Geography Matters More for Small Firms: Frontier Skill Supply in Startup Labor Markets*

Paul J. Hamilton David H. Hsu Prasanna Tambe Zhiwei Wang

March 17, 2026

DRAFT. Please do not post, quote, or circulate.

Abstract

Entrepreneurial firms face greater friction in hiring talent due to reputation scarcity and resource constraints. These constraints are salient in high-technology settings where talent supply is often scarce and geographically concentrated. Using data from a large online labor market platform for high-growth startups, we jointly track job listings and applications for more than 1,000 emerging technologies across U.S. regions. We document two key findings. First, pioneer regions—where technologies were first developed—are initially the tightest labor markets for small firms (few applicants per listing), but this pattern fully reverses as experienced talent accumulates, with pioneer regions crossing from a relative talent deficit to a talent advantage roughly 20 years after emergence. Large firms, with established brands, formal recruiting operations, and broader search capacity, are insulated from local talent conditions and absorb scarce workers early in the technology lifecycle. Second, the mechanism underlying this pattern is an asymmetry in diffusion: job listings for new technologies spread rapidly across regions, but applications lag substantially, generating the lifecycle in labor market tightness that disproportionately binds small firms. These findings highlight how firm size mediates the interaction between technology diffusion and talent availability, with implications for startup location strategy and the geography of entrepreneurial growth.

Keywords: *Technology diffusion, firm size, startup labor markets, geographic frictions, labor market tightness*

*We are grateful for valuable feedback from seminar participants at the Conference on Information Systems and Technology, and the Workshop for Information Systems & Economics. Contact information: Paul Hamilton (Harvard Business School, phamilton@hbs.edu), David H. Hsu (Wharton School, University of Pennsylvania, dhsu@wharton.upenn.edu), Prasanna Tambe (Wharton School, University of Pennsylvania, tambe@wharton.upenn.edu), and Zhiwei Wang (Wharton School, University of Pennsylvania, zwangcj@wharton.upenn.edu).

1 Introduction

Despite the relative abundance of sought after skills within tech clusters, these labor markets were exceptionally tight in the late 2010s and exhibited very low unemployment rates. Thus, many businesses located in these talent clusters struggled to get the workers they wanted especially if they lacked a brand name like Apple or Netflix that attracts employees.

Kerr and Robert-Nicoud (2020)

Recent work in strategy and entrepreneurship argues that startups face unique challenges in recruiting talent, shaped by frictions such as reputation scarcity, resource constraints and career uncertainty faced by workers (Hsu et al., 2026; Roach and Sauermann, 2015; Baron and Hannan, 2002). However, this work has largely abstracted away from the location choice of startups. Geographic location is particularly relevant in frontier technological settings, where the skilled talent startups need are often scarce and concentrated in regional clusters. A large literature across economics, strategy, and innovation studies documents that innovative activity and technology adoption tend to cluster geographically, reflecting localized learning, knowledge spillovers, and agglomeration economies (Marshall, 1890; Jaffe et al., 1993; Moretti, 2011).

For startups in high-technology sectors, locating in regional centers of innovation provides access to a thicker pool of specialized talent. But these same clusters attract established firms that are better positioned to compete for scarce workers through brand recognition, compensation and structured career paths. These early innovation clusters may thus simultaneously host the deepest talent pools and the tightest labor markets for startups. It is unclear whether locating at innovation clusters expand or constrain talent access for startups (Kerr and Robert-Nicoud, 2020).

This tension between a deeper talent pool and greater competition can evolve as the technology diffuses geographically. Recent work shows that labor demand for emerging technologies is initially concentrated in a small number of pioneering regions and only gradually spreads to other locations (Kalyani et al., 2025). An open question is that if the supply of skilled talent follows the same path as job demand. If talent supply diffuses more slowly — because workers face frictions in localized

learning, mobility and career uncertainty — talent competition in pioneering regions can alleviate over the technology life cycle. As competition eases, startups, which lack the brand recognition and resources compared to established firms, stand to benefit disproportionately from a deepening local talent pool.

In this paper, we ask whether the hiring frictions that startups face in early innovation clusters ease or intensify over the technology life cycle. We study this question using data from Wellfound (formerly AngelList Talent), a large online labor market platform for startups in high-growth technology sectors. The data record both job listings and applications, allowing us to observe both sides of the market for more than 1,000 emerging technologies across U.S. regions from 2012 to 2023. For each technology, region, and year, we track where firms post jobs and where workers apply. This joint observation of listings and applications distinguishes our approach from prior work, which typically infers labor market adjustment indirectly from wages, employment levels, or migration patterns. Following Kalyani et al. (2025), we define "pioneer regions" as those that housed the majority of early patenting activity for each technology.

Our key results are as follows. First, the labor market consequences of technology diffusion fall disproportionately on the smallest startups. Pioneer regions—where new technologies first emerge—are initially tight labor markets for firms with fewer than 500 employees: small firms there receive fewer applicants per listing than comparable firms elsewhere. Over time, this pattern reverses as local talent pools deepen, and pioneer regions become the places where startups enjoy the greatest recruiting advantage. The implied crossover from talent deficit to talent advantage occurs roughly 20 years after a technology's emergence. Larger firms show no comparable pattern. With established recruiting operations, stronger employer brands, and broader search capacity, large firms appear able to insulate themselves from local talent conditions and absorb scarce workers early in the technology lifecycle. The key boundary condition on these labor market frictions is organizational scale: geography matters most for the firms least able to recruit around it.

Second, the mechanism behind this pattern is an asymmetry in how quickly the two sides of the labor market adjust. Job listings for new technologies spread rapidly across regions, but appli-

cant supply follows more slowly. Conditional on posted wages, the pioneer–non-pioneer gap in job listings converges toward zero as technologies mature, while the corresponding gap in applicant interest persists for much longer. This divergence generates the lifecycle in labor market tightness that disproportionately affects small firms.

We note three caveats. First, our data capture application behavior on a single platform that specializes in startup hiring; informal recruiting, referrals, and off-platform hiring may follow different dynamics. Second, our outcomes measure where workers direct their applications, not where they reside or whether they are ultimately hired. Third, pioneer status is not randomly assigned. Our primary interest is in the differential evolution of tightness over technology age—how the gap between pioneer and non-pioneer regions changes as technologies mature—not the causal effect of pioneer designation.

Our work relates to several literatures. We build directly on Kalyani et al. (2025), who document the geographic diffusion of technology-related job postings. We extend their analysis by observing the supply side—worker applications—and showing that talent diffusion lags job diffusion in ways that differ sharply by firm size. We also contribute to research on agglomeration and the geography of innovation (Marshall, 1890; Jaffe et al., 1993; Moretti, 2011), by showing that persistent regional advantage operates partly through the accumulation of experienced talent, and that this advantage is mediated by firm scale. We connect to the strategy and entrepreneurship literature on how labor market frictions shape startup growth (Baron and Hannan, 2002; Roach and Sauermann, 2015; Hsu and Tambe, 2024), by providing direct evidence that small firms are most exposed to local talent dynamics.

The remainder of the paper is structured as follows. Section 2 develops the conceptual framework. Section 3 describes the data. Section 4 presents the results. Section 5 discusses implications and limitations. Appendix B presents a two-region search model that formalizes the key mechanisms.

2 Conceptual Framework

This section outlines the logic connecting technology diffusion, labor market frictions, and firm size. The argument proceeds in three steps: technologies diffuse faster on the firm side than the worker side; this asymmetry generates a lifecycle pattern in regional labor market tightness; and the tightness pattern binds differentially by firm size. Appendix B formalizes these mechanisms in a two-region search model.

2.1 Asymmetric Diffusion of Jobs and Talent

New technologies emerge in a small number of regions before spreading (Kalyani et al., 2025).[1] On the firm side, posting jobs tied to a new technology is a low-cost way to experiment with adoption. Startups in non-pioneer regions can begin expressing demand for new skills quickly, even before they have the organizational capacity to fully utilize the technology (Baron and Hannan, 2002). As a result, job listings diffuse rapidly.

On the worker side, several frictions slow adjustment. Frontier skills are often tacit and acquired through hands-on experience in environments where the technology is actively used (Zucker et al., 2003; Rosenkopf and Almeida, 2003). Workers face relocation costs, uncertainty about employer quality, and career risks associated with moving to unfamiliar firms (Moretti, 2011). In startup settings, these frictions are amplified: startups offer steeper learning curves but also higher risk and weaker signaling of firm quality (Roach and Sauermann, 2015; Hsu et al., 2026). Evidence that technical labor mobility occurs disproportionately *within* technology clusters rather than across them is consistent with these frictions binding in practice (Fallick et al., 2006).

The result is an asymmetry: firms across regions quickly express demand for new skills, but workers with those skills remain geographically concentrated. This asymmetry generates the lifecycle pattern we document empirically. Pioneer regions are initially tight—many listings, few applicants—

[1] The term *pioneering regions* was introduced by these authors as "the urban areas that housed the majority of early patenting for the new technology."

and become progressively slacker as the local talent pool deepens through learning, turnover, and reputation accumulation (Owen-Smith and Powell, 2004).

2.2 Firm Size and Labor Market Competition

The central finding of this paper is that the lifecycle pattern is concentrated among the smallest startups. We argue this reflects differences in how firms of different sizes compete for scarce talent.

When a technology first emerges, the pool of experienced workers is small and geographically concentrated. Large firms—even within the startup ecosystem—have advantages in this competition: established employer brands, formal recruiting operations, broader geographic search capacity, and the ability to hold vacancies longer or offer more competitive packages (Tambe, 2014). These scale advantages allow large firms to attract scarce workers regardless of local conditions. Small firms lack these buffers. They depend heavily on local labor markets, rely on informal signals of skill and fit, and face higher costs from failed hires.

Early in the technology lifecycle, this creates a crowding-out dynamic in pioneer regions. Large firms absorb a disproportionate share of the limited talent pool, leaving small firms in the tightest part of the market. As the technology matures and the pool of experienced workers grows—through learning-by-doing, job-to-job mobility, and the accumulation of technology-specific human capital—the constraint on small firms relaxes. The lifecycle reversal we document is, in this framing, the gradual deepening of the talent pool past the point where large-firm absorption leaves small firms starved.

In non-pioneer regions, the dynamic differs: fewer experienced workers are available, but fewer large firms compete for them either, so the firm-size differential in talent access is less pronounced. The firm-size heterogeneity we observe is therefore a joint product of geographic concentration of talent and within-market competition mediated by organizational scale.

3 Data and Key Measures

3.1 AngelList Data on Job Postings and Applications

We leverage tens of millions of job applications submitted to full-time, non-remote job postings on AngelList Talent (now branded as "Wellfound"), a platform for job listings in the high-tech startup ecosystem. Since its inception in 2011, AngelList Talent has attracted over 62,000 companies to post over 215,000 jobs. The site has had approximately 3.6 million unique job candidates upload profiles and has connected about 1 million applicants to jobs. We are unaware of other papers investigating both sides of the technical job market at this scale.

We have data on all job postings employers made on the platform from 2012 to 2023, including a description of the firm, the skills required for the job, and compensation details (wages and equity). From the applicant profiles, we have their demographic information, data on prior work experience, and the complete set of job listings they applied to on the platform.

We also observe the full matching funnel between applicants and jobs. The main dataset includes *pairings* between listings and applicants. A pairing is generated when the applicant *views* a job listing, indicating that the listing entered the worker's consideration set. Conditional on viewing, applicants separately decide whether to submit an *application*. The *applicant-to-listing ratio* (our main outcome) counts workers who completed the application step. We also observe whether the firm expressed mutual interest in each applicant, which allows us to observe the number of *matches* between workers and firms. We do not observe what happens off-platform after workers and firms match, so we cannot observe the interview process or know whether matches ultimately lead to hiring. Figure 1 plots the geographic distribution of job listings, demonstrating that our sample has broad geographic coverage of major population centers in the U.S.

3.2 Pioneering Regions of New Technologies

For our analysis, we need to (i) identify the core technologies reflected in each job listing's requirements and (ii) identify the pioneering locations of those technologies. To do so, we use text embeddings to match the skill requirements in each listing to the data from Kalyani et al. (2025).

Kalyani et al. (2025) identifies 1,286 unique technologies (e.g., cloud computing, autonomous vehicles) and maps them to the Core-Based Statistical Areas (CBSAs) where they were first developed. Their data construction process is roughly as follows. New technologies are identified by first obtaining bi-grams appearing in U.S. patent filings since 1976. Among these bi-grams, only those appearing in patents with at least 1,000 citations are retained. Technology-related bi-grams are then filtered through matching with Wikipedia pages. For each technology, the authors define the "emergence year" as the year in which the technology reaches 100 cite-weighted patents (and is followed by at least 10% annual growth in patent filings over the next five years). The "pioneering regions" for a technology are then defined as the regions that account for the majority of patenting activity around that emergence period.

We match our job listings to the technologies from Kalyani et al. (2025) using a text-embedding approach. For each listing, we concatenate the required skills into a single text string. We then use the OpenAI text embeddings API to generate embeddings for both (i) these job-listing skill strings and (ii) the technology descriptions in the Kalyani et al. (2025) data. For each listing, we identify all technologies whose cosine similarity with the listing exceeds a chosen threshold and treat these as matches. Under this approach, a job listing may be matched to zero, one, or multiple technologies. Our results are robust to alternative cosine-similarity thresholds. Table 1 presents example technologies and the top 10 CBSAs ranked by the number of technologies they pioneered.

3.3 Measurement of Labor Market Outcomes

Our main outcome measures labor market tightness defined at the region-technology-year level. A natural starting point is the number of applications per listing in region r, technology k, and year t.

However, because technologies differ substantially in their overall scale across years, we normalize both vacancies and applicants within each technology-year market before comparing regions. For each technology k and year t, we first compute region r's share of listings:

$$\text{Regional Share (Listings)}_{r,k,t} = \frac{\text{Num. Listings}_{r,k,t}}{\text{Num. Listings}_{k,t}}$$

and its share of applicants:

$$\text{Regional Share (Applicants)}_{r,k,t} = \frac{\text{Num. Applicants}_{r,k,t}}{\text{Num. Applicants}_{k,t}}.$$

Here, $\textit{Num. Applicants}_{r,k,t}$ denotes the number of unique applicants in year t to jobs requiring technology k in region r, while $\textit{Num. Applicants}_{k,t}$ denotes the total number of unique applicants across all regions for technology k in year t. We then use these measures to construct the regional applicant-to-listing ratio as:

$$\text{Applicant-to-Listing Ratio}_{r,k,t} = \frac{\text{Regional Share (Applicants)}_{r,k,t}}{\text{Regional Share (Listings)}_{r,k,t}}.$$

This measure compares a region's share of applicants to its share of vacancies within a given technology-year market. Values below one indicate that a region accounts for a larger share of listings than of applicants, implying a tighter market and more difficult hiring. Values above one indicate that a region attracts a larger share of applicants than its share of listings would predict, implying a slacker market and easier hiring.

This measure is closely related to the notion of labor market tightness in standard search-and-matching models (Mortensen, 2011). In those models, the job-filling rate is increasing in the ratio of job seekers to vacancies. Our applicant-to-listing ratio is the empirical analogue of $\theta_{r,t}$ in the search model developed in Appendix B, with the caveat that it captures the effective supply of talent observed on Wellfound rather than the full market, since startups may also recruit through other channels such as company websites.

It is useful to note that our main outcome measure can be rewritten as:

$$\text{Applicant-to-Listing Ratio}_{r,k,t} = \frac{\text{Num. Applicants}_{r,k,t}}{\text{Num. Listings}_{r,k,t}} \times \frac{\text{Num. Listings}_{k,t}}{\text{Num. Applicants}_{k,t}}.$$

The second term varies only at the technology-year level. Therefore, in specifications with technology-year fixed effects, variation in the applicant-to-listing ratio is equivalent to variation in applicants per listing within a region-technology-year market, up to a technology-year-specific scaling factor.

4 Results

In this section we first describe the diffusion of job postings from pioneer regions to other labor markets, then quantify how both listings and applications change with technology age, and finally show how the balance between the two — applicants per listing, our inverse measure of labor-market tightness — systematically changes over the technology lifecycle.

4.1 Listings and Applicants

Before analyzing labor market tightness, we first document the distinct geographic trajectories of job listings and applicants. Understanding how these two sides of the market diffuse is critical for interpreting the tightness reversal we document later.

We begin by visualizing the spatial diffusion of job listings using Machine Learning (ML) as a representative case study. In the early stages of the technology's lifecycle (2012–2014), hiring activity was geographically concentrated in a small number of coastal innovation hubs. However, as the technology matured, the geography of listing expanded substantially. By the end of our sample period, ML-related job postings had proliferated into secondary markets and interior regions, reflecting a broad "democratization" of demand as firms across the country adopted these tools.

Does the supply of talent follow this same dispersive path? To answer this, we estimate separate dynamic specifications for the volume of all job listings and applicants in our sample. Defining $y_{r,k,t}$ as the number of active job listings (or alternatively, job applicants) for technology k in region r at

year t, we estimate:

$$\ln\left(y_{r,k,t}\right) = \beta_1 \text{Pioneer}_{r,k} + \beta_2 \left(\text{Pioneer}_{r,k} \times \text{Years Since Emergence}_{k,t}\right) + \mathbf{X}_{r,k,t}\mathbf{\Phi} + \gamma_{k,t} + \lambda_{r,t} + \epsilon_{r,k,t}$$

The variable $\text{Pioneer}_{r,k}$ is an indicator equal to one if region r is a pioneer region for technology k. The variable $\text{Years Since Emergence}_{k,t}$ measures the age of the technology. The vector $\mathbf{X}_{r,k,t}$ includes the median posted salary of jobs in each market as a control. We note that posted salary is a plausible mediator of tightness rather than a purely exogenous confounder (firms adjust compensation in response to hiring difficulty), so our primary specifications treat salary as a control for composition and we verify that results are robust to its exclusion.

Crucially, the main specification includes Technology × Year ($\gamma_{k,t}$) and Region × Year ($\lambda_{r,t}$) fixed effects. The inclusion of $\gamma_{k,t}$ absorbs all technology-specific time trends, such as the global "hype cycle" of a new technology or aggregate shifts in labor supply for specific technical skills. The inclusion of $\lambda_{r,t}$ absorbs all region-specific time trends, such as local macroeconomic conditions, housing costs, or broader regional labor market tightness. Therefore, our estimates of β_1 and β_2 are identified from the differential evolution of pioneer vs. non-pioneer regions *within* the same technology-year and region-year.

Two identification caveats are worth stating up front. First, pioneer status is not randomly assigned: pioneer regions have durable, technology-specific attributes—research institutions, firm networks, alumni pipelines—that independently attract both postings and applications. This selection means $\hat{\beta}_1$ (the level difference at emergence) is descriptive; it should not be interpreted as the causal effect of pioneer designation. Our primary estimand is $\hat{\beta}_2$ (the differential trend over technology age), which is identified from within-technology-year variation in how pioneer and non-pioneer tightness evolve and is less sensitive to time-invariant selection. Second, our outcome captures the geography of worker *application behavior*—where workers direct their applications—rather than the stock of skilled workers residing in a region, so it should be interpreted as a measure of labor supply revealed through applications rather than local talent availability per se. The relevant object

is therefore effective labor supply to firms on the platform, not the underlying regional endowment of skilled workers.

The results reveal a fundamental asymmetry in the speed of adjustment. On the demand side (Figure 3, left panel), we observe a rapid convergence. The positive marginal effect of being in a pioneer region decays steeply, confirming that firms in follower regions are quick to post vacancies once a technology becomes proven. On the supply side (Figure 3, right panel), however, the diffusion appears more sluggish. While the concentration of applicant interest in pioneer regions does decline, it does so at a much slower rate than listings. We note that the level difference between pioneers and non-pioneers in both panels partly reflects selection—pioneer regions have durable technology-specific attributes that attract both firms and workers independently—so the primary object of interest is the *differential trajectory*: how quickly each side converges and how the balance between them evolves over technology age.

This divergence implies that, even conditional on wages, the promise of new work spreads quickly while the workforce interested in (and capable of) performing that work remains sticky. The result is a widening structural gap: as technologies mature, follower regions generate ample demand but struggle to attract a commensurate share of applicant interest, while pioneer regions—despite losing their monopoly on job postings—retain a disproportionate share of the talent supply. This asymmetric pace of diffusion is the mechanism underlying the lifecycle reversal formalized in Appendix B: postings (driven by firm surplus) diffuse faster than applicants (held in place by the learning premium). It sets the stage for the reversal in labor market tightness we examine next.

4.2 Main Results

Our framework predicts that the labor market consequences of technological pioneering should be most pronounced for startups. Because small firms have weaker employer brands and more limited recruiting infrastructure, they are less able to recruit around local labor market frictions. We therefore focus first on firms with fewer than 500 employees and ask how labor market tightness evolves for these firms in pioneer versus non-pioneer regions over the technology lifecycle. We then

formally test whether the corresponding lifecycle pattern differs for larger firms.

Our main outcome is the applicant-to-listing ratio, which measures labor market tightness from the firm's perspective. A higher ratio indicates that firms receive more applicants per vacancy and therefore face a slacker market, while a lower ratio indicates a tighter market. This measure is especially relevant for startups, for whom local shortages of skilled labor should bind most strongly.

Table 4 reports estimates from a fully interacted specification that allows the pioneer lifecycle pattern to differ by firm size. We use 500 employees as the cutoff because it corresponds to the Small Business Innovation Research (SBIR) program's definition of a small business.[2] In this parameterization, the coefficients on *Pioneer Region* and *Pioneer Region* $\times$ *Years Since Emergence* capture the lifecycle pattern for firms with fewer than 500 employees, while the interactions with the large-firm indicator test whether the corresponding effects differ for firms with 500 or more employees.

The estimates show a strong pioneer lifecycle pattern for small firms. At emergence, pioneer regions are significantly tighter labor markets for firms with fewer than 500 employees: the estimated coefficient on *Pioneer Region* is negative, indicating that small firms in pioneer hubs receive fewer applicants per listing than comparable firms in non-pioneer regions. Over time, however, this gap closes. The interaction of *Pioneer Region* with *Years Since Emergence* is positive and statistically significant, implying that the applicant-to-listing ratio in pioneer regions rises steadily relative to non-pioneer regions as the technology matures.[3] Thus, for startups, the regions where new technologies first emerge are initially the hardest places to hire, but they become progressively more favorable over the technology lifecycle.

The interaction terms with the large-firm indicator show that this lifecycle pattern is significantly weaker for firms with 500 or more employees. In other words, the main effect is not a pooled average that applies uniformly across firms: it is fundamentally a small-firm phenomenon. Large

[2]The 500-employee cutoff is also supported by the data: when we estimate the full specification separately by firm-size bucket, the estimated pioneer effect changes sign around this threshold.

[3]Figure 4 reports a nonparametric binned event-study specification, replacing the continuous interaction with five-year event-time bins. The estimated pioneer–non-pioneer gap rises monotonically with technology age, indicating that the linear specification is not masking important nonlinearities.

firms appear much less exposed to the local supply dynamics of pioneer regions, consistent with the idea that they can recruit more broadly and better absorb early-stage hiring frictions.

Taken together, these estimates imply a clear lifecycle in startup hiring conditions. Early in a technology's life, pioneer regions combine concentrated labor demand with relatively scarce applicant supply, producing unusually tight markets for small firms. Over time, however, applicant supply in pioneer regions deepens faster than local demand, so the initial congestion dissipates and eventually reverses. The estimates imply that pioneer regions cross from a relative "talent deficit" to a relative "talent advantage" roughly 20 years after emergence. This reversal is the central empirical fact of the paper: the regions that are initially the most difficult places for startups to hire frontier talent later become the places where startups enjoy the greatest recruiting advantage.

A natural concern is that raw applicant counts may reflect quantity rather than quality. The long-run increase in applicants per listing in pioneer regions could, in principle, be driven by lower-quality candidates entering the market rather than by a genuine deepening of the relevant talent pool. To address this concern, Table 4 also reports estimates using the log of experience-weighted applicants per listing as the dependent variable. This measure weights applicants by their prior years of experience, placing greater weight on more seasoned workers.

The experience-weighted results closely mirror the baseline estimates. For firms with fewer than 500 employees, the estimated pioneer penalty at emergence and the rate of recovery over the lifecycle are nearly identical to those in the unweighted specification. The interaction terms again show that the corresponding lifecycle pattern is much weaker for larger firms. This indicates that the long-run improvement in pioneer-region labor markets for startups does not simply reflect greater applicant volume or low-quality "spam." Instead, it is consistent with the accumulation of experienced human capital in the regions where technologies first take root.

To further assess whether this pattern appears in downstream recruiting outcomes, Table 5 examines platform matches as a secondary outcome. A match is recorded when an applicant applies to a listing and the firm subsequently expresses mutual interest. Matches therefore reflect both worker application behavior and firm-side evaluation.

The match results again point to a startup-specific pioneer lifecycle pattern. First, in the logit specification, which estimates whether a region-technology-year cell generates any matches at all, the interaction of *Pioneer Region* with *Years Since Emergence* is positive and statistically significant for small firms. As technologies mature, pioneer regions become more likely to generate at least one successful match for startups. Second, in the OLS specification, estimated conditional on at least one match, the corresponding interaction is negative. Thus, mature pioneer markets are more likely to generate matches for startups, but conditional on matching occurring, the number of matches per listing declines. Taken together with the rise in applications per listing, this suggests that as pioneer labor markets deepen, startups attract more applicants and become more likely to generate successful recruiting outcomes, but those additional applicants do not translate proportionally into more mutual-interest matches. One interpretation is that thicker pioneer labor markets bring both more applications and more screening or congestion.

These results show that the main findings are not simply about the geographic diffusion of technologies. Rather, they are about which firms are exposed to the resulting labor market frictions. The pioneer lifecycle pattern is concentrated among startups: the regions where frontier technologies first emerge are initially difficult places for small firms to hire, but over time those same regions become a source of recruiting advantage. Larger firms, by contrast, appear substantially less tethered to the local labor supply dynamics of pioneer regions.

4.2.1 Remote and Non-Remote Job Listings

Our main estimates are based on non-remote job listings only. This is the natural baseline for studying geographic frictions, since the importance of local labor supply should be greatest when firms require workers to be physically present. At the same time, the platform also includes listings with remote-work options. Table 6 compares our baseline estimates for non-remote jobs to corresponding estimates for listings that allow some remote work.

If the pioneer lifecycle effect were driven entirely by geographic constraints, we would expect it to weaken substantially for listings with remote-work options. That is what we find: the estimated

effect is attenuated for remote-capable listings relative to the non-remote baseline. This attenuation is consistent with the view that geographic frictions play an important role in shaping firms' access to skilled labor.

At the same time, the effect does not disappear entirely for listings with remote-work options. One plausible explanation is that such listings do not necessarily correspond to fully location-independent work. This interpretation is consistent with evidence that remote work often reflects hybrid arrangements rather than fully location-independent work: Barrero et al. (2021) report that the typical post-pandemic plan for jobs feasible for working from home involves about two days per week at home, with many employers still expecting workers onsite three or more days per week. Accordingly, the platform's remote category may capture only a partial relaxation of geographic constraints rather than their complete removal.

4.2.2 Additional Analyses

We report four additional analyses in the appendix that speak to the interpretation and robustness of the main results. First, Table 8 restricts the sample to the pre-COVID period. The results are, if anything, slightly stronger than in the full sample: for small firms, pioneer regions begin with lower applicant-to-listing ratios and converge more quickly over time, while the large-firm estimates remain much weaker and move in the opposite direction. The same pattern also holds for the experience-weighted applicant-to-listing ratio. These estimates show that the main findings are not driven by pandemic-era shifts in remote work or recruiting behavior.

Second, Table 9 re-estimates the baseline specification without salary controls. The results are nearly unchanged: the all-firm estimates remain close to the baseline, and the core lifecycle pattern continues to be driven by small firms, for which the Pioneer $\times$ Years Since Emergence coefficient remains positive and statistically significant. This indicates that the main results are not an artifact of conditioning on posted compensation, and alleviates the concern that salary is a "bad control" that mechanically absorbs part of the equilibrium response to local labor market tightness.

Third, Table 10 examines posted salary itself as an outcome. Here, we find little evidence that

the pioneer lifecycle pattern is primarily reflected in wages. The coefficients on Pioneer and Pioneer × Years Since Emergence are small and statistically insignificant in the full sample and among firms with fewer than 500 employees. Only among larger firms do we see weak evidence of higher posted salaries in pioneer regions over time. Overall, these estimates suggest that the main applicant-to-listing results are not simply a wage-setting phenomenon.

Finally, Table 11 adds region × technology fixed effects, so identification comes from within-region-technology changes over the lifecycle of a technology. This is the most demanding specification because it absorbs time-invariant differences across local technology markets. In these estimates, the lifecycle pattern remains strong for small firms: the coefficient on Pioneer × Years Since Emergence is positive and statistically significant, while the corresponding estimate for large firms is imprecise and not statistically distinguishable from zero. This reinforces the conclusion that the lifecycle reversal is concentrated among smaller startups rather than reflecting static cross-sectional differences between pioneer and non-pioneer markets.

4.3 Instrumental Variables: OSRD Funding as an Instrument for Pioneer Status

Our primary object of interest is the equilibrium relationship between pioneer regions and labor market tightness over the technology lifecycle. That is, we are not primarily trying to estimate the causal effect of a region becoming a pioneer in some technology. Rather, we ask a more descriptive but economically central question: when a new technology becomes established in a region—for whatever underlying reason—how does labor market tightness in that region evolve as the technology matures? From this perspective, the OLS specification is the natural starting point, because it directly characterizes how applicant supply and labor demand co-evolve in pioneer versus non-pioneer regions over time.

At the same time, one may still ask whether the estimated pioneer lifecycle pattern partly reflects deeper differences between regions that go on to pioneer technologies and those that do not. The interacted fixed effects ($\gamma_{k,t}$ and $\lambda_{r,t}$) absorb common shocks at the technology-year and region-year

levels, but they do not eliminate all possible sources of non-random selection into pioneer status at the region-technology level.

To provide additional evidence on this point, we also estimate an instrumental variables specification that instruments for $\text{Pioneer}_{r,k}$ using the log OSRD funding rate from Gross and Sampat (2023), defined as the log ratio of Office of Scientific Research and Development (OSRD) contracts to total patents at the CBSA $\times$ NBER patent-category level during the 1940s. The relevance of this instrument is motivated by Gross and Sampat (2023)'s argument that OSRD funding priorities were driven primarily by short-run military needs rather than long-run commercial promise, and that research performers were chosen for their ability to deliver high-quality results quickly on novel wartime problems. In their setting, they also show that more and less intensively treated clusters exhibit parallel prewar trends, with divergence emerging only after the wartime shock.

In our setting, the idea is that these wartime investments created durable region-technology research capacity that predicts which regions later pioneer new technologies in related fields. For the IV estimates to admit a causal interpretation here, one must additionally assume that this historical research capacity affects contemporary labor market tightness only through its effect on pioneering status, conditional on the region $\times$ year and technology $\times$ year fixed effects. That exclusion restriction is demanding. Historical OSRD investments may have had persistent effects on region-technology ecosystems—including local institutions, networks, firm formation, and industrial capabilities—that continue to shape present-day labor markets through channels other than pioneer designation itself. Indeed, Gross and Sampat (2023) emphasize that OSRD funding catalyzed long-run agglomeration, firm in-migration, entry, and employment growth, which is precisely why the instrument is relevant but also why its exclusion is not automatic in our context. We therefore interpret the IV estimates as suggestive complementary evidence rather than as a definitive causal design.

Table 7 reports the IV results. Panel A reproduces the OLS estimates for comparison, while Panel B reports the 2SLS estimates. The IV estimates preserve the main qualitative pattern from OLS: pioneer regions begin with tighter labor markets and then converge over the technology lifecy-

cle, with the effect concentrated among smaller firms. The 2SLS coefficients are generally larger in magnitude than their OLS counterparts. Corresponding IV estimates for the additional outcomes, including the experience-weighted applicant-to-listing ratio and the match-to-listing ratio, are reported in Appendix A.2.

First-stage diagnostics and Wu-Hausman tests, reported in Appendix A.2, indicate that the instrument is strong and that OLS and 2SLS estimates differ statistically. At the same time, these diagnostics do not address the exclusion restriction, so we continue to treat the IV results as supportive but secondary evidence and place primary weight on the OLS estimates as the relevant characterization of equilibrium tightness in pioneer regions over time.

5 Discussion

We show that technology diffusion and talent diffusion do not travel together in startup labor markets. Startups across regions quickly begin posting jobs tied to emerging technologies, but workers respond more slowly, producing an early period in which pioneering regions look talent constrained. As technologies mature, the pattern flips: pioneering regions become relatively talent abundant. The sharpest boundary condition is organizational scale. The reversal appears primarily among the smallest startups. The headline implication is that regional advantage in new technologies evolves less through where jobs appear and more through where relevant experience accumulates and remains legible to both workers and firms.

5.1 Implications for Entrepreneurial Strategy

For entrepreneurial strategy, the findings shift attention from location as a static choice to location as exposure to a dynamic labor market. Smaller startups do not compete for talent in an abstract national market. They instead compete in local markets whose thickness, learning opportunities, and credibility of signals change over a technology's life cycle. Early in a technology's emergence, being in a pioneering region can be strategically costly because it places a small startup in the

most intense competition for scarce talent. Over time, the same location becomes an asset because it provides access to a deepening pool of workers who have already learned the technology and whose experience is more likely to be seen by the local market.

The size heterogeneity effect is consequential for theory. The fact that the reversal concentrates among the smallest startups suggests that regional conditions bind most when organizations lack the capabilities and slack to substitute away from local scarcity. Larger startups can buffer local constraints through broader search, stronger employer visibility, more structured hiring processes, or the ability to hold vacancies longer. This implies that regional advantage in talent is not uniformly shared across entrepreneurial firms; it is mediated by organizational scale, and therefore likely shapes entry, survival, and growth in ways that standard agglomeration or diffusion accounts do not capture. Put differently, geography interacts with organizational constraints to determine who can credibly pursue frontier technologies outside early hubs.

5.2 Implications for the Geography of Innovation

A core premise in the economics and geography of innovation is that agglomeration persists because knowledge is localized, spillovers travel through people, and thick markets support matching and recombination. Our results sharpen that logic by distinguishing the diffusion of opportunities (job postings) from the diffusion of capability (applicant behavior). Labor demand linked to new technologies disperses quickly, suggesting that awareness of the opportunity set and the organizational intention to adopt can spread broadly. Yet labor supply remains geographically sticky, particularly for technical roles, and it does so in a way that generates a dynamic regional advantage: pioneer regions initially exhibit tighter markets, then become the places where talent congregates as experience accumulates.

This dynamic helps reconcile two facts that can otherwise seem in tension. First, many technologies appear to diffuse widely in their geographic locus of application, as reflected in job postings outside early hubs. Second, the advantages of pioneering regions often persist. Our evidence suggests that persistence can arise even when demand disperses because pioneering regions build a

stock of experienced workers and a local ecology that makes that experience easier to acquire, observe, and redeploy. In this sense, the mechanism is not simply "more jobs in the hub," but "more accumulated, validated experience in the hub," which can attract subsequent waves of applicants and ease matching frictions for future startups. The attenuation for remote jobs also clarifies a boundary condition: when work can be decoupled from place, geography matters less, consistent with the idea that co-location operates partly through learning and interaction rather than only through information access.

5.3 Limitations

Our study has several limitations that shape the interpretation of the results and point to important opportunities for future research. First, the analysis relies on data from a single online labor market platform that specializes in startup hiring. While this platform plays a central role in matching workers to entrepreneurial firms and provides a rare joint view of job postings and applications, it does not capture all hiring activity. Informal recruiting, referrals, and off-platform hiring may follow different dynamics.

Second, our primary outcome measures capture worker interest rather than realized hires. The applicant-to-listing ratio reflects effective labor supply from the firm's perspective, but we do not observe which applicants receive offers, accept positions, or remain with firms over time. As a result, the analysis cannot directly assess how differences in applicant supply translate into hiring outcomes or subsequent performance. Although prior research links application volume to hiring frictions, future work should examine how these relationships operate in startup settings.

Third, the study focuses on U.S. startups and technologies identified through patent-based measures of emergence. Labor market adjustment may differ in other institutional contexts and different mobility regimes. Similarly, some emerging technologies may diffuse primarily through non-patented channels, such as open-source communities or standards-based ecosystems, which could exhibit different talent-diffusion dynamics.

Finally, the organizing framework is stylized: it imposes stationarity, abstracts from within-

region heterogeneity, and treats knowledge capital as exogenous. More broadly, the analysis is primarily descriptive. While the patterns are consistent with mechanisms based on localized learning, signaling, and organizational constraints, the data do not allow us to causally identify why individual workers choose to apply to certain jobs or why startups adopt particular hiring strategies. The results therefore motivate, rather than settle, questions about the micro-level processes that generate persistent regional differences in talent availability.

5.4 Future Research Directions

The results point to several high priority directions for future research that would deepen our understanding of how labor markets shape entrepreneurial strategy and the geography of innovation. A first and most important direction is to link labor market tightness to realized organizational outcomes. This study measures effective labor supply through applicant behavior, which captures an early and strategically relevant constraint faced by startups. Future work should examine how variation in applicant supply translates into realized hiring and downstream organizational effects. Doing so would clarify when abundant applicant pools ease capability building and when they merely reflect excess interest without improving match quality. Connecting labor market conditions to employee quality and retention would substantially advance research on startup scaling under constraint.

A second priority is to study how startups adapt strategically to persistent talent frictions. The findings suggest that small startups outside pioneering regions may face enduring disadvantages in accessing experienced talent even after technologies diffuse. Future research could examine which organizational responses mitigate these constraints, such as via internal training, alternative job design, equity compensation, or hybrid/remote work structures. Understanding which adaptations expand access to frontier talent versus which simply broaden applicant pools would sharpen theories of entrepreneurial capability formation and competitive positioning.

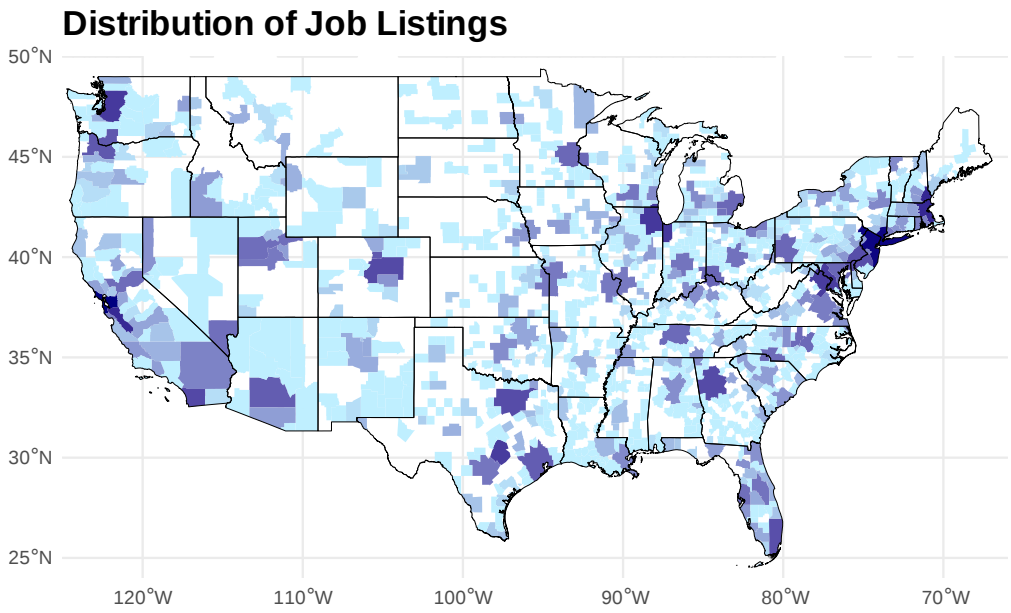
A third promising direction concerns worker careers and the diffusion of experience. The patterns documented here are consistent with localized learning and cumulative skill acquisition, but

future work should trace individual workers over time to examine how experience in pioneering regions affects mobility, career outcomes, and the geographic spread of skills. Such research would clarify whether pioneering regions function primarily as training grounds that seed other ecosystems or as magnets that retain experienced talent, reinforcing regional advantage.

5.5 Conclusion

Jobs travel faster than talent. Startups can pursue new technologies across regions quickly, but the skills needed to execute those technologies accumulate slowly and unevenly. As a result, early technology hubs move from being talent constrained to talent advantaged over time, with the effect concentrated among the smallest startups. The strategic implication is straightforward: regional advantage in frontier technologies is shaped less by where firms post jobs and more by where experience is built. This dynamic helps explain why agglomeration persists even as technologies diffuse, and why entrepreneurial competition remains tightly linked to the geography of human capital.

Figure 1: Distribution of All Job Listings in the Sample.



Job listings on the platform span a broad cross-section of U.S. labor markets, with activity concentrated in major metropolitan areas but present nationwide. This figure maps all job listings in the sample to their corresponding CBSAs, shading regions according to listing density.

Table 1: Technologies and Pioneer Regions.

Panel A: Example Technologies

Technology	Emergence Year	Top 3 Pioneer CBSAs	Job Listings
Component-based software engineering	1997	Colorado Springs, CO; San Francisco-Oakland-Hayward, CA; San Jose-Sunnyvale-Santa Clara, CA	2,624
Data communication	1994	New York-Newark-Jersey City, NY-NJ-PA; San Diego-Carlsbad, CA; San Jose-Sunnyvale-Santa Clara, CA	1,815
Machine learning	2007	Boston-Cambridge-Newton, MA-NH; Los Angeles-Long Beach-Anaheim, CA; New York-Newark-Jersey City, NY-NJ-PA	4,046
Mobile device	1997	Cedar Rapids, IA; San Francisco-Oakland-Hayward, CA; San Jose-Sunnyvale-Santa Clara, CA	1,805
Social networking service	2009	San Francisco-Oakland-Hayward, CA; San Jose-Sunnyvale-Santa Clara, CA	2,358
Targeted advertising	1999	New York-Newark-Jersey City, NY-NJ-PA; Philadelphia-Camden-Wilmington, PA-NJ-DE-MD; San Jose-Sunnyvale-Santa Clara, CA	2,433

Panel B: Top Pioneer CBSAs

Region (CBSA)	Technologies Pioneered	Job Listings
San Jose-Sunnyvale-Santa Clara, CA	545	7,226
San Francisco-Oakland-Hayward, CA	534	37,062
New York-Newark-Jersey City, NY-NJ-PA	438	30,144
Boston-Cambridge-Newton, MA-NH	323	7,904
Los Angeles-Long Beach-Anaheim, CA	281	12,062
Washington-Arlington-Alexandria, DC-VA-MD-WV	173	2,738
San Diego-Carlsbad, CA	172	1,626
Chicago-Naperville-Elgin, IL-IN-WI	171	3,524
Seattle-Tacoma-Bellevue, WA	165	3,306
Dallas-Fort Worth-Arlington, TX	104	1,357

Panel C: Composition of Pioneer and Non-Pioneer Regions

Population	Non-Pioneer	Pioneer
<500k	730 (15.7%)	6 (0.4%)
500k–1M	530 (11.4%)	8 (0.5%)
1M+ (Hub)	3,378 (72.8%)	1,456 (99.0%)

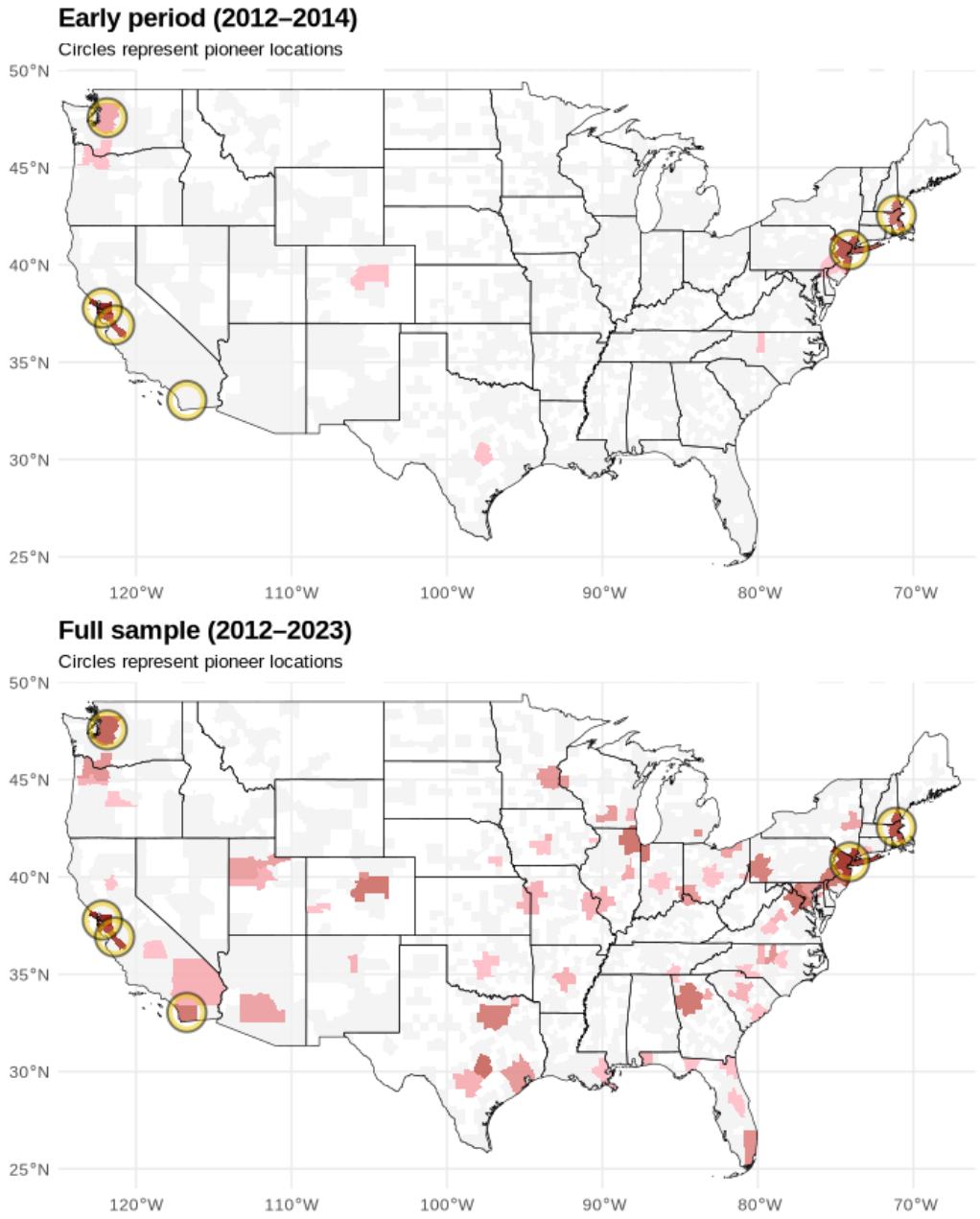
Most technologies are born in a small set of population centers. Panel A reports example technologies used in the analysis, including each technology's emergence year, its top three pioneer CBSAs — defined as the metropolitan areas accounting for the largest share of early patenting activity from Kalyani et al. (2025) — and the total number of job listings associated with that technology in the sample. *Panel B* summarizes the most influential pioneer CBSAs overall by reporting, for each region, the total number of distinct technologies for which it is identified as a pioneer as well as the total number of job listings observed in that region on the platform. Together, the two panels illustrate both (i) the heterogeneity in where new technologies originate and (ii) the concentration of technological pioneering activity in a small set of metropolitan areas. *Panel C* counts region–technology pairs by city size, with percentages computed within pioneer and non-pioneer groups. Pioneer regions are overwhelmingly large metropolitan hubs, whereas non-pioneer regions include both hubs and smaller cities.

Table 2: Sample Composition by Firm Size.

Firm Size (Num. Employees)	Firms (n, %)	Job Listings (n, %)	Applications (n, %)
1-10	11,177 (41.0%)	23,577 (22.0%)	4,080,892 (19.4%)
11-50	8,440 (30.9%)	36,845 (34.4%)	7,542,799 (35.9%)
51-200	3,542 (13.0%)	27,026 (25.2%)	5,599,392 (26.7%)
201-500	960 (3.5%)	8,323 (7.8%)	1,704,971 (8.1%)
501-1000	415 (1.5%)	3,402 (3.2%)	719,493 (3.4%)
1001-5000	322 (1.2%)	3,043 (2.8%)	577,088 (2.7%)
5000+	191 (0.7%)	652 (0.6%)	142,424 (0.7%)
MISSING	2,238 (8.2%)	4,198 (3.9%)	627,267 (3.0%)
Total	27,285 (100%)	107,066 (100%)	20,994,326 (100%)

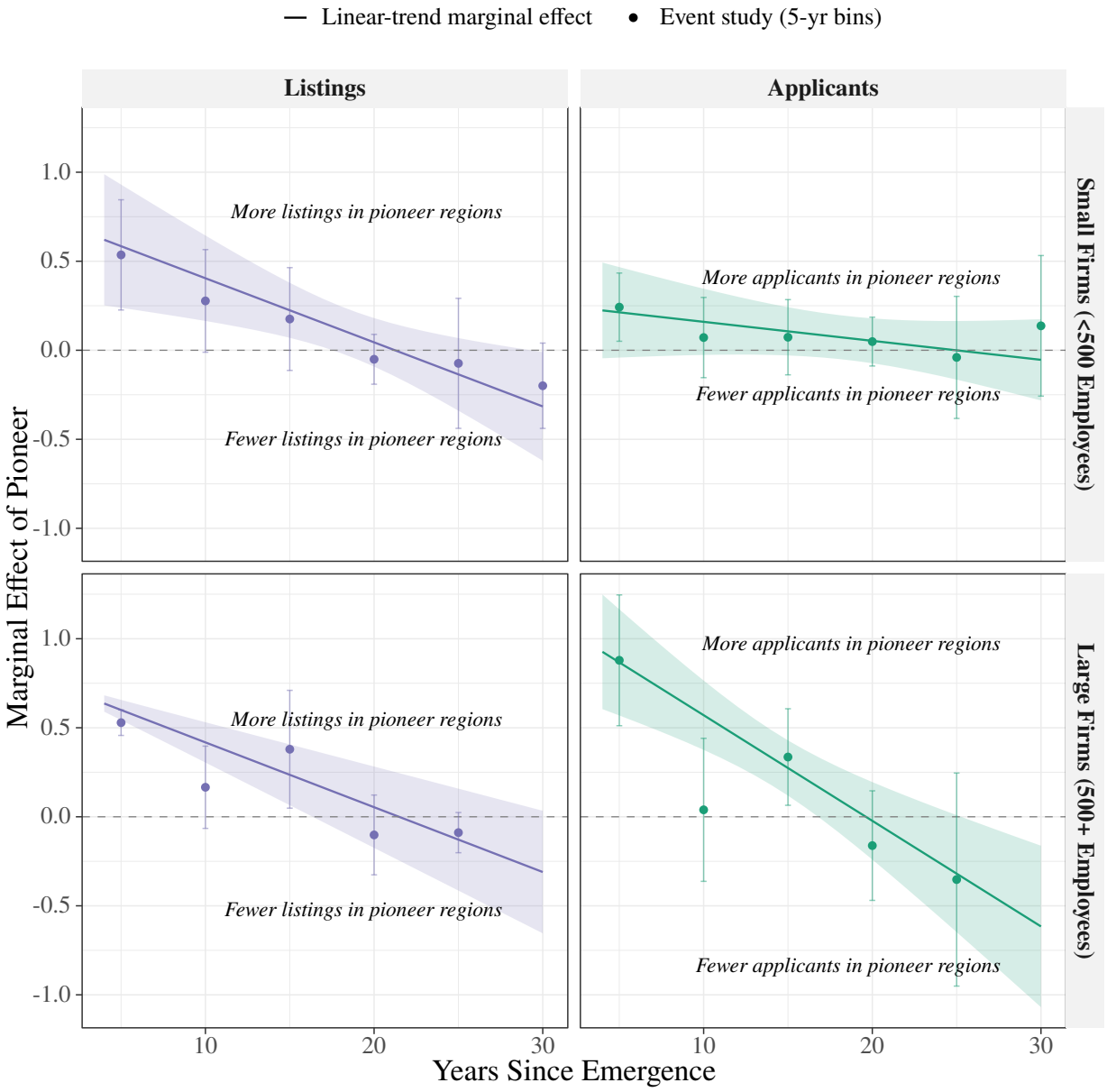
The Wellfound sample spans firms of many sizes but is concentrated among small startups, especially those with fewer than 50 employees. This table reports the distribution of job listings across firm size categories in the Wellfound sample from 2012–2023. Firm size is measured by the number of employees reported at the time of the listing.

Figure 2: Spatial Diffusion of Machine Learning Job Listings.



For a given technology (e.g. Machine Learning), job listings initially cluster in pioneer regions and subsequently diffuse outward to a wide range of secondary markets over time. The top plot maps the geographic distribution of machine-learning (ML) job listings during the early period of the sample (2012–2014). Shaded regions represent CBSAs posting ML-related job openings, with darker shading indicating higher listing volume. Black-outlined circles denote the pioneer locations—the CBSAs that first adopted machine-learning technologies according to Kalyani et al. (2025). The figure shows that ML hiring activity in the early 2010s was highly concentrated in a small number of coastal technology hubs. The bottom plot shows the cumulative distribution of ML job listings by the end of the sample period (2012–2023), using the same shading scale but computed within-panel. While the pioneer regions remain major centers of ML hiring, the geographic footprint of ML job postings expands substantially, with adoption spreading into many interior and southern CBSAs over the decade. The comparison illustrates the diffusion of ML-related labor demand outward from initial technology leaders into a broader set of regional labor markets.

Figure 3: Estimated Pioneer–Non-Pioneer Gaps in Job Listings and Applicants Over the Technology Lifecycle.



Listings diffuses across regions faster than applicants, helping explain the reversal in pioneer-region labor market tightness over time. This figure plots estimated pioneer–non-pioneer differences separately for job listings and applicants from Table 3. The left panel reports the estimated gap in the number of listings, and the right panel reports the estimated gap in the number of applicants. The horizontal axis measures years since technology emergence, and the vertical axis reports the estimated difference between pioneer and non-pioneer regions. Together, the panels show that the early concentration of listings in pioneer regions declines more quickly than the concentration of applicant interest. All estimates come from specifications with technology × year and region × year fixed effects. Shaded bands denote 95% confidence intervals, with standard errors two-way clustered by technology and region.

Table 3: The Pioneer Lifecycle: Job Listings and Applicants.

	$\log(\text{Listings}_{r,k,t})$		$\log(\text{Applicants}_{r,k,t})$		$\log(\text{Experience-Weighted Applicants}_{r,k,t})$	
	<500 Employees	500+ Employees	<500 Employees	500+ Employees	<500 Employees	500+ Employees
$\text{Pioneer Region}_{r,k}$	0.764**	0.782***	0.266	1.164***	0.241	1.340***
	(0.234)	(0.010)	(0.168)	(0.215)	(0.157)	(0.218)
$\text{Pioneer Region}_{r,k} \times \text{Years Since Emergence}_{k,t}$	-0.036**	-0.036***	-0.011	-0.059***	-0.010	-0.071***
	(0.012)	(0.006)	(0.008)	(0.014)	(0.008)	(0.014)
$\log(\text{Median Posted Salary}_{r,k,t})$	0.046*	0.018	0.151***	0.056	0.163***	0.075
	(0.022)	(0.027)	(0.032)	(0.040)	(0.034)	(0.049)
$\text{IQR Posted Salary Ratio}_{r,k,t,s}$	0.256***	0.058+	0.336***	0.095*	0.302***	0.089+
	(0.069)	(0.031)	(0.083)	(0.043)	(0.078)	(0.045)
Technology × Year ($\gamma_{k,t}$)	X	X	X	X	X	X
Region × Year ($\lambda_{r,t}$)	X	X	X	X	X	X
N	28471	7096	28471	7096	26552	6864
R^2	0.904	0.854	0.814	0.735	0.821	0.719
Adj. R^2	0.888	0.825	0.783	0.683	0.792	0.666

$^{+}\ p < 0.1$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

Pioneer regions initially have more job listings and more applicants for emerging technologies, but both gaps shrink as technologies mature. This table reports estimates of pioneer–non-pioneer differences in the number of job listings and applicants over the technology lifecycle. The dependent variables are the log number of listings and the log number of applicants, measured at the region × technology × year level. Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. *Pioneer Region* indicates that region r is a pioneering location for technology k, and its interaction with *Years Since Emergence* captures how pioneer–non-pioneer differences evolve as the technology ages. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 4: The Pioneer Lifecycle: Applicant-to-Listing Ratio.

	log(Applicants-to-Listing Ratio$_{r,k,t,s}$)			log(Experience-Weighted Applicants-to-Listing Ratio$_{r,k,t,s}$)		
	<500 Employees	500+ Employees	All Firms	<500 Employees	500+ Employees	All Firms
Pioneer Region$_{r,k}$	-0.498***	0.382	-0.498***	-0.501***	0.567+	-0.501***
	(0.115)	(0.259)	(0.115)	(0.123)	(0.296)	(0.123)
Pioneer Region$_{r,k}$ × **1**{500+ Employees}$_s$			0.880**			1.068**
			(0.302)			(0.334)
Pioneer Region$_{r,k}$ × Years Since Emergence$_{k,t}$	0.025***	-0.023+	0.025***	0.025***	-0.035*	0.025***
	(0.006)	(0.013)	(0.006)	(0.007)	(0.014)	(0.007)
Pioneer Region$_{r,k}$ × Years Since Emergence$_{k,t}$ × **1**{500+ Employees}$_s$			-0.048**			-0.060***
			(0.014)			(0.016)
log(Median Posted Salary$_{r,k,t,s}$)	0.105***	0.038	0.105***	0.115***	0.052	0.115***
	(0.030)	(0.039)	(0.030)	(0.030)	(0.046)	(0.030)
IQR Posted Salary Ratio$_{r,k,t,s}$	0.080***	0.037	0.080***	0.058**	0.028	0.058**
	(0.022)	(0.029)	(0.022)	(0.021)	(0.031)	(0.021)
Additional Terms			X			X
Technology × Year ($\gamma_{k,t}$)	X	X		X	X	
Region × Year ($\lambda_{r,t}$)	X	X		X	X	
Technology × Year ($\gamma_{k,t}$) × Firm Size (σ_s)			X			X
Region × Year ($\lambda_{r,t}$) × Firm Size (σ_s)			X			X
N	28471	7096	41054	26552	6864	38712
R^2	0.393	0.480	0.431	0.400	0.456	0.433
Adj. R^2	0.291	0.376	0.330	0.302	0.352	0.334

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Pioneer regions begin with tighter labor markets, but this disadvantage fades and reverses over the technology lifecycle, with the effect concentrated among small firms. This table reports estimates of pioneer–non-pioneer differences in the applicant-to-listing ratio over the technology lifecycle. The dependent variable is the log applicant-to-listing ratio measured at the region × technology × year level. Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. *Pioneer Region* indicates that region r is a pioneering location for technology k, and its interaction with *Years Since Emergence* captures how pioneer–non-pioneer differences in labor market tightness evolve as the technology matures. "Additional Terms" indicates that the specification includes the full set of interactions among *Pioneer Region*, *Years Since Emergence*, firm size, and posted salary controls; coefficients on these auxiliary interactions are omitted for brevity. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 5: Alternative Outcome - Matches.

	(Logit) $\mathbf{1}\{\text{Matches}_{r,k,t,s} > 0\}$			(OLS) $\log(\text{Matches-to-Listing Ratio}_{r,k,t,s}) \mid \text{Matches}_{r,k,t,s} > 0$		
	<500 Employees	500+ Employees	All Firms	<500 Employees	500+ Employees	All Firms
Pioneer Region$_{r,k}$	-0.127*	0.199*	-0.127*	0.182+	-0.229	0.182+
	(0.050)	(0.100)	(0.050)	(0.095)	(0.377)	(0.095)
Pioneer Region$_{r,k} \times \mathbf{1}\{\text{500+ Employees}\}_s$			0.327***			-0.410
			(0.092)			(0.360)
Pioneer Region$_{r,k} \times$ Years Since Emergence$_{k,t}$	0.007**	-0.009+	0.007**	-0.010*	0.004	-0.010*
	(0.003)	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)
Pioneer Region$_{r,k} \times$ Years Since Emergence$_{k,t} \times \mathbf{1}\{\text{500+ Employees}\}_s$			-0.016***			0.015***
			(0.005)			(0.004)
log(Median Posted Salary$_{r,k,t,s}$)	0.035***	0.016+	0.035***	-0.046+	-0.078	-0.046+
	(0.009)	(0.009)	(0.009)	(0.026)	(0.063)	(0.026)
IQR Posted Salary Ratio$_{r,k,t,s}$	0.078***	0.019*	0.078***	-0.061***	0.029	-0.061***
	(0.017)	(0.008)	(0.017)	(0.017)	(0.043)	(0.017)
Additional Terms			X			X
Technology $\times$ Year ($\gamma_{k,t}$)	X	X		X	X	
Region $\times$ Year ($\lambda_{r,t}$)	X	X		X	X	
Technology $\times$ Year ($\gamma_{k,t}$) $\times$ Firm Size (σ_s)			X			X
Region $\times$ Year ($\lambda_{r,t}$) $\times$ Firm Size (σ_s)			X			X
N	28471	7096	41054	20182	4413	27938
R^2	0.384	0.512	0.427	0.347	0.497	0.414
Adj. R^2	0.198	0.311	0.243	0.233	0.396	0.306
BIC	68751.721	16095.765	104188.119	78624.568	17068.641	111515.700

$^{+}\ p < 0.1$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

Differences in applicant flows across pioneer and non-pioneer regions also appear in downstream matching outcomes. This table reports estimates of pioneer–non-pioneer differences in matches over the technology lifecycle. A match is recorded when an applicant applies to a listing and the firm subsequently expresses mutual interest. Columns (1)–(3) report logit estimates where the dependent variable is an indicator for whether a region $\times$ technology $\times$ year cell records any matches. Columns (4)–(6) report OLS estimates where the dependent variable is the log match-to-listing ratio, conditional on the cell having at least one match. Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. *Pioneer Region* indicates that region r is a pioneering location for technology k, and its interaction with *Years Since Emergence* captures how pioneer–non-pioneer differences in matching outcomes evolve as technologies mature. "Additional Terms" indicates that the fully interacted specification includes the full set of interactions among *Pioneer Region*, *Years Since Emergence*, firm size, and posted salary controls; coefficients on these auxiliary interactions are omitted for brevity. All specifications include technology $\times$ year and region $\times$ year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 6: Heterogeneity by Remote versus Non-Remote Listings.

	<500 Employees			500+ Employees		
	Non-Remote Listings	Remote Listings	All Listings	Non-Remote Listings	Remote Listings	All Listings
Pioneer Region$_{r,k}$	-0.498***	-0.200*	-0.498***	0.382	-2.088*	0.382
	(0.115)	(0.090)	(0.114)	(0.259)	(0.969)	(0.259)
Pioneer Region$_{r,k} \times \mathbf{1}\{\text{Remote}\}_l$			0.299*			-2.470*
			(0.151)			(1.135)
Pioneer Region$_{r,k} \times$ Years Since Emergence$_{k,t}$	0.025***	0.010*	0.025***	-0.023+	0.063	-0.023+
	(0.006)	(0.005)	(0.006)	(0.013)	(0.042)	(0.013)
Pioneer Region$_{r,k} \times$ Years Since Emergence$_{k,t} \times \mathbf{1}\{\text{Remote}\}_l$			-0.015+			0.086+
			(0.008)			(0.048)
log(Median Posted Salary$_{r,k,t,l}$)	0.105***	0.121***	0.105***	0.038	-0.062	0.038
	(0.030)	(0.031)	(0.030)	(0.039)	(0.040)	(0.039)
IQR Posted Salary Ratio$_{r,k,t,l}$	0.080***	0.103***	0.080***	0.037	0.008	0.037
	(0.022)	(0.027)	(0.022)	(0.029)	(0.027)	(0.029)
Additional Terms			X			X
Technology $\times$ Year ($\gamma_{k,t}$)	X	X		X	X	
Region $\times$ Year ($\lambda_{r,t}$)	X	X		X	X	
Technology $\times$ Year ($\gamma_{k,t}$) $\times$ Remote (ℓ_l)			X			X
Region $\times$ Year ($\lambda_{r,t}$) $\times$ Remote (ℓ_l)			X			X
N	28471	20920	49391	7096	1510	8606
R^2	0.393	0.334	0.369	0.480	0.776	0.526
Adj. R^2	0.291	0.220	0.263	0.376	0.687	0.418

$^{+}\ p < 0.1$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

The pioneer lifecycle pattern is weaker for listings with remote-work options, consistent with geographic frictions playing an important role in the main results. This table compares baseline estimates for non-remote listings with corresponding estimates for listings that allow some remote work. The dependent variable is the log applicant-to-listing ratio measured at the region $\times$ technology $\times$ year level. *Pioneer Region* indicates that region r is a pioneering location for technology k, and its interaction with *Years Since Emergence* captures how pioneer–non-pioneer differences evolve over the technology lifecycle within each listing type. "Additional Terms" indicates that the fully interacted specification includes the full set of interactions among *Pioneer Region*, *Years Since Emergence*, remote status, and posted salary controls; coefficients on these auxiliary interactions are omitted for brevity. All specifications include technology $\times$ year and region $\times$ year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 7: Instrumental Variables: OLS versus 2SLS Estimates of the Pioneer Lifecycle.

	Panel A: OLS			Panel B: 2SLS		
	All Firms	<500 Emp.	500+ Emp.	All Firms	<500 Emp.	500+ Emp.
Pioneer Region$_{r,k}$	-0.295*	-0.503***	0.381			
	(0.138)	(0.116)	(0.258)			
$\widehat{\text{Pioneer Region}}_{r,k}$				-1.389***	-0.943*	-1.283
				(0.385)	(0.447)	(1.048)
Pioneer Region$_{r,k}$ × Years Since Emergence$_{k,t}$	0.014+	0.026***	-0.023+			
	(0.007)	(0.007)	(0.013)			
$\widehat{\text{Pioneer Region}}_{r,k}$ × Years Since Emergence$_{k,t}$				0.064***	0.048**	0.031
				(0.019)	(0.018)	(0.054)
Technology × Year ($\gamma_{k,t}$)	X	X	X	X	X	X
Region × Year ($\lambda_{r,t}$)	X	X	X	X	X	X
N	41054	28471	7096	41054	28471	7096
First-stage F (Pioneer)				413.3	172.5	203.5
First-stage F (Pioneer × YSE)				636.9	338.5	194.6
Wu-Hausman p-value				<0.001	0.077	<0.001
Adj. R^2	0.186	0.290	0.376	0.168	0.286	0.328

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The IV estimates preserve the main pioneer lifecycle pattern from OLS but are larger in magnitude, especially for small firms. This table compares OLS and instrumental variables estimates of pioneer–non-pioneer differences in the applicant-to-listing ratio over the technology lifecycle. Panel A reports the baseline OLS estimates, and Panel B reports 2SLS estimates instrumenting for *Pioneer Region* using the log OSRD funding rate from Gross and Sampat (2023). Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. *Years Since Emergence* measures the age of the technology, and its interaction with *Pioneer Region* captures how pioneer–non-pioneer differences evolve as technologies mature. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

References

James N. Baron and Michael T. Hannan. Organizational Blueprints for Success in High-Tech Start-Ups: Lessons from the Stanford Project on Emerging Companies. *California Management Review*, 44(3):8–36, April 2002. ISSN 0008-1256, 2162-8564. doi: 10.2307/41166130. URL https://journals.sagepub.com/doi/10.2307/41166130.

Jose Maria Barrero, Nicholas Bloom, and Steven J. Davis. Why Working from Home Will Stick. Technical report, April 2021.

Bruce Fallick, Charles A Fleischman, and James B Rebitzer. Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster. *Review of Economics and Statistics*, 88(3):472–481, August 2006. ISSN 0034-6535, 1530-9142. doi: 10.1162/rest.88.3.472. URL https://direct.mit.edu/rest/article/88/3/472-481/57616.

Daniel P. Gross and Bhaven N. Sampat. America, Jump-Started: World War II R&D and the Takeoff of the US Innovation System. *American Economic Review*, 113(12):3323–3356, December 2023. ISSN 0002-8282. doi: 10.1257/aer.20221365. URL https://pubs.aeaweb.org/doi/10.1257/aer.20221365.

David H. Hsu and Prasanna B. Tambe. Remote Work and Job Applicant Diversity: Evidence from Technology Startups. *Management Science*, page mnsc.2022.03391, April 2024. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.2022.03391. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2022.03391.

David H. Hsu, Liinus Hietaniemi, and Torben Hsu. Towards a Strategic Research Agenda on Startup Labor Markets . preprint, SSRN, 2026. URL https://www.ssrn.com/abstract=6194179.

A. B. Jaffe, M. Trajtenberg, and R. Henderson. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, 108(3):577–598, August 1993. ISSN 0033-5533, 1531-4650. doi: 10.2307/2118401. URL https://academic.oup.com/qje/article-lookup/doi/10.2307/2118401.

Aakash Kalyani, Nicholas Bloom, Marcela Carvalho, Tarek Hassan, Josh Lerner, and Ahmed Tahoun. The Diffusion of New Technologies. *The Quarterly Journal of Economics*, 140(2):1299–1365, April 2025. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjaf002. URL https://academic.oup.com/qje/article/140/2/1299/7959830.

William R. Kerr and Frederic Robert-Nicoud. Tech Clusters. *Journal of Economic Perspectives*, 34(3):50–76, August 2020. ISSN 0895-3309. doi: 10.1257/jep.34.3.50. URL https://pubs.aeaweb.org/doi/10.1257/jep.34.3.50.

Alfred Marshall. *The Principles of Economics*. McMaster University Archive for the History of Economic Thought, 1890. URL https://EconPapers.repec.org/RePEc:hay:hetboo:marshall1890.

Enrico Moretti. Local Labor Markets. In *Handbook of Labor Economics*, volume 4, pages 1237–1313. Elsevier, 2011. ISBN 978-0-444-53452-1. doi: 10.1016/S0169-7218(11)02412-9. URL https://linkinghub.elsevier.com/retrieve/pii/S0169721811024129.

Dale T Mortensen. Markets with Search Friction and the DMP Model. *American Economic Review*, 101(4):1073–1091, June 2011. ISSN 0002-8282. doi: 10.1257/aer.101.4.1073. URL https://pubs.aeaweb.org/doi/10.1257/aer.101.4.1073.

Jason Owen-Smith and Walter W. Powell. Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. *Organization Science*, 15(1):5–21, February 2004. ISSN 1047-7039, 1526-5455. doi: 10.1287/orsc.1030.0054. URL https://pubsonline.informs.org/doi/10.1287/orsc.1030.0054.

Michael Roach and Henry Sauermann. Founder or Joiner? The Role of Preferences and Context in Shaping Different Entrepreneurial Interests. *Management Science*, 61(9):2160–2184, September 2015. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.2014.2100. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2014.2100.

Lori Rosenkopf and Paul Almeida. Overcoming Local Search Through Alliances and Mobility. *Management Science*, 49(6):751–766, June 2003. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.49.6.751.16026. URL https://pubsonline.informs.org/doi/10.1287/mnsc.49.6.751.16026.

Prasanna Tambe. Big Data Investment, Skills, and Firm Value. *Management Science*, 60(6):1452–1469, June 2014. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.2014.1899. URL http://pubsonline.informs.org/doi/10.

1287/mnsc.2014.1899.

Lynne G Zucker, Michael R Darby, and Marilynn B Brewer. Intellectual Human Capital and the Birth of US Biotechnology Enterprises. In John R M Hand and Baruch Lev, editors, *Intangible Assets*, pages 185–206. Oxford University PressOxford, February 2003. ISBN 978-0-19-925693-8 978-1-383-03975-7. doi: 10.1093/oso/9780199256938.003.0008. URL https://academic.oup.com/book/52488/chapter/421380019.

A Additional Analyses

This appendix presents robustness checks and extended results that complement the main findings in section 4.

A.1 Robustness

We probe the robustness of the main results along four dimensions: functional form, salary controls, fixed-effects structure, and sample period.

Functional form. The main specification imposes a linear trend on the pioneer–non-pioneer gap over technology age. Figure 4 relaxes this restriction by estimating a nonparametric binned event-study specification, replacing the continuous Years Since Emergence interaction with indicators for five-year bins. The monotonically increasing pattern confirms that the linear specification is not masking nonlinear dynamics: the pioneer–non-pioneer gap in the applicant-to-listing ratio rises smoothly with technology age, consistent with the parametric estimates.

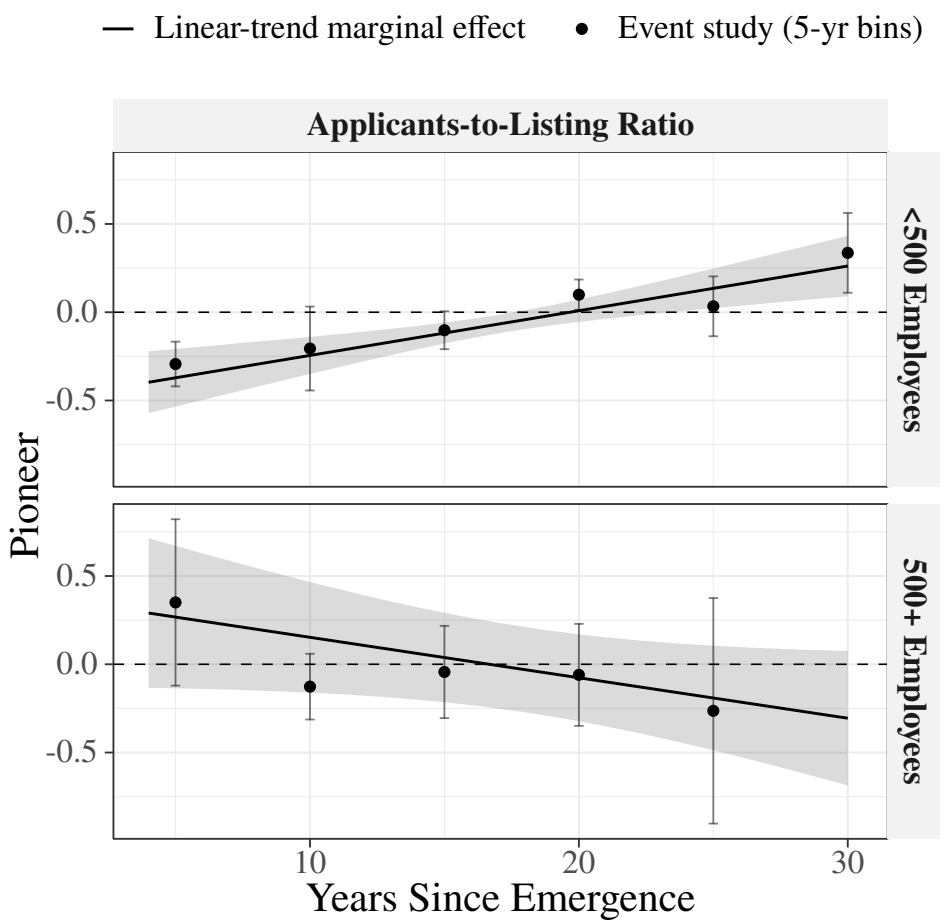
Pre-COVID sample. A natural concern is that the post-2020 shift toward remote work may confound the geographic patterns. Table 8 restricts the sample to 2012–2019. The small-firm lifecycle pattern is unchanged. The large-firm estimates become more precise and move in the opposite direction—large firms in pioneer regions start with an applicant advantage that erodes over time—consistent with the idea that the pandemic-era remote-work shock introduced noise into the large-firm geographic signal.

Salary controls. The main specification controls for the log median posted salary and its interquartile range. Because posted salary may itself respond to local tightness, it is a plausible "bad control." Table 9 re-estimates the specification without salary controls. The results are nearly unchanged, indicating that the lifecycle pattern is not an artifact of conditioning on compensation.

Posted salary as an outcome. Table 10 estimates the same specification with the log of median posted salary as the dependent variable. The pioneer lifecycle coefficients are small and statistically insignificant for the pooled sample and for small firms. Only among larger firms is there weak evidence of higher posted salaries in pioneer regions over time. The main applicant-to-listing results are therefore not simply a wage-setting phenomenon.

Region × technology fixed effects. Table 11 replaces region × year fixed effects with region × technology fixed effects. This is the most demanding specification: it absorbs time-invariant differences across local technology markets entirely, including the pioneer main effect ($\hat{\beta}_1$), and identifies $\hat{\beta}_2$ from within-region-technology variation over time. The lifecycle pattern remains strong for small firms, with the Pioneer × Years Since Emergence coefficient positive and statistically significant. The large-firm estimate is imprecise and not distinguishable from zero.

Figure 4: Nonparametric Pioneer Lifecycle: Binned Event-Time Estimates.



A nonparametric event-study yields the same qualitative lifecycle pattern as the linear specification. This figure reports point estimates and 95% confidence intervals from a nonparametric event-study specification that interacts *Pioneer Region* with five-year bins of *Years Since Emergence*. Panels split the sample by firm size. The estimates trace how pioneer–non-pioneer differences in the applicant-to-listing ratio evolve over the technology lifecycle without imposing a linear trend. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 8: Applicant-to-Listing Ratio - Pre-Covid (Year <= 2019)

	log(Applicants-to-Listing Ratio$_{r,k,t,s}$)			log(Experience-Weighted Applicants-to-Listing Ratio$_{r,k,t,s}$)		
	<500 Employees	500+ Employees	All Firms	<500 Employees	500+ Employees	All Firms
Pioneer Region$_{r,k}$	-0.549***	0.437+	-0.549***	-0.536***	0.623*	-0.536***
	(0.131)	(0.250)	(0.131)	(0.136)	(0.291)	(0.136)
Pioneer Region$_{r,k}$ × **1**{500+ Employees}$_s$			0.986**			1.159***
			(0.300)			(0.328)
Pioneer Region$_{r,k}$ × Years Since Emergence$_{k,t}$	0.028***	-0.026*	0.028***	0.027***	-0.038**	0.027***
	(0.007)	(0.012)	(0.007)	(0.008)	(0.014)	(0.008)
Pioneer Region$_{r,k}$ × Years Since Emergence$_{k,t}$ × **1**{500+ Employees}$_s$			-0.054***			-0.065***
			(0.015)			(0.016)
log(Median Posted Salary$_{r,k,t,s}$)	0.104***	0.040	0.104***	0.111***	0.050	0.111***
	(0.029)	(0.041)	(0.029)	(0.029)	(0.050)	(0.029)
IQR Posted Salary Ratio$_{r,k,t,s}$	0.060*	0.042	0.060*	0.057*	0.028	0.057*
	(0.025)	(0.029)	(0.025)	(0.024)	(0.031)	(0.024)
Additional Terms			X			X
Technology × Year ($\gamma_{k,t}$)	X	X		X	X	
Region × Year ($\lambda_{r,t}$)	X	X		X	X	
Technology × Year ($\gamma_{k,t}$) × Firm Size (σ_s)			X			X
Region × Year ($\lambda_{r,t}$) × Firm Size (σ_s)			X			X
N	21875	6529	33722	21393	6417	32973
R^2	0.390	0.478	0.432	0.396	0.449	0.431
Adj. R^2	0.300	0.387	0.342	0.308	0.353	0.341

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The main pioneer lifecycle pattern is not driven by pandemic-era changes in recruiting or remote work. This table re-estimates the baseline specification using only observations through 2019. The dependent variables are the log applicant-to-listing ratio and the log experience-weighted applicant-to-listing ratio, measured at the region × technology × year level. Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. *Pioneer Region* indicates that region r is a pioneering location for technology k, and its interaction with *Years Since Emergence* captures how pioneer–non-pioneer differences in labor market tightness evolve over the technology lifecycle in the pre-COVID sample. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

A.2 Instrumental Variables: Extended Results

This subsection reports extended IV results that complement Table 7 in the main text.

Table 13 estimates the IV specification for two additional outcomes—the experience-weighted applicant-to-listing ratio and the match-to-listing ratio—split by firm size. The IV estimates preserve the qualitative pattern from OLS across all outcomes: the lifecycle effect is concentrated among small firms, with large-firm estimates statistically insignificant throughout.

Table 12 reports the first-stage regressions. The OSRD funding rate is a strong predictor of both Pioneer and Pioneer $\times$ YSE, with first-stage F-statistics well above conventional thresholds in all specifications. The Wu-Hausman tests (reported in Table 7) reject OLS consistency, consistent with attenuation from measurement error in the binary pioneer indicator or omitted-variable bias from time-varying region-technology confounders. As discussed in subsection 4.3, the exclusion restriction is demanding, and we interpret the IV estimates as complementary evidence rather than a definitive causal design.

Table 9: Robustness: Dropping Salary Controls.

	Baseline	Without Salary Controls		
	All Firms	All Firms	<500 Employees	500+ Employees
Pioneer Region$_{r,k}$	-0.290*	-0.295*	-0.503***	0.381
	(0.136)	(0.138)	(0.116)	(0.258)
Pioneer Region$_{r,k}$ × Years Since Emergence$_{k,t}$	0.014+	0.014+	0.026***	-0.023+
	(0.007)	(0.007)	(0.007)	(0.013)
log(Median Posted Salary$_{r,k,t,s}$)	0.067**			
	(0.025)			
IQR Posted Salary Ratio$_{r,k,t,s}$	0.083***			
	(0.021)			
Technology × Year ($\gamma_{k,t}$)	X	X	X	X
Region × Year ($\lambda_{r,t}$)	X	X	X	X
N	41054	41054	28471	7096
R^2	0.270	0.269	0.391	0.479
Adj. R^2	0.187	0.186	0.290	0.376

$^{+}\ p < 0.1$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

The main pioneer lifecycle pattern does not depend on conditioning on posted compensation. This table re-estimates the baseline specification without salary controls. The dependent variable is the log applicant-to-listing ratio measured at the region × technology × year level. Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. *Pioneer Region* indicates that region r is a pioneering location for technology k, and its interaction with *Years Since Emergence* captures how pioneer–non-pioneer differences in labor market tightness evolve over the technology lifecycle. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 10: Posted Salary as an Outcome.

	$\log(\text{Median Posted Salary}_{r,k,t})$		
	All Firms	<500 Employees	500+ Employees
$\text{Pioneer Region}_{r,k}$	-0.021	0.002	-0.053+
	(0.022)	(0.025)	(0.028)
$\text{Pioneer Region}_{r,k} \times \text{Years Since Emergence}_{k,t}$	0.001	0.000	0.003+
	(0.001)	(0.001)	(0.001)
Technology × Year ($\gamma_{k,t}$)	X	X	X
Region × Year ($\lambda_{r,t}$)	X	X	X
N	41054	28471	7096
R^2	0.408	0.430	0.400
Adj. R^2	0.341	0.335	0.281

$^{+}\ p < 0.1$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

Differences in labor market tightness across pioneer and non-pioneer regions are not primarily reflected in posted salaries. This table examines posted salary as an outcome using the same pioneer lifecycle specification as in the main analysis. The dependent variable is log median posted salary measured at the region × technology × year level. Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. *Pioneer Region* indicates that region r is a pioneering location for technology k, and its interaction with *Years Since Emergence* captures whether compensation in pioneer regions evolves differently as technologies mature. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 11: Robustness: Region × Technology Fixed Effects.

	Baseline	Region × Technology FE		
	All Firms	All Firms	<500 Employees	500+ Employees
Pioneer Region$_{r,k}$	-0.290*			
	(0.136)			
Pioneer Region$_{r,k}$ × Years Since Emergence$_{k,t}$	0.014+	-0.038	0.042***	-0.198
	(0.007)	(0.057)	(0.011)	(0.156)
Pioneer Region$_{r,k}$ × Years Since Emergence$^2_{k,t}$		0.001	-0.001*	0.004
		(0.002)	(0.000)	(0.004)
log(Median Posted Salary$_{r,k,t,s}$)	0.067**	0.009	0.029+	-0.003
	(0.025)	(0.015)	(0.017)	(0.023)
IQR Posted Salary Ratio$_{r,k,t,s}$	0.083***	0.045***	0.033**	0.014
	(0.021)	(0.013)	(0.010)	(0.021)
Technology × Year ($\gamma_{k,t}$)	X	X	X	X
Region × Year ($\lambda_{r,t}$)	X	X	X	X
Region × Technology ($\zeta_{r,k}$)		X	X	X
N	41054	41054	28471	7096
R^2	0.270	0.465	0.671	0.803
Adj. R^2	0.187	0.286	0.501	0.712

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The pioneer lifecycle pattern remains present when identification comes from within-region-technology changes over time rather than cross-sectional differences across markets. This table re-estimates the baseline specification replacing region × year fixed effects with region × technology fixed effects. The dependent variable is the log applicant-to-listing ratio measured at the region × technology × year level. Because region × technology fixed effects absorb the level effect of *Pioneer Region*, the coefficient of interest is the interaction of *Pioneer Region* with *Years Since Emergence*. Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. All specifications include technology × year fixed effects. Standard errors are two-way clustered by technology and region. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 12: Instrumental Variables: First-Stage Results.

	$\text{Pioneer}_{r,k}$			$\text{Pioneer}_{r,k} \times \text{Years Since Emergence}_{k,t}$		
	All Firms	<500 Emp.	500+ Emp.	All Firms	<500 Emp.	500+ Emp.
$\text{OSRD Rate}_{r,c(k)}$	0.046*	0.035*	0.069*	2.144**	2.228**	0.554
	(0.019)	(0.015)	(0.026)	(0.774)	(0.724)	(0.875)
$\text{OSRD Rate}_{r,c(k)} \times \text{Years Since Emergence}_{k,t}$				-0.064	-0.083*	0.061
				(0.042)	(0.038)	(0.052)
Technology × Year ($\gamma_{k,t}$)	X	X	X	X	X	X
Region × Year ($\lambda_{r,t}$)	X	X	X	X	X	X
N	41054	28471	7096	41054	28471	7096
Adj. R^2	0.606	0.572	0.617	0.590	0.555	0.624
Within R^2	0.019	0.012	0.033	0.030	0.023	0.052

Standard errors two-way clustered by technology and region.

Historical OSRD funding strongly predicts which region-technology pairs are later classified as pioneers. This table reports the first-stage estimates for the instrumental variables specification. The endogenous regressors are *Pioneer Region* and its interaction with *Years Since Emergence*, instrumented using the log OSRD funding rate from Gross and Sampat (2023). Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

Table 13: Instrumental Variables: Extended Outcomes by Firm Size.

	Panel A: OLS				Panel B: 2SLS			
	Exp-Wt (<500)	Exp-Wt (500+)	Match (<500)	Match (500+)	Exp-Wt (<500)	Exp-Wt (500+)	Match (<500)	Match (500+)
Pioneer Region$_{r,k}$	-0.504***	0.565+	0.185+	-0.222				
	(0.125)	(0.294)	(0.094)	(0.373)				
$\widehat{\text{Pioneer Region}}_{r,k}$					-1.189*	-1.262	1.003	-2.059+
					(0.534)	(1.143)	(0.862)	(1.061)
Pioneer Region$_{r,k}$ × Years Since Emergence$_{k,t}$	0.025***	-0.035*	-0.011*	0.004				
	(0.007)	(0.014)	(0.004)	(0.004)				
$\widehat{\text{Pioneer Region}}_{r,k}$ × Years Since Emergence$_{k,t}$					0.063**	0.032	-0.041	0.102+
					(0.019)	(0.059)	(0.052)	(0.051)
Technology × Year ($\gamma_{k,t}$)	X	X	X	X	X	X	X	X
Region × Year ($\lambda_{r,t}$)	X	X	X	X	X	X	X	X
N	26552	6864	20182	4413	26552	6864	20182	4413
Adj. R^2	0.300	0.352	0.233	0.396	0.291	0.306	0.219	0.360

Standard errors two-way clustered by technology and region.

The IV analysis yields the same qualitative pioneer lifecycle pattern for additional recruiting outcomes. This table reports 2SLS estimates for additional dependent variables, including the experience-weighted applicant-to-listing ratio and the match-to-listing ratio. *Pioneer Region* is instrumented using the log OSRD funding rate from Gross and Sampat (2023), and its interaction with *Years Since Emergence* captures how pioneer–non-pioneer differences evolve over the technology lifecycle. Columns compare all firms, firms with fewer than 500 employees, and firms with 500 or more employees. All specifications include technology × year and region × year fixed effects. Standard errors are two-way clustered by technology and region. Observations are weighted using the technology-level weights from Kalyani et al. (2025).

B Formal Model

This appendix provides the formal details of the two-region search model referenced in Section 2. We characterize the environment, derive equilibrium wages and tightness under Nash bargaining, and verify the lifecycle reversal result.

B.1 Environment

Regions, workers, and time. The model has two regions, $r \in \{P, N\}$, where P denotes the pioneer region and N the non-pioneer region. There is a single type of skilled worker. Time is discrete, $t = 1, 2, \ldots$, representing years since the technology's emergence.

Timing convention. The model is solved as a sequence of within-period static equilibria indexed by state variables $(K_{r,t}, N_{r,t})$, which evolve exogenously across periods reflecting the diffusion of the technology (declining $K_{P,t}/K_{N,t}$, growing $N_{N,t}$). The Bellman equations characterize within-period equilibrium holding these state variables fixed. The lifecycle reversal (Section B.6) characterizes how equilibrium tightness changes as the state evolves across periods. This sequence-of-equilibria approach is standard in the DMP literature when the stationary block of the model is used to derive comparative statics along a slowly-moving state path.

Two parameters govern within-period payoffs:

- *Separation probability* $\delta_m \in [0,1)$: probability a worker-firm match dissolves each period.
- *Discount factor* $\beta \in (0,1)$.

Define the composite discounting parameter:

$$B \equiv 1 - \beta(1 - \delta_m) \in (0,1]$$

Production technology. Output in region r at time t is:

$$Y_{r,t} = A_t \, K_{r,t}^{\phi} \, N_{r,t}^{1-\phi}, \qquad \phi \in (0,1) \qquad \text{(Production Function)}$$

where A_t is aggregate technology common across regions, $K_{r,t}$ is region-specific knowledge capital, and $N_{r,t}$ is skilled employment. The marginal revenue product is:

$$MRP_{r,t} = (1 - \phi) \, A_t \, K_{r,t}^{\phi} \, N_{r,t}^{-\phi} \qquad \text{(MRP)}$$

Own-type diminishing returns: $MRP_{r,t}$ is strictly decreasing in $N_{r,t}$.

B.2 Matching

Matching function. Following the DMP framework, the flow of new matches is:

$$M_{r,t} = \mu \, U_{r,t}^{\alpha} \, V_{r,t}^{1-\alpha}, \qquad \alpha \in (0,1), \, \mu > 0 \qquad \text{(Matching Function)}$$

where $U_{r,t}$ is applicants, $V_{r,t}$ is vacancies, μ is matching efficiency, and α is the applicant-side elasticity.

Key rates. Define tightness $\theta_{r,t} \equiv U_{r,t}/V_{r,t}$ (applicants per vacancy; high θ = slack market). The vacancy fill rate and job-finding probability are:

$$q(\theta_{r,t}) \equiv \frac{M_{r,t}}{V_{r,t}} = \mu\, \theta_{r,t}^{\alpha} \qquad \text{(Fill Rate)}$$

$$f(\theta_{r,t}) \equiv \frac{M_{r,t}}{U_{r,t}} = \mu\, \theta_{r,t}^{\alpha-1} \qquad \text{(Find Rate)}$$

A higher $\theta_{r,t}$ (more applicants per vacancy) makes filling positions easier (q increasing in θ) and finding a job harder (f decreasing in θ since $\alpha < 1$).

B.3 Value Functions

Firm: filled vacancy. Let $J_{r,t}$ be the value to a firm of a filled job. The firm earns flow profit $MRP_{r,t} - w_{r,t}$ and the match dissolves with probability δ_m. Under the within-period stationarity of the sequence-of-equilibria approach:

$$B\, J_{r,t} = MRP_{r,t} - w_{r,t} \qquad \text{(B1)}$$

Firm: free entry. Firms post vacancies until the expected value of a vacancy equals the per-period posting cost $c > 0$:

$$q(\theta_{r,t})\, J_{r,t} = c \qquad \text{(B2)}$$

Worker value functions. Let $W_{r,t}$ be the value of employment in the focal technology and $UW_{r,t}$ the value of non-employment:

$$W_{r,t} = w_{r,t} + \beta\big[(1-\delta_m)W_{r,t+1} + \delta_m\, UW_{r,t+1}\big]$$
$$UW_{r,t} = b_{r,t} + \beta\big[f(\theta_{r,t})W_{r,t+1} + (1 - f(\theta_{r,t}))UW_{r,t+1}\big]$$

Worker surplus $\Lambda_{r,t} \equiv W_{r,t} - UW_{r,t}$ satisfies:

$$\Lambda_{r,t} = \frac{w_{r,t} - b_{r,t}}{1 - \beta(1 - \delta_m - f(\theta_{r,t}))}$$

B.4 Outside Options and the Learning Premium

Outside options are uniform across regions: $b_{r,t} = \bar{b}_t$. This ensures that $K_{r,t}$ affects tightness only through the firm-side MRP, making the model's comparative statics transparent.

the learning premium (Technology-Specific Learning Premium) enters through workers' continuation values. A worker employed in the pioneer region in period t enters period $t+1$ with an enhanced outside option $\tilde{b}_{t+1} = \bar{b}_{t+1} + \ell^{(t)}\, K_{P,t}$ (compared with $\bar{b}_{t+1}$ for a non-pioneer worker), reflecting technology-specific human capital accumulated through frontier employment. This raises $W_{P,t}$ above $W_{N,t}$: even when current wages are equalized across regions under $\eta = 0$, pioneer employment is strictly preferred in expected value terms because it yields a higher continuation payoff.

Workers therefore direct applications toward pioneers, sustaining applicant flow $U_{P,t}$ even as pioneer vacancy postings $V_{P,t}$ decline and generating the divergence between listings and applicants documented in Section 4.2.

B.5 Equilibrium Wages and Tightness

Wages under Nash bargaining. Under Nash bargaining with worker bargaining power $\eta \in [0,1]$:

$$w_{r,t} = \eta\, MRP_{r,t} + (1-\eta)\, \bar{b}_t$$

Firm surplus is $(1-\eta)(MRP_{r,t} - \bar{b}_t)$ and the free-entry condition gives:

$$q(\theta_{r,t})\, \frac{(1-\eta)(MRP_{r,t} - \bar{b}_t)}{B} = c$$

Taking logs:

$$\log\, \theta_{r,t} = -\frac{1}{\alpha}\left[\log(MRP_{r,t} - \bar{b}_t) + \log(1-\eta) - \log(Bc/\mu)\right]$$

Taking the pioneer-minus-non-pioneer difference, the $\log(1-\eta)$ term cancels exactly. The tightness comparison condition ($\star$) is therefore *invariant to* η: bargaining power does not affect whether or when pioneer markets are tighter or looser. For brevity we report the $\eta = 0$ (competitive) case throughout.

Equilibrium tightness. Setting $\eta = 0$ and combining (B2) with (B1):

$$B\, \frac{c}{q(\theta_{r,t})} = MRP_{r,t} - \bar{b}_t \;\Rightarrow\; \frac{Bc}{\mu}\, \theta_{r,t}^{-\alpha} = MRP_{r,t} - \bar{b}_t$$

Taking logs:

$$\log\, \theta_{r,t} = -\frac{1}{\alpha}\left[\log(MRP_{r,t} - \bar{b}_t) - \log(Bc/\mu)\right] \tag{1}$$

Tightness is strictly decreasing in net firm surplus $MRP_{r,t} - \bar{b}_t$: higher surplus induces more vacancy posting, driving $\theta_{r,t} = U_{r,t}/V_{r,t}$ down and making the market tighter.

Derivation of condition ($\star$). The pioneer market is looser ($\theta_{P,t} > \theta_{N,t}$) iff $\log \theta_{P,t} > \log \theta_{N,t}$, which from (1) requires:

$$\log(MRP_{P,t} - \bar{b}_t) < \log(MRP_{N,t} - \bar{b}_t)$$

Since $b_{r,t} = \bar{b}_t$ is common across regions, this simplifies to:

$$\Delta_r \log MRP_t \;<\; 0 \tag{$\star$}$$

That is, pioneer markets are looser exactly when pioneer MRP falls below non-pioneer MRP. (The net surplus comparison $MRP_{P,t} - \bar{b}_t \gtrless MRP_{N,t} - \bar{b}_t$ and the MRP comparison $MRP_{P,t} \gtrless MRP_{N,t}$ have the same sign for any $\bar{b}_t < \min_r MRP_{r,t}$.) This is the key equilibrium condition of the model.

B.6 Application: The Labor Market Reversal

MRP decomposition. From (MRP):

$$\Delta_r \log MRP_t = \underbrace{\phi\, \Delta_r \log K_t}_{(A)\text{ Knowledge advantage}} - \underbrace{\phi\, \Delta_r \log N_t}_{(B)\text{ Own-type crowding}} \tag{2}$$

Condition ($\star$) says pioneer markets are looser when $\Delta_r \log MRP_t < 0$, i.e. when $(B) > (A)$: $\Delta_r \log N_t > \Delta_r \log K_t$.

Early phase. By the employment dynamics conditions(i), $\Delta_r \log K_0$ is large. By the employment dynamics conditions(ii), $\Delta_r \log N_0$ is small (pioneers have first-mover advantage in employment but the gap has not yet accumulated). Term (A) dominates: pioneer MRP is high, pioneer vacancies are highly productive, and firms accept a low fill rate. Pioneer markets are *tight*: few applicants per listing. the learning premium sustains applicant flow to pioneers.

Later phase. As the technology matures, two forces shift the balance:

(i) *Knowledge advantage erodes* (A falls): The pioneer K advantage declines as the technology becomes more widely understood and non-pioneer firms adopt it.

(ii) *Own-type crowding grows* (B rises): $\Delta_r \log N_t$ increases as pioneers accumulate experienced workers. Diminishing returns reduce the incremental MRP of each additional hire, depressing pioneer MRP below non-pioneer MRP.

Postings shift toward non-pioneers. Applicants continue targeting pioneers for the learning premium (the learning premium), so $U_{P,t}$ falls more slowly than $V_{P,t}$. The pioneer applicant-to-listing ratio $\theta_{P,t}$ rises relative to $\theta_{N,t}$: markets transition from tight to loose.

Proof of the lifecycle reversal. Condition ($\star$) requires $\Delta_r \log MRP_t = \phi(\Delta_r \log K_t - \Delta_r \log N_t) < 0$, i.e. $\Delta_r \log K_t < \Delta_r \log N_t$.

At $t = 0$: By the employment dynamics conditions(i)–(ii), $\Delta_r \log K_0$ is large and $\Delta_r \log N_0$ is small, so $\Delta_r \log K_0 > \Delta_r \log N_0$: pioneer markets are tight.

As $t \to \infty$: the employment dynamics conditions(i) gives $\Delta_r \log K_t \to 0$. the employment dynamics conditions(iii) gives $\liminf(\Delta_r \log N_t - \Delta_r \log K_t) > 0$, so eventually $\Delta_r \log N_t > \Delta_r \log K_t$: pioneer markets are loose.

Existence of t^*: The function $g(t) \equiv \Delta_r \log K_t - \Delta_r \log N_t$ is continuous, starts positive, and is eventually negative. By the intermediate value theorem, there exists a finite t^* where $g(t^*) = 0$, i.e. $\Delta_r \log K_{t^*} = \Delta_r \log N_{t^*}$. □

The crossover $t^* \approx 20$ years corresponds empirically to $\hat{\beta}_1/\hat{\beta}_2 \approx 0.259/0.013$.

B.7 Endogenizing the Separation Rate

The baseline treats δ_m as exogenous. A natural extension allows the separation hazard to rise with market thickness (more poaching opportunities in slack markets):

$$\delta_{r,t}(\theta_{r,t}) = \delta^0 + \rho\, f(\theta_{r,t}) = \delta^0 + \rho\, \mu\, \theta_{r,t}^{\alpha-1}, \qquad \rho \geq 0$$

The effective discounting parameter becomes $B(\theta_{r,t}) \equiv A_0 + A_1\,\theta_{r,t}^{\alpha-1}$ where $A_0 = 1 - \beta + \beta\delta^0 > 0$ and $A_1 = \beta\rho\mu \geq 0$. The equilibrium condition:

$$H(\theta) \;\equiv\; A_0\,\theta^{-\alpha} + A_1\,\theta^{-1} = \frac{c}{\mu(MRP_t - b_t)}$$

is continuous, strictly decreasing in θ, and ranges over $(0,\infty)$, guaranteeing a unique equilibrium for any positive net surplus.

The elasticity of tightness with respect to net surplus becomes:

$$\frac{\partial \log\theta}{\partial \log(MRP - b)} = -\frac{A_0 + A_1\theta^{\alpha-1}}{\alpha A_0 + A_1\theta^{\alpha-1}} \;\in\; \left(-\frac{1}{\alpha},\, -1\right)$$

With exogenous separations ($A_1 = 0$) this equals $-1/\alpha$. With endogenous separations ($A_1 > 0$ and $\alpha < 1$), the magnitude falls strictly between 1 and $1/\alpha$: endogenous churn dampens the tightness response in magnitude but leaves the direction, existence, and timing of the crossover intact.