Competing for Talent:

Large Firms and Startup Growth

By James Bessen (TPRI, BU), Felix Poege (Bocconi, ICRIOS), and Ronja Röttger (TPRI, BU)

Working paper
September, 2024

Abstract: This paper explores the impact of large firms' hiring in local labor markets on the salaries offered by startups and on startup growth and performance. We analyze firm data matched to job advertisements and find strong evidence of "crowding out." A standard deviation increase in the share of ads posted by large firms raises startup pay offers by 5-10% for critical managerial, STEM, and sales jobs, and it reduces expected startup growth by 36%. Crowding is diminished by employee mobility and by spillovers to startups in closely related businesses. It is increased by big firm markups, which may have a large effect on startups. Results are robust to a shift-share instrumental variable strategy. Crowding has important implications for firm strategy, regional policy, and for understanding the slowdown in the aggregate growth of startup firms.

Keywords: startups, labor market crowding, market structure, regional economics
JEL codes: M13, J31, J42, R11, L25

Thanks to excellent research assistance from Erich Denk, comments from Po-hsuan Hsu, Michael Impink, Mike Meurer, Tim Simcoe, Thorsten Grohsjean, and participants in the TPRI and Bocconi seminars, Munich Summer Institute, DRUID, BSE Summer Forum, and the Academy of Management. This work was supported in part through a grant from the Alfred P. Sloan Foundation (Grant #G-2021-16968).

Introduction

The "Talent War" is a popular theme in the business press. Articles discuss how large firms aggressively compete for talent, offering exceptionally high wages for technical jobs. Such competition for talent is an important way workers participate in economic growth. For startups, however, this talent competition is seen as "today's toughest challenge"[1] because new ventures need to recruit high-quality employees to grow (see e.g., J. Chen, Hshieh, and Zhang 2021; Dimmock, Huang, and Weisbenner 2022). Startups may be particularly susceptible to competition from large firms because large firms pay higher wages for the same jobs (e.g. Brown and Medoff 1989), perhaps due to rent sharing, and because large firms tend to recruit more high quality, technically advanced workers that startups need. On the other hand, many startups choose to locate in clusters – which are often anchored around large firms – because they gain from agglomeration benefits (Shaver and Flyer 2000; Sorenson and Audia 2000; Stuart and Sorenson 2003; McCann and Folta 2011; Guzman 2024). Understanding the factors that affect the challenge posed by large firm hiring may be critical to startup strategy. And getting the balance right in a region may be important for policymakers who want to encourage a vibrant entrepreneurial economy.

This paper analyzes how large firms influence startup success through local labor markets. We build a dataset linking job advertisements (help-wanted ads) to small and large firms. Using these data, we show that large firm hiring "crowds out" startups in the labor market and forces them to offer higher wages, which in turn leads to less growth and less startup success. We explore factors that might affect these impacts, including employee mobility and markups. Finally, we explore how different groups of startups might be affected, including those in closely related businesses, who might benefit from spillovers, and firms founded or led by women or minorities, who may lack networks of contacts to aid hiring.

Standard economics tells us that greater demand for labor from any firm, large or small, will tend to raise equilibrium wages. Our empirical analysis, however, concerns the *share* of large firm hiring, not the level. That is, large firm hiring poses a particular challenge to startups, so the share of hiring from large firms matters. We explore why this might be so in a monopsony model of labor market competition between large firms and startups. Successful hiring by startups and large firms depends on the preferences of prospective employees as well as the ability of the firms to offer attractive positions. Some workers might prefer to work at large firms because those firms offer greater security or better amenities; other workers might prefer the opportunities for advancement offered by rapidly growing startups. Firms with higher relative productivity or

[1] See, for example, Devaney, Tim and Tom Stein, "The Talent Wars: Today's Toughest Startup Challenge," ReadWriteWeb, June 11, 2012, and Cade Metz, "Tech Giants Are Paying Huge Salaries for Scarce A.I. Talent," New York Times, Oct. 22, 2017.

markups will be able to offer higher wages. In a Hotelling-type model of labor market competition between a large firm and a startup, productivity and demand shocks affecting large firms lead them to increase local hiring, which increases startup wage premiums and decreases startup profits. The magnitude of the effect, however, depends on a variety of factors including big firm markups, spillovers from large firms, and employee mobility. The magnitude of the crowding effect will also tend to be greater for those types of workers that both large firms and startups both demand such as technologically capable occupations.

To explore the crowding effect and the importance of these different influences, we build a dataset linking job advertisements to startups and large firms and to local labor markets. Using Crunchbase, we identify some 141,117 US startups founded between 2010 and 2020, and we identify large firms as publicly listed firms with over 1,000 employees. We then link the large and small Crunchbase firms to job advertisements listed by Burning Glass (now EMSI). The job ads identify the prospective location for each job, which we link to US commuting zones – our measure of local labor markets. This setup allows us to construct our key explanatory variable, the yearly share of job ads posted by large firms in each commuting zone - the Big Firm Share (BFS). Salaries (or salary ranges) listed on many job ads, as well as the number of ads posted by individual startups - as a measure of firm growth - serve as our outcome variables. Reassuringly, we find that cumulative ads posted are closely related to reported firm employment and various performance metrics including funding, survival, and exit by IPO or acquisition.

A key challenge for the empirical analysis is endogeneity, in particular in the form of omitted variables. Startup decisions to locate in a particular commuting zone, to advertise jobs, and the advertised salaries are influenced by time-varying characteristics of the commuting zone. These same factors may also influence decisions of large firms to hire there, leading to a spurious correlation. We use a shift-share instrumental variable strategy to identify the impact of large firms' hiring independently of these confounders in local labor markets. Because large firms often respond to demand or productivity shocks by increasing hiring in multiple labor markets, we construct an instrumental variable by multiplying each large firm's lagged share of hiring in a labor market by the growth of that firm's hiring in *other* labor markets. Aggregating the predicted share of hiring by large firms in each commuting zone, we obtain a reasonably strong instrumental variable that plausibly meets the exclusion restriction.

We empirically test whether startups react to labor demand by large firms, and show that large firms "crowd out" startups in the labor market forcing them to offer higher wages. We relate the salary posted in job advertisements to the big firm share, while controlling for detailed job characteristics requested, such as education, experience, and occupation. We find that a standard deviation increase in the big firm share of ads in a commuting zone raises salaries offered by all firms by about 2% and salaries of startups by about 4%; for select occupations (managers, STEM, and sales), startup salaries increase by 5-10%. We find similar effects with measures of big firm residual wages, which is the excess salary that big firms in a commuting zone-year offer relative to a benchmark based on job characteristics. Finally, we find that increased labor mobility in a commuting zone abates the effect of the big firm share.

We trace the effect of big firm hiring on startup growth and startup success, and find large negative effects as well. To capture startup growth and success, we study job-advertising behavior. Following Lee and Kim (2023), we assume that startups initially begin in an experimental or developmental phase, and then, when they are ready to grow, they begin posting job ads. Many startups never proceed to the growth phase. With this in mind, we first estimate the hazard that a startup will post its first job ad in a given year (extensive measure). We then estimate the expected number of ads that a startup posts, conditional on it having begun advertising jobs (intensive measure). Combining these estimates, we calculate that a standard deviation increase in the big firm share of job ads decreases the total expected number of ads that a startup will post by 36 percent using our instrumental variable estimates (47 percent using OLS estimates). Further, we verify that the startup wage premium - excess salaries paid by startups in a commuting zone relative to job characteristics – negatively affects startup growth. Crowding does, indeed, appear to substantially reduce startup growth.

Spillovers and markups affect crowding by changing the relative productivities of startups and big firms. Our model predicts that spillovers from large firms raise the productivity of startups that operate businesses that are close to the large firm, but big firm markups increase the revenue productivity of large firms without raising startup productivity. Consequently, spillovers to proximate firms should soften the crowding-out effect whereas markups should increase it. To test this empirically, we first measure markups using data from Compustat and the U.S. Bureau of Labor Statistics (BLS), and find that markups are negatively related to posted salaries. Second, we measure business proximity by analyzing the texts of business descriptions. Our markup-adjusted proximity measure is the scaled cosine similarity of the texts (a weighted average over the different firms in the commuting zone) divided by the markup. Consistent with the model, spillovers from large firms appear to substantially benefit startups with proximate business models and reduce the impact of large firm hiring on startup wage premiums and ad growth. This finding contrasts with labor market "Kill Zones" (Kamepalli, Rajan, and Zingales 2020). While big firm acquisitions might diminish venture capital financing for proximate startups, this occurs against a background where big firms seem to strongly benefit proximate startups.

Startup growth in job ads is only one factor in a larger process that determines growth in startup employment. For example, the networks of founders present a major channel of growth that job ads do not capture. This is an important aspect, as for example the literature consistently finds that women/minority-led firms tend to have smaller networks of contacts (Abraham 2015; Neumeyer et al. 2019; Kyrgidou et al. 2021; Woodwark, Wood, and Schnarr 2021). We first investigate whether cumulative ads through 2020 track employment figures listed in Crunchbase in 2021. Both numbers closely track each other, and we find that for each job ad that was run, a startup had about 1.5 employees. Finding a number larger than one is consistent with the idea of network-based hiring in initial growth stages, as well as the presence of other channels in addition to job advertising. However, it turns out that this relationship is different for startups that are founded or led by women or minorities, which only have about 1.16 employees for each ad posted, suggesting that these firms rely more heavily on advertising. Moreover, we find that these firms are significantly more affected by crowding, putting them at a distinct disadvantage.

This paper contributes the first theoretical and empirical analysis of the interaction between large firm hiring and startup pay, growth, and performance. We find that large firms' hiring in a local labor market substantially impacts startup pay offers, especially in key managerial, STEM, and sales occupations. Furthermore, we find that large firm hiring is associated with substantially slower startup growth and performance. Moreover, we identify key factors influencing these crowding effects: spillovers, markups, and employee mobility. We also identify key differences in the labor market activity of firms founded or led by women or minorities.

These findings are important not only for entrepreneurs and managers at large firms, but also for policymakers at the regional and national level. Regional policymakers frequently aim to lure large firms with subsidies to spur the local entrepreneurial economy. We reinforce the argument that such "smokestack chasing" brings the danger of crowding out entrepreneurs (Chatterji et al., 2014). Yet, a better understanding of the factors that affect labor market crowding may provide complementary policies to lessen the impact of crowding.[2] On a macroeconomic scale, the relationship between startups and large firms is particularly relevant as the slower growth of startups appears to account for a substantial portion of the slowdown in aggregate productivity growth (Decker et al. 2020). Haltiwanger and Davis (2014) proposed labor market frictions as a possible reason, and large firm crowding might create a particularly important labor market friction. This friction may be especially important both because markups have been rising (De Loecker, Eeckhout, and Unger 2020) and because the large firm share of employment has been rising—the aggregate US share of employment at firms with more than 1,000 employees rose from 37% in 1988 to 44% in 2022 (data from the Current Population Survey).

Our contribution is related to several literatures where we add insights about the importance of large firms on startup growth and performance. We contribute to the research on startup formation and startup growth and the availability of skilled human capital in the local labor market by highlighting the role of large firms (Glaeser and Kerr 2009; Dahl and Klepper 2015; Karahan, Pugsley, and Şahin 2024; Balsmeier et al. 2020; Guzman and Stern 2020; Chen, Hshieh, and Zhang 2021; Dimmock, Huang, and Weisbenner 2022). We contribute to the literature on the startup wage premium (Burton, Dahl, and Sorenson 2018; Kim 2018; Babina et al. 2019; Sorenson et al. 2021), finding that large firm hiring is a major determinant of startup pay premiums.

[2] For example, Agrawal et al. (2014) find that regions are most innovative when they include both a sizable population of small firms and large labs. Consistent with this, we find that large firms provide important spillover benefits to startups.

While other papers look at the effects of large firms, we add insights about the specific effects of large firms on startup growth and performance. For one, many startups arise as spin-outs from big firms (Klepper 2009; Sevilir 2010; Babina and Howell 2022). Consistent with our results, studies on the arrival of "Million Dollar Plants" find that these large plants tend to raise wages, suggesting labor market crowding (Greenstone and Moretti 2003; Greenstone, Hornbeck, and Moretti 2010; Gupta 2022; Slattery and Zidar 2020; Qian and Tan 2021; Bhardwaj et al. 2023). Gupta (2022) in particular considers the impact of large plants on startup creation. However, relative to these prior contributions, we add specific insights on the effects of crowding on startups and startup growth and performance after entry.

Our research also highlights the importance of looking at the specific role of large firms within clusters and not just the role of clusters themselves. An extensive literature looks at the interaction between geographic clusters and entrepreneurship (see review in Chatterji, Glaeser, and Kerr 2014). Regional interactions in clusters can both aid and harm startups to different degrees (Pe'er and Keil 2013). Some startups, especially new startups or startups with few resources, tend to benefit from more accessible resources, including skilled labor, social ties, suppliers, and customers (Shaver and Flyer 2000; Sorenson and Audia 2000; Stuart and Sorenson 2003; McCann and Folta 2011; Guzman 2024). In contrast, clusters can harm the performance of more established startups because of more local competition (Shaver and Flyer 2000; Sorenson and Audia 2000; Stuart and Sorenson 2005). Fairlie and Chatterji (2013) found that a crowded labor market in Silicon Valley was associated with lower startup entry. Our research suggests that the effects of clusters might vary depending on the role of large firms within clusters; it pays to "look under the hood." On the one hand, large firms are responsible for much of the agglomeration benefits of clusters; on the other, large firm crowding can reduce startup growth.

Finally, we contribute to the literature on minority/women-founded startups, which finds that such startups face difficulties at entry (Fairlie, Robb, and Robinson 2022; Fairlie and Robinson 2023; Bates, Bradford, and Seamans 2018; Bennett and Robinson 2023) and in the long run, partially due to a reduced labor supply (Kacperczyk, Younkin, and Rocha 2023) or differences in networks (Abraham 2015). We contribute evidence that firms founded or led by women or minorities rely more on formal labor markets in hiring.

Model

Basic Setup

A growing literature argues that labor markets are imperfect (Bhaskar, Manning, and To 2002; Manning 2021; Card 2022). One possible reason for monopsonistic labor markets is that workers may have heterogeneous preferences for working at different employers. For instance, workers could have different preferences for working at startups compared to large firms. Large established firms may offer better amenities and long-term security; startups have a greater risk of failing but also a greater possibility of rapid growth that potentially promises greater personal opportunity and greater reward. Following Bhaskar, Manning, and To (2002), we can analyze these heterogeneous preferences in a Hotelling type model. In a literal Hotelling model, workers might face transport costs to commute to work and they might differ in their commuting distance from different employers. Here, we interpret "distance" to be in preference space and a corresponding "transport cost" to reflect the disutility of distance.

To keep things simple, we consider a duopsony; while our empirical application includes many firms, basic insights to guide empirical specifications can be derived from a two-firm model. In each labor market, there is a startup firm, labelled *s*, and a large incumbent, designated *L*. These firms offer wages w_s and w_L, respectively. In addition to wages, the jobs they offer differ in other ways and workers differ in their preferences regarding these non-wage characteristics. Let the mass of workers in each market be 1 and assume that workers can be ranked on their relative preferences by $x, 0 \leq x \leq 1.$ We assume that in each labor market workers are uniformly distributed over $x.$

A worker at location x will experience a disutility of tx accepting a job with firm *s* and a disutility of $t(1-x)$ accepting a job with firm *L*. The value of t represents the disutility of a unit of distance in preference space, the "transport cost," and it depends on labor market conditions that can vary from market to market. In other words, the worker's utility from working at firm *s* is $u_s = w_s - tx$ and at firm *L* is $u_L = w_L - t(1-x)$. Define the break-even location as $\tilde{x}: u_s = u_L$. Then, for a given set of wage offers, all workers with $x < \tilde{x}$ will prefer to work at the startup and all workers with $x > \tilde{x}$ will prefer the big firm. Solving $u_s = u_L$ for x we get firm *s*'s share of the labor supply,

$$S_s(w_s, w_L) = \tilde{x} = \frac{w_s - w_L + t}{2t}$$

and the big firm share is (see figure)

$$S_L(w_s, w_L) = \tilde{y} = 1 - \tilde{x}.$$

Then let the firm's revenue per worker in the given occupation be $p_i, i = s, L.$ Given the supply functions, firm profits per worker are, respectively,

$$\pi_s(w_s, w_L) = (p_s - w_s)S_s(w_s, w_L), \quad \pi_L(w_s, w_L) = (p_L - w_L)S_L(w_s, w_L).$$

In the Appendix, we solve for the Nash equilibrium in wages.[3] The Nash equilibrium wages and profits are functions of the revenue productivities, $p_i, i = s, L.$ In our empirical analysis, local labor markets will have different revenue productivities for different firms at different times. While we do not observe

[3] We assume that $p_0 - p_1 + 3t > 0$ in order to guarantee an interior solution.

revenue productivities,[4] we do observe the market shares of big firms. In the model, the equilibrium big firm market share is a function of revenue productivities,

$$\hat{y} = \frac{p_L - p_s}{6t} + \frac{1}{2}. \tag{1}$$

The big firm share varies with the relative revenue productivity differences of the big firm and startup. Using (1), we can use $\hat{y}$ as a proxy for productivity differences and write the equilibrium startup wage and profit as

$$\hat{w}_s = p_s - 2t(1 - \hat{y}) \tag{2}$$

$$\hat{\pi}_s = 2t(1 - \hat{y})^2. \tag{3}$$

So, variation in productivity across local labor markets and time produces variation in the big firm share and corresponding variation in startup wages and profits. These two equations express the "crowding out" hypothesis: a larger share of big firm hiring in a labor market is associated with higher startup wages and lower startup profits, all else equal.

These equations allow us to treat the big firm share of hiring as our central independent variable. Our key empirical assumption is that p_L is subject to exogenous productivity and demand shocks. These translate into changes in $\hat{y}$ that we measure, letting variation in p_s fall into the error term (we assume t is constant). Because of endogeneity concerns—for example, p_s might be correlated with p_L—we use instrumental variable estimation so that variation in p_L is orthogonal to the error term. Furthermore, note that p_L can be

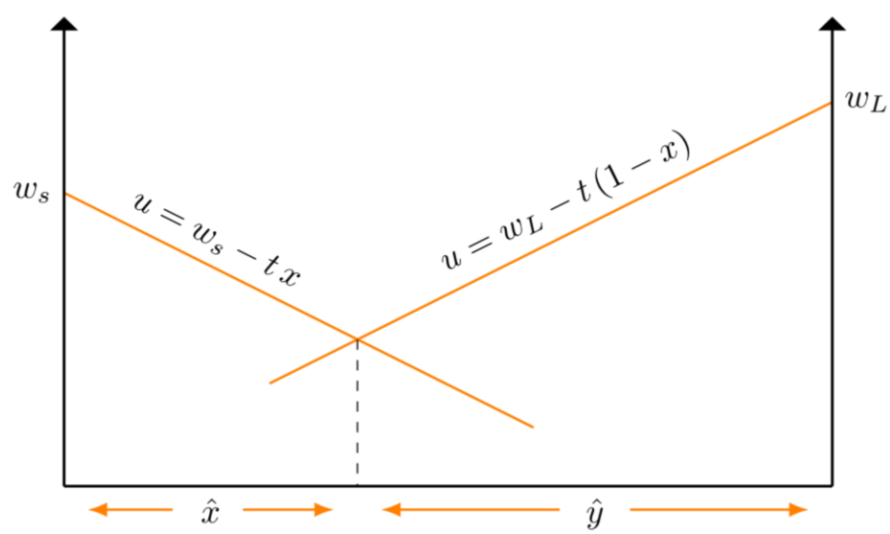


Figure 1: Competition on the Hotelling line

[4] We observe revenue per employee for Compustat firms, however, that does not provide the revenue productivity of specific occupations at the firm. We also proxy productivity with wage residuals.

decomposed into two terms: $p_L = \omega m$, where ω is the output per worker and m is the markup (taking the marginal cost as numeraire). This means that we cannot interpret a rise in the big firm share of ads as a simple reflection of big firm productivity; it might, instead, reflect big firm markups. Below we will explore this possibility. In our empirical analysis, we interpret profits as being indicative of startup performance generally, relating it to firm hiring growth, financing, and successful exits.

Spillovers and markups

So far, the model does not consider spillovers. When a big firm increases its activity in a local labor market, there may be additional knowledge spillovers to startup firms, leading to greater revenue productivity of the startup, p_S. But this benefit applies only to certain firms within the local labor market, specifically those that are in closely related industries. Hence, we model the spillover as a product of the big firm's output per worker, ω, and the proximity of the startup, λ, so that startup revenue per worker is a base level plus a spillover term, $p_S = \overline{p} + \lambda\omega$. Taking $p_L = \omega m$, startup revenue productivity is then

$$p_S = \overline{p} + \frac{\lambda}{m} p_L.$$

Plugging this into (1),

$$\hat{y} = \frac{p_L\left(1 - \frac{\lambda}{m}\right) - \overline{p}}{6t} + \frac{1}{2}.$$

This equation captures the idea that beneficial spillovers from big firms can ameliorate detrimental crowding. Greater spillovers, realized by startups in proximity to the big firm (high λ), reduce the big firm share and, from (2) and (3), they reduce startup pay and increase startup profits. On the other hand, higher big firm markups dilute this effect. Markups increase big firm revenue per worker without a corresponding increase in spillovers. Hence, holding proximity constant, higher markups tend to raise startup wages and lower their profits.

It is widely recognized that markups represent market power that can reduce output and raise prices. Here, we identify another implication of markups: they can reduce the profits and growth of innovative startups via labor market crowding. There are two effects. Markups increase big firm revenue productivity, directly increasing the crowding effect, and markups further reduce the benefits of spillovers to closely related startups.

Distance and mobility

So-called transport costs feature prominently in equations (1) through (3). Yet transport costs might be arguably related to regional policy differences. For instance, to the extent that changes in the startup and big firm shares of hiring involve workers changing employers—that is the extent to which new hires are recruited from other firms—then regional policies that affect job-to-job mobility will affect *t*. For instance, policies that enhance labor market frictions such as noncompete agreements or occupational licensing restrictions might reduce job mobility and hence increase transport costs, *t*.

Of course, many other factors affect transport costs and transport costs likely influence firm entry and the local equilibrium firm shares. However, a little analysis predicts how changes in the big firm share are affected by transport costs/mobility:

$$\frac{\partial\,\hat{y}}{\partial t} < 0 \text{ if } p_L > p_s, \quad \frac{\partial^2 \widehat{w}_s}{\partial t\,\partial \hat{y}} > 0, \quad \frac{\partial^2 \hat{\pi}_s}{\partial t\,\partial \hat{y}} < 0.$$

The intuition is that higher transport costs "soften" competition, shrinking the big firm share when $p_L > p_s$. And while softer competition can decrease the *level* of startup pay and increase the level of startup profits for a given y, the effect is to increase the marginal variation of wages and profits with respect to y. An increase in t steepens the utility curves in Figure 1 so that an increase in y is associated with a larger increase in startup wage and a larger drop in startup profit.

To summarize, we will take the following predictions to the data: First, we predict that large firms crowd out startups in the labor market. The big firm share of ads and big firm markups are positively associated with startup pay and negatively associated with startup growth and performance. Second, we predict that greater spillovers from large firms to proximate startups reduce the crowding effect. Third, we predict that higher job mobility in a local labor market also reduces the crowding effect.

Data

Source data

Our key datasets are from Crunchbase, Compustat, and Burning Glass Technologies (BGT, now called EMSI). Crunchbase, a large database of corporations, includes data on firm founding, financing rounds, exits, headquarters location, and business description. Crunchbase primarily focuses on collecting information on high-growth startups, but to cover all parties in financing and exits, its coverage extends to a much broader range of firms. This sample is more selective than the universe of newly created firms.

Compustat, from Standard & Poors, provides information on publicly listed firms, including employees, industry codes, the headquarter's location, and data to compute markups.

BGT scrapes, deduplicates, and cleans the near universe of online job advertisements to achieve a coverage of 60-70% of all job openings and 80-90% of openings requiring a bachelor's degree or more (Carnevale, Jayasundera, and Repnikov 2014). The data include the advertised salary, firm name, industry, occupation, required education and experience, and geographic location of the job. We omit job advertisements that are missing a firm name, as these are largely ads by recruiters. Our sample spans from January 2010 to December 2020. Depending on the analysis, we focus on ad-level information or aggregate the ads by firm and year. We assign geographic locations to commuting zones.

We attempt to link each ad from BGT to a corresponding entry in Crunchbase. To do so, we design a flexible algorithm that – based on training data – calculates a match score for pairs of entries, taking into account the similarity of name, location, industry, size, and year. We validate the match using manually collected validation data, which we also use to set a threshold for the minimum match score we accept. In this set of accepted algorithmic matches, around 90% coincided with the manual choice (precision); and of the manual choices, around 90% could be recovered (recall).[5] We describe the matching algorithm in greater detail in Bessen et al. (2023). After additional refinement, we will make the data available to the scientific community.

We similarly create a link between Compustat and Crunchbase based on (a) stock market ticker information (b) probabilistic links based on the same model used for the Crunchbase-Burning Glass match and (c) manual corrections.

Key variables and measures

We report summary statistics in the Appendix.

Startups. Startups are young firms, many of whom aim to achieve high growth rates. We identify startups in Crunchbase as firms that were founded since 2010, corresponding to the coverage of the labor demand data (see below). However, the founding date itself is not sufficient to identify startups, as companies resulting from M&A and spinoff activity look the same according to this criterion. We manually inspect the largest 'startups' to exclude such cases.[6]

Local labor markets (commuting zones). We assign all startups to local labor markets identified by commuting zones. Commuting zones are defined to delineate labor markets by grouping counties with strong within-cluster and weak between-cluster commuting ties using a hierarchical cluster analysis and the Census Bureau's "journey to work" data. The county groupings of commuting zones were first defined beginning in the 1980s and are slightly adjusted every 10 years. We selected the commuting zones defined in 1990 and utilize Dorn's crosswalk file to map counties to commuting zones (Autor and Dorn 2013). For

---

[5] Much of the incorrect matches is related to cases where the firm information was ambiguous and the human coders refrained from a choice, whereas the algorithm always

[6] In firm-level analysis, which Crunchbase has typically been used for in the literature, the comparatively small number of such cases is likely inconsequential. On the other hand, such firms are often responsible for a large quantities of ad-posting, which can manifest as outliers and could systematically bias analyses in ad-level datasets.

more details on the construction of commuting zones, see Tolbert and Sizer (1996). We assign all startups to the location listed in their Crunchbase profile in 2021, via geocoding and geographic assignment.[7]

Large firms. We are concerned with the influence of large firms. While other definitions are possible, we follow much of the literature and focus on firms that were listed on the stock market and recorded in Compustat. We consider a firm as large if it was publicly listed between 2010 and 2020 and had at least 1000 employees.[8] Similar to startups, we assign large firm headquarters to commuting zones according to the location listed in Compustat. This approach excludes privately held large firms, so our results might not pertain to *all* large firms. Nevertheless, public large firms make up the bulk of employment of large firms and their influence is interesting by itself. We note that the large tech firms, GAFAM (Google, Apple, Facebook, Amazon, Microsoft) only account for 3.0% the ads of large firms. Our big firm effects are not mainly about the big tech firms.

Industry and business proximity. We leverage business descriptions to understand the business positioning of firms in industries as well as relative to each other. We transform the business descriptions into a numerical representation using BERT models specifically trained to capture text similarity (Reimers and Gurevych, 2019).[9] In a first application, we use the numerical representation to assign startups to NAICS industries, which we primarily use for fixed effects. With this, we can overcome a limitation of Crunchbase, where the categorization of firms is comparatively coarse and not unique. We achieve this classification by assigning each business description to the most similar description of a NAICS code.

Second, we calculate cosine text similarity scores between startups and large firms, which we deploy as a measure of business proximity.[10] We use this measure to proxy for the influence of direct spillovers – ideas, technologies, product markets, network-based recruiting – from big firms in the same commuting zone and startups. Presumably, spillovers will be strongest between large firms and nearby startups. In line with the model, we adjust the business proximity by the big firm markups, but results are robust to using the standard proximity as well.

Firms founded/led by minorities or women. We use information contained in Crunchbase as to the gender or ethnicity of the startup's leadership team. Crunchbase contains tags for startups that are

---

[7] Crunchbase lists the current location, which may deviate from historical information. In Crunchbase, we can test for location changes by using older vintages for those firms already previously included. Descriptively, between 2013 and 2021, we record a move between commuting zones for 4.5% of startups (3.2% between 2019 and 2021). Vintages between 2013 and 2019 are not available to us. We provide robustness checks excluding startups with recorded, see below.

[8] Our results are robust to changing the size of "large firms" (see Robustness Checks).

[9] Standard BERT models are inappropriate for calculating similarity scores. We verify that the Crunchbase business descriptions are within the length of texts admissible by BERT (512 tokens).

[10] With this approach, we follow the spirit of Guzman and Li (2023)'s differentiation score, where they measure the average website similarity to the closest five public firms.

women-founded or led, or minority-founded or led. However, this variable is missing for most observations, and it might be biased because it is self-reported. Crunchbase also provides a listing of persons associated with startups and their gender. We combine these two measures to construct an extended identifier that allows us to identify many more observations (see comparison in the Appendix).

Commuting zone level mobility. We calculate commuting zone-level job mobility using data from the Current Population Survey. We measure job mobility as the share of workers who report that their current employer is not the same employer they had the previous month.

Markups. Following the literature (Hall 1988; De Loecker, Eeckhout, and Unger 2020), we estimate markups as the revenue elasticity of variable factors times revenue over cost of goods sold. Revenue and cost of goods sold come from Compustat. The output elasticities are calculated from 2-digit industry BLS data as an index of the cost share of labor and intermediates. For each commuting zone each year, we calculated "big firm markups" as an average for big firms weighted by their ads in that labor market. As an alternative to the BLS-based estimates of output elasticities, we use estimates of markups derived by DeLoecker, Eeckhout, and Unger (DEU markups) using production function estimation (see Robustness Checks).

Uses of help wanted ads. We use help wanted ads to measure 1) the relative labor demand of startups and large firms, 2) the growth of startups, and 3) the wage premium that startups pay. We capture labor demand of large firms as the share of help wanted ads posted in each commuting zone each year.

To use job ads to capture startup growth, we follow Lee and Kim (2023). Conceptually, help-wanted ads capture neither the founding team nor hiring of early-stage employees. Instead, they pinpoint the time when growth of a startup exceeds the possibility to fill positions using information available from networks, i.e., we capture startup scaling. This is consistent with observations in the data, where the propensity to post ads is low in the first years, but continually rises over time (see Figure A1 in the Appendix). In startup-level analyses, we focus on ads from the headquarter commuting zone only. With this, we focus on the correspondence between the local incidence of the big firm share. Further, ads in other parts of the US are suspect to a greater degree of measurement error. Below we show that the number of home commuting zone ads is closely related to actual firm employment.

For our analysis, we infer startup wage premiums from advertised salaries. If a salary range is reported, we retain the midpoint. Startup ads advertise a salary only in 29% of cases. If reporting would be related to the big firm share, it may impact our analysis results. In Robustness Checks, we show that variation in the share of ads reporting salaries on the commuting zone level largely is driven by variation across occupations, which we control for – so that omitted variable bias is unlikely. Further, the results are robust to including a control for the share of ads advertising a salary.

Empirical Methods

Implementation

In the main analysis, we take the existence of startup projects as given and discuss the salary offered by ads posted by these startups (ad-level analysis), or we analyze the incidence of ad posting and other startup-level outcomes (startup panel analysis). We assume that large firms are subject to exogenous productivity and demand shocks that influence their hiring in local labor markets. These changes then possibly affect startup pay and growth. In equations (2) and (3), our model shows how startup wages and startup profits (hence growth) are related to the large firms' share of hiring.

Thus, our main variable of interest is the 'big firm share' ($\hat{y}$ in the model, BFS in what follows), which is the share of ads in a commuting zone posted by large, influential, and often highly productive firms. For ad counts *N*, firm *f*, year *t*, and commuting zone *c*, we define BFS as:

$$BFS_{ct} = \sum_f Big_f \cdot s_{fct}\,, \quad \text{where } s_{fct} \equiv \frac{N_{fct}}{N_{ct}}$$

We standardize the variable by dividing by the standard deviation. In the ad-level analysis, we aim to understand how big firms affect startups differentially from other firms. To do so, we adapt equation (2) to estimate

$$\ln w_{it} = \beta_1 Startup_i + \beta_2 BFS_{ct} + \beta_3 Startup_i \cdot BFS_{ct} + \gamma X_{it} + \mu_c + \delta_t + \epsilon_{it} \tag{4}$$

where we include commuting zone fixed effects (μ), year fixed effects (δ), as well as additional fixed effects and control variables (*X*). We expect $\beta_2, \beta_3 > 0$.

In the startup-panel analysis, we focus only on startup outcomes; an interaction specification is not required, and we directly model the effect of big firm share on startup-related outcomes.

Identification

The main identification concern is that unobserved factors may affect the success of local startups and big firms at the same time. In general, startups have little control over the big firm share, and commuting zone fixed effects allow us to focus on time-varying factors. However, local shocks – for example, variations in local demand for products or local labor supply – may induce a spurious relationship between the big firm share and startup salaries and outcomes.

To alleviate these concerns, we focus on variation from big firm-level shocks, such as variation in worldwide demand that affects the labor demand of large firms while only minimally affecting local startups. The argument relies on the idea that many decisions about firm growth and hiring are taken centrally (Hazell et al. 2022; Hjort, Li, and Sarsons 2020) and, given the increasingly wide spread of local branches of big firms, disseminate through geographic space (Hsieh and Rossi-Hansberg 2023). In parallel, local shocks in other locations propagate through firms' establishment networks (Giroud and Mueller 2019; Bena, Dinc, and Erel 2022). We argue that changes to national big firm labor demand affect local startups' growth and labor demand decisions only through changes in the local labor demand of the big firm.

Specifically, we instrument the big firm share in a commuting zone-year cell with the predicted big firm share based on big firm hiring in other commuting zones.[11] The prediction is based on the last year's BFS and the firm-level growth rate between years, leaving out the focal commuting zone.[12] Consequently, we construct the following shift-share instrumental variable:

$$IV_{ct}^{BFS} = \sum_f \mathrm{Big}_f \ \hat{s}_{fct} = \sum_f \mathrm{Big}_f \frac{g_{f,-c,t}^*}{g_{-c,t}^*} s_{fc,t-1}$$

$$IV_{ct}^{BFS} = \sum_f \mathrm{Big}_f \hat{s}_{fct} = \sum_f \mathrm{Big}_f \frac{g_{f,-c,t}^{\star}}{g_{-c,t}^{\star}} s_{fc,t-1}$$

where Big_f is an indicator for big firms, g_f^* is the firm-level growth in other commuting zones, and g* is the national growth of ad-posting.[13]

To justify our instrumental variable, we assume that shocks to the (global) labor demand of big firms $g_{f,-c,t}^*$ are random, in contrast to their location decisions $s_{fc,t-1}$, which may be driven by local factors that are common with startup activity. The assumption of exogenous shocks and endogenous shares follows Borusyak et al. (2022) and justifies our use of shift-share instrumental variables to identify effects. The instrument relevancy is strong; a clear linear relationship exists between the commuting zone-level big firm share and the instrument (see Table A.2 in the Appendix). The first stage F-statistic of 75 or greater in the default specification is large relative to conventional thresholds.

Throughout the paper, we cluster standard errors at the level of the instrument variation or higher, which, in our case is the commuting zone.

---

[11] As a robustness check, we provide instrumental variable results based on variation in hiring in the commuting zone of the big firm's headquarter instead of the whole-firm variation in hiring (see Robustness Checks). Further, to ensure that the instrument variation is driven by big-firm level shocks rather than the lagged big firm shares, we tested whether the IV regressions are robust to including lagged big firm shares. Results are similar, albeit with less precision.

[12] All regressions include commuting zone fixed effects, so that we focus on changes in (predicted) BFS over time, not on the level.

[13] With this approach, we focus on changes in the intensive margin of hiring of the big firms and abstract from changes due to the new arrival of big firms in a commuting zone. We winsorize the firm-level growth factors $g_{f,-c,t}^*$ at factor 3 to avoid noisy predictions resulting from extreme changes in smaller firms.

Findings

Crowding

Table 1 explores the impact of the big firm share of hiring on advertised salaries as in equation (4). Column 1 regresses log salaries (times 100 to show percentage changes) against the share of ads posted by big firms in each commuting zone each year. The regression controls for education, experience, experience squared, and part-time jobs, with fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone. A standard deviation increase in the big firm share of ads in a commuting zone is associated with an increase in advertised salaries of all firms of about 2 percent, all else equal. Column 2 shows a similar estimate using instrumental variable estimation.

These findings show that large firm hiring tends to crowd all firms in the labor market to some extent. The remainder of the table explores the specific effect on startups. Existing evidence on startup wages is mixed. Some studies show that startups pay wage premiums while others find wage discounts (Kim 2018; Burton, Dahl, and Sorenson 2018; Babina et al. 2019; Sorenson et al. 2021). Column 3 shows about a 9 percent premium for startups in advertised salaries after controlling for the abovementioned factors. Columns 4 and 5 estimate the crowding effect for startups compared to other firms by interacting the big firm share with startup status, using OLS and IV estimation, respectively. A standard deviation increase in the big firm share is associated with a 4 percentage point increase in startup salaries as opposed to a 2 percentage point increase in the salaries of other firms. Also note that in these regressions, aside from the crowding effect, startups now pay only 2-3 percent more than other firms.

Why might the effects of crowding be greater for startups? For one, startups may be fundamentally more vulnerable, but startups may also hire more in those occupations that have greater overlap with big firm hiring and/or occupations with higher "transport" costs, t. Column 6 shows a regression on the sample of startup firms where the big firm share is interacted with dummy variables for different groups of occupations. Managerial/financial jobs, STEM, and sales jobs all have higher coefficients, from 5 to 10 percent. These three categories account for 46 percent of startup help-wanted ads. Put differently, a higher big firm share of ads for all occupations in a given local labor market and year is associated with startups paying higher wages specifically for managerial/financial jobs, STEM, and sales jobs.[14] Thus, crowding is greater in specific occupations.

These findings suggest a substantial relationship between big firm revenue productivity and startup pay. The big firm share of advertising provides a proxy for relative revenue productivity in an extensive

---

[14] Results are robust to alternatively using big firm shares on the commuting zone-year-*occupation group* level, which even allows including commuting zone-year fixed effects.

dimension (equation 1). However, other factors might influence big firm share. We can corroborate our analysis by focusing on an intensive measure using a proxy variable directly related to big firm revenue productivity. The model suggests that big firm wages will rise in response to productivity shocks, all else equal.[15] To derive an intensive measure, we estimate a Mincer wage equation[16] on log salaries to calculate a residual for each ad and then take the mean residual for big firms in each commuting zone each year. We standardize this variable by dividing by its standard deviation. While this variable likely has greater measurement error than the big firm share, we should expect it also to be positively associated with startup wage premiums.[17]

Table 2 reports salary regressions using big firm residuals and other variables that help corroborate our model. Column 1 finds that a standard deviation increase in the big firm residual is associated with a 1 percent increase in salary offers for all firms. We exclude big firms from the estimation sample to avoid a spurious correlation between big firm wages and residuals. Column 2 shows the instrumental variable estimation using the same instrument as in Table 1 (predicted BFS should be correlated with BF residuals). While the first stage regression is not very strong (F statistic of 11.2), the coefficient is larger. Column 3 interacts the big firm residual with the startup dummy. As in Table 1, we find that startups pay a significantly higher wage premium as before.

We can also decompose the big firm residual (which is a measure of log salary) into a productivity and a markup component because, as above, $p_L = \omega m$. That is,

$$\ln productivity = BFresidual - \ln markup.$$

Column 4 shows the regression with both terms, which have been standardized. These coefficients are somewhat higher than the coefficient for the combined big firm residual in Column 1. The crowding effect appears to arise equally from productivity shocks as from changes to markups, consistent with the model. To check the robustness of this finding, we repeat the regression using markups estimated from production functions by DeLoecker, Eeckhout, and Unger (2020). Results are similar but larger.

Crowding arises not just from productivity advantages of big firms, but also from markups. We can gain some sense of the magnitude of the markup coefficient with a back-of-the-envelope calculation. DeLoecker, Eeckhout, and Unger estimate that markups of Compustat firms have risen from 1.21 to 1.61 from 1980 to 2016. That corresponds to about two standard deviations in our big firm residual, corresponding to a very substantial increase in startup pay if we extrapolate our regression to their sample.

---

[15] In the Appendix, $\hat{w}_L = \frac{2p_L + p_S}{3} - t$.

[16] We regress log salary against controls for education, experience, experience squared, and part time jobs, with fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone.

[17] Measurement error might be greater because the Mincer residual is estimated on a sub-population of those ads that list salary information and might thus be subject to sampling variance.

Finally, Column 5 explores the effect of job-to-job mobility on crowding. Job mobility may restrict job changing and thus increase effective transport costs. We find a significant negative relationship between log salary and employment mobility of workers times the big firm share as predicted by the model.

To summarize, we find evidence that a greater share of hiring by large firms in a local labor market plausibly causes a significant increase in wage premiums, especially for startups. These wage effects appear to be larger for select occupations, and they are closely related to higher wages paid by big firms and to big firm markups. The wage effects are diminished in commuting zones where employees switch jobs more often.

Startup growth in advertising

The model implies that a higher share of big firm hiring should also be associated with poorer profits for startups. This happens for two reasons: 1) the crowding effect means that startups will pay higher wages, reducing margins, and 2) greater big firm share implies smaller startup share, hence smaller output. The effect of crowding on profits can be sizeable—a 5 percent increase in wages for a startup that earns a 15 percent operating margin will substantially reduce expected profits, reducing investment, and hence slowing growth. Although we do not observe startup profits, we infer that more profitable startups will be more likely to grow, all else equal. And a growing startup will hire more workers, hence it will place more help-wanted ads, all else equal. In addition, crowding may increase labor market search frictions, slowing the rate at which startups can hire.

We can begin exploring the relationship between big firm share of ads and startup growth by estimating how the mean number of ads posted each year by startups is associated with changes in the big firm share. Following Lee and Kim (2023), we assume that startups take some time to develop their business. Once they demonstrate that it is profitable, they begin hiring, typically placing ads to do so (we will consider other channels of hiring below). Startups that don't make that transition fail, and most startups do not make that transition.[18] Because of this qualitative difference, we estimate ad growth in two phases in our startup sample. First, we use a linear probability model of the hazard of placing the first ad of the form

$$p_{it} = \alpha_1 BFS_{ct} + \beta_1 X_{it} + \epsilon_{it}, \quad t \leq t_0$$

where p_{it} equals 1 the first year that startup *i* advertises and zero otherwise; the sample excludes observations past that first year, t_0. X_{it} is a vector of fixed effects for firm age, firm, and industry x year. The second estimation is a similar linear model of the number of ads the firm posts in a given year, n_{it}, conditional on it having posted its first ad,

$$n_{it} = \alpha_2 BFS_{ct} + \beta_2 X_{it} + \epsilon_{it}, \quad t \geq t_0.$$

[18] There will, of course, be some startups that advertise and then fail. To the extent we can trust Crunchbase's data on closures, that number appears to be small. Of all firms that posted at least one job ad, only 2.6% are reported to have closed by 2021.

Because Crunchbase data on firm closures is not highly accurate, this approach largely eliminates the downward bias that would arise if we estimated the unconditional number of ads posted. With both estimates in hand, the mean number of ads posted per firm is the probability of placing the first ad times the expected number of ads posted conditional on having placed at least one ad. The marginal effect of an increase in the big firm share can be calculated from α_1 and α_2.

The first column of Table 3 explores the relationship between startup wages and the likelihood of posting an ad. The explanatory variable is the mean Mincer wage residual for the startup, calculated as described above, and standardized (divided by its standard deviation). This captures the effect of crowding on wages as well as other factors that might boost wages. We find that a higher startup wage residual is significantly associated with a smaller likelihood that a startup will post an ad. Columns 2 and 3 use the big firm share, instead, as a measure of crowding, performing OLS and IV estimations respectively. Here the effects are statistically significant and quite substantial. In the OLS regression, a standard deviation increase in the big firm share of ads decreases the likelihood of advertising by 34 percent; in the IV regression, the reduction is 25 percent.

Columns 4 and 5 use the intensive measure as the dependent variable, namely the number of ads the startup places each year beginning with the first year it places an ad.[19] Here a standard deviation increase in the big firm share is associated with a 19 percent decrease in ads per year with the OLS estimate; the IV estimate is a 14 percent decrease. Using the instrumental variable estimates, the combined effect of a standard deviation increase in the big firm share, multiplying the reduced ad hazard by the reduced expected number of ads run, is a 36% reduction in the expected number of ads;[20] for the OLS estimates the reduction is 47%. Hiring by large firms in a local labor market appears to substantially increase the wages that startups pay and substantially decrease their ad growth.

Importantly, these effects appear to have some significant heterogeneity depending on how close the business of the startup is to the large firms in its labor market. In the model, startups that are closer to the local large firms should gain from spillovers that come from greater large firm activity. These spillovers raise revenue per worker which diminishes the crowding effect, so that the impact on both startup pay and advertising is diminished. These spillovers to startups could arise from knowledge developed by the big firms, or because big firms create networks of customers that increase the market for proximate startups, or because

[19] These regressions winsorize the top 1 percent of the dependent variable. A small number of startups have very large employment growth, typically fueled by large infusions of debt or investment. Such business models appear atypical of our sample and so we treat these firms as outliers. Winsorization does not substantially change our OLS coefficient, but it does affect our IV estimate.

[20] The expected number of ads is $E[ads] = P[advertising] * E[ads \mid advertising]$. A standard deviation increase in BFS decreases $P[advertising]$ by .25 and $E[ads \mid advertising]$ by .142, so the decrease in $E[ads]$ is 1 – (1 - .25)*(1 - .142) = 0.36. Similarly for the OLS estimates.

big firms develop deeper markets for the workers and inputs that startups need. Table 4 looks at the big firm effects on wages and advertising of startups by differences in the markup-adjusted measure of business proximity. In the Appendix we derive a regression specification

$$\pi_{it} = \beta_\pi y_{c,t} + \delta \frac{\lambda}{m} y_{c,t} + \gamma X_{i,c,t} + \epsilon_{i,c,t},$$

where λ is a measure of business proximity of the startup to big firms, m is mean big firm markup for the commuting zone, and $X_{i,c,t}$ consists of control variables and fixed effects.

We rank startups by their markup-adjusted proximity to big firms and divide them into quartiles. Column 1 repeats the salary regression of Table 1, Column 1, interacting the big firm share of ads with a dummy for quartile group.[21] The crowding effect is about seven times larger for distant compared to close startups and the difference is statistically significant (F-tests of the null hypothesis that the coefficients are equal are shown in the last row). Columns 2 and 3 add the proximity interaction to the OLS regressions similar to Table 3, Columns 2 and 4 respectively. The reduction in advertising hazard and count associated with an increase in the big firm share of ads is substantially less for close startups than for distant ones and these differences are statistically significant. While large firm labor market activity still impacts close startups, the effects are much weaker, implying that crowding is partially offset by substantial positive externalities from large firm activity. It might be the case that large firm acquisitions inhibit the growth of proximate startups (Kamepalli, Rajan, and Zingales 2020; Koski, Kässi, and Braesemann 2023), but proximate startups also appear to benefit from substantial positive spillovers from large firms in the same labor market.

Startup performance

Does the reduced level of advertising that we observe with increased big firm share actually represent decreased growth of employment and worse startup performance? For several reasons, advertising might diverge from actual employment. For one thing, not all jobs are advertised; some new hires are found through word-of-mouth or other networks. Table 5 explores the relationship between startup employment in 2021 and the cumulative number ads posted over the previous 10 years with controls for firm age in 2021. The 2021 Crunchbase "snapshot" includes an intervaled measure of employment; we use the mean of the interval as our dependent variable. Column 1 shows that cumulative ads posted is highly correlated with actual employment, but that there are about 1.45 employees per each job posted. Note that if startups would only hire via ad postings, then we would expect a coefficient close to 1. This suggests that startups do indeed hire through networks and other channels.

---

[21] Because we only measure proximity for startups, the sample is just startups. We group the middle quartiles into one group along with separate groups for the top and bottom quartiles.

We find some support for this hypothesis by looking at the difference in the relationship between employees and ads for startup firms that are founded or led by minorities or women. A number of researchers argue that women and minorities have smaller networks or less social capital to draw on (Neumeyer et al. 2019; Kyrgidou et al. 2021; Woodwark, Wood, and Schnarr 2021). If so, we should expect a closer relationship between employees and cumulative ads for these firms. Column 2 shows that these firms have only 1.16 employees per ad posted, implying much smaller use of networks in hiring.

If firms led by minorities and women rely relatively more on help wanted advertising, it raises the question whether these firms are more affected by labor market crowding. Columns 3-5 explore this possibility by repeating the log salary regression (Table 1, Column 1), the ad hazard regression (Table 3, Column 2), and the ad count regression (Table 3, Column 4), adding an interaction with firm type. We test whether the coefficient of the big firm share is equal for minority/women led firms. In all three cases, the interaction term is larger for firms founded/led by minorities or women. For the salary regression, the difference is economically and statistically significant, but it is not statistically significant for the ad regressions. These firms appear to be at a disadvantage because they rely more heavily on advertising and are more adversely affected by crowding.

Finally, Table 6 shows that the cumulative ads measure not only is closely related to employment growth, but also to a variety of metric of startup performance, namely: 1) whether the startup ever received funding, 2) the number of funding rounds, 3) the cumulative amount of funds raised, 4) whether the startup had an operating website in 2021 (a measure of survival), 5) whether the startup is reported as having closed, 6) whether the startup was acquired, and 7) whether it had an IPO. The regressions include controls for firm age. The cumulative number of ads correlates strongly with all of these measures (negatively for closure), suggesting it is a reasonable measure of firm performance.

Robustness Checks

We perform a variety of robustness checks. First, an increased big firm share may induce firms to post wages more selectively, which may lead to the increased average posted salaries observed for commuting zones with larger big firm share. In Table C.1, we address the concern and show that results on posted salaries in ads are robust to controlling for the share of posted salaries on a commuting zone-year level.

We show that the results are robust to an alternative definition of large firms and startups. First, in Table C.2, we increase the threshold for large firms to 10,000 from the default 1,000. In unreported tests, we verified that the results are consistent for lower thresholds. Second, in Table C.3, we show that results are robust when focusing on tech startups, which we define as belonging to a Crunchbase category in Software, AI, IT, Internet services, Messaging or Platforms.

We show that results are robust to an alternative instrumental variable and the exclusion of relevant subgroups, s. In Table C.3, we construct a variable following section "Identification" that, instead of leave-one-out variation in whole-firm hiring, uses only variation in ad posting in the commuting zone of the big firm's headquarter (except for the headquarter's commuting zone itself). While the instrument relevancy declines, results are in line with the baseline results.

We show that results are robust to subsamples and variables definitions that are less likely to be affected by endogenous choices of startups. For this, we leverage Crunchbase snapshots from earlier years, most importantly 2013, but also 2019. In Table C.3, we first exclude startups that moved between 2013/2019 and 2021. Second, we show that results are consistent when only focusing on startups also included in the 2013 snapshot. It is especially reassuring that the proximity analysis of Table 4 is consistent when using business descriptions of startups from 2013. In these analyses, the set of startups and their orientation is less likely to be affected by the large firms.

In Table C.7, we show that the proximity analysis of Table 4 is robust to alternative proximity definitions. In Table 4, we use the markup-adjusted business proximity, which is the (cosine) similarity-weighted big firm share relative to the overall big firm share and the big firm markup. Results are similar when using the absolute business proximity (similarity-weighted big firm share) or the relative business proximity (similarity-weighted relative to overall big firm share).

We also check our identification of whether firms are founded or led by minorities or women. We identify these firms using different data items from Crunchbase: Crunchbase allows firms to self-report whether they are founded led by blacks, Hispanics, or women and Crunchbase also provides its own listing of the gender of firm founders and CEOs. We combined these to create a variable that is 1 if any of the Crunchbase variables indicate a minority or female founder/CEO and 0 otherwise. In the Appendix (Table C.4) we compare the different components. In Table C.5 we check the regressions in Table 5 using only the self-reported identifiers. The results are similar, but the composite identifier produces much larger regression samples and more statistically significant results.

Finally, one concern is that our estimates of markups are based on an assumed equality between cost shares of input factors and their output elasticities. There are reasons why cost shares might diverge from output elasticities, and there might also be measurement errors in the BLS industry aggregates we use to calculate cost shares. Alternatively, we can use output elasticities derived from production function estimation by DeLoecker, Eeckhout, and Unger (2020) for Compustat firms until 2016. As a robustness check, we repeat the regression of Table 2, Column 4, using the BLS-based markup measure for this period and the DeLoecker, Eeckhout, and Unger (DEU) markups. The coefficients using the DEU markups are similar but slightly larger.

Conclusion

Our analysis shows that large firm hiring can crowd out startups in the labor market with substantial negative impacts on startup growth. Greater hiring by large firms raises wages for startups, especially in managerial/financial, STEM, and sales occupations. While higher wages provide beneficial rewards for talent, they substantially slow the growth of startups, likely diminishing startup performance and the prospects of building a vibrant startup ecosystem. Crowding especially disadvantages firms led or founded by women and minorities.

Regional policymakers frequently seek to attract large firms with tax and other incentives, and we show that in doing so, they face a tradeoff. Large firms can increase agglomeration economies in a region, but crowding might diminish or even reverse these benefits. However, policy can also alter the tradeoff between agglomeration benefits and crowding out. For one, policies that improve employee mobility, such as non-enforcement of employee non-compete agreements, might diminish crowding. More important, our analysis highlights the importance of regional specialization. To the extent that firms in a region specialize in a limited number of industries, spillovers will likely limit the effects of crowding. This rationale provides one reason for policymakers to encourage specialization.

Our analysis also raises possibly important issues for macroeconomic analysis. Markups at large firms appear to increase the crowding of startups, suggesting a feedback loop: greater markups slow the growth of competitive firms, reducing competition and leading to even greater markups. Aside from this possibility, there is a more straightforward concern that growing employment at large firms and growing markups contribute to slower growth of startups overall. Growing markups increase the crowding effect without increasing spillover benefits. Future research will be needed to explore how much these trends contribute to slower aggregate growth of productive startups and, with that, slower aggregate productivity growth.

References

Abraham, Mabel. 2015. "Explaining Unequal Returns to Social Capital among Entrepreneurs." *Academy of Management Proceedings* 2015 (1): 11264. https://doi.org/10.5465/ambpp.2015.227.

Agrawal, Ajay, Iain Cockburn, Alberto Galasso, and Alexander Oettl. 2014. "Why Are Some Regions More Innovative than Others? The Role of Small Firms in the Presence of Large Labs." *Journal of Urban Economics* 81:149–65.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–97.

Babina, Tania, and Sabrina T Howell. 2022. “Entrepreneurial Spillovers from Corporate R&D.” *Journal of Labor Economics*, November. https://doi.org/10.1086/723501.
Babina, Tania, Wenting Ma, Christian Moser, Paige Ouimet, and Rebecca Zarutskie. 2019. “Pay, Employment, and Dynamics of Young Firms.” SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.3425596.
Balsmeier, Benjamin, Lee Fleming, Matt Marx, and Seungryul Ryan Shin. 2020. “Skilled Human Capital and High-Growth Entrepreneurship: Evidence from Inventor Inflows.” w27605. National Bureau of Economic Research. https://doi.org/10.3386/w27605.
Bates, Timothy, William D. Bradford, and Robert Seamans. 2018. “Minority Entrepreneurship in Twenty-First Century America.” *Small Business Economics* 50 (3): 415–27. https://doi.org/10.1007/s11187-017-9883-5.
Bena, Jan, Serdar Dinc, and Isil Erel. 2022. “The International Propagation of Economic Downturns through Multinational Companies: The Real Economy Channel.” *Journal of Financial Economics* 146 (1): 277–304. https://doi.org/10.1016/j.jfineco.2021.09.005.
Bennett, Victor Manuel, and David T. Robinson. 2023. “Why Aren’t There More Minority Entrepreneurs?” SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.4360750.
Bessen, James, Erich Denk, Felix Poege, and Ronja Roettger. 2023. “CrunchGlass: Matching Crunchbase and Burning Glass Technologies.” TPRI Working Paper.
Bhardwaj, Abhishek, Devaki Ghose, Saptarshi Mukherjee, and Manpreet Singh. 2023. “Large Plant Openings and Retail Prices.” SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.4560164.
Bhaskar, Venkataraman, Alan Manning, and Ted To. 2002. “Oligopsony and Monopsonistic Competition in Labor Markets.” *Journal of Economic Perspectives* 16 (2): 155–74.
Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. “Quasi-Experimental Shift-Share Research Designs.” *The Review of Economic Studies* 89 (1): 181–213. https://doi.org/10.1093/restud/rdab030.
Brown, Charles, and James Medoff. 1989. “The Employer Size-Wage Effect.” *Journal of Political Economy* 97 (5): 1027–59.
Burton, M. Diane, Michael S. Dahl, and Olav Sorenson. 2018. “Do Start-Ups Pay Less?” *ILR Review* 71 (5): 1179–1200. https://doi.org/10.1177/0019793917747240.
Card, David. 2022. “Who Set Your Wage?” *American Economic Review* 112 (4): 1075–90. https://doi.org/10.1257/aer.112.4.1075.
Chatterji, Aaron, Edward Glaeser, and William Kerr. 2014. “Clusters of Entrepreneurship and Innovation.” *Innovation Policy and the Economy* 14 (January):129–66. https://doi.org/10.1086/674023.
Chen, Jun, Shenje Hshieh, and Feng Zhang. 2021. “The Role of High-Skilled Foreign Labor in Startup Performance: Evidence from Two Natural Experiments.” *Journal of Financial Economics* 142 (1): 430–52.
Cook, Lisa D., Matt Marx, and Emmanuel Yimfor. 2022. “Funding Black High-Growth Startups.” Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w30682.
Dahl, Michael S., and Steven Klepper. 2015. “Whom Do New Firms Hire?” *Industrial and Corporate Change* 24 (4): 819–36. https://doi.org/10.1093/icc/dtv026.
Davis, Steven J, and John Haltiwanger. 2014. “Labor Market Fluidity and Economic Performance.” National Bureau of Economic Research.
De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. “The Rise of Market Power and the Macroeconomic Implications.” *The Quarterly Journal of Economics* 135 (2): 561–644.

Decker, Ryan A, John C Haltiwanger, Ron S Jarmin, and Javier Miranda. 2020. "Changing Business Dynamism and Productivity: Shocks vs. Responsiveness." *American Economic Review* 110 (12): 3952–2990.
Dimmock, Stephen G, Jiekun Huang, and Scott J Weisbenner. 2022. "Give Me Your Tired, Your Poor, Your High-Skilled Labor: H-1b Lottery Outcomes and Entrepreneurial Success." *Management Science* 68 (9): 6950–70.
Fairlie, Robert, and Aaron K. Chatterji. 2013. "High-Technology Entrepreneurship in Silicon Valley." *Journal of Economics & Management Strategy* 22 (2): 365–89. https://doi.org/10.1111/jems.12015.
Fairlie, Robert, Alicia Robb, and David T. Robinson. 2022. "Black and White: Access to Capital Among Minority-Owned Start-Ups." *Management Science* 68 (4): 2377–2400. https://doi.org/10.1287/mnsc.2021.3998.
Fairlie, Robert, and David T. Robinson. 2023. "Racial Differences in Access to Capital for Innovative Start-Ups." *Entrepreneurship and Innovation Policy and the Economy* 2 (January):149–66. https://doi.org/10.1086/723238.
Giroud, Xavier, and Holger M. Mueller. 2019. "Firms' Internal Networks and Local Economic Shocks." *American Economic Review* 109 (10): 3617–49. https://doi.org/10.1257/aer.20170346.
Glaeser, Edward L., and William R. Kerr. 2009. "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" *Journal of Economics & Management Strategy* 18 (3): 623–63. https://doi.org/10.1111/j.1530-9134.2009.00225.x.
Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy* 118 (3): 536–98.
Greenstone, Michael, and Enrico Moretti. 2003. "Bidding for Industrial Plants: Does Winning a 'Million Dollar Plant' Increase Welfare?" Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w9844.
Gupta, Rahul R. 2022. "Does Goliath Help David? Anchor Firms and Startup Clusters."
Guzman, Jorge. 2024. "Go West Young Firm: The Impact of Startup Migration on the Performance of Migrants." *Management Science* 70 (7): 4824–46.
Guzman, Jorge, and Scott Stern. 2020. "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988–2014." *American Economic Journal: Economic Policy* 12 (4): 212–43. https://doi.org/10.1257/pol.20170498.
Hall, Robert E. 1988. "The Relation between Price and Marginal Cost in US Industry." *Journal of Political Economy* 96 (5): 921–47.
Hazell, Jonathon, Christina Patterson, Heather Sarsons, and Bledi Taska. 2022. "National Wage Setting." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w30623.
Hjort, Jonas, Xuan Li, and Heather Sarsons. 2020. "Across-Country Wage Compression in Multinationals." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w26788.
Hsieh, Chang-Tai, and Esteban Rossi-Hansberg. 2023. "The Industrial Revolution in Services." *Journal of Political Economy Macroeconomics* 1 (1): 3–42.
Kacperczyk, Olenka, Peter Younkin, and Vera Rocha. 2023. "Do Employees Work Less for Female Leaders? A Multi-Method Study of Entrepreneurial Firms." *Organization Science* 34 (3): 1111–33. https://doi.org/10.1287/orsc.2022.1611.
Kamepalli, Sai Krishna, Raghuram G. Rajan, and Luigi Zingales. 2020. "Kill Zone." SSRN Scholarly Paper ID 3555915. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3555915.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin. 2024. "Demographic Origins of the Start-up Deficit." *American Economic Review* 114 (7): 1986–2023.
Kim, J. Daniel. 2018. "Is There a Startup Wage Premium? Evidence from MIT Graduates." *Research Policy* 47 (3): 637–49. https://doi.org/10.1016/j.respol.2018.01.010.
Klepper, Steven. 2009. "Spinoffs: A Review and Synthesis." *European Management Review* 6 (3): 159–71. https://doi.org/10.1057/emr.2009.18.
Koski, Heli A., Otto Kässi, and Fabian Braesemann. 2023. "Killers on the Road of Emerging Start-Ups - Implications for Market Entry and Venture Capital Financing." SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.4390546.
Kyrgidou, Lida, Naoum Mylonas, Eugenia Petridou, and Evdokia Vacharoglou. 2021. "Entrepreneurs' Competencies and Networking as Determinants of Women-Owned Ventures Success in Post-Economic Crisis Era in Greece." *Journal of Research in Marketing and Entrepreneurship* 23 (2): 211–34. https://doi.org/10.1108/JRME-08-2020-0105.
Lee, Saerom, and J Daniel Kim. 2023. "When Do Startups Scale? Large-Scale Evidence from Job Postings." *Strategic Management Journal*.
Manning, Alan. 2021. "Monopsony in Labor Markets: A Review." *ILR Review* 74 (1): 3–26. https://doi.org/10.1177/0019793920922499.
McCann, Brian T., and Timothy B. Folta. 2011. "Performance Differentials within Geographic Clusters." *Journal of Business Venturing* 26 (1): 104–23. https://doi.org/10.1016/j.jbusvent.2009.04.004.
Neumeyer, Xaver, Susana C. Santos, António Caetano, and Pamela Kalbfleisch. 2019. "Entrepreneurship Ecosystems and Women Entrepreneurs: A Social Capital and Network Approach." *Small Business Economics* 53 (2): 475–89. https://doi.org/10.1007/s11187-018-9996-5.
Pe'er, Aviad, and Thomas Keil. 2013. "Are All Startups Affected Similarly by Clusters? Agglomeration, Competition, Firm Heterogeneity, and Survival." *Journal of Business Venturing* 28 (3): 354–72.
Qian, F., and R. Tan. 2021. "The Effects of High-Skilled Firm Entry on Incumbent Residents.," Stanford Institute for Economic Policy Research (SIEPR) Working Paper, 21-039, .
Sevilir, Merih. 2010. "Human Capital Investment, New Firm Creation and Venture Capital." *Journal of Financial Intermediation* 19 (4): 483–508. https://doi.org/10.1016/j.jfi.2009.08.002.
Shaver, J. Myles, and Fredrick Flyer. 2000. "Agglomeration Economies, Firm Heterogeneity, and Foreign Direct Investment in the United States." *Strategic Management Journal* 21 (12): 1175–93. https://doi.org/10.1002/1097-0266(200012)21:12<1175::AID-SMJ139>3.0.CO;2-Q.
Slattery, Cailin, and Owen Zidar. 2020. "Evaluating State and Local Business Incentives." *Journal of Economic Perspectives* 34 (2): 90–118. https://doi.org/10.1257/jep.34.2.90.
Sorenson, Olav, and Pino G. Audia. 2000. "The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the United States, 1940–1989." *American Journal of Sociology* 106 (2): 424–62. https://doi.org/10.1086/316962.
Sorenson, Olav, Michael S. Dahl, Rodrigo Canales, and M. Diane Burton. 2021. "Do Startup Employees Earn More in the Long Run?" *Organization Science*, January. https://doi.org/10.1287/orsc.2020.1371.
Stuart, Toby, and Olav Sorenson. 2003. "The Geography of Opportunity: Spatial Heterogeneity in Founding Rates and the Performance of Biotechnology Firms." *Research Policy*, Special Issue on Technology Entrepreneurship and Contact Information for corresponding authors, 32 (2): 229–53. https://doi.org/10.1016/S0048-7333(02)00098-7.
———. 2005. "Social Networks and Entrepreneurship." In *Handbook of Entrepreneurship Research: Interdisciplinary Perspectives*, edited by Sharon A. Alvarez, Rajshree Agarwal, and Olav

Sorenson, 233–52. International Handbook Series on Entrepreneurship. Boston, MA: Springer US. https://doi.org/10.1007/0-387-23622-8_11.

Tolbert, Charles M, and Molly Sizer. 1996. "US Commuting Zones and Labor Market Areas: A 1990 Update. ERS Staff Paper Number 9614. Washington, DC: Economic Research Service." *Rural Economy Division, US Department of Agriculture*.

Woodwark, Meredith, Alison Wood, and Karin Schnarr. 2021. "Standing on the Shoulders of Giantesses: How Women Technology Founders Use Single and Mixed Gender Networks for Success and Change." *International Journal of Gender and Entrepreneurship* 13 (4): 420–48. https://doi.org/10.1108/IJGE-10-2020-0159.

Tables

Table 1. Salary premium (percent)/ The effect of the big firm share in hiring on salary

	OLS	IV	OLS	OLS	IV	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
	Ln Salary	Ln Salary	Ln Salary	Ln Salary	Ln Salary	Ln Salary
	b/se	b/se	b/se	b/se	b/se	b/se
BFS	1.87***	2.20***				
	(0.32)	(0.65)				
Startup			9.43***	1.51	3.01	
			(0.60)	(3.34)	(3.91)	
Other x BFS				1.85***	2.15***	
				(0.32)	(0.66)	
Startup x BFS				3.95***	3.87***	
				(0.91)	(1.01)	
Occupation Groups x BFS						
Managerial, finance x BFS						5.08***
						(1.28)
STEM x BFS						5.98***
						(1.53)
Other professionals x BFS						-0.15
						(1.95)
Health x BFS						-5.43***
						(1.53)
Service x BFS						4.90***
						(1.61)
Sales x BFS						9.69***
						(2.11)
Administrative x BFS						3.16***
						(1.22)
Production, transportation x BFS						0.89
						(1.20)
N	23986809	23284079	23986813	23986809	23284079	363369
R-squared	0.438	0.045	0.438	0.438	0.046	0.491
F statistic, 1st stage		75.2			37.7	
Sample	All w salary	All w salary	All w salary	All w salary	All w salary	Startup w salary

Note: The table shows the relationship between the big firm share in hiring and log salaries included in online job postings. The dependent variable, log salary, is multiplied by 100 in order to have the estimated coefficients show percentage changes. BFS represents the big firm share of hiring and is standardized. All regressions control for education, experience, experience squared and part-time work. In addition, we include fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone. The unit of observation is at the ad level. In Column 1 to Column 5, the estimation sample consists of all online job postings that include wage salaries. Thus, all firms that post salaries are included in the estimation. In Column 6, the estimation sample is restricted to only include startups. Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Table 2. Salary premium (percent), Other measures
Dependent variable: Ln Salary x 100

	(1) OLS b/se	(2) IV b/se	(3) OLS b/se	(4) OLS b/se	(5) OLS b/se
BF residual	1.13***	4.42**			
	(0.24)	(1.94)			
Other x BF residual			1.11***		
			(0.24)		
Startup x BF residual			3.10***		
			(0.58)		
Startup			10.08***		
			(0.60)		
Ln productivity				1.68***	
				(0.37)	
Ln markup				1.72***	
				(0.34)	
BFS					4.15***
					(1.07)
Mobility x BFS					-0.42**
					(0.18)
N	20642101	20186696	20642101	20642026	18985708
R-squared	0.447	0.041	0.448	0.447	0.446
Test coef. equal (pval)			0.001		
F-test, 1st stage		11.2			
Sample	Excl. big firms	Excl. big firms	Excl. big firms	Excl. big firms	All w salary

Note: The table shows the relationship between the big firm salary residual and log salaries included in online job postings. The dependent variable, log salary, is multiplied by 100 in order to have the estimated coefficients in percentage changes. The big firm salary residual is estimated from the following regression: We regress log salary against controls for education, experience, experience squared, and part-time jobs, with fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone. Subsequently, we calculate a residual for each ad and take the mean residual for big firms in each commuting zone each year. This measure is standardized by dividing by its standard deviation. In Column 4, the big firm residual is decomposed into a productivity and a markup component (also standardized). Column 5 includes a mobility measure of commuting zones from the Current Population Survey (see Data section for more details). BFS represents the big firm share of hiring. All regressions control for education, experience, experience squared, and part-time work. In addition, we include fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone. The unit of observation is at the ad level. The estimation sample consists of all online job postings that include wage salaries, excluding big firms in Columns 1-4 to avoid a spurious correlation between big firm wages and big firm residuals. Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Table 3. Big Firm Share and Ad Growth

	OLS (1) Ad hazard b/se	OLS (2) Ad hazard b/se	IV (3) Ad hazard b/se	OLS (4) Ad count b/se	IV (5) Ad count b/se
Startup wage residual	-0.189***				
	(0.050)				
BFS		-1.260***	-0.930***	-1.079***	-0.811**
		(0.173)	(0.343)	(0.179)	(0.342)
N	842380	860561	845557	138837	138671
R-squared	0.013	0.298	0.001	0.649	0.001
Age FE	yes	yes	yes	yes	yes
Firm FE		yes	yes	yes	yes
Industry x year FE		yes	yes	yes	yes
Mean dep. Variable		3.71	3.72	5.69	5.69
Net effect		-1.26	-0.93	-1.08	-0.81
Net percent		-34.0%	-25.0%	-19.0%	-14.2%
F statistic, 1st stage			77.7		55.7

Note: The table shows the association between the big firm share of hiring and ad growth of startups. The first three Columns estimate linear probability models of the hazard of placing the first ad. The dependent variable, the ad hazard, is a binary variable equal to 1 in the first year that a startup advertises and 0 otherwise. The sample excludes observations past the first year of advertising and only includes startups. The startup wage residual in Column 1 is estimated as follows: We regress log salary against controls for education, experience, experience squared, and part-time jobs, with fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone. Subsequently, we calculate a residual for each ad and take the mean residual for startups in each commuting zone each year. We divide this measure by its standard deviation to standardize it. BFS represents the big firm share of hiring. Columns 4 and 5 use the number of ads a startup posts each year conditional on having posted its first ad as a dependent variable. We winsorize the top 1 percent of the ad count variable. The regressions include fixed effects for firm age, firm, and industry x year, as indicated. The unit of observation is at the firm x year level. The sample includes all startups that post ads. Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Table 4. Proximity

	(1) Ln salary (pct) b/se	(2) Ad hazard b/se	(3) Ad count b/se
Distant x BFS	3.511***	-1.382***	-1.296***
	(1.131)	(0.092)	(0.167)
Mid x BFS	0.578	-1.106***	-0.932***
	(0.857)	(0.093)	(0.166)
Close x BFS	0.476	-0.949***	-0.471***
	(0.904)	(0.095)	(0.163)
N	354818	866631	142692
R-squared	0.493	0.033	0.074
Unit of observation	Ad	firm x year	firm x year
Sample	Startups	Startups	Advertising startups
Baseline distant		3.1	3.9
Baseline close		4.7	7.3
Effect on distant	3.5%	-44.3%	-33.6%
Effect on close	0.5%	-20.4%	-6.5%
Test close=distant (pval)	0.000	0.000	0.000

Note: The table shows the effect of the big firm share on wages and advertising of startups by differences in the markup-adjusted measure of business proximity. Startups are divided into those in the bottom quartile, middle quartiles, and top quartiles of markup-adjusted proximity to big firms. See the text for more details on the construction of the markup-adjusted proximity measure. BFS represents the big firm share of hiring. In Column 1, the dependent variable (log salary) is multiplied by 100 in order to have the estimated coefficients in percentage changes. The regression controls for education, experience, experience squared and part-time work. In addition, we include fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone. The unit of observation is at the ad level. The sample includes only advertising startups. In Column 2, the dependent variable, ad hazard, is a binary variable equal to 1 in the first year that a startup advertises and 0 otherwise. The sample excludes observations past the first year of advertising and only includes startups. The regression includes fixed effects for firm age, commuting zone, and industry x year. Note that we cannot control for firm fixed effects in this regression because the proximity measure is based on firms. The unit of observation is at the firm x year level. In Column 3, the dependent variable is the number of ads a startup posts each year, conditional on having posted its first ad. The regression includes fixed effects for firm age, commuting zone, and industry x year as indicated. The unit of observation is at the firm x year level. The sample includes only advertising startups. In all columns, standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Table 5. Employment

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Employment	Employment	Ln Salary	Ad hazard	Ad count
	b/se	b/se	b/se	b/se	b/se
Cumulative Ads	1.45***				
	(0.02)				
Interactions with dummy for firms led/founded by minorities or women					
Other firms x cum. Ads		1.48***			
		(0.02)			
Min./fem. x cum. Ads		1.16***			
		(0.07)			
Other firms x BFS			0.91**	-1.07***	-0.67***
			(0.36)	(0.10)	(0.17)
Min./fem. x BFS			2.39***	-1.32***	-0.91***
			(0.49)	(0.18)	(0.20)
Min./fem. Led/founded			-5.67***	2.52***	1.12
			(1.73)	(0.89)	(0.92)
N	131104	127791	10831544	845373	139947
R-squared	0.034	0.035	0.462	0.033	0.070
Test coef. difference (pval)			0.000	0.147	0.203

Note: The table shows the relationship between startup employment in 2021 and the cumulative number of ads posted over the previous 10 years. Column 1 and Column 2 use the mean of an intervaled employment measure from the 2021 Crunchbase "snapshot" as the dependent variable. The main independent variable is the cumulative number of ads posted over the previous 10 years. Minority/female-founded/led startups is a binary variable equal to 1 whenever a startup is founded or led by a woman and 0 otherwise. The regressions control for firm age in 2021. The sample only includes startups. In Column 3, we repeat the analysis of Table 1, Column 1, for minority/female-founded or led startups. In Column 4, we repeat the analysis of Table 3, Column 2, for minority/female-founded or led startups. In Column 5, we repeat the analysis of Table 3, Column 4, for minority/female-founded or led startups. In all columns, standard errors are shown in parentheses (*** p<0.01, ** p<0.05, * p<0.10). In Columns 3-5, standard errors are clustered at the commuting zone level.

Table 6. Ads and Outcomes (2021)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ever funded	Funding rounds	Funds received ($1k)	Website operating	Closed	Acquired	IPO
Cumulative Ads	5.05***	0.52***	57305.79***	2.48***	-1.06***	1.51***	1.35***
(100s)	(0.20)	(0.01)	(822.94)	(0.19)	(0.10)	(0.10)	(0.04)
N	141117	50629	39046	137313	141117	141117	141117
R-squared	0.009	0.075	0.111	0.003	0.013	0.014	0.010

Note: The table shows the relationship between the cumulative ads measure and a variety of startup performance measures. The independent variable of interest is the cumulative number of ads posted over the previous 10 years, rescaled to hundreds of ads (divided by 100). In Column 1, the dependent variable is a binary variable equal to 1 whenever the startup has received funding at any point in time and 0 otherwise. In Column 2, the dependent variable captures the number of funding rounds. In Column 3, the dependent variable is the cumulative amount of funds raised (in thousands of dollars). In Column 4, the dependent variable is whether the startup had an operating website in 2021 (a measure of survival). In Column 5, the dependent variable is whether the startup is reported as having closed. In Column 6, the dependent variable is whether the startup was acquired. In Column 7, the dependent variable is whether the startup had an IPO. All binary dependent variables are multiplied by 100. Each regression controls for startup age. The sample is restricted to startups. Standard errors are shown in parentheses (*** p<0.01, ** p<0.05, * p<0.10).

Appendix

Appendix A: Summary Tables

Table A.1: Summary statistics

Summary statistics: Ad-level

	Mean	SD	Min	50 pct	75 pct	Max
Salary (000 USD)	54.60	40.12	10.01	41.60	66.00	350.00
Startup	0.02	0.15	0.00	0.00	0.00	1.00
Big firm share (%)	0.24	0.06	0.00	0.23	0.28	0.82
Big firm share (Pred., %)	0.24	0.06	0.00	0.22	0.27	3.00
Education: High school (%)	0.30	0.46	0.00	0.00	1.00	1.00
Education: College+ (%)	0.20	0.40	0.00	0.00	0.00	1.00
Education: Missing (%)	0.16	0.36	0.00	0.00	0.00	1.00
Experience: Years	1.21	2.11	0.00	0.00	2.00	15.00
Experience: Missing (%)	0.58	0.49	0.00	1.00	1.00	1.00
Part-time job (%)	0.14	0.35	0.00	0.00	0.00	1.00
Salary posted (%)	0.24	0.10	0.01	0.26	0.33	1.00
Labor market tightness	0.90	0.42	0.12	0.93	1.17	3.70

Legend: N=31.7 million.

Note: The total number of observations is 31.7 million. The unit of observation is at the ad-level. Only ads with salary information are considered.

Summary statistics: Start-year level

	Mean	SD	Min	50 pct	75 pct	Max
Ad posting (extensive)	0.09	0.28	0.00	0.00	0.00	1.00
Ad posting (intensive)	5.59	15.13	0.00	1.00	3.00	105.00
Founding Year	2012.98	2.35	2010.00	2013.00	2015.00	2020.00
Big firm share (%)	0.27	0.06	0.01	0.27	0.31	0.66
Big firm share (Pred., %)	0.26	0.06	0.00	0.27	0.30	0.71
Women/Minority founded/led	0.12	0.32	0.00	0.00	0.00	1.00
Markup-adjusted business proximity	0.16	0.05	0.07	0.15	0.19	0.26
Startup salary premium	0.18	0.26	-0.81	0.15	0.29	5.79
Tech startup	0.51	0.50	0.00	1.00	1.00	1.00
In 2013 Crunchbase	0.20	0.40	0.00	0.00	0.00	1.00

Legend: N=977,848.

Note: The total number of observations is 977,848. The unit of observation is at the startup-level. Intensive margin is restricted to ad posting after the first ad (N=144,514), and winsorized at the 95th percentile (original maximum is >2000).

Table A.2: Instrumental variable, first stage regressions

	(1) Big firm share b/se	(2) Big firm share b/se	(3) Big firm share b/se
Big firm share (Pred.)	0.50***	0.37***	0.37***
	(0.03)	(0.04)	(0.05)
N	23471952	845557	138671
R-squared	0.848	0.897	0.917
Sample	Ads w salary	Startups, ext.	Startups, int.

Note: The table shows the first stage regressions of the instrumental variable approach for the different regression analyses at the ad- and startup level. Column 1 shows the first stage at the ad level, Column 2 shows the first stage for the startup level until the first ad, and Column 3 shows the first stage for the startup level once the first ad has been posted and after. In Column 1, we control for education, experience, experience squared, and part-time work. In Columns 2 and 3, we include fixed effects for firm age, firm, and industry x year. Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Appendix B: Model

Basic Nash equilibrium

Solving the two first-order conditions, $\frac{\partial \pi_i}{\partial w_i} = 0$, we obtain reaction functions for w_s and w_L respectively,

$$r_s = \frac{1}{2}(p_s - w_L + t), \quad r_L = \frac{1}{2}(p_L - w_s + t). \tag{A1}$$

Then, setting $r_s = w_s, \ r_L = w_L$, we solve for the Nash equilibrium,

$$\widehat{w}_s = \frac{2p_s + p_L}{3} - t, \quad \widehat{w}_L = \frac{2p_L + p_s}{3} - t \tag{A2}$$

with equilibrium profits

$$\hat{\pi}_s = \frac{(p_s - p_L + 3t)^2}{18t}, \quad \hat{\pi}_L = \frac{(p_L - p_s + 3t)^2}{18t}. \tag{A3}$$

and big firm share,

$$\hat{y} = 1 - \frac{\widehat{w}_s - \widehat{w}_L + t}{2t} = \frac{p_L - p_s}{6t} + \frac{1}{2}. \tag{A4}$$

We obtain equation (2) from (A2) and (A4)

$$\widehat{w}_s - p_s = \frac{p_L - p_s}{3} - t = 2t\hat{y} - 2t.$$

Similarly, for (3).

Spillovers and markups

To implement an empirical analysis of spillovers and markups, it is helpful to add some stylized features to our base model. Let us suppose that within the local labor market there are now two startup firms in two distinct industries, *k* and *n*, where startup *k* benefits from big firm spillovers and startup *n* does not. These two startups each hire workers with experience in their respective industries, so they do not directly compete in the labor market. However, the big firm does hire from both industries and so competes with each startup. There are, effectively, two submarkets: S_k workers with *k* industry experience and $1 - S_k$ workers without. In each submarket, there is a separate equilibrium with big firm shares $\hat{y}^k$ and $\hat{y}^n$, so that the combined big firm share is $\hat{y} = S_k\hat{y}^k + (1 - S_k)\hat{y}^n$.

From the text, we have

$$\hat{y} = \frac{p_L\left(1-\frac{\lambda}{m}\right)-\overline{p}}{6t} + \frac{1}{2} \text{ and } \hat{y} = S_k\hat{y}^k + (1-S_k)\hat{y}^n. \tag{A5}$$

Assuming that $\lambda = 0$ for industry *n*,

$$\hat{y}^k = \frac{p_L\left(1-\frac{\lambda}{m}\right)-\overline{p}}{6t} + \frac{1}{2}, \quad \hat{y}^n = \frac{p_L-\overline{p}}{6t} + \frac{1}{2}. \tag{A6}$$

Combining (A5) and (A6) and rearranging,

$$\hat{y}^k = \hat{y} - (1-S_k)\frac{\frac{\lambda}{m}p_L}{6t}. \tag{A7}$$

We can relate $\hat{y}^k$ to startup profits as in (3),

$$\hat{\pi}_s^k = 2t(1-\hat{y}^k)^2 \tag{A8}$$

however, we do not observe $\hat{y}^k$; we only observe $\hat{y}$. But

$$\frac{\partial\, \hat{\pi}_s^k}{\partial\, \hat{y}} = \frac{\partial\, \hat{\pi}_s^k}{\partial\, \hat{y}^k}\frac{\partial\, \hat{y}^k}{\partial\, \hat{y}}. \tag{A9}$$

Taking the derivative of (A7) and using (A6), we get

$$\frac{\partial\, \hat{y}^k}{\partial\, \hat{y}} = 1-(1-S_k)\frac{\frac{\lambda}{m}}{1-\frac{\lambda}{m}} \approx 1-(1-S_k)\frac{\lambda}{m} \tag{A10}$$

the approximation holding as long as $\frac{\lambda}{m}$ is not too large (it is not in the data). Combining (A9) and (A10), we can write

$$\frac{\partial\, \hat{\pi}_s^k}{\partial\, \hat{y}} \approx \beta_\pi + \delta\frac{\lambda}{m}$$

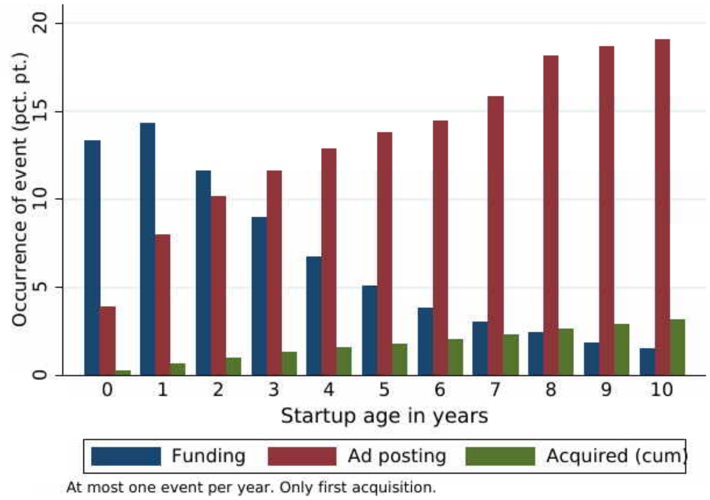
yielding a regression specification

$$\pi_{it} = \beta_\pi y_{cz,t} + \delta\frac{\lambda}{m}y_{cz,t} + \gamma X_{i,cz,t} + \epsilon_{i,cz,t}, \tag{A11}$$

where λ is a measure of business proximity for the startup relative to big firms, *m* is mean big firm markup for the commuting zone, and $X_{i,cz,t}$ consists of control variables and fixed effects. Given positive spillovers, we expect $\delta > 0 > \beta_\pi$. We obtain a similar equation for startup wages and in the empirical analysis we proxy profits with firm growth.

Appendix C: Supplemental Figures and Tables

Figure C.1: Ad-posting, funding, and acquisitions over the startup life-cycle



Note: The figure shows yearly funding, yearly ad posting, and whether a startup has previously been acquired over the startup age. The y-axis shows the event's occurrence in percentage points, and the x-axis shows the startup age in years. At most, one event per year is shown (funding, ad posting), and only the first acquisition is included.

Table C.1 Salary premium (percent) table with controls for salary posting

	OLS (1) Ln Salary b/se	IV (2) Ln Salary b/se	OLS (3) Ln Salary b/se	IV (4) Ln Salary b/se
Other x BFS	1.61***	1.63***	1.58***	1.65**
	(0.29)	(0.51)	(0.38)	(0.83)
Startup x BFS	3.53***	3.52***	3.50***	3.53***
	(0.81)	(0.89)	(0.85)	(1.10)
Startup	1.70	1.91	1.70	1.91
	(3.18)	(3.44)	(3.18)	(3.44)
Salary posted (%)			-0.87	0.31
			(6.00)	(8.20)
N	23986809	23471952	23986809	23471952
R-squared	0.438	0.045	0.438	0.045
Net effect	3.529	3.517	3.495	3.534
Net std err	0.808	0.893	0.852	1.096
F stat		202.5		138.8
Sample	All w salary	All w salary	All w salary	All w salary

Note: The table shows the relationship between the big firm share of hiring and startup wages offered, additionally controlling for the share of ads in which salary is being posted. The dependent variable, log salary, is multiplied by 100 in order to have the estimated coefficients show percentage changes. BFS represents the big firm share of hiring and is standardized. The salary posted is the share of ads in a commuting zone and year that listed a salary. All regressions control for education, experience, experience squared, and part-time work. In addition, we include fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone. The unit of observation is at the ad level. The estimation sample consists of all online job postings that include wage salaries. Thus, all firms that post salaries are included in the estimation. Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Table C.2: Robustness: Big firm share with threshold of 10,000

	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
	Ln Salary	Ln Salary	Hazard	Hazard	Count	Count
	b/se	b/se	b/se	b/se	b/se	b/se
Other x BFS	1.8***	1.5***				
	(0.3)	(0.5)				
Startup x BFS	4.0***	4.1***				
	(0.7)	(0.9)				
BFS			-21.0***	-18.1***	-16.4***	-13.0**
			(3.0)	(5.5)	(3.4)	(5.2)
Startup	1.2	-0.1				
	(2.8)	(3.3)				
N	23986809	23471952	860561	845557	138837	138671
R-squared	0.438	0.045	0.298	0.001	0.649	0.001
F stat		227.332		116.167		59.890

Note: The table shows a robustness check, defining large firms as employing 10,000 employees instead of 1,000 (default). For detailed descriptions, see Table 1 (Columns 1-2) and Table 3 (Columns 3-6). Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Table C.3: Robustness of Big Firm Share and Ad Growth-Relationship

	IV: HQ growth only	Tech startups	Excl. moved startups	2013 snapshot	2013 relative proximity
	(1)	(2)	(3)	(4)	(5)
	Ad hazard	Ad hazard	Ad hazard	Ad hazard	Ad hazard
	b/se	b/se	b/se	b/se	b/se
BFS	-1.192***	-1.032***	-1.263***	-0.463**	
	(0.298)	(0.188)	(0.164)	(0.209)	
BFS x Q1 (Distant)					-0.529**
					(0.207)
BFS x Q2					-0.502**
					(0.211)
BFS x Q3					-0.424*
					(0.220)
BFS x Q4 (Close)					-0.362*
					(0.219)
N	845557	439027	833122	163027	158877
R-squared	0.001	0.309	0.300	0.063	0.064
Age	yes	yes	yes	yes	yes
Firm	yes	yes	yes		
Czone				yes	yes
industry x year	yes	yes	yes	yes	yes
Mean dep. variable	3.72	3.61	3.64	2.96	2.91
F stat	75.9				

Note: The table shows robustness tests to an alternative instrumental variable and exclusion of relevant subgroups. The dependent variable is whether a startup posts ads. In Column 1, we use an alternative instrument that uses only variation in ad posting in the headquarter commuting zone of the big firm. In Column 2, we focus on tech startups, defined as being in the Crunchbase category of Software, AI, IT, Internet services, Messaging, or Platforms. In Column 3, we exclude startups that changed their locations between Crunchbase snapshots from the sample. In Column 4, we focus on startups that were already included in the 2013 Crunchbase snapshot. Finally, in Column 5, we use business descriptions of startups in 2013 to calculate proximity scores. The unit of observation is at the startup-year level. Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Measures of firms founded or led by women or minorities

We seek to identify startups that were founded or are led by minorities or women. Crunchbase provides identifiers for startups self-identified as having been founded or led by women, Hispanics, or blacks. That identifier is, however, missing for most firms. Among startups, 0.5 percent are identified as founded or started by blacks, 0.4 percent by Hispanics, and 6.7 percent by women. Taking the union of these identifiers, 7.4 percent are founded or led by minorities or women.

There is a concern, however, that these measures might exhibit a selection bias because they are self-identified (see Cook, Marx, and Yimfor 2022 fn. 4). We can, however, construct an extended measure that should correct any selection bias, at least for women who make up most of the group of firms identified. We use additional data that Crunchbase provides on the gender of startup founders and leaders. This alternative measure takes on the value of 1 whenever a startup is founded or led by a woman based on the additional gender variable provided.

Table C.4 in the Appendix shows how much the two measures for female-founded or led firms overlap. The "original" variable refers to the Crunchbase tag. The "alternative" variable refers to the alternative measure based on the additional data on gender and first names. The first row shows the number of observations where the original variable is non-missing, however the alternative variable has missing observations. The second row shows the number of observations where the original variable is missing but the alternative is non-missing. We can see that among startups, about 28.5 percent of observations are non-missing for the "original" variable but missing for the alternative. The majority of observations are non-missing for both variables among startups.

"Overlap" shows the number of observations of the two measures for female founders or leaders where one of the two measures has a missing observation as well as the number of observations where both measures are non-missing (this essentially adds up row 1 and 2). 51 percent of the observations are non-missing among startups. Thus, for non-missing observations, the table shows that the two measures for female-founded or led startups strongly overlap. Almost 96 percent of observations are both equal to 1 or both equal to 0, which suggests that our "original" measure does relatively well in identifying female-founded or led startups.

Based on this analysis, we construct an alternative indicator of whether a firm is founded/led by minorities or women. Where the original identifier is missing, we replace it with the alternative name-based identifier. This expands the number of firms where the new identifier is non-missing, but the share of firms founded or led by minorities or women increases only slightly from 7.4

percent as above to 7.9 percent, suggesting little selection bias. In any case, we use the extended identifier in the main text (Table 5). To check robustness, we run the same analysis with the original identifier, see Table C.5. The coefficients are similar; however, the sample size increases with the new identifier, improving the accuracy of some estimates.

Table C.4: Overlap between female founded/led startups and alternative measure

	(1) Non-startups Col %	(2) Startups Col %	(3) Total Col %
Original and Alternative			
Original non-missing	2.26	28.50	6.29
Alternative non-missing	29.65	9.08	26.49
Both non-missing	2.56	39.85	8.29
Both missing	65.52	22.58	58.93
Total	100.00	100.00	100.00
N	(1136251)	(206212)	(1342463)
Overlap			
either one missing	92.58	48.53	79.83
both non-missing	7.42	51.47	20.17
Total	100.00	100.00	100.00
N	(391733)	(159648)	(551381)
Overlap among non-missing			
not equal	4.20	4.15	4.16
both equal 1 or 0	95.80	95.85	95.84
Total	100.00	100.00	100.00
N	(29069)	(82167)	(111236)

Note: The original and alternative header refers to the "original" variable identifying female founded or led firms given by the Crunchbase tag. The "alternative" variable refers to our second measure of female-founded or led firms based on the additional gender and first name information. The first row shows the number of observations where the original variable is non-missing but the alternative variable has missing observations. The second row shows the number of observations where the original variable is missing, but the alternative is non-missing. The third row shows the number of observations where both variables have non-missing observations, and the fourth row shows the number of observations where both measures have missing observations. Overlap is defined as either one missing or both non-missing. Overlap among non-missing shows the number of observations when both measures are either equal to 1 or 0 or are not equal.

Table C.5: Minority/female led/founded firms, using original Crunchbase identifier

	(1)	(2)	(3)
	Ln Salary	Ad hazard	Ad count
	b/se	b/se	b/se
Other firms x BFS	1.04**	-1.07***	-0.65***
	(0.52)	(0.10)	(0.17)
Min./fem. x BFS	1.17	-1.31***	-0.92***
	(1.63)	(0.18)	(0.20)
Min./fem. Led/founded	14.25**	2.53***	1.29
	(5.74)	(0.92)	(0.93)
N	1052629	823675	136257
R-squared	0.514	0.034	0.071
Test coef. difference (pval)	0.933	0.160	0.153

Note: The table shows the relationship between the minority/female-led/founded startup identifier and startup variables, including salary, ad hazard, and ad count. We use the original Crunchbase tag that identifies minority and women-founded or led startups as the main independent variable. Minority/female-founded/led startups is a binary variable equal to 1 whenever a startup is founded and/or led by a woman and 0 otherwise. BFS represents the big firm share of hiring and is standardized. In Column 1, we repeat the analysis of Table 1, Column 1, for the original minority/female-founded or led startup identifier. In Column 2, we repeat the analysis of Table 3, Column 2, for the original minority/female-founded or led startup identifier. In Column 3, we repeat the analysis of Table 3, Column 4, for the original minority/female-founded or led startup identifier. In all Columns, standard errors are shown in parentheses and are clustered at the commuting zone level. (*** p<0.01, ** p<0.05, * p<0.10).

Table C.6: BLS-based markup vs DEU markup

Dep. var.: Ln Salary	(1) BLS-based Markup	(2) DEU markup
Ln productivity	3.536***	
	(0.434)	
Ln BF BLS-based markup	3.222***	
	(0.528)	
Ln productivity (alt)		4.723***
		(0.580)
Ln BF DEU markup		4.429***
		(0.586)
N	6694829	6694829
R-squared	0.403	0.403

Note: The table shows a replication of Table 2, Column 4, using two different alternative measures for markups. The dependent variable, log salary, is multiplied by 100 in order to have the estimated coefficients in percentage changes. Similar to Table 2, Column 4, the big firm residual is decomposed into a productivity and a markup component. The regressions control for education, experience, experience squared, and part-time work. In addition, we include fixed effects for year, 2-digit industrial sector, 6-digit occupation code, and commuting zone. The unit of observation is at the ad level. The estimation sample consists of all online job postings that include wage salaries between 2010 and 2016 (due to data limitations in the markup variable). Thus, all firms that post salaries are included in the estimation. Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).

Table C.7: Robustness to alternative proximity measures

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln Salary	Ad hazard	Ad count	Ln Salary	Ad hazard	Ad count
	b/se	b/se	b/se	b/se	b/se	b/se
BFS x Q1 (Distant)	2.323**	-1.701***	-1.956***	3.000**	-1.339***	-1.139***
	(1.031)	(0.096)	(0.185)	(1.191)	(0.094)	(0.156)
BFS x Q2	0.249	-1.461***	-1.649***	1.150	-1.146***	-0.971***
	(0.901)	(0.097)	(0.186)	(0.960)	(0.097)	(0.165)
BFS x Q3	0.280	-1.341***	-1.399***	0.624	-0.996***	-0.738***
	(0.933)	(0.096)	(0.186)	(0.934)	(0.096)	(0.159)
BFS x Q4 (Close)	1.415	-1.241***	-1.066***	1.062	-0.908***	-0.365**
	(0.956)	(0.095)	(0.162)	(0.942)	(0.098)	(0.164)
N	354818	866632	142692	354818	866632	142692
R-squared	0.492	0.033	0.074	0.491	0.033	0.075
Unit	Ad	firm x year	firm x year	Ad	firm x year	firm x year
Sample	Startups	Startups	Startups	Startups	Startups	Startups
Proximity	Absolute	Absolute	Absolute	Relative	Relative	Relative

Note: The table shows the effect of the big firm share on wages and advertising of startups by differences in the business proximity. Business proximity is operationalized as absolute proximity (Columns 1-3) and relative proximity (Columns 4-6). For markup-adjusted business proximity and comprehensive notes, see Table 4. Standard errors are shown in parentheses and are clustered at the commuting zone level (*** p<0.01, ** p<0.05, * p<0.10).