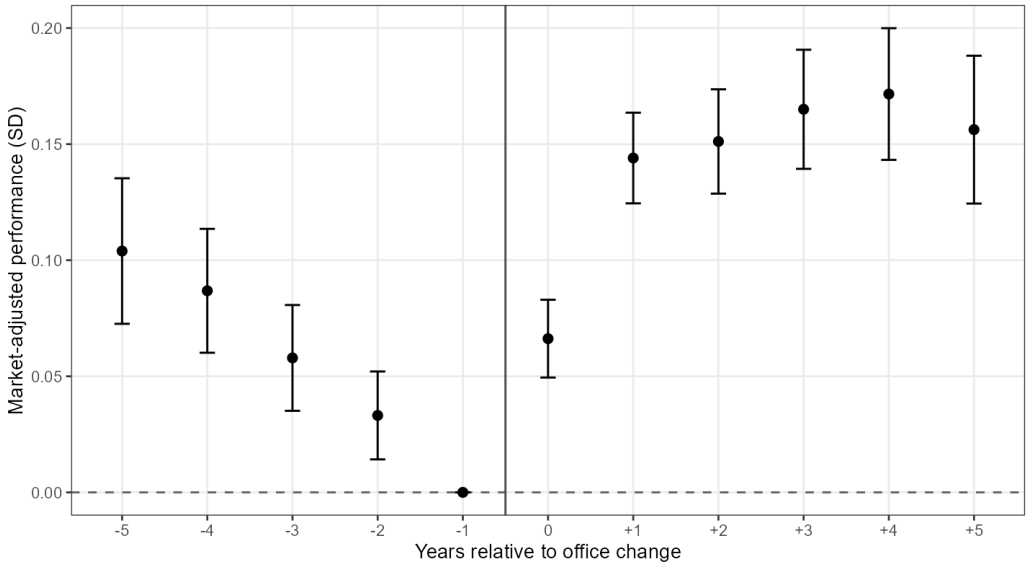


Figure 3. Performance around a voluntary office change. Event-study coefficients from equation (2) with 95% confidence intervals. The dependent variable is market-adjusted performance in SD units, the year before the move is the omitted baseline, agents who never move are the controls, and standard errors are clustered by agent.

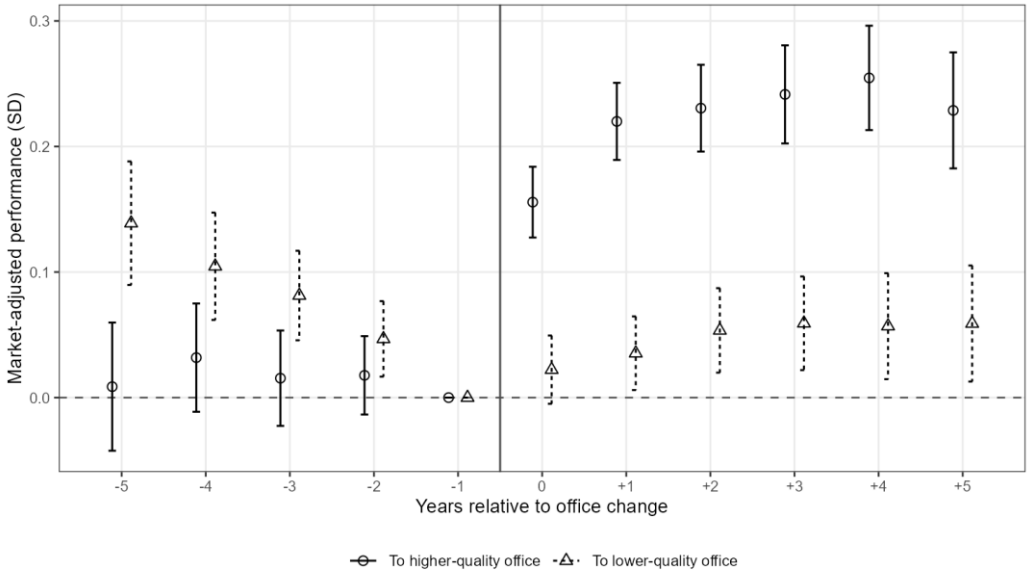


Figure 4. Performance around a move, by direction of office-quality change. Event-study coefficients with 95% confidence intervals for the 4,765 agents who moved to a higher-quality office (circles) and the 4,830 who moved to a lower-quality office (triangles). Office quality is the average market-adjusted performance of the office's other agents, excluding the mover. Specification as in Figure 3.

Table 1. Descriptive statistics.

Variable	N	Mean (SD)	Min	P25	Median	P75	Max
Panel A. Transactions							
Sale price ($000)	1,898,013	366.4 (429.3)	10.0	183.9	280.0	421.6	39,600.0
Days on market	1,891,506	66.9 (97.5)	0.0	11.0	36.0	82.0	3,554.0
Sale price / original list price	1,879,476	0.962 (0.086)	0.500	0.933	0.979	1.000	1.500
Buyer agent recorded (share)	1,898,013	0.902 (0.297)	0.000	1.000	1.000	1.000	1.000
Panel B. Agent-years (the unit of analysis)							
Transactions closed per year	571,295	6.06 (14.02)	1.00	1.00	3.00	7.00	2,126.00
Dollar volume per year ($000)	571,295	2,181.6 (5,433.8)	10.0	370.0	875.0	2,268.5	1,002,415
Average price per transaction ($000)	571,295	367.2 (354.0)	10.0	199.9	290.0	429.0	22,610.0
Months active during the year	571,295	3.54 (2.88)	1.00	1.00	2.00	5.00	12.00
Colleagues in the same office	571,295	95.59 (164.13)	1.00	7.00	35.00	111.00	1,265.00
Market-adjusted performance (SD)	434,881	0.00 (1.00)	-4.99	-0.71	-0.00	0.69	5.79
Panel C. Agents							
Years observed	113,670	5.03 (4.55)	1.00	1.00	3.00	7.00	17.00
Transactions over career	113,670	30.5 (84.5)	1.0	2.0	7.0	26.0	6,467.0
Dollar volume over career ($000)	113,670	10,964.6 (31,343.7)	10.0	630.0	2,254.0	9,121.4	1,743,519
Offices worked in	113,670	1.71 (1.09)	1.00	1.00	1.00	2.00	11.00
Ever changed office (share)	113,670	0.411 (0.492)	0.000	0.000	0.000	1.000	1.000
Observed five or more years (share)	113,670	0.398 (0.489)	0.000	0.000	0.000	1.000	1.000
Panel D. Offices and office-years							
Agents per office-year	84,308	6.78 (24.53)	1.00	1.00	2.00	4.00	1,265.00
Transactions per office-year	84,308	41.1 (154.9)	1.0	2.0	6.0	23.0	7,112.0
Years observed	15,368	5.49 (4.91)	1.00	1.00	4.00	8.00	17.00
Franchise-affiliated (share)	15,368	0.121 (0.326)	0.000	0.000	0.000	0.000	1.000

Notes. Panel A is the main analysis sample of 1,898,013 sold transactions, Panel B the agent-year panel used in all production and mobility analyses, and Panels C and D aggregate that panel to the agent and the office. Monetary variables are in thousands of dollars. The listing-price ratio excludes the 1.0% of listings outside 0.5 to 1.5, and days on market is missing for 0.3%. Market-adjusted performance is log annual volume net of ZIP-by-year conditions and tenure, standardized to unit variance, for agents observed at least five years (434,881 agent-years). Maxima in Panels B and C reflect institutional accounts rather than individual producers (Appendix E.3).

Table 2. Agent-year variance decomposition of annual dollar volume and its two components.

Sample or outcome	Agent	Office	Franchise	Residual
Panel A. Market-adjusted log annual dollar volume, by minimum tenure				
All agents (1+ years)	47.0	8.2		44.7
Agents with 3+ years	44.7	8.6		46.8
Agents with 5+ years (main)	43.6	9.0		47.4
Franchise subsample (5+ years)	48.7	4.5	1.2	45.6
Panel B. The two components of annual volume (agents with 5+ years)				
Number of transactions closed (log count)	46.0	9.2		44.8
Average price per transaction (log price)	25.8	7.0		67.2

Notes. Entries are Shapley shares (percent) of the variance of each row's outcome after removing ZIP-by-year conditions and agent tenure; rows sum to 100 with the residual. Panel A decomposes log annual dollar volume at three minimum-tenure cutoffs, and its final row adds the franchise level among franchise-affiliated offices. Panel B decomposes the two components of log volume separately, since log volume equals log transactions plus log average price. Its rows are shares of different variances (0.91 for count, 0.15 for price) and do not partition the Panel A shares. Panel B uses agents observed at least five years (434,881 agent-years).

Table 3. From the conventional firm effect to the corrected components.

Step	What it removes	Agent (%)	Office (%)
(1) One level at a time, uncorrected	nothing	54.6	39.7
(2) One level at a time, corrected	estimation noise	48.4	32.4
(3) Both levels jointly, uncorrected	composition	52.7	9.9
(4) Both levels jointly, corrected	composition and noise	44.4	6.4

Notes. Entries are percentages of the total variance of market-adjusted log annual volume among agents observed at least five years. Rows 1 and 2 estimate one level at a time, so each level is credited with whatever the other contributes; row 2 applies the sample-splitting correction, the covariance of effects estimated on two halves of each career that share no years (299,487 agent-years; 32,204 agents). Rows 3 and 4 estimate agent and office effects jointly, which removes composition; row 4 applies the leave-out correction of Kline et al. (2020) on the largest leave-out-connected set (154,434 agent-years; 17,109 agents; 2,880 offices). The agent-office covariance is -4.3% uncorrected and -1.8% corrected, and implied agent reliability is .89. The Shapley shares in Table 2 divide jointly explained variance between levels and are not comparable to these rows (Appendix C.4).

Table 4. Forced versus voluntary movers (dependent variable: market-adjusted performance, SD).

Years relative to move	Forced movers	Voluntary movers
-5	0.001 (0.033)	0.105 (0.017)
-4	-0.026 (0.031)	0.085 (0.014)
-3	0.027 (0.025)	0.057 (0.012)
-2	-0.019 (0.021)	0.034 (0.010)
0 (move or closure year)	0.093 (0.019)	0.069 (0.009)
+1	0.038 (0.023)	0.152 (0.010)
+2	0.040 (0.025)	0.159 (0.012)
+3	0.058 (0.027)	0.171 (0.013)
+4	0.058 (0.029)	0.177 (0.015)
+5	0.066 (0.032)	0.161 (0.017)
Pooled post-move effect (t+1 to t+3)	0.045 (0.022)	0.161 (0.010)
Post jointly zero, p	0.177	$p < .001$
Equality of forced and voluntary post effects, joint p	$p < .001$	

Notes. Event-study coefficients from a single stacked specification of equation (2) with agent and year fixed effects; agents who never move are the controls and the year before the move (t-1) is the omitted baseline. Standard errors clustered by agent in parentheses. Forced movers are the 2,399 agents present in the final active year of one of 249 screened office closures (Appendix D.2); voluntary movers are agents who changed office exactly once. Years beyond five before or after the move are absorbed in the endpoint bins. The pooled effect averages the t+1 to t+3 coefficients.

APPENDIX

APPENDIX A: THE INDIVIDUAL AT THE TRANSACTION LEVEL

The body of the paper measures production at the agent-year level. A separate question is whether the individual affects the product the seller receives, meaning the price a given house fetches. This is a question about product quality rather than production, and it indicates whether the individual's imprint reaches down to the single transaction or is confined to deal generation. Because the analysis requires the physical characteristics of the house, it is the one result in the paper estimated on the hedonic subsample rather than the full transaction population.

We work at the level of the individual sale and first remove house and market characteristics from price. We residualize log sale price on a hedonic specification containing six continuous property characteristics in logs (lot size, bedrooms, bathrooms, parking, homeowners association fee, and living area), seventeen binary indicators for features such as new construction, garage, pool, view, waterfront, and utility connections, and fixed effects for year, month, year built, and property subtype. Together these absorb 74.8% of the variance in log sale price, and the residual measures whether, given this house at this time, the sale over- or under-performed. Table A1 describes the underlying transactions.

Table A1. Listing-level descriptives, hedonic subsample.

Variable	Mean	SD	P25	Median	P75
Sale price (USD)	337,655	355,923	178,450	268,000	400,000
Days on market	68.6	108.8	7.0	31.0	84.0
Bedrooms	3.57	1.17	3.00	3.00	4.00
Bathrooms	2.28	0.82	1.75	2.00	2.75
Living area (sq. ft.)	2,305	1,211	1,498	2,033	2,834
Lot size (acres)	0.73	45.06	0.11	0.19	0.28
Year built	1988	26	1976	1997	2006

Notes. Hedonic subsample, N = 775,376 sold listings. Binary characteristics (share of listings): new construction 6.7%, homeowners association 27.2%, cooling 84.2%, heating 88.7%, fireplace 42.8%, garage 70.6%. The lot-size mean is inflated by a small number of extreme records; the median is representative.

Decomposing the price residual across ZIP, agent, office, and franchise (Table A2), the individual does leave a mark on price: the agent share is 16.3%, ahead of the office's 10.5%, and stable as the minimum-listings bar rises, at 16.3%, 15.6%, and 15.1% for thresholds of 5, 15, and 30 listings. The ZIP market is the largest single component at 21.3%, unsurprising when the outcome is price. More important for the argument, the individual's price effect of 16.3% is much smaller than the individual's production effect of 43.6%. Individuals differ mostly in deal flow and only secondarily in the price extracted from a given home, which is why the main decomposition reports a larger individual share at the annual level than appears here.

Table A2. Variance decomposition of the hedonic-adjusted log sale price (Shapley shares, %).

Sample	ZIP	Agent	Office	Franchise	Residual
Full (N = 747,805)	21.3	16.3	10.5		51.9
Franchise subsample (N = 372,272)	21.0	16.1	8.9	1.9	52.1

Notes. Shapley shares (percent) of the variance of the hedonic-adjusted log sale price, for agents with at least five listings. The hedonic first stage absorbs 74.8% of the variance of log sale price. The agent share is 16.3, 15.6 and 15.1 at minimum-listing thresholds of 5, 15 and 30.

Two qualifications apply to this analysis specifically. The hedonic subsample over-represents Utah and better-documented listings, so a state-stratified replication and a representativeness analysis of the excluded listings are natural extensions. And the price residual is a noisier object than

annual production, since it contains everything about a single sale that the hedonic specification does not capture.

APPENDIX B: THE DEAL-FLOW ADVANTAGE AND THE SCOPE OF DURABILITY

B.1 Ruling Out Market Assignment and Specialization

The main decomposition reports that individuals differ far more in how many deals they close than in the prices they achieve. The concern this raises is that deal count might reflect where an agent works rather than who the agent is: an agent in a busy market, or one who has found an easy niche, will close more transactions without being more skilled. Neither explanation accounts for much.

Market conditions explain a modest share. Our performance measure has already netted ZIP-by-year conditions out of the count effect, so the market cannot reenter through the residual. To size its role in context, we reintroduce it and ask how much of the total unadjusted variation in agents' annual transaction counts the market accounts for. The answer is 15.2% (Figure B1, Panel A). The large majority of count variation arises within markets, between agents selling in the same places at the same times.

Observable specialization explains less still. For each agent we build concentration indices over property types, price tiers, and ZIP codes, together with the mean and dispersion of listed home sizes and the agent's tenure, and we regress the market-adjusted count effect on these observables, entering them in blocks. Tenure alone explains 2.6%, and the cumulative share rises to at most 8.6% as the specialization measures are added, with the full model explaining 8.2% (Figure B1, Panel B). Each measure contributes little individually, and they remain a small share when combined. About 91% of the variation in the count effect is irreducible to either the market or any observable strategy.

High-count agents do behave visibly differently, and the pattern is monotonic across every measure. Moving from the lowest to the highest quintile of the count effect, average days on market falls from 71.1 to 54.6, photographs per listing rise from 11.4 to 18.2, the ratio of sale price to

original list price rises from 0.954 to 0.974, and the share of listings receiving a price reduction falls from 0.687 to 0.562 (Figure B1, Panel C). We read these behaviors as the visible symptoms of skill rather than its cause, since a skilled agent both prices accurately and closes more.

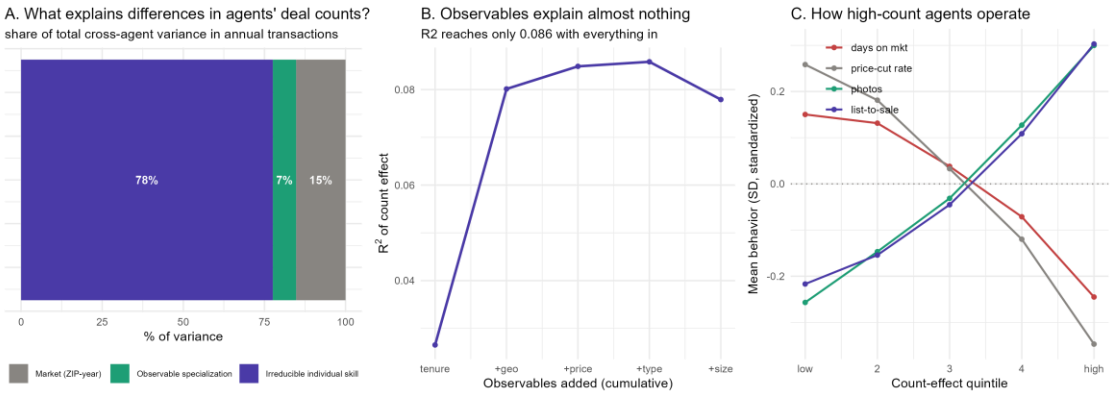


Figure B1. The deal-flow advantage is individual skill. (A) Of the total cross-agent variance in annual deal counts, the market explains 15.2% and observable specialization a smaller share, leaving the large remainder as individual skill. (B) Cumulative share of the agent count effect explained as specialization observables are added, which never exceeds 9%. (C) Standardized operating behavior across count-effect quintiles.

B.2 Whom the Durability Result Describes

The split-half test of the durability analysis can only examine agents observed in both halves of a career, which means it conditions on survival. The concern is that the persistence we measure might be a property of survivors rather than of agents.

The condition turns out to be narrow in the way that matters. The durability sample contains 79.4% of agents but 98.3% of all listings and 98.1% of all dollar volume (Figure B2, Panel A). The agents it excludes are overwhelmingly one-year entrants, with a mean tenure of 1.14 years and a mean of 3.6 transactions, who are also lower performers while active, at negative 0.37 SD against negative 0.05 SD for the retained group (Figure B2, Panel B). The survival condition drops many people but almost none of the industry's activity.

Panel C provides the diagnostic that distinguishes the two explanations. If survivorship were inflating the durability estimate, then demanding still more survival should inflate it further, and sharply. Instead, raising the minimum-tenure bar raises the correlation only gently, along the same noise-reduction gradient documented in Figure 1: from .816 at four or more years to .857 at six, .876 at eight, and .888 at ten or more. Durability describes agents who build an actual career in the industry, and that population does essentially all of the industry's business.

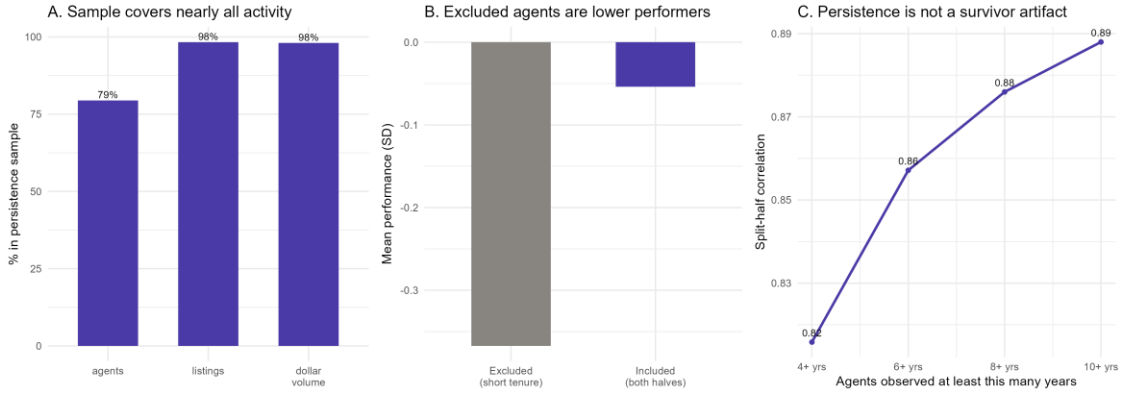


Figure B2. Scope of the durability result. (A) The durability sample is 79% of agents but over 98% of activity. (B) Excluded, short-tenure agents are lower performers while active. (C) Durability rises only gently as the minimum-tenure bar increases.

APPENDIX C: DIAGNOSTICS FOR THE CORRECTED DECOMPOSITION

The results section summarizes three diagnostics for the assumptions underlying the corrected decomposition and reports their results. This appendix gives their construction. The diagnostics follow the practice recommended by Bonhomme et al. (2026), whose review of this estimator class identifies limited mobility bias, additivity, and exogenous mobility as the three assumptions that require testing rather than assertion.

C.1 Limited Mobility Bias: The Sub-Sampling Diagnostic

The identity behind the correction is that the variance measured across estimated effects equals the true variance plus the average squared standard error of those effects. For office effects, the relevant standard error is governed by the number of agents who moved through each office, since only a

mover holds ability constant while the organization changes, and an office's effect is estimated from the differences those movers exhibit. Each such difference contains the mover's own idiosyncratic good or bad year, which the estimator must charge to the office. The inflation in the plug-in office share should therefore scale as the inverse of movers per office.

We test this prediction by manufacturing the problem (Andrews et al., 2012). We draw random sub-samples of agents at shares of 10%, 20%, 30%, 50%, 75%, and 100%, taking two independent draws at each share below the full sample. Within each draw we rebuild the leave-out connected set from scratch and re-estimate both the plug-in and the corrected components, so that every estimate faces the connectivity its own sub-sample provides. We report results against average movers per office, which is the quantity that governs the bias, rather than against the sub-sample share.

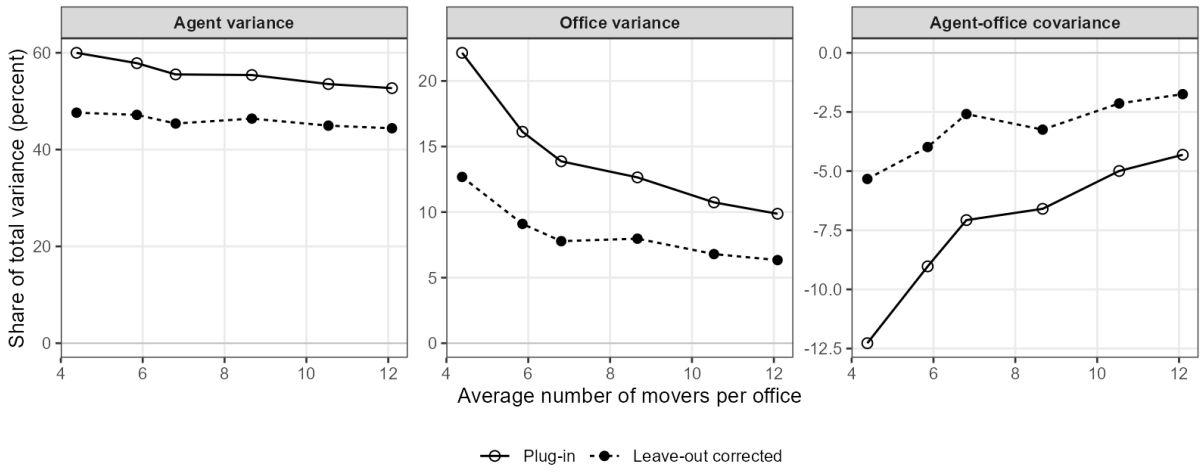


Figure C1. Limited mobility bias and the sub-sampling diagnostic. Variance shares against the average number of movers per office, across random sub-samples of agents. Open circles: plug-in estimates. Filled circles: leave-out corrected estimates. The plug-in office share more than halves as the mobility network thickens, and the corrected share falls by about half as much.

Table C1 reports the results. The plug-in office share falls from 22.1% at 4.4 movers per office to 9.9% at 12.1, slightly more than halving, and the plug-in agent share falls from 60.0% to 52.7%. The corrected estimates are steadier but not flat: the agent share moves from 47.6% to 44.4% and the office share from 12.7% to 6.3%, so the correction roughly halves the sensitivity rather than

removing it. The full-sample row of Table C1 is a separate draw of the randomized leave-out approximation, which is why it reads 6.3 where Table 3 reads 6.4: the two are the same estimator on the same connected set and differ only by Monte Carlo error in the projection. The corrected covariance moves from negative 5.3% to negative 1.7% while its plug-in counterpart moves from negative 12.3% to negative 4.3%.

Table C1. Sub-sampling diagnostic for limited mobility bias.

Movers per office	Offices	Agent, plug-in	Agent, corrected	Office, plug-in	Office, corrected	Covariance, plug-in
4.4	702	60.0	47.6	22.1	12.7	-12.3
5.9	1,116	57.9	47.2	16.1	9.1	-9.0
6.8	1,464	55.5	45.4	13.9	7.8	-7.1
8.7	1,954	55.4	46.4	12.6	8.0	-6.6
10.5	2,457	53.5	44.9	10.7	6.8	-5.0
12.1	2,880	52.7	44.4	9.9	6.3	-4.3

Notes. Each row averages two independent random sub-samples of agents (one draw at the full sample), drawn at 10, 20, 30, 50, 75 and 100 percent. The leave-out connected set is rebuilt within each draw before estimation. Entries are percentages of the total variance of market-adjusted log annual volume. Movers per office is the average number of agents observed at more than one office, per office, and is the quantity that governs limited mobility bias. The corrected covariance runs from -5.3 to -1.7 across the same rows and is omitted from the table for width; it is plotted in Figure C1. The plug-in office share falls by slightly more than half across the range while the corrected office share falls by about half as much, and the corrected agent share varies between 44.4 and 47.6.

The inflation obeys the predicted scaling. Regressing the log of the plug-in minus corrected gap on the log of movers per office gives a slope of negative 0.93 (SE 0.03, R-squared .99) for the office component and -0.92 (SE 0.04, R-squared .99) for the covariance, against a theoretical value of negative 1. Extrapolating the inflation observed in the thinnest sub-sample by the inverse rule predicts 3.4 percentage points at the full sample against 3.5 observed. The agent component scales more gently, at negative 0.37, which is expected: agent effects are identified from all of an agent's years rather than from mobility, so sub-sampling agents does not shorten careers, and the decline that does appear is inherited from the office side through the shared design.

Two qualifications apply. The corrected office share is not flat, so the correction halves the sensitivity to mobility density rather than eliminating it, and the estimate we report is the one

computed on the fullest network, which is the least affected. And the diagnostic speaks only to sensitivity to mobility density. It does not establish that the corrected estimates are unbiased on some other dimension.

C.2 Additivity

Additivity requires that the office contributes the same increment to every agent. Without it, an average office share is a weighted average across agents rather than an amount every agent receives, because the office's contribution depends on who is in it. Card et al. (2016) reject additivity in wage data by showing that firm effects differ between men and women, so the restriction cannot be taken for granted. We test it two ways. The quartile event study reported first compares moves between the extremes of the office quality distribution and is included for completeness. The direct test on movers reported second is the one the paper relies on, because it is far better powered and does not require comparing small cells without a control group.

The quartile event study follows Card et al. (2013). We rank offices into quartiles by leave-out quality, defined as the mean market-adjusted performance of the office's other agents, so that an office's rank is never set by the mover whose move is being studied. We keep agents who move exactly once and are observed for at least two years on each side, and trace mean performance at event time negative two through positive two for each origin-by-destination quartile cell. Additivity implies symmetry. A move from the bottom quartile to the top should gain what the reverse move loses. Among 10,143 movers with both office qualities estimable, of whom 3,643 have a balanced window, the upward corner move gains 0.241 SD and the downward corner move loses 0.181 SD, summing to positive 0.060 SD, roughly a quarter of the upward move and a third of the downward one. Two cautions attach to the comparison, and they are why we treat it as corroboration rather than as the primary test. It isolates the largest available change in office quality, a rise or fall of three quartiles, and the corner cells are small, with 116 movers going up and 138 going down. And it is a

descriptive profile with no control group, so nothing common to all agents over the period is netted out. Neither caution bears on the symmetry itself, which is the quantity additivity restricts, but the downward figure should not be read as what happens to agents who move to worse offices in general. The direction split, which uses the full sample of downward movers with agents who never move as controls, finds no decline for them at all, and in fact a small gain.

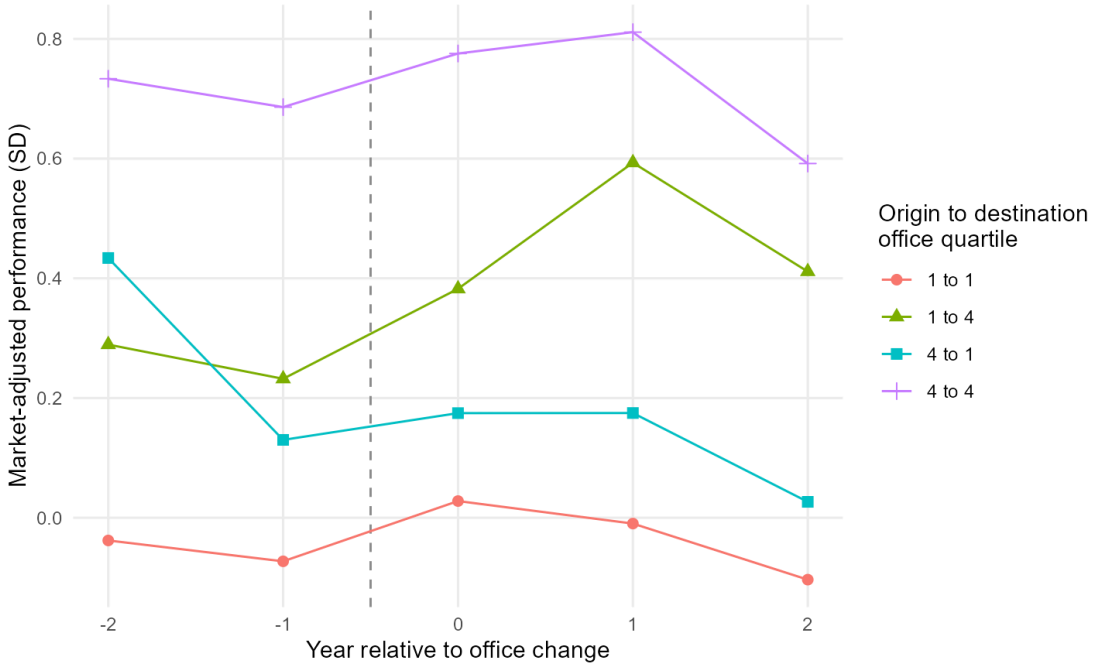


Figure C2. Mover event study by origin and destination office quartile. Mean market-adjusted performance around an office change, for the four corner transitions. Symmetry between the 1-to-4 and 4-to-1 paths is the additivity test; the pre-move path is the exogenous-mobility test.

One feature of this test is worth reporting. The cells in which agents neither improve nor worsen their office are close to flat, drifting by negative 0.001 SD for bottom-quartile stayers and negative 0.008 SD for top-quartile stayers, so there is no common career-stage drift for the comparison to net out and the positive 0.060 SD asymmetry stands as measured. Gains modestly exceed losses, which is the mild signature of selection on match quality. The symmetry itself is not overturned at these magnitudes, but the direct test below, which is far better powered, does reject the restriction.

The direct test asks the same question of movers, which avoids estimating office effects twice and is far better powered. We regress each mover's change in performance, the mean of years positive one and positive two minus the mean of years negative two and negative one, on the change in leave-out office quality, interacted with the agent's own type. Additivity implies the interaction is zero. The design requires one precaution. Agent type is measured only on years at event time negative three or earlier, which shares no observation with the outcome window. Measuring type on the immediately preceding years would place the same noise on both sides of the equation and manufacture an interaction through mean reversion alone.

Among 2,567 movers meeting these requirements, the return to office quality is 0.396 for low-type agents, and the extra return for high-type agents is negative 0.220 (SE 0.073, p = .003), which leaves 0.176 for high-type agents. Entering type continuously gives an interaction of negative 0.139 (SE 0.038), agreeing in sign. Office quality is therefore worth roughly twice as much to a weak agent as to a strong one, and additivity is rejected. The 95% interval around the binary interaction runs from negative 0.362 to negative 0.078 and around the continuous one from negative 0.213 to negative 0.065, so the data rule out equality and also rule out a return that rises with agent quality. The violation runs toward substitution: a good office does most for the agents who bring the least, and splitting by direction of move gives interactions of negative 0.518 (SE 0.158) for upward moves and negative 0.188 (SE 0.149) for downward ones. The consequence for the paper is that the average office share in Table 2 and Table 3 is a weighted average of returns that differ across agents rather than an increment every agent receives, so it describes the typical agent rather than all of them.

We also attempted the more direct test of estimating office effects separately within agent-type subgroups and comparing them, which is the approach Card et al. take by gender. It has no power in our data, and we report the failure because it is informative about what this design can support.

Splitting the sample halves the movers identifying each office effect and leaves only 209 offices with every required effect estimable, so the subgroup estimates are almost entirely sampling noise, with reliabilities of .327 and .043. Attenuation from measurement error then drives any regression slope toward zero whether or not additivity holds, and correcting the attenuation requires dividing by a signal variance estimated near zero, which produced an implied correlation above three and a bootstrap interval on the slope of negative 3.6 to 3.2. The movers-based test above is the same hypothesis asked in a form the data can answer.

C.3 Exogenous Mobility

Exogenous mobility requires that moves are unrelated to transitory shocks. This is the assumption our data strain, and the voluntary mover analysis reaches the same conclusion.

The quartile event study shows that the average pre-move change is negative 0.077 SD across all cells. The average conceals the pattern that matters. The dip is concentrated in the steepest downward move: agents moving to a bottom-quartile office from the top fell 0.304 SD in the year before the move, against 0.072 SD for the same destination from the third quartile and 0.057 SD for agents moving from the bottom quartile to the top. Almost every cell declines a little, but the steep downward move declines three to six times as much.

The regression version does not confirm it. Regressing each mover's pre-move trend, the change from event time negative two to negative one, on the change in office quality gives a coefficient of 0.035 (SE 0.033, $p = .288$), about a tenth the magnitude of the estimated return to office quality and indistinguishable from zero. On this test the pre-move trend is not systematically graded by how much an agent's office quality changes. What remains is the multi-year dip documented in the voluntary event study, which is common to movers rather than graded by the size of the quality change, together with the concentration of that dip among the steepest downward moves.

We do not regard this violation as resolved by these tests, and the paper handles it by design, through the closure sample, rather than by assumption. What the diagnostic adds is the direction of the resulting bias. Charging an agent's own continuing decline to the destination office makes organizations appear more consequential than they are. The violation therefore inflates the office share and the return to office quality rather than shrinking them. It works against the paper's central claim that the organization adds little, and it means the return to office quality should be read as an upper bound.

C.4 Why the Corrected Components Start From a Larger Raw Variance

Table 2 and Table 3 report the individual component on different bases, and the two should not be read as competing estimates of the same quantity. Table 2 reports Shapley shares, which divide variance that two levels jointly explain evenly between them. The corrections in Table 3 operate instead on the raw between-agent variance, which assigns the jointly explained, sorting-related portion to the agent and is mechanically larger as a result. That is why the corrected sequence begins at 54.6% rather than at the 43.6% Shapley figure. It appears in Table 3 because it is the quantity the corrections start from, and the corrected endpoints, 48.4% and 44.4%, are the numbers we interpret.

APPENDIX D: SPECIFICATION CHECKS FOR THE MOBILITY ANALYSIS

D.1 The Voluntary Event Study

Two features of the panel could distort the voluntary event study. The first is mechanical. An agent's first and last observed calendar years are incomplete, since an agent licensed in July or exiting in March shows an artificially low count for that year, and if movers are disproportionately observed in such years the move effect would be contaminated by an accounting artifact rather than by economics. Our main specification drops each agent's first and last observed year. Table D1 restores them in column 2 and adds explicit entry and exit indicators in column 3. Those indicators carry the large negative coefficients the partial-year problem predicts, negative 0.381 and negative 0.378,

which confirms the problem is real, but the post-move estimates stay positive, precisely estimated and of similar magnitude across all three columns, so the trimming is not what produces the result.

The second is that a single average post-move coefficient can be a misleading summary when different move cohorts have different effects, because conventional two-way fixed effects can weight some cohorts negatively. Column 4 reports the interaction-weighted estimator of Sun and Abraham (2021), which is robust to that problem. It traces a very similar path, at 0.143, 0.140, and 0.143 against 0.144, 0.151, and 0.165 in the main specification.

Table D1. Voluntary mover event studies: specification checks (dependent variable: market-adjusted performance, SD).

Years relative to move	(1) Main	(2) All years kept	(3) All years plus entry and exit controls	(4) Sun and Abraham
-5	0.104 (0.016)	0.024 (0.015)	0.080 (0.014)	0.131
-4	0.087 (0.014)	0.017 (0.012)	0.066 (0.012)	0.109
-3	0.058 (0.012)	0.011 (0.011)	0.055 (0.010)	0.076
-2	0.033 (0.010)	-0.013 (0.009)	0.022 (0.009)	0.042
0 (move year)	0.066 (0.009)	0.072 (0.008)	0.028 (0.008)	0.065
+1	0.144 (0.010)	0.135 (0.009)	0.123 (0.009)	0.143
+2	0.151 (0.011)	0.122 (0.010)	0.124 (0.010)	0.140
+3	0.165 (0.013)	0.107 (0.012)	0.122 (0.012)	0.143
+4	0.172 (0.014)	0.111 (0.013)	0.134 (0.013)	0.148
+5	0.156 (0.016)	0.101 (0.014)	0.131 (0.014)	0.134
First-year indicator			-0.381 (0.005)	
Last-year indicator			-0.378 (0.006)	
Pre-trend joint p (t-5 to t-2)	< 0.001	0.045	< 0.001	
Post joint p (t+1 to t+5)	< 0.001	< 0.001	< 0.001	
Observations	221,457	295,564	295,564	221,457

Notes. Dependent variable is market-adjusted performance in standard deviations. All models include agent and year fixed effects; the omitted period is the year before the move; agents who never move serve as controls. The window is symmetric from t-5 to t+5, with periods outside it absorbed by endpoint bins, so column (1) reproduces Figure 3. Standard errors clustered by agent in parentheses. Column (1) drops each agent's mechanically partial first and last observed calendar year; columns (2) and (3) keep them, and column (3) adds first-year and last-year indicators whose large negative coefficients confirm the partial-year problem. Column (4) reports interaction-weighted Sun and Abraham (2021) estimates aggregated across move cohorts; cohort-level standard errors are not aggregated.

D.2 Forced Versus Voluntary Movers

The forced-mover analysis identifies office closures with four screens. First, the office must have been active for at least three years, so that it has an operating history. Second, it must have had at least three agents in its final year and at least half of its recent peak size, which excludes offices that withered away rather than closing. Third, its displaced agents must have dispersed to at least two destination offices, which excludes rebrandings and acquisitions recorded as closures. Finally, the closure must precede the last fully observed calendar year, since an office whose final activity falls in the incomplete final year of the extract cannot be distinguished from one still operating. Two hundred and forty-nine closures pass all four.

Figure D1 plots the stacked event study reported in Table 4, comparing forced with voluntary movers against agents who never move. The contrast the figure makes visible is the one the design turns on. The forced path begins from a flat pre-move window and rises modestly, while the voluntary path declines into the move year and then rises far more steeply.

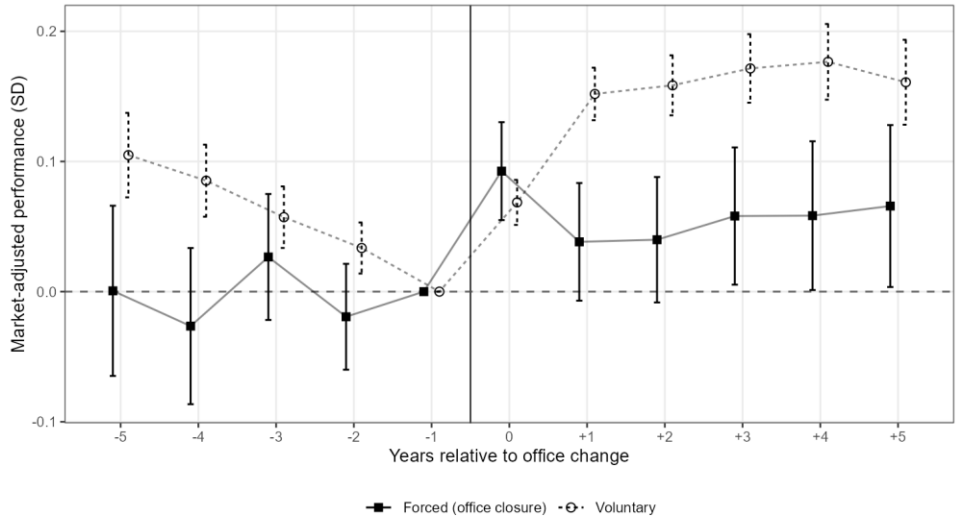


Figure D1. Forced versus voluntary movers. Event-study coefficients with 95% confidence intervals from the stacked specification of Table 4. Squares: forced movers displaced by office closures. Circles: voluntary movers.

D.3 Robustness of the Return to Office Quality

The direction analysis estimates that a mover's post-move gain rises by 0.300 SD for every standard deviation improvement in office quality, and reports a credible range of roughly 0.15 to 0.30 across constructions. Table D2 reports every construction, and this section explains what each addresses.

The first threat is the timing problem documented above: if movers are measured against a depressed baseline year, the estimated change is inflated. Row 5 addresses it. It recomputes each mover's change against a baseline drawn from two years before the move and earlier, excluding the year on which movers are ranked, and adds fixed effects for the mover's pre-move performance decile so that comparisons are made within groups of similarly situated agents. The problem is visible in our data. Measured against the single preceding year, the naive post-move change falls almost mechanically from 1.058 SD for bottom-decile movers to negative 0.698 for top-decile movers. Measured against the reversion-safe baseline, the office-quality slope is 0.276 (SE 0.028), close to the baseline estimate rather than collapsing toward zero. Relatedly, mobility is only mildly selective on prior performance, with the share of movers in each pre-move decile ranging from 7.9% to 11.8% against the 10% that would obtain if moving were unrelated to performance. In the direction split, movers to higher-quality offices show no pre-trend ($p = .534$) while movers to lower-quality offices show a clear one ($p < .001$). The asymmetry reported in the direction split is not driven by the group with the pre-trend, since it is the up-movers who gain.

The second threat is that origin and destination quality are measured over mismatched periods, so that an apparent quality difference partly reflects when each office was observed. Rows 3 and 4 address it. They measure both offices' quality strictly within a common pre-move window of three and five years respectively, giving slopes of 0.146 (SE 0.021) and 0.157 (SE 0.021). These are the most conservative constructions we estimate and place the lower bound of the credible range.

The third threat is that office quality is itself an estimate, so measurement error in the regressor attenuates the slope while error correlated with the mover's own outcomes inflates it. Row 6 and the

three rows beneath it address it. Row 6 restricts to movers whose office quality is estimable purely leave-out, giving 0.257 (SE 0.022). Deleting one destination office at a time keeps the slope between 0.245 and 0.262 across all 2,037 deletions, so no single office drives it. A bootstrap over movers gives a 95% confidence interval of 0.212 to 0.303, and rebuilding office quality fifty times from random halves of each office's other agents yields a mean slope of 0.226 with a standard deviation of 0.009, smaller as expected when quality is measured from half the data but stable across every re-split.

Appendix C.3 adds a fourth consideration bearing on which end of the range to prefer. Agents moving to worse offices were already declining before the move, so part of their measured loss belongs to their own trajectory rather than to the destination. The slope is therefore an upper bound on the causal return to office quality, and the fixed-window constructions at the bottom of the range are the safer reading.

Table D2. The return to office quality: constructions and robustness.

Specification	Slope on change in office quality	SE	N (movers)
(1) Baseline: post-move change on change in office quality	0.300	(0.020)	9,595
(2) Destination and origin entered separately	0.330 / -0.266	(0.026) / (0.025)	9,595
(3) Fixed pre-move window, K = 3 years	0.146	(0.021)	7,237
(4) Fixed pre-move window, K = 5 years	0.157	(0.021)	7,335
(5) Reversion-safe change, pre-move decile fixed effects	0.276	(0.028)	5,526
(6) Strict leave-out subsample	0.257	(0.022)	9,044
Delete-one-office jackknife range on (6)	[0.245, 0.262]		2,037 offices
Bootstrap 95% confidence interval on (6), 500 replications	[0.212, 0.303]		
Random re-splits of office quality (mean, SD), 50 replications	0.226 (0.009)		

Notes. The dependent variable is each mover's post-move minus pre-move mean market-adjusted performance; the pre-move baseline uses even-year observations. Row (2) reports destination and origin coefficients separately, and the test that they are equal in size is rejected at p = .037. Rows (3) and (4) measure both offices' quality strictly within a common pre-move window. Row (5) baselines the change on t-2 and earlier, with fixed effects for the mover's pre-move within-year decile. Row (6) uses only movers whose office quality is estimable purely leave-out. Robust standard errors in parentheses.

APPENDIX E: SAMPLE CONSTRUCTION AND PANEL DESCRIPTIVES

E.1 From Raw Records to the Analysis Sample

The data section summarizes the construction of the analysis sample. This appendix documents each step and reports the descriptive exhibits.

The raw extract is the ATTOM national MLS listing-analytics file, from which we keep every record whose listing state is Utah, Arizona, or Idaho: 14,208,564 listing-status records. A status record is a single reported state of a listing (active, pending, sold, expired, withdrawn), so one listing generates several records over its life. We collapse this status panel to one row per listing, keeping the final status, the originating office and agents, prices, and time on market. This yields 4,206,982 distinct listings, geocoded to five-digit ZIP codes and merged to 2020 USDA rural-urban commuting area codes.

MLS member and office identifiers are assigned independently by each source MLS, so the same numeric identifier can denote different people in different systems. Before any other construction we namespace every agent and office identifier by its MLS source, so that identity is unique within and across systems. The same step allows a person's listing-side and buyer-side records, which carry the same member identifier within an MLS, to resolve to a single individual. Failing to namespace would merge distinct agents who happen to share an identifier across systems, which would attenuate the individual component we estimate. Four listing services additionally route transactions whose counterparty is not one of their members to a single non-member identifier, which the feed labels NON-MLS or NON-MEMBER. These identifiers accumulate tens of thousands of credits across hundreds of cities and, in one case, three states, so they denote a bucket rather than a person. The largest alone carried 36,532 credits over seventeen years and was the most active account in the

panel. We drop listings credited to these identifiers and treat their buyer-side credits as unidentified, which removes 46 accounts and 1.7% of agent-transaction credits.

The transaction population consists of sold listings attributed to a valid listing office, after excluding listings flagged in the public remarks as existing purely for appraisal or comparable purposes. These are records an agent enters to document an off-market or non-arm's-length transaction rather than a property that was marketed and sold, so they do not reflect the production we measure. This leaves 2,760,944 sold listings. Because the underlying feed is status-level, a single sold listing can appear more than once; we retain one record per listing, and after removing listings before 2007 and those without a usable ZIP code 2,472,825 sold listings remain. Seven further screens narrow this to completed home sales. We drop leases and rentals, which carry a monthly rent rather than a sale price in the price field; non-residential property types such as land, commercial and farm listings; listings credited to placeholder member identifiers rather than to a person; sale prices below $10,000, which are non-arm's-length transfers; listings reporting negative or implausibly long marketing durations; properties recorded twice because two listing services cover the same area; and the final calendar year of the extract, which is mechanically incomplete. The main analysis sample contains 1,898,013 sold transactions. Requiring a complete set of property characteristics, which only the price decomposition of Appendix A needs, leaves 775,376 listings.

Table E1. Sample construction.

Stage	Unit	N	% of raw
Raw MLS status records (UT, AZ, ID)	status records	14,208,564	100.0
Distinct listings (status panel collapsed)	listings	4,206,982	29.6
Sold, valid office, non-appraisal (transaction population)	listings	2,760,944	19.4
Sold listings, de-duplicated, 2007+, valid ZIP	listings	2,472,825	17.4
Main analysis sample (completed home sales, cleaned screens)	listings	1,898,013	13.4
Hedonic subsample (complete property characteristics; Appendix A only)	listings	775,376	5.5
Agent-year panel (both sides, all agents)	agent-years	571,295	

Notes. The fourth row removes duplicate status rows for the same listing, listings before 2007, and listings without a usable ZIP code. The fifth row applies the seven cleaned-sample screens described above: leases and rentals, non-residential property types, placeholder and non-member identifiers, sale prices below $10,000, impossible marketing durations, properties recorded by two overlapping listing services, and the mechanically incomplete final calendar year. The hedonic subsample retains only listings recording every property characteristic the price regression needs (bathrooms, bedrooms, lot size, year built, living area and days on market); it is 40.9% of the main analysis sample and is used only for the transaction-level price decomposition of Appendix A. Every other result in the paper is estimated on the main analysis sample.

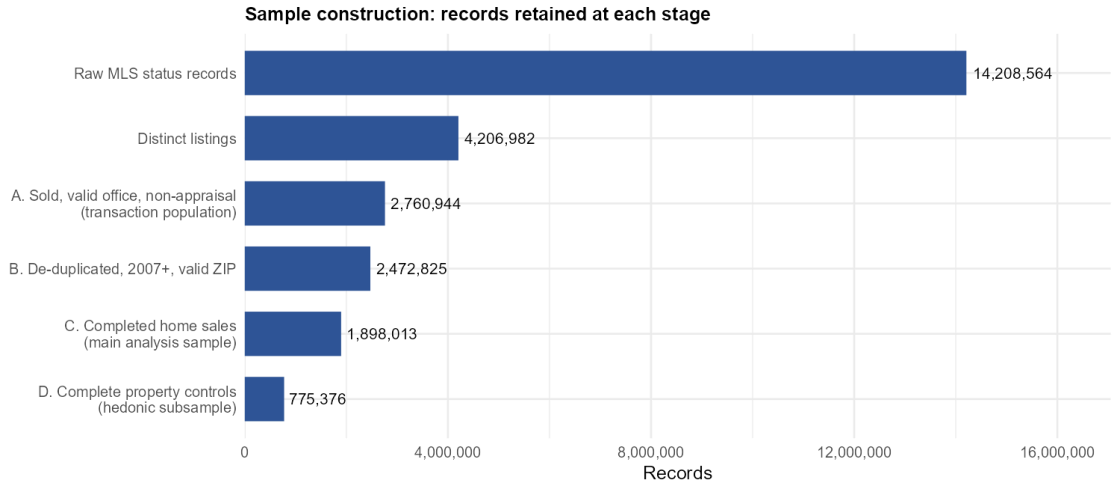


Figure E1. Sample construction. Records retained at each stage, from raw status records down to the transaction population and the samples used for analysis.

Figures E2 and E3 report the two descriptive patterns discussed with the descriptive statistics. Figure E2 shows that annual transactions and the count of active agents both track the housing cycle. Both climb steadily to a 2021 peak and fall sharply afterward, and the agent count first falls through a 2009 trough during the housing bust before expanding through the following decade. Figure E3 shows the heterogeneity that motivates the paper: career transactions per agent and

average agents per office-year are both heavily right-skewed, and the office distribution is concentrated among small offices, which is the condition under which individual heterogeneity remains visible in unit performance.

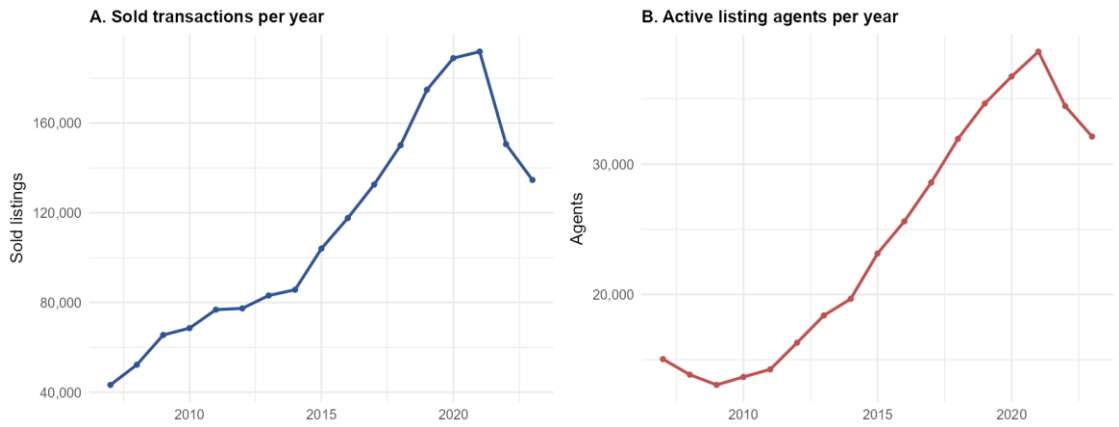


Figure E2. The sample over time (the panel ends in 2023; the partial 2024 records are excluded from the sample). (A) Sold transactions per year. (B) Active agents per year.

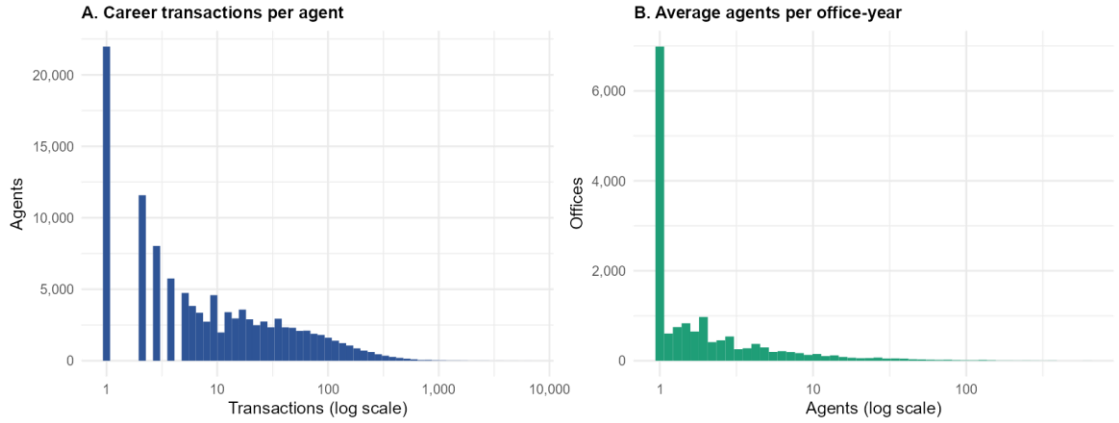


Figure E3. Heterogeneity in activity (log scales). (A) Career transactions per agent. (B) Average agents per office-year.

E.2 Assigning a Market to an Agent-Year

The market adjustment described earlier absorbs a ZIP-by-year fixed effect, which requires one ZIP per agent-year, but agents commonly sell in several. We assign each agent-year the ZIP in which it closed the most transactions, breaking ties toward the lower ZIP code, and apply the same rule to

the office and the franchise. No transaction is dropped: volume and transaction counts are summed across every ZIP in which the agent sold.

The rule binds unevenly. Only 35.0% of agent-years fall in a single ZIP, the median spans two and the mean 3.8, and listing-weighted only 37.8% of transactions occur in their own agent-year's modal ZIP. Concentration falls with activity. Agent-years closing one or two transactions are 71.0% single-ZIP, while those closing twenty-six or more span 17.9 ZIPs on average and place 29.3% of their listings in the modal one. The same rule applied to offices is nearly exact, since 95.3% of agent-years involve a single office and the modal office accounts for 98.4% of listings.

Because the highest-volume agent-years are the most coarsely labelled, the natural concern is that the adjustment under-corrects for them and credits the resulting market variation to the agent. Table E2 tests this. Row 2 replaces the modal indicator with the agent-year's full portfolio, each ZIP entering weighted by its share of that agent-year's listings, so the modal specification is the special case in which the weight is one at the mode. The weighted control fits better, raising the stage-1 R-squared from .305 to .314, and the two residuals correlate at .971. The agent share rises from 43.6% to 44.2% and the office share is unchanged. Row 3 instead uses the modal county, a coarser market the rule measures far more accurately, and raises both components. The paper's specification therefore returns the smallest individual share of the three: the modal assignment is a simplification, but correcting it moves the headline number up rather than down.

Table E2. The market adjustment and the choice of ZIP.

Market adjustment	Stage-1 R2	Agent	Office	Residual
(1) Modal ZIP-by-year (the paper)	.305	43.57	9.05	47.38
(2) Exposure-weighted ZIP-by-year	.314	44.22	9.04	46.75
(3) Modal county-by-year	.246	45.09	10.21	44.69

Notes. Shapley shares (percent) of the variance of market-adjusted log annual volume, for the 434,881 agent-years of agents observed at least five years, using the same decomposition as Table 2. Only the market control differs. Row 2 enters every ZIP an agent-year sold in, weighted by its share of that agent-year's listings, across 10,567 ZIP-by-year cells; row 1 is the special case in which the weight is one at the modal ZIP. Row 3 absorbs less market variation, which raises both components mechanically.

E.3 The Upper Tail of the Activity Distribution

A member identifier denotes a license rather than a natural person. Most identifiers are individuals, but the extreme upper tail is not. After removing the non-member buckets described in E.1, the most active agent-year closes 2,126 transactions, and the accounts above roughly one sale per calendar day are institutional: iBuyer brokerages, homebuilder sales arms, and distressed-asset disposition contracts, each of which transacts through one agent of record. Thirty-eight identifiers exceed that threshold and hold 2.4% of all agent-transaction credits.

We do not remove them. They are real licenses conducting real transactions, no reproducible rule separates them from a high-volume team without also removing genuine builder and property-management employees, and the identification of institutional status rests on company names rather than on anything observable in the production data. Instead, we report what the decomposition does without them. Table E3 removes the tail two ways: dropping every agent who ever exceeds one sale per day and trimming only those agent-years. Neither moves the individual share by more than a tenth of a percentage point, so the paper's central estimate does not rest on the accounts whose status as individuals is most doubtful.

Table E3. The individual share with the upper tail removed.

Specification	Agent-years	Agent (%)	Office (%)	Max/year
Screened sample (reported)	434,881	43.57	9.05	2,126
Drop agents ever above one sale per day	434,467	43.47	9.03	365
Trim agent-years above one sale per day	434,801	43.55	9.03	365

Notes. Shapley shares of the variance of market-adjusted log annual volume among agents observed at least five years, as in Table 2. One sale per calendar day is 365 transactions in a year. The 38 identifiers above that threshold hold 2.4% of credits.

APPENDIX F: THE SCREENED OFFICES

The screening analysis scores each office on the within-agent differences of the agents who worked there. For each agent observed at more than one office, we compare the agent's

performance at a given office with that agent's own average across the other offices where they worked. An office's score is the mean of these differences across its agents, and each agent's own contribution is excluded from the score they are compared against, so no agent helps set the benchmark by which they are measured.

F.1 Do the Screened Offices Cause the Boost?

Even a within-agent boost could reflect timing rather than development if these offices recruit agents at a low point, and the mobility results show that this concern is real. We therefore estimate an event study around arrival at one of the 69 offices, compared with arrival at an ordinary office, using agents who never move as controls (Figure F1). Arrivals at ordinary offices gain 0.033 to 0.047 SD. Arrivals at the 69 offices gain 0.589, 0.543, and 0.523 SD in the three years after arrival, each with a standard error below 0.098. The pre-arrival coefficients of negative 0.278 and negative 0.318 do show the pattern a rebound would produce, so we test it directly by dropping the arrivals whose performance three years before arrival was unusually low relative to their own history. That leaves 239 of 293 arrivals, and the post-arrival gains remain large at 0.536, 0.476, and 0.393 SD, while the pre-arrival coefficient turns positive at 0.237. The remaining arrivals were, if anything, above their own baseline before they arrived. Mean reversion does not produce this pattern.

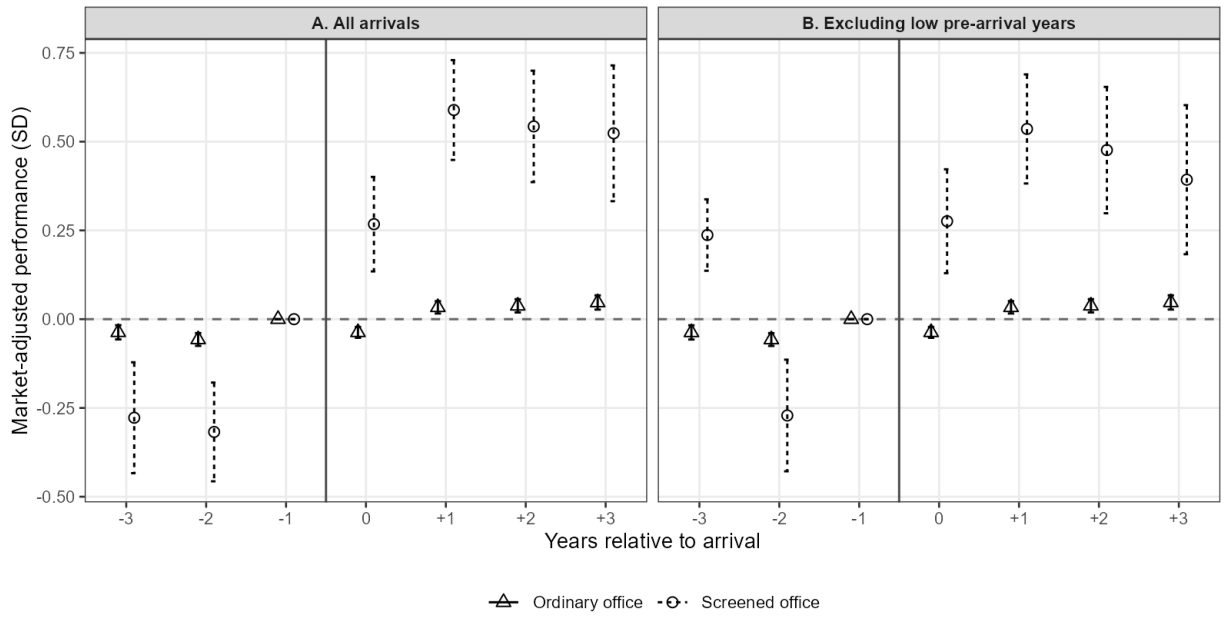


Figure F1. Arrival at screened versus ordinary offices. Event-study coefficients with 95% confidence intervals; the year before arrival is the omitted baseline and agents who never move are the controls. Panel A covers all arrivals. Panel B excludes arrivals whose performance three years before arrival was in the bottom quartile of their own history.

F.2 How the Screened Offices Differ

Table F1 compares the 69 screened offices with the other 1,603 offices for which both screens can be computed. The first three rows are mechanical, since they are the screens themselves: the screened offices have a within-agent boost of 0.260 SD against roughly zero elsewhere, and their quality reproduces across both halves of the panel. The remaining rows are not mechanical and are more informative. The screened offices are smaller in headcount, averaging 15.6 agents per office-year against 22.4 elsewhere, yet they transact far more, at 283.7 transactions and $110.4 million per office-year against 135.6 and $50.3 million. They also employ fewer distinct agents over the panel, 57.4 against 86.0, so they are not simply churning through people. They are no more likely to be franchise-affiliated, at 33.3% against 30.3%, which is the one dimension on which they are indistinguishable. The picture is of small, stable, high-volume offices rather than large or branded ones.

Table F1. The 69 screened offices compared with other offices.

Variable	Screened (69)	Other (1,603)	Difference	p
Within-agent boost (SD)	0.260 (0.168)	-0.009 (0.144)	0.269	< 0.001
Office quality, odd years (SD)	0.907 (0.251)	0.029 (0.360)	0.879	< 0.001
Office quality, even years (SD)	0.918 (0.245)	0.020 (0.350)	0.898	< 0.001
Agents per office-year	15.58 (15.39)	22.44 (39.50)	-6.85	.001
Transactions per office-year	283.7 (420.8)	135.6 (229.3)	148.1	.005
Dollar volume per office-year ($000)	110,386.6 (166,576.0)	50,296.9 (99,851.3)	60,089.7	.004
Years observed	10.17 (4.49)	11.60 (4.40)	-1.43	.012
Distinct agents ever employed	57.4 (82.6)	86.0 (177.8)	-28.6	.010
Franchise-affiliated (share)	0.333 (0.475)	0.303 (0.460)	0.031	.599

Notes. Office-level means with standard deviations in parentheses, across the 1,672 offices for which both screens of Section 7 can be computed. The first three rows are the screening variables themselves and differ by construction. Monetary values are in thousands of dollars. The final column reports the p-value of a two-sided t-test of equal means.