Centre for Economic Performance

Discussion Paper

ISSN 2042-2695

No. 2231
November 2026

Migrant entrepreneurship, high-growth firms and cities

Tom Kemeny
Max Nathan
Ceren Ozgen
Guido Pialli
Anna Rosso
Mateo Seré
Anna Valero

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

UKRI Economic and Social Research Council

Abstract
High-growth firms represent a small share of businesses in OECD countries, but disproportionately drive national economic activity. This paper examines migrants' role in founding high-growth firms, and the urban nature of high-growth entrepreneurship. Using novel founder- and company-level data on over 12,000 UK companies and 15,000 founders, we analyse the characteristics of migrant founders and the firms that they build. We establish six new facts. First, relative to the UK migrant population, migrant entrepreneurs are substantially over-represented among founders of high-growth potential firms. Second, compared to UK-born founders, migrant founders are positively selected on human capital, but negatively selected on entrepreneurial experience. Third, high-growth potential firms are highly urbanised, and those with a migrant (co-)founder display a distinctive urban geography. Fourth, firms with mixed native-migrant founding teams show faster employment and revenue growth than all-UK teams. All-migrant teams show weaker growth premia. Fifth, compared with UK-born founded firms, firms with mixed native-migrant founding teams have more distinctive strategic positioning, with fewer peers. Sixth, for co-founded firms, our results vary by birth countries, first language and urban location. Migrant-founded firms in London and conurbations exhibit the largest performance premia.

Keywords: entrepreneurship; immigration; diversity; cities; high-growth firms
JEL codes: J61; L26; O31; R10; R12

This paper was produced as part of the Growth Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Thanks to Sameera Siddiqui for outstanding research assistance. Thanks to Annabelle Brenes, Chigusa Okamoto, Catherine van der List, Yi Wu and audiences at UCL, European UEA 2026 and Reading WUEEG 2026 for helpful comments. We thank the UKRI for funding under grant ES/W010232/1. The project was reviewed and approved by the UCL Research Ethics Board, application 22883/00. Thanks to Henry Whorwood at Beauhurst and Filipe Mesquita at Diffbot for guidance on data. Other usual disclaimers apply. LLMs were used to adjudicate edge cases in founder-person matching and, alongside RA classifications, to validate a sample of company description text. Claude Code was also used to develop parts of the codebase for distinctiveness measures and alternative metrics used in sensitivity checks. All AI-assisted outputs were reviewed and edited by the authors.

Tom Kemeny, University of Toronto. Max Nathan, University College London and Centre for Economic Performance at the London School of Economics. Ceren Ozgen, University of Birmingham. Guido Pialli, University College London and University of Turin. Anna Rosso, University of Milan. Mateo Seré, University College London and University of Montevideo. Anna Valero, Centre for Economic Performance at the London School of Economics.

Published by
Centre for Economic Performance
London School of Economic and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

© H. Breinlich and M. Magli, submitted 2026.

DOI: 10.21953/researchonline.lse.ac.uk.00141909

1 Introduction

What role do migrants play in founding high-growth firms, and how does this shape the spatial distribution of high-growth entrepreneurship? In many countries, young, high-growth companies make a disproportionate contribution to national employment, innovation, and productivity (Calvino et al., 2018; Haltiwanger et al., 2013). These companies account for only a small share of all firms — around five percent in the UK in 2024 (Office for National Statistics, 2025).[1] In a recent review, Botelho et al. (2026) describe these firms as 'innovation-driven startups' whose founders bring novel business models, products and processes into the market, driving growth through Schumpetarian processes of entry, competition and knowledge spillovers (Aghion and Howitt, 1992; Klette and Kortum, 2004; Lentz and Mortensen, 2008). Channeling Jane Jacobs, Duranton and Puga (2001) argue that big cities are 'nurseries' for startups: high-growth-orientated firms are notably concentrated in high-tech clusters and urban cores (Guzman and Stern, 2015; Chatterji et al., 2014).

Understanding who founds these firms, which ventures succeed and why, is therefore important to understanding drivers of economic growth more broadly.[2] Migrants are over-represented in firm formation, including high-growth companies, and play prominent roles in high-tech clusters (Azoulay et al., 2022; Kerr and Robert-Nicoud, 2020). As we explain in Section 2, there are three intersecting explanations for these geographies of high-growth migrant entrepreneurship. A first line of argument emphasises migrants' positive selection on entrepreneurial skills and attitudes (Lee et al., 2025; Beine et al., 2024). But since both migrants and entrepreneurs differ from the average worker, migrant-native differences may disappear once other founder observables are controlled for (Manning, 2025). A second set of explanations emphasises co-founding team dynamics: innovative startups benefit from co-founders (Howell and Hall, 2026; Gans et al., 2019) and migrants may bring cognitive diversity to founding teams, gains from co-ethnic networks, or both (Kerr and Mandorff, 2023; Hong and Page, 2004). Conversely, team homophily might lead to suboptimal outcomes (via groupthink) and diversity might create frictions between team members. Migrant populations are highly urbanised, and a third suggestion is that concentrations of skilled migrants create high-end enclave effects, supporting innovation and entrepreneurship (Saxenian, 2006). Against this, urban-level concentrations of migrant entrepreneurs might simply reflect individual and team-level, rather than area-level processes.

[1] High growth businesses are defined as firms with 10 or more staff and an average growth in employment of greater than 20%, per annum, over a three-year period.

[2] Our focus in this paper is entrepreneurship by choice, not necessity (Jones et al., 2018). We focus on high-growth potential ventures, similarly to Botelho et al. (2026)'s notion of 'innovation-driven entrepreneurship'.

What we do. We explore these issues using a novel dataset covering over 12,000 UK companies and 15,000 founders. Our analysis focuses on a highly selective but economically important sample: UK companies identified as high-growth-potential based on observable lifecycle triggers such as academic spinouts, winning large innovation-related grants or receiving external finance. Combining geocoded founder and company-level information with online profiles from a global knowledge graph, we build a novel linked founder–firm dataset, including detailed information on founders' education and career histories, dimensions that are remain rare in the entrepreneurship literature (Botelho et al., 2026). We describe our sources in detail in Section 3. We show that while companies in our data are unlike average UK firms, our sample is very largely selected on high-growth *potential*, and exploit the variation between actual and unrealised firm performance for our analysis.[3]

Specifically, we examine how the characteristics of migrant entrepreneurs, the composition of founding teams, and the spatial environments in which firms operate relate to company outcomes. Our analysis is guided by four questions.

1. How do migrant founders differ from UK-born founders in terms of observable characteristics such as age, gender, education, and career histories?

2. How are high-growth firms and their founders distributed across space, and how urbanised are these patterns?

3. How does founding team composition vary across firms founded by UK-born entrepreneurs, migrant entrepreneurs, and mixed teams combining both groups?

4. How do firms with different founding team compositions compare across a range of outcomes, including novel market positioning, access to external finance, employment and revenue growth, and exit events?

Main results. We provide new evidence on high-growth migrant entrepreneurship and its geographies. We establish several new facts: we highlight six of these. **First,** relative to the UK's migrant population, migrant entrepreneurs are substantially overrepresented among founders of high-growth and high-growth-potential firms. **Second,** compared to UK-born founders, migrant founders are positively selected on human capital, but negatively selected on entrepreneurial experience. We find no evidence of selection on other entrepreneurial skills. **Third,** high-growth potential firms are highly urbanised, and those with a migrant founder or co-founder display a distinctive urban geography. **Fourth,** firms with mixed native-migrant founding teams follow faster revenue growth trajectories than all-UK teams, and initially, also scale faster on headcount. These results

[3]Around nine percent of our sample are selected on the basis of actual employment or revenue growth: we run robustness tests excluding these companies.

are largely unaffected by excluding companies selected on jobs/revenue growth, although revenue premia are less precisely estimated. We also find evidence for growth premia in all-migrant teams, although this is more sensitive to sample construction. **Fifth,** as compared with firms founded by natives, firms with mixed native-migrant founding teams have more distinctive strategic positioning, with fewer peers. **Sixth**, birth country, language and urban location channels help explain our main results. For co-founded firms, performance gains are generally larger when founders share a birth country and/or first language. Conversely, mixed teams grow faster when founders have different languages, and firms with all-migrant founders have fewer peers when founders come from different countries. Migrant-founded firms in London and conurbations exhibiting the largest performance premia relative to all-UK counterparts. Given we find weak evidence for positive selection of migrant entrepreneurs, and few significant effects for solo migrant founders, our findings are consistent with migrant-driven team synergies and founder/firm sorting into urban cores, notably London. By contrast, we find few differences when comparing more high-tech or knowledge-intensive firms with the pooled sample.

Our current findings are descriptive, not causal. One constraint is that we are unable to observe companies forming, or co-founding decisions: these likely reflect factors also influencing firm outcomes, such as early exposure to entrepreneurship, parental wealth, attitudes or reputation (Akcigit et al., 2025; Bernstein et al., 2022; Hegde and Tumlinson, 2021; Lindquist et al., 2015). Conditional on founding, however, we observe a very rich array of individual characteristics, background and experience. In tests on co-founded firms, we also show that natives who co-found with migrants (vs. other natives) show no significant mean differences on relevant observables. The same pattern emerges when we look at migrant selection. A second constraint is that we do not directly control for cross-firm differences in top team or wider workforce composition, which may also influence firm outcomes (Certo et al., 2006; Kenney and Patton, 2015; Triana et al., 2019; Mack et al., 2025). We use rich controls, detailed fixed effects and diagnostic tests to mitigate these endogeneity challenges. A third issue is that founders likely select into high-potential urban locations (Tareque et al., 2024; Bryan and Guzman, 2021). Our current analysis does not control for such sorting.

Related literature. This paper contributes to several literatures, detailed in Section 2 and summarised here. First, we add to the large literature on migrant entrepreneurs by providing new evidence on their role shaping high-growth firms. The closest papers to this one are Jin et al. (2025), who explore founding team synergies using similar US data, and Lee et al. (2025), who link US survey and administrative data to explore migrant entrepreneurship and innovation. Using similar data to Jin et al, Gompers and Woo (2026) descriptively compare expat, returnee and local founders across multiple countries, focusing on the expat / returnee premium on external finance and successful

exits. Our paper differs in its focus on high-growth entrepreneurship, on geographies of entrepreneurial activity, and in exploring firm positioning alongside conventional measures of performance. Given the distinctive role of skilled migrant entrepreneurs there, US-based findings may also not generalise. Second, we contribute to firm-level work on innovation-driven startups (Akcigit et al., 2025; Botelho et al., 2026), by linking firm outcomes to unusually rich information on founders. Third, we build on a large body of work on urban entrepreneurship (Duranton and Puga, 2001; Glaeser et al., 2010a; Guiso and Schivardi, 2011; Andersson and Larsson, 2014; Chatterji et al., 2014; Glaeser et al., 2015) by directly considering the role of migrant founders in cities. Finally, from a methodological perspective, we contribute to a growing literature that combines online and firm-level microdata to explore an array of research questions (see Gray et al. (2026) and Dahlke et al. (2025) for reviews).

The rest of the paper proceeds as follows. Section 2 gives a brief conceptual framework. Section 3 describes the data, build and key variables. Sections 4 and 5 present descriptives for founders, companies and cities, and compare migrant vs. UK-born founders (Facts 1-3). Section 6 explores founder demography and firm outcomes (Fact 4). Section 7 analyses founding team composition and firm distinctiveness (Fact 5). Section 8 looks at the role of sending country, language and location across the UK urban system (Fact 6). Section 9 concludes.

2 Framework

Classic models of entrepreneurship (Knight, 1921; Schumpeter, 1962; Lucas, 1978; Rosen, 1982; Evans and Jovanovic, 1989) view entry and outcomes as primarily influenced by individual attitudes, such as high risk tolerance; resources, including family wealth; and entrepreneurial abilities. These latter 'skills' may include formal qualifications, years of experience, number of jobs held, previous entrepreneurial activity or management capabilities (Akcigit et al., 2025; Hegde and Tumlinson, 2021; Lazear, 2004); for innovation-driven founders, STEM qualifications also matter (Burton et al., 2002). More broadly, entry costs and learning opportunities influence both entrepreneurship decisions and venture performance. These vary across locations (Guiso and Schivardi, 2011; Andersson and Larsson, 2014; Guiso et al., 2015), with dense urban areas offering lower entry costs and greater learning opportunities (Glaeser et al., 2010a,b, 2015; Davis and Dingel, 2019; Duranton and Puga, 2023). Founders are therefore likely to sort into areas more promising for entrepreneurship. Entrepreneurs also decide whether to work alone or in teams, and this decision reflects both individuals' own assessment of their abilities, and the availability of co-founders. Developing distinctive ideas has high payoffs but is risky, complex and time-consuming, often requiring specialist skills and knowledge (Howell and Hall,

2026; Gans et al., 2019; Rajan, 2012; Luttmer, 2011). High-growth potential firms may therefore perform better with teams than solo-founders; however, poorly matched teams will damage the success of the venture (Botelho et al., 2026; Honoré, 2022; Calder-Wang et al., 2021; D'Acunto et al., 2020).

A large literature also documents the outsized role of **migrants in entrepreneurship**. Migrants are overrepresented among the self-employed, among founders of employer firms, and among founders of high-growth companies. They also play prominent roles in high-technology clusters such as Silicon Valley (for a review, see Chodavadia et al. (2024)). However, there is less agreement about the mechanisms underlying these patterns. Given existing theories of entrepreneurship, why might *migrant entrepreneurs* differ from other entrepreneurs in their propensity to start firms or in the trajectories of the ventures they create?

One set of explanations focuses on **individual entrepreneurs' characteristics**. Migrants are often positively selected on human capital (Hunt and Gauthier-Loiselle, 2010; Lee et al., 2025; Beine et al., 2024), and migrant entrepreneurs may also differ in attitudes toward risk, ambition, or opportunity recognition (Kleinhempel et al., 2023). They may possess informational advantages, including knowledge of international markets or the ability to identify underserved niches in host economies. If migrants disproportionately possess such endowments, the firms they found may achieve outsize success (Azoulay et al., 2022). At the same time, apparent migrant advantages may reflect **compositional effects** rather than causal differences. Entrepreneurs are not a random draw from the workforce: they tend to be disproportionately male, middle-aged, and relatively experienced. If migrants share similar demographic profiles, differences in entrepreneurial representation or venture outcomes may simply reflect these underlying characteristics. Recent work emphasises that once observable characteristics such as age, gender, and human capital are taken into account, migrants may not be inherently more entrepreneurial or more successful than natives (Manning, 2025). Moreover, migrants may enter entrepreneurship because barriers in mainstream labour markets limit alternative opportunities (Kloosterman and Rath, 2003; Jones et al., 2018). Negatively-sleected entrepreneurship may reflect necessity rather than opportunity, and discrimination may constrain firm survival and growth (Ewens, 2022; Fairlie et al., 2022).

A second set of mechanisms emphasises the role of **founding teams and networks**. Many ventures are created by teams rather than individual founders, and as above, team composition may shape both the entry process and subsequent firm performance. Specifically, founder sameness and founder difference might affect firm positioning and performance. For instance, co-ethnic networks among migrant founders may lower entry costs by facilitating access to labour, finance, and market information. Shared ethnic and migrant backgrounds can support trust and knowledge exchange within entrepreneurial communities (Aldrich and Waldinger, 1990; Kerr, 2010; Kerr and Mandorff, 2023). On

the other hand, diversity within founding teams may generate advantages. Given costs and complexity of entrepreneurial search, differences in experience and perspectives can enhance problem solving and idea generation (Hong and Page, 2004), potentially improving both innovation and firm management. Founding teams combining migrants and natives may therefore benefit from complementarities in knowledge and networks, and these might be reflected in firm positioning and firm outcomes. However, per the literature on suboptimal teams, diversity may also introduce coordination frictions, making the overall effect of team composition theoretically ambiguous.

These mechanisms operate within **spatial contexts**. A third strand of the literature highlights the role of **location characteristics**, particularly those associated with cities, in shaping entrepreneurial activity and firm performance. Complex economic activities tend to concentrate in urban areas where dense labour markets and localised knowledge spillovers facilitate learning and collaboration (Balland et al., 2020; Kemeny and Storper, 2020; Koster and Ozgen, 2021). These dynamics may be especially important for startups and young firms, for which large cities can act as "nurseries" (Duranton and Puga, 2001).

Taken together, these perspectives suggest that the formation and trajectories of migrant-founded firms may reflect the interaction between *individual founder characteristics, founding team composition, and spatial context*. Yet relatively little empirical work studies these mechanisms jointly, particularly for high-growth firms and outside the United States. The UK provides a particularly useful setting in which to examine these questions. While there is extensive evidence on migrant entrepreneurship in the United States, much less is known about the composition and performance of migrant-founded high-growth firms in the UK, despite the country's openness to skilled migration and the strong spatial concentration of innovative activity, especially in London and other large cities. Moreover, most existing work in the UK focuses either on individual founders or on firms as homogeneous units, paying less attention to the demographic composition of founding teams and the potential complementarities between migrant and native entrepreneurs (Hyde et al., 2023; Ives et al., 2023).

3 Data

We develop rich, linked information on high-growth companies and their founders. Our company-level data come from Beauhurst, a commercial provider that tracks high-growth and high-growth potential companies in the United Kingdom.[4] We link founders described in Beauhurst to individual founder profiles in Diffbot, a very large commercial knowledge graph database constructed from the public web.[5] Below we provide more details of these data.

[4]See `www.beauhurst.com` for more detail.
[5]See `www.diffbot.com` for more detail.

Beauhurst is a dataset of 'high-growth potential' UK companies, including company level characteristics (age, industry, headcount, etc), geolocated HQ and plant trading addresses, detailed rich-text descriptions and financial performance measures (external finance, revenues). Companies are selected on the basis of one or more lifecycle events, which Beauhurst treats as 'growth signals'. Some of these signals are early-stage (being an academic spin-out, attending an accelerator programme), others typically mid-stage (receiving angel or VC finance) and others later stage (high-revenue growth or employment growth episodes). Section 4.2 gives more detail of these signals. Beauhurst start with the UK's open companies registry, Companies House, the population of companies in the UK, then for tracked firms, enrich this through multiple, validated, other sources.[6] Our raw dataset covers 44,100 active and extinct companies, of which 33,400 companies are founded 2000-2020 inclusive. The dataset also covers 257.9k 'key individuals', including firm founders, company officers, other top employees and investors/shareholders. For these individuals we have names, ages, gender, nationalities, and information on previous founder / employee / board roles. Beauhurst identifies 33.5k of these individuals as founders, based on their job titles.

Diffbot is a global knowledge graph database built using factual information from the public internet, using feature extraction and supervised learning build an array of derived variables (Mesquita et al., 2019). Graph databases work with factual data, linking entities (such as firms, founders, workers and places) through meaningful relationships (such as founding activity, employment histories and trading addresses). As of January 2025, Diffbot's graph described 231.3M individuals with employment histories globally, of which 10.6M were in the UK. For these individuals, who include company founders, Diffbot profiles include variables covering names, age, gender and detailed education and career histories, including details of job spells, roles and work locations. Diffbot also contains granular 'skills' for individuals, based on profiles, and leveraging a validated typology of 32,000 skills. Importantly, Diffbot's graph includes company identifiers from open company registers from around the world, including Companies House covering the UK. Building on methods developed in Gray et al. (2026), we use this feature to link organisations and people in Diffbot to companies and founders in Beauhurst. We detail our workflow below.

3.1 Defining founders

We consider company founders as entrepreneurs—individuals who set up companies which then employ others. Because 'founder' has no legal definition, actually identifying com-

[6]Beauhurst's combination of structured sampling frame and validated data is an important advantage over crowd-sourced alternatives such as Crunchbase or Pitchbook, where the sampling frame is implicit, and where coverage across fields is variable (Botelho et al., 2026).

pany founders in data is not straightforward. Across the full timespan of our data, we identify 33,486 company founders and co-founders in Beauhurst by searching on validated roles and job titles. This is a conservative approach which precisely identifies a set of self-identified entrepreneurs. However, not all companies in our data list 'founders' or 'co-founders'.[7] We drop these companies from our analysis, reducing our sample of companies to 23,944. An alternative approach involves identifying the first-appointed director of a company, which identifies founding board members. We run tests showing this approach generates inflated founder counts.[8] Overall, we deem the significant reduction in company sample size acceptable given the corresponding gain in precision.

3.2 Founder — company sample

Beauhurst person profiles contain limited information on individuals. For identified founders, we first gapfill missing data by scraping Companies House for updated profile information. We then search for founders on Diffbot and extract person profiles, providing us with much richer information. To accomplish this, we start by searching for companies in Diffbot; for matched companies we further search for founders. We initially keep only matches where companies are matched by identifier or name and UK location, and where founders are matched by name. A research assistant manually validated matches for a random 1% sample of founders.[9] Finally, we use an LLM-based workflow to adjudicate edge cases where founders are matched on first name or surname only, using contextual information to resolve underlying matches/non-matches.[10]

We then use cleaning routines detailed in Gray et al. (2026) to build detailed Diffbot founder profiles, which we match to Beauhurst founders and companies. Our final sample consists of Beauhurst companies founded in 2000 or later, active and extinct,[11] which have identified founders, and whose founder/s are well-matched to Diffbot profiles. We further

[7]For example, some entrepreneurs may not see themselves as 'founders', rather as 'someone who owns a business' (email exchange with Beauhurst, December 2025).

[8]Companies may be running for some while, including with employees, before they formalise governance arrangements; many companies appoint whole boards at single time points; some individual directors sit on many boards. Taken together these issues imply inflated counts. Applying the first director definition to our data generates 122.5k 'founders', around four times as many as using the self-identified basis. Further, the median firm now has six 'founders' and the largest 'founding team' has 143 members. These numbers are too big to be credible.

[9]The RA classified 160 Diffbot profiles and corresponding Beauhurst founders. 94.4% were confirmed matches.

[10]Two examples illustrate the issue. Example 1: Edward E Johnson in Beauhurst / Woody Johnson in Diffbot. Both listed as a director of ABC Import Export, a company based in London with CRN GB123456. Our algorithm matches on the company and the surname. We accept the match: Woody is a known diminutive of Edward, and other contextual information checks out. Example 2: Luke Antonidis in Beauhurst / Diana Antonidis in Diffbot. Both listed as a director of XYS Strategies, a consultancy based in Manchester with CRN GN654321. Our algorithm matches on the company and the surname. We reject the match: Luke and Diana are very unlikely to be the same person; this is likely a family firm.

[11]In our final company dataframe 649 companies are extinct, or 5.33% of the sample.

restrict our sample to companies where we can identify the founder(s) as either migrants or UK-born, using a rules-based approach we define below.

From our initial sample of 33,486 founders we match 31,767 to Diffbot profiles, and keep 26,379 matches, an overall match rate of 78.8%. Restricting to firms founded in 2000-2020 further reduces our sample to 23,641 founders. Restricting to founders where we can identify migrant/native status reduces our sample to 15,505 founders. Correspondingly, our starting sample of 23,944 companies with identified founders reduces to 22,689 where founders are matched to Diffbot, 19,739 with well-matched founders, and to 16,684 companies founded between 2000 and 2020. Taking out companies where any founder is non-identifiable as a migrant or native reduces the sample to 12,182 companies.

3.3 Defining migrants

A fundamental constraint in migrant entrepreneurship research is that founder data usually excludes migrant status. In our case, neither Beauhurst or Diffbot directly observes country of birth. Both Beauhurst and the UK companies registry contain nationality information, which is often used as a proxy for migrant status (Acosta and Marinoni, 2025). However, nationality is endogenous. Like most countries, the UK offers migrants paths to citizenship: in 2021, 43% of UK migrants held UK passports (Migration-Observatory, 2023). In theory, using nationality as a noisy proxy might give a lower bound on true migrant counts. One problem with this approach is that a migrant's decision to become a citizen may be directly related to their current labour market status, including founder activity. A further, UK-specific problem is that the 2016 Brexit vote led many EU migrants based in the UK to apply for British nationality, potentially introducing more structural bias into the measure during our sample period.[12]

Instead, and in line with Lee and Glennon (2023), Jin et al. (2025) and Beine et al. (2026), we proxy for migrant status by using founders' country of education, focusing on the lowest observed level of education in Diffbot profiles (typically undergraduate university education). In these studies, which use US data, country of education is viewed as giving a lower bound because US universities bring in large numbers of foreign students. Because UK universities also attract more foreign students than UK students study abroad—in 2022/3, 26% of all UK students were born abroad (HESA, 2024)—this proxy is likely to give a lower bound on migrant founders in our UK setting too. We further test our approach using a bespoke online survey of UCL staff. We find that school / university education gives a nearly 99% correct attribution of true migrant status, as

[12]Prior to Brexit, EU migrants to the UK had freedom of movement, and had little incentive to become UK citizens compared to migrants from the rest of the world. The Brexit vote and subseuqent exit agreement substantially shifted those incentives for EU arrivals, although not for migrants from other countries.

well as the predicted lower bound.[13]

Where missing, we fill in education location information by combining Diffbot text and UK / global place-name dictionaries (see Gray et al. (2026) for details). However, 23.7% of founders cannot be identified as migrants or UK-born, almost entirely because their Diffbot profiles contain no education information. We run balancing tests to check how these sub-samples differ. Appendix Table A1 shows the results of balancing tests, where the dependent variable is either having identifiable migrant/native status, or having education information in Diffbot. Those included in our sample are younger, female, dual-nationals, or speak more than one language; they also have more jobs in their career history and are more likely to have management or tech experience. However, the size of these differences is small: around 4-6 percent when controlling for founder location, slightly larger otherwise. Controlling for founder sorting across locations, using city fixed effects, shrinks these differences further.

3.4 Outcome measures

As discussed in Section 2, theory and evidence on high-growth firms emphasises the importance of novel business models, processes and products: the resulting markups translate into higher revenue and employment. We therefore explore two types of company outcomes: business performance metrics, drawing on company accounts, and measures of company 'distinctiveness', which we build from company text data.

Business performance. Beauhurst contains an array of business performance data. Firms in Beauhurst have many potential growth paths, from sales or revenue-maximisation to VC-led scaling (building up user base / market share with or without revenue). To reflect this breadth, and as customary in the entrepreneurship literature, we construct simple dummy variables covering whether the firm has any revenue post-incorporation, whether the firm has any employees, whether the firm has any external finance, and whether the firm has had a successful exit, through acquisition or IPO. In our data, all exits are acquisitions. For the subsample of firms posting revenue or employment, we track these outcomes over time. See Section 6 for more detail.

Distinctiveness. We build measures of firm 'distinctiveness' relative to peers, based on companies' own descriptions of their strategies, products and services, taken by Beauhurst from firm websites. Here we follow a recent literature using natural language processing to explore company characteristics (Draca et al., 2023; Guzman and Li,

[13]We surveyed a large UCL department, including senior, mid-career and early career researchers, PhD students and administrative staff. Respondents were asked to give their countries of school and (if any) university education, country of birth, nationality. While limited in scope - we survey a single education institution, rather than a larger set of companies - we achieve a response rate of 35 percent, which is defensible in online surveys.

2022; Hoberg and Phillips, 2016). Appendix Figure A1 gives an example of the text.[14] Guzman and Li (2022), who use similar text data, refer to such text as describing firms' 'positioning' or 'value proposition'.[15]

Adapting the workflows from Draca et al. (2023) and Guzman and Li (2022), we clean the text from firms' own descriptions (using Beauhurst analysts' text as a fallback), train a Word2Vec model on the resulting text, and compute semantic similarities between each pair of companies using cosine similarity metrics. We also build alternative metrics using TF-IDF instead of word vectorisation. Results are shown in robustness tests and are very similar.

Because 'novelty' does not straightforwardly translate to firm positioning, we build three types of metrics. First, we count the number of peers for each firm against all other firms in the sample. We count pairs across all cohorts; across previous cohorts, and across firms founded in the same year. To define peers we use a data-driven similarity threshold, which we set as the 95th percentile of all pairwise similarities for each timeframe. We run robustness tests with more conservative thresholds. Next, we build alternate peer counts conditioning on pairs in the same SIC3 industry. Finally we derive two non-threshold-based metrics: 'originality', defined as (1 - maximum similarity of any company founded in earlier years), and 'trendiness', defined as a company's mean similarity with all others founded in the same year.

More distinctive companies will have fewer peers, both overall and within nearby industry space. They are also likely to have more 'original' positioning and have less in common with peers founded in the same year. Per Section 1, pursuing such original business models, products and services gives an *a priori* ambiguous link to subsequent firm outcomes. More novel companies are more likely to obtain external finance, especially from venture capital firms (Botelho et al., 2026; Kerr et al., 2014). Novelty may then allow innovative entrants to charge higher markups, building revenue and market share, raising headcount and raising the likelihood of successful exit (Klette and Kortum, 2004; Gans et al., 2019). Conversely, novel positioning is more complex and risky. 'Excess' novelty might be *less* likely to receive external finance, pick up market share and allow firms to scale (Rajan, 2012). Founders therefore have a choice about the level of novelty vs. adopting others' ideas. With this framework in mind, we benchmark our metrics by regressing measures of company distinctiveness to *company-level* performance measures, after controlling for founder and company characteristics, industry, location and cohort

[14]We validate this interpretation of the text by sampling 100 company descriptions, then asking an RA and two different LLMs to classify the text as describing either 1) strategy 2) products and services 3) both 4) neither or 5) unable to classify. Both LLMs and the RA agree, classifying between 86 and 93 percent of the descriptions in classes 1)- 3).

[15]Note that Guzman and Li (2022) use historic website text to describe founders' original positioning; Beauhurst uses current/recent text, which likely describes the company's contemporaneous positioning or positioning at the point of tracking. To get closer to founders' initial decisions, we run robustness tests using the youngest firms in our sample.

effects.

Appendix Table A2 shows the extensive margin for linear probability models where the outcome is a dummy taking the value one if the firm achieves revenue, hires staff, attracts external finance, or achieve successful exits (IPO or acquisition). More novel companies are *more* likely to get some external finance after founding, but are *less* likely to add staff, achieve any revenue or exit. Appendix Figure A2 shows the intensive margin for employment and revenue, two outcomes we can track over time, across each distinctiveness measure. Conditional on receiving *any* employment or revenue, more novel firms have *higher* employment and revenue growth than their less distinctive counterparts.

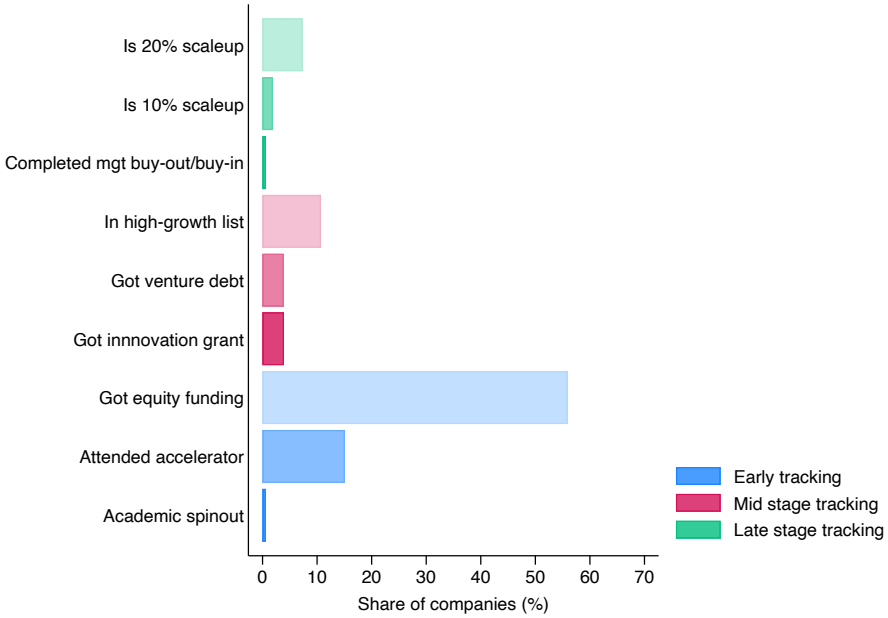
3.5 Sample benchmarking

By design, companies in Beauhurst are selected on observables that the firm's analysts believe positively predict future performance. Benchmarking against business population data confirms that our company sample is – not surprisingly – unlike the UK business population as a whole, and differs on observable characteristics. Appendix Table A3 provides a high-level comparison of our sample against the UK business population for 2010 (the first year demographics are available), 2015, and 2020. The sample comprises between 0.1-0.14% of all private sector firms and consistently has around four times as many employers. The share of SMEs in the sample is four to five times lower in 2010 and 2015, but has largely converged in 2020, the last sample year (where there will be proportionately more new/small companies in the data). In 2020, where public data is available, the sample has a third more high-growth firms than the business population.

Given these differences, we might worry that our sample selects on the performance outcomes we care about. Alternatively, if our sample is selected on high-growth *potential*, rather than actual high performance, this gives us meaningful variation to work with. We provide three pieces of evidence in support of this second scenario.

First, Beauhurst's growth signals – events that bring companies into the data – are predominantly early or mid-stage events, based on average company founding year for each signal. Specifically, Figure 1 shows that over 65 percent of companies are tracked at early stage (blue bars), around 25 percent at mid stage (red), and just under 10 percent at late stage (green). Of the latter, 9.3 percent of companies are tracked on the basis of being 'gazelles': exhibiting three years of 10 percent or 20 percent growth in employment or revenue. Appendix Table A4 provides more detail. Second, actual company performance varies substantially. Summary statistics (Table 5, below) show that around 15% of firms never record headcount; over 80% never post revenue; nearly 30% do not raise external finance, and the vast majority do not exit. Third, in Section 6.3 we run robustness tests on company performance, excluding firms selected on the basis of employment/revenue growth. Results are largely unaffected, although revenue growth premia are fitted on a

Figure 1: Sample selection: breakdown by tracking reasons



Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive. Figure shows the sample breakdown by reasons for tracking. We bin these reasons into three groups based on the mean ages of firms. We define companies tracked on the basis of attending an accelerator or being an academic spinout as tracked at early stage. We define companies tracked on the basis of receiving major innovation grants, equity or debt finance, or cited in high-growth lists as mid stage. We define companies that have completed buyouts/buyins or are scale-ups as late stage. 41 companies have no tracking information. See Table A4 for other notes.

smaller sample and are less precisely estimated.

4 Descriptives

4.1 Founders

Founder characteristics. Table 1 details core characteristics of founders in our data. Consistent with the wider literature on successful entrepreneurs, the average founder age for these high-growth companies is over 47. Over 80% of founders are male. In both cases, there are modest raw difference between migrant and UK-born founders. We estimate that 26.5% of founders are migrants, measured by our country of education proxy. This is almost double the share of migrants in the UK working-age population and as explained above, is likely to be a lower bound. 32% of founders are non-UK nationals. Just under 1% of founders are dual nationals. Virtually all of these raw migrant-native differences are statistically significant.

Figure 2 visually summarises key characteristics of migrant and native founders. As well as similar age and gender profiles, migrant founders have similar years of labour market experience to native founders. However, they differ on levels and types of qualifi-

Table 1: Summary statistics and t-tests of mean differences by migrant status

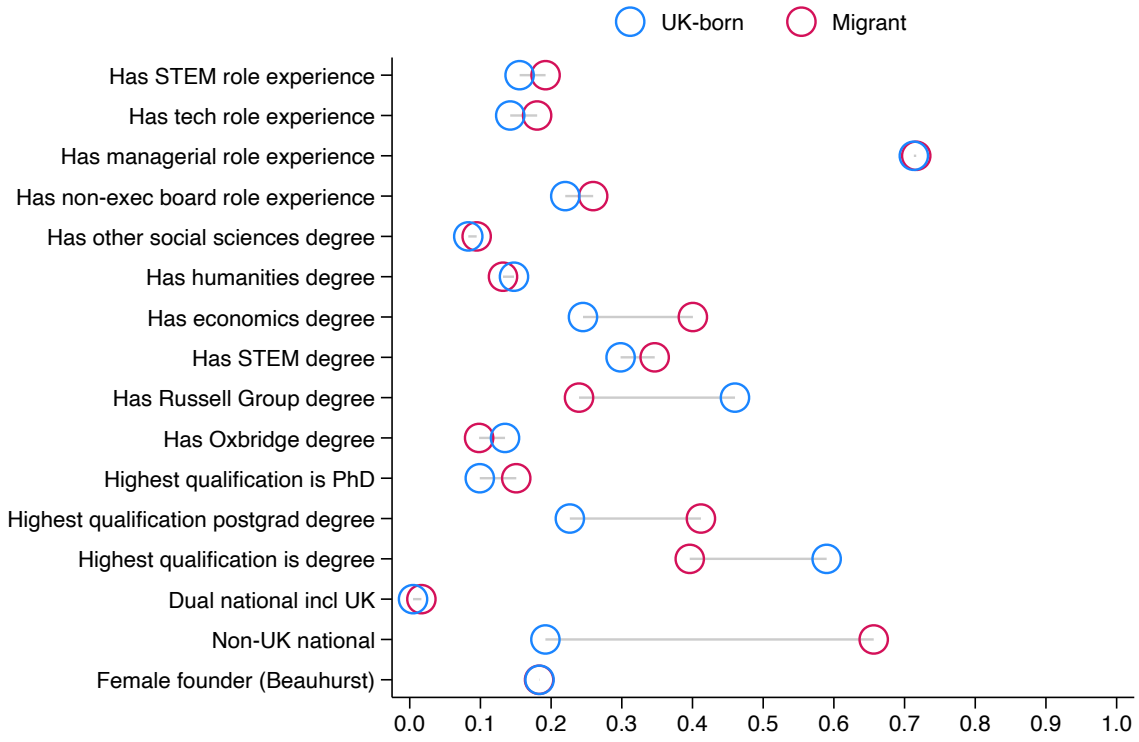
	All	UK-born	Migrant	Diff
Founder age	47.075	47.482	45.946	1.536***
Founder age (Beauhurst)	43.290	43.656	42.189	1.466***
Female founder	0.180	0.174	0.196	-0.022***
Female founder (Beauhurst)	0.178	0.178	0.180	-0.003
Non-UK national	0.309	0.188	0.654	-0.466***
Migrant founder	0.259	0.000	1.000	-1.000
Dual national incl UK	0.008	0.004	0.016	-0.012***
Highest qualification is degree	0.538	0.587	0.397	0.191***
Highest qualification postgrad degree	0.275	0.227	0.412	-0.185***
Highest qualification is PhD	0.115	0.102	0.151	-0.049***
Has Oxbridge degree	0.128	0.138	0.097	0.041***
Has Russell Group degree	0.406	0.464	0.238	0.226***
Has STEM degree	0.312	0.300	0.345	-0.046***
Has economics degree	0.288	0.247	0.403	-0.156***
Has humanities degree	0.143	0.146	0.133	0.013**
Has other social sciences degree	0.086	0.083	0.094	-0.011**
Number of languages spoken	1.980	1.672	2.860	-1.189***
Years of experience	10.536	10.645	10.227	0.418***
Number of jobs held	7.242	7.008	7.901	-0.893***
Has non-exec board role experience	0.238	0.228	0.268	-0.040***
Has managerial role experience	0.714	0.713	0.717	-0.004
Has tech role experience	0.152	0.143	0.179	-0.037***
Has STEM role experience	0.165	0.156	0.191	-0.035***
Is currently a founder	0.454	0.439	0.498	-0.060***
Serial founder	0.072	0.077	0.059	0.018***
Observations	16164			

Source: Beauhurst, Diffbot. Sample is 15,505 founders of Beauhurst companies founded 2000-2020 inclusive. Cells show mean. Diff = mean difference in UK-born vs migrant founder means. Stars from a two-tailed t-test. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

cations, and types of professional experience. As Manning (2025) points out, such simple differences explain little if migrants and natives structurally differ on many dimensions (e.g. age, gender, qualifications). In section 5 we directly test selection on entrepreneurial capabilities (for example, management experience, years in the labour market or serial founding), controlling for other individual and company level characteristics.

(Note that in Table 1 and Figure 2 we include migrant-native comparisons by shares of founders attending elite UK university types: Oxbridge and the Russell Group (which includes UCL, LSE and Birmingham, among others). We find significant differences between migrant and native founders, which are quantitatively large for the Russell Group. However, these differences largely come from undergraduate attendance, which is also our main way of distinguishing migrants from natives (Section 3). So these raw differences should be viewed with caution. In Section 5 we test migrant-native differences both with and without these covariates.)

Figure 2: Founder characteristics by origin



Source: Beauhurst, Diffbot. Sample is 15,505 founders of companies founded 2000-2020 inclusive.

Founder location. Founders and companies are not evenly distributed: large minorities of both are in cities, especially London. Table 2 shows founders' *current* location, derived using Diffbot profile information. Note that this field is missing for around 50% of founders so should be interpreted with caution. Appendix Table A5 gives results of a selection test, suggesting some minor observable differences between founders with/out

this information.[16] Panel A shows all founders. Panel B shows migrant founders. Overall, founders are most likely located in London, big cities or university cities. Migrant founders have a somewhat different footprint, and are most concentrated in London, Oxbridge and Edinburgh. Notably, some migrant founders are not in the UK any more, with small groups in the US, India, Australia, South Africa, France, Italy, Germany, Switzerland and Russia.

Table 2: Founder current location: Top 20 locations

All founders			Migrant founders		
N	Share (%)	City	N	Share (%)	City
3258	21.01	London	971	23.95	London
172	1.11	Manchester	54	1.33	Cambridge
168	1.08	Edinburgh	38	0.94	Edinburgh
150	0.97	Cambridge	27	0.67	Oxford
143	0.92	Bristol	23	0.57	Manchester
112	0.72	Glasgow	20	0.49	New York City
90	0.58	Leeds	19	0.47	Paris
82	0.53	Oxford	18	0.44	Sydney
79	0.51	Birmingham	17	0.42	Glasgow
67	0.43	Newcastle Upon Tyne	13	0.32	Birmingham
56	0.36	Hove	13	0.32	Bristol
54	0.35	Cardiff	12	0.30	Moscow
43	0.28	Southampton	11	0.27	Melbourne
40	0.26	Nottingham	10	0.25	Milan
39	0.25	Sheffield	10	0.25	Johannesburg
39	0.25	Aberdeen	8	0.20	Athens
36	0.23	Bath	7	0.17	Newcastle Upon Tyne
31	0.20	New York City	7	0.17	Hove
31	0.20	Liverpool	7	0.17	Bengaluru
29	0.19	Reading	7	0.17	Zurich

Source: Beauhurst, Diffbot. Sample is 15,505 founders of companies founded 2000-2020 inclusive. Location data comes from Diffbot, and includes founders of UK companies who are now no longer in the UK. Area units are cities (em_city in Diffbot). Table shows top 20 locations for each founder type, by frequency.

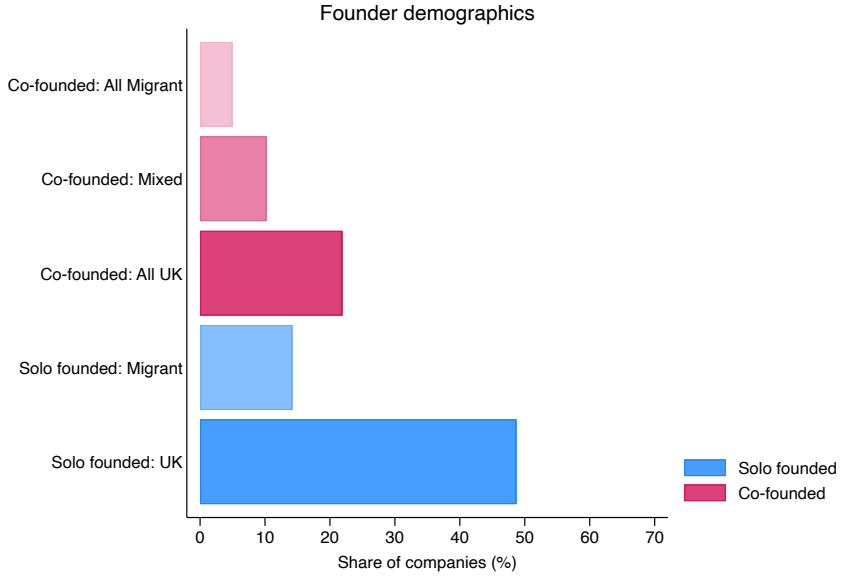
4.2 Companies

Companies by founder types. Figure 3 shows the distribution of solo and co-founded companies and by level of migrant involvement. Note that contrary to the high-growth

[16] Founders with Diffbot city location show no difference in migrant / sending country / nationality / gender compared to those without. However, the former are slightly younger, more likely to speak multiple languages, have slightly more labour market experience, are more likely to be a founder, and have management and tech experience; are less likely to have board experience. Oster tests suggest unobservables are unlikely to play a major role.

entrepreneurship literature, which emphasises the benefits of founding teams, almost two thirds of companies are solo-founded rather than have founding teams. This is consistent with our sample being predominantly selected on high-growth *potential*, as discussed in Section 3.5. However, migrants are more likely to be involved in co-founded firms than operate as solo founders.

Figure 3: Company types by founder demographics



Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive.

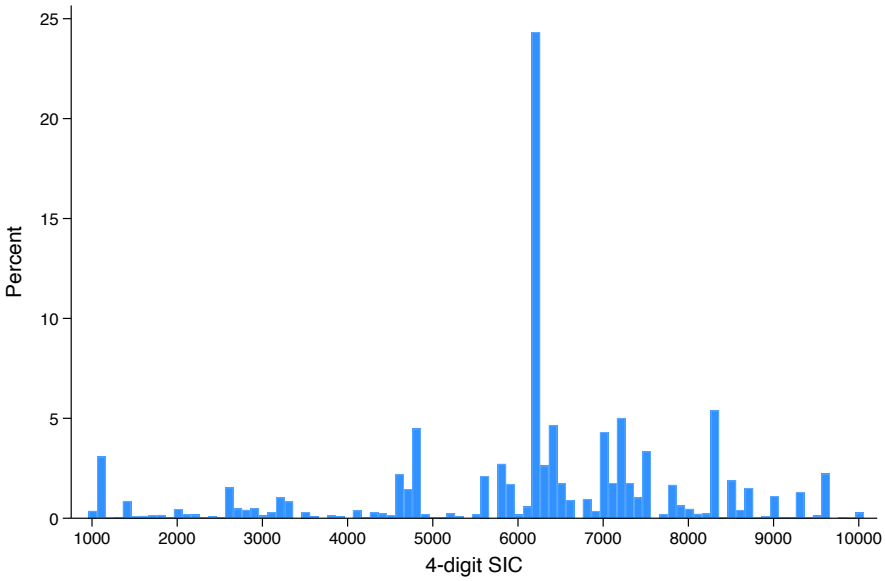
Table 3: Founding team composition

	Share (%)	N
At least one migrant founder	29.40	3,583
Sole founded company	62.90	7,661
Migrant founder	22.50	1,727
UK-born founder	77.50	5,934
Co-founded company	37.10	4,521
All migrant founders	13.50	609
All UK-born founders	58.90	2,665
Mixed founding team	27.60	1,247
Total	100.00	
N		12,182

Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive and their founders.

Table 3 decomposes the company sample by founding team types - note that here we show shares within solo/co-founded blocs. 62.9% of companies are solo-founded; 37.1% are co-founded. Only 22% of solo-founded firms have a migrant founder. By contrast,

Figure 4: Companies by 4-digit industry

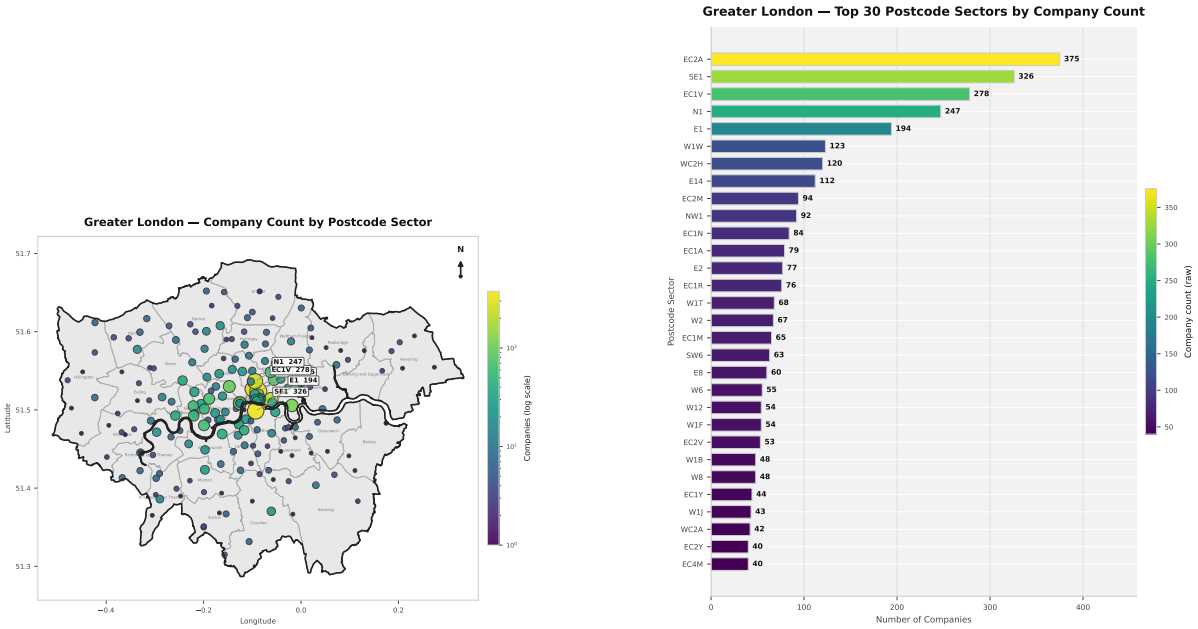


Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive. Histogram shows bins of 4-digit SIC07 codes (e.g. 6200-6299). Some bins contain only one code (e.g. 8000-8999).

over 40% of co-founded firms have at least one migrant founder; of these just over 27% have a mix of migrant and native founders, and over 13% have all migrant founders.

Companies by sector. Figure 4 is a binned histogram showing the distribution of companies by 3-digit industry code. The largest bin of companies is in 620 / ICT activities - specifically 6201 / Computer programming (13.15%), 6209 / Other information technology and computed service activities (7.49%) and 6202 / Computer consultancy (3.04%). Other large categories are 8299 / "Other business support service activities" (5.36%) and 7490 / "Other professional, scientific and technical activities" (3.08%). We run sensitivity tests with/out ICT activities. We also run extensions exploring sector heterogeneity across broader blocks of 'high-tech' and 'science and technology' activities. See Section 8 for details.

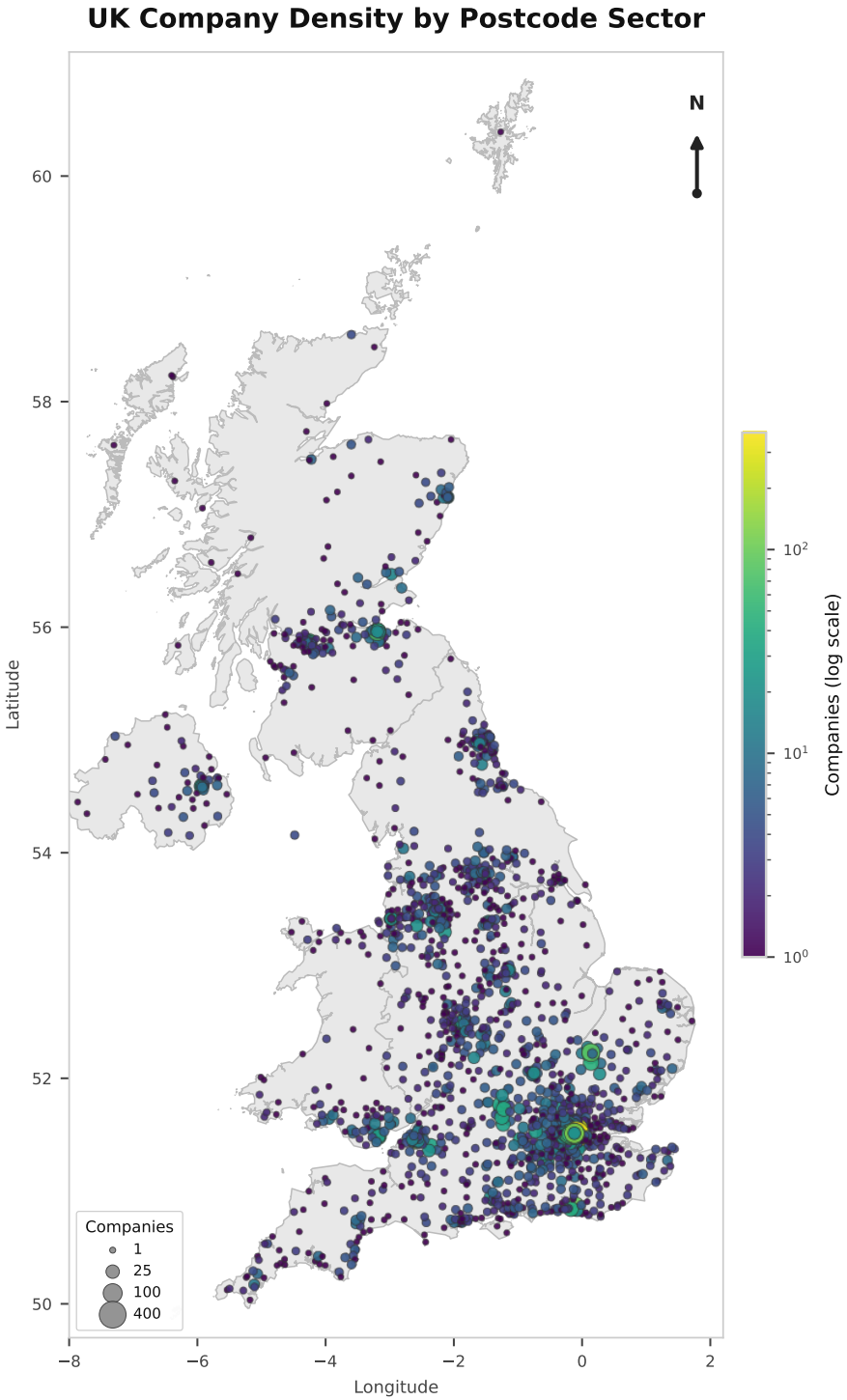
Figure 5: Company trading locations: London



(a) (b)

Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive. Company trading addresses aggregated to postcode sector.

Figure 6: Company trading locations: UK



Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive. Company trading addresses aggregated to postcode sector.

Table 4: Company Locations: Top 20 TTWAs

A. All companies			B. Migrant-founded		
N	Share (%)	City	N	Share (%)	City
5374	44.22	London	2161	60.50	London
437	3.60	Manchester	126	3.53	Cambridge
359	2.95	Edinburgh	104	2.91	Oxford
344	2.83	Cambridge	80	2.24	Manchester
310	2.55	Bristol	78	2.18	Edinburgh
274	2.25	Oxford	76	2.13	Slough & Heathrow
257	2.11	Slough & Heathrow	55	1.54	Belfast
251	2.07	Birmingham	47	1.32	Bristol
207	1.70	Glasgow	46	1.29	Guildford & Aldershot
184	1.51	Cardiff	42	1.18	Birmingham
179	1.47	Leeds	41	1.15	Glasgow
178	1.46	Belfast	34	0.95	Brighton
177	1.46	Guildford & Aldershot	32	0.90	Newcastle
170	1.40	Newcastle	31	0.87	Milton Keynes
168	1.38	Brighton	24	0.67	Southampton
101	0.83	Luton	23	0.64	Luton
100	0.82	Reading	22	0.62	Cardiff
99	0.81	Aberdeen	20	0.56	Leeds
92	0.76	Southampton	18	0.50	Leamington Spa
90	0.74	Milton Keynes	14	0.39	Aberdeen

Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive. Trading address data. Area units are Travel to Work Areas. Samples conditional on companies where trading address is known. Migrant-founded companies defined as companies that have a solo migrant founder, mixed migrant-native founding teams, or all-migrant founding teams.

Company location. In-sample companies are present in almost all parts of the UK, but are heavily concentrated in major cities, Oxford, Cambridge and Greater London. Figure 5 gives detail for the capital, based on main trading address postcode sectors, which contain an average of 3,000 individual addresses.[17] Companies are very strongly concentrated in central London neighbourhoods (Panel A), with raw counts of firms dominated by a few central postcode sectors (Panel B). Figure 6 shows individual company trading locations across the UK.

Table 4 shows company counts and shares at the Travel to Work Area (TTWA) level. TTWAs are self-contained commuting zones, which we use here as proxies for local spatial economies. The table focuses on the top 20 TTWAs by company location counts. Panel A shows all companies. Panel B shows migrant-founded companies (defined as companies with at least one migrant founder). Results are broadly consistent with the founder location analysis above.[18] Nearly 40% of companies are in London, with

[17] Beauhurst provides full trading addresses; we aggregate to postcode sector level for speed.

[18] Founder and company-based location measures differ on two dimensions. First, coverage: founder location has over 50% missings; company location coverage is close-to-complete. Second, what is measured:

other concentrations in contiguous 'mega-London' TTWAs such as Slough and Heathrow, Guildford and Aldershot, Luton and Reading. The next largest shares are in a mix of large conurbations (Manchester, Bristol, Birmingham, Leeds, Newcastle, Glasgow, Cardiff), and university cities (Cambridge, Edinburgh, Oxford, Aberdeen, Southampton). Migrant-founded companies show a much stronger London footprint, with nearly 60% in the capital, and rather higher shares in Cambridge and Oxford. Shares in other TTWAs are lower than for companies as a whole.

4.3 Company outcomes

Table 5 reports summary statistics for the 12,182 companies in our sample, founded between 2000 and 2020.

The upper panel shows company demographics. The middle panel shows business performance outcomes. Headcount data are available for 85% of firms, while revenue data are much sparser, observed for around 18%. Nearly three-quarters of firms received some form of external finance over the sample period. By contrast, exit events — acquisitions or IPOs — are relatively rare: acquisitions each occur for around six percent of firms, and there are no IPOs in the sample. The lower panel describes our text-based measures of firm distinctiveness. The mean originality score —measuring textual distinctiveness relative to previously founded firms — is 0.71, with relatively low dispersion, suggesting most firms occupy moderately novel positions in idea space. Around 40% of firms are unique within their three-digit SIC industry bin, while the average similarity to the rest of the founding cohort is low (0.024), consistent with substantial heterogeneity across firms. Firms face on average around 1,358 peers in the full sample, but far fewer—around 5—within their narrower industry bin. Appendix Table A7 reports unconditional mean outcomes by urban system category (from the ONS 2011 Urban-Rural typology) and team type at each firm's last observed year. The raw means show lower employment and revenue for all-migrant firms outside major conurbations, with the gap narrowing inside major conurbations. We do not interpret these unconditional differences as causal: they conflate location with industry mix, cohort, and team-level characteristics that we know vary across the urban hierarchy. The unconditional patterns motivate the regression analysis that follows, which asks whether team-type effects differ inside and outside London once these confounders are absorbed.

we observe founders where they are *currently*, which may not be the same as where their companies trade. These metrics give us different pictures of migrant founders in cities like London. In the 2021 Census, London's migrant population share was 41%. in our sample, the share of - observed - migrant founders currently based in London is just under 24%. Using company location as a proxy for founder location, just under 63% of migrant founders are London-based.

Table 5: Companies: summary statistics

	N	Mean	SD	Min	Max
Year founded	12178	2013.302	4.540	2000.000	2020.000
Active	12178	0.947	0.225	0.000	1.000
Number of child companies	12178	0.549	2.594	0.000	119.000
#companies in corporate structure	12178	2.510	11.856	1.000	456.000
Multisite	12178	0.543	0.498	0.000	1.000
Any revenue	12177	0.177	0.382	0.000	1.000
Revenue £m	2443	16.630	103.547	-0.107	3756.800
Any headcount	12177	0.854	0.353	0.000	1.000
Headcount	10397	24.538	119.446	1.000	5086.900
Any external finance	12177	0.721	0.449	0.000	1.000
Total external finance raised £m	8621	6.362	40.520	0.003	1314.490
Any Exit	12177	0.057	0.232	0.000	1.000
Any acquisition	12177	0.057	0.232	0.000	1.000
Any IPO	12177	0.000	0.000	0.000	0.000
10% scaleup	12178	0.019	0.137	0.000	1.000
20% scaleup	12178	0.073	0.261	0.000	1.000
Number of peers	12177	2012.482	2649.389	0.000	16332.161
Number of prior peers	12177	1326.516	1948.884	0.000	14720.000
Number of cohort peers	12177	113.851	152.297	0.000	1021.000
Number of SIC3 peers	12154	101.001	256.769	0.000	2785.000
Number of SIC3 prior peers	12145	61.632	163.919	0.000	2463.000
Number of SIC3 cohort peers	12177	7.332	20.252	0.000	253.000
Originality vs. prior firms	12177	0.098	0.033	-0.000	0.357
Trendiness (mean cohort similarity)	12177	0.657	0.059	0.370	0.788

Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive. Per main text, not all companies report headcount, employment or external finance.

5 Comparing Migrant and UK-born founders

Are migrant and UK-born founders different, and if so how? In our raw data (Section 4.1), when comparing migrant and UK-born *entrepreneurs* we find some significant differences in individual characteristics and substantial clustering of founders and firms, especially in London, big cities and university cities. However, these differences might be highly correlated with each other, or explained by other individual, company or location characteristics. Because everyone in our data is a founder, we cannot directly test whether these characteristics affect entry into entrepreneurship. However, conditional on entry we can formally compare migrant and UK founders' selection on observable entrepreneurial skills as described in Section 2. We run two descriptive exercises to explore this. Migrants and natives who co-found with each other may also be different from those who co-found only with other migrants or natives. For co-founded firms, we run selection regressions to compare the characteristics of these groups.

5.1 Entrepreneurial skills

First, we look at correlates of migrant entrepreneurship. For founder i, company f, industry j and trading location a we estimate:

$$\text{Migrant}_{ifaj} = a + \mathbf{X}b_i + \mathbf{X}'c_f + A_a + J_j + e_{ifja} \quad (1)$$

Where **X** is a vector of individual level characteristics, **X'** is a vector of company level characteristics, J is a 3-digit industry dummy and A is a Travel To Work Area (TTWA) dummy for the main trading location. We estimate equation (1) as a linear probability model (LPM). The vector of $\hat{b}_i$ shows the salience of individual characteristics on migrant founders, relative to UK-born founders, controlling for company characteristics and location.

Challenges here include: first, that migrant founders may be selected on unobservable characteristics (such as being risk-loving); and second, that founding decisions may be influenced by the availability of co-founders, which may be correlated to industry and/or area-level conditions (milieux effects). To mitigate the first issue we run Oster tests (Oster, 2019) to gauge the relative importance of unobservables. To handle other challenges, we estimate models with dummies for co-founded companies and the number of founders, and include detailed industry and TTWA dummies.[19] Individual covariates are summarised in Figure 7. Migrant entrepreneurship is positively correlated with founders having higher-level qualifications, especially post-graduate and PhD level; having an economics degree; number of languages spoken, and (as expected) being a non-UK-national. It is negatively correlated with being a serial founder and having a Russell Group university degree.[20] Appendix Table A8 gives full results including company controls.

Second, we look more precisely at potential migrant selection on observable entrepreneurial skills, which we define as as a vector **E** including: years of work experience; number of employments; qualifications; being a current founder; serial entrepreneurship; management experience; non-executive board member experience.

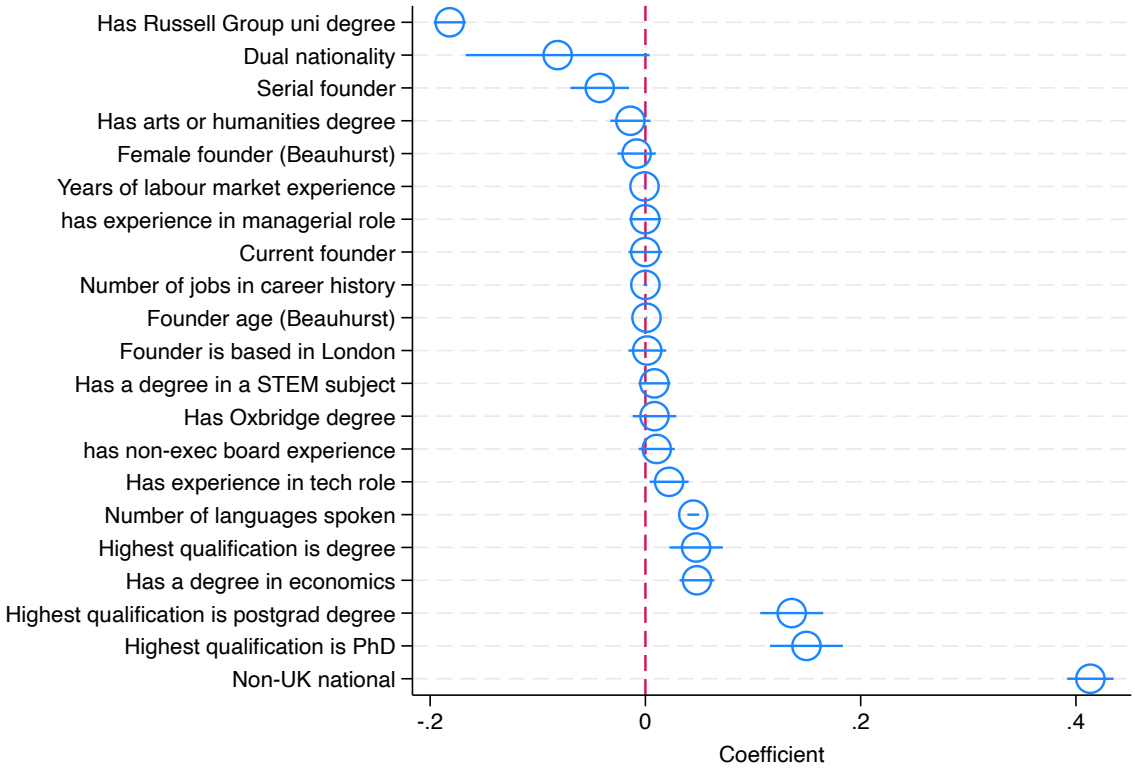
For each variable E in **E** we estimate:

$$E_{ifja} = a + b_i\text{Migrant}_i + \mathbf{Z}b2_i + \mathbf{X}'c_f + J_j + A_a + u_{ifja} \quad (2)$$

[19] Ideally we would want to include company fixed effects. For companies with a solo founder, company fixed effects are entirely collinear to the outcome so we lose these observations.

[20] Because we typically define migrant status on where founders studied at university, undergraduate degree and Oxbridge/Russell group dummies are arguably bad controls. In sensitivity tests we remove these predictors. The overall pattern of other results remains identical. Results available on request.

Figure 7: Predictors of migrant entrepreneurship



Source: Beauhurst, Diffbot. Sample is 15,505 founders of companies founded 2000-2020 inclusive. Figure shows coefficients and 95 % confidence intervals of a linear probability model where the dependent variable is being a migrant founder, and individual-level predictors are shown above. Regression also includes controls for number of founders in the company, company incorporation year, Beauhurst tracking stage, 3-digit industry and Travel To Work Area (TTWA) location dummies.

We estimate equation 2 using OLS or an LPM for binary outcomes. $\hat{b1}_i$ picks up the relative 'effect' of being a migrant, rather than a native, on founders' level of qualifications; years of experience or number of jobs, or on the probability of having management experience or being a serial entrepreneur, conditional on other individual level characteristics (**Z**), company characteristics (**X'**), and company industry and location as before.

Results are shown in Table 6. Migrant founders are significantly more likely to have higher-level qualifications than the UK-born, but less likely to be serial entrepreneurs. On other entrepreneurial skills, we find no evidence of significant migrant-native founder differences. However, these regressions are almost all underpowered, with Oster tests suggesting that unobservables—for example, differences in attitudes—explain much of the result.[21]

[21]Seemingly Unrelated Regressions, which explore migrant selection on all entrepreneurial skills simultaneously, confirm these patterns. Results available on request.

Table 6: Entrepreneurial skill selection

	(1) Postgrad	(2) PhD	(3) Experience	(4) Number of jobs
Panel A				
Migrant	0.097***	0.044***	0.122	0.050
	(0.011)	(0.008)	(0.206)	(0.158)
N	14671	14671	14671	14671
R^2	0.186	0.242	0.145	0.100
F-statistic	144.045	57.628	87.028	52.632
Oster delta, Migrant	1.577	4.600	-0.472	0.147
	Current Founder	Serial founder	Non-exec	Management
Panel B				
Migrant	-0.005	-0.023**	0.014	-0.002
	(0.019)	(0.009)	(0.012)	(0.011)
N	6684	14671	14671	14671
R^2	0.132	0.053	0.066	0.078
F-statistic	24.541	8.176	25.871	40.459
Oster delta, Migrant	-0.152	5.294	1.080	-0.419

Source: Beauhurst, Diffbot. Sample is 15,505 founders of companies founded 2000-2020 inclusive. Experience = years of career. Non-exec = has/had a non-executive role in career. Management = has/had a management role in career. All regressions fit founder level controls (age, gender, non-UK national, dual nationality, highest qualification dummies [degree, post-graduate degree, PhD], degree type dummies [STEM, arts and humanities, economics, other social science], experience in tech role, based in London) and company controls (number of founders, company incorporation year, TTWA location, 3-digit industry and tracking stage). Standard errors clustered on founder. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.2 Testing for founder sorting

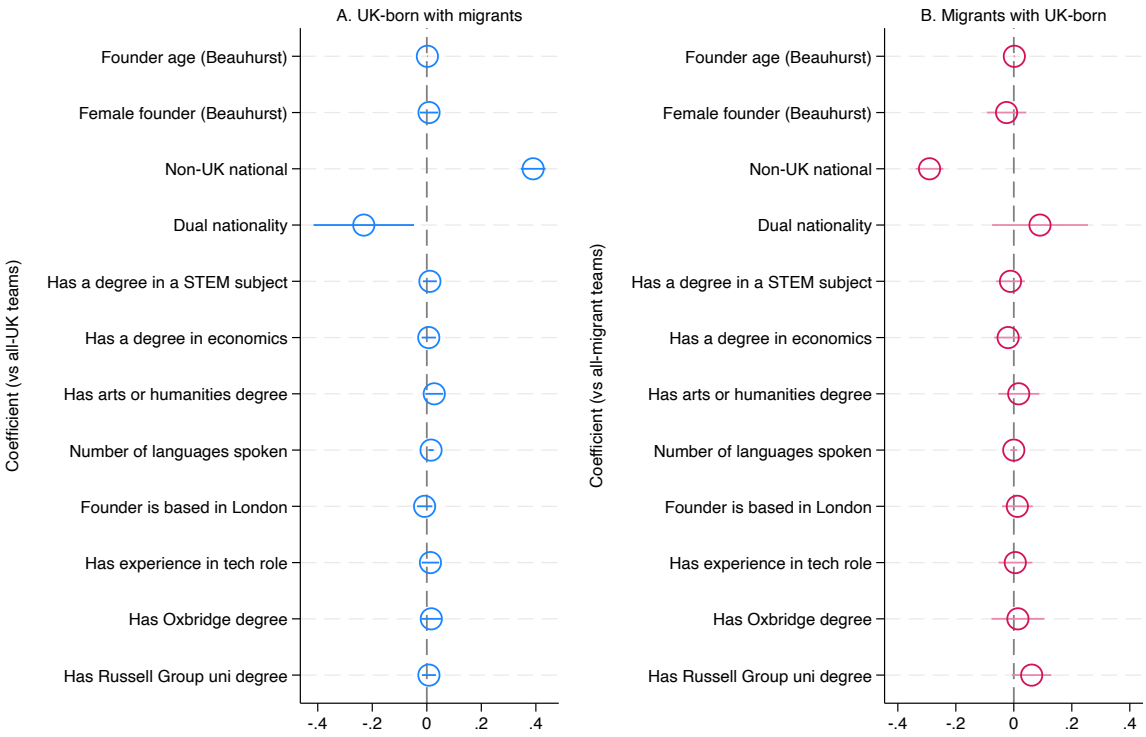
Everyone in our data is a founder, so we cannot directly observe either the founding decision or the decision to co-found. This is an important omitted variable: an alternative explanation of performance differences in migrant-founded firms is high-quality founder matching rather than team synergies. For instance, UK-born founders who set up with other natives may be systematically different from those who co-found with migrants. Since co-founding is a two-sided decision, *migrants* co-founding with other migrants may also differ from those in mixed teams. We run tests to assess the likely extent of founder sorting, for both natives and migrants.

Appendix Figure A3 shows selected mean differences in observables for UK-born founders in co-founded firms, comparing those co-founding with migrants against those co-founding with other UK-born. Quantitatively, raw differences in human capital and experience between those who co-found with migrants vs. those who co-found with other natives are very small.

We then run selection tests for co-founded teams, focusing on on native (migrant) founders' decision to co-found with natives (migrants) or not. We estimate linear prob-

ability models where the dependent variable is a dummy taking the value 1 for being in a mixed team and 0 for being in all-UK (all-migrant) teams. We include individual-level predictors **X**, company controls **X'**, TTWA and SIC3 dummies per Equation 1. Figure 8 plots coefficients and 95 % confidence intervals for native co-founders (Panel A) and migrant co-founders (Panel B). Reassuringly, almost all coefficients are small and non-significant. Appendix Table A6 gives results. Oster tests on betas of non-UK nationals, the predictor with the largest influence, suggest unobservables play little role in our main findings.

Figure 8: Native / migrant selection in co-founded firms



Source: Beauhurst, Diffbot. Sample is founders of co-founded firms, 2000-2020 inclusive. Figure shows coefficients and 95 % confidence intervals of linear probability models where the dependent variable is co-founding in mixed teams, and individual-level predictors are given above. Panel A shows results for UK-born founders, where the alternative is founding in all-UK teams. Panel B shows results for migrant founders, where the alternative is founding in all-migrant teams. Regressions also include controls for number of founders in the company, company incorporation year, Beauhurst tracking stage, 3-digit industry and Travel To Work Area (TTWA) location dummies.

6 Company performance and founder characteristics

6.1 Post-founding events

We now turn to company-level analysis. We start with simple cross-section models regressing the probability of post-founding performance events on founder characteristics. Many relevant outcomes happen either rarely (receiving external finance) or once (exit). We therefore set up outcomes as the post-founding probability of the event occurring; we can think of these regressions as giving the extensive margin of performance. For company f in 3-digit industry j, location a, founded in year t, we estimate:

$$Pr(Y)_{ifjta} = a + \mathbf{TYPE}b1_f + \mathbf{Z}b2_i f + \mathbf{X}'c_f + JT_j t + A_a + u_{ifjta} \tag{3}$$

where Y is one of a set of dummies which take the value 1 if a firm has, at any point after founding: any employment, any revenue, has received any external finance or has exited. **TYPE** is a set of dummies capturing founder demography: we include dummies for solo migrant founders, mixed migrant-native teams and all-migrant teams. Coefficients of these parameters pick up the relative 'effect' of these founder structures, relative to all UK-led firms.

Earlier comparisons of the entrepreneurial skills of migrant and UK-born founders suggest that unobservables are important influences on differences between the two groups (Section 5). Equation 3 does not, therefore, identify causal links because it does not fully control for individual unobservables influencing entrepreneurial entry, or driving the decision to co-found. We mitigate these challenges by including rich founder-level controls in **Z**, as well as 3-digit industry-by-founding year dummies to pick up sectoral trends, and TTWA dummies based on firms' trading location to cover milieu effects. Note also that our regressions pool solo-founded and co-founded firms. To force a like-for-like comparison, **X'** includes a dummy for co-founded firms, and the count of the number of founders, alongside other company-level controls as before.

Table 7 shows the results of linear probability models for equation (3). Solo migrant founders are 4.4 percent more likely to access external finance than solo-UK founders, significant at $p < 0.001$. All-migrant and mixed migrant-native founding teams demonstrate an advantage in accessing external finance relative to UK-only teams. Relative to all-UK founded-firms, mixed teams are 3.7 percent more likely to receive external funding, compared to 7.2 percent for migrant-only teams, with results significant at $p < 0.05$ and $p < 0.001$ respectively. Direct comparison shows a 5% significant difference between these coefficients. We find no significant differences in any other outcomes, although all-migrant teams are slightly less likely than UK-founded teams to post any headcount.

Table 7: Founder demography and Firm Outcomes

	Post-founding outcome			
	(1)	(2)	(3)	(4)
	Employ	Revenue	Ext. Fin	Exit
Solo migrant	-0.014	0.008	0.044***	-0.007
	(0.011)	(0.010)	(0.014)	(0.007)
All migrant team	-0.032*	-0.009	0.072***	0.001
	(0.017)	(0.015)	(0.018)	(0.012)
Mixed team	-0.011	0.014	0.037**	0.008
	(0.011)	(0.013)	(0.015)	(0.008)
Mixed - Migrant	0.021	0.022	-0.035	0.006
Std. Error	0.018	0.016	0.017	0.012
p-value	0.248	0.160	0.047	0.599
Mean Dep. Var.	0.853	0.159	0.730	0.055
Observations	9,940	9,940	9,940	9,940
Adj. R^2	0.148	0.170	0.172	0.047

Source: Beauhurst, Diffbot. Linear probability models comparing future outcomes for migrant-founded companies and mixed / all-migrant founding teams, versus UK-founded companies. Sample is companies founded 2000-2020. Mixed - Migrant shows the difference in coefficients (H0: $b_{mixed} = b_{all-migrant}$). Controls: founder gender, founder age, education, serial founder, founder years of experience, founder managerial / tech / stem background, co-founded firm, team size, multisite firm. Fixed effects: 3-digit SIC Industry × Incorporation Year, Trading TTWA. Standard errors clustered at industry-year level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.2 Firm growth

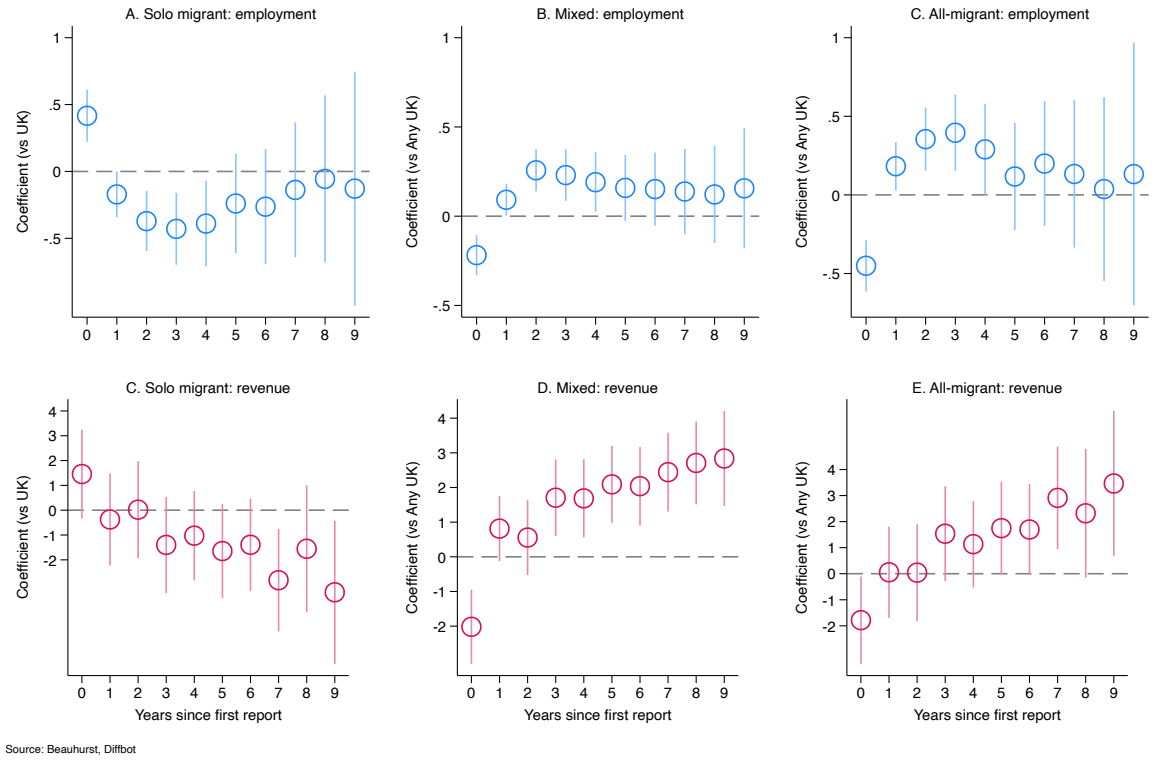
To more directly explore the association between founding team composition and firm growth, we shift to panel data, focusing on the available time-varying outcomes: employment and revenue.

Our specification for company f in 3-digit industry j, located in travel-to-work area a, incorporated in year t_0 and observed in calendar year t, is:

$$\ln Y_{fjat} = a + \mathbf{TYPE}_{fjat(0)}\, b1 + (t * \mathbf{TYPE})_{fjat(s)}\, b2 + \mathbf{X}' c_{fjat0} + JT_{jt} + A_a + u_{fjat} \quad (4)$$

where Y is either log employment or revenue observed for firm f in calendar year t. **TYPE** is the same vector of founder-demography dummies as in equation (3): solo migrant founders, mixed migrant-native teams and all-migrant teams, with all-UK founders and teams the omitted category. Firm age is $s = t - r_f$, measured as years since the firm's first reported outcome r_f. For each category in **TYPE** this gives us two coefficients of interest: $\hat{b1}$s which estimate base year 'effects' of founder types in *t0*, and $\hat{b2}$s which

Figure 9: Firm Outcomes: Intensive margin



Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-202s0. Figure shows coefficients and 95% confidence intervals from Equation 4. The dependent variable is log employment (top row) or log revenue (bottom row). $\hat{b1}$s, levels effects, are denoted in year 0. $\hat{b2}$s, growth effects, are shown from years 1 - 9. Year 10 is excluded due to small cell sizes. All regressions include controls for share female, mean age, any Oxbridge, any managerial experience, any tech background, any STEM, any serial founder, and maximum prior founding experience, number of founders and a co-founded indicator, 3-digit industry-by-year and TTWA fixed effects respectively. Standard errors are clustered at the firm level.

estimate growth 'effects' for subsequent years from s. $\mathbf{X}_f$ collects founder- and team-level controls (share female, mean age, any Oxbridge, any managerial experience, any tech background, any STEM, any serial founder, and maximum prior founding experience). As before we also include controls for number of founders and a co-founded indicator to force clean comparisons between solo-founded and co-founded firms. JT_{jt} and A_a are 3-digit industry-by-year and TTWA fixed effects respectively. Standard errors are clustered at the firm level. Per Table 5, more firms achieve employment than revenue. Our baseline employment sample includes 34,214 observations for around 10k firms and our baseline revenue sample 8,162 observations for around 2.1k firms.

Figure 9 reports $\hat{b1}$s and $\hat{b2}$s for all-migrant teams, mixed teams and solo migrant founders. (Appendix Table A11 gives detailed results.) We find that all-migrant and mixed teams initially report lower employment and revenue than UK-born teams. Solo migrant founders post higher numbers, although this is only significant for employment. Growth trajectories then move in opposite directions, with solo migrant founders growing

less rapidly than UK-born solo founders, although in most years differences are not significant. Conversely, both mixed and all-migrant teams have faster employment growth than all-UK teams in the first four years, after which the premium becomes non-significant. For revenue, mixed teams post an increasing growth premium from year 3; all-migrant teams post slower growth premia which become significant later.

6.3 Robustness

We run our extensive and intensive margin tests through a series of robustness checks.

Results for post-founding events survive a range of sensitivity tests.[22] We then re-run the analysis restricting the sample to co-founded firms. The overall pattern of results is unchanged (Appendix Table A9). We then run leave-one-out tests for the c25% of companies in ICT activities to explore their influence. Figure A4 shows results for the main sample and excluding these companies. Overall results are almost identical, suggesting little influence of this block of companies. In Section 8 we explore heterogeneity across ICT, high-tech and science and technology activities more broadly.

As a last check, we re-run the extensive margin results excluding companies selected on actual high-growth activity. As explained in Section 3.5 around 90 percent of our company sample is selected on high-growth potential: Beauhurst starts tracking these firms at early or mid points in their lifecycle (Figure 1). However, 9.3 percent of firms are selected on the basis of employment or revenue growth (Table A4). This raises the possibility that for our performance analysis, a minority of our sample is selected on the dependent variable. We therefore test whether our results are robust to removing these companies from the data. The new sample is 11,037 companies compared to 12,182 in the main analysis. Appendix Table A10 reproduces the extensive margin analysis. Results are almost identical.

Firm growth results are run through two substantive tests. We first explore the influence of ICT activities, as above. Figure A5 gives results: estimates without ICT firms are essentially identical to the main results. We then re-run our results excluding companies selected on employment and revenue growth, as above. Appendix Figure A6 reproduces the intensive margin analysis, for the subsample of firms reporting employment or revenue. Employment tests run on 9,284 companies, around 11% fewer than the main regressions; growth premia for mixed teams and all-migrant teams are almost identical to before. Revenue tests run on 1,263 companies, around 45% fewer. Mixed team growth premia are similar but much less precisely estimated. All-migrant team premia coefficients are much less stable and almost all insignificant.

[22]Specifically: adding controls for firm shares of non-UK nationals and dual nationals; removing potential bad controls (degree share, Oxbridge and Russell Group shares); and fitting alternate combinations of 3-digit industry x TTWA fixed effects x year fixed effects.

7 Company distinctiveness and founder characteristics

Over and above standard metrics, migrant-founded firms may differ on other dimensions of firm performance. Per our Section 2 framework, we focus on founder strategy (Botelho et al., 2026; Guzman and Li, 2022; Gans et al., 2019; Rajan, 2012), in particular the way founders position firms and their products and services. In validation tests on our text-based measures, more novel firms are more likely to attract external finance, and conditional on having some revenue or headcount, grow faster on both dimensions than less novel firms (Section 3.4).

How might founder demography correlate with firm distinctiveness? Per Section 1, teams with complementary skills and backgrounds may be more likely to develop distinctive propositions. If migrant and UK-born founders bring different qualities, experiences and knowledge to a team, migrant-native synergies may increase firm distinctiveness. Against this, more homogeneous teams (all-migrant and/or all-UK) may work effectively together, with fewer frictions.

Here we run OLS/LPM estimations regressing current company distinctiveness measures on founding team characteristics. Each regression compares outcomes for migrant-only and mixed founding teams relative to all UK-born founding teams. For company f in 3-digit industry j, location a, founded in year t, we estimate:

$$D_{ifjta} = a + \mathbf{FOUNDERS}b1_f + \mathbf{Z}b2_i f + \mathbf{X}'c_f + JT_j t + A_a + u_{ifjta} \tag{5}$$

Where D is one our five dummy/continuous distinctiveness measures, and other parameters are defined as in Equation 3.

Table 8 gives our main results. Solo migrant founders are not distinctive on any measure compared to solo UK-born founders. In contrast, mixed migrant-native founding teams have over 157 fewer peers across all industries and years than all-UK teams, a difference of around 7.8 percent on the average firm (Panel A, Col 1). This result holds when we compare only with earlier-founded firms (Col 2) but not with firms in the same founding-year cohort (Col 3). All-migrant teams, in contrast, have significantly fewer within-industry peers than all-UK teams (Col 4) - a narrower measure of distinctiveness and one that does not hold for earlier or same-cohort comparisons (Panel B, Cols 1-2). Non-threshold-based measures - originality and trendiness - show no differences between any migrant and non-migrant founded firms (Panel B, Cols 3-4). In general, we find no difference in outcomes for all-migrant versus mixed teams.

We run four robustness checks. First, Table A12 gives results for co-founded firms, which are very similar. Second, Figure A7 shows results from the ICT leave-out test.

Table 8: Founder demography and Firm Distinctiveness

	(1) #Peers	(2) # Prior	(3) # Cohort	(4) # SIC3 peers
Panel A				
Solo migrant	-73.049	-24.534	0.414	-7.300
	(86.047)	(67.398)	(5.904)	(11.158)
Migrant-only team	-178.450	-76.262	-4.129	-26.087**
	(143.551)	(133.987)	(9.676)	(12.313)
Mixed team	-157.508**	-135.875**	-5.666	-8.970
	(77.312)	(57.802)	(4.883)	(8.628)
Mixed - Migrant	-20.942	59.613	1.536	-17.117
Std. Error	146.721	129.425	10.167	13.099
p-value	0.887	0.645	0.880	0.192
Mean Dep. Var.	2010.135	1365.979	119.633	111.811
Observations	9,940	9,940	9,940	9,940
Adj. R^2	0.112	0.110	0.127	0.163
	# SIC3 prior	# SIC3 cohort	Originality	Trendiness
Panel B				
Solo migrant	-3.223	-0.136	-0.002	0.001
	(7.817)	(1.134)	(0.001)	(0.002)
Migrant-only team	-16.970	-1.754*	0.001	-0.005
	(11.370)	(1.039)	(0.002)	(0.003)
Mixed team	-9.304	-0.499	0.001	-0.002
	(5.754)	(0.664)	(0.001)	(0.002)
Mixed - Migrant	-7.666	-1.255	0.000	-0.003
Std. Error	9.914	1.180	0.002	0.003
p-value	0.440	0.288	0.988	0.364
Mean Dep. Var.	67.729	8.264	0.098	0.657
Observations	9,936	9,940	9,940	9,940
Adj. R^2	0.121	0.191	0.063	0.213

Source: Beauhurst, Diffbot. Linear probability models comparing future outcomes for migrant-founded companies and mixed / all-migrant founding teams, versus UK-founded companies. Sample is companies founded 2000-2020. Mixed - Migrant shows the difference in coefficients (H0: $b_{mixed} = b_{all-migrant}$). Controls: founder gender, ounder age, education, serial founder, founder years of experience, founder managerial / tech / stem background, co-founded firm, team size, multisite firm. Fixed effects: 3-digit SIC Industry × Incorporation Year, Trading TTWA. Standard errors clustered at industry-year level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Again, results are almost identical, suggesting little influence of these firms. Third, Table A13 shows results using alternative metrics: peer counts built using less precise TF-IDF metrics, and peer counts using alternative, less conservative thresholds for peer definition. TF-IDF metrics give similar results but with smaller migrant-UK differences, consistent with the less granular natural language processing. Using lower thresholds to define ij peers generates more peers per firm but will erode precision: results bear this out, with increasingly large but imprecisely-estimated results. Finally, Table A14 gives results for younger firms.[23] Per Guzman and Li (2022), the intuition is to pick up firm positioning closer to the point of founding, rather than in some post-founding period as in our main analysis. Reassuringly, these results have similar signs and magnitudes as our main results, but are less precisely estimated given the much smaller N.

8 Extensions: heterogeneity by sector, country, language, location

In this section we extend our main results. We first focus on sectoral differences, specifically whether migrant-performance links differ in 'knowledge-intensive' activities. Next we look at team channels, specifically the role of modelled birth country homogeneity (for all-migrant founding teams) and shared language (for all-migrant and mixed teams). These results give us insight into the relative importance of homophily and diversity on a range of company outcomes. Then, we look at how results vary across different parts of the UK urban system, comparing findings from pooled data to those for firms in London, other major conurbations, and smaller cities and towns.

8.1 Sector differences

We first explore sectoral differences in migrant-performance relationships. Existing literature suggests that diversity-performance linkages amongst workers are strongest in 'knowledge-intensive' settings such as science and technology sectors (Ozgen, 2021; Berliant and Fujita, 2012). Innovation-driven entrepreneurship is also dominated by STEM activity (Botelho et al., 2026). Conversely, company formation and entrepreneurial discovery are inherently complex (Gans et al., 2019); sectoral settings may be less important than the core processes of starting and building a firm.

We use three definitions of 'knowledge-intensive' activity. The narrowest is SIC 620, Computer programming, consultancy and related activities. Bakhshi et al. (2015) provide a definition of 'high-tech' activities across manufacturing (e.g. and services (e.g.

[23] Cross-checks using firms founded 2017-2020 give almost identical results. For firms founded 2019-2020 and 2020 only, sample sizes are small (704 and 138 observations respectively) and all coefficients are insignificant.

chemicals, pharma, telecoms, ICT, computer games). Harris (2015) provide a a broader definition of 'science and technology' sectors, covering ICT, life sciences, high-tech manufacturing, high-tech services (such as architecture) and publishing/broadcasting (such as advertising and computer games).

Overall, we find weak evidence that migrant founders in knowledge-intensive sectors drive our main results, with most results very similar to the pooled sample.[24] Extensive margin regressions are shown in Figure A8. Coefplots show the 'effect' of solo migrant, all-migrant and mixed teams vs UK founders, across the pooled sample, ICT, high-tech and sci-tech sub-samples. Mixed teams in ONS sci-tech activities are more likely to get external funding than the average mixed team, relative to UK founding teams; other subsample results are very similar to the pooled sample. Intensive margin results are shown in Figure A9. We find no differences in employment or revenue premia in knowledge-intensive subsectors. Distinctiveness results are shown in Figure A10. All-migrant teams in sci-tech sectors have fewer all-year peers and are more distinctive in their cohort than other all-migrant teams, relative to all-UK teams. Mixed teams in high-tech activity also have fewer peers than other mixed teams.

8.2 Common birth countries

We next vary all-migrant groups by modelled country of origin (proxied by lowest observed conutry of education, per Section 3). Specifically, in Equations 3 and 4 we vary **TYPE** to include separate dummies for all-migrant groups from the same sending country, or different sending countries. Overall, we find that all-migrant groups from the same sending country are more likely to receive external finance, and post faster revenue growth (conditional on having revenue). However, all-migrant groups from different countries drive our distinctiveness result, with substantially fewer peers.

For the extensive margin analysis we focus on our main result: all-migrant and mixed teams are more likely to get external finance than all-UK counterparts, with funding premia of 7.2% and 3.7% respectively. Appendix Figure A11 gives results for the sending-country decomposition. All-migrant founders from the same country are even more likely to get external finance, with no significant link for all-migrant founders from different countries.

For the intensive margin analysis, Appendix Figure A12 splits all-migrant founding teams into those from the same birth country vs. different birth countries. All-migrant teams from different sending countries post higher, significant early employment growth premia compared to same-country teams. By contrast, same-country teams post rapid and robust revenue growth premia compared with different-country teams.

[24] We also run regressions breaking out the individual ONS science and technology categories. Sci-tech differences are driven by companies in the two largest groups, digital technology and life sciences/health.

For the distinctiveness analysis, Panel A of Figure A13 subsets all-migrant founding teams into those with different birth countries vs. the same. All-migrant teams from different sending countries have fewer peers, though the effect is marginally significant. By contrast, we find no significant link for all-migrant firms from the same sending country.

8.3 Shared language

Next, we explore the role of cultural proximity using first language as a proxy (Ferrucci and Lissoni, 2019; Ozgen, 2021). We now vary **TYPE** to include dummies for all-migrant and mixed teams where founders share the same language, or have different languages.[25] Overall, results suggest that linguistic proximity operates differently from birth country, depending on the outcome. Teams sharing the same language are more likely to obtain external finance. For all-migrant teams, shared language also explains employment and revenue growth. For mixed teams, shared language drives up distinctiveness, but revenue growth is faster when teams have different first languages.

In Appendix Figure A14, we show that the probability a firm gets external finance is significantly higher and more robust for all-migrant and mixed-teams where founders speak the same language, compared to teams where first languages are different (coefficients are positive but only significant at 10%). We also find that mixed teams where founder languages differ are slightly more likely to have an IPO or be acquired.

Appendix Figure A15 gives results for company scaling. For all-migrant teams, our main employment and revenue results are explained by growth premia for same-language teams. For mixed teams, where founders include migrants and UK-born, results differ by outcome. Both same and different-language teams help explain employment growth premia. For revenues, different-language teams drive the result.

Distinctiveness results are given in Figure A13 Panel B. For all-migrant groups, there is no significant role of linguistic proximity in the number of peers. By contrast, for mixed teams, our main result is explained by teams sharing the same first language: the effect size is larger and more robust (significant at 1%).

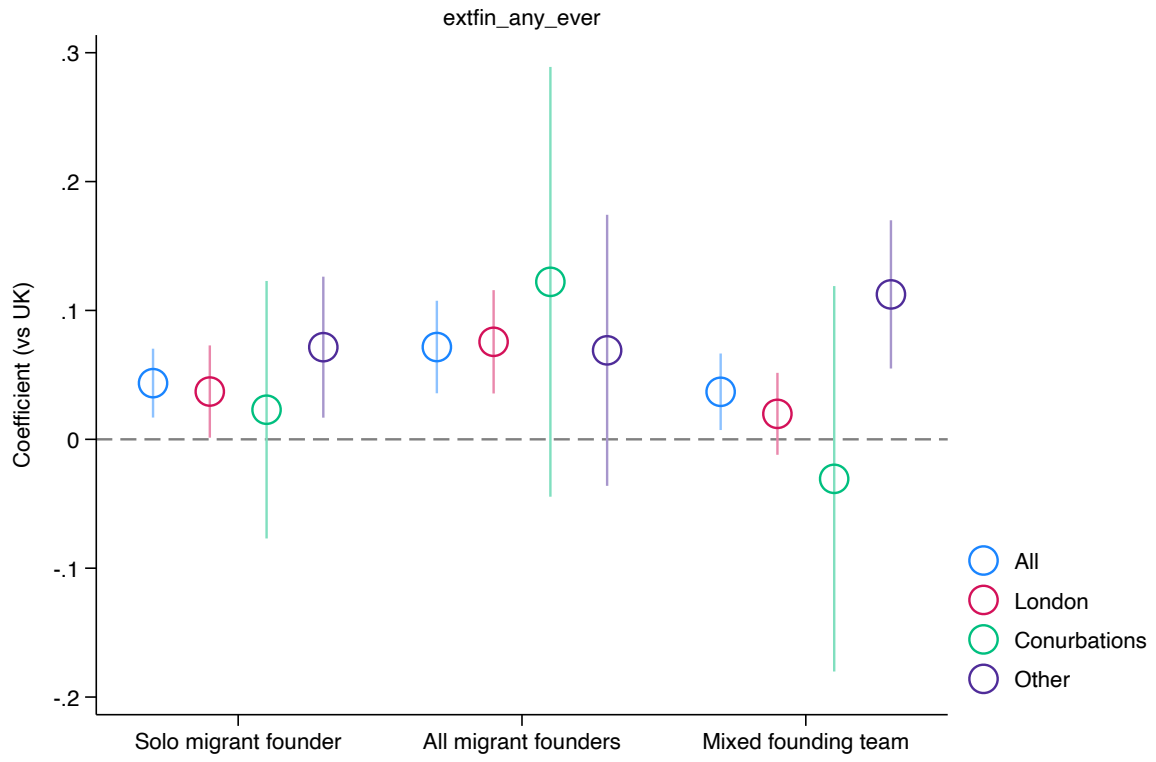
8.4 Results across the urban system

We finally explore differences across locations. In this version of the paper we re-run Equations 3, 4 and 5 subsetting our data into firms based in the Greater London TTWA, other large UK conurbations (Greater Manchester, Birmingham, Liverpool, Leeds, Newcastle, Edinburgh, Glasgow and Aberdeen TTWAs) and all other locations (predominantly smaller cities and towns, see Figure 6). These regressions estimate like-for-like

[25] Sensitivity tests using language family branches - derived from Ethnologue - give identical results.

comparisons of migrant and UK-founded firms in these areas.[26]

Figure 10: Extensive margin of external finance split by area types



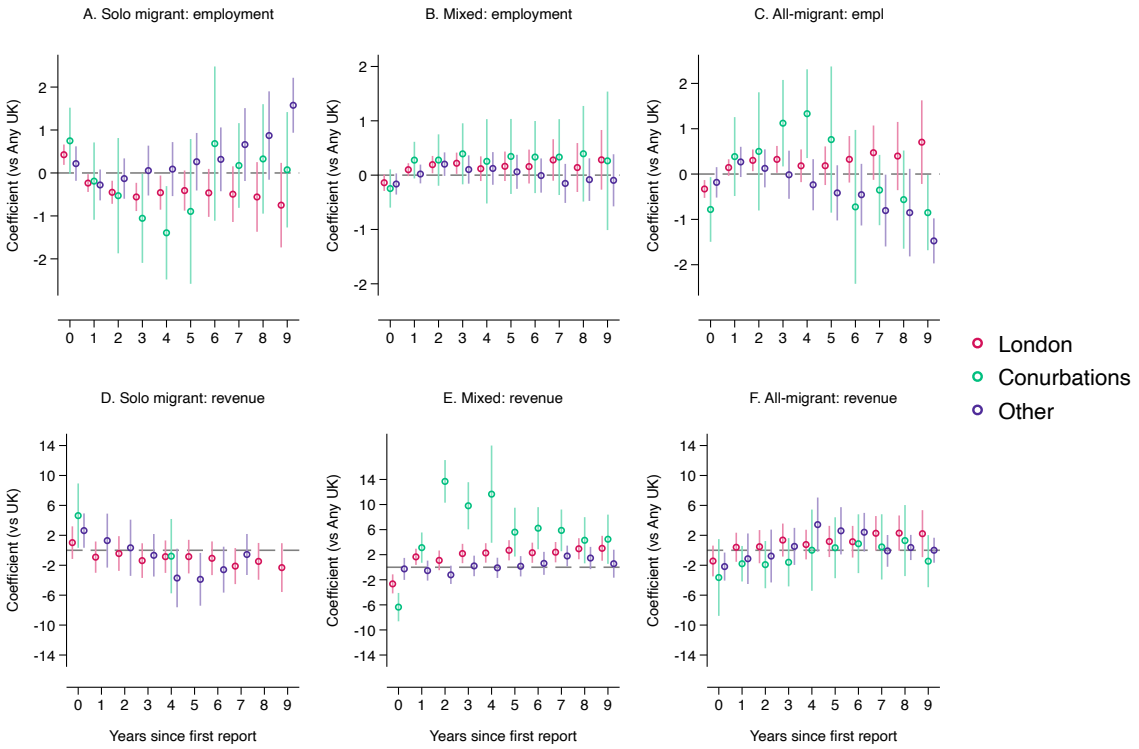
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals of linear probability models where the dependent variable is the probability of any employment post-founding, any revenue, any external finance, IPO or acquisition. Other notes per Table 7.

Results for the probability of obtaining external finance are shown in Figure 10. Solo migrant founders based in London or outside conurbations have the highest probability of getting external finance. For all-migrant firms, main results are explained by firms in London. For mixed teams, our main result is explained by firms outside London and the conurbations, who have the highest probability of receiving external support. Results are robust to varying London boundaries to include contiguous TTWAs (Slough and Heathrow, Guildford and Aldershot) that are part of the larger functional urban area.

Figure 11 re-runs our company scaling analysis for employment and revenue. The results imply that migrant founder correlates of employment and revenue growth vary substantially both across locations, and across founding team types. On employment, solo migrants have an L-shape trajectory in London, and a u-shape trajectory in conurbations and smaller cities. By contrast, mixed teams' premia move evenly across the urban system, with little place variation. All-migrant teams have a much more dynamic n-shape trajectory in conurbations and smaller cities. On revenue, solo migrant firms have

[26]Future versions of the paper will keep the pooled data and add interactions with continuous population / employment density.

Figure 11: Firm outcomes: intensive margin split by area types



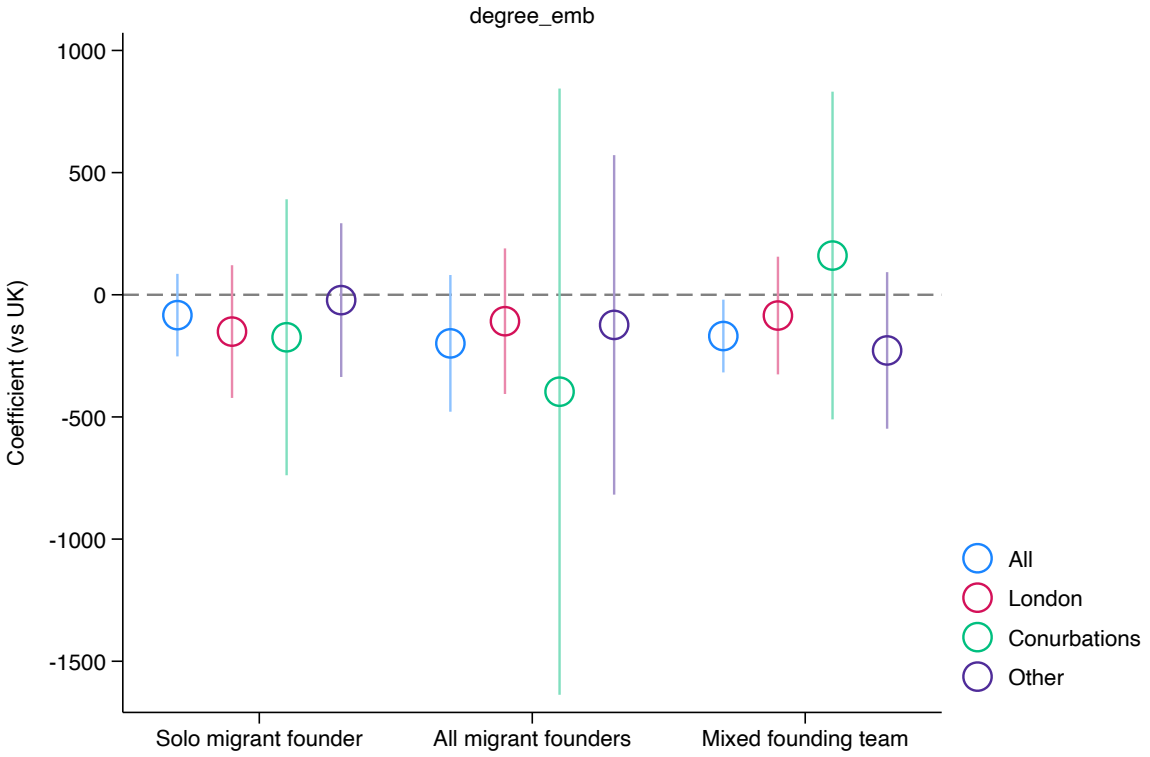
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals from Equation 4. The dependent variable is log employment (top row) or log revenue (bottom row). $\hat{b1}$s, levels effects, are denoted in year 0. $\hat{b2}$s, growth effects, are shown from years 1 - 9. Year 10 is excluded due to small cell sizes. All regressions include controls for share female, mean age, any Oxbridge, any managerial experience, any tech background, any STEM, any serial founder, and maximum prior founding experience, number of founders and a co-founded indicator, 3-digit industry-by-year and TTWA fixed effects respectively. Standard errors are clustered at the firm level.

an L-shape trajectory across urban system, with revenue premia declining over time, and relatively little variation across places. Note that solo migrants in conurbations are a small sample and do not post revenue every year post first-report. For mixed founding teams, revenue premia are higher in London and especially, other conurbations. All-migrant teams show relatively little variation across places. Results are robust to varying London boundaries to include contiguous TTWAs (Slough and Heathrow, Guildford and Aldershot) that are part of the larger functional urban area.

Figure 12 replays our main distinctiveness result, cross-industry and year peer counts, across locations. Area splits suggest that differences are larger in London and larger cities for migrant-founded or mixed teams vs. all-UK teams; but results are not precisely estimated.

Overall, this array of results shows founder and team-level channels vary substantively across more or less urban places. As expected given its position in the UK urban system,

Figure 12: Number of peer firms split by area types



Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals of OLS models where the dependent variable is the number of peers across all industries. Other notes per Table 8.

London appears to offer the most affordances: solo migrants and all-migrant teams are most likely to obtain external finance; there are positive employment growth premia for all-migrant and mixed teams compared to all-UK teams; and a positive revenue premium for mixed teams. London companies with mixed founding teams have fewer peers compared to all-UK teams. In non-London conurbations, all-migrant teams are most likely to add headcount compared to all-UK teams, and mixed teams are most likely to add revenue. Perhaps more surprisingly, in smaller towns and cities mixed teams are most likely to get external finance; and solo migrant founders add employment compared with all-UK teams.

9 Conclusions

In this paper, we use novel founder- and firm-level data on over 12,000 UK companies and 15,000 company founders to establish several new facts about migrant entrepreneurship, high-growth firms and cities. **First,** relative to the UK's migrant population, migrant entrepreneurs are substantially over-represented among founders of high-growth and high-growth-potential firms. **Second,** compared to UK-born founders, migrant founders are

positively selected on human capital, but negatively selected on entrepreneurial experience. We find no evidence of selection on other entrepreneurial skills. **Third,** high-growth potential firms are highly urbanised, and those with a migrant founder or co-founder display a distinctive urban geography. **Fourth,** firms with mixed native-migrant founding teams follow faster revenue growth trajectories than all-UK teams, and initially, also scale faster on headcount. We also find evidence for growth premia in all-migrant teams, although this is more sensitive to sample construction. **Fifth,** as compared with firms founded by natives, firms with mixed native-migrant founding teams have more distinctive strategic positioning, with fewer peers. **Sixth**, shared birth country, shared first language and urban location help explain our results. Migrant-founded firms in London and conurbations exhibit the largest performance premia relative to all-UK counterparts. In contrast, 'knowledge intensive' firms show little difference to the average firm.

References

Acosta, C. and Marinoni, A. (2025). Uncertainty and immigrant entrepreneurship: Evidence from brexit. *Organization Science*, 36(3):1072–1087.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2):323–351.

Akcigit, U., Alp, H., Pearce, J., and Prato, M. (2025). Transformative and subsistence entrepreneurs: Origins and impacts on economic growth. Report, NBER.

Aldrich, H. E. and Waldinger, R. (1990). Ethnicity and entrepreneurship. *Annual Review of Sociology*, 16:111–135.

Andersson, M. and Larsson, J. P. (2014). Local entrepreneurship clusters in cities. *Journal of Economic Geography*, 16(1):39–66.

Azoulay, P., Jones, B. F., Kim, J. D., and Miranda, J. (2022). Immigration and entrepreneurship in the united states. *American Economic Review: Insights*, 4(1):71–88.

Bakhshi, H., Davies, J., Freeman, A., and Higgs, P. (2015). The geography of the uk's creative and high-tech economies. Technical report, Nesta.

Balland, P.-A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P. A., Rigby, D. L., and Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. *Nature Human Behaviour*, 4(3):248–254.

Beine, M., Peri, G., and Raux, M. (2024). The contribution of foreign master's students to us start-ups. Report, NBER.

Beine, M., Peri, G., Raux, M., and Timricht, L. (2026). The contribution of international master students to entrepreneurship in the world. CESifo Working Paper 13014, CESifo, Munich.

Berliant, M. and Fujita, M. (2012). Culture and diversity in knowledge creation. *Regional Science and Urban Economics*, 42(4):648–662.

Bernstein, S., Mehta, K., Townsend, R. R., and Xu, T. (2022). Do startups benefit from their investors' reputation? evidence from a randomized field experiment. Report, NBER.

Botelho, T. L., Fehder, D. C., and Hochberg, Y. V. (2026). Innovation-driven entrepreneurship. *Journal of Economic Literature*, 64(1):89–140.

Bryan, K. and Guzman, J. (2021). Entrepreneurial migration. Report.

Burton, M., Sørensen, J. B., and Beckman, C. M. (2002). *7. Coming from good stock: Career histories and new venture formation*, volume 19, page 0. Emerald Group Publishing Limited.

Calder-Wang, S., Gompers, P. A., and Huang, K. (2021). Diversity and performance in entrepreneurial teams. *National Bureau of Economic Research Working Paper Series*, No. 28684.

Calvino, F., Criscuolo, C., and Menon, C. (2018). A cross-country analysis of start-up employment dynamics. *Industrial and Corporate Change*, 27(4):677–698.

Certo, S. T., Lester, R. H., Dalton, C. M., and Dalton, D. R. (2006). Top management teams, strategy and financial performance: A meta-analytic examination. *Journal of Management Studies*, 43(4):813–839.

Chatterji, A., Glaeser, E., and Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innovation Policy and the Economy*, 14(1):129–166.

Chodavadia, S. A., Pekkala Kerr, S., Kerr, W. R., and Maiden, L. J. (2024). Immigrant entrepreneurship: New estimates and a research agenda. *National Bureau of Economic Research Working Paper Series*, No. 32400.

D'Acunto, F., Tate, G. A., and Yang, L. A. (2020). Entrepreneurial teams: Diversity of skills and early-stage growth. Report.

Dahlke, J., Schmidt, S., Lenz, D., Kinne, J., Dehghan, R., Abbasiharofteh, M., Schütz, M., Kriesch, L., Hottenrott, H., Kanilmaz, U. N., Grashof, N., Hajikhani, A., Liu, L., Riccaboni, M., Balland, P.-A., Wörter, M., and Rammer, C. (2025). The we-bai paradigm of innovation research: Extracting insight from organizational web data through ai. Report, ZEW.

Davis, D. R. and Dingel, J. I. (2019). A spatial knowledge economy. *American Economic Review*, 109(1):153–70.

Department for Business, Enterprise, Innovation and Skills (2020). Business population estimates for the UK and regions 2020: Statistical release. Technical report, BEIS, London.

Department for Business, Innovation and Skills (2010). Business population estimates for the UK and regions 2010: Statistical release. Technical report, BIS, London.

Department for Business, Innovation and Skills (2015). Business population estimates for the UK and regions 2015: Statistical release. Technical report, BIS, London.

Draca, M., Nathan, M., Nguyen, V., Oliviera-Cunha, J., Rosso, A., and Valero, A. (2023). Economic growth goes fractal.

Duranton, G. and Puga, D. (2001). Nursery cities: Urban diversity, process innovation and the life cycle of products. *American Economic Review*, 91(5):1454–1477.

Duranton, G. and Puga, D. (2023). Urban growth and its aggregate implications. *Econometrica*, 91(6):2219–2259.

Evans, D. S. and Jovanovic, B. (1989). An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy*, 97(4):808–827.

Ewens, M. (2022). Race and gender in entrepreneurial finance. Report, NBER.

Fairlie, R., Robb, A., and Robinson, D. T. (2022). Black and white: Access to capital among minority-owned start-ups. *Management Science*, 68(4):2377–2400.

Ferrucci, E. and Lissoni, F. (2019). Foreign inventors in europe and the united states: Diversity and patent quality. *Research Policy*, 48(9):103774.

Gans, J. S., Stern, S., and Wu, J. (2019). Foundations of entrepreneurial strategy. *Strategic Management Journal*, 40(5):736–756.

Glaeser, E. L., Kerr, S. P., and Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics*, 97(2):498–520.

Glaeser, E. L., Kerr, W., and Ponzetto, G. A. M. (2010a). Clusters of entrepreneurship. *Journal of Urban Economics*, 67(1):150–168.

Glaeser, E. L., Kerr, W. R., and Ponzetto, G. A. M. (2010b). Clusters of entrepreneurship. *Journal of Urban Economics*, 67(1):150–168.

Gompers, P. A. and Woo, H. (2026). Global entrepreneurial talent flows: Widely funded, selectively successful? Working Paper 35943, National Bureau of Economic Research, Cambridge, MA.

Gray, S., Kemeny, T., Nathan, M., Ozgen, C., Pialli, G., Rosso, A. C., Reades, J., Seré, M., Siddiqui, S., and Valero, A. (2026). Graph-ee: Building employer - employee panels from a knowledge graph and company microdata. *CEP Discussion Paper CEPDP2156*.

Guiso, L., Pistaferri, L., and Schivardi, F. (2015). Learning entrepreneurship from other entrepreneurs? *National Bureau of Economic Research Working Paper Series*, No. 21775.

Guiso, L. and Schivardi, F. (2011). What determines entrepreneurial clusters? *Journal of the European Economic Association*, 9(1):61–86.

Guzman, J. and Li, A. (2022). Measuring founding strategy. *Management Science*, 0(0):null.

Guzman, J. and Stern, S. (2015). Where is Silicon Valley? *Science*, 347(6222):606–609.

Haltiwanger, J., Jarmin, Ron, S., and Miranda, J. (2013). Who creates jobs? small versus large versus young. *The Review of Economics and Statistics*, 95(2):347–361.

Harris, J. (2015). Identifying science and technology businesses in official statistics. Technical report, Office for National Statistics, London.

Hegde, D. and Tumlinson, J. (2021). Information frictions and entrepreneurship. *Strategic Management Journal*, 42(3):491–528.

HESA (2024). Higher education student statistics: Uk, 2022/23 - where students come from and go to study. Report, HESA.

Hoberg, G. and Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465.

Hong, L. and Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, 101(46):16385–16389.

Honoré, F. (2022). Joining forces: How can founding members' prior experience variety and shared experience increase startup survival? *Academy of Management Journal*, 65(1):248–272.

Howell, T. and Hall, T. A. (2026). Lone genius or lonely fool? Exploring the viability of solo-founding in entrepreneurship. *Strategic Management Journal*, 47(10):2646–2678.

Hunt, J. and Gauthier-Loiselle, M. (2010). How much does immigration boost innovation? *American Economic Journal: Macroeconomics*, 2(2):31–56.

Hyde, F., Katsande, A., Robinson, D., Valencia, B., Whorwood, H., and Yip, E. (2023). Gender and diversity in uk tech: women founders and other underrepresented entrepreneurs. Report, Beauhurst / AWS.

Ives, E., Kocer, D., and Salter, P. (2023). Job creators 2023. Report, The Entrepreneurs Network.

Jin, Z., Kermani, A., and McQuade, T. (2025). Native-immigrant entrepreneurial synergies. NBER Working Paper 33804.

Jones, T., Ram, M., and Villares-Varela, M. (2018). Diversity, economic development and new migrant entrepreneurs. *Urban Studies*, 56(5):960–976.

Kemeny, T. and Storper, M. (2020). Superstar cities and left-behind places: disruptive innovation, labor demand, and interregional inequality. LSE International Inequalities Institute Working Paper 41.

Kenney, M. and Patton, D. (2015). Gender, ethnicity and entrepreneurship in initial public offerings: illustrations from an open database*. *Research Policy*, 44(9):1773–1784.

Kerr, William, R. and Mandorff, M. (2023). Social networks, ethnicity, and entrepreneurship. *Journal of Human Resources*, 58(1):183.

Kerr, W. (2010). Breakthrough inventions and migrating clusters of innovation. *Journal of Urban Economics*, 67(1):46–60.

Kerr, W. R., Nanda, R., and Rhodes-Kropf, M. (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives*, 28(3):25–48.

Kerr, W. R. and Robert-Nicoud, F. (2020). Tech clusters. *Journal of Economic Perspectives*, 34(3):50–76.

Kleinhempel, J., Klasing, M. J., and Beugelsdijk, S. (2023). Cultural roots of entrepreneurship: Evidence from second-generation immigrants. *Organization Science*, 34(5):1800–1819.

Klette, T. and Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):986–1018.

Kloosterman, R. and Rath, J. (2003). *Immigrant Entrepreneurs: Venturing Abroad in the Age of Globalisation*. Berg, Oxford.

Knight, F. (1921). *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston.

Koster, H. R. A. and Ozgen, C. (2021). Cities and tasks. *Journal of Urban Economics*, 126:103386.

Lazear, E. P. (2004). Balanced skills and entrepreneurship. *American Economic Review*, 94(2):208–211.

Lee, K. M., Kim, M. J., Brown, J. D., Earle, J. S., and Liu, Z. (2025). Are immigrants more innovative? evidence from entrepreneurs. *Journal of Economics & Management Strategy*, n/a(n/a).

Lee, S. and Glennon, B. (2023). The effect of immigration policy on founding location choice: Evidence from Canada's Start-up Visa program. NBER Working Paper 31634.

Lentz, R. and Mortensen, D. T. (2008). An empirical model of growth through product innovation. *Econometrica*, 76(6):1317–1373.

Lindquist, M. J., Sol, J., and Van Praag, M. (2015). Why do entrepreneurial parents have entrepreneurial children? *Journal of Labor Economics*, 33(2):269–296.

Lucas, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, 9(2):508–523.

Luttmer, E. G. J. (2011). On the mechanics of firm growth. *The Review of Economic Studies*, 78(3):1042–1068.

Mack, D. Z., Chen, G., Hsu, P.-H., Lee, Y. T., and George, G. (2025). Interfaces, social information processing, and diversity cascades: How board diversity influences invention output. *Research Policy*, 54(1):105148.

Manning, A. (2025). *Why Immigration Policy is Hard: And How To Make It Better*. Polity, Cambridge.

Mesquita, F., Cannaviccio, M., Schmidek, J., Mirza, P., and Barbosa, D. (2019). Knowledgenet: A benchmark dataset for knowledge base population. Report, Association for Computational Linguistics.

Migration-Observatory (2023). Citizenship and naturalisation for migrants in the uk. Report, Migration Observatory.

Office for National Statistics (2025). High growth businesses. Technical report, ONS, Newport.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Ozgen, C. (2021). The economics of diversity: Innovation, productivity and the labour market. *Journal of Economic Surveys*, 35(4):1168–1216.

Rajan, R. G. (2012). Presidential address: The corporation in finance. *The Journal of Finance*, 67(4):1173–1217.

Rosen, S. (1982). Authority, control, and the distribution of earnings. *The Bell Journal of Economics*, 13(2):311–323.

Saxenian, A.-L. (2006). *The New Argonauts: Regional Advantage in a Global Economy*. Harvard University Press, Cambridge, MA.

Schumpeter, J. (1962). *The Theory of Economic Development*. Springer, Berlin.

Tareque, I. S., Guzman, J., and Wang, D. (2024). High-skilled immigration enhances regional entrepreneurship. *Proceedings of the National Academy of Sciences*, 121(37):e2402001121.

Triana, M. d. C., Richard, O. C., and Su, W. (2019). Gender diversity in senior management, strategic change, and firm performance: Examining the mediating nature of strategic change in high tech firms. *Research Policy*, 48(7):1681–1693.

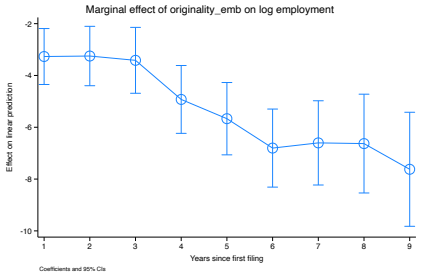
A Additional figures and tables

Figure A1: Example company description text provided by Beauhurst

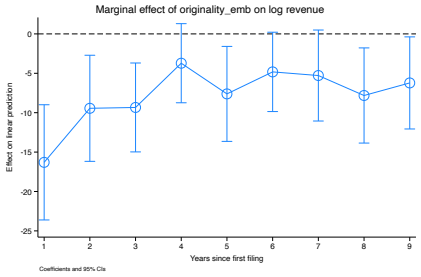
Company	Company description text	Beauhurst analyst text
Deliveroo	Deliveroo is on a mission to transform the way the world thinks about food delivery. It's not a chicken chow mein and a night on the sofa anymore, it's your favourite local restaurant, it's a dinner party, a date. We are five years in, and along the way our team have taken hundreds of ideas from brainstorms to global roll-outs, like Deliveroo Editions – bespoke kitchens designed to host a locally-curated selection of restaurants. Editions are our solution to ensuring that our customers have access to the best of the food-scene, no matter where they live. And that's just what we're like at Deliveroo, no compromise allowed and lots of food-inspired challenges to get your teeth into. Out-of-the-box thinking is actively encouraged and we move quickly to make great ideas happen. We're energetic, fast-paced and blow off steam with free-for-all Friday lunches. It's a formula that's working too – we're bringing great food to customers in 12 countries and over 800 cities.	Deliveroo provides delivery services for restaurants, using technology to predict the time taken to prepare meals and efficient ways of delivering orders using the location of restaurants, customers and riders.

Source: Beauhurst. *Note:* Example text for Deliveroo, a leading UK food delivery service. The figure shows the company's own description of its positioning - strategy, products and services - and Beauhurst's own summary. In our analysis we use companies' own words rather than Beauhurst's, because we are interested in how companies seek to position and differentiate themselves versus competitors.

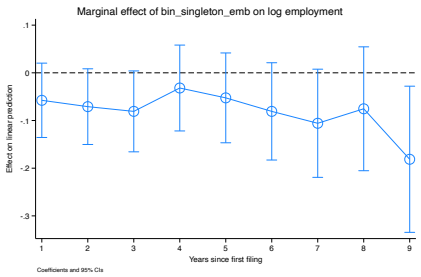
Figure A2: Validation: distinctiveness and employment / revenue dynamics



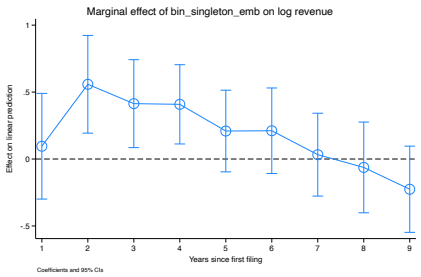
(a)



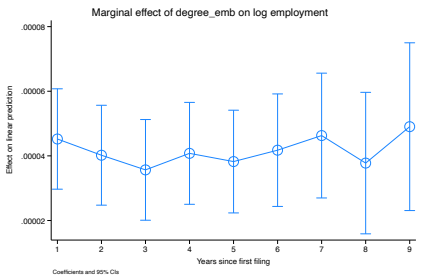
(b)



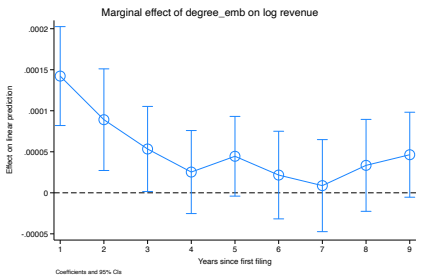
(c)



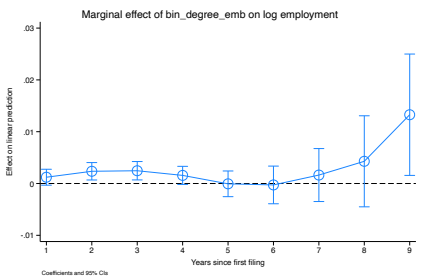
(d)



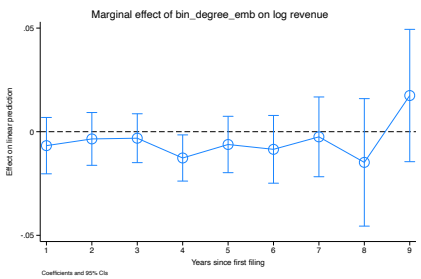
(e)



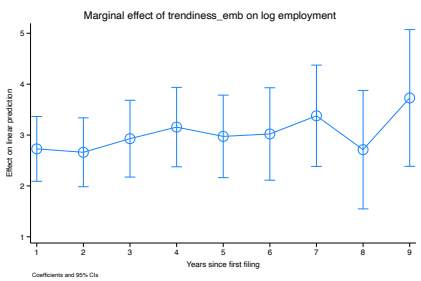
(f)



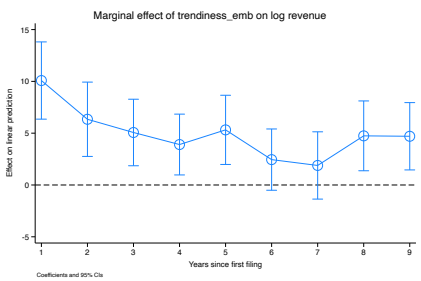
(g)



(h)



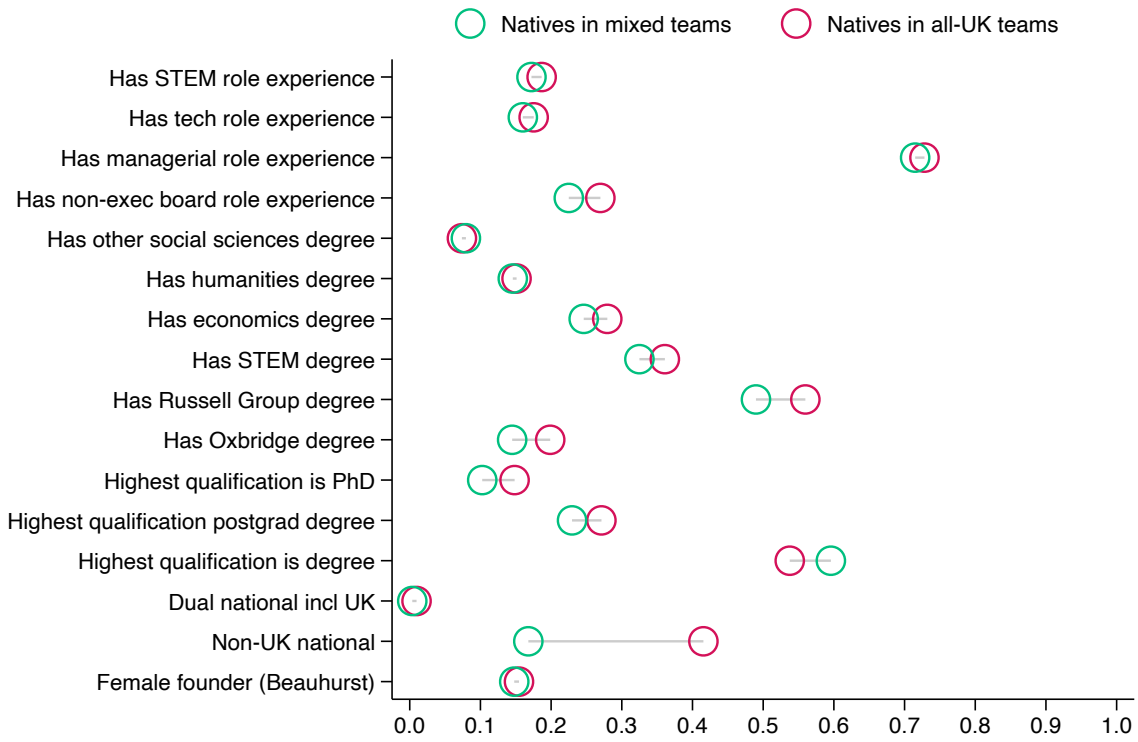
(i)



(j)

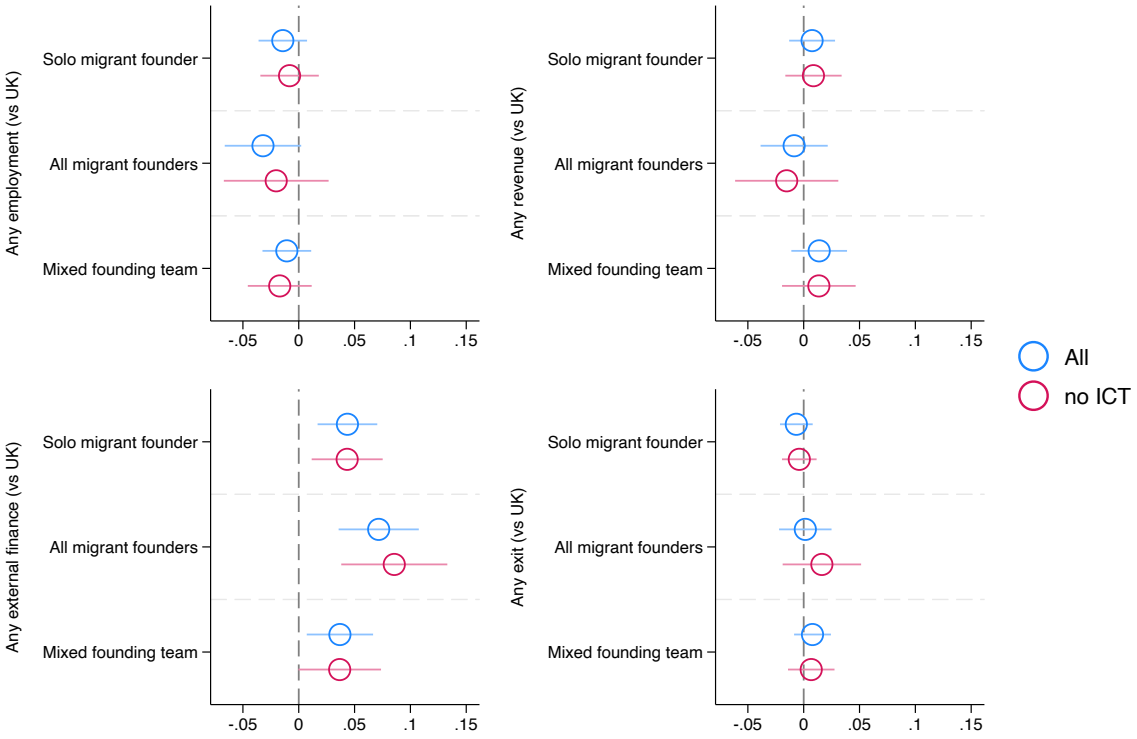
Source: Beauhurst, Diffbot. *Note:* Figures show marginal effects from regressions of firm novelty metrics on employment and revenue growth for 9,264 firms reporting headcount and 1,811 firms reporting revenue. Graphs show change from the first reported filing. Controls, fixed effects and standard errors as in Table A2.

Figure A3: Native selection in co-founded firms



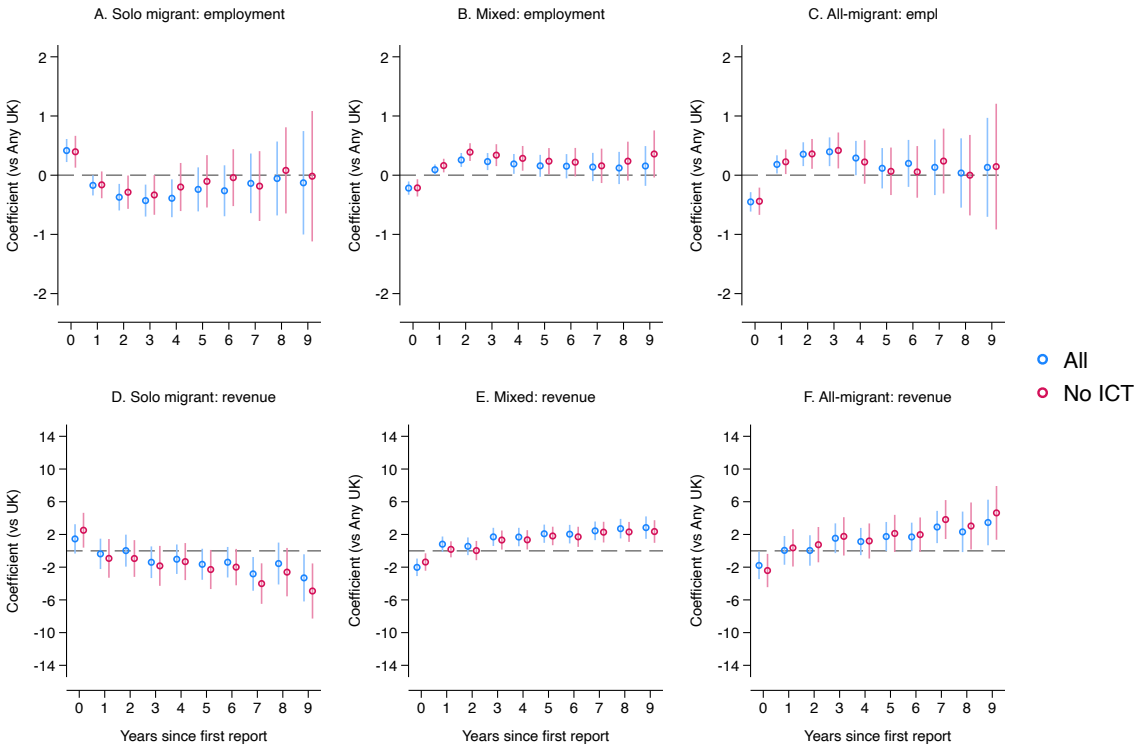
Source: Beauhurst, Diffbot. Sample UK-born founders in either mixed teams or all-UK teams.

Figure A4: Extensive margin - ICT leave-out test



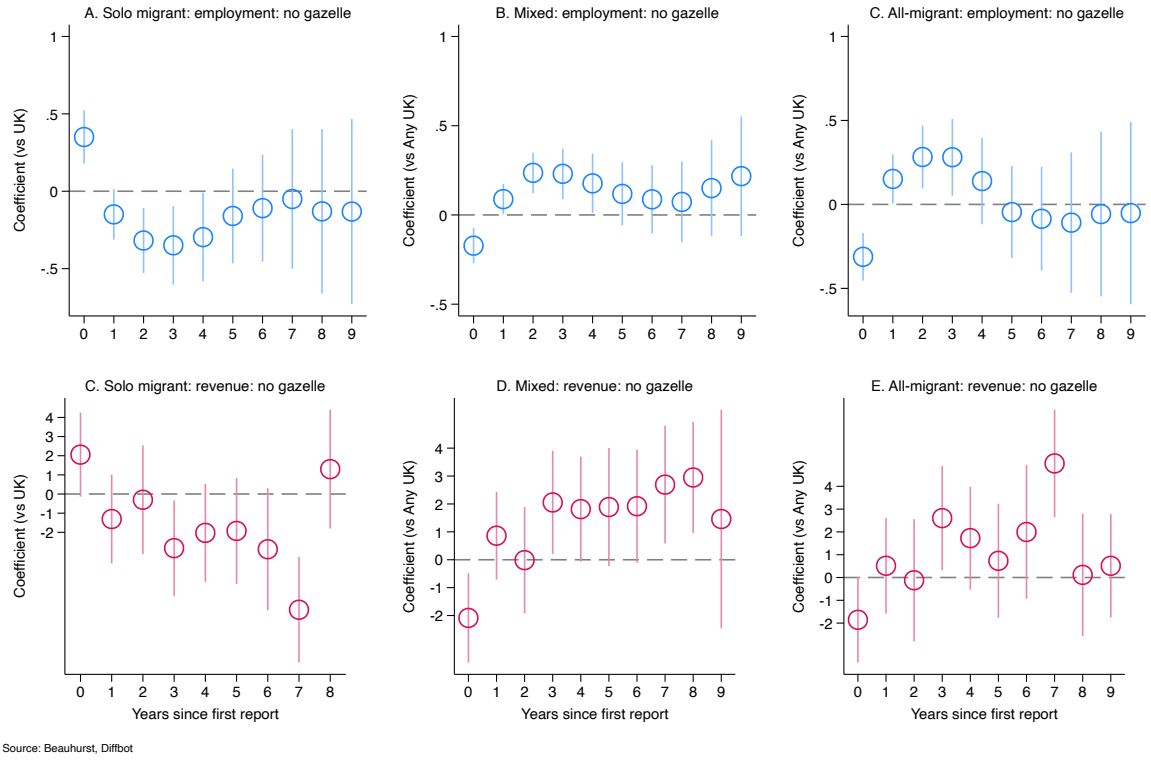
Source: Beauhurst, Diffbot. ICT defined as SIC07 code 620, "Computer programming, consultancy and related activities".

Figure A5: Intensive margin - ICT leave-out test



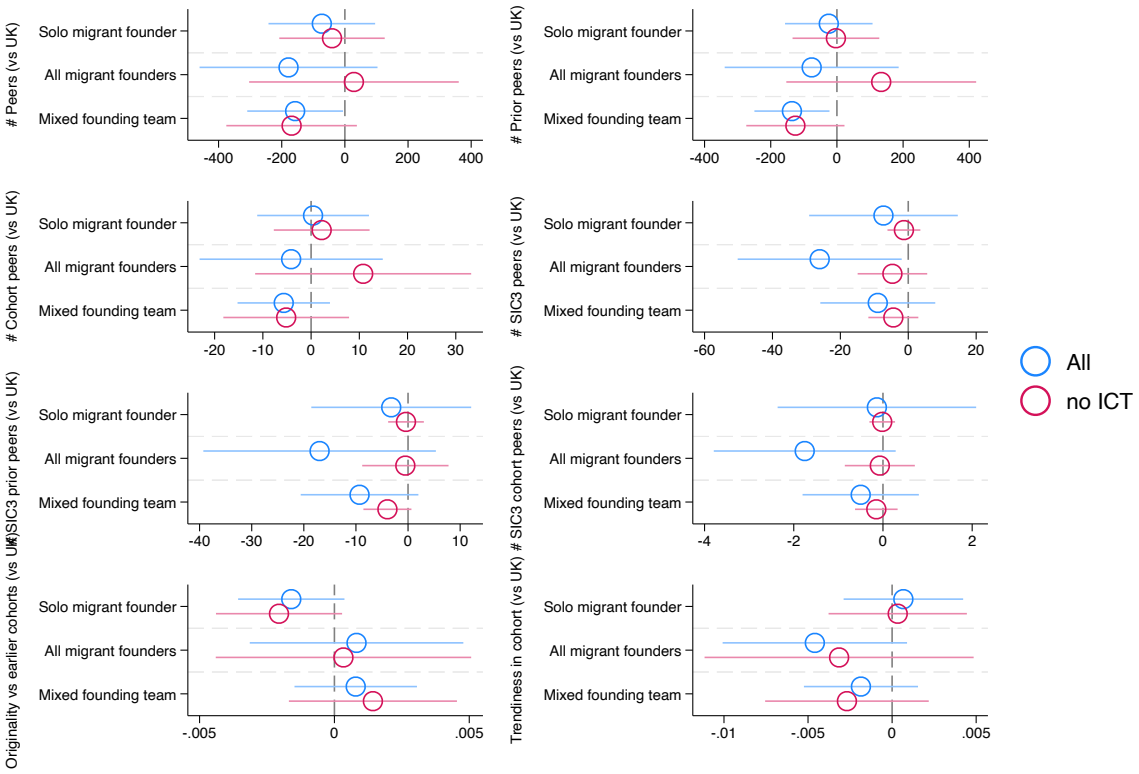
Source: Beauhurst, Diffbot. ICT defined as SIC07 code 620, "Computer programming, consultancy and related activities".

Figure A6: Intensive margin: robustness test removing scaleups



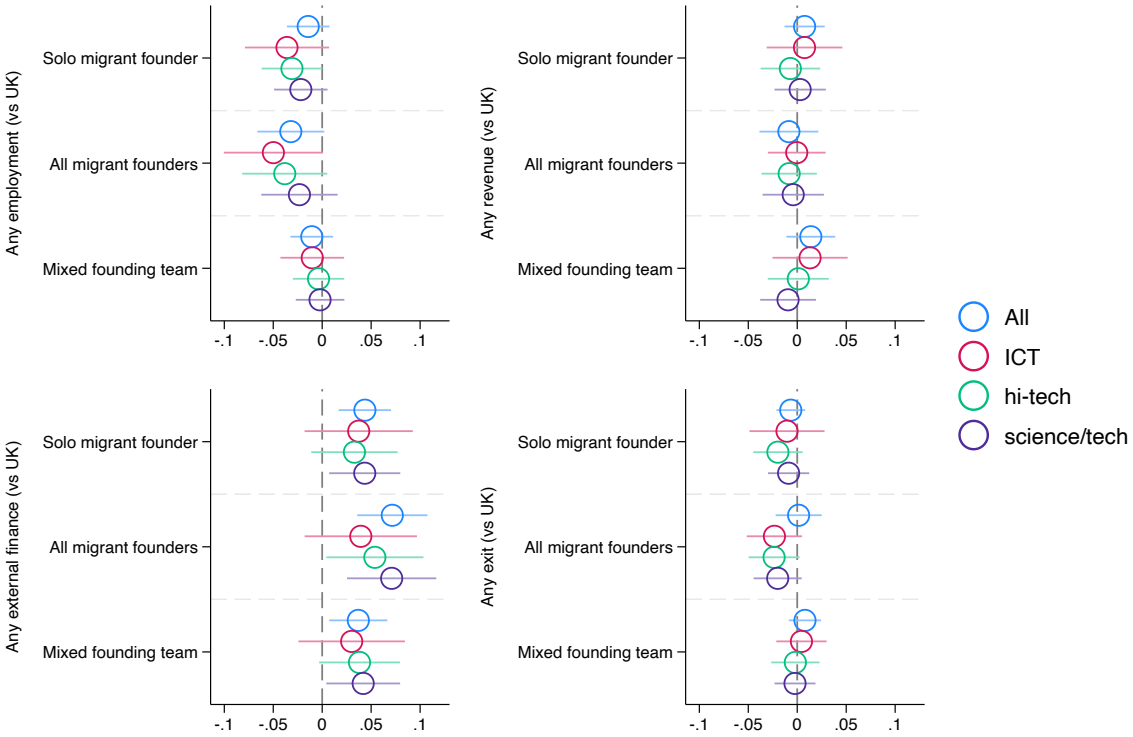
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020, excluding companies tracked on the basis of employment or revenue growth. Figure shows coefficients and 95% confidence intervals from Equation 4. The dependent variable is log employment (top row) or log revenue (bottom row). $\hat{b1}$s, levels effects, are denoted in year 0. $\hat{b2}$s, growth effects, are shown from years 1 - 9. Year 10 is excluded due to small cell sizes. All regressions include controls for share female, mean age, any Oxbridge, any managerial experience, any tech background, any STEM, any serial founder, and maximum prior founding experience, number of founders and a co-founded indicator, 3-digit industry-by-year and TTWA fixed effects respectively. Standard errors are clustered at the firm level.

Figure A7: Firm distinctiveness - ICT leave-out test - W2V



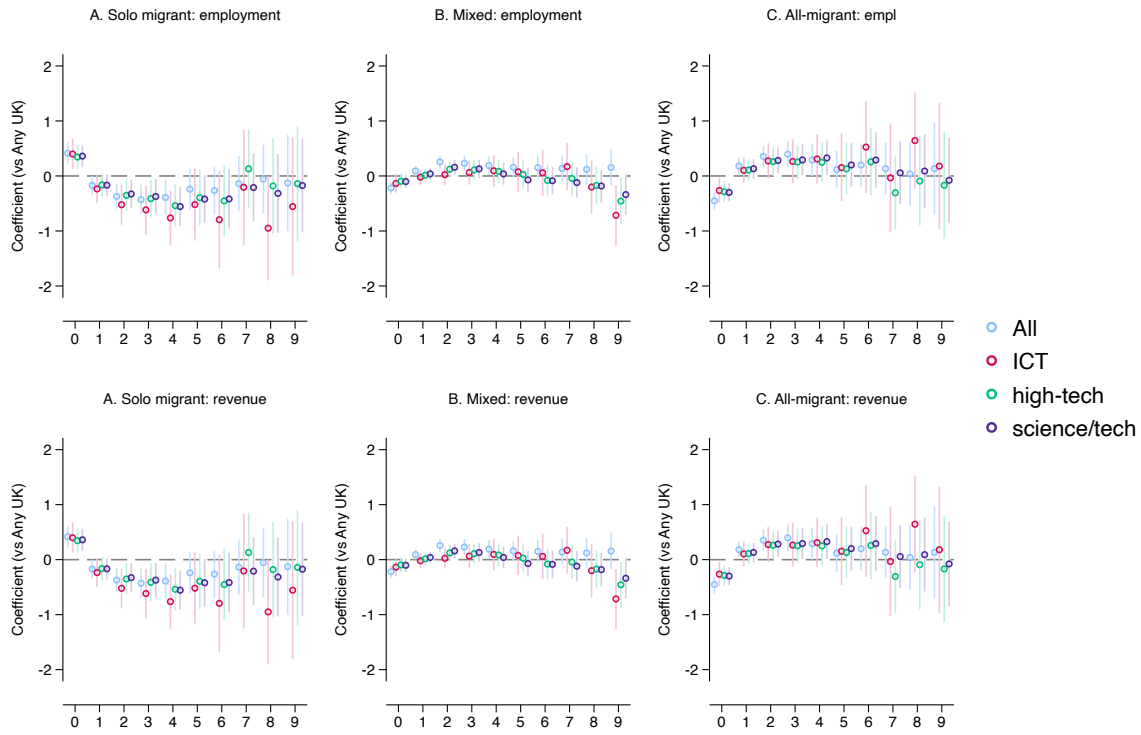
Source: Beauhurst, Diffbot. ICT defined as SIC07 code 620, "Computer programming, consultancy and related activities".

Figure A8: Extensive margin - sector splits



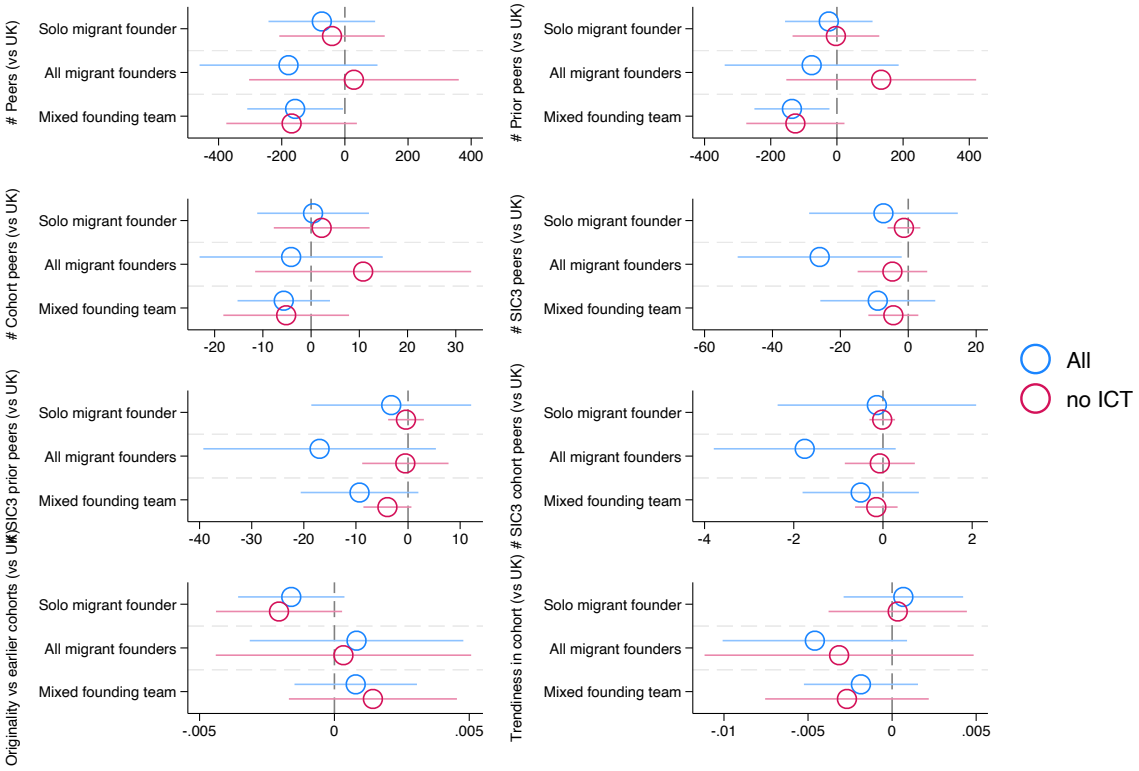
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals of linear probability models where the dependent variable is the probability of any employment post-founding, any revenue, external finance or exit. Other notes per Table 7.

Figure A9: Intensive margin: sector splits



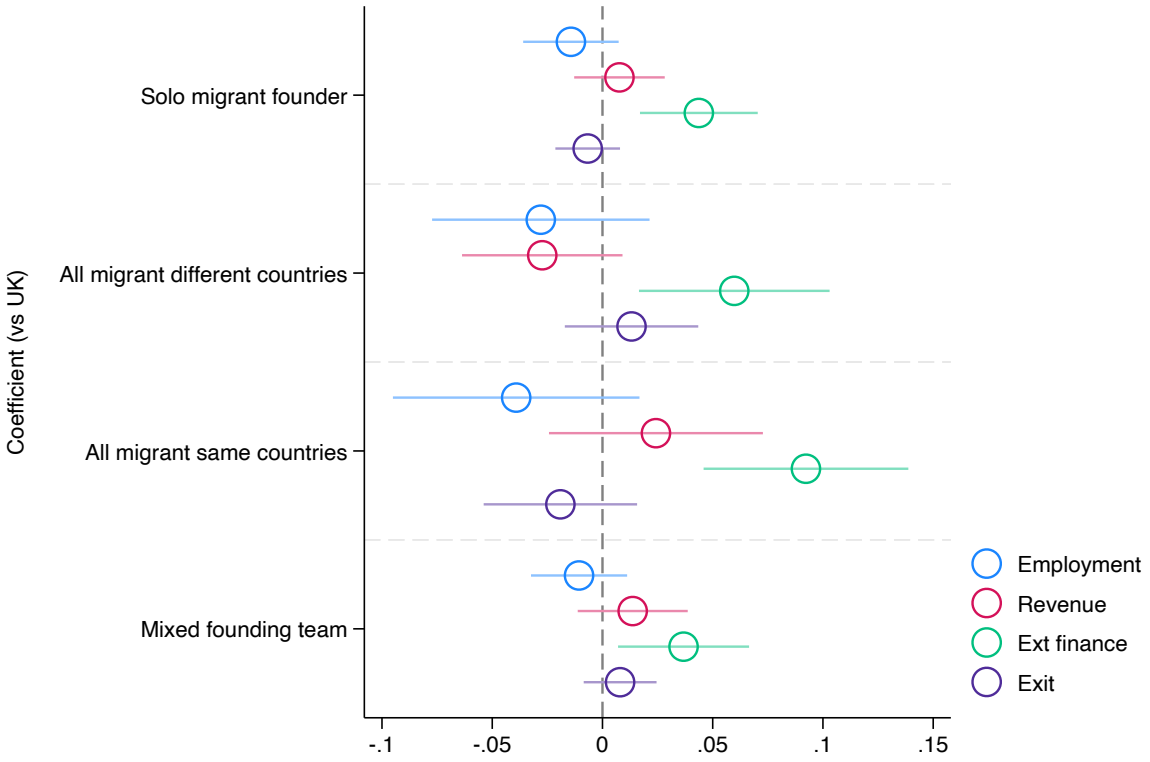
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020, excluding companies tracked on the basis of employment or revenue growth. Figure shows coefficients and 95% confidence intervals from Equation 4. The dependent variable is log employment (top row) or log revenue (bottom row). $\hat{b1}$s, levels effects, are denoted in year 0. $\hat{b2}$s, growth effects, are shown from years 1 - 9. Year 10 is excluded due to small cell sizes. All regressions include controls for share female, mean age, any Oxbridge, any managerial experience, any tech background, any STEM, any serial founder, and maximum prior founding experience, number of founders and a co-founded indicator, 3-digit industry-by-year and TTWA fixed effects respectively. Standard errors are clustered at the firm level.

Figure A10: Firm Distinctiveness - sector splits



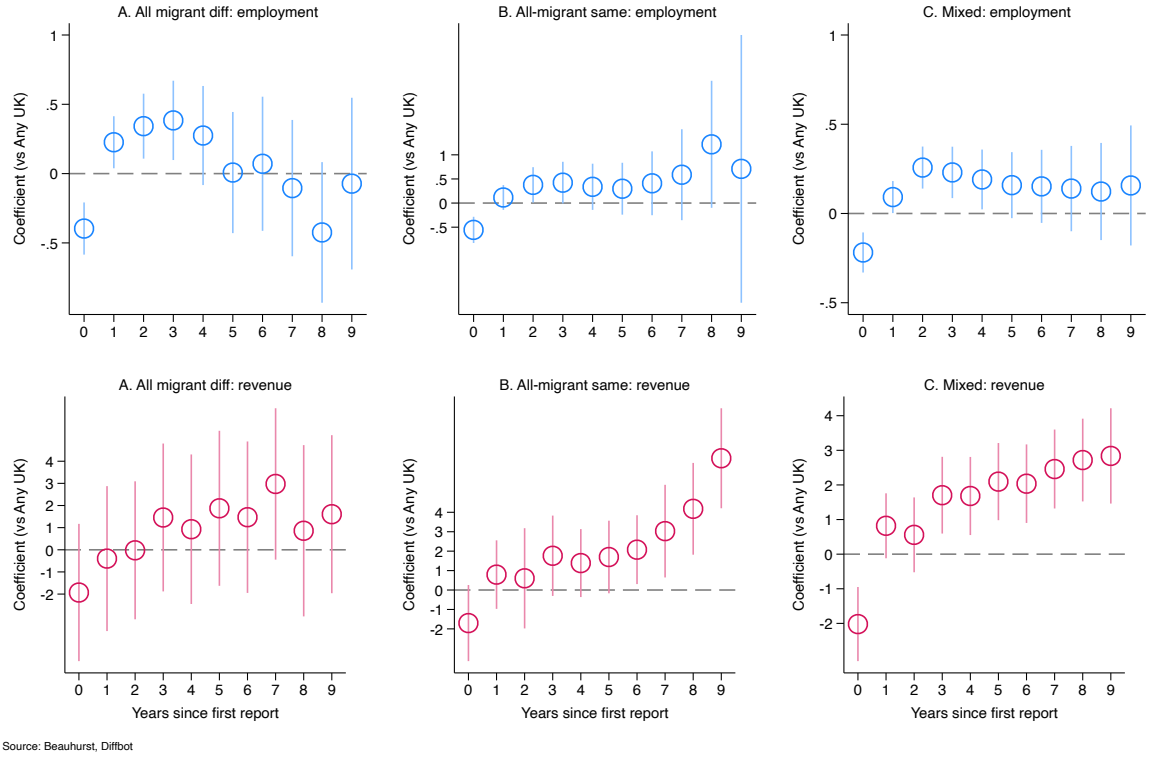
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals from Equation 5.

Figure A11: Firm Outcomes: Extensive margin split by sending country



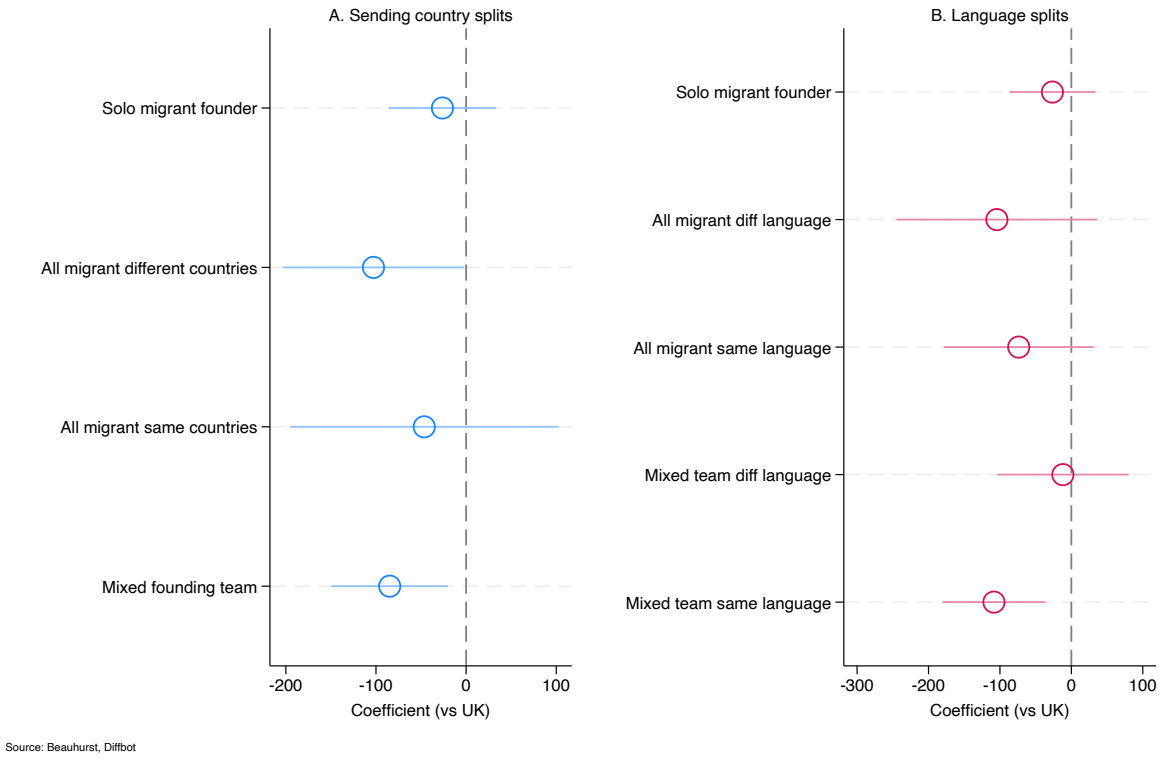
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals of linear probability models where the dependent variable is the probability of any employment post-founding, any revenue, external finance or exit. Other notes per Table 7.

Figure A12: Firm Outcomes: intensive margin split by sending country



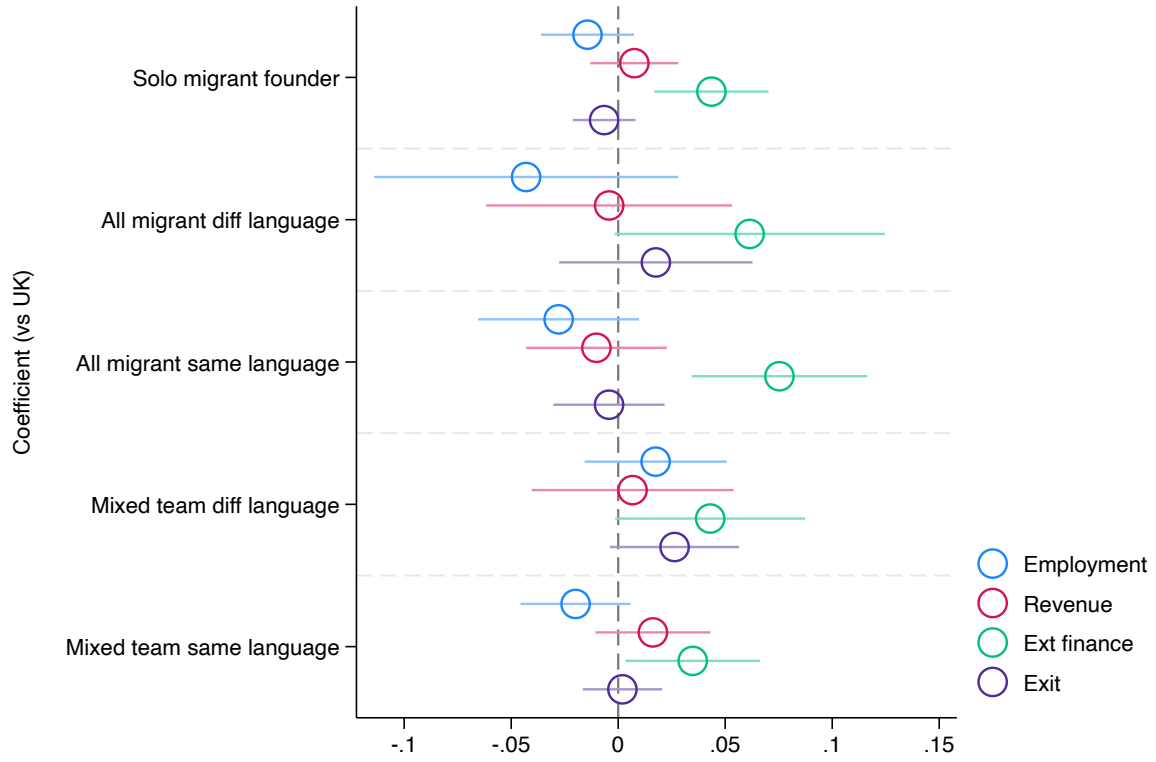
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals of Equation 4. The dependent variable is log employment (top) or log revenue (bottom). Other notes per Figure 9.

Figure A13: Firm distinctiveness: split by sending country and first language, Word2Vec



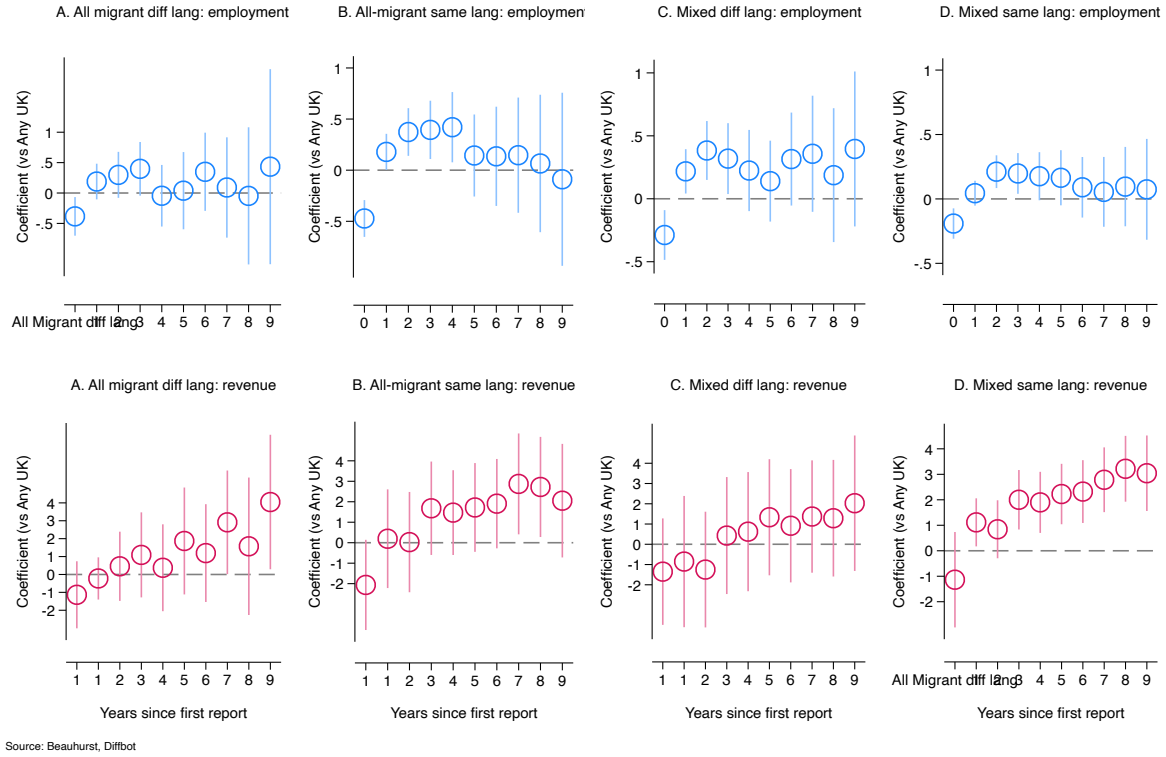
Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals from Equation 5.

Figure A14: Firm Outcomes: Extensive margin split by founder first language



Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals of linear probability models where the dependent variable is the probability of any employment post-founding, any revenue, external finance or exit. Other notes per Table 7.

Figure A15: Firm Outcomes: intensive margin split by founder first language



Source: Beauhurst, Diffbot. *Note:* Sample is companies founded 2000-2020. Figure shows coefficients and 95% confidence intervals of Equation 4. The dependent variable is log employment (top) or log revenue (bottom). Other notes per Figure 9.

Table A1: Founder balancing test. Dependent variable = Pr(founder in sample)

	(1) Status	(2) + city FE	(3) Education	(4) + city FE
Founder age (Beauhurst)	-0.006***	-0.004***	-0.005***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Female founder (Beauhurst)	0.015**	0.011	0.013**	0.018*
	(0.007)	(0.011)	(0.006)	(0.009)
Non-UK national	0.029***	0.009	0.022***	0.009
	(0.006)	(0.010)	(0.006)	(0.009)
Dual nationality	0.055***	0.063***	0.032	0.035
	(0.021)	(0.024)	(0.020)	(0.023)
Speaks more than 1 language	0.079***	0.056***	0.081***	0.061***
	(0.006)	(0.008)	(0.005)	(0.007)
Years of labour market experience	0.002***	0.001	0.002***	0.001**
	(0.000)	(0.001)	(0.000)	(0.001)
Number of jobs in career history	0.014***	0.009***	0.013***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Serial founder	0.008	0.008	0.009	0.004
	(0.012)	(0.016)	(0.010)	(0.015)
Has board experience	0.001	-0.009	-0.005	-0.019*
	(0.007)	(0.011)	(0.006)	(0.010)
Has management experience	0.024***	0.019**	0.032***	0.016*
	(0.006)	(0.010)	(0.006)	(0.009)
Has experience in tech role	0.022***	0.003	0.022***	0.005
	(0.007)	(0.011)	(0.006)	(0.009)
Observations	19113	7095	19113	7095
R^2	0.082	0.107	0.090	0.107
F-statistic	135.733	23.320	141.297	27.785
Oster delta	2.786	2.720	2.919	2.720

Notes: Source: Beauhurst, Diffbot. Sample is 23640 founders of companies founded 2000-2020. Models with city FE run on 8685 founders where current location is known. *** = significant at 1%, ** = 5%, * = 10%.

Table A2: Validation: Distinctiveness Measures and Firm Outcomes

	(1) Employ	(2) Revenue	(3) Ext Fin	(4) Exit
Panel A: Originality vs older firms				
All-time distinctiveness	-0.345***	-0.391***	0.138	-0.287***
	(0.116)	(0.115)	(0.146)	(0.067)
R^2	0.149	0.171	0.170	0.048
F-statistic	5.089	4.316	27.773	5.447
Panel B: Uniqueness				
No peers	-0.019***	-0.016**	0.009	-0.009*
	(0.007)	(0.008)	(0.010)	(0.005)
R^2	0.149	0.170	0.170	0.047
F-statistic	4.979	4.018	28.281	4.830
Panel C: Uniqueness in SIC3 bin				
No SIC3 peers	-0.019***	-0.016**	0.009	-0.009*
	(0.007)	(0.008)	(0.010)	(0.005)
R^2	0.149	0.170	0.170	0.047
F-statistic	4.979	4.018	28.281	4.830
Panel D: Number of peers				
Number of peers	0.000***	0.000***	-0.000***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.149	0.171	0.174	0.047
F-statistic	5.556	4.444	30.005	4.802
Panel E: Number of peers in SIC3 bin				
Number of SIC3 peers	0.001***	0.001**	-0.001**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.149	0.170	0.171	0.047
F-statistic	9.100	4.048	27.953	5.916
Panel F: Similarity within cohort				
Product similarity within cohort	0.284***	0.224***	-0.175**	0.104**
	(0.071)	(0.065)	(0.085)	(0.044)
R^2	0.150	0.171	0.170	0.047
F-statistic	6.439	4.433	28.530	4.952
Observations	9940	9940	9940	9940

Source: Beauhurst, Diffbot. Linear probability models. Each panel reports a separate regression of the outcome on one distinctiveness measure. Originality measures distinctiveness relative to all previously founded companies (higher = more original). Bin singleton is a dummy = 1 if the firm has no peers (= similar firms based on cosine similarity) in its own SIC3 industry. Degree is a count of a firm's peers in any industry. Bin degree is a count of a firm's peers in its own SIC3 industry. Trendiness measures similarity to firms founded in the same year (higher = more similar to cohort). Controls: founder gender, founder age, education, serial founders, founder years of experience, founder managerial/tech/stem background, co-founded firm, team size, multisite firm. Fixed effects: 3-digit SIC Industry × Incorporation Year, trading TTWA. Standard errors clustered at industry-year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Sample benchmarking against administrative data

2010		
Metric	BPE	Sample
Number of businesses	4,484,535	688
Number of employers	1,193,965	632
Share employers (%)	26.6	91.9
Number of SMEs	4,478,595	78
Share SMEs (%)	99.8	11.3
	ONS	Sample
Share of high-growth firms (%)	–	28.5
2015		
Metric	BPE	Sample
Number of businesses	5,389,450	5,888
Number of employers	1,311,865	5,418
Share employers (%)	24.3	92.1
Number of SMEs	5,382,485	1,142
Share SMEs (%)	99.8	19.4
	ONS	Sample
Share of high-growth firms (%)	–	13.6
2020		
Metric	BPE	Sample
Number of businesses	5,980,520	7,960
Number of employers	1,412,745	7,128
Share employers (%)	23.6	89.5
Number of SMEs	5,972,685	6,640
Share SMEs (%)	99.8	83.4
	ONS	Sample
Share of high-growth firms (%)	4.4	6.1

Sources: Department for Business, Innovation and Skills (2010, 2015); Department for Business, Enterprise, Innovation and Skills (2020); Office for National Statistics (2025). *Note:* This table compares our company estimation sample in 2010, 2015 and 2020 with administrative data for firms for the same year, or as close as is available. Comparison data drawn from Business Population Estimates (BPE), which covers the economy at the start of each year, and ONS high-growth business estimates (ONS) for the whole of 2020. Businesses and employers defined as private sector firms/companies (BPE Table 1). SMEs defined as firms/companies with 0–249 employees (BPE Table 24). High-growth firms/companies defined as having 10+ employees and average 3-year employment growth of 20% or more.

Table A4: Companies by reason for tracking

A. Tracking reason	Frequency	Percent
Academic spinout	73	0.60
Attended accelerator	1835	15.08
Got major innovation grant	479	3.94
Got equity funding	6801	55.92
Got venture debt	472	3.88
In high-growth list	1303	10.71
Completed managed buyout/in	74	0.61
Is 10% scaleup	232	1.91
Is 20% scaleup	894	7.35
N	12,163	100.00
B. Tracking stage	**Share**	**Avg. founding year**
Early stage trigger	0.655	2015
Mid stage trigger	0.246	2012
Late stage trigger	0.099	2008

Source: Beauhurst, Diffbot. Sample is 12,182 companies founded 2000-2020 inclusive. Panel A shows companies in our sample by Beauhurst 'growth signal'. Panel B bins these into three growth stages based on company founding year. The intuition is the younger (older) a company is when it hits a growth signal, the earlier (later) is the corresponding growth stage. We define academic spinouts and companies attending accelerators as tracked at early stage. We define companies receiving major innovation grants, equity or debt finance, or cited in high-growth lists as tracked at mid stage. We define companies that have completed buyouts/buyins or are scale-ups as tracked at late stage. 19 companies have no tracking information.

Table A5: Founder balancing test. Dependent variable = Pr(founder has em_city info)

	(1) All	(2) Origin country FE
Founder age (Beauhurst)	-0.005***	-0.005***
	(0.001)	(0.001)
Female founder (Beauhurst)	0.011	0.011
	(0.017)	(0.017)
Non-UK national	-0.002	-0.007
	(0.018)	(0.019)
Dual national incl UK	0.012	0.002
	(0.077)	(0.076)
Speaks more than 1 language	0.052***	0.048***
	(0.014)	(0.014)
Migrant founder	-0.006	0.000
	(0.017)	(.)
Years of experience	0.005***	0.005***
	(0.001)	(0.001)
Number of jobs held	-0.000	-0.001
	(0.002)	(0.002)
Is currently a founder	0.123***	0.123***
	(0.013)	(0.013)
Serial founder	0.024	0.025
	(0.029)	(0.029)
Has non-exec board role experience	-0.054***	-0.050***
	(0.018)	(0.018)
Has managerial role experience	0.080***	0.079***
	(0.014)	(0.014)
Has tech role experience	0.056***	0.055***
	(0.020)	(0.021)
Observations	6606	6576
R^2	0.045	0.059
F-statistic	21.638	22.301
Oster delta	4.159	5.136

Notes: Source: Beauhurst, Diffbot. Sample is 15,505 founders of companies founded 2000-2020. Models with sending country FE based on modelled migrant status, uses dummies for UK and other predicted sending countries. *** = significant at 1%, ** = 5%, * = 10%.

Table A6: Native sorting test: co-founded firms

	(1)
Founder age (Beauhurst)	0.002***
	(0.001)
Female founder (Beauhurst)	0.008
	(0.017)
Non-UK national	0.390***
	(0.023)
Dual nationality	-0.231**
	(0.094)
Has a degree in a STEM subject	0.011
	(0.013)
Has a degree in economics	0.007
	(0.013)
Has arts or humanities degree	0.027
	(0.017)
Number of languages spoken	0.016***
	(0.005)
Founder is based in London	-0.008
	(0.014)
Has experience in tech role	0.013
	(0.016)
Has Oxbridge degree	0.016
	(0.018)
Has Russell Group uni degree	0.008
	(0.013)
Observations	5278
$R\hat{2}$	0.196
F-statistic	29.303
Oster delta, Non-UK national	3.705

Source: Beauhurst, Diffbot. *Notes:* Sample restricted to native founders in co-founded firms (2+ founders) incorporated 2000–2020. The table shows results for a linear probability model where the outcome is a dummy taking the value 1 for UK-born founders co-founding with migrants, and 0 for UK-born co-founding with other natives. Company-level controls include incorporation year, dummy for tracking at mid stage, dummy for late stage tracking, co-founded company dummy, number of founders, plus 3-digit industry dummies and TTWA dummies. Standard errors in parentheses, clustered on founder. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Spatial Descriptives: Firm Outcomes by Urban Category and Founding Team Type

		Team type			
Urban category	Outcome	All-UK	All-Migrant	Mixed	Total
	Employment (mean)	27.24	31.41	24.49	26.63
	Revenue, £m (mean)	21.17	24.62	37.51	29.45
Major conurbation	N (employment)	995	312	980	2,287
	Employment (mean)	39.08	9.68	19.96	30.30
	Revenue, £m (mean)	35.22	2.10	10.03	26.32
Other urban	N (employment)	453	53	279	785
	Employment (mean)	38.83	—	18.94	30.03
	Revenue, £m (mean)	109.69	—	30.70	73.23
Rural town/fringe	N (employment)	52	5	31	88
	Employment (mean)	24.90	—	32.82	27.34
	Revenue, £m (mean)	13.51	—	9.00	11.05
Rural	N (employment)	92	6	66	164
	Employment (mean)	30.85	27.60	23.84	27.62
	Revenue, £m (mean)	28.28	22.27	31.75	29.15
Total	N (employment)	1,592	376	1,356	3,324

Notes: Source: Beauhurst, Companies House, ONS Postcode Directory. Sample restricted to co-founded firms (2+ founders) incorporated 2000–2020, at their last observation in the panel. Urban categories based on ONS 2011 Rural-Urban Classification. Employment is mean number of employees. Revenue is mean annual revenue expressed in £ millions. N reports number of non-missing employment observations. — indicates suppressed cells with fewer than 10 observations. * indicates cells with fewer than 10 observations; interpret with caution.

Table A8: Predictors of migrant founders

	All companies	Co-founded firms only
	(1)	(2)
Years of labour market experience	−0.001* (0.000)	−0.001 (0.001)
Number of jobs in career history	0.000 (0.001)	−0.001 (0.002)
Current founder	−0.000 (0.008)	0.006 (0.016)
Serial founder	−0.042*** (0.014)	−0.051* (0.026)
Has non-exec board experience	0.011 (0.009)	0.026 (0.017)
Has experience in managerial role	−0.000 (0.008)	−0.018 (0.015)
Highest qualification is degree	0.047*** (0.013)	0.033 (0.030)
Highest qualification is postgrad degree	0.136*** (0.015)	0.116*** (0.033)
Highest qualification is PhD	0.150*** (0.017)	0.150*** (0.038)
Founder age (Beauhurst)	0.001*** (0.000)	−0.001 (0.001)
Female founder (Beauhurst)	−0.008 (0.009)	−0.032 (0.021)
Non-UK national	0.413*** (0.011)	0.115*** (0.026)
Dual nationality	−0.081* (0.044)	−0.009 (0.084)
Has a degree in a STEM subject	0.008 (0.008)	0.002 (0.015)
Has a degree in economics	0.048*** (0.008)	0.055*** (0.016)
Has arts or humanities degree	−0.014 (0.010)	−0.043** (0.019)
Has Oxbridge degree	0.008 (0.010)	−0.003 (0.022)
Has Russell Group uni degree	−0.182*** (0.008)	−0.220*** (0.017)
Number of languages spoken	0.045*** (0.003)	0.044*** (0.005)
Founder is based in London	0.002 (0.009)	0.011 (0.017)
Has experience in tech role	0.022** (0.009)	−0.002 (0.018)
Observations	14680	6521
R^2	0.368	0.667
F-statistic	221.453	19.017
Oster delta, PhD	4.162	
Company controls	Yes	No
Company FE	No	Yes

Source: Beauhurst, Diffbot. *Notes:* Company controls include incorporation year, dummy for tracking at mid stage, dummy for late stage tracking, co-founded company dummy, number of founders. All regressions include 3-digit industry dummies and TTWA dummies. Standard errors in parentheses, clustered on founder. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Founder demography and Firm Outcomes: Co-founded firms

	Post-founding outcome			
	(1)	(2)	(3)	(4)
	Employ	Revenue	Ext. Fin	Exit
Migrant-only tean	-0.032	-0.013	0.070***	0.003
	(0.020)	(0.017)	(0.021)	(0.013)
Mixed team	-0.003	0.016	0.040**	0.006
	(0.011)	(0.015)	(0.016)	(0.009)
Mixed - Migrant	0.029	0.028	-0.030	0.003
Std. Error	0.020	0.016	0.020	0.012
p-value	0.150	0.081	0.127	0.800
Mean Dep. Var.	0.868	0.142	0.825	0.061
Observations	3,254	3,254	3,254	3,254
Adj. R^2	0.182	0.135	0.113	0.067

Source: Beauhurst, Diffbot. Linear probability models comparing future outcomes for migrant-founded companies and mixed / all-migrant founding teams, versus UK-founded companies. Sample is companies founded 2000-2020. Mixed - Migrant shows the difference in coefficients (H0: $b_{mixed} = b_{all-migrant}$). Controls: founder gender, founder age, education, serial founder, founder years of experience, founder managerial / tech / stem background, co-founded firm, team size, multisite firm. Fixed effects: 3-digit SIC Industry × Incorporation Year, Trading TTWA. Standard errors clustered at industry-year level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Extensive margin: robustness test excluding scaleups

	Firm Performance Outcomes			
	(1) Employ	(2) Revenue	(3) Ext. Fin	(4) Exit
Solo migrant	-0.017	0.002	0.045***	-0.006
	(0.012)	(0.010)	(0.014)	(0.007)
all migrant team	-0.036**	-0.016	0.064***	0.007
	(0.018)	(0.014)	(0.019)	(0.012)
Mixed team	-0.013	0.014	0.038**	0.013
	(0.012)	(0.012)	(0.016)	(0.009)
Mixed - Migrant	0.023	0.030	-0.026	0.006
Std. Error	0.019	0.015	0.018	0.013
p-value	0.234	0.050	0.142	0.672
Mean Dep. Var.	0.842	0.106	0.741	0.048
Observations	9,075	9,075	9,075	9,075
Adj. R^2	0.151	0.072	0.173	0.036

Source: Beauhurst, Diffbot. Linear probability models comparing future outcomes for migrant-founded companies and mixed / all-migrant founding teams, versus UK-founded companies. Sample is companies founded 2000-2020, minus companies tracked on the basis of employment and/or revenue growth. Mixed - Migrant shows the difference in coefficients (H0: $b_{mixed} = b_{all-migrant}$). Controls: founder gender, founder age, education, serial founder, founder years of experience, founder managerial / tech / stem background, co-founded firm, team size, multisite firm. Fixed effects: 3-digit SIC Industry × Incorporation Year, Trading TTWA. Standard errors clustered at industry-year level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Growth trajectories by founding team type

	Employment	Revenue
Level effects (omitted: all-UK)		
All migrant team	-0.451***	-1.783**
	(0.083)	(0.854)
Mixed team	-0.218***	-2.018***
	(0.057)	(0.545)
1.solo_migrant	0.417***	1.459
	(0.099)	(0.913)
All-migrant team × year		
Year 1	0.183**	0.056
	(0.078)	(0.896)
Year 2	0.355***	0.042
	(0.103)	(0.948)
Year 3	0.396***	1.537*
	(0.124)	(0.927)
Year 4	0.290*	1.135
	(0.148)	(0.848)
Year 5	0.117	1.750*
	(0.174)	(0.906)
Year 6	0.199	1.698*
	(0.202)	(0.890)
Year 7	0.133	2.912***
	(0.239)	(1.005)
Year 8	0.037	2.321*
	(0.299)	(1.262)
Year 9	0.132	3.465**
	(0.426)	(1.419)
Mixed team × year		
Year 1	0.092**	0.815*
	(0.046)	(0.478)
Year 2	0.257***	0.556
	(0.060)	(0.551)
Year 3	0.230***	1.706***
	(0.073)	(0.565)
Year 4	0.190**	1.686***
	(0.085)	(0.575)
Year 5	0.159*	2.092***
	(0.094)	(0.567)
Year 6	0.152	2.040***
	(0.104)	(0.578)
Year 7	0.138	2.443***
	(0.122)	(0.580)
Year 8	0.122	2.709***
	(0.139)	(0.609)
Year 9	0.156	2.837***
	(0.171)	(0.699)
Observations	29,455	5,739
Firms	9,114	1,624
Adj. R^2	0.338	0.351

Notes: Each column reports a panel regression of ln(employment) on the founder-demography vector **TYPE** interacted with years since first report (years 1–9; year 0 captured by the level effects), the **X** vector of founder and firm controls plus fixed effects as reported in Equation 4. The omitted year is firm age 0; the omitted firm type is all-UK. To conserve space the table reports level effects and the team-by-year interactions only; year main effects, solo-migrant-by-year interactions, and individual coefficients on the founder and team controls in **X** are estimated but not displayed. Year 10 is excluded from estimation owing to small cell sizes. Standard errors in parentheses are clustered at the firm level. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Founder Demography and Firm Distinctiveness: Teams

	Distinctiveness outcomes							
	(1) # Peers	(2) # Prior	(3) # Cohort	(4) # SIC3 peers	(5) # SIC3 prior	(6) # SIC3 cohort	(7) Originality	(8) Trendiness
Migrant-only team	-112.687	-47.809	-2.604	-11.575	-8.125	-1.036	0.000	-0.003
	(152.024)	(140.978)	(10.431)	(13.220)	(11.276)	(1.232)	(0.002)	(0.003)
Mixed team	-159.820*	-122.265*	-4.670	-8.917	-8.346	-0.421	0.001	0.000
	(89.706)	(66.404)	(5.738)	(9.815)	(5.889)	(0.731)	(0.001)	(0.002)
Mixed − Migrant-only	47.132	74.457	2.066	-2.658	0.221	-0.615	-0.001	-0.003
Std. error	161.546	144.433	11.340	15.875	11.459	1.508	0.002	0.003
p-value	0.771	0.606	0.856	0.867	0.985	0.684	0.788	0.417
Mean dep. var.	1978.684	1454.540	131.511	128.319	80.647	10.329	0.097	0.658
Observations	3,253	3,253	3,253	3,253	3,253	3,253	3,253	3,253
Adj. R^2	0.104	0.093	0.115	0.138	0.076	0.133	0.048	0.209

Source: Beauhurst, Diffbot. OLS models comparing distinctiveness of companies with mixed migrant–native and all-migrant founding teams, versus all-UK founding teams. Sample is co-founded companies founded 2000–2020. Mixed − Migrant-only shows the difference in coefficients (H_0: $b_{mixed} = b_{all\text{-}migrant}$). Controls: founder gender, founder age, education, serial founder, founder years of experience, founder managerial / tech / STEM background, team size, multisite firm. Fixed effects: 3-digit SIC industry × incorporation year, trading TTWA. Standard errors clustered at industry-year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Founder demography and Firm Distinctiveness: alternative metrics: peer counts check

	Firm Performance Outcomes			
	(1) Main	(2) TF-IDF	(3) 10% Threshold	(4) 25% Threshold
Solo migrant	-73.049	-26.408	-75.587	-60.536
	(86.047)	(30.504)	(138.928)	(242.938)
Migrant-only team	-178.450	-82.249*	-368.035	-765.051*
	(143.551)	(47.423)	(252.490)	(449.533)
Mixed team	-157.508**	-84.572**	-256.340*	-359.730
	(77.312)	(32.921)	(130.892)	(242.283)
Mixed - Migrant	-20.942	2.323	-111.695	-405.321
Std. Error	146.721	49.433	256.574	460.092
p-value	0.887	0.963	0.663	0.379
Mean Dep. Var.	2010.135	1345.067	4183.352	10876.822
Observations	9,940	9,941	9,940	9,940
Adj. R^2	0.112	0.161	0.134	0.168

Source: Beauhurst, Diffbot. Linear probability models comparing future outcomes for migrant-founded companies and mixed / all-migrant founding teams, versus UK-founded companies. Mixed - Migrant shows the difference in coefficients (H0: $b_{mixed} = b_{all-migrant}$). Controls: founder gender, founder age, education, serial founder, founder years of experience, founder managerial / tech / stem background, co-founded firm, team size, multisite firm. Fixed effects: 3-digit SIC Industry × Incorporation Year, Trading TTWA. Standard errors clustered at industry-year level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Founder Demography and Firm Distinctiveness: Younger Firms

	Distinctiveness outcomes							
	(1) # Peers	(2) # Prior	(3) # Cohort	(4) # SIC3 peers	(5) # SIC3 prior	(6) # SIC3 cohort	(7) Originality	(8) Trendiness
Solo migrant	-100.922	-91.706	0.706	-5.533	-3.525	0.132	-0.003	0.006
	(175.119)	(203.834)	(9.830)	(18.592)	(20.678)	(1.342)	(0.003)	(0.004)
Migrant-only team	197.128	171.641	16.471	-1.753	-3.986	0.351	-0.003	0.003
	(294.402)	(350.938)	(17.424)	(27.490)	(35.836)	(2.765)	(0.006)	(0.006)
Mixed team	-185.791	-221.711	-5.046	-16.452	-13.991	-0.342	-0.001	0.003
	(128.152)	(146.946)	(6.974)	(10.985)	(12.980)	(0.596)	(0.002)	(0.004)
Mixed − Migrant-only	382.919	393.352	21.517	14.699	10.004	0.693	-0.003	-0.000
Std. error	283.773	344.865	16.833	23.581	29.292	2.289	0.006	0.006
p-value	0.179	0.256	0.203	0.534	0.733	0.763	0.636	0.991
Mean dep. var.	1494.592	1907.156	97.365	90.874	117.769	9.866	0.102	0.645
Observations	1,712	1,712	1,712	1,712	1,712	1,712	1,712	1,712
Adj. R^2	0.051	0.060	0.077	0.078	0.105	0.153	0.061	0.143

Source: Beauhurst, Diffbot. OLS models comparing distinctiveness of migrant-founded companies and mixed / all-migrant founding teams, versus UK-founded companies. Sample is companies founded 2018–2020. Mixed − Migrant-only shows the difference in coefficients (H_0: $b_{mixed} = b_{all\text{-}migrant}$). Controls: founder gender, founder age, education, serial founder, founder years of experience, founder managerial / tech / STEM background, co-founded firm, team size, multisite firm. Fixed effects: 3-digit SIC industry × incorporation year, trading TTWA. Standard errors clustered at industry-year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.